Studies in Computational Intelligence 908

Utku Kose
Jafar Alzubi *Editors*

Deep Learning for Cancer Diagnosis

Springer

Studies in Computational Intelligence

Volume 908

Series Editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland

The series "Studies in Computational Intelligence" (SCI) publishes new developments and advances in the various areas of computational intelligence—quickly and with a high quality. The intent is to cover the theory, applications, and design methods of computational intelligence, as embedded in the fields of engineering, computer science, physics and life sciences, as well as the methodologies behind them. The series contains monographs, lecture notes and edited volumes in computational intelligence spanning the areas of neural networks, connectionist systems, genetic algorithms, evolutionary computation, artificial intelligence, cellular automata, self-organizing systems, soft computing, fuzzy systems, and hybrid intelligent systems. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution, which enable both wide and rapid dissemination of research output.

The books of this series are submitted to indexing to Web of Science, EI-Compendex, DBLP, SCOPUS, Google Scholar and Springerlink.

More information about this series at http://www.springer.com/series/7092

Utku Kose · Jafar Alzubi
Editors

Deep Learning for Cancer Diagnosis

Springer

Editors
Utku Kose
Department of Computer Engineering
Süleyman Demirel University
Isparta, Turkey

Jafar Alzubi
Faculty of Engineering
Al-Balqa Applied University
Balqa, Jordan

ISSN 1860-949X ISSN 1860-9503 (electronic)
Studies in Computational Intelligence
ISBN 978-981-15-6320-1 ISBN 978-981-15-6321-8 (eBook)
https://doi.org/10.1007/978-981-15-6321-8

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Foreword by Dr. Omer Deperlioglu

As the field of artificial intelligence is affecting all fields of the modern life, the need for processing big amount of data has been solved effectively with deep learning techniques. With the start of the twenty-first century, computational technologies have started to gain a momentum to solve time complexity appeared while using traditional techniques of machine learning/artificial intelligence. At this point, the deep learning has become a strong actor in contributing to the unstoppable rise of intelligent systems. As the future of artificial intelligence has many more surprises for all of us, current using alternatives of deep learning takes researchers' interest too much. Among all its employments within different fields, maybe the field of medical is one of the most vital platforms where the future of humankind is shaped rapidly.

In this book, titled as *Deep Learning for Cancer Diagnosis*, you will find lots of recent works focusing on the essentials of deep learning, its application types for cancer diagnosis, and even general overview of how the future of deep learning will change and improve in the next time period. As cancer is one of the most important diseases threatening the existence of humankind, it has been a very long time medical and many supportive fields are working hardly for developing effective solutions. Except from medicine works, which are generally focused on the treatment process of cancer, it is too important to design efficient methodologies for early diagnosis of different cancer types. As each different type of cancer is affecting different parts and organs of our body, it has been always too difficult to get something like a multi-cancer diagnosis system. But in my opinion, it is now easier—as not done like before—to run computationally effective and accurate automated systems thanks to artificial intelligence and the subfield of deep learning. You will see in this book that deep learning can be a single or hybrid solution (with especially image processing) for effective detection of tumors and diagnosis of different cancer types. As deep learning is really a deep subfield with many combinations of deeper neural network types, you will see that there is a great effort in designing unique, alternative models to diagnose different cancer types such as brain tumors/gliomas, breast cancer, lung cancer, melanoma, and fibrosarcoma. It is remarkable that each of the chapter included in this book uses pure language to

explain all technical details to a wide audience of beginners to advanced researchers enrolled in the context of artificial intelligence, cancer diagnosis, and biomedical-based topics. I suggest all readers to consider the following points: (1) Starting sections of each chapter explain the essentials of deep learning approaches, methods, and techniques from different perspective. So, you may read the whole book firstly to understand the essentials of deep learning. (2) After that, you can read further sections to get adapted to the employment of different techniques for diagnosis processes. Except from these, it is also easy to get informed about how specific types of cancers are diagnosed thanks to deep learning-based methods, by reading the associated chapters separately. After finishing that book, you will be ready to work on cancer diagnosis via deep learning and/or improve your knowledge skills for further research, in which the literatures of artificial intelligence and medical are both still needing. I suggest all the colleagues that they can use that work in their lectures about especially artificial intelligence and its use within medical–biomedical problems. In this context, the book opens a view window for a great collaboration of computer science and the medical, by considering even a specific but remarkable enough research topic.

I would like to thank all authors for their valuable contributions to create that book as a great reference for studies on cancer diagnosis and effective use of deep learning for all phases of that. Just turn the pages to make a start to your journey!

Dr. Omer Deperlioglu
Afyon Kocatepe University
Afyonkarahisar, Turkey
e-mail: deperlioglu@aku.edu.tr

Foreword by Dr. Jose Antonio Marmolejo-Saucedo

Machine learning is known as the locomotive of the artificial intelligence as it is the key factor of building autonomous systems. After artificial intelligence started to take an active part in especially daily life, the future view for the world in science fiction movies has been started to be discussed widely. Here, more active role in daily life caused machine learning solutions to be included in many devices and software systems we are using. By the way, the capabilities of intelligent systems to learn from samples and even learning something in an evolutionary way caused researchers in all fields to transform their works into innovative forms including running of intelligent solution mechanisms on the background. As the automation has been already transforming in a technological manner, the age of Industry 4.0 has become the main title of innovations appeared over shoulders of especially machine learning.

It is remarkable that machine learning is actually an optimization process of artificial intelligence. In general, machine learning techniques employ optimization processes, which are called as training or learning (Because of that, especially intelligent optimization algorithms can be used directly for training machine learning and designing hybrid systems in this way). Since that optimization has been important as the mathematical background of machine learning, big data caused researchers to think about designing something new. Called as deep learning, more complicated versions of well-known artificial neural networks are currently among the most effectively used techniques for solving today's advanced problems. In other terms, the advanced optimization, which is required by bigger problems, is solved by deep learning nowadays.

It is a pleasure for me to write a foreword for a work of using deep learning and explaining its active role in the field of medical. Titled as *Deep Learning for Cancer Diagnosis*, this book is a timely support for the literature of medical as we are currently experiencing a terrible pandemic by COVID-19 and that is an important sign that we need more active use of technology and artificial intelligence for early predictions and diagnosis of diseases treating our life. As we all know, cancer is the enemy of the living organisms' cells and it causes deaths at the end if it is not treated effectively. Currently, there are different vital types of cancers, and they all

need as early as possible diagnosis for successful treatment at the end. In this sense, deep learning techniques have been accurately used for diagnosing cancer, and the field of medical is highly interested in effective collaborations for automated systems for ensuring high-level solutions that can diagnose and treat better than humans. In this book, readers will be able to read about how deep learning is applied for the diagnosis process, as including pre-work and post-work approaches changing according to the medical data considered. The associated literature is too active in medical diagnosis by artificial intelligence, and it seems that the future will be more enrolled in early diagnosing, worldwide solutions to ensure a sustainable, disease-free life for humankind.

The chapters included in the book target especially important types of cancers, and according to me, all these chapters have key points, which can enable the readers to get necessary 'know-how' about applying deep learning in medical diagnosis. In this way, the readers will get informed about using artificial intelligence for not only cancer but also for different diseases, which are caused by viruses, genes, and even harmful environmental factors resulting in evolutionary changes leading to new types of diseases. Hopefully, the science has always solutions and that book will train open minds for better solutions for the future.

Dr. Jose Antonio Marmolejo-Saucedo
Universidad Panamericana
Mexico City, Mexico
e-mail: jmarmolejo@up.edu.mx

Preface

The field of artificial intelligence has a very wide of diversity in terms of different computational and logical solutions for real-world problems. As long as we can conduct the exact modeling of a real-world problem, it is always possible to run an effective artificial intelligence algorithm technique for successful results. The need for the artificial intelligence is generally associated with more accurate and efficient results for problems, which humankind cannot solve with high success rates or provide an exact solution yet. So, artificial intelligence has been a great tool for technological developments and that has been feeling too much by the community, especially in 2000s. As like many other technologies are evolving in time, that field of intelligent systems has been evolving too, and as a result, we have different kinds of solutions introduced to the literature. Recently, there is a storm of deep learning, which is a collection—at least for now—of advanced neural network models with specific layer types that enable us to deal with complex, big amount, and real-time data. Because the results by deep learning are too successful, it has been employed within all fields of modern life rapidly. The field of medical is among them.

While we are writing that preface, humankind is experiencing a fatal pandemic caused by COVID-19 type virus, and there is a great emergency state within every country as including breaks in educational institutions, curfews, and valuable medical efforts to find diagnosis and treatment solutions against that virus. Here, the role of technology is too remarkable that there is a race against time as deaths caused by COVID-19 are increasing fast if there is no immediate action taken. Moving from that fact, it is possible to indicate that the future of humankind will be associated with important developments for especially well-being, ensuring a stable health state within all over the world, and keeping sustainability oriented green solutions to keep the balance of the life at a certain level. That includes research efforts for the active run of technology within problems of medical and even its associated fields such as biology, genetics, and chemistry.

In this context, we would like to introduce our edited work: Deep Learning for Cancer Diagnosis, as the latest collections of a total of 16 valuable chapters targeting different aspects of how deep learning is applied for deriving successful solutions for another vital enemy of humankind: cancer. According to the latest

statistics, there are more than 100 different types of cancer threatening human life, and these different cancer types always require early diagnosis and effective treatments for desired success. As it is important to understand how deep learning is applied with alternative ways to diagnose different cancer types, we have also focused on gathering the chapters explaining essentials of different deep learning techniques and their characteristics making them to be applied further for medical diagnosis including the cancer. In this way, valuable readers are enabled to read a comprehensive enough work to learn and understand connections among artificial intelligence, medical, diagnosis, algorithms, and data better. As Dr. Deperlioglu mentions in his valuable foreword that the book can be used collaboratively within different fields as in different courses in them, and it will give essential idea about the current state of the science and rise insights about the future. Additionally, Dr. Marmolejo-Saucedo indicates in his wide picture foreword that the book will be a benefit for performing more about medical diagnosis and developing the future of medical with artificial intelligence and its key elements.

Here is how each of the chapters contributed to that book project in terms of cancer diagnosis via deep learning-oriented solutions:

Chapter 1 considers common use of deep learning and image processing for cancer diagnosis with a wide-open introduction to essentials of their uses and then evaluates a well-known deep learning technique: convolutional neural network (CNN) for breast cancer diagnosis.

Chapter 2 evaluates the performance of popular machine/deep learning techniques in the context of breast cancer and skin cancer diagnosis. In this context, the chapter employs support vector machines (SVM), random forest (RF), recurrent neural network (RNN), and convolutional neural network (CNN) for diagnosis processes.

Chapter 3 provides a great overview for the role of deep learning in processing medical image for cancer diagnosis, by discussing challenges and future scope. Additionally, the chapter also evaluates the diagnosis applications against CT/MR brain and abdomen images, mammogram images, histopathological images, and even diabetic retinopathy in this manner.

Chapter 4 focuses on diagnosis regarding canine fibroma and fibrosarcoma and employs the convolutional neural network (CNN) for that purpose. In detail, it evaluates the introduced model FibroNet with some other well-known models: VGG16, ResNET50, MobileNet-V2, and the Inception-V3.

Chapter 5 recalls the recent of subject of big data and uses different models of convolutional neural network (CNN) for performing diagnosis regarding Melanoma, by applying it over skin lesions data.

Chapter 6 introduces a hybrid use of intelligent techniques for diagnosing a tumor as oligodendroglioma (benign tumors) or astrocytoma (malignant tumors), by considering both radiology and pathology images. In detail, a hybrid combination of Inception-V3 and the support vector machines (SVM) is used for the diagnosis task.

Chapter 7 provides a general discussion regarding how deep learning is applied for a better diagnosis of cancer. In this sense, it introduces the technical background in the scope of diagnosis with deep learning and evaluates different techniques in terms of different evaluation metrics.

Chapter 8 directs its scope to the lung cancer, and it explains how deep learning-oriented research has been done in terms of diagnosing lung cancer over especially computed tomography (CT). The chapter briefly focuses on the most recent works for informing the readers about the current state of the literature.

Chapter 9 ensures a review of how cancer diagnosis is done effectively thanks to a strong relation between deep learning and image processing. The chapter briefly discusses deep learning techniques and active use of image processing against cancer diagnosis (over medical imaging), in terms of the associated literature.

Chapter 10 proposes a lightweight deep learning model for ensuring robust diagnosis of breast cancer over mammography images. In detail, the chapter compares the developed model with alternative techniques from the deep learning domain.

Chapter 11 considers the subject of brain tumor segmentation and provides an overview for the role of deep learning in ensuring successful segmentation. In order to achieve that, the chapter informs about the background, discusses brain tumor visualization, and then evaluates the role of deep learning in brain tumor segmentation.

Chapter 12 runs the model of convolutional neural network (CNN) for diagnosing lung cancer. For the diagnosis tasks, the chapter uses chest X-ray images and evaluates the success of the CNN with some other competitors.

Chapter 13 introduces the mechanism of deep learning in cancer diagnosis, by considering a view of the literature, and then provides a direct prediction for the future. In, especially, predictions regarding the future, the chapter discusses not only the role of deep learning but also medical imaging approaches methods.

Chapter 14 employs the 2D-UNET, which is a model of convolutional neural network (CNN) for brain tumor segmentation and considers the gliomas on diagnosing detection efforts done over the MRI images.

Chapter 15 follows another review approach for evaluating how deep learning is used for cancer diagnosis and enables readers to make a pure start from essentials of the field of artificial intelligence to the current end state of how popular techniques of deep learning are used within research works for cancer diagnosis.

Chapter 16 considers the subject of diagnosis of gliomas and gives a wide overview of applying deep learning techniques for detecting the gliomas over MRI image data. In detail, the chapter also discusses limitations as well as challenges of deep learning and even potential future of the deep learning techniques in prognosis, diagnosis, and eventually decision making.

As the readers, you are all welcome deep inside to read further about each chapter, in order to be up to date about the literature of deep learning for cancer diagnosis, and also gain knowledge about how different techniques can be applied further for fighting with that ruthless disease. Valuable feedback and contributive ideas from all readers about how we can make use of intelligent technologies for

diagnosing cancers and even dealing with all medical diseases (as like the current COVID-19 virus) are all welcome. As the editors, we would like thank you for your attention in our work and wish to see your in our future works as well.

Editors
Dr. Utku Kose
utkukose@sdu.edu.tr
http://www.utkukose.com/

Dr. Jafar Alzubi
j.zubi@bau.edu.jo
http://www.utkukose.com/

Acknowledgements As the editors, we would like to thank Loyola D'Silva, Swetha Divakar, and the Springer Team for their valuable efforts and great support on pre-organizing the content and publishing of the book.

About This Book

As it is known, artificial intelligence has taken many steps away for effective solutions in the field of medical. In this context, deep learning is a recent and remarkable subfield, which can deal with huge data for more accurate results. As a vital research topic, medical diagnosis is among research efforts in which deep learning-oriented solutions are often employed. Considering that state, the objective of this edited book is to provide recent advanced applications of deep learning for diagnosing different types of cancer disease. The book consists of a total of 16 chapters, which are a collection of the recent research efforts for understanding the current state of the literature, diagnosing cancer effectively, and giving deep enough insights into the future. The target audience of the book covers scientists, experts, M.Sc. and Ph.D. students, postdocs, and anyone interested in the related subjects covered. The book is suitable to be used as a reference work in the courses associated with artificial intelligence, medical, and biomedical.

Contents

Contributors

V. Ajantha Devi AP3 Solutions, Chennai, India

H. Ajay Kumar Mar Ephraem College of Engineering and Technology, Elavuvilai, India

Bekir Aksoy Faculty of Technology Mechatronics Engineering, Isparta University of Applied Sciences, Isparta, Turkey

Gokhan Altan Iskenderun Technical University, Computer Engineering Department, Hatay, Turkey

K. Bagyalakshmi Department of Electrical and Electronics Engineering, Sri Ranganathar Institute of Engineering and Technology, Coimbatore, India

Kathleen M. Curran School of Medicine, University College Dublin, Dublin, Ireland

D. A. A. Deepal Department of Computer Science, Faculty of Applied Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda, Sri Lanka

T. G. I. Fernando Department of Computer Science, Faculty of Applied Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda, Sri Lanka

Fabrizio Frezza Department of Information Engineering, Electronics and Telecommunications (DIET), Sapienza University of Rome, Rome, Italy

Balazs Gulyas Nanyang Technological University, Singapore, Singapore

M. Hanefi Calp Department of Management Information Systems, Karadeniz Technical University, Trabzon, Turkey

John J. Healy School of Electrical and Electronic Engineering, University College Dublin, Dublin, Ireland

L. R. Jonisha Miriam Mar Ephraem College of Engineering and Technology, Elavuvilai, India

F. Kayaalp Düzce University, Düzce, Turkey

İsmail Kırbaş Burdur Mehmet Akif Ersoy University, Burdur, Turkey

Pinar Koc Institute of Education Sciences, Computer and Instructional Technologies Education, Suleyman Demirel University, Isparta, Turkey

Ambeshwar Kumar School of Computing, SASTRA Deemed University, Thanjavur, India

S. N. Kumar Amal Jyothi College of Engineering, Kanjirappally, Kerala, India

A. Lenin Fred Mar Ephraem College of Engineering and Technology, Elavuvilai, India

R. Manikandan School of Computing, SASTRA Deemed University, Thanjavur, India

Iosif Mporas School of Engineering and Computer Science, University of Hertfordshire, Hatfield, UK

Khushboo Munir Department of Information Engineering, Electronics and Telecommunications (DIET), Sapienza University of Rome, Rome, Italy

Prasitthichai Naronglerdrit Department of Computer Engineering, Faculty of Engineering at Sriracha, Kasetsart University Sriracha Campus, Chonburi, Thailand

Anand Nayyar Graduate School, Faculty of Information Technology, Duy Tan University, Da Nang, Viet Nam

Özlem Özmen Burdur Mehmet Akif Ersoy University, Burdur, Turkey

Koray Özsoy Senirkent Vocational School Electric and Energy Department, Isparta University of Applied Sciences, Isparta, Turkey

Parasuraman Padmanabhan Nanyang Technological University, Singapore, Singapore

R. Parameshwaran Department of ECE, National Institute of Technology, Tiruchirappalli, India

Rizwan Patan Department of Computer Science and Engineering, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India

J. Prassanna School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India

Robbi Rahim Department of Management, Sekolah Tinggi Ilmu Manajemen Sukma, Medan, Indonesia

Nikhil Ramesh PSG College of Technology, Coimbatore, Tamil Nadu, India

N. Ravitha Rajalakshmi Department of Information Technology, PSG College of Technology, Coimbatore, Tamil Nadu, India

Antonello Rizzi Department of Information Engineering, Electronics and Telecommunications (DIET), Sapienza University of Rome, Rome, Italy

Osamah Khaled Musleh Salman Faculty of Technology Mechatronics Engineering, Isparta University of Applied Sciences, Isparta, Turkey

B. Sangeetha Department of Information Technology, PSG College of Technology, Coimbatore, Tamil Nadu, India

K. R. Sekar School of Computing, SASTRA Deemed University, Thanjavur, India

Amira Serifovic Trbalic Faculty of Electrical Engineering, University of Tuzla, Tuzla, Bosnia and Herzegovina

M. Sinan Basarslan Doğuş University, Istanbul, Turkey

R. Vidhyapriya Department of Information Technology, PSG College of Technology, Coimbatore, Tamil Nadu, India

Cihan Yalcin Institute of Education Sciences, Computer Engineering Department, Suleyman Demirel University, Isparta, Turkey

Chapter 1
Fusion of Deep Learning and Image Processing Techniques for Breast Cancer Diagnosis

V. Ajantha Devi and Anand Nayyar

Abstract Deep learning has the capacity to gain great accuracy of diagnosing of numerous types of cancers, along with lung, cervical, colon, and breast cancer. It builds an efficient algorithm based on multiple processing (hidden) layers of neurons. Manual assessment of Cancer using Medical Image (CT images) requires expensive human labors and can easily cause the misdiagnose of any type of cancer. The Researcher focus on automatically diagnosing cancer by using the deep learning technique. Breast cancer is a particularly common sickness among women and maximum associated cause of female mortality. The survival rate of breast cancer patients can be expanded with the aid of powerful treatment, which can initiate upon early prognosis of the disease. This chapter introduces Deep Learning under medical image processing to analysis and diagnosis of Cancer (Ehteshami Bejnordi et al., in Deep learning-based assessment of tumor-associated stroma for diagnosing breast cancer in histopathology images, pp. 929–932, 2017 [1]). Identification of most cancer might facilitate in sparing a massive wide variety of lives over the globe community and deep neural networks may be correctly used for intelligent image analysis. The essential structure of how this deep learning takes a shot at medical image processing (Litjens et al. in A survey on deep learning in medical image analysis, 2017, [2]; Rezaeilouyeh et al. in J Med Imaging 3(4):044501, 2016 [3]) is furnished in this research, i.e., pre-processing, image segmentation and post-processing. The following piece of this part depicts the rudiments of the field of disease conclusion, which incorporates steps of malignant growth determination followed by the regular arrangement strategies utilized by specialists, giving a verifiable thought of disease grouping methods to the readers. Next an attempt has been made to classify the extracted features from mammograms as benign or malignant by using Convolutional neural network (CNN)

V. Ajantha Devi
AP3 Solutions, Chennai, India
e-mail: ap3solutionsresearch@gmail.com

A. Nayyar (✉)
Graduate School, Duy Tan University, Da Nang 550000, Viet Nam
e-mail: anandnayyar@duytan.edu.vn

Faculty of Information Technology, Duy Tan University, Da Nang 550000, Viet Nam

U. Kose and J. Alzubi (eds.), *Deep Learning for Cancer Diagnosis*,
Studies in Computational Intelligence 908,
https://doi.org/10.1007/978-981-15-6321-8_1

(Cireşan in Mitosis detection in breast cancer histology images with deep neural networks. Springer, Berlin, pp. 411–418, 2013 [4]; LeCun et al. in International symposium on circuits and systems, pp. 253–256, 2010 [5]; Huynh et al. in J Med Imaging 3(3):034501, 2016 [6]) is applied to classify cancer using optimal features obtained from cell segmented images. Performance improvised of the approaches by varying various parameters is studied.

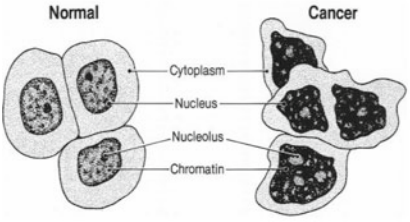
Keywords Deep learning · Medical image processing · Image preprocessing · Image segmentation · Image classification · Object detection · Post processing

1.1 Cancer

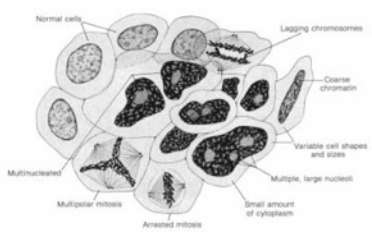
The cellular growing in living organisms can occur through ordered processes (hyperplasia, dysplasia and metaplasia) or disordered processes (neoplasia) [7]. Cancer 1 is a common designation for more than one hundred diseases which share the characteristic of accelerated and disordered cellular replication.

As an abnormal cellular division process [8], the neoplasia produces an accumulation of abnormal cells, a tissue mass [9, 10] ordinarily called tumor (or neoplasm). Figure 1.1a shows a simplified cellular structural organization in normal and cancerous cells. Disease cells, described as huge size of the core contrasted with the absolute cell size, little cytoplasm, numerous cores, different and huge nucleoli, and coarse chromatin, which can be classified as benign [11] or malignant. In Fig. 1.1b cancer cells are illustrated undergoing cell division, showing abnormal mitosis.

Other types of neoplasm are characterized by aggressive behavior, showing very fast growth, invading adjacent tissues and organs (local destructive capacity) and, eventually, even migrating to non-contiguous and distant sites (metastasis) [12]. This aggressive type of tumor is named malignant and cancer is a synonym. In order to feed



(a) Structure of normal and cancer cells



(b) Illustration of normal and cancer cells side-by-side. Cell division process on cancerous cells are leading to nuclear abnormalities.

Fig. 1.1 Major cellular characteristic structures in normal and cancer cells [10]

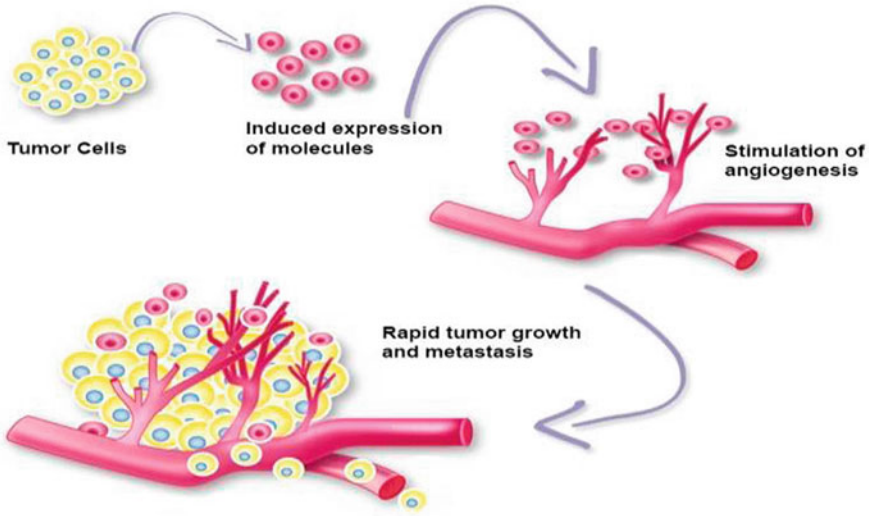


Fig. 1.2 Schematic presentation of tumor-induced angiogenesis

the accelerated growth of its cancerous cells, the tumor itself releases substances to keep blood vessels opened and it also induces the formation (angiogenesis-inducing molecules) of brand-new vessels, in a process known as neoangiogenic or tumoral angiogenesis. Acquisition of capacity of angiogenesis by cancer cells is considered the most critical step in tumor growth and metastasis [13]. Figure 1.2 shows a simplified schematic view of a tumor-associated angiogenesis.

1.1.1 Pathology

Congregating essential science and clinical practice, pathology is centered around the investigation of basic and useful changes in cells, tissues and organs brought about by ailments, being a discerning systematization to help the analysis, anticipation and treatment of such illnesses [7, 12].

Considering sample types, pathology can be divided into two main branches: (a) histopathology and (b) cytopathology.

Histopathology

Histopathology is the study of illness indicatives using by microscopic inspection of tissue samples prepared and fixated onto glass slides. These samples came from puncture biopsies or surgical excisions.

A standard procedure in histopathology is the visual analysis [14, 15] of tissue sections under the light microscope. However, tissues are normally too thick for

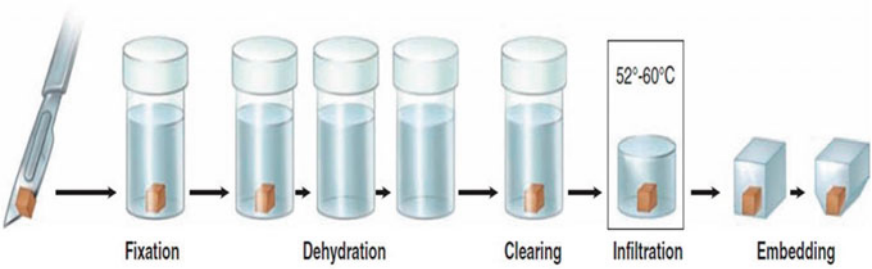


Fig. 1.3 Schematic presentation of sectioning fixed and paraffin-embedded tissue [16]

light to pass through them and they must be sliced to obtain appropriate thin translucent sections for the microscopic examination. The whole procedure, from fixation to observing a tissue in a light microscope, may take from half to two-and-half days [16]. The principle objective is to save the first tissue structure and the subatomic piece. The preparation includes the following steps: (a) fixation, (b) dehydration, (c) clearing, (d) infiltration, embedding and (f) trimming. Figure 1.3 shows the fundamental advances utilized in tissue groundwork for light microscopy.

First (a), solutions of chemicals conserve the proteins and inhibit the action of degradative enzymes. Then, all the tissue water is removed by the use of increasingly concentrated alcohol (70–100% ethanol) solutions (b), the alcohol is removed (c), the tissue is immersed in melted paraffin (d) and placed in a small mold in order to harden (e). Finally, the resulting paraffin block is trimmed to expose the tissue (f) and a microtome is used for sectioning the block.

Cytopathology

Cytopathology is focused on study and diagnosis of diseases at the cellular level [10], analyzing the structure, the functionality and the chemistry of cells (Fig. 1.4).

It is possible see a pinkish red color identifying the cytoplasm [17, 18] and the nuclei highlighted in a darker tone of blue. Cytologic exams are extremely useful in malignant neoplasia diagnosis and their precursor lesions, as well as detecting the presence of infectious and parasitic agents [19].

Normally, specimen to cytology tests are collected from patients using minimally-invasive biopsy methods, such as smears, scrapes, puncture, centrifugation of liquids and others.

1.1.2 Breast Cancer

Like different sorts of malignant growth, breast cancer disease is the irregular, quick and unordered multiplication of cells, right now, mammary tissue [20]. Individual

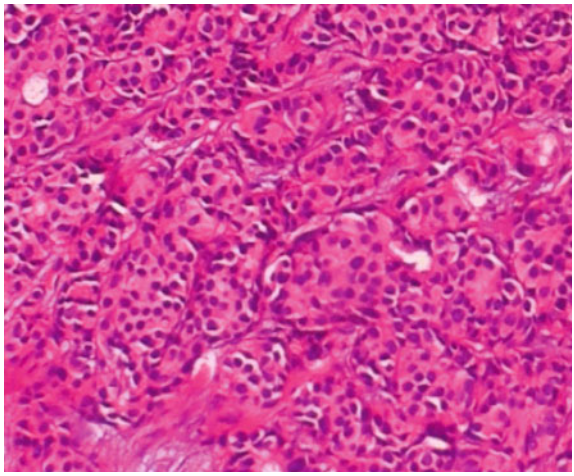


Fig. 1.4 Detail of HE sections of a ductal carcinoma (at 100× magnification)

genetic mutations—caused by several factors—are responsible for this disease that can begin in different breast regions.

Female Breast

The female mammary gland is an organ that has a dynamic behavior: demonstrates morphologic alteration throughout the reproductive life cycle (menstruation, pregnancy, lactation, menopause, etc.) and the age of women [21, 22]. Contrasting to other glands, the breast is functional during the lactation period.

In general, in mammal animals, including the human species, the mammary gland is constituted by lobes, lobules, milk ducts, connective tissue, fat, blood vessels and lymphatic vessels. Figure 1.5 presents the main anatomical structures [23] of the female breast (in cross section).

Forming the normal female breast there are between 15 and 20 independent lobes or segments [24], separated by fibrous tissue, radially distributed from the nipple. The lobe is a well-defined part of an organ (the brain, the breast, the kidney, the liver, the lung, etc.), delimited by sulci, fissures, connective tissues or other anatomic structures [23]. A lobe is visible without a microscope. Each breast lobe is composed of many tiny lobules, at the end of which there are sacs (alveoli) where milk is produced in response to hormonal signals. Lobules are connected to the nipple through thin tubes (diameter of 1–2 mm) which are the milk ducts that carry milk from the alveoli toward the dark area of the skin in the center of the breast (areola). From the areola, the ducts join together into larger ducts (up to 4 mm) ending at the nipple.

The fibrofatty tissue, fibrous connective and adipose tissue, forming the major components of the breast tissue, filling spaces between lobules and ducts. The generic name stroma is given to this tissue type. Connective tissues and ligaments provide support to the breast and give it its shape and volume. Younger women might have denser and less fatty breast tissue, different from older women. The breast itself has no muscle tissue, it lies underneath the breasts, separating them from the chest wall.

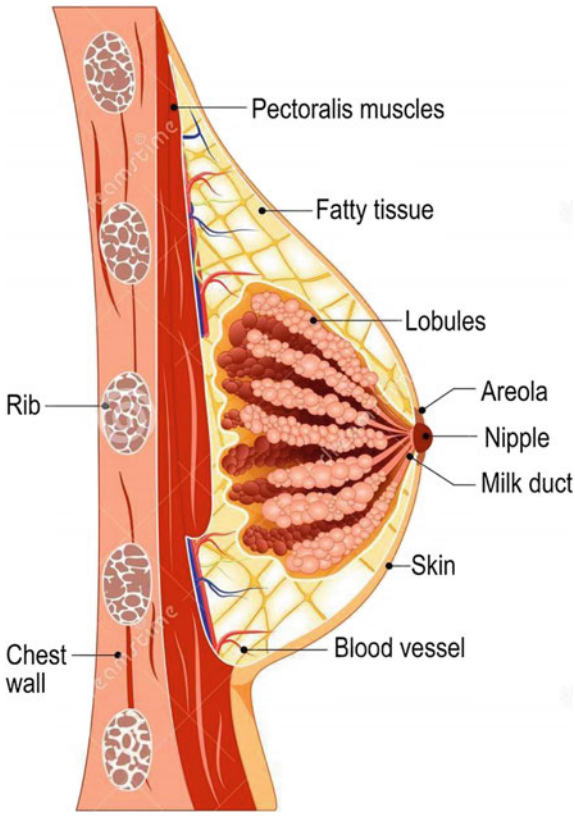
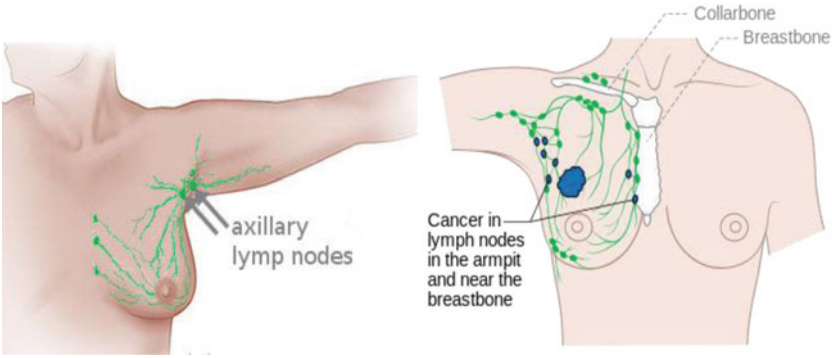


Fig. 1.5 Normal anatomy of a female mammary gland, in a cross-section scheme

In the breast lymphatic vessels are still observed (Fig. 1.6), which transport a colorless fluid, rich in defense cells, named lymph. Distributed thru the lymphatic system there are small structures called lymph nodes or lymphatic nodes. These bean shaped structures store lymphocytes A great part of the breast lymphatic vessels conduct towards the lymph nodes located in the axilla (axillary lymph nodes), highlighted in Fig. 1.6a. If cancer cells reach these lymph nodes (Fig. 1.6b), the probability that the disease has spread to other organs rises considerably [12].

Male Breast Cancer

In general, testosterone hormone causes involution of male mammary organ. An ordinary male breast cancer [27] is fundamentally made out of ductal structures inside collagenized stroma, with no or uncommon lobular components contrasted with the female breast [12, 19] cancer. Albeit uncommon, men can likewise be affected with breast cancer, representing under 1% of all instances of the infection. Hazard elements, pathology, and anticipation of male breast cancer growth medicines are very like those saw in the female populace, anyway breast cancer disease in men [12, 28–30] is frequently analyzed at later stages.



Lymphatic vessels of female breast and axillary lymph nodes. Breast cancer invading lymph nodes.

Fig. 1.6 Lymphatic vessels of a female breast and axillary lymph nodes [25, 26]

The medications for breast cancer disease in men are like those in ladies [11], except for careful choices. Given the breast cancer volume and tumor area, the standard system for men is to have a mastectomy medical procedure, instead of a lumpectomy. Male breast cancer diseases [27, 28, 30] are typically hormone receptor positive tumors and hormonal treatment is likewise a typical piece of the fundamental treatment. Hereditary testing ought to be considered for men who create breast cancer disease. The endurance rates and visualizations for men are not on a par with for ladies. Men have a 25% higher death rate than ladies. As referenced already, this is accepted to be expected partially to men getting analyzed at a later phase of the infection.

1.2 Diagnosis of Breast Cancer

The primary tests for the underlying analysis of the breast cancer are

- imaging tests
- malignancy, grading and tumor staging, can only be established through the biopsy
- the suspected area, which is analyzed by pathologists using anatomopathological exams.

1.2.1 Imaging Exams

The primary imaging tests [31] applied in breast cancer malignant growth recognition [4, 6] include advances, for example, demonstrative mammograms (x-ray),

Magnetic Resonance Imaging (MRI), breast cancer ultrasound (sonography), and thermography.

Mammography [14] is the present standard test for breast cancer disease screening, just as for analytic examination of findings at the physical (manual) assessment [14, 15, 32]. Indeed, mammographic screening has moved the range of breast cancer [33] pathology away from for the most part enormous tumors, effortlessly envisioned and effectively obvious, toward ever littler and every now and again noninvasive tumors [34].

Mammography is performed by a hardware which uses low-vitality X- ray to create high-goals radiological [35] images of the breast inward structure. The mammographic images are a reflection of the breast life structures and its incidental change by the neurotic procedures. The mammography image must present a high-differentiate goal, low clamor and follow severe conventions, taking into consideration legitimate differentiation between ordinary tissue and potential variations from the norm. In Fig. 1.7 a mammogram featuring a non-substantial tumor is appeared. By and large, the advanced mammography has significant benefits over customary mammography, including decoupling of image procurement and capacity.

In the early 1990s, the digital mammography [14] became widely used to automatically detect breast lesion [36] areas which were indicative of cancer [9, 37, 38]. Therefore, asymmetry tests between the left and right breasts, and executed pattern recognition algorithms are performed to evaluate breast tissue texture [39] (microcalcifications, dense structures, lesions).

MRI is a test on imaging, complementary to sonography and mammography, which apply radio-frequency pulses (instead X-rays) and electromagnetic waves to produce the images. Routinely for cancer detection and presurgical evaluation,

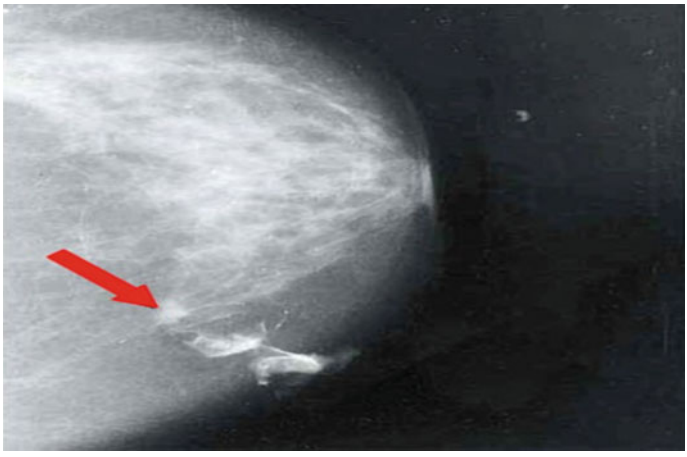


Fig. 1.7 Mammogram showing non-palpable tumor (arrow) near right border of mammary parenchyma

this test is done after an intravenous application in the patient of a gadolinium-based contrast agent [35, 40]. The contrast agent allows the enhancement of the visualization of certain structures, and highlights neovascularization in the case of mammary tumors, as can be seen in Fig. 1.8.

In (A) breast image pre-contrast application. In (B, C) breast image post applied contrast. Note in (B) the highlighting of the tumor and in (C) the breast tissue neovascularization inducted by the tumor is also highlighted.

MRI test also allows detailed evaluation of nodes, because it reaches deeper regions of the breast tissue [41]. It is extensively used to monitor the integrity of breast implants and recommended for screening in high-risk patients, such as those with a confirmed or suspected family history, with known genetic predisposition to breast cancer or who have already been affected by the disease [42].

The medical thermography, in particular Infrared Thermograph (IRT), is a rapid, passive and non-invasive method which has been successfully used to diagnose disorders [1, 33] such as breast cancer, diabetes, neuropathy and peripheral vascular issues [43]. Since 1982 thermography is approved by the FDA as a complementary exam to mammography for breast cancer diagnosis [31]. However, the FDA itself emphatically alerts that thermography is not a substitute for mammography and should not be used by itself for breast cancer screening or diagnosis [1, 33].

Breast thermography [44] or Digital Infrared Imaging (DII) [45] is a procedure to map the thermal emission of the skin over the area of the breasts, in order to show thermal asymmetries that are indicative of the presence of cysts, infections, breast cancer or other diseases. While other tests such as mammography and ultrasound detect anatomical changes already established in breast, thermography [44] has the

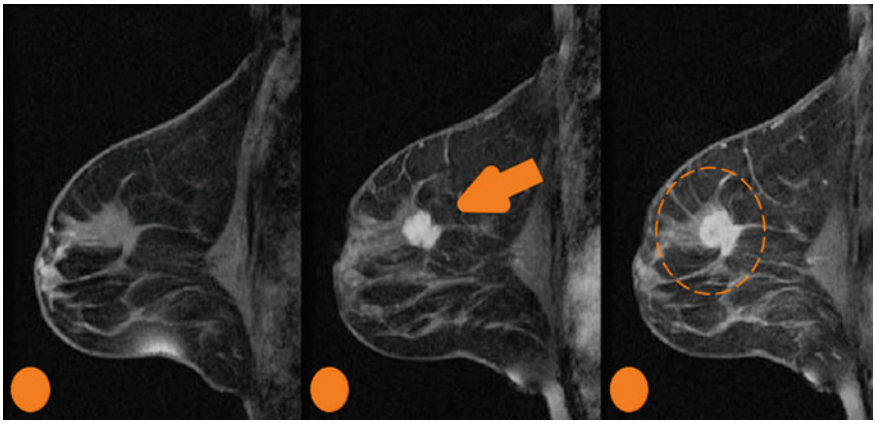
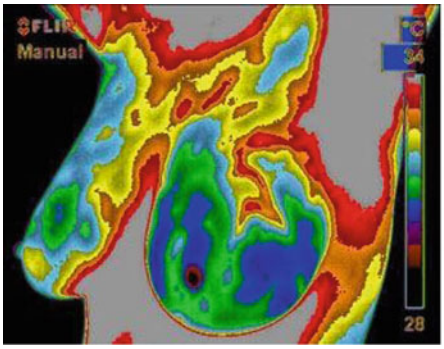
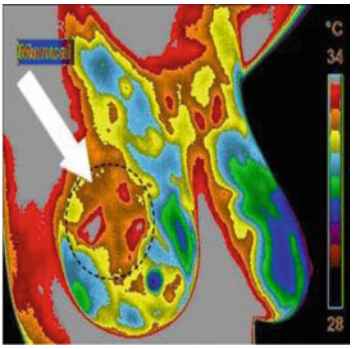


Fig. 1.8 Sequence of breast MRI images [40]



Left breast presenting thermal pattern compatible with normal vascularization



Right breast showing apparent temperature rise and vascularization (angiogenesis). The lump in the upper outer quadrant of the breast is highlighted.

Fig. 1.9 Typical breast thermography image [44, 46]

advantage of being a functional examination, which studies metabolic and vascular abnormalities of the breast (Fig. 1.9).

Anatomopathological Exams

With respect to malignant growth, if the physical assessment (contact) identifies tangible bumps or imaging tests find suspicious tissue territories, anatomopathological tests are required. Pathological exam analyses cellular and tissue microscopic alterations present in samples collected from biopsies or surgeries. The pathologist responsible for pathological examination can make the correlation with clinical and imaging tests. Generally, the pathological diagnosis is considered definitive, but it may be inconclusive, due to limiting factors such as insufficient material collected or even if the collected sample is unrepresentative of the suspicious lesion.

Currently, pathologists determine the tumor grading [14] by assessing the spatial organization of the tissue (e.g., distribution of cancer cells, nuclei morphological properties, interaction with the stroma, etc.). These parameters are evaluated in small sample regions of the microscopic slide in order to give a score considering some "scoring system" such as Nottingham Histologic Score System1 Complementing the final decision, for prognosis and clinical intervention, respective hormone receptor status by IHC is also analyzed in the IHC-stained sections.

Breast Biopsies

Once screening tests, such as mammography or breast ultrasound, have found suspicious changes, it is recommended to biopsy the region. Breast biopsy is the removal of a small amount of breast tissue for pathologic evaluation to determine whether it is cancerous or non-cancerous (benign). Surgical biopsy techniques [47] can be done by

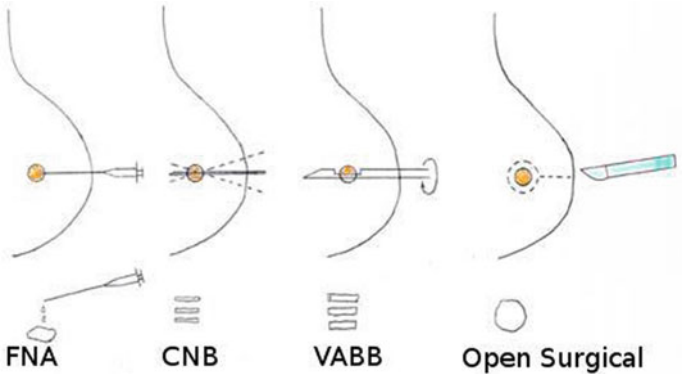


Fig. 1.10 Schematic representation of different breast biopsy types [47]

incision or excision. The more appropriate biopsy method depends on several factors, such as, how suspicious the lesion is; the size, shape, and location of the lesion; the number of abnormalities present; the patient's medical history, etc. Most often, a needle biopsy is done first and then, if needed, a surgical biopsy is done. Figure 1.10 shows a schematic representation of the four main breast biopsy techniques. FNA, CNB, and VABB are percutaneous methods [48], i.e., a needle is inserted through the skin [47].

Surgical biopsy yields the largest breast tissue sample of all the breast biopsy methods, and the accuracy of a diagnosis is better. However, the surgical biopsy method is much more invasive than percutaneous methods, it requires stitches and can leave a scar.

FNA was introduced in 1930 and became popular in the 70s. In this technique a very thin needle (20–21G 11) is used. In general, the needle used during FNA is smaller than a needle that is normally used to collect blood samples [19]. A 10 or 20 ml syringe attached to the needle allows to aspirate fluids and clusters of cells from the puncture site. It is a fast and low-cost procedure usually requiring no anesthesia. Several needle insertions are usually needed to guarantee that an adequate tissue sample is taken. The collected material is deposited on slides (samples are smeared on a microscope slide) for a cytological study.

1.3 Deep Learning and CNNs

Deep learning models as a rule embrace progressive structures to interface their layers. The yield of a lower layer can be viewed as the contribution of a higher layer by means of straightforward direct or nonlinear computations. These models can change low-level highlights of the information into elevated level dynamic highlights. Owning to this trademark, Deep learning models [49] can be more grounded than shallow AI models in include portrayal. The exhibition of conventional AI

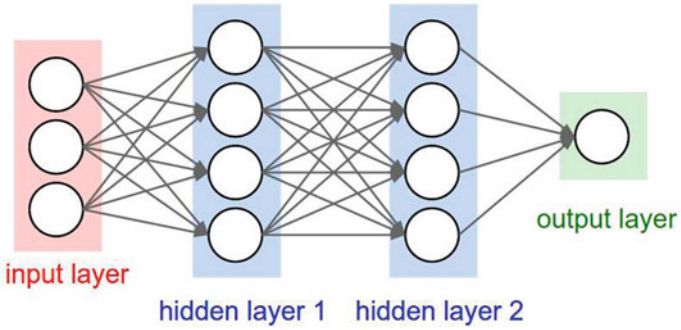


Fig. 1.11 The architecture of Deep Neural network [59]

strategies for the most part depends on clients' encounters, while Deep learning approaches depend on the information. With the advancement of PC innovation, PCs' presentation is quickly improved. In the meantime, data on the Internet is likewise regurgitating. These elements give a solid driving force to Deep figuring out how to create and make Deep learning become the common technique in AI. The idea of Deep learning was advanced in 2006 from the start. Deep learning began moderately late yet grew quickly at home.

Deep learning has had a wild development in computer vision, such as object detection, object tracking, and image segmentation. Object detection aims to recognize [4, 24] a class of objects from a large number of images. The traditional object detection methods mainly include candidate region selection, feature extraction and classification [3, 50]. This manual feature extraction method needs users to design what features they should extract. And these processes are often high-cost and time-consuming. Deep learning has the ability of unsupervised feature learning [50–53], and it can extract the features of images without any human intervention Thus, it is gradually attracted more and more attention by researchers.

DL has indicated extraordinary portrayal learning can find powerful highlights just as their mapping from information for given undertakings. In other words, with artificial neural networks of multiple nonlinear layers, referred to as deep learning architectures, hierarchical representations of data [54–56] can be discovered with increasing levels of abstraction. The deep learning architectures classify into four groups. They are Deep Neural network [57, 58]. Convolutional neural networks, Recurrent Neural network and emergent architectures. The deep learning architectures [59] is shown in Fig. 1.11.

1.3.1 CNN Architecture

The simplest example of a deep learning model is the feedforward deep network or MLP. It consists of a collection of neurons connected in an acyclic graph organized

into a succession of neuron layers. A neuron computes a function on inputs from the preceding layer and passes the result (neuron's activation) on to outputs in the succeeding layer. Within each layer, all neurons compute the same function, but individual neurons may have distinct sets of inputs and outputs and may assign different weights to their inputs.

The MLP receive as input a single vector in its input layer. Then, this input is transformed thorough a sequence of intermediary layers (hidden layers). In regular NN each hidden layer is fully-connected. The fundamental structure of CNNs consists of Convolutional layers, nonlinear layers and pooling layers shown in Fig. 1.12.

CNNs are essentially layers of convolutions followed by subsampling and completely associated layers. Convolutional layers, pooling layers, and completely associated layers are three head neural layers of a CNN [4, 5, 60]. Each layer has an alternate assignment. Convolutions and subsampling layers fill in as highlight extraction layers while a completely associated layer order which class current information has a place with utilizing separated highlights. The assignment of a pooling layer is limiting the elements of highlight maps and system parameters. Pooling layers' calculations think about neighboring pixels, so they are change stable. A forward advance and a regressive advance are utilized to prepare the system. The forward advance plans to delineate the info picture with the present parameters (loads and inclination) in each layer.

Then again, assessment of the misfortune cost with the ground truth names is produced using the expectation yield. At that point the regressive advance computes the inclinations of each parameter with chain governs based on the misfortune cost. Every parameter is recalculated by inclinations, and prepared for the following forward assessment. A completely associated layer is a convolutional layer with w × h channels where w and h are the spatial elements of the volume in the past layer and no cushioning, i.e., channels spread the whole volume, bringing about component maps with size 1 × 1. The yield layer of a convolutional arrange is in every case completely associated with however many component maps as would be prudent classes. It empowers us to take care of forward the neural system into a vector with a predefined length. A case of ordinary CNN [38, 60] engineering with two component stages is appeared in Fig. 1.12.

Convolution with an assortment of filters, similar to the educated filters (additionally named highlight maps or initiation maps) in Fig. 1.12. improves the portrayal: at the first layer of a CNN the highlights go from singular pixels to basic natives,

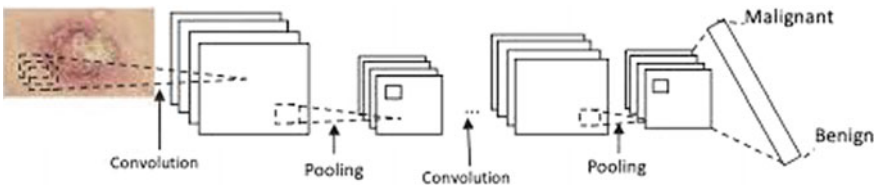


Fig. 1.12 A case of a run of the mill CNN design with two component stages

similar to flat and vertical lines, circles, and fixes of shading. Rather than ordinary single-channel picture handling filters, these CNN filters are processed over the entirety of the information channels. Because of its interpretation invariant property, convolutional filters yield a high reaction any place an element is identified.

The addition of a pooling (subsampling) layer between two progressive convolutional layers is normal. The primary target of this training is to continuously lessen the spatial size of the portrayal. Along these lines, lessening the quantity of parameters and calculations required by the system enables the overfitting to control. The pooling layer down examples the volume spatially in every profundity cut of the info volume autonomously. Accordingly, the pool administrator resizes the contribution along the width and the stature, disposing of the initiations. Practically speaking, the maximum pooling capacity, which applies a window capacity to the info fix, and registers the most extreme in that area, has indicated better outcomes.

In a completely associated layer, neurons have full associations with all actuations in the past layer and their initiations can be registered utilizing a lattice increase followed by an inclination offset. This sort of layer is standard in an ordinary NN. The last completely associated layer holds the net yield, for example, likelihood circulations over classes.

1.4 Experimental Methodology

Generally, the Experimental methodology comprises of four stages as shown in Fig. 1.13:

1. Image pre-processing [61]: It is one of the most significant strides in the methodology framework and decides its exactness. The objective is to diminish spot clamor and improve image quality without devastating the significant highlights of the images. Speckle commotion is a type of multiplicative clamor produced by various disperses with irregular stage inside the goals cell of ultrasound pillar. The principle past related-work with respect to dot clamor decrease procedures classified into three gatherings: filtering systems, wavelet space methods, and aggravating methodologies.
2. Image segmentation: Its goal is separating the image into non covering locale. In this way, it will separate the zone of enthusiasm from the foundation. It is viewed as one of the most difficult undertakings of picture handling and example acknowledgment. By and large, histogram Thresholding and dynamic shape model are two most famous methods in related research.
3. Feature extraction: It plans to find a lot of interesting highlights of breast cancer disease sores that can recognize the sore and non-injury.
4. Classification: It chooses whether the suspicious area is amiable or harmful. Breast Cancer (mammography) Images are acquired [54, 62]. Each image is 'pre-processed' as in Fig. 1.14 to prepare its ROI selection. Framework for every ROIs is produced utilizing the pixel esteems as in Fig. 1.15.

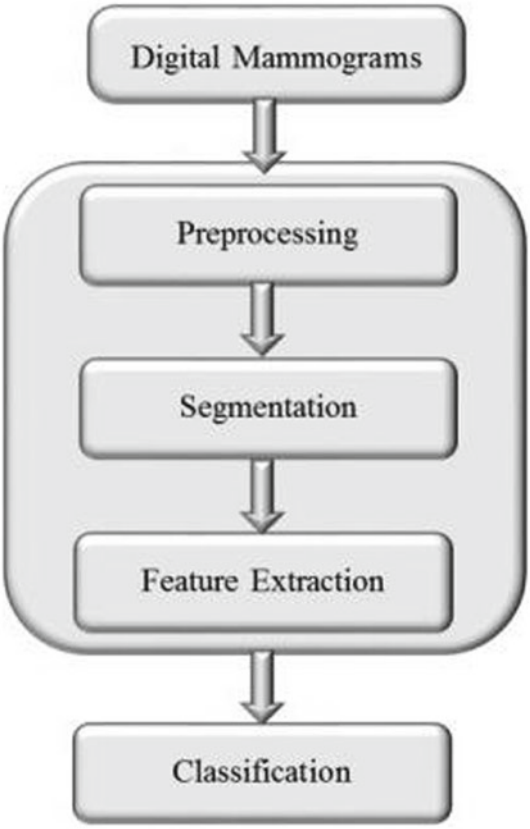


Fig. 1.13 Block diagram of the experimental methodology

Highlight esteems (tumor territory, mean pixel esteem, and so on) are removed, which are broke down for exact evaluation as in Fig. 1.16. GPU-helped superior registering is applied for quicker and less expensive examination/arrangement.

5. Step 1: PROBLEM STATEMENT

Predicting if the cancer diagnosis is benign or malignant based on several observations/features 30 features are used, examples such as radius (mean of distances from center to points on the perimeter), texture (standard deviation of gray-scale values), perimeter, area, smoothness (local variation in radius lengths), compactness (perimeterˆ2/area − 1.0), concavity (severity of concave portions of the contour), concave points (number of concave portions of the contour), symmetry, fractal dimension ("coastline approximation"—1).

6. Step 2: IMPORTING DATA

Breast Cancer Wisconsin (Diagnostic) Database is openly accessible and is normally used to consider the Breast Cancer growth characterization issue. This dataset

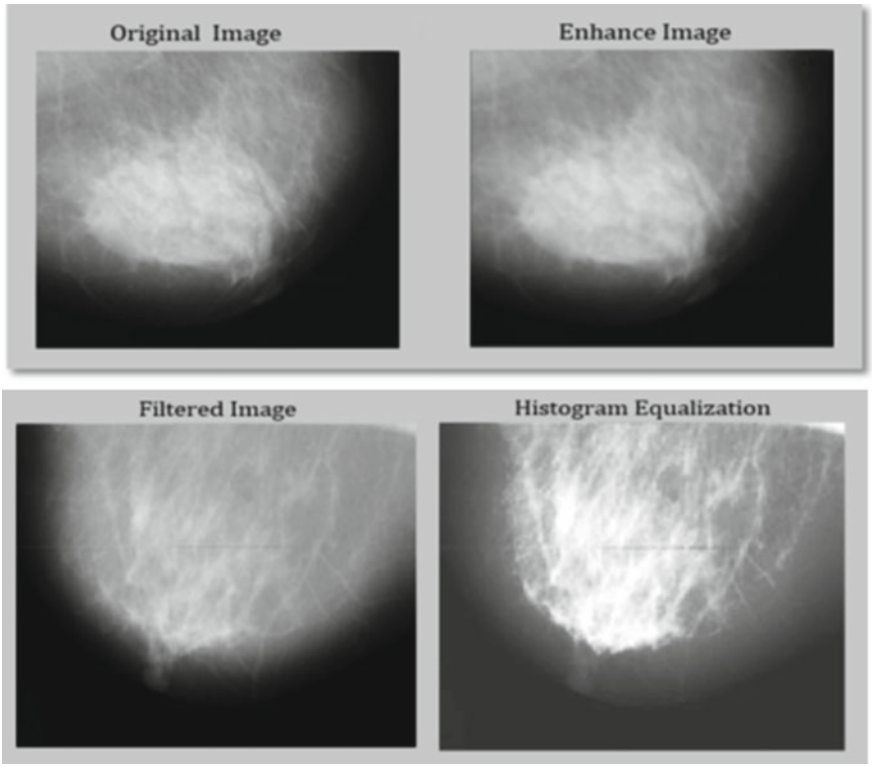


Fig. 1.14 Preprocessing of the image

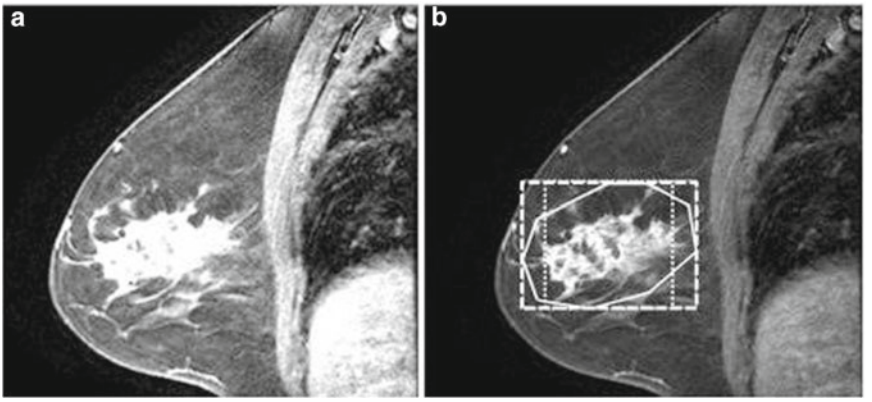


Fig. 1.15 Analysis for mammography of breast cancer: **a** original image; **b** a user-defined region-of-interest (ROI)

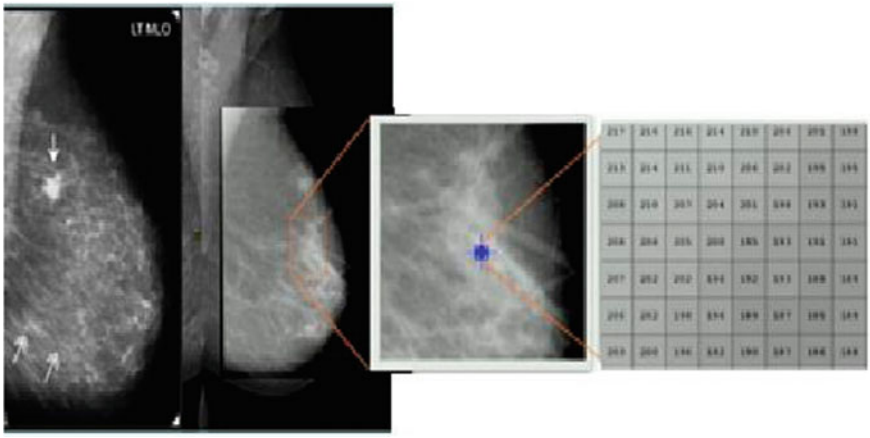


Fig. 1.16 Feature extraction

contains 569 examples each falling inside two primary classes: benign or malignant. The benign subset contains 357 examples and the malignant subset contains 212 examples. The examples are gathered from 30 patients with various amplification factors including 40×, 100×, 200×, 400×. A portion of the model pictures with a 400× amplification factor is appeared in Fig. 1.17. Each class has four sub class, the four sorts of kind malignancy [6] are Adenosis (A), Fibroadenoma (F), Tubular Adenoma (TA), and Phyllodes Tumor (PT). The four subclasses of

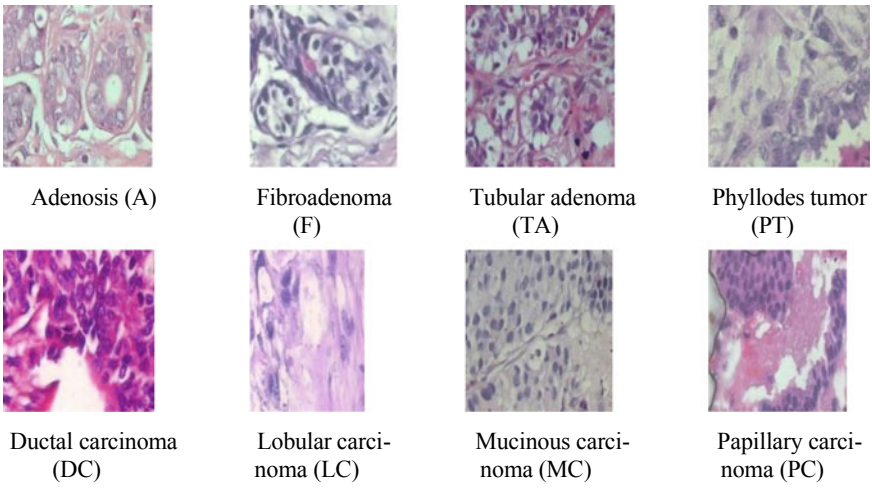


Fig. 1.17 The first row shows the four types of benign tumors, and the second row shows the malignant tumors. The magnification factor of these images is 400×

threatening malignant growth are Ductal Carcinoma (DC), Lobular Carcinoma (LC), Mucinous Carcinoma (MC), and Papillary Carcinoma (PC).

7. STEP 3: VISUALIZING THE DATA

The correlation between the variables are Strong correlation between the mean radius and mean perimeter, mean area and mean perimeter (Fig. 1.18).

8. STEP 4: MODEL TRAINING (FINDING A PROBLEM SOLUTION)

The pre-trained model may appear to resemble a black box from the start; however, it is effectively interpretable. So as to make the pretrained models progressively interpretable, we have envisioned the component vectors, yields of each layer as shown in Fig. 1.19.

The underlying layers deal with the large highlights like edges, shapes, and so on:

Now the model is fit to be retrained (with the new classifier). Number of ages, optimizer, learning rate, energy, and the remainder of initializations are performed

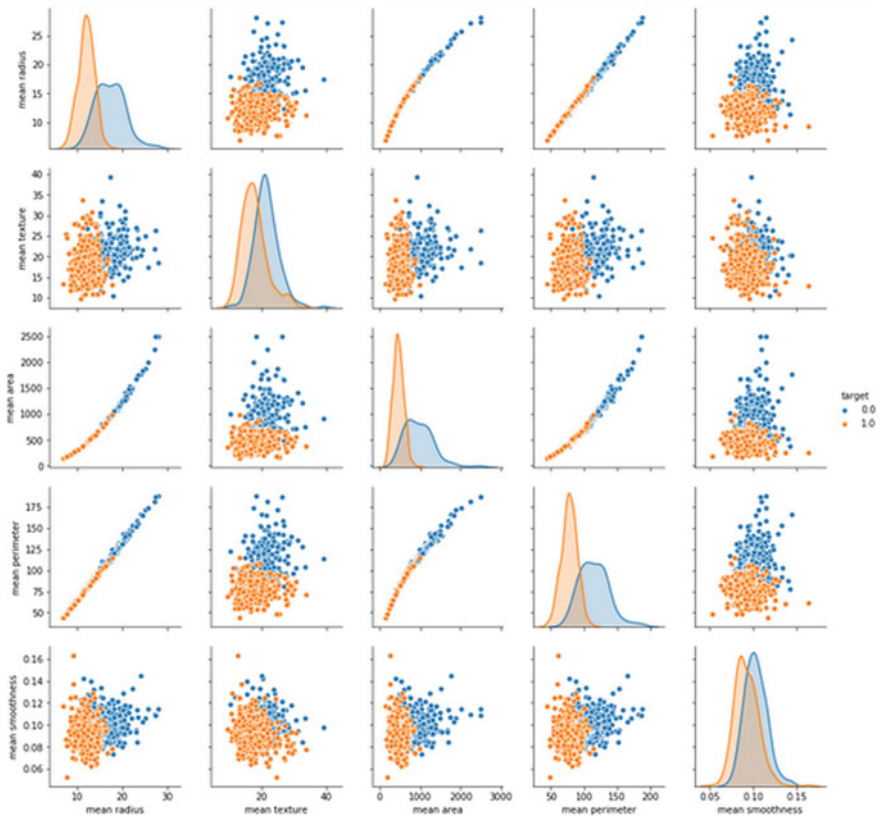


Fig. 1.18 Data visualization

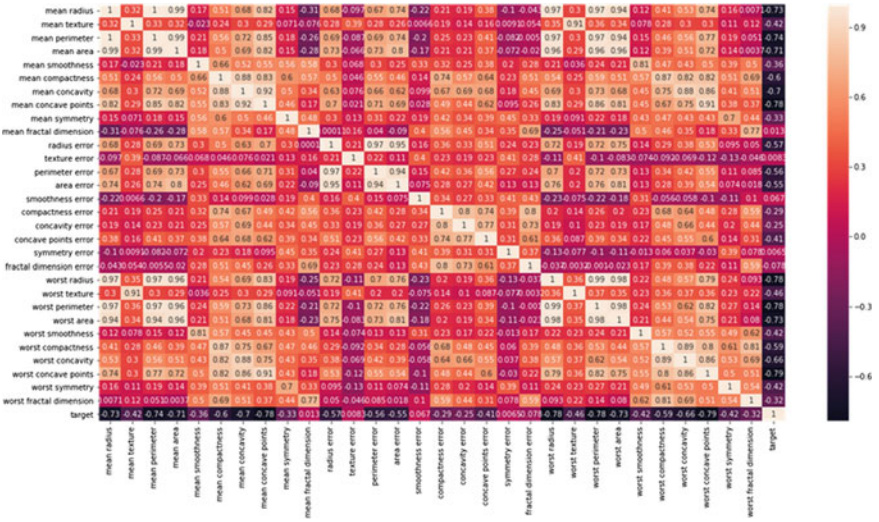


Fig. 1.19 Pre-trained model

lastly model is called with its capacity signature esteems which were initialized before.

9. STEP 5: EVALUATING THE MODEL

Confusion Matrix

Confusion Matrix is a significant metric while dissecting misclassification. Each line of the framework speaks to the examples in an anticipated class while every segment speaks to the occurrences in a real class. The diagonals speak to the classes that have been accurately characterized. This aide as we probably are aware which classes are being misclassified as well as what they are being misclassified as in Fig. 1.20.

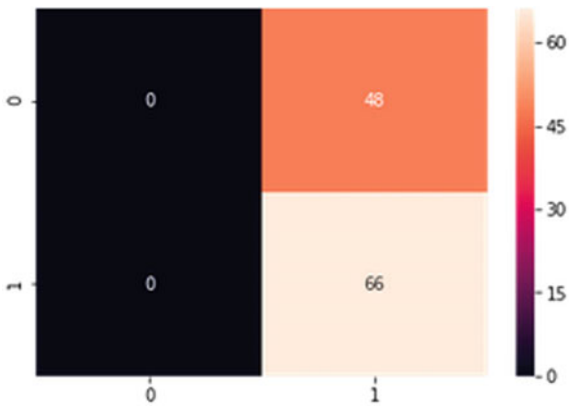


Fig. 1.20 Confusion matrix

10. STEP 6: IMPROVING THE MODEL

mean radius	6.981000
mean texture	9.710000
mean perimeter	43.790000
mean area	143.500000
mean smoothness	0.052630
mean compactness	0.019380
mean concavity	0.000000
mean concave points	0.000000
mean symmetry	0.106000
mean fractal dimension	0.049960
radius error	0.111500
texture error	0.362100
perimeter error	0.757000
area error	6.802000
smoothness error	0.001713
compactness error	0.002252
concavity error	0.000000
concave points error	0.000000
symmetry error	0.007882
fractal dimension error	0.000950
worst radius	7.930000
worst texture	12.020000
worst perimeter	50.410000
worst area	185.200000
worst smoothness	0.071170
worst compactness	0.027290
worst concavity	0.000000
worst concave points	0.000000
worst symmetry	0.156500
worst fractal dimension	0.055040

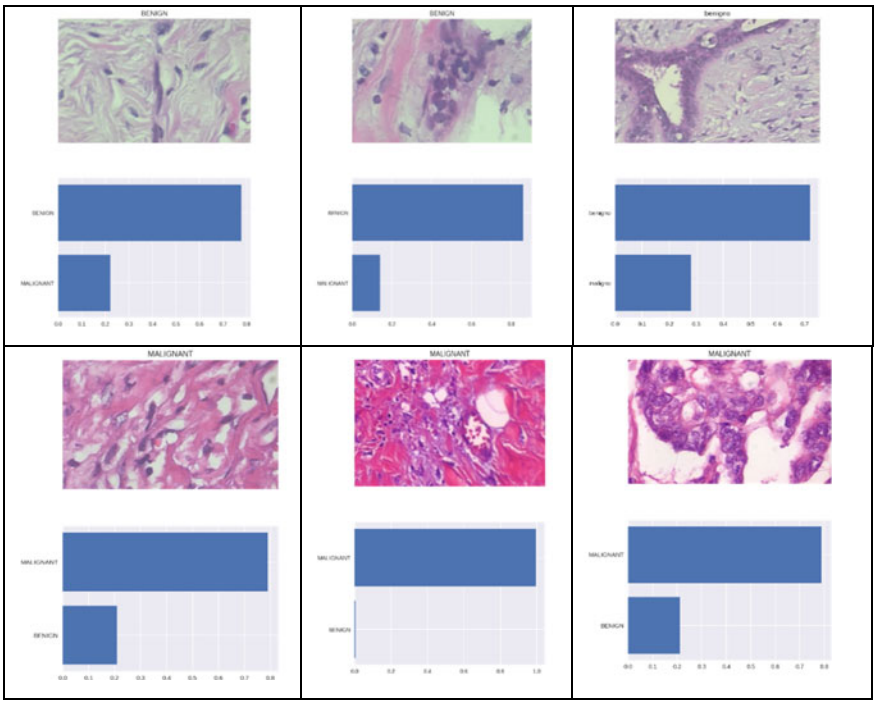
Trained Model

mean radius	21.129000
mean texture	29.570000
mean perimeter	144.710000
mean area	2355.500000
mean smoothness	0.110770
mean compactness	0.326020
mean concavity	0.426800
mean concave points	0.201200
mean symmetry	0.198000
mean fractal dimension	0.045790
radius error	2.761500
texture error	4.522900
perimeter error	21.223000
area error	518.798000
smoothness error	0.029417
compactness error	0.133148
concavity error	0.396000
concave points error	0.052790
symmetry error	0.071068
fractal dimension error	0.028890
worst radius	25.190000
worst texture	37.520000
worst perimeter	170.390000
worst area	3246.800000
worst smoothness	0.129430
worst compactness	1.030710
worst concavity	1.105000
worst concave points	0.291000
worst symmetry	0.420900
worst fractal dimension	0.152460

Testing Performance

11. STEP 7: IMPROVING THE MODEL

train accuracy: 91.62679425837321 % On training set: True Positive: 148 True Negative: 235 False Negative: 28 False Positive: 7 True Positive Rate / Recall: 84.09% Precision: 95.48% False Positive Rate / Fallout: 2.89%	test accuracy: 91.94630872483222 % On Test set: True Positive: 29 True Negative: 108 False Negative: 6 False Positive: 6 True Positive Rate / Recall: 82.86% Precision: 82.86% False Positive Rate / Fallout: 5.26%



1.5 Conclusion

The acquired outcome demonstrates that determination of appropriate ways to deal with plan a calculation for diagnosing the breast cancer growth to the precision, affectability, and false positive recognizable pieces of proof. To evacuate foundation commotion and pectoral muscle, district developing and thresholding techniques are end up being acceptable. The nature of the mammography was upgraded and Mass in mammography is removed with appropriate stamping utilization of Segmentation.

Validation on the training is performed which gives a precision of 86%. It is concluded in reason that it was useful to utilize a pre-prepared system for this utility on the grounds that the loads picked up for ImageNet dataset end up being working for the dataset we accommodated preparing. Image preprocessing is fundamental before we feed the system with the pictures.

It is basic to take note of that in the event that we had a bigger dataset, the model must be finetuned prior to require to be prepared. Since models requiring picture preparing require long handling time, it is advantageous to store the model in memory so it tends to be stacked again for use in future.

References

1. B. Ehteshami Bejnordi, J. Lin, B. Glass, M. Mullooly, G.L. Gierach, M.E. Sherman, N. Karssemeijer, J. van der Laak, A.H. Beck, Deep learning-based assessment of tumor-associated stroma for diagnosing breast cancer in histopathology images (2017), pp. 929–932
2. G.J.S. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A.W.M. van der Laak, B. van Ginneken, C.I. Sánchez, A survey on deep learning in medical image analysis. CoRR, abs/1702.05747 (2017)
3. H. Rezaeilouyeh, A. Mollahosseini, M.H. Mahoor, Microscopic medical image classification framework via deep learning and shearlet transform. J. Med. Imaging **3**(4), 044501 (2016)
4. D.C. Cireşan, A. Giusti, L.M. Gambardella, J. Schmidhuber, *Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks* (Springer, Berlin, 2013), pp. 411–418. ISBN 978-3-642-40763-5
5. Y. LeCun, K. Kavukcuoglu, C. Farabet, Convolutional networks and applications in vision, in *International Symposium on Circuits and Systems* (2010), pp. 253–256
6. B.Q. Huynh, H. Li, M.L. Giger, Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. J. Med. Imaging **3**(3), 034501 (2016)
7. A.J. Mendez, P.G. Tahoces, M.J. Lado, M. Souto, J.J. Vidal, Computer-aided diagnosis: automatic detection of malignant masses in digitized mammograms. Med. Phys. **25**, 957–964 (1998)
8. M. Kowal, P. Filipczuk, A. Obuchowicz, J. Korbicz, R. Monczak, Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images. Comput. Biol. Med. **43**(10), 1563–1572 (2013)
9. J. Dheeba, N.A. Singh, S.T. Selvi, Computer-aided detection of breast cancer on mammograms: a swarm intelligence optimized wavelet neural network approach. J. Biomed. Inform. **49**, 45–52 (2014). https://doi.org/10.1016/j.jbi.2014.01.010
10. A.H. Fischer, K.A. Jacobson, J. Rose, R. Zeller, Hematoxylin and eosin staining of tissue and cell sections. In: *Cold Spring Harbor Protocols* (2008). https://doi.org/10.1101/pdb.prot4986
11. A.S. Becker, M. Marcon, S. Ghafoor, M.C. Wurnig, T. Frauenfelder, A. Boss, Deep learning in mammography: diagnostic accuracy of a multipurpose image analysis software in the detection of breast cancer. Invest. Radiol. (2017)
12. M. Heath, D. Kopans, R. Moore, P. Kegelmeyer Jr., The digital database for screening mammography. Br. J. Psychiatry 212–218 (2000). (*Proceedings of the 5th International Workshop on Digital Mammography*)
13. S.K. Pal, S.K. Lau, L. Kruper, U. Nwoye, C. Garberoglio, R.K. Gupta, B. Paz, V. Lalit, E. Guzman, A. Artinyan, G. Somlo, Papillary carcinoma of the breast: an overview. Breast Cancer Res. Treat. **122**(3), 637–645 (2010). https://doi.org/10.1007/s10549-010-0961-5
14. S. Doyle, S. Agner, A. Madabhushi, M. Feldman, J. Tomaszewski, Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features,

in *Proceedings of the 5th IEEE International Symposium on Biomedical Imageing (ISBI): From Nano to Macro* (Paris, France), vol. 61 (IEEE, 2008), pp. 496–499. https://doi.org/10.1109/isbi.2008.4541041
15. J.E. Joy, E.E. Penhoet, D.B. Petitti (eds.), *Saving Women's Lives: Strategies for Improving Breast Cancer Detection and Diagnosis* (National Academies Press, Washington, D.C., 2005). ISBN: 978-0-309-09213-5
16. B.B. Lahiri, S. Bagavathiappan, T. Jayakumar, J. Philip, Medical applications of infrared thermography: a review. Infrared Phys. Technol. **55**(4), 221–235 (2012)
17. V.A.F. Alves, M. Bibbo, F.C.L. Schmitt, F. Milanezi, A. Longatto Filho, Comparison of manual and automated methods of liquid-based cytology: a morphologic study. Acta Cytol. **2**(48), 187–193 (2004)
18. R.K. Gibb, M.G. Martens, The impact of liquid-based cytology in decreasing the incidence of cervical cancer. Rev. Obstet. Gynecol. **4**, s2–s11 (2011)
19. D.B. Kopans, An open letter to panels that are deciding guidelines for breast cancer screening. Breast Cancer Res. Treat **151**(1), 19–25 (2015) (Online). http://www.ncbi.nlm.nih.gov/pubmed/25868866
20. K.I. Satoto, O.D. Nurhayati, R. Isnanto, Pattern recognition to detect breast cancer thermogram images based on fuzzy inference system method. Int. J. Comput. Sci. Technol. **2**(3), 484–487 (2011)
21. W. Lotter, G. Sorensen, D. Cox, A multi-scale CNN and curriculum learning strategy for mammogram classification, in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* (2017)
22. N. Rahimi, The ubiquitin-proteasome system meets angiogenesis. Mol. Cancer Ther. **11**(3), 538–548. https://doi.org/10.1158/1535-7163.mct-11-0555. http://mct.aacrjournals.org/content/11/3/538.full
23. S.S. Brandt, G. Karemore, N. Karssemeijer, M. Nielsen, An anatomically oriented breast coordinate system for mammogram analysis. IEEE Trans. Med. Imaging **30**(10), 1841–1851. https://doi.org/10.1109/tmi.2011.2155082
24. M.U. Dalmış, G. Litjens, K. Holland, A. Setio, R. Mann, N. Karssemeijer, A. Gubern-Mrida, Using deep learning to segment breast and fibroglandular tissue in MRI volumes. Med. Phys. **44**(2), 533–546 (2017)
25. Cancer Research UK (CRUK). Diagram 3 of 3 showing stage 3C breast cancer CRUK 401. 2014. url: http://commons.wikimedia.org/wiki/File:Diagram_3_of_3_showing_stage_3C_breast_cancer_CRUK_401.svg#/media/File:Diagram_3_of_3_showing_stage_3C_breast_cancer_CRUK_401.svg
26. National Cancer Institute (NCI), Male Breast Cancer Treatment (PDQ®) (2014). url: http://www.cancer.gov/cancertopics/pdq/treatment/malebreast/patient
27. Breastcancer.org. Male Breast Cancer (2014). http://www.breastcancer.org/symptoms/types/male_bc
28. A.-A. Nahid, K. Yinan, Histopathological breast-image classification using local and frequency domains by convolutional neural network. Information **19**(9), 1–26. https://doi.org/10.3390/info9010019
29. National Cancer Institute (NCI), SEER Cancer Statistics Factsheets: Breast Cancer (2014). http://seer.cancer.gov/statfacts/html/breast.html
30. A. Janowczyk, A. Madabhushi, Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. J. Pathol. Inform. **7** (2016)
31. S.W. Duffy, L. Tabar, H.H. Chen, M. Holmqvist, M.F. Yen, S. Abdsalah, B. Epstein, E. Frodis, E. Ljungberg, C. Hedborg-Melander, A. Sundbom, M. Tholin, M. Wiege, A. Akerlund, H.M. Wu, T.S. Tung, Y.H. Chiu, C.P. Chiu, C.C. Huang, R.A. Smith, M. Rosen, M. Stenbeck, L. Holmberg, The impact of organized mammography service screening on breast carcinoma mortality in seven Swedish counties. Cancer **95**(3), 458–469 (2002)
32. L. Tabár, T. Tot, P.B. Dean, *Breast Cancer: The Art and Science of Early Detection with Mammography—Perception, Interpretation, Histopathologic Correlation* (Tyeme Verlag, New York, 2004)

33. N. Bayramoglu, J. Kannala, J. Heikkilä, Deep learning for magnification independent breast cancer histopathology image classification, in *23rd International Conference on Pattern Recognition* (Cancun, Mexico), vol. 1 (IEEE 2016), pp. 2440–2445. https://doi.org/10.1109/icpr.2016.7900002
34. I. Domingues, J.S. Cardoso, Mass detection on mammogram images: a first assessment of deep learning techniques (2013)
35. S. Azizi, F. Imani, S. Ghavidel, A. Tahmasebi, J.T. Kwak, S. Xu, B. Turkbey, P. Choyke, P. Pinto, B. Wood et al., Detection of prostate cancer using temporal sequences of ultrasound data: a large clinical feasibility study. Int. J. Comput. Assist. Radiol. Surg. **35**(4), 1–10 (2016)
36. J. Arevalo, F.A. González, R. Ramos-Pollán, J.L. Oliveira, M.A.G. Lopez, Representation learning for mammography mass lesion classification with convolutional neural networks. Comput. Methods Programs Biomed. **127**, 248–257 (2016)
37. Z. Chen, A. Oliver, E. Denton, R. Zwiggelaar, Automated mammographic risk classification based on breast density estimation, in *Pattern Recognition and Image Analysis* (Madeira, Portugal), vol. 7887 (2013), pp. 237–244. https://doi.org/10.1007/978-3-642-38628-2_28
38. P. Fonseca, J. Mendoza, J. Wainer, J. Ferrer, J. Pinto, J. Guerrero, B. Castaneda, Automatic breast density classification using a convolutional neural network architecture search procedure, in *Proceedings of SPIE Medical Imaging 2015: Computer-Aided Diagnosis* (Orlando, USA), vol. 9414. Society of Photo-Optical Instrumentation Engineers (SPIE) (2015). https://doi.org/10.1117/12.2081576. ISBN: 9781628415049
39. S. Kothari et al., Pathology imaging informatics for quantitative analysis of whole-slide images. J. Am. Med. Inform. Assoc. **20**(6), 1099–1108 (2013)
40. I. Gardiner, U. Kuusk, B.B. Forster, A. Spielmann, Breast magnetic resonance imaging. BC Med. J. **47**(10), 543–548 (2005)
41. P. Carmeliet, R.K. Jain, Angiogenesis in cancer and other diseases. Nature **407**, 249–257 (2000)
42. M.E. Sanders, J.F. Simpson, Breast pathology, in *Proceedings of the 5th international*, vol. 6 (Demos Medical, New York, 2014), pp. 212–218. ISBN: 978-1-936287-68-0
43. S.K. Apple, Sentinel lymph node in breast cancer: Review article from a pathologist's point of view. J. Pathol. Transl. Med. **50**(2), 83 (2016)
44. Instituto Nacional do Câncer (INCA), Ministério da Saúde (MS). Diretrizes brasileiras para o rastreamento do câncer do colo do útero. Tech. rep. Rio de Janeiro: Instituto Nacional do Câncer (INCA) (2011)
45. N. Arora, D. Martins, D. Ruggerio, E. Tousimis, A.J. Swistel, M.P. Osborne, R.M. Simmons, Effectiveness of a noninvasive digital infrared thermal imaging system in the detection of breast cancer. Am. J. Surg. **196**(4), 523–526 (2008)
46. W.C. Amalu (ed.), Case Studies. Nov 2012. http://www.breastthermography.com/case_studies.htm
47. Breastcancer.org. Biopsy (2012). http://www.breastcancer.org/symptoms/testing/types/biopsy
48. Breastcancer.org. Types of Breast Cancer (2013). http://www.breastcancer.org/symptoms/types
49. A. Bhardwaj, A. Tiwari, Breast cancer diagnosis using Genetically Optimized Neural Network model. Expert Syst. Appl. **42**(10), 4611–4620. https://doi.org/10.1016/j.eswa.2015.01.065
50. V. Ajantha Devi, Supervised learning approach to object identification and recognition. Int. J. Pure Appl. Math. **119**(12), 15969–15976 (2018). ISSN: 1314-3395
51. B.W. Wei, Z. Han, X. He, Y. Yin, Deep learning model based breast cancer histopathological image classification, in *Proceedings of the 2017 IEEE 2nd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)* (Chengdu, China) (IEEE, 2017), pp. 348–353. https://doi.org/10.1109/icccbda.2017.7951937
52. J. Alzubi, A. Nayyar, A. Kumar, Machine learning from theory to algorithms: an overview. J. Phys. Conf. Ser. **1142**(1). (IOP Publishing)
53. F. Saeed, A. Paul, P. Karthigaikumar, A. Nayyar, Convolutional neural network based early fire detection. Multimedia Tools Appl. 1–17. (Springer USA)
54. M.R.K. Mookiah, U.R. Acharya, E.Y.K. Ng, Data mining technique for breast cancer detection in thermograms using hybrid feature extraction strategy, Quant. InfraRed Thermography J. **9**(2), 151–165. https://doi.org/10.1080/17686733.2012.738788

55. N.H. Motlagh, M. Jannesary, H. Aboulkheyr, P. Khosravi, O. Elemento, M. Totonchi, I. Hajirasouliha, Breast Cancer Histopathological Image Classification: A Deep Learning Approach, in bioRxiv (2018). https://doi.org/10.1101/242818
56. A.G. Taghian, B.L. Smith, J.K. Erban (eds.), *Breast Cancer: a Multidisciplinary Approach to Diagnosis and Management*. (Demos Medical, New York, 2010). ISBN: 978–1–933864–44–0
57. J. Schmidhuber, Deep learning in neural networks: an overview. Neural Netw. **61**, 85–117 (2015)
58. A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, in *Advances in Neural Information Processing Systems*, vol. 25, ed. by F. Pereira, C.J.C. Burges, L. Bottou, K.Q. Weinberger (Curran Associates, Inc., 2012), pp. 1097–1105
59. D. Eigen, J.T. Rolfe, R. Fergus, Y. LeCun, Understanding deep architectures using a recursive convolutional network. CoRR, abs/1312.1847 (2013)
60. A.T. Azar, S.A. El-Said, Probabilistic neural network for breast cancer classification. Neural Comput. Appl. **23**(6), 1737–1751 (2012)
61. M.C. Araújo, R.C.F. Lima, R.M.C.R. de Souza, Interval symbolic feature extraction for thermography breast cancer detection. Expert Syst. Appl. **41**(152014), 6728–6737. https://doi.org/10.1016/j.eswa.2014.04.027
62. T. Kooi, G. Litjens, B. van Ginneken, A. Gubern-Mérida, C.I. Sánchez, R. Mann, A. den Heeten, N. Karssemeijer, Large scale deep learning for computer aided detection of mammographic lesions. Med. Image Anal. **35**, 303–312 (2017)

Chapter 2
Performance Evaluation of Classification Algorithms on Diagnosis of Breast Cancer and Skin Disease

M. Sinan Basarslan and F. Kayaalp

Abstract Health is so important for human beings. Thanks to the technological developments both in medicine and information technologies, the success percentages of both medical diagnosing and medical treatment systems are increasing day by day. Cancer is the most common causes of death in today's world and is generally diagnosed at the last stages. Cancer has many types such as breast cancer, skin cancer, leukemia and etc. Diagnosis of cancer at early stages is very important for the success of medical treatments. The aim of this study was to evaluate the classification performances of some popular algorithms on the way to design an efficient computer aided breast and/or skin cancer diagnosing system to support the doctors and patients. For this purpose, same machine learning and deep learning algorithms were applied on immunotherapy dataset and breast cancer Coimbra dataset from UCI machine learning data repository. Feature selection by information gain and reliefF were applied on datasets before classification in order to increase the efficiency of classification processes. Support Vector Machines (SVM), Random Forest (RF), Recurrent Neural Network (RNN) and Convolutional Neural Network (CNN) algorithms were used for classification experiments. Accuracy values are used for performance metric. According to these results, RNN has shown the best performance among the others with 92% on both datasets. This shows that deep learning algorithms especially RNN has great potential to diagnose the cancer from dataset with high success ratios.

Keywords Deep learning · Breast cancer Coimbra · Immunotherapy · Attribute selection

M. Sinan Basarslan (✉)
Doğuş University, Istanbul, Turkey
e-mail: mbasarslan@dogus.edu.tr

F. Kayaalp
Düzce University, Düzce, Turkey

U. Kose and J. Alzubi (eds.), *Deep Learning for Cancer Diagnosis*,
Studies in Computational Intelligence 908,
https://doi.org/10.1007/978-981-15-6321-8_2

2.1 Introduction

There is nothing more important than the health of people. Especially the diseases affect people's family life, business life and all kinds of social life. Cancer disease is the most common health problem affecting people in recent years.

Cancer disease is the uncontrolled or abnormal growth and reproduction of cells as a result of DNA damage in cells [1]. According to the World Health Organization (WHO) data, the incidence of cancer increased by 4 millions and increased to 18 millions in 2018. It is predicted that there will be 27 millions new cancer cases in 2030, and 40 millions in 2040 [2]. There are many types of cancer such as breast cancer, skin cancer, leukemia and so on. Diagnosis of cancer in the early stages is crucial for the success of medical treatments.

The aim of this study is to evaluate the classification performance of some popular algorithms to design an effective computer-aided breast and/or skin cancer diagnostic system to support doctors and patients. Classification algorithms are applied on breast cancer Coimbra and immunotherapy datasets. Before the classifier model was designed, the performance of the models were increased by applying the attribute selection processes. For the attribute selection process, reliefF and information gain are used. Classification algorithms used in the study were Support Vector Machines (SVM), Random Forest, Recurrent Neural Network (RNN) and Convolutional Neural Network (CNN).

In section two, information about breast and immunotherapy datasets are given. Feature selection, classification algorithms and performance criteria are given in the third section. In section four, results of the experiments are shown and the results are concluded in section five.

The steps followed in the study are shown in Fig. 2.1.

2.2 Datasets

Immunotherapy [3] (Table 2.1) and Breast Cancer Coimbra data sets (Table 2.2) are given information about the attributes [4].

Table 2.1 contains information about the attributes of the data set for the treatment of patients with hand and foot warts in a hospital in Iran. 90 of 180 patients were treated with immunotherapy. It was determined whether they responded positively or not to this treatment [3].

2.3 Method

In this section, information about attribute selection, classification algorithms, and performance evaluation criteria are given.

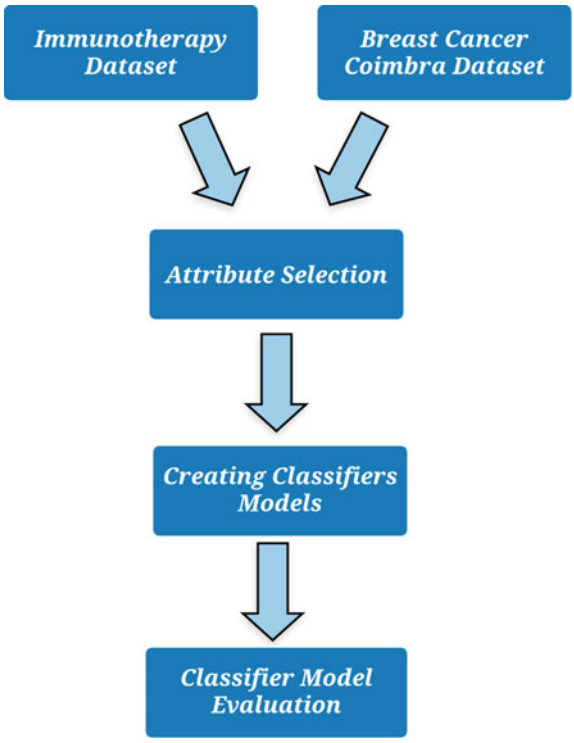


Fig. 2.1 The steps of the study

Table 2.1 Immunotherapy dataset

Attribute number	Attribute values	Data type
1	Sex	Numerical
2	Age (year)	Numerical
3	Time elapsed before treatment (mounth)	Numerical
4	Number of warts	Numerical
5	Type of warts	Numerical
6	Surface area of the warts (mm^2)	Numerical
7	Result of treatment	Categorical

2.3.1 Deep Learning

Deep Neural Networks is the expanded number of layers of the artificial neural network, which works like neurons in the human brain. Deep learning is one of the artificial intelligence techniques often used in advanced applications in today's technologies. In this study, CNN and RNN algorithms were used.

Table 2.2 Breast cancer Coimbra dataset

Attribute number	Attribute values	Data type
1	Age	Numerical
2	BMI	Numerical
3	Glucose	Numerical
4	Insulin	Numerical
5	HOMA	Numerical
6	Leptin	Numerical
7	Adiponectin	Numerical
8	Resistin	Numerical
9	MCP-1	Numerical
10	Class	Categorical

2.3.2 Attribute Selection

The main method used to increase the learning rate of the classification models is the feature selection feature. With the feature selection process, more efficient results can be obtained in terms of the performance of the classification algorithms to be applied on the data set. For this reason, information gain and reliefF attribute selection methods were used in the study.

Information Gain Attribute Selection The information gain (IG) algorithm tends to choose variables with different values, despite the poor performance. To overcome the problems in this case, the Gain Ratio is used. Equations of the calculations are shown in Eqs. (1) and (2) [5].

$$SplitInfo_A(D) = \sum_{j=1}^{v} \frac{D_j}{|D|} * \log_2\left(\frac{|D_j|}{D}\right) \tag{1}$$

$$GainRatio(A) = \frac{Gain(A)}{SplitInfo(A)} \tag{2}$$

ReliefF Attribute Selection ReliefF is an improved version of the Relief algorithm. This algorithm selects a model by taking a sample from the dataset and creates a model based on its proximity to the samples in its class and its distance from different classes as shown in Eq. (3) [6].

Kononenko notes that ReliefF attempts to approximate the following difference of probabilities for the weight of an attribute X [7].

$$\begin{aligned} W_X = {} & P(different\ value\ of\ X | nearest\ instance\ of\ different\ class) \\ & - P(different\ value\ of\ X | nearest\ instance\ of\ different\ class). \end{aligned} \tag{3}$$

Table 2.3 Selected attributes

Dataset name	Selected Attributes with IG	Selected Attributes with ReliefF
Breast cancer Coimbra (10 attributes)	6, 1, 3, 4, 7, 10 total 5 attributes	6, 3, 2, 1, 5, 7, 10 total 6 attributes
Immunotherapy (7 attributes)	3, 5, 7 total 2 attributes	2, 3, 5, 7 total 3 attributes

By removing the context sensitivity provided by the "nearest instance" condition, attributes are treated as independent of one another; Eq. 4 then becomes [7]

$$Relief_X = P(different\ value\ of\ X|different\ class) - P(different\ value\ of\ X|same\ class), \tag{4}$$

which can be reformulated as

$$Relief_X = \frac{Gini'X \sum x \in X P(x)^2}{\left(1 - \sum c \in C P(c)^2\right) \sum c \in C P(c)^2}, \tag{5}$$

where C is the class variable.

Table 2.3 shows the features selected after the two feature selection algorithms applied on the datasets.

Target classes are marked as red in Table 2.3. As seen in Table 2.3, 5 attributes were selected with IG and 6 with ReliefF in Breast Cancer Coimbra dataset. In the immunotherapy dataset, 2 attributes were selected with IG and 3 with ReliefF.

2.3.3 Classification Algorithms

In the study, Support Vector Machines, Random Forest, which is one of the machine learning algorithms, was used on the data sets, while Recurrent Neural Network, and Convolutional Neural, one of the deep learning algorithms, were used.

Support Vector Machines SVM is a vector space-based data mining method. It sets the target boundary between two points furthest from the randomly selected point on the training data [8]. One of the main assumptions of SVM is that all samples in the training set are independent and similarly distributed [9].

Random Forest Developed by Leo Breiman. A single decision was developed to combine the decisions of multivariate decision trees trained with different sets of education instead of producing a single set of training trees. In this algorithm, random property and boot selection are used when creating different training clusters [10].

Table 2.4 Confusion matrix

Actual/predicted as	Positive	Negative
Positive	T_P	F_N
Negative	F_P	T_N

Recurrent Neural Network RNN were first introduced in the 1980s, and these networks have regained popularity thanks to new hardware innovations in the recent period. RNN is a neural network model developed to take the advantages of sequential information and learn the patterns existing in this information [11]. This network architecture is expected to perform the same task for all elements of a series [12].

Convolutional Neural Network CNN consists of different layers such as input, output and hidden layers. Each layer has different functions [9]. Hidden layers are usually related to activation functions, normalization, pooling, and complete connectivity. Each convolutional layer executes a convolution function on its inputs and transfers the result to the next layer. Each neuron is responsible for the available data area [13].

2.3.4 Performance Criteria

Confusion matrix was used to compare the performance of classifier models in the study. As shown in Table 2.4, the confusion matrix describes the number of correct and incorrectly classified samples per class based on a binary classification. In the confusion matrix, tp, fp, fn, and tn represent true positive (tp), false positive (fp), false negative (fn) and true negative (tn) numbers, respectively [14].

Accuracy, sensitivity, precision, f-measure used in the study are given between Eqs. (6) and (9), respectively.

$$Accuracy = \frac{T_P + T_N}{T_P + F_N} \tag{6}$$

$$Sensitivity = \frac{T_P}{T_P + F_N} \tag{7}$$

$$Precision = \frac{T_P}{T_P + F_P} \tag{8}$$

$$F\text{-}measure = \frac{2 * Precision * Sensitivity}{Precision + Sensitivity} \tag{9}$$

2.4 Experimental Results

The performance values of classification processes on breast cancer Coimbra and Immunotherapy datasets obtained by SVM, CNN, RNN and RF classification algorithms are given in Tables 2.5, 2.6, 2.7 and 2.8, respectively. Performance criteria used in the evaluation of these classifier models; accuracy, sensitivity, sensitivity and F-measure. Data sets are divided into training and test data with 5-fold cross-validation.

Table 2.5 shows the performance of SVM classifier on raw data and on datasets obtained by IG and ReliefF attribute selection processes. The performance of SVM on dataset obtained by reliefF has best values in all performance criteria.

Table 2.6 shows the performance of Random Forest classifier on raw data and on datasets obtained by IG and ReliefF attribute selection processes. The performance

Table 2.5 Results with SVM

Dataset	Accuracy	Sensitivity	Precision	F-measure
Raw data	0.71	0.68	0.67	0.68
IG selection	0.76	0.76	0.783	0.72
ReliefF selection	0.79	0.77	0.79	0.78

Table 2.6 Results with RF

Dataset	Accuracy	Sensitivity	Precision	F-measure
Raw data	0.66	0.65	0.670	0.58
IG selection	0.66	0.66	0.683	0.62
ReliefF selection	0.68	0.67	0.69	0.68

Table 2.7 Results with CNN

Dataset	Accuracy	Sensitivity	Precision	F-measure
Raw data	0.724	0.830	0.842	0.834
IG selection	0.830	0.885	0.892	0.826
ReliefF selection	0.840	0.919	0.894	0.902

Table 2.8 Results with RNN

Dataset	Accuracy	Sensitivity	Precision	F-measure
Raw data	0.835	0.840	0.851	0.841
IG selection	0.89	0.886	0.910	0.856
ReliefF selection	0.92	0.948	0.935	0.926

of Random Forest on dataset obtained by reliefF has best values in all performance criteria.

Table 2.7 shows the performance of CNN classifier on raw data and on datasets obtained by IG and ReliefF attribute selection processes. The performance of CNN on dataset obtained by reliefF has best values in all performance criteria.

Table 2.8 shows the performance of RNN classifier on raw data and on datasets obtained by IG and ReliefF attribute selection processes. The performance of RNN on dataset obtained by reliefF has best values in all performance criteria.

When the performance criteria are examined in all the tables given above, it is seen that:

- Classifications with all algorithms on raw data has shown the worst performance,
- Classifications with all algorithms on the dataset obtained by ReliefF attribute selection process has shown the best performance,
- Classifications with Deep Learning algorithms (CNN and RNN) has shown better performance then Machine Learning algorithms (SVM and Random Forest),
- Among Machine Learning algorithms, classifications with SVM has shown better performance then Random Forest,
- Among Deep Learning algorithms, classifications with RNN has shown better performance then CNN.

2.5 Conclusion and Discussion

In this study, deep learning and machine learning algorithms were applied on two human health related datasets. By applying the attribute selection process before classification, the effect of attribute selection processes on the performance of the created classification models was examined. SVM, Random Forest, CNN and RNN are used for classification. Information gain and relief algorithms are used in attribute selection processes.

As a result of the classification models, RNN has given the best performance in all performance evaluation criteria. With this study, it was again observed that selecting attributes positively affects the performance of classifier models. Since the classifier models created from raw data without the feature selection process have been applied, the feature selection process gives more inefficient results than the data set applied, it is seen that the feature selection has a positive effect on the performance criteria. The right decisions to be taken early for treatment are very important for human health. The aim of this study was to help doctors make right decision about human health and it is observed that this can be done with RNN.

In the next studies, it is aimed to work on big data especially about health. Early decisions are very important especially on cancer to achieve patient health. The purpose of these studies is to help doctors make the right decision about human health.

References

1. R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2020. CA Cancer J. Clin. **70**(1), 7–30 (2020)
2. F. Bray, J. Ferlay, I. Soerjomataram et al., Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. **68**(6), 394–424 (2018)
3. F. Khozeimeh, R. Alizadehsani, M. Roshanzamir, A. Khosravi, P. Layegh, S. Nahavandi, An expert system for selecting wart treatment method. Comput. Biol. Med. **81**, 167–175 (2017)
4. Breast Cancer Coimbra Dataset, http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Coimbra
5. L. Venkataramana, S.G. Jacob, S. Shanmuganathan, V.V.P. Dattuluri, Benchmarking gene selection techniques for prediction of distinct carcinoma from gene expression data: a computational study, in *Nature Inspired Computing for Data Science* (Springer, Cham, 2020), pp. 241–277
6. M. Robnik-Šikonja, I. Kononenko, Theoretical and empirical analysis of ReliefF and RReliefF. Mach. Learn. **53**(1–2), 23–69 (2003)
7. I. Kononenko, On biases in estimating multi-valued features, in *Ijcai*, vol. 95 (1995), pp. 1034–1040
8. Ç. Elmas, *Yapay Sinir Ağları (Kuram, Mimari, Eğitim, Uygulama)* (Seçkin Yayıncılık, Ankara, 2003)
9. O. Song, W. Hu, W. Xie, *Robust Support Vector Machine with Bullet Hole Image Classification* (IEEE, 2002)
10. L. Breiman, Some properties of splitting criteria. Mach. Learn. **24**(1), 41–47 (1996)
11. S. Pattanayak, *Pro Deep Learning with TensorFlow* (Apress, New York, 2017), pp. 153–278. ISBN 978-1-4842-3095-4
12. A. Sherstinsky, Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Phys. D Nonlinear Phenom. 132306 (2020)
13. Ş. Kayıkçı, A convolutional neural network model implementation for speech recognition. Düzce Üniversitesi Bilim ve Teknoloji Dergisi **7**(3), 1892–1898 (2019)
14. H. Gunduz, Deep learning-based Parkinson's disease classification using vocal feature sets. IEEE Access **7**, 115540–115551 (2019)

Chapter 3
Deep Learning Algorithms in Medical Image Processing for Cancer Diagnosis: Overview, Challenges and Future

S. N. Kumar, A. Lenin Fred, Parasuraman Padmanabhan, Balazs Gulyas, H. Ajay Kumar, and L. R. Jonisha Miriam

Abstract Health care sector is entirely different from other industrial sector owing to the value of human life and people gives the highest priority. Medical image processing is a research domain where advance computer-aided algorithms are used for disease prognosis and treatment planning. Machine learning comprises of neural networks and fuzzy logic algorithms that have immense applications in the automation of a process. The deep learning algorithm is a machine learning technique that does not relies on feature extraction unlike classical neural network algorithms. Computer-aided automatic processing is in high demand in the medical field due to the improved accuracy and precision. The coupling of machine learning algorithms with high-performance computing gives promising results in medical image analysis like fusion, segmentation, registration and classification. This chapter proposes the applications of deep learning algorithms in cancer diagnosis specifically in the CT/MR brain and abdomen images, mammogram images, histopathological images and also in the detection of diabetic retinopathy. The overview of deep learning algorithms in cancer diagnosis, challenges and future scope is also highlighted in this work. The pros and cons of various types of deep learning neural network architectures are also stated in this work. The intelligent machines in future will be using the deep learning algorithms for the disease diagnosis, treatment planning and surgery.

S. N. Kumar (✉)
Amal Jyothi College of Engineering, Kanjirappally, Kerala, India
e-mail: appu123kumar@gmail.com

A. Lenin Fred · H. Ajay Kumar · L. R. Jonisha Miriam
Mar Ephraem College of Engineering and Technology, Elavuvilai, India

P. Padmanabhan · B. Gulyas
Nanyang Technological University, Singapore, Singapore

U. Kose and J. Alzubi (eds.), *Deep Learning for Cancer Diagnosis*,
Studies in Computational Intelligence 908,
https://doi.org/10.1007/978-981-15-6321-8_3

3.1 Introduction

Machine learning and artificial intelligence has evolved rapidly in recent years and has immense applications in computer vision, computer-aided diagnosis and medical image processing for treatment planning [1]. The classical machine learning algorithms like decision tree, back propagation, SVM, KNN are widely used in many applications [2].

The general architecture of the neural network is depicted in Fig. 3.1. The input layer comprises of multiple neurons that accept features as input data and activation functions are used to sum the data, feed to the hidden layer and finally output layer generates the output. The number of hidden layers used relies on the complexity of the application and deep learning architecture comprises of multiple hidden layers [3]. The general architecture of ANN is depicted in Fig. 3.1.

Least square support vector machine (LSSVM) is a fully automatic approach employed for the breast cancer diagnosis in Wisconsin breast cancer dataset (WBCD). The dataset comprises of 683 samples. Out of these, 444 and 239 belongs to benign and malignant class respectively. In LSSVM, the medical data was examined in a detailed manner with less time complexity. Accuracy, sensitivity, specificity and confusion matrix are used for validation. This method uses 10-fold cross-validation that yields 98.53% classification accuracy [4]. For the breast cancer diagnosis, hybrid machine learning technique based on the fuzzy artificial immune system and the k-NN algorithm was proposed for the analysis of WBCD dataset. The classification accuracy is 99.14% when compared with 8 other methods this proposed system provides better in 10-fold cross-validation [5]. The feature selection is combined with support vector machine in the breast cancer detection for achieving maximum positive prediction and reducing computational complexity. In this, the features are measured with the help of F-score and SVM parameters are optimized by the grid search algorithm. Experiments were conducted on WBCD and the performance was evaluated with sensitivity, specificity, positive predictive value, negative predictive value, confusion

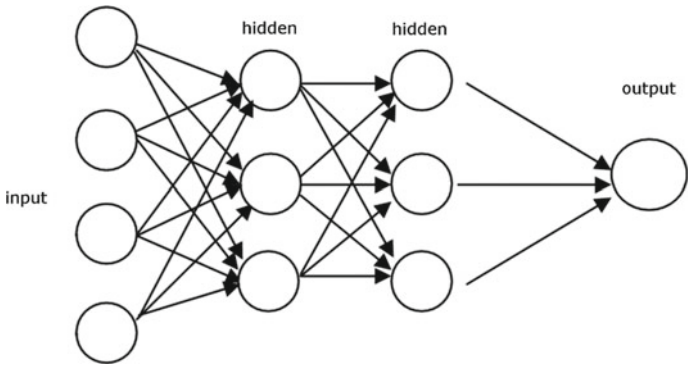


Fig. 3.1 General architecture of neural network

matrix and ROC curves. The 99.51% of classification accuracy was achieved for the SVM model that comprises of five features [6]. Support vector machine-based classifier was compared with Bayesian classifier and Artificial Neural Network for the automatic diagnosis and prognosis of breast cancer. The Wisconsin prognostic breast cancer dataset and WBCD cancer dataset was used, then the achieved accuracy is 96.91% and 97% respectively. The comparative study reveals that SVM is superior in terms of sensitivity, specificity and accuracy [7].

The authors explored the applicability of three well-known classifiers named SVM, kNN and PNN in two publicly available FNAB and gene microarray benchmark dataset for the characterization of benign and malignant tumours of the breast. The 692 specimens of FNAB and 295 microarrays are used for classification 98.8% and accuracy 96.33% was achieved against both benchmark datasets [8]. For the classification of gene data, ensemble methods like bagging and boosting are compared with single supervised learning. These algorithms are applied for 7 datasets and the performance is estimated by specificity, sensitivity, positive predicted value and predictive accuracy. Comparative analysis shows that bagged and boosted decision tree outperforms better than a single decision tree [9]. In order to overcome the drawbacks of local minima, improper learning rate and over fitting, extreme learning machine was focussed. This approach works on GCM, Lung and Lymphoma dataset and the performance of this approach was compared with ANN, SANN, and SVM algorithms. The comparative analysis indicates that the proposed methodology based on ELM accomplish better accuracy with reduced complexity and training time [10]. Single model schemes like Neural fuzzy, K-nearest neighbour, Quadratic classifier and their combined ensemble methods are demonstrated for the classification of breast cancer diagnostic medical dataset. The dataset contains 699 samples. From these 458 samples belongs to benign and 241 samples belong to malignant. The accuracies of Neural fuzzy, K-NN, Quadratic classifier and their combined ensemble methods are 94.28%, 96.42%, 94.50% and 97.14% respectively. It shows that the ensemble method provides superior accuracy when compared to single models [11]. Classification and Regression Tree (CART) is a decision tree which was implemented to enhance the accuracy and to remove unreliable branches from the decision tree. In this, CART is compared with ID3 and C4.5 algorithms. Accuracy, computation time and size of the tree are used to analyze the performance of breast cancer. If CART without feature selection is used then the achieved accuracy of Breast Cancer are as follows; for Breast Cancer Wisconsin (Original) and Breast Cancer Wisconsin (Diagnostic) dataset, accuracy is 69.23%, 94.84% and 92.97% respectively. In CART with feature selection, the SVM Attribute Eval method with 73.03% accuracy is best for Breast Cancer dataset, Principal Components Attribute Eval method with 96.99% accuracy is best for Breast Cancer Wisconsin (Original) dataset and Symmetric Uncert Attributeset Eval method with 94.72% accuracy is best for Breast Cancer Wisconsin (Diagnostic) dataset. As a result, CART has better accuracy than ID3 and C4.5 algorithms [12]. Machine learning classifiers (MLC) was applied for the breast cancer diagnosis on craniocaudal (CC) and/or mediolateral oblique (MLO) mammography images. The analysis is taken out in 286 cases. When merging feature vectors are extracted from CC and MLO views, classifiers

achieve an area under the ROC curve of 0.996 [13]. C4.5 approach was adapted for the classification of patients into either carcinoma in situ or malignant potential. For training, nearly random samples of 500 records were taken from SEER dataset. It yields ~94% and ~93% accuracy during training and testing phase respectively [14]. Multiclassifiers like Decision tree (J48), MLP, Naive Bayes (NB), SMO, and Instance-based for K-NN (IBK) are used for the diagnosis and prognosis of breast cancer on three databases namely WBC, WDBC and WPBC. Experimental analysis is carried out on a WEKA data mining tool and it uses 10-fold validation. The performance is assessed with classification accuracy and confusion matrix. The fusion of SMO, J48, NB and IBK provides 97.2818%, 97.7153% and 77.3196% accuracies for WBC, WDBC and WPBC datasets respectively [15]. The authors have performed several data mining techniques and soft computing approaches for the diagnosis and prognosis of breast cancer on UCI machine learning and SEER dataset. From the experimental review, it is clear that decision tree obtains 93.62% accuracy in both of the datasets [16]. In order to predict the recurrence of breast cancer, Decision Tree (C4.5), SVM, and ANN is developed on Iranian Center for Breast Cancer dataset which comprises of 1189 records, 22 predictor variables and one outcome variable. Performance of these three methods is compared with the help of sensitivity, specificity and accuracy using 10-fold cross-validation. The accuracies of C4.5, SVM and ANN are 93.6%, 95.7% and 94.7% respectively. Results suggest that SVM classifier is the best predictor classifier among these methods [17].

The Bayesian deep learning approach combined with active learning approach is used for the skin cancer diagnosis form the lesion images. This approach which performs better BALD and uniform acquisition function on ISIC 2016 melanoma dataset with an accuracy rate of 75% [18]. For the diagnosis of the prostate cancer diagnosis, CNN based deep learning XmasNet is used for the classification of the lesion in the prostate. The network architecture was trained using 3D multiparametric data. The XmasNet performance was evaluated with PROSTATEx and outperformed over 69 methods with an AUC of 0.84 [19]. The computer-aided diagnosis with deep learning framework was suggested for the detection of cancer on the lung nodules for the Kaggle Data Science Bowl 2017 challenge. The training of the framework for the detection of the lung nodule was done with LUNA16 and cancer classification with KDSB17 datasets. In this Kaggle Data Science Bowl, 2017 challenge deep learning with ResNet18 architecture placed top 3% of 1972 teams with high accuracy [20]. For the recognition of Adenocarcinoma and squamous cell carcinoma in the histopathological slides of the lung tissues is difficult for an expert by visual analysis hence CNN model with inception v3 was used. Using the Cancer Genome Atlas (TCGA) histopathological datasets this model achieves an AUC of 0.97%. The network is also trained for predicting mutated genes in the lungs and it predicts six types of genes in the pathology genes with an accuracy of 0.733–0.856 [21]. The diagnosis of lung cancer is challenging to improve the classification rate of cancer. The deep CNN (DCNN) was used for the classification of different types of cancer adenocarcinoma, squamous cell carcinoma, and small cell carcinoma in the lungs. The threefold cross-validation was used for the evaluation of the classification and estimates 71% accuracy in the lung microscopic cytology and pathology

images [22]. Using the deep neural network, the benign and malignant lung nodule detection and classification in the LIDC-IDRI database was performed. The CNN, DNN, and SAE deep neural network architecture are suggested and compared, where CNN deep neural network outperforms in the accuracy, sensitivity and specificity with values of 84.15%, 83.96% and 84.32% respectively with other neural network architectures in the prediction of lung cancers [23].

A computerized lung cancer detection algorithm was proposed with automatic feature learning from multichannel ROI and the deep neural network (CNN, DBN and SDAE). The proposed algorithm extracts the sub patches form the multichannel ROI for learning from the morphological and texture features. The comparison with the multichannel and single channel with different neural networks states that CNN performs better in accuracy and AUC in ROI and nodule detection [24]. Breast cancer detection was done in the Image Retrieval in Medical Applications (IRMA) mammogram images using the deep learning convolutional neural network. A BC-DROID framework was recommended for the automatic detection of ROI and CNN detects and classifies the ROI with benign and malignant in the breast mammogram images and it achieves a classification and detection accuracy of 93.5% and 90% respectively [25]. Improving the diagnostic accuracy of Multipurpose Image Analysis Software, the deep learning Artificial Neural Network was involved and trained using Breast Cancer Digital Repository dataset. The deep neural network improves the performance of the software with the ROC curve of training and testing are 0.81 and 0.79 respectively [26]. The deep learning algorithm was used for the recognition of Lymph node metastases in Breast cancer of women in the CAMELYON16 challenge competition. Initially, the network was trained using with and without metastases by immune histochemical staining of total 270 datasets. The receiver operating characteristic curve (AUC) was best for deep learning with GoogleNet architecture is 0.994 and compared with Pathologist diagnosis [27]. The breast cancer diagnosis using CADx with pre-trained CNN using fused radiomic features of MRI, Mammogram and Ultrasound were proposed. The fusion features improve the AUC curve with the traditional CADx in the recognition of breast lesion cancer. For the classification, the SVM classifier is used and was evaluated with fivefold cross-validation [28].

Deep learning allows computers to generate complex concepts out of simpler concepts and it comprises of multiple processing layers to learn the data with multiple levels of abstraction. A wide number of supervised and unsupervised learning neural networks are there like back propagation, radial basis function, group method data handling, self-organizing map etc. [29]. The deep learning is a supervised neural network comprises of multiple layers and now extends to 1000 for handling the big data. The deep learning neural network algorithms are further classified into various types such as networks CNN, deep neural network (DNN), deep belief network (DBN), deep autoencoder, deep Boltzmann machine (DBM), deep conventional extreme machine learning (DC-ELM) recurrent neural network (RNN), Recursive Neural Networks and it's variant like BLSTM and MDLATM etc. [30].

Table 3.1 Deep learning architectures for breast cancer diagnosis

Reference number	Application area	Medical imaging modality	Deep learning architecture	Validation metrics
[58]	Breast	Mammogram	Unsupervised deep learning	AUC—0.59 (0.56–0.62)
[59]	Breast	Pathology	CNN	F measure—0.7345
[60]	Breast	Mammogram	CNN	Accuracy—85% Sensitivity—85%
[61]	Breast	Mammogram	Cascade of deep learning and random forest classifiers	True positive Rate—0.96
[62]	Breast	Mammogram	CNN	AUC—0.826
[63]	Breast	Histology	CNN (Aggnet)	AUC—0.8695
[64]	Breast	Mammogram	Ensemble classifier	AUC—0.86
[65]	Breast	Pathology	CNN	F score—0.94 (0.92–0.94)

3.2 Stages in Cancer Diagnosis Using Medical Imaging

The medical images such as CT/MRI and Ultrasound are taken for disease analysis as well as to study the anatomical organs healthy function. The medical imaging modalities role is pivotal in cancer diagnosis and treatment planning. The first step in the processing of medical images is preprocessing or filtering. The objective is to filter is to remove the noise in the image incurred during the acquisition phase or to improve the quality of the image for better analysis [31]. The segmentation refers to the extraction of the region of interest and in the case of medical images, ROI represents anatomical organ or its anomalies like a tumor, cyst etc. The classification stage comprises usually a machine-learning algorithm to categorize the severity of cancer. The compression refers to the usage of computer-aided algorithms for reducing the file size for data storage and transfer. The machine learning algorithms can be employed in any one of the phases of cancer diagnosis and is depicted in Table 3.1 [32].

3.3 Types of Deep Learning Neural Network Architectures

The deep learning neural network architectures are of different types based on the learning strategy and are categorized as follows; supervised, semi-supervised, unsupervised and reinforcement learning [33]. The classification of deep learning neural networks is depicted in Fig. 3.2.

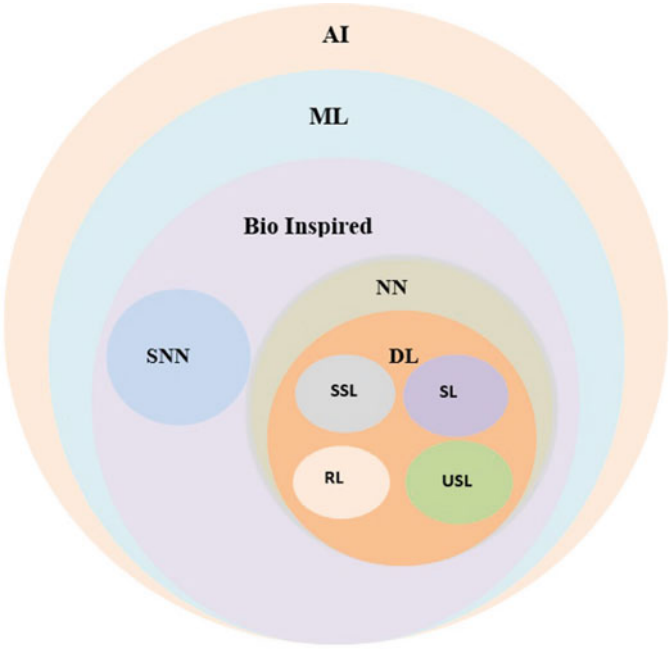


Fig. 3.2 Classification of deep learning architectures based on learning strategy

3.3.1 Deep Supervised Learning Architectures

The supervised learning architectures use predefined data for training. The network is taught with the possible input combinations and the target outputs. The testing phase utilizes the trained phase information for validation. Some of the typical approaches in supervised learning are convolution neural network, recurrent neural network, long short term memory, deep neural networks and gated recurrent units [34, 35].

3.3.2 Deep Semi-supervised Learning Architectures

The deep supervised learning architectures use partly labelled data for the training phase. In some scenarios, deep reinforcement learning, RNN, LSTM, GRU and Generative Adversarial Networks (GAN) are used as semi-supervised learning architectures [36].

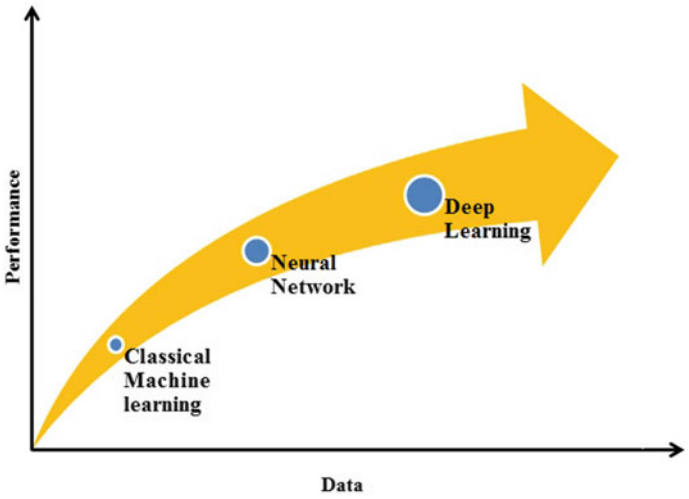


Fig. 3.3 Performance of traditional machine and deep learning architecture

3.3.3 Deep Unsupervised Learning Architectures

The deep unsupervised learning architectures do not use any labelled data and analyze the internal representation of data using some features. The unsupervised approaches meant for clustering and dimensionality reduction. The deep unsupervised learning architectures are Auto-Encoders (AE), Restricted Boltzmann Machines (RBM) [37]. The performance of traditional machine learning and deep learning architecture is depicted in Fig. 3.3.

3.4 Typical Deep Learning Architectures

3.4.1 Convolution Neural Network

The CNN was efficient for the processing of 2D and 3D images. Most of the CNN algorithms are trained with a gradient-based algorithm. The number of parameters to be tuned is less, when compared with the other neural network architectures. The CNN architecture comprises of; feature extractors and classification [38]. The feature extraction layer accepts the input from the previous layer and feeds to the next layer. The CNN architecture comprises of three types of layers; convolution, max-pooling and classification. The convolution layers are represented by even numbers and max-pooling layers are represented by odd numbers. The classification layer is a fully connected layer and it forms the last stage of architecture. Back propagation architecture is employed in the classification stage for better accuracy. Different types

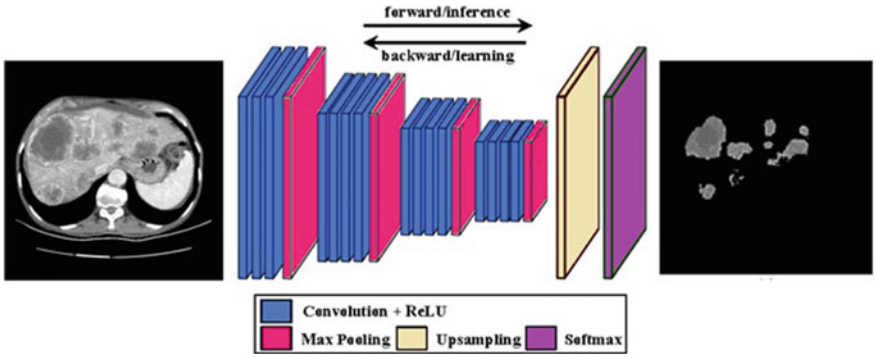


Fig. 3.4 Convolution neural network architecture

of pooling operations are there; max pooling, min pooling, average pooling and global average pooling. In the convolution layer, feature maps are generated by convolving the inputs with kernel comprising of the linear or nonlinear activation function. The examples of activation functions are sigmoid, hyperbolic tangent, Softmax, rectified linear, and identity functions. The pooling layer is also termed as subsampling layer and here downsampling operation takes place. The number of classification layers varies in accordance with the application and in some cases two to four layers are observed in LeNet [39], AlexNet [40], and VGG Net [41]. The convolution neural network architecture is depicted in Fig. 3.4.

3.4.2 Multi-scale Convolution Neural Network

The classical convolution neural network was modified to formulate multi-scale convolution neural network [42], it comprises of 3 convolution layer, rectified linear unit and the max-pooling layer and two fully connected layers. The downsampling is performed on the input image and feature extraction is done and delivered to the multi-scale convolution neural network.

3.4.2.1 LeNet-5

It is a seven-level convolution layer neural network and was used for handwritten digits digit classification. Input images of size 32×32 were used and for a complex application, the number of convolution layers is used. The LeNet architecture is depicted in Fig. 3.5, it comprises of 2 convolution layers, subsampling layers and fully connected layers. A single output layer with Gaussian connectivity was employed [39].

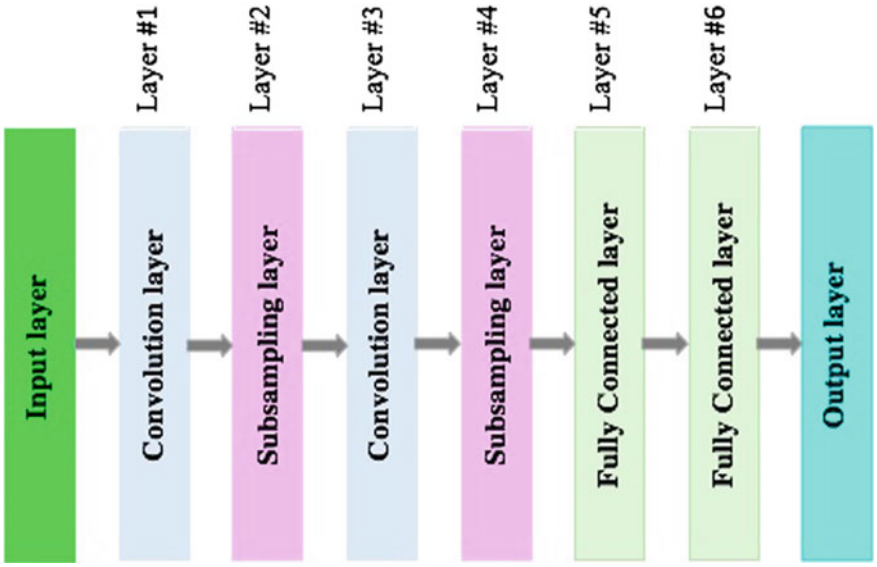


Fig. 3.5 LeNet architecture

3.4.2.2 AlexNet

The architecture of Alexnet is similar to LeNet, but it has deeper layers with more number of filters per layer and coupled convolution layers. The ReLU activation function was put after each convolution and fully connected layer. It was a challenge winning architecture in 2012 with a minimized error of 15.3% from 26%. It comprises of 11 × 11, 5 × 5, and 3 × 3 convolutional kernels, max pooling, dropout, data augmentation, and ReLU activations [40]. The AlexNet architecture is depicted in Fig. 3.6.

3.4.2.3 ZFNet

The architecture of ZFNet was similar to AlexNet, however, the parameters are tuned and it is a challenge winning architecture in 2013%. The error was minimized to 14.8%. It uses 7 × 7 kernels instead of 11 × 11 kernels thereby minimizes the number of weights. The number of parameters to be tuned is also minimized and hence, the accuracy is improved [41].

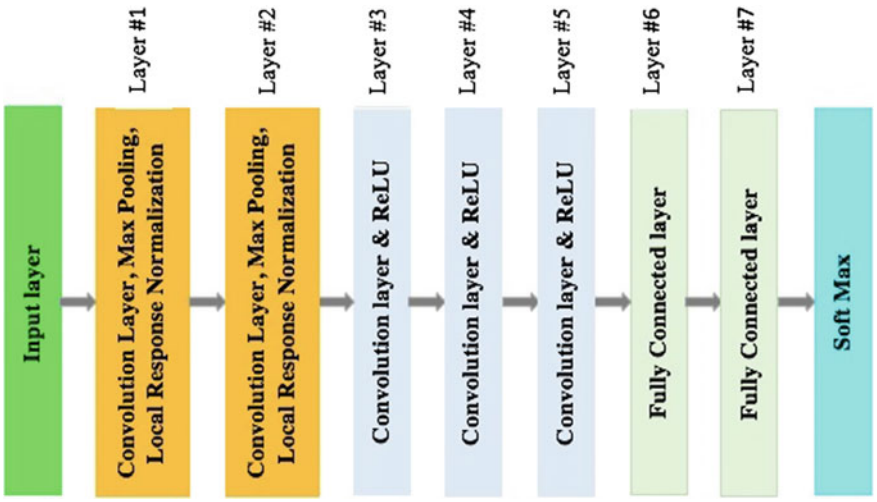


Fig. 3.6 AlexNet architecture

3.4.2.4 GoogleNet

The architecture of GoogleNet comprises of LeNet with an inception structure. It comprises of 22 layers and the error rate was minimized progressively from 6.67% to 3.6% during testing. This architecture was a winner of ILSVRC 2014 [42]. It has low computation complexity when compared with the classical CNN architecture. It was less used when compared with other architectures; AlexNet, VGG [43]. The GoogleNet architecture is depicted in Fig. 3.7.

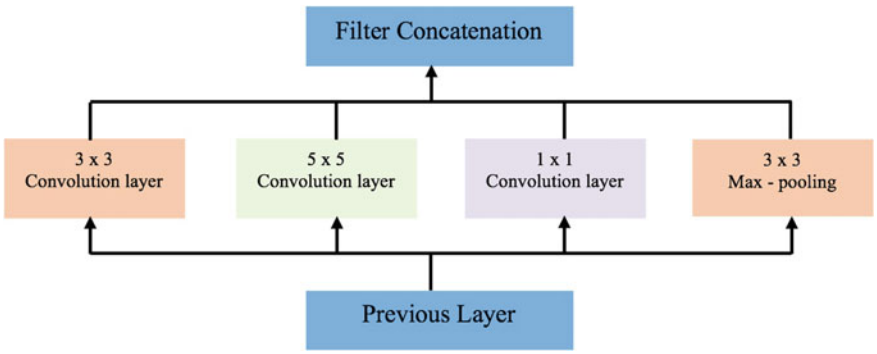


Fig. 3.7 GoogLeNet

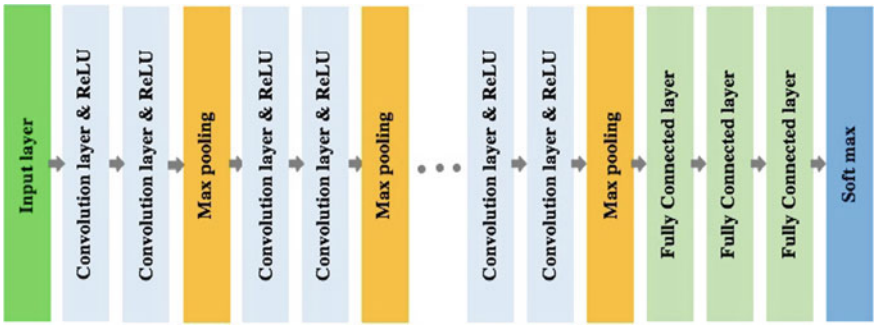


Fig. 3.8 VGG net architecture

3.4.2.5 VGGNet

The VGGNet was the winner of ILSVRC 2014 and comprises of 16 convolution layers with multiple filters [44]. The feature extraction was found to be efficient with this architecture; however, the parameter tuning is a crucial one. Three VGG models were proposed; VGG-11, VGG-16 and VGG-19 and have 11, 16 and 19 layers. All the VGG models end with 3 fully connected layers. The VGGNet architecture is depicted in Fig. 3.8.

3.4.2.6 ResNet

The ResNet was the winner of ILSVRC 2015 and it uses skip connections and batch normalization [45]. The computation complexity was less when compared with the VGGNet and gated recurrent units are used as skip connections. It comprises of 152 layers and the error was minimized to 3.57%. It resolves the problem of vanishing gradient problem. It is a classical feed-forward neural network with a residual connection [46]. It comprises of several residual blocks and based on the architecture, its operation varies. The residual network is depicted in Fig. 3.9.

3.4.2.7 Fully Convolutional Networks (FCNs)

The architecture of fully convolution network differs from the classical CNN such that fully convolution layer was replaced by the upsampling layer and a deconvolution layer, fully connected layer is depicted in Fig. 3.10.

The architecture is proposed in such a way that fully convolution layer and deconvolution layer forms the backward versions of pooling and convolution layers. The accuracy of the architecture was improved by the incorporation of upsampling layer and deconvolution layer [33, 47].

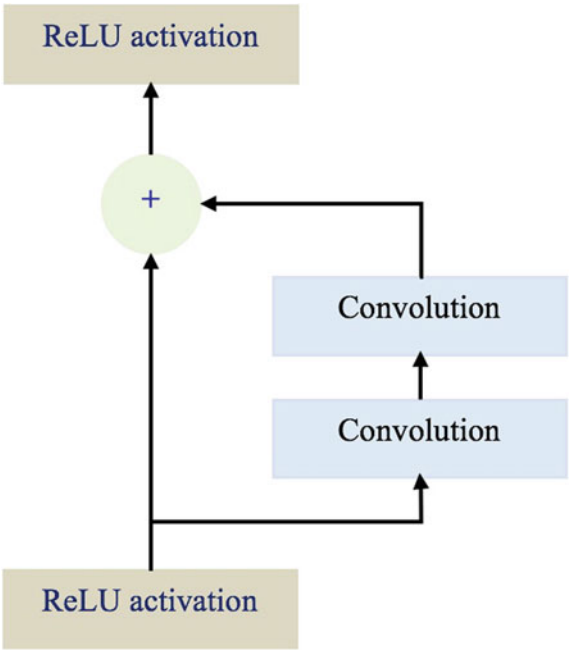


Fig. 3.9 Residual network

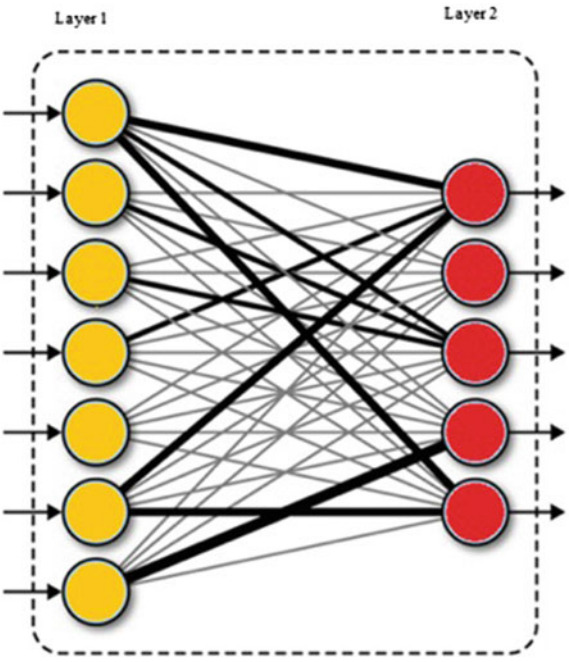


Fig. 3.10 Fully convolutional network architecture

3.4.2.8 U-Net Fully Convolutional Neural Network

The U-Net was developed for the medical image segmentation and it comprises of two paths. The first path comprises encoder that captures the context in an image. The second path comprises a decoder and it comprises of transposed convolutions [48, 49]. The U-Net Fully Convolutional Neural Network is depicted in Fig. 3.11.

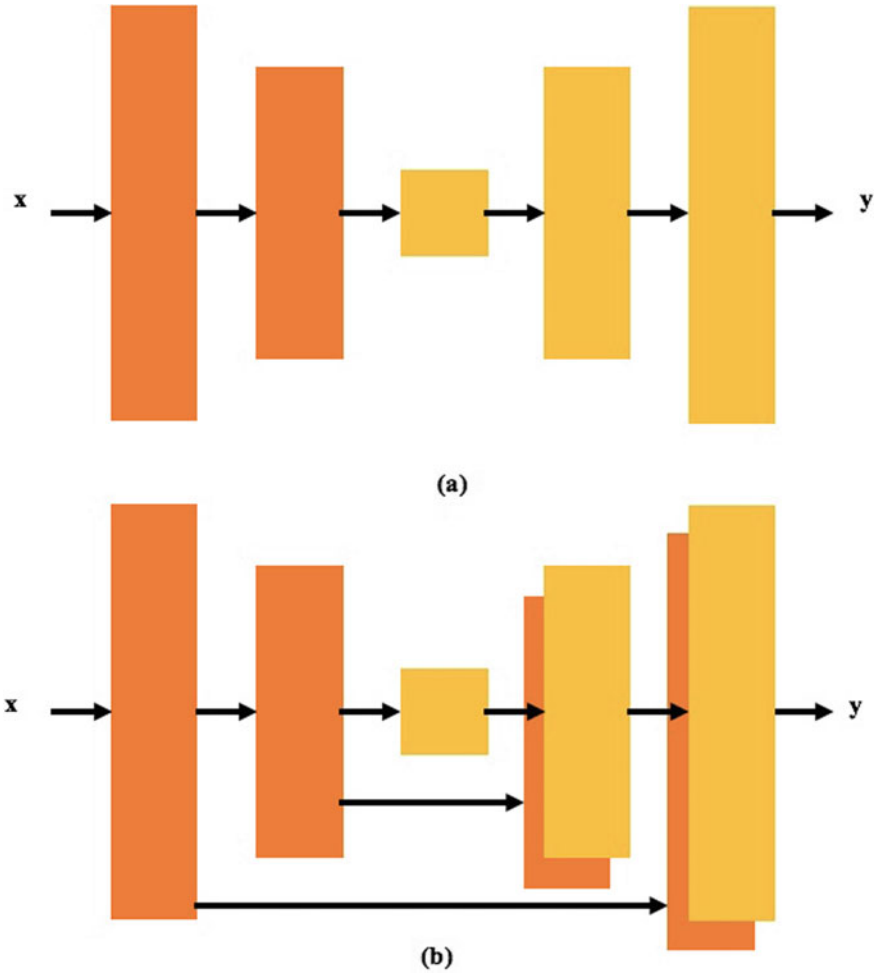


Fig. 3.11 **a** Encoder-Decoder, **b** U Net Fully Convolutional Network architecture

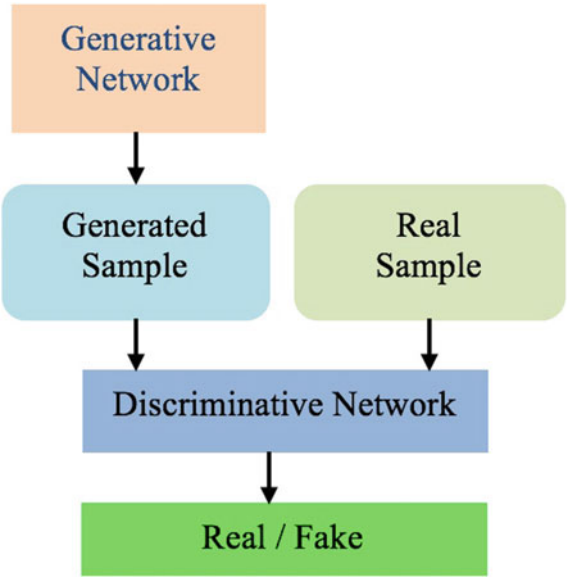


Fig. 3.12 Generative adversarial network architecture

3.4.2.9 Generative Adversarial Networks (GANs)

Generative adversarial network (GAN) is a generative model of unsupervised learning which consists of two neural networks namely discriminative network (Discriminator) and generative network (Generator). The aim of the discriminator is to distinguish real and fake samples and the objective of the generator is to generate fake samples of image data and it fools the discriminator. This architecture is trained in two ways. The discriminator is trained while the generator is idle. On the other hand, the generator is trained while the discriminator is idle. There are five different types of GAN viz Vanilla, Conditional, deep convolutional, Laplacian and super-resolution. Face ageing, Photo blending, Photo inpainting, 3D object generation, Video generation are the applications of GAN [50, 51]. The Generative Adversarial Network Architecture is depicted in Fig. 3.12.

3.4.2.10 Recurrent Neural Networks

The basic structure of the RNN is depicted in Fig. 3.13. Different versions of RNN architecture are described in [52]. The recurrent neural network comprises multiple functional blocks depicted in Fig. 3.14. The vanishing gradient problem exists in recurrent neural networks. In the recurrent neural network, previous states are fed as input to the current state of the network which indicates that it requires memory. It uses sequential information and the connection between the nodes form a directed graph. The purpose of RNN is to map the input sequence to a fixed-sized vector. Extension of effective pixel neighbourhood is achieved by the combination of RNN

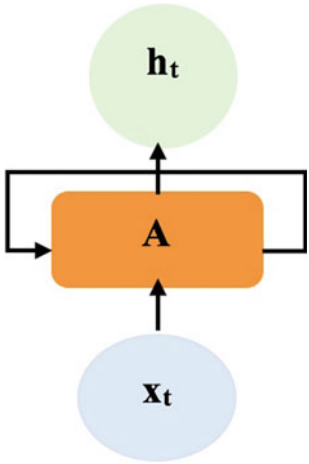


Fig. 3.13 Basic functional block

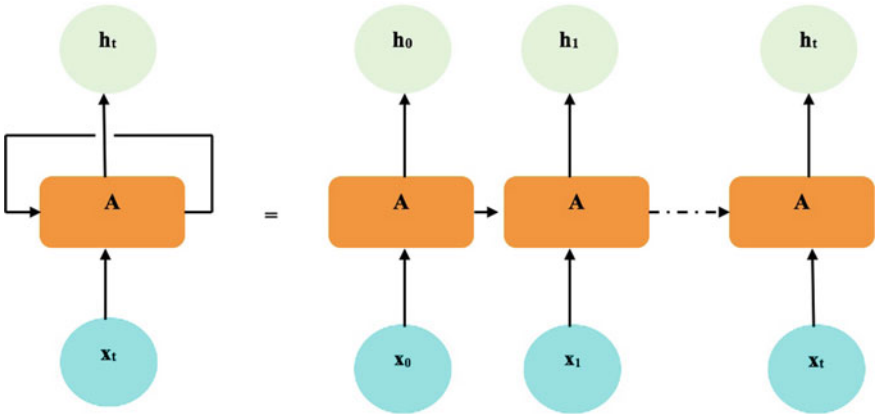


Fig. 3.14 Unrolled RNN

with the convolutional layer. It finds its application in NLP, Time series prediction and machine translation. Long short-term memory network (LSTM) is a type of RNN [53]. The basic functional block of R and unrolled RNN are depicted in Figs. 3.13 and 3.14.

3.4.2.11 Autoencoders

Autoencoder is an efficient unsupervised learning architecture comprises of three layers viz encoder, code and decoder. The purpose of encoder is the mapping of input into smaller representation. Thus the compressed image is the distorted part of the input. Code represents the compressed input. Also, the layer between encoder

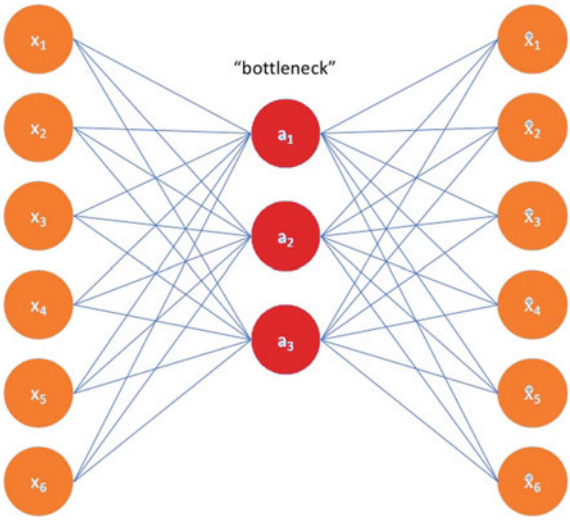


Fig. 3.15 Autoencoder architecture

and decoder is termed as a bottleneck. The autoencoder architecture is depicted in Fig. 3.15.

The decoder maps the code to a reconstruction of the original input. Data specific and lossy are the main properties. Before training the architecture four hyperparameters like code size, number of layers, number of nodes per layer and loss function should be tuned. Dimensionality reduction, image compression, image denoising, feature extraction are the application areas of autoencoder [54, 55].

3.4.2.12 Deep Belief Networks

It comprises of restricted Boltzmann machine (RBM) for the pretrained phase and a feed-forward network for the fine-tuning phase. The RBM extracts the features from the input data vectors and feeds it to the feed-forward network. The back propagation architecture with lower learning rate was employed in deep belief network. In addition, it is having multiple hidden layers. The layer-by-layer learning strategies are main advantage of the deep belief network and its capability to learn from higher-level features are from previous layers [56, 57]. The deep belief network is depicted in Fig. 3.16. The Y represents the input layer and H1, H2 and H3 represents the hidden layers.

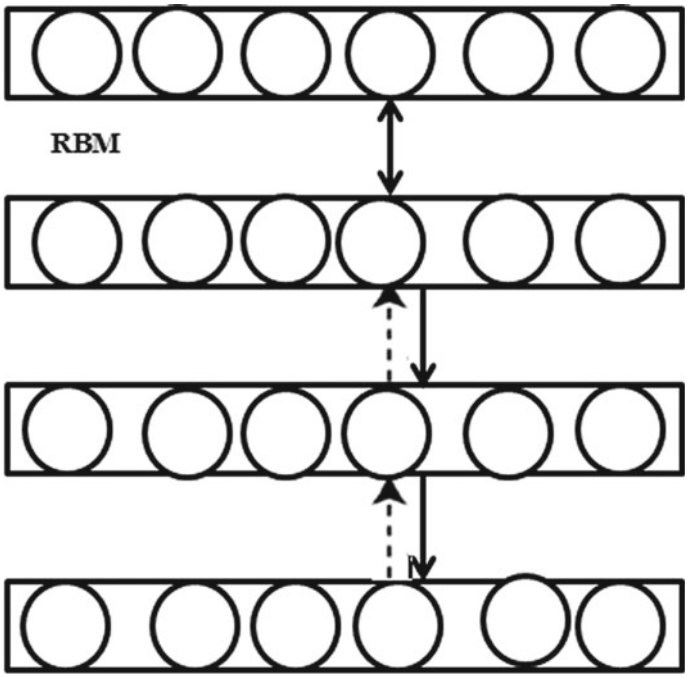


Fig. 3.16 Deep belief networks

3.5 Deep Learning Architectures for Cancer Diagnosis

The deep learning architectures are having immense applications in disease diagnosis and some of the works related to cancer diagnosis of various anatomical organs are listed below. In each table, the application area indicates the anatomical organ. For the validation of deep learning architectures, different types of metrics can be used and the metrics used in each work is also highlighted. The various Deep Learning architectures for cancer diagnosis are depicted in Tables 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9 and 3.10.

The sparse encoder with a convolution architecture was proposed in [58] termed as convolution sparse encoder, applicable for 3D data sets. In [59], lesion segmentation was done by coupling global fixed thresholding and local dynamic thresholding, detection of tumour stage was performed by CNN with handcrafted features. In [60], three CNN architectures were evaluated; AlexeNet, VGG Net and GoogLeNet. The GoogLeNet was found to be efficient with an accuracy of 85% when compared with the AlexeNet and VGG Net having an accuracy of 84% and 82%. The hybrid machine learning classifiers are employed in [61] that comprises of four modules. The first stage comprises of multiscale deep belief network with Gaussian mixture model, the second stage comprises of cascade connection of deep CNN with SVM classifier

Table 3.2 Deep learning architectures for breast cancer diagnosis

Reference number	Application area	Medical imaging modality	Deep learning architecture	Validation metrics
[66]	Breast	Mammogram	CNN (Alexnet and VGG) and SVM	Accuracy—99.38%
[67]	Breast	Histology	Deep cascaded networks	F score—0.788
[68]	Breast	Mammogram	CNN	AUC—79.9–86%
[69]	Breast	Mammogram	Faster R-CNN	AUC—0.72
[70]	Breast	Mammogram	Deep CNN	Sensitivity—91%
[71]	Breast	Mammogram	CNN (ImageNet)	Accuracy—96.7%
[72]	Breast	Pathology	CNN	AUC—above 97%, Sensitivity—73.2%
[73]	Breast	Pathology	CNN (Convnet)	Dice Coefficient —75.86% Positive predictive Value—71.62% Negative predictive Value—96.77%
[74]	Breast	Ultrasound	CNN (Alexnet)	TPF—0.98 Fps/image—0.16 F measure—0.91

Table 3.3 Deep learning architectures for liver cancer diagnosis

Reference number	Application area	Medical imaging modality	Deep learning architecture	Validation metrics
[75]	Liver	CT/3D	DNN	Accuracy—99.38%, Jaccard index—98.18%
[76]	Liver	CT	BPNN	Accuracy—73.23%
[77]	Liver	CT	CNN	Dice—80.06% Precision—82.67% Recall—84.34%

and the third stage comprises of cascade connection of two random forest classifiers. The proposed CNN3 architecture [62] generates superior results in terms of AUC (0.826) when compared with other techniques DeCAF (0.836), CNN2 (0.821), HCfeats (0.799), HOG (0.796), HGD (0.793). The CNN architecture (AggNet) generates efficient results when compared with the AM-MV, AM-GLAD and AM-GT models for the detection of breast cancer in histology images [63].

The tumour classification by SVM with the classical method of feature extraction, CNN features as input to SVM and ensemble classifier, a combination of above two

Table 3.4 Deep learning architectures for lung cancer diagnosis

Reference number	Application area	Medical imaging modality	Deep learning architecture	Validation metrics
[78]	Lung	CT	Autoencoder with deep features	Accuracy—75.01%
[79]	Lung	CT	Deep Belief Networks	Accuracy—81.19%
[80]	Lung	CT/3D	DNN	Accuracy—82.1% Sensitivity—78.2% Specificity—86.13%
[81]	Lung	CT/3D	DNN	Sensitivity—78.9%
[82]	Lung	CT/3D	DCNN	True Positive—93%
[20]	Lung	CAT Scans/3D	Modified Resnet	Sensitivity—0.538 F1—0.33
[21]	Lung	Histopathology image	DCNN	AUC—0.97
[23]	Lung	CT/3D	CNN	Accuracy—84.15%, sensitivity—83.96%, specificity 84.32%
[83]	Lung	CT	Unet + Resnet	Accuracy—84%
[84]	Lungs	PET/CT images/3D	CNN	Sensitivity—>90%

Table 3.5 Deep learning architectures for chest cancer diagnosis

Reference number	Application area	Medical imaging modality	Deep learning architecture	Validation metrics
[85]	Chest	X-ray images	CNN (Imagenet)	AUC—0.87
[86]	Chest	X-ray images	CNN (Imagenet)	AUC—0.93 (Right Pleural Effusion) AUC—0.89 (Enlarged heart) AUC—0.79 (classification between healthy and abnormal chest x-ray)

was tested on mammogram images. The ensemble classifier yields efficient results when compared with the other two techniques [64]. The pretrained CNN with SVM was found to be effective for the classification of the breast tumour in mammogram images when compared with the generic classifier based on texture feature extraction [65].

The pretrained CNN models AlexNet and VGG-F with SVM were employed for the breast tumour classification. The pretrained CNN VGG-F, when coupled

Table 3.6 Deep learning architectures for skin cancer diagnosis

Reference number	Application area	Medical imaging modality	Deep learning architecture	Validation metrics
[87]	Skin (basal cell carcinoma)	Histopathology image/3D	Convolutional auto-encoder DL	F-measure—89.4% Accuracy—91.4%
[88]	Skin	Dermoscopy images	Deep belief net	Accuracy—89% Sensitivity—90% Specificity—88.3%
[89]	Skin	Standard camera image	DCNN	Accuracy—98.55 Sensitivity—95%
[90]	Skin	Dermoscopy images	CNN (ReLU)	Accuracy—86.67%

Table 3.7 Deep learning architectures for histopathology image analysis

Reference number	Application area	Medical imaging modality	Deep learning architecture	Validation metrics
[91]	Colon	Histopathology image	Shallow neural network	Accuracy—84%
[92]	Colon	Histopathology image	SC-CNN +NEP	AUC—0.917 F1—0.784
[93]	Colon	TS-QPI images/3D	Deep Neural Network	AUC—95%
[94]	Colon	Capsule Endoscopy (CE)	REFDL	Sensitivity—94.2% Accuracy—96.5%

Table 3.8 Deep learning architectures for bladder cancer diagnosis

Reference number	Application area	Medical imaging modality	Deep learning architecture	Validation metrics
[95]	Bladder	CT urography/3D	Deep learning CNN	Jaccard Index—76.2
[96]	Bladder	CT	Deep learning CNN	AUC—0.73
[97]	Bladder	Pathology	CNN (cystonet)	Sensitivity—90.9% Specificity—98.6%

with SVM, generates superior results [66]. The CNN architecture proposed in [67] generates robust results when compared with the models in the 2014 ICPR MITOS-ATYPIA challenge. The CNN architecture proposed in [68] outperforms the DeCAF CNN2 HOG HGD HCfeats models. The fast R-CNN comprises of a region proposal network (RPN), a deep fully convolutional network that performs better than the classical CNN architecture [69]. The DCNN based CAD system was found to be proficient when compared with the feature-based CAD system [70]. The convolution neural network with a decision mechanism outperforms the feature-based

Table 3.9 Deep learning architectures for cervical and prostate cancer diagnosis

Reference number	Application area	Medical imaging modality	Deep learning Architecture	Validation metrics
[98]	Cervical	Cervigram	CNN (Alexnet)	Sensitivity—87.83% Specificity—90%
[99]	Prostate	Multiparametric MRI/3D	CNN (Xmasnet)	AUC—0.84
[100]	Prostate	Multiparametric MRI	DCNN	AUC—0.897
[101]	Prostate	Photoacoustic (PAI)	DNN (ReLU)	Accuracy—95.04

Table 3.10 Deep learning architectures for brain cancer diagnosis

Reference number	Application area	Medical imaging modality	Deep learning architecture	Validation metrics
[102]	Brain	MRI	Input Cascade CNN	Dice—0.84 Specificity—0.88 Sensitivity—0.84
[103]	Brain	MRI	Multi-layer stacked denoising Autoencoder network	Average Accuracy—98.04
[104]	Brain	MRI	U-Net	HGG—0.88 LGG—0.84
[105]	Brain	MRI	conditional Generative Adversarial Network (cGAN)	Dice—0.68 Sensitivity—0.99 Specificity—0.98
[106]	Brain	MRI	DNN	Classification rate—96.9% Recall 0.97 Precision—0.97 F measure—0.97 AUC—0.984
[107]	Brain	MRI	Fully convolutional neural networks (FCNNs)	Dice—0.86
[108]	Brain	MRI	Multi-view deep learning framework (MvNet) (Mb-FCRN)	Dice—0.88 Accuracy—0.55
[109]	Brain	MRI	Deep Wavelet Autoencoder (DWA)-DNN	Overall accuracy—96% Average accuracy—93%

classifiers [71]. The CNN architecture proposed in [72] detects small tumour in gigapixel pathology slides. The CNN architecture proposed in [73] detects invasive tumour in whole slide images and minimizes the human effort with reduced time complexity. The different deep learning architectures Patch-based LeNet, U-Net, Transfer Learning FCN-AlexNet are tested on the breast ultrasound lesion images; AlexNet and Patch-based LeNet generates efficient results [74].

In [75], prior to tumour classification by DNN, ROI extraction was done by watershed and Gaussian mixture model (GMM) algorithm. The fully convolution neural network architecture U net was proposed for the segmentation of liver tumour, post-processing was performed by 3D connected component labelling to improve the segmentation result [76]. The deep CNN was found to be superior when compared with the other classifiers like AdaBoost, Random Forests (RF), and support vector machine (SVM) [77].

The Autoencoder was used for the classification of lung nodules and was able to detect minute lesions not observable by human experts [78]. Three deep learning architectures Convolutional Neural Network (CNN), Deep Belief Networks (DBNs), Stacked Denoising Autoencoder (SDAE) were used for cancerous nodule detection; DBN has improved accuracy of 0.8119 [79]. The DNN was found to be efficient for the segmentation of cancer nodules in 3D CT images; applicable for small lung nodule segmentation. The efficiency of DNN increases with an increase in training data [80]. The CNN proposed in [81] comprises of two stages; the first phase extracts volumetric features and the second phase performs classification. The deep learning architecture was coupled with SVM classifier for lung nodule detection; the false positive was minimized by the rule-based technique [82]. The improved ResNet architecture yields efficient results for lung nodule detection than the classical ResNet architecture [20]. An inception v3 CNN model was proposed for the non-small cell lung cancer prediction and classification and yields efficient results when compared with the classical CNN architecture [21].

Decap5 baseline descriptor is employed for chest pathology detection. With the help of CNN and GIST, the best performance was achieved in [85]. For the identification of various categories of pathologies, CNN was implemented for X-ray images. The fusion of Decap and PiCodes provides the best performance in terms of sensitivity, specificity and AUC [86].

For the detection of cancer in the Basal cell, the deep learning convolutional autoencoder was used for the automatic detection of classification of Basal Cell Carcinoma in the histopathology images [87]. The deep belief architecture with self-supervised SVM was used in the detection of melanoma in the dermoscopic images [88]. And also deep convolutional neural network was used for the melanoma detection using the standard camera images [89]. The convolutional CNN was suggested for the detection of skin lesion border in the dermoscopy images [90].

The high dimensional gene data is processed using a shallow neural network for detection of cancer in the colon of the histopathological images [91]. The Neighboring Ensemble Predictor (NEP) coupled with CNN for the detection of nuclei in the colorectal adenocarcinoma images [92]. The regularized ensemble framework

deep learning for the detection of bowel cancer with an overall accuracy of 96.5% [93].

Deep learning convolutional neural network uses leave-one-case-out for 3D segmentation of bladder. The 3D information from computed tomography provides more accurate information for calculating tumour size change in response to treatment [95]. Deep learning algorithm was demonstrated for augmented cystoscopic detection of bladder cancer. The diagnostic yield of cystoscopy is enhanced by high sensitivity and specificity [96]. In [97], authors explore the applicability of deep learning CNN to differentiate inside and outside of the bladder. Jaccard index for DL-CNN-based likelihood map and level sets is better when compared to Haar-feature-based likelihood map and level sets and CLASS with LCR) method.

The deep learning CNN framework for the diagnosis of cervical dysplasia from multimodal clinical and image features and this model gives performance metrics of sensitivity 87.83% [98]. The XmasNet based on CNN was used for the classification of prostate cancer lesions in the multiparametric MRI images [99]. The deep convolutional neural network for the detection of prostate cancer in the MRI images attained an AUC of 0.897 [100]. In the photoacoustic images, the deep learning neural net was used for the classification of the malignant, benign and normal prostate cancer [101].

The deep fully connected convolutional neural network is used for the segmentation of the brain tumour in the BRATS dataset and acquired high performance with cascaded architecture [102]. For the classification of the brain tumour using the deep learning Multi-layer stacked denoising Autoencoder network and achieved an accuracy of 98.4% [103]. The automatic detection and segmentation of the brain (HGG and LGG) tumour using U-Net based deep convolutional network and combined detection achieved dice similarity index is 0.86 [104]. The conditional Generative Adversarial Network (cGAN) used for the detection of brain tumour, with a U-Net generative model and Markovian GAN" discriminative model [105]. The features extracted with PCA and deep learning classifier combined with DWT classifies normal, glioblastoma, sarcoma and metastatic bronchogenic carcinoma tumours and achieves a classification ratio 96.9% [106]. For the brain tumour segmentation, the fully convolutional neural networks (FCNNs) and Conditional Random Fields (CRFs) are integrated for the fast image patch segmentation in the Multimodal Brain Tumor Image Segmentation Challenge (BRATS) [107]. In [108, 109], a detailed analysis of different types of deep learning neural network architectures have been discussed and these works motivated to perform this study on applications of deep learning architectures in medical image processing.

3.6 Conclusion

The machine learning algorithms play an inevitable role in disease diagnosis and for therapeutic applications. Out of the various machine learning techniques, deep learning architectures are gaining much prominence due to their superior features

and is widely used in tumour segmentation and classification. This chapter initially highlights the machine learning algorithms architectures in cancer diagnosis, and then the different types of deep learning architectures are described. A comparative analysis of the pros and cons of various deep learning architectures and the works related to deep learning architectures in tumour segmentation and classification of various anatomical organs are described in detail. The outcome of this work is beneficial for researchers working in machine learning algorithms for disease diagnosis and treatment planning.

Acknowledgements The authors would like to acknowledge the support provided by Nanyang Technological University under NTU Ref: RCA-17/334 for providing the medical images and supporting us in the preparation of the manuscript. Parasuraman Padmanabhan and Balazs Gulyas also acknowledge the support from Lee Kong Chian School of Medicine and Data Science and AI Research (DSAIR) centre of NTU (Project Number ADH-11/2017-DSAIR) and the support from the Cognitive NeuroImaging Centre (CONIC) at NTU.

References

1. W.H. Wolberg, W.N. Street, O.L. Mangasarian, Machine learning techniques to diagnose breast cancer from image-processed nuclear features of fine needle aspirates. Cancer Lett. **77**(2–3), 163–171 (1994)
2. H. Bhavsar, A. Ganatra, A comparative study of training algorithms for supervised machine learning. Int. J. Soft Comput. Eng. (IJSCE) **2**(4), 2231–2307 (2012)
3. J.G. Lee, S. Jun, Y.W. Cho, H. Lee, G.B. Kim, J.B. Seo, N. Kim, Deep learning in medical imaging: general overview. Korean J. Radiol. **18**(4), 570–584 (2017)
4. K. Polat, S. Güneş, Breast cancer diagnosis using least square support vector machine. Digit. Sig. Proc. **17**(4), 694–701 (2007)
5. S. Şahan, K. Polat, H. Kodaz, S. Güneş, A new hybrid method based on fuzzy-artificial immune system and k-NN algorithm for breast cancer diagnosis. Comput. Biol. Med. **37**(3), 415–423 (2007)
6. M.F. Akay, Support vector machines combined with feature selection for breast cancer diagnosis. Expert Syst. Appl. **36**(2), 3240–3247 (2009)
7. I. Maglogiannis, E. Zafiropoulos, I. Anagnostopoulos, An intelligent system for automated breast cancer diagnosis and prognosis using SVM based classifiers. Appl. Intell. **30**(1), 24–36 (2009)
8. A. Osareh, B. Shadgar, Machine learning techniques to diagnose breast cancer, in *2010 5th International Symposium on Health Informatics and Bioinformatics* (IEEE, 2010), pp. 114–120
9. A.C. Tan, D. Gilbert, Ensemble machine learning on gene expression data for cancer classification, in *Proceedings of New Zealand Bioinformatics Conference*, Te Papa, Wellington, New Zealand, 13–14 Feb 2003
10. R. Zhang, G.B. Huang, N. Sundararajan, P. Saratchandran, Multicategory classification using an extreme learning machine for microarray gene expression cancer diagnosis. IEEE/ACM Trans. Comput. Biol. Bioinf. **4**(3), 485–495 (2007)
11. S.L. Hsieh, S.H. Hsieh, P.H. Cheng, C.H. Chen, K.P. Hsu, I.S. Lee, Z. Wang, F. Lai, Design ensemble machine learning model for breast cancer diagnosis. J. Med. Syst. **36**(5), 2841–2847 (2012)
12. D. Lavanya, D.K. Rani, Analysis of feature selection with classification: Breast cancer datasets. Indian J. Comput. Sci. Eng. (IJCSE) **2**(5), 756–763 (2011)

13. R. Ramos-Pollán, M.A. Guevara-López, C. Suárez-Ortega, G. Díaz-Herrero, J.M. Franco-Valiente, M. Rubio-Del-Solar, N. González-De-Posada, M.A. Vaz, J. Loureiro, I. Ramos, Discovering mammography-based machine learning classifiers for breast cancer diagnosis. J. Med. Syst. **36**(4), 2259–2269 (2012)
14. K. Rajesh, S. Anand, Analysis of SEER dataset for breast cancer diagnosis using C4. 5 classification algorithm. Int. J. Adv. Res. Comput. Commun. Eng. **1**(2), 2278–1021 (2012)
15. G.I. Salama, M. Abdelhalim, M.A. Zeid, Breast cancer diagnosis on three different datasets using multi-classifiers. Breast Cancer (WDBC) **32**(569), 2 (2012)
16. S. Kharya, Using data mining techniques for diagnosis and prognosis of cancer disease (2012). arXiv preprint arXiv:1205.1923
17. L.G. Ahmad, A.T. Eshlaghy, A. Poorebrahimi, M. Ebrahimi, A.R. Razavi, Using three machine learning techniques for predicting breast cancer recurrence. J. Health Med. Inform. **4**(124), 3 (2013)
18. Y. Gal, R. Islam, Z. Ghahramani, Deep bayesian active learning with image data, in *Proceedings of the 34th International Conference on Machine Learning*, vol. 70 (2017), pp. 1183–1192. JMLR.org
19. S. Liu, H. Zheng, Y. Feng, W. Li, Prostate cancer diagnosis using deep learning with 3D multiparametric MRI, in *Medical Imaging 2017: Computer-Aided Diagnosis*, vol. 10134 (International Society for Optics and Photonics, 2017), p. 1013428
20. K. Kuan, M. Ravaut, G. Manek, H. Chen, J. Lin, B. Nazir, C. Chen, T.C. Howe, Z. Zeng, V. Chandrasekhar, Deep learning for lung cancer detection: tackling the kaggle data science bowl 2017 challenge (2017). arXiv preprint arXiv:1705.09435
21. N. Coudray, P.S. Ocampo, T. Sakellaropoulos, N. Narula, M. Snuderl, D. Fenyö, A.L. Moreira, N. Razavian, A. Tsirigos, Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. Nat. Med. **24**(10), 1559–1567 (2018)
22. A. Teramoto, T. Tsukamoto, Y. Kiriyama, H. Fujita, Automated classification of lung cancer types from cytological images using deep convolutional neural networks. BioMed Res. Int. **2017** (2017)
23. Q. Song, L. Zhao, X. Luo, X. Dou, Using deep learning for classification of lung nodules on computed tomography images. J. Healthc. Eng. **2017** (2017)
24. W. Sun, B. Zheng, W. Qian, Automatic feature learning using multichannel ROI based on deep structured algorithms for computerized lung cancer diagnosis. Comput. Biol. Med. **1**(89), 530–539 (2017)
25. R. Platania, S. Shams, S. Yang, J. Zhang, K. Lee, S.J. Park, Automated breast cancer diagnosis using deep learning and region of interest detection (bc-droid), in *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics* (2017), pp. 536–543
26. A.S. Becker, M. Marcon, S. Ghafoor, M.C. Wurnig, T. Frauenfelder, A. Boss, Deep learning in mammography: diagnostic accuracy of a multipurpose image analysis software in the detection of breast cancer. Invest. Radiol. **52**(7), 434–440 (2017)
27. B.E. Bejnordi, M. Veta, P.J. Van Diest, B. Van Ginneken, N. Karssemeijer, G. Litjens, J.A. Van Der Laak, M. Hermsen, Q.F. Manson, M. Balkenhol, O. Geessink, Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. Jama **318**(22), 2199–2210 (2017)
28. N. Antropova, B.Q. Huynh, M.L. Giger, A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. Med. Phys. **44**(10), 5162–5171 (2017)
29. J. Schmidhuber, Deep learning in neural networks: an overview. Neural Netw. **1**(61), 85–117 (2015)
30. C.C. Aggarwal, Neural Networks and Deep Learning, vol. 10 (Springer, Berlin, 2018), pp. 978-983
31. G.C. Pereira, M. Traughber, R.F. Muzic, The role of imaging in radiation therapy planning: past, present, and future. BioMed Res. Int. **2014** (2014)

32. G. Litjens, T. Kooi, B.E. Bejnordi, A.A. Setio, F. Ciompi, M. Ghafoorian, J.A. Van Der Laak, B. Van Ginneken, C.I. Sánchez, A survey on deep learning in medical image analysis. Med. Image Anal. **1**(42), 60–88 (2017)
33. J. Dai, Y. Li, K. He, J. Sun, R-fcn: object detection via region-based fully convolutional networks, in *Advances in Neural Information Processing Systems* (2016), pp. 379–387
34. M.I. Razzak, S. Naz, A. Zaib, Deep learning for medical image processing: overview, challenges and the future, in *Classification in BioApps* (Springer, Cham, 2018), pp. 323–350
35. S.K. Zhou, H. Greenspan, D. Shen (eds.), *Deep Learning for Medical Image Analysis* (Academic Press, 2017)
36. A. Oliver, A. Odena, C.A. Raffel, E.D. Cubuk, I. Goodfellow, Realistic evaluation of deep semi-supervised learning algorithms, in *Advances in Neural Information Processing Systems* (2018), pp. 3235–3246
37. R. Raina, A. Madhavan, A.Y. Ng, Large-scale deep unsupervised learning using graphics processors, in *Proceedings of the 26th Annual International Conference on Machine Learning* (2009), pp. 873–880
38. W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, F.E. Alsaadi, A survey of deep neural network architectures and their applications. Neurocomputing **19**(234), 11–26 (2017)
39. Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition. Proc. IEEE **86**(11), 2278–2324 (1998)
40. M.Z. Alom, T.M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M.S. Nasrin, B.C. Van Esesn, A.A. Awwal, V.K. Asari, The history began from alexnet: a comprehensive survey on deep learning approaches (2018). arXiv preprint arXiv:1803.01164
41. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition (2014). arXiv preprint arXiv:1409.1556
42. H.T. Mustafa, J. Yang, M. Zareapoor, Multi-scale convolutional neural network for multi-focus image fusion. Image Vis. Comput. **1**(85), 26–35 (2019)
43. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A.C. Berg, Imagenet large scale visual recognition challenge. Int. J. Comput. Vis. **115**(3), 211–252 (2015)
44. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2015), pp. 1–9
45. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 770–778
46. S. Targ, D. Almeida, K. Lyman, ResNet in ResNet: generalizing residual architectures (2016). arXiv preprint arXiv:1603.08029
47. J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2015), pp. 3431–3440
48. O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation, in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer, Cham, 2015), pp. 234–241
49. Ö. Çiçek, A. Abdulkadir, S.S. Lienkamp, T. Brox, O. Ronneberger, 3D U-Net: learning dense volumetric segmentation from sparse annotation, in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer, Cham, 2016), pp. 424–432
50. Z. Wang, Q. She, T.E. Ward, Generative adversarial networks: a survey and taxonomy (2019). arXiv preprint arXiv:1906.01529
51. A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, A.A. Bharath, Generative adversarial networks: an overview. IEEE Sig. Process. Mag. **35**(1), 53–65 (2018)
52. S. Hochreiter, The vanishing gradient problem during learning recurrent neural nets and problem solutions. Int. J. Uncertainty Fuzziness Knowl. Based Syst. **6**(02), 107–116 (1998)
53. P. Liu, X. Qiu, X. Huang, Recurrent neural network for text classification with multi-task learning (2016). arXiv preprint arXiv:1605.05101

54. M. Loey, A. El-Sawy, H. El-Bakry, Deep learning autoencoder approach for handwritten arabic digits recognition (2017). arXiv preprint arXiv:1706.06720
55. S.A. Thomas, A.M. Race, R.T. Steven, I.S. Gilmore, J. Bunch, Dimensionality reduction of mass spectrometry imaging data using autoencoders, in *2016 IEEE Symposium Series on Computational Intelligence (SSCI)* (IEEE, 2016), pp. 1–7
56. M.A. Keyvanrad, M.M. Homayounpour, A brief survey on deep belief networks and introducing a new object oriented toolbox (DeeBNet) (2014). arXiv preprint arXiv:1408.3264
57. G.E. Hinton, Deep belief networks. Scholarpedia **4**(5), 5947 (2009)
58. M. Kallenberg, K. Petersen, M. Nielsen, A.Y. Ng, P. Diao, C. Igel, C.M. Vachon, K. Holland, R.R. Winkel, N. Karssemeijer, M. Lillholm, Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring. IEEE Trans. Med. Imaging **35**(5), 1322–1331 (2016)
59. H. Wang, A.C. Roa, A.N. Basavanhally, H.L. Gilmore, N. Shih, M. Feldman, J. Tomaszewski, F. Gonzalez, A. Madabhushi, Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features. J. Med. Imaging **1**(3), 034003 (2014)
60. M.G. Ertosun, D.L. Rubin, Probabilistic visual search for masses within mammography images using deep learning, in *2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (IEEE, 2015), pp. 1310–1315
61. N. Dhungel, G. Carneiro, A.P. Bradley, Automated mass detection in mammograms using cascaded deep learning and random forests, in *2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA)* (IEEE, 2015), pp. 1–8
62. J. Arevalo, F.A. González, R. Ramos-Pollán, J.L. Oliveira, M.A. Lopez, Representation learning for mammography mass lesion classification with convolutional neural networks. Comput. Methods Programs Biomed. **1**(127), 248–257 (2016)
63. S. Albarqouni, C. Baur, F. Achilles, V. Belagiannis, S. Demirci, N. Navab, Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images. IEEE Trans. Med. Imaging **35**(5), 1313–1321 (2016)
64. B.Q. Huynh, H. Li, M.L. Giger, Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. J. Med. Imaging **3**(3), 034501 (2016)
65. R. Turkki, N. Linder, P.E. Kovanen, T. Pellinen, J. Lundin, Antibody-supervised deep learning for quantification of tumor-infiltrating immune cells in hematoxylin and eosin stained breast cancer samples. J. Pathol. Inform. **7** (2016)
66. J. Gallego-Posada, D.A. Montoya-Zapata, O.L. Quintero-Montoya, Detection and diagnosis of breast tumors using deep convolutional neural networks. Med. Phys. **43**, 3705 (2016)
67. H. Chen, Q. Dou, X. Wang, J. Qin, P.A. Heng, Mitosis detection in breast cancer histology images via deep cascaded networks, in *Thirtieth AAAI Conference on Artificial Intelligence* (2016)
68. J. Arevalo, F.A. González, R. Ramos-Pollán, J.L. Oliveira, M.A. Lopez, Convolutional neural networks for mammography mass lesion classification, in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (IEEE, 2015), pp. 797–800
69. A. Akselrod-Ballin, L. Karlinsky, S. Alpert, S. Hasoul, R. Ben-Ari, E. Barkan, A region based convolutional network for tumor detection and classification in breast mammography, in *Deep Learning and Data Labeling for Medical Applications* (2016, Springer, Cham), pp. 197–205
70. R.K. Samala, H.P. Chan, L. Hadjiiski, M.A. Helvie, J. Wei, K. Cha, Mass detection in digital breast tomosynthesis: Deep convolutional neural network with transfer learning from mammography. Med. Phys. **43**(12), 6654–6666 (2016)
71. Z. Jiao, X. Gao, Y. Wang, J. Li, A deep feature based framework for breast masses classification. Neurocomputing **12**(197), 221–231 (2016)
72. Y. Liu, K. Gadepalli, M. Norouzi, G.E. Dahl, T. Kohlberger, A. Boyko, S. Venugopalan, A. Timofeev, P.Q. Nelson, G.S. Corrado, J.D. Hipp, Detecting cancer metastases on gigapixel pathology images (2017). arXiv preprint arXiv:1703.02442

73. A. Cruz-Roa, H. Gilmore, A. Basavanhally, M. Feldman, S. Ganesan, N.N. Shih, J. Tomaszewski, F.A. González, A. Madabhushi, Accurate and reproducible invasive breast cancer detection in whole-slide images: a deep learning approach for quantifying tumor extent. Sci. Rep. **18**(7), 46450 (2017)
74. M.H. Yap, G. Pons, J. Martí, S. Ganau, M. Sentís, R. Zwiggelaar, A.K. Davison, R. Martí, Automated breast ultrasound lesions detection using convolutional neural networks. IEEE J. Biomed. Health Inform. **22**(4), 1218–1226 (2017)
75. A. Das, U.R. Acharya, S.S. Panda, S. Sabut, Deep learning-based liver cancer detection using watershed transform and Gaussian mixture model techniques. Cogn. Syst. Res. **1**(54), 165–175 (2019)
76. P. Devi, P. Dabas, Liver tumour detection using artificial neural networks for medical images. Int. J. Innov. Res. Sci. Technol. **2**(3), 34–38 (2015)
77. W. Li, Automatic segmentation of liver tumour in CT images with deep convolutional neural networks. J. Comput. Commun. **3**(11), 146 (2015)
78. D. Kumar, A. Wong, D.A. Clausi, Lung nodule classification using deep features in CT images, in *2015 12th Conference on Computer and Robot Vision* (IEEE, 2015), pp. 133–138
79. W. Sun, B. Zheng, W. Qian, Computer aided lung cancer diagnosis with deep learning algorithms, in *Medical Imaging 2016: Computer-Aided Diagnosis*, vol. 9785 (International Society for Optics and Photonics, 2016), p. 97850Z
80. R. Gruetzemacher, A. Gupta, Using deep learning for pulmonary nodule detection & diagnosis, in *Twenty-Second Americas Conference on Information Systems*, San Diego (2016)
81. R. Golan, C. Jacob, J. Denzinger, Lung nodule detection in CT images using deep convolutional neural networks, in *2016 International Joint Conference on Neural Networks (IJCNN)* (IEEE, 2016), pp. 243–250
82. K. Hirayama, J.K. Tan, H. Kim, Extraction of GGO candidate regions from the LIDC database using deep learning, in *2016 16th International Conference on Control, Automation and Systems (ICCAS)* (IEEE, 2016), pp. 724–727
83. S. Bhatia, Y. Sinha, L. Goel, Lung cancer detection: a deep learning approach, in *Soft Computing for Problem Solving* (Springer, Singapore, 2019), pp. 699–705
84. A. Teramoto, H. Fujita, O. Yamamuro, T. Tamaki, Automated detection of pulmonary nodules in PET/CT images: ensemble false-positive reduction using a convolutional neural network technique. Med. Phys. **43**(6Part1), 2821–2827 (2016)
85. Y. Bar, I. Diamant, L. Wolf, S. Lieberman, E. Konen, H. Greenspan, Chest pathology detection using deep learning with non-medical training, in *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)* (IEEE, 2015), pp. 294–297
86. Y. Bar, I. Diamant, L. Wolf, H. Greenspan, Deep learning with non-medical training used for chest pathology identification, in *Medical Imaging 2015: Computer-Aided Diagnosis*, vol. 9414 (International Society for Optics and Photonics, 2015), p. 94140V
87. A.A. Cruz-Roa, J.E. Ovalle, A. Madabhushi, F.A. Osorio, A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection, in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer, Berlin, 2013), pp. 403–410
88. A. Masood, A. Al-Jumaily, K. Anam, Self-supervised learning model for skin cancer diagnosis, in *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)* (IEEE, 2015), pp. 1012–1015
89. M.H. Jafari, N. Karimi, E. Nasr-Esfahani, S. Samavi, S.M. Soroushmehr, K. Ward, K. Najarian, Skin lesion segmentation in clinical images using deep learning, in *2016 23rd International Conference on Pattern Recognition (ICPR)* (IEEE, 2016), pp. 337–342
90. P. Sabouri, H. Gholam Hosseini, Lesion border detection using deep learning, in *2016 IEEE Congress on Evolutionary Computation (CEC)* (IEEE, 2016), pp. 1416–1421
91. H. Chen, H. Zhao, J. Shen, R. Zhou, Q. Zhou, Supervised machine learning model for high dimensional gene data in colon cancer detection, in *2015 IEEE International Congress on Big Data* (IEEE, 2015), pp. 134–141

92. K. Sirinukunwattana, S.E. Raza, Y.W. Tsang, D.R. Snead, I.A. Cree, N.M. Rajpoot, Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. IEEE Trans. Med. Imaging **35**(5), 1196–1206 (2016)
93. C.L. Chen, A. Mahjoubfar, L.C. Tai, I.K. Blaby, A. Huang, K.R. Niazi, B. Jalali, Deep learning in label-free cell classification. Sci. Rep. **15**(6), 21471 (2016)
94. X. Yuan, L. Xie, M. Abouelenien, A regularized ensemble framework of deep learning for cancer detection from multi-class, imbalanced training data. Pattern Recogn. **1**(77), 160–172 (2018)
95. K.H. Cha, L. Hadjiiski, R.K. Samala, H.P. Chan, E.M. Caoili, R.H. Cohan, Urinary bladder segmentation in CT urography using deep-learning convolutional neural network and level sets. Med. Phys. **43**(4), 1882–1896 (2016)
96. K.H. Cha, L.M. Hadjiiski, R.K. Samala, H.P. Chan, R.H. Cohan, E.M. Caoili, C. Paramagul, A. Alva, A.Z. Weizer, Bladder cancer segmentation in CT for treatment response assessment: application of deep-learning convolution neural network—a pilot study. Tomography **2**(4), 421 (2016)
97. E. Shkolyar, X. Jia, T.C. Chang, D. Trivedi, K.E. Mach, M.Q. Meng, L. Xing, J.C. Liao, Augmented bladder tumor detection using deep learning. Eur. Urol. **76**(6), 714–718 (2019)
98. T. Xu, H. Zhang, X. Huang, S. Zhang, D.N. Metaxas, Multimodal deep learning for cervical dysplasia diagnosis. in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer, Cham, 2016), pp. 115–123
99. S. Liu, H. Zheng, Y. Feng, W. Li, Prostate cancer diagnosis using deep learning with 3D multiparametric MRI, in *Medical Imaging 2017: Computer-Aided Diagnosis*, vol. 10134 (International Society for Optics and Photonics, 2017), p. 1013428
100. Y.K. Tsehay, N.S. Lay, H.R. Roth, X. Wang, J.T. Kwak, B.I. Turkbey, P.A. Pinto, B.J. Wood, R.M. Summers, Convolutional neural network based deep-learning architecture for prostate cancer detection on multiparametric magnetic resonance images, in *Medical Imaging 2017: Computer-Aided Diagnosis*, vol. 10134 (International Society for Optics and Photonics, 2017), p. 1013405
101. A.R. Rajanna, R. Ptucha, S. Sinha, B. Chinni, V. Dogra, N.A. Rao, Prostate cancer detection using photoacoustic imaging and deep learning. Electron. Imaging **2016**(15), 1–6 (2016)
102. M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.M. Jodoin, H. Larochelle, Brain tumor segmentation with deep neural networks. Med. Image Anal. **1**(35), 18–31 (2017)
103. Z. Xiao, R. Huang, Y. Ding, T. Lan, R. Dong, Z. Qin, X. Zhang, W. Wang, A deep learning-based segmentation method for brain tumor in MR images, in *2016 IEEE 6th International Conference on Computational Advances in Bio and Medical Sciences (ICCABS)* (IEEE, 2016), pp. 1–6
104. H. Dong, G. Yang, F. Liu, Y. Mo, Y. Guo, Automatic brain tumor detection and segmentation using u-net based fully convolutional networks, in *Annual Conference on Medical Image Understanding and Analysis* (Springer, Cham, 2017), pp. 506–517
105. M. Rezaei, K. Harmuth, W. Gierke, T. Kellermeier, M. Fischer, H. Yang, C. Meinel, A conditional adversarial network for semantic segmentation of brain tumor, in *International MICCAI Brainlesion Workshop* (Springer, Cham, 2017), pp. 241–252
106. H. Mohsen, E.S. El-Dahshan, E.S. El-Horbaty, A.B. Salem, Classification using deep learning neural networks for brain tumors. Future Comput. Inform. J. **3**(1), 68–71 (2018)
107. X. Zhao, Y. Wu, G. Song, Z. Li, Y. Zhang, Y. Fan, A deep learning model integrating FCNNs and CRFs for brain tumor segmentation. Med. Image Anal. **1**(43), 98–111 (2018)
108. K. Munir, H. Elahi, A. Ayub, F. Frezza, A. Rizzi, Cancer diagnosis using deep learning: a bibliographic review. Cancers **11**(9), 1235 (2019)
109. M.Z. Alom, T.M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M.S. Nasrin, M. Hasan, B.C. Van Essen, A.A. Awwal, V.K. Asari, A state-of-the-art survey on deep learning theory and architectures. Electronics **8**(3), 292 (2019)

Chapter 4
Classification of Canine Fibroma and Fibrosarcoma Histopathological Images Using Convolutional Neural Networks

İsmail Kırbaş and Özlem Özmen

Abstract Within the scope of the study, a high-performance convolutional network model that can classify canine fibroma and fibrosarcoma tumors based on 200 high resolution real histopathological microscope images has been developed. In order to determine the network performance, the well-known network models (VGG16, ResNET50, MobileNet-V2 and Inception-V3) were subjected to training and testing according to the same hardware and training criteria. While comparing the models, 13 different performance criteria were used and performance calculations were made for each model separately. The results obtained seem extremely satisfactory. Compared to its counterparts, the proposed network model (FibroNet) contains fewer trainable items, while achieving a much higher performance value and training time is shorter than others. Thanks to low prediction error rate achieved with FibroNET network using real data, it seems possible to develop an artificial intelligence-based reliable decision support system that will facilitate surgeons' decision making in practice.

Keywords Canine · Convolutional neural networks · Fibroma · Fibrosarcoma · FibroNET · Histopathology · Tumor classification

4.1 Introduction

Most of the canine neoplasms localized to the skin [1]. Fibromas are benign fibroid tumours and they are composed of fibrous tissue (fibroblasts, fibrocytes and collagen fibers) mainly fibrocytes that produced collagen is abundant in collagenous stroma [2]. Tumours arising from mesenchymal connective tissues. While these tumours diagnosed in all domestic species, they primarily occur in aged dogs. The larger dogs such as Boxers, Golden Retrievers and Doberman Pinschers are predisposed

İ. Kırbaş (✉) · Ö. Özmen
Burdur Mehmet Akif Ersoy University, Burdur, Turkey
e-mail: ismailkirbas@mehmetakif.edu.tr

Ö. Özmen
e-mail: ozlemoz@mehmetakif.edu.tr

U. Kose and J. Alzubi (eds.), *Deep Learning for Cancer Diagnosis*,
Studies in Computational Intelligence 908,
https://doi.org/10.1007/978-981-15-6321-8_4

breeds for benign and malign connective tissue tumours. The main localization area of the tumours are head and extremities of the dogs. Fibromas (benign fibrous tissue tumours) are generally solid, firm, raised, often hairless masses and originating under the skin. The most commonly tumours are round to oval in shape and arise from intradermal or subcutaneous tissues. Generally, these tumours have firm, rubbery and whitish cut surface. Fibromas classified as fibroma durum (firm) and fibroma molle (soft) types according to their consistency especially amount of collagen. They are well-circumscribed single solid tumours without a capsule [2, 3]. Fibromas are connective tissue tumours composed of fibrocytes and fibroblasts that produce different amounts of collagen fibers. The collagen is generally look like interwoven fascicles and seldomly like as whorls. The neoplastic fibrocytes are spindle shape, uniform, with normochromatic big oval to round nuclei and an indistinct cytoplasm. At the histopathological examination mitotic figures are seldomly observed. Generally, the collagen fibers are brightly eosinophilic and easily distinguishable at the intercellular areas. Fibromas are slow growing and benign tumours, for that reason treatment is optional. But if the appearance or growth changes occur, they must be surgically removed [4].

Fibrosarcomas are malignant tumours of the fibrous connective tissue that generally arise from the skin and subcutaneous connective tissues [5]. These tumours have variable presentations depending on etiopathogenesis, species, localisation and age of the host. Fibroblasts are the main cell type of fibrosarcomas. They are the most common malignant soft-tissue tumours in pet animals especially cats and dogs. Most of the tumours are solid and can occur everywhere on the body, but main localisation areas are head and extremities in dogs. Predisposed breeds are Golden Retrievers, Gordon Setters, Brittany Spaniels, Irish Wolfhounds, and Doberman Pinschers. At the clinical examination size and appearance of the fibrosarcomas may change case to case. They are generally slow-growing, malignant tumours most often arising under the skin surface may look lumpy and appear nodular. These tumours developed in the subcutaneous soft tissues, or subjacent fat may require palpation to identify. Fibrosarcomas generally appear as firm masses, fleshy lesions involving the dermis and subcutaneous connective tissue and tumoral cells generally invade musculature along fascial planes. If tumours occurred as multiple masses, they are often localised within the same anatomic region [4].

Fibrosarcomas usually not encapsulated and often invade underlying muscles. Generally obvious interwoven fascicular pattern is seen at the cut surface of the tumour. These tumours can be well differentiated and consist oval to round tumoral cells that arranged in herringbone or interwoven patterns. Cytoplasm of the cells is scant, and nuclei are spindle shaped with inconspicuous nucleoli. In poorly differentiated and anaplastic tumours, anisocytotic or pleomorphic cells may be seen commonly. In aggressive tumours multinucleated giant cells with ovoid or polygonal shaped and prominent nucleoli are characteristic. Numerous mitotic figures may be seen in undifferentiated tumours. Lymphocytes and mononuclear inflammatory cells infiltrations may be seen in some fibrosarcomas [2].

In dogs, fibrosarcomas are invasive tumours that spread into surrounding tissues; only about 10% of these tumours metastasizes to the body of the effected animal.

Generally, the tumour easily diagnosed because of the characteristic appearance. However, in some cases differentiation from leiomyosarcomas and peripheral nerve sheath tumours, can be problematic. At the differential diagnosis amount of the collagenous stroma may be used and it is more abundant in fibrosarcomas than the other tumours. For this aim Masson's trichrome stain may be used for to make easily visible of the collagen. In contrast to fibrosarcomas, leiomyosarcomas have more rounded shape of nuclei at the histopathological examination, but this feature is not very reliable [2]. For treatment, wide and deep surgically extirpation of the mass is the best choice. Because it is hard to determine the tumour's edges during surgery, complete removal may be difficult and recurrence is common. More than 70% relapse occurs during the 1-year period from the initial surgery even if surgical removal is successful. Although invasion is common, metastasis of the fibrosarcomas are rare. Followup radiation therapy and chemotherapy generally essential after surgical removal of the mass especially for invasive and inoperable tumours [6].

4.2 Data and Methodology

The original dataset used in this paper is provided by Burdur Mehmet Akif Ersoy University, Faculty of Veterinary Medicine, Department of Pathology in JPG format. The dataset consists of 200 high resolution (1920 × 1200) histology microscopic images with 4× to 40× magnification. These images are labeled with two classes: fibroma and fibrosarcoma, and each class consists of 100 images.

Totally 27 tumoral mass (13 fibroma and 9 fibrosarcoma) were examined in this study. The materials collected from archive of the Department of Pathology between 2000 and 2019. Age of the dogs were changing 1–14 years and different breeds with both genders. Figure 4.1 shows examples of fibroma and fibrosarcoma removed surgically.

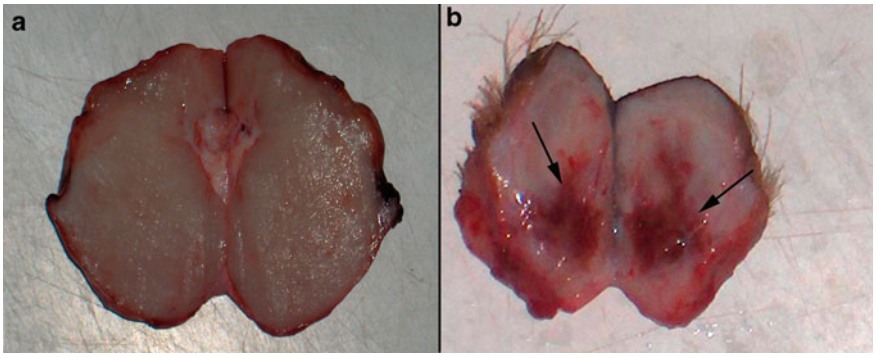


Fig. 4.1 **a** Cut surface of the fibroma no necrosis and hemorrhage; **b** cut surface of the fibrosarcoma hemorrhagic areas (arrows)

The relationship and the pattern between the classes are tried to be determined without any future selection process. For this purpose, the data on hand must be derived and reproduced without losing the information they carry. Keras library [7] can automatically perform image augmentation by taking various parameters as input. In our study, we performed zoom in-out, rotation, mirroring, horizontal and vertical flipping operations on the microscope images and augmentation operation was performed using these techniques. A balanced distribution was achieved by using an equal number of samples in both classes. Since the learning process takes place as unsupervised in deep learning algorithms, many examples are needed. Table 4.1 gives details of the number of pictures produced for both classes.

Images obtained from microscopes are 1920 × 1200 pixels. However, they were reduced to 224 × 224 dimensions since it will take a long time to process such large data and will require high amount of processor power. Then the scaled photographs were augmented by techniques such as rotating, flipping, mirroring, zooming in and out. Figure 4.2 shows the scaling and picture augmentation processes of the original histopathological images.

Accordingly, a total of 1462 images were produced for training purposes for the fibroma class, while 368 images were produced for use in testing phase. Similarly,

Table 4.1 Augmented fibroma and fibrosarcoma classes

Classes	Augmented training images	Augmented test images	Total images
Fibroma	1445	349	1794
Fibrosarcoma	1462	378	1840
Total	2907	727	3634

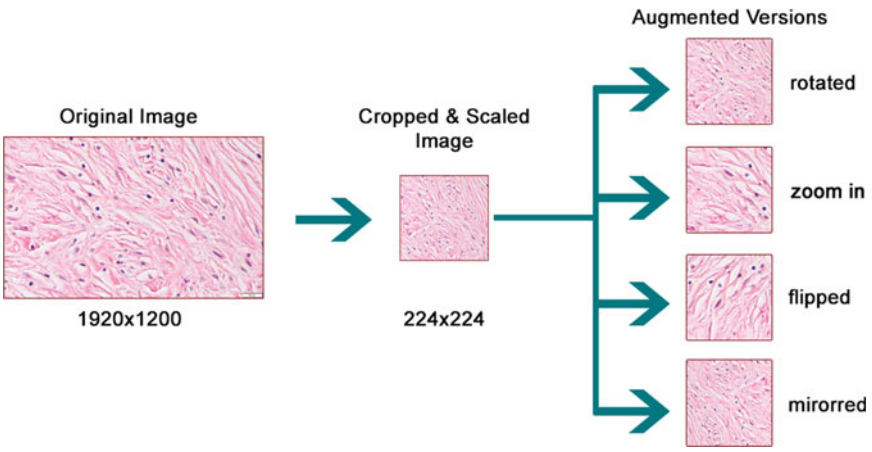


Fig. 4.2 Histopathological image preprocessing operations including scaling and augmentation operations

1445 pictures were produced for training in the fibrosarcoma class and 359 pictures were used for testing purposes.

Deep neural networks are usually created by connecting ten or more layers of artificial neural networks in succession. There are dozens of neurons in each layer, and after the neurons in the layers perform the activation process, it transmits the result obtained to the neurons in the next layer [8]. In other words, the output value of each layer becomes the input value for the layer after it. If the output value of each neuron in a layer is sent as the input value for all neurons in the next layer, then fully connected is used for the layers. However, this increases the load of the process and can cause excessive memorization of the data. Overfitting the data is a situation for the goal of revealing and learning the relations or patterns among the data. For this reason, some connections between layers are interrupted from time to time during the training process so that the excessive memorization can be prevented. This technique is called drop out [9].

Convolutional neural networks are often used for image classification problems [10]. Pictures to be processed on these networks are often made much smaller than their original size (usually 224 × 224 pixels). If the colour information is insignificant for the problem that is tried to be solved, the original pictures are made black and white even if they are coloured. This ensures that fewer neurons are used in the input layer and the processing load is reduced. The images, which are pre-processed by shrinking and generally black and white, are then subjected to a matrix operation called convolution, transferring data from the first layer to a second layer of neurons with fewer neurons.

Within the scope of the study, networks which are frequently used in the literature such as VGG 16 [11], Resnet50 (He et al. 2016), MobileNETv2 [12] and Inception V3 [13] were handled and used with transfer learning technique. In the transfer learning method, the networks previously trained with different data and for different classification procedures are adapted for the target dataset by adapting the output layer number to the desired problem. The aim is to quickly adapt an existing CNN solution with a complex layer architecture to another problem at low processing and time cost.

In addition to the existing networks, a unique CNN, called FibroNET, was created within the scope of the study and performance comparisons were made for five different network structures.

4.3 Proposed Convolutional Neural Network FibroNET

The proposed neural network architecture FibroNET is shown in Fig. 4.3. While feature extraction is carried out with convolutional layers, max pooling layers are used for size reduction. The global average pooling means that a 3 dimensions tensor is sliced and computed the average the slices in order to have only 1-dimension tensor or a vector so that the dense layer deal with only a vector. Dense layer is connected

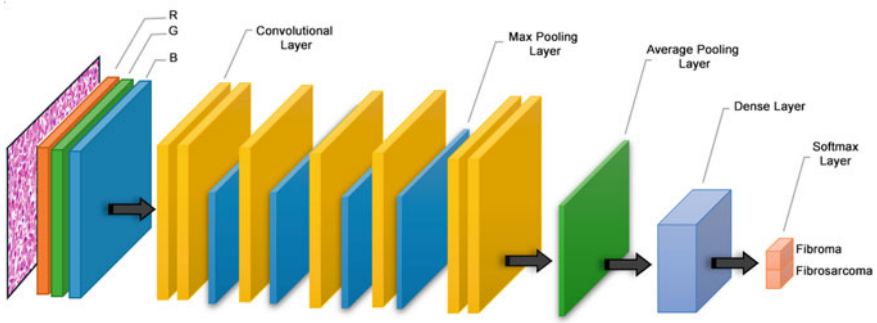


Fig. 4.3 The proposed Convolutional Neural Network (FibroNET) architecture

Table 4.2 The FibroNET layer architecture

1	Layer type	Filter size (kernel size, filter, activation)	Number of parameter
0	Input	224, 224, 3	0
1	Convolution 2D	3, 16, relu	448
2	Convolution 2D	3, 30, relu	4350
3	Max pooling	2	0
4	Convolution 2D	3, 60, relu	16,260
5	Max pooling	2	0
6	Convolution 2D	3, 90, relu	48,690
7	Max pooling	2	0
8	Convolution 2D	3, 110, relu	89,210
9	Max pooling	2	0
10	Convolution 2D	3, 130, relu	128,830
11	Convolution 2D	1, 40, relu	5240
12	Global average pooling	40	0
13	Dense	2, softmax	82

to the last layer that uses the softmax function and has two outputs fibroma and fibrosarcoma.

The FibroNET network consists of 14 layers in total. 7 of these layers are called convolutional layers and they make feature extracting. After convolution layers, max pooling layers are used to reduce the size. Global average pooling layer also has been added behind the last convolution layer so that the dense layer input can be converted to a vector. Table 4.2 shows the FibroNET network structure and its layer layouts in detail.

Because of providing high performance for the training of deep learning networks, Google Colab environment was used and TPU was selected for computing opera-

Table 4.3 Performance evaluation results for five different CNNs

Performance metrics	VGG-16	ResNET50	MobileNETv2	Inception v3	Proposed model (FibroNET)
R2 score	−1.0364	−0.0048	0.0811	0.6975	0.9595
Mean absolute error	0.5089	0.4992	0.4764	0.2221	0.0328
Mean squared error	0.5089	0.2511	0.2296	0.0755	0.0100
ACC score	0.4910	0.5089	0.6905	0.9367	0.9889
ROC AUC score	0.5	0.5	0.7611	0.9684	0.9994
LOG loss	17.5781	0.6954	0.6522	0.2972	0.0435
Train accuracy	0.4991	0.5009	0.5057	0,6708	0.9856
Train loss	7.6817	0.6956	0.6940	0,6547	0.0438
Test accuracy	0.4910	0.5089	0.5158	0,6905	0.9890
Test loss	7.8057	0.6945	0.6923	0,6522	0.0436
Training time (s)	3551	1754	926	4142	191
Epochs	21	21	21	61	136
Total params	14,764,866	25,638,714	3,540,986	23,853,786	293,110
Trainable params	14,764,866	25,585,594	3,506,874	23,819,354	293,110

tions. Performing calculations with TPU, provides much shorter processing times compared to CPU and GPU alternatives.

First of all, five deep learning models are trained to maximize the validation accuracy parameter. Number of training epoch is determined by early stopping technique. When 1.10–3 advances in validation accuracy become unavailable after 20 trials between the iterations, the training iteration was terminated and the most successful values obtained during the training period were recorded with the model structure and weights. Afterwards, the model obtained was tested with the test data and the performance criteria given in Table 4.3 were calculated and compared.

Simonyan and Zisserman [11] proposed a convolutional neural network model called VGG16 for ImageNet competition and it has been achieved 92.7% top-5 test accuracy. The model achieves in ImageNet dataset consists of over 14 million images belonging to 1000 classes and VGG16 was trained for weeks and was using NVIDIA Titan Black GPU's. Unlike VGG16, which is a simple network model, it uses 2 or 3 of the convolution additions. It is converted to an attribute vector with 4096 neurons in the fully connected (FC) layer. Softmax performance of 1000 classes is calculated at the output of two FC layers. Approximately 138 million parameters are calculated in the network. As in other models, the height and width dimensions of the matrices

from the input to the output decrease while the depth value (number of channels) increases.

ResNet50 [14], which has a different logic than the previous models, where the network model started to really deepen; residual value consists of adding the blocks (residual block) that feed to the next layers. With this feature, ResNet50 is no longer a classic model. Theoretically, the performance is thought to increase as the number of layers increases in the model. However, it has been experienced that this is not really the case. Based on this, ResNet50 model was developed within Microsoft.

MobileNet [15] network structure has been developed especially for use in image recognition processes on mobile devices. It requires high processing power to perform the convolution process. Since this process power is very difficult to obtain in mobile devices, the convolution process in MobileNET network structure is performed by dividing it into two sub-processes. Unlike version 1, MobileNetV2 has a second connection type called residual connection in the residual block and has fewer trainable parameters than MobileNetV1.

Szegedy et al. developed [13] the GoogLeNet network with Google support which is known as Inception-v1. Inception-v3 architecture also consists of batch normalization and ReLU functions are used after convolution. It is 42 layers deep and its computation cost is 2.5 times higher than GoogLeNet.

In Table 4.3, there is a detailed comparison of five different CNN models with 13 different metrics in the same hardware configurations. The best values are underlined according to the criteria in Table 4.3.

4.4 Results and Discussion

In this study, evaluation of histopathological images obtained from canines with deep artificial neural networks and determination of tumor type were studied. For this purpose, 200 high resolution microscope images were used, and the dataset was made up of 3634 images in total after the pre-processing and augmentation process.

Within the scope of the study, a convolutional neural network (FibroNET) was developed. It has significant advantages compared to its counterparts. It was determined that the training time is low, the number of trainable parameters is low, and the test accuracy is very high.

When Table 4.3 is examined, the proposed model has proven itself and it has been shown in numbers that it is far superior to other well-known models. VGG-16, ResNet50 and MobileNetV2 models used in the study did not produce a successful result within the specified training criteria. While the Inception-V3 model has a higher performance than the counted models, it has the longest training period. Our proposed model FibroNET, has fewer items than alternatives and therefore it can be trained in a shorter time. Its test and train accuracy values are at the top when compared to the others.

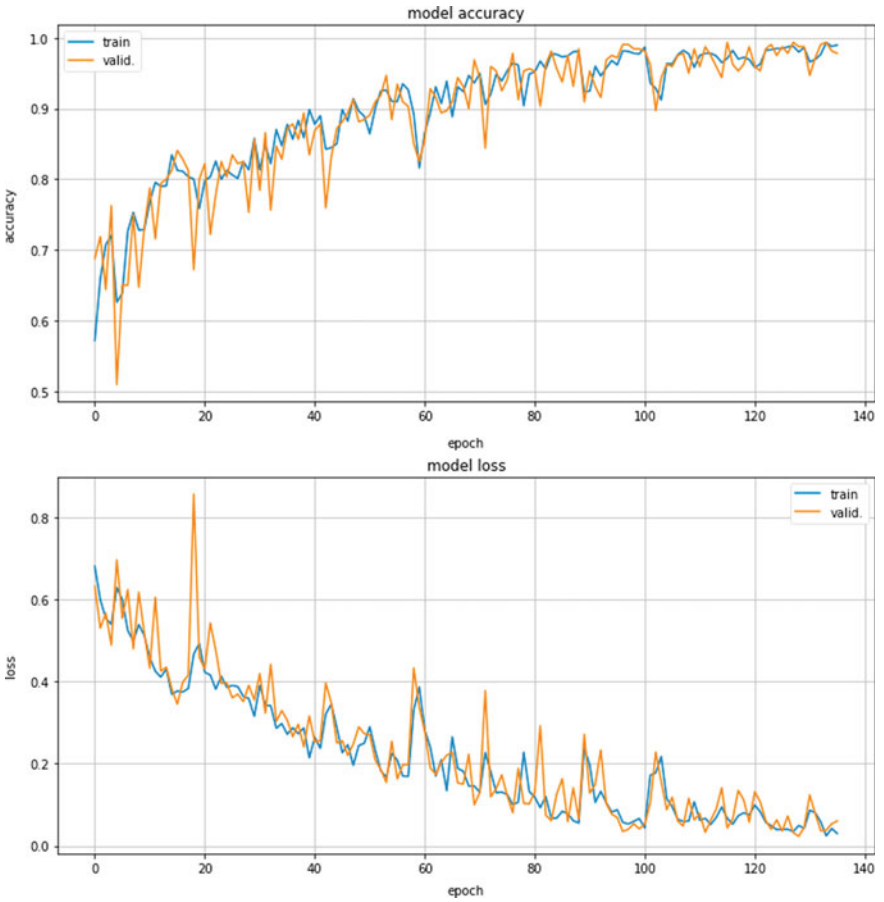


Fig. 4.4 Model accuracy and model loss graphics for FibroNET

Figure 4.4 shows the model accuracy and model loss graphs calculated during the training and validation stages of the FibroNET network. As can be seen in Fig. 4.4, as the number of training iterations (number of epoch) increases, the value of the loss parameter converges to 0, while the accuracy value approaches 1. This indicates that the proposed model is well trained and the prediction error rate will be low.

In Fig. 4.5, the confusion matrix results are obtained by testing the FibroNET network with the test data. According to result values, while the model correctly predicted all fibrosarcoma results, it incorrectly labelled as fibroma only 8 results that should be detected as fibrosarcoma.

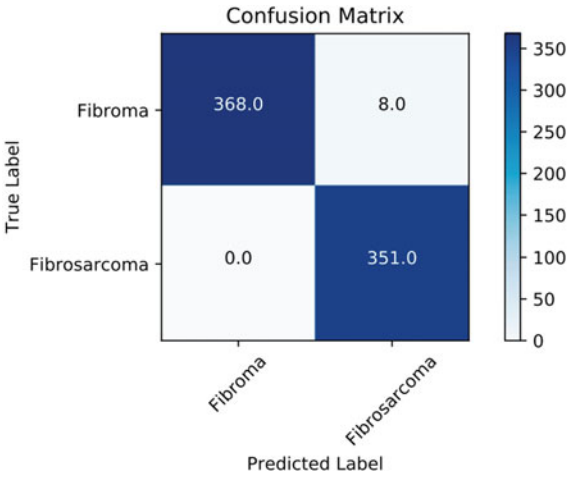


Fig. 4.5 Confusion matrix results for FibroNET

The tumor classification problem seems to be a very difficult problem even for specialists. The conducted study shows that the proposed model (FibroNET) can achieve high performance on real data. We believe that, in future studies, by developing a larger dataset, it will be possible to develop an artificial intelligence-based reliable decision support system that will facilitate surgeons' decision making in practice.

References

1. D.E. Bostock, Neoplasms of the skin and subcutaneous tissues in dogs and cats. Br. Vet. J. **142**(1), 1–19 (1986)
2. M.H. Goldschmidt, M.J. Hendrick, Tumors of the skin and soft tissues, in *Tumors in Domestic Animals*, ed. by D.J. Meuten (Iowa State Press, Ames, Iowa, USA, 2008), pp. 45–117
3. F. Fernandes, H.N. Soufen, B.M. Ianni, E. Arteaga, F.J. Ramires, C. Mady, Primary neoplasms of the heart. Clinical and histological presentation of 50 cases. Arquivos brasileiros de cardiologia **76**(3), 231–237 (2001)
4. Vet Manual Merck & Co. Veterinary Manual. Library Catalog: www.msdvetmanual.com
5. C.E. Doige, S.E. Weisbrode, Disease of bone and joints, in *Thomson's Special Veterinary Pathology*, ed. by W.W. Carlton, M.D. McGavin (Mosby, St. Louis, 1995), p. 446
6. S. Mukaratirwa, J. Chipunza, S. Chitanga, M. Chimonyo, E. Bhebhe, Canine cutaneous neoplasms: prevalence and influence of age, sex and site on the presence and potential malignancy of cutaneous neoplasms in dogs from Zimbabwe. J. South Afr. Vet. Assoc. **76**(2), 59–62 (2005)
7. R. Atienza, *Advanced Deep Learning with Keras: Apply Deep Learning Techniques, Autoencoders, GANs, Variational Autoencoders, Deep Reinforcement Learning, Policy Gradients, and More* (Packt Publishing, 2018)
8. N. Buduma, N. Locascio, *Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms*, 1st edn. (O'Reilly Media, Inc., 2017)
9. A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, in *Advances in Neural Information Processing Systems*, vol. 25, ed. by F.

Pereira, C.J.C. Burges, L. Bottou, K.Q. Weinberger (Curran Associates, Inc., Red Hook, 2012), pp. 1097–1105
10. F. Chollet, *Deep Learning with Python*, 1st edn. (Manning Publications Co., USA, 2017)
11. K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:1409.1556 [cs], April 2015
12. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L. Chen, MobileNetV2: inverted residuals and linear bottlenecks, in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2018), pp. 4510–4520. ISSN: 1063-6919
13. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2015), pp. 1–9
14. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016), pp. 770–778
15. A. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications* (2017)

Chapter 5
Evaluation of Big Data Based CNN Models in Classification of Skin Lesions with Melanoma

Prasitthichai Naronglerdrit and Iosif Mporas

Abstract That chapter presents a method regarding diagnosis of pigmented skin lesions using convolutional neural networks. The architecture is modeled over convolutional neural networks and it is evaluated using new CNN models as well as re-trained modification of pre-existing CNN models were used. The experimental results showed that CNN models pre-trained on big datasets for general purpose image classification when re-trained in order to identify skin lesion types offer more accurate results when compared to convolutional neural network models trained explicitly from the dermatoscopic images. The best performance was achieved by re-training a modified version of ResNet-50 convolutional neural network with accuracy equal to 93.89%. Analysis on skin lesion pathology type was also performed with classification accuracy for melanoma and basal cell carcinoma being equal to 79.13 and 82.88%, respectively.

5.1 Introduction

Over the last years skin cancer cases are becoming to a greater extent common with more than 5 million people been diagnosed with it in the USA. Roughly three quarters of the skin cancer cases resulting in death have been caused from melanomas thus making it the most dangerous type of cancer of the skin [1]. The region of the spot over the skin, which is affected, is named as the area of lesion, and the lesions of the skin are the first clinical symptoms of diseases like melanoma, chickenpox and others [2]. In the sense of the melanoma, the growth rate within most cases is slow

P. Naronglerdrit (✉)
Department of Computer Engineering, Faculty of Engineering at Sriracha, Kasetsart University Sriracha Campus, Chonburi, Thailand
e-mail: prasitthichai@eng.src.ku.ac.th

I. Mporas
School of Engineering and Computer Science, University of Hertfordshire, Hatfield AL10 9AB, UK
e-mail: i.mporas@herts.ac.uk

U. Kose and J. Alzubi (eds.), *Deep Learning for Cancer Diagnosis*,
Studies in Computational Intelligence 908,
https://doi.org/10.1007/978-981-15-6321-8_5

enough to allow to be removed relatively easily and with low hazard and cost if the melanoma lesion has been detected in an initial stage. When the pathology of the skin is found in an early stage then the rate of patients to survive is more than 95% [3, 4], but on the other hand if the pathology of the skin is found at a late stage then treatment is more difficult with the survival rate being dramatically reduced to approximately 13% [5].

Currently visual examination is considered as the standard clinical procedure for detection and identification of skin lesions [6]. One of the most popular protocols for medical assessment of the lesions on the skin is the 'ABCD' protocol, according to which the dermatologist examines the (A)symmetry, the (B)order, the (C)olour, and the (D)iameter of the spot on the skin. As regards asymmetry it is examined whether segments of the skin lesion area diverge in shape or not. Regarding border, it is examined if the edges of the corresponding area of the skin are irregular or blurred and, in some cases, if they have notches. The colour might be uneven and with various colourizations of black, brown and pink. As regards the diameter of the skin lesion, most melanomas have diameter of at least six millimetres while any change of the size, the shape or the diameter being essential information the corresponding medical staff needs to be aware of [7]. Medical examination of the skin of patients is carried out by doctors, general practitioners or with expertise in dermatology, and typically it is an examination requiring a lot of time. Moreover, the diagnosis of lesions of skin is a very subjective procedure [8], as diagnosis can be imprecise or incorrect or could outcome to quite dissimilar diagnosis even if it has been done by dermatologists with lots of experience. In addition to subjectivity, diagnosis supported by AI-based computational tools can result in reduced diagnostic accuracy when diagnostic tools are utilized by doctors with no previous appropriate training or enough experience as shown in previous study [2].

Due to the limits and the difficulty of the clinical assessment of skin lesions as described above and also due to the number of cases related to skin diseases which are rising every year the development of accurate computer-aided tools for automated dermatoscopic image processing and for classification of skin lesions for the analysis of skin spots related to melanomas or different skin pathologies is necessary. Furthermore, the utilization of the cutting edge technological achievements in the areas of digital image processing, AI with emphasis in deep learning (machine learning) and big data applied to achieve skin cancer detection can make it feasible to allow successful diagnosis for skin lesions avoiding the need for body contact with the skin of the patient [2] or even perform the diagnosis from distance by sending to the dermatologist a photo of the corresponding skin area through the Internet. Moreover, the cost of making a diagnosis and treating of other than melanoma types of skin cancer is noteworthy as for instance only in 2010 it costed AU$ 511 million to Australia [9] and at the same time the overall cost to the healthcare system of the United States for the cases of melanoma is calculated at approximately $ seven billion annually [10]. In countries such as Australia, and the United Kingdom (UK), there has been a lot of attention given in general practitioners to improve their diagnostic

performance in order to be able to precisely identify and diagnose skin pathologies and cancers. That is especially due to the cost caused to national healthcare systems and aiming at retaining their sustainability [11].

Dermatoscopy (which is also reported in the literature as the epiluminescence microscopy, or dermoscopy), is a skin imaging method which is non-invasive. Dermatoscopic images are illuminated and also magnified images of an area over the skin which offer high resolution analysis of the lesions on the skin [8], nevertheless software-aided analysis of skin images has a number of issues making it a difficult task, like the usually not quite high contrast between the region of the spot and the normal skin region around it which results in automatic segmentation of the skin lesion areas with limited accuracy. Except this, another problem is that quite often the melanoma and the non-melanoma skin spots appear to be visually similar between them thus increasing making it difficult to distinguish a lesion corresponding to melanoma from a non-melanoma area of the skin. In addition, the variation of the characteristics of the skin from one patient to another, e.g. the skin colour and the presence of veins and natural hairs, result in melanoma cases with highly different colour and texture characteristics [3].

Different approaches for automatic image-based skin lesion classification have been published in the literature over the last decade. Rubegni et al. [12] used neural networks to perform automated classification of pigmented skin lesions, specifically melanoma vs nevi, and 13 image features representing the lesions were used parameterizing the geometry, the colourization, the textures and colour clusters of the skin lesions. In [4] Sheha et al. presented a classification architecture of malignant and benign skin lesions and evaluated the performance of classifiers like neural networks, support vector machines (SVMs) and k-nearest neighbours when using geometry based, chromatic and texture image features. In [10] Alcón et al. performed skin lesion binary classification with two skin lesion types, malignant and benign lesions, with the corresponding images being collected from standard digital cameras and evaluated decision trees using the AdaBoost metaclassifier and image features related to the symmetry, edges/boundaries, colour and texture of the skin spot area. Refianti et al. presented in [13] a binary classification methodology for melanoma skin cancer and non-melanoma skin cancer and the methodology relied on a convolutional neural network (CNN) structure. In [14] Prigent and colleagues proposed classification methodology for the skin hyper-pigmentation-lesions, by employing support vector machines for classification, and the multi-spectral analysis of skin images done in the context of the classification. In [11], Chang et al. introduced a software diagnostic algorithm to process malignant and benign lesions of skin which is extracting shape, texture and colour features of the lesion area and was tested using support vector machines as classifier on a dataset consisting of typical digital photos of skin. In [15] Kawahara et al. classified the skin images thanks to a model of convolutional neural networks with deep features extracted from the CNNs. Capdehourat et al. presented an approach for classification in [16], by using decision trees, and also support vector machines over pigmented skin lesions. They evaluated their approach over a set of

dermatoscopic images. In [2], Sumithra et al. presented an evaluation regarding k-nearest neighbours, and the support vector machines, in the context of skin image segmentation and lesions classification with parameterization of the lesion area (by using texture, colour, and the histogram-based image features). Lehman et al. in [17] proposed an approach which is fusing three convolutional neural networks, each of them using distinct image transformations. In [3] Li et al. presented an architecture for classification of skin lesion method which was utilizing parameters extracted from deep convolutional neural networks and were combined with colour, morphological and textural features of the skin lesion area. Mahdiraji et al. in [6] calculated boundary features which were processed by neural networks to perform classification of skin lesions using sets of images acquired from conventional cameras. Mishra et al. [18] and Jafari et al. [19] presented segmentation approaches which can isolate the skin spot (i.e. the region of interest) from the remaining non-lesion part of the skin images with the employment of convolutional neural networks. In [20] Rehman et al. utilized convolutional neural networks to perform image segmentation followed by classification of the extracted image segments using artificial neural networks. Kaymak et al. in [21] presented a two-step deep learning algorithm for hierarchical classification of the lesions of malignant-pigmented skin with the first step being classification of the skin images to non-melanocytic or melanocytic, and the second stage being classification of the malignant types.

In the present paper we examine different models of convolutional neural networks for identification of skin lesion types captured from digital skin images using dermatoscopy. Except this we evaluate well known CNN models for image classification which have been pre-trained using big data and for the needs of skin lesion classification have been modified and re-trained. The remaining of this chapter is organized as follows. In Sect. 5.2 the block diagram and steps of the evaluated methodology for the automatic classification and diagnosis of lesions of skin is described. Section 5.3 presents the evaluation setup and in Sect. 5.4 the experimental evaluation results are presented. Finally, in Sect. 5.5 conclusions are given.

5.2 Classification and Diagnosis of Skin Lesions Using CNNs

For the classification and diagnosis of skin lesions convolutional neural networks were utilized for the classification stage after using image processing algorithms for image pre-processing and segmentation of the dermatoscopic image. The dermatoscopic images are in a first step pre-processed, with the pre-processing step consisting of filtering of the image using a median filter and then removing of any detected hair. Following the pre-processing step image segmentation is performed to the skin images in order to detect and segment the area of the skin lesion, i.e. the region of interest (ROI) in the image. Once detecting the skin region of interest the corresponding bounding box is defined and then used to crop the skin lesion image segment

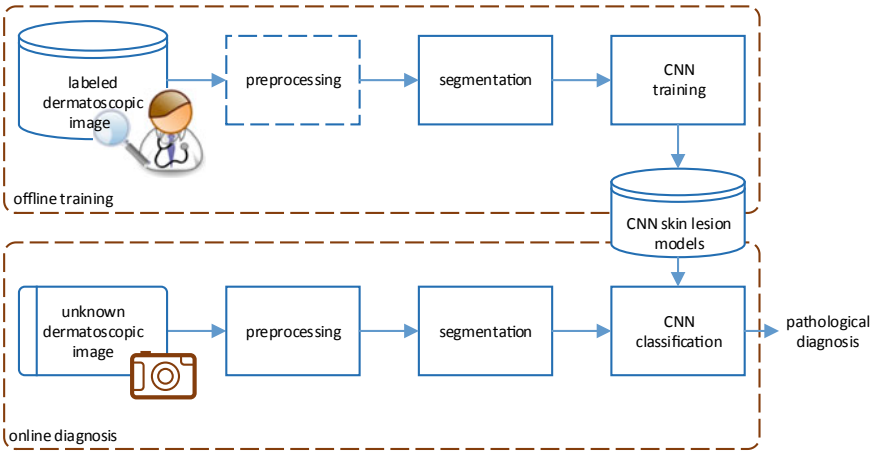


Fig. 5.1 Block diagram regarding the considered architecture for classification of skin lesions using convolutional neural networks

which is subsequently converted to a predefined image size and sent to a convolutional neural network for identification of the corresponding skin lesion type. The block diagram of the dermatoscopic images-based skin lesion classification using convolutional neural networks as classifier is illustrated in Fig. 5.1.

At the time of the offline training phase, a set of dermatoscopic images with known skin pathology labels, clinically verified by dermatologists, are pre-processed, segmented and the extracted ROIs used to train CNN models for classification of skin lesions. In the online diagnosis phase a new/unknown dermatoscopic image is processed following the same steps as in offline training and the CNN models assign a label to the corresponding skin lesion, thus performing online diagnosis. The pre-processing, segmentation and classification steps of the architecture are described in more detail below.

5.2.1 Dermatoscopic Image Pre-processing

Firstly, we performed pre-processing of the dermatoscopic images for the purpose of hair removal [16, 22]. To automatically remove hair from the images hair detection is needed and subsequently is followed by in painting of the images. In our implementation to detect the pixels of the dermatoscopic images that contained natural hair we utilized a median filter for smoothing and then applied bottom-hat filtering as in [23]. The bottom-hat (also referred to as black top-hat) filter is performing an image transformation, based on a morphological operation which is defined as

$$B_{\text{hat}}(f) = (f \cdot b) - f \tag{5.1}$$

here f corresponds to the grayscale image, b means the structuring element, which was selected to be disk-shaped in the present evaluation. As f and b are both discrete, and $\cdot$ means the morphological operator of closing, defined as

$$f \cdot b = (f \oplus b) \ominus b \tag{5.2}$$

with $\oplus$ is the dilation operator, and $\ominus$ is the erosion operator [24].

5.2.2 Dermatoscopic Image Segmentation

After pre-processing the skin images to detect and remove hair resizing to 256 × 256 pixels was applied and subsequently Gaussian filtering for smoothing of the pixel values of the images. As a next step, we used the active contour method as proposed by Chan-Vese [25] in order to segment the skin images to foreground and background. The active contour algorithm extracted the borders of the skin lesion and it was empirically found that 300 iterations was a fair trade-off between the achieved accuracy of segmentation vs the time needed for computations.

After the active contour segmentation, the detected skin lesions bounding box was found which had variable sizes. With the purpose of using the regions of interest, i.e. the segmented skin images, as input to a convolutional neural network all segmented images needed to be resized to a fixed size which was empirically selected equal to 64 × 64 pixels. Except this, pixels' values were rescaled to the range [0, 1] for each of the three colours (RGB) to decrease big errors during subsequent processing from the activation function (ReLU in our implementation) of the convolutional neural network.

5.2.3 CNN-Based Image Classification

The pre-processed, segmented and value-normalized skin images corresponding to the skin lesion areas (ROIs) were then sent as input to convolutional neural networks to classify them. The CNN classifier will assign a skin pathology label to each unknown dermatoscopic image so that automatic pathological diagnosis will be performed.

Table 5.1 Types of skin pathology, and the number of instances in the HAM10000 dataset [26, 27]

Type of the pathology	Total images
Actinic keratosis (akiec)	327
Basal cell carcinoma (bcc)	514
Dermatofibroma (df)	115
Melanoma (mel)	1113
Nevus (nv)	6705
Pigmented benign keratosis (bkl)	1099
Vascular lesions (vasc)	142
Total	10,015

5.3 Experimental Setup

5.3.1 Data Description

To evaluate the skin lesions classification accuracy, the HAM10000 [26, 27] dataset was used. The HAM10000 database is a big dataset of digital images with skin lesions, which consist of 10015 skin images and all images have been labelled with their corresponding pathology type as shown in Table 5.1. The dermatoscopic images of the dataset was gathered through a period of 20 years from two different sites, which are the skin cancer practice of Cliff Rosendahl in Australia-Queensland, and the Department of Dermatology at the Medical University of Vienna, in Austria.

5.3.2 New CNN Models

As a baseline classification algorithm to classify the dermatoscopic images we used convolutional neural networks (CNNs) trained explicitly from the HAM10000 dataset images. The new CNN models were constructed after optimizing the convolutional layers number as presented in Table 5.2. As seen in the Table 5.2, the CNN models were built by using the pipeline of convolution layers and pooling layers. In the sense of the convolution layers, the filter size used was 3×3 and stride equal to 1. The pooling layers were used for reducing the size of the input regarding the next convolution layer and have filter size 2×2 and stride equal to 2.

In addition, the other hyper-parameters which were used in the setup of the CNN architecture consist of ReLU activation function, pooling layer implemented using max pooling, the optimizer using Adam and the cross-entropy (categorical) used as loss function. After the last pooling layer and dense layer, a dropout layer was used to avoid any over-fitting in the neural network. The classification of pigmented skin lesions was performed by a fully connected layer and softmax activation function.

Table 5.2 New CNN model architectures evaluated for skin lesion classification

Layers	Output	Model 1	Model 2	Model 3	Model 4
Convolution	64 × 64	3 × 3 conv, stride 1	3 × 3 conv, stride 1	3 × 3 conv, stride 1	3 × 3 conv, stride 1
Pooling	32 × 32	2 × 2 max pool, stride 2	2 × 2 max pool, stride 2	2 × 2 max pool, stride 2	2 × 2 max pool, stride 2
Convolution	32 × 32		3 × 3 conv, stride 1	3 × 3 conv, stride 1	3 × 3 conv, stride 1
Pooling	16 × 16		2 × 2 max pool, stride 2	2 × 2 max pool, stride 2	2 × 2 max pool, stride 2
Convolution	16 × 16			3 × 3 conv, stride 1	3 × 3 conv, stride 1
Pooling	8 × 8			2 × 2 max pool, stride 2	2 × 2 max pool, stride 2
Convolution	8 × 8				3 × 3 conv, stride 1
Pooling	4 × 4				2 × 2 max pool, stride 2
Flatten		65536	16384	4096	1024
Dropout	0.5 dropout				
Dense	128 fully connected				
Dropout	0.5 dropout				
Classification	7 fully-connected, softmax				

The number of the model, i.e. 'Model-N', indicates the number of convolution layer and pooling layer sets in the CNN model

The CNN models were constructed using Tensorflow and Keras packets, and the training process used the CUDA GPU computing for acceleration. The CNN models were trained for 50 epochs and batch normalization size equal to 16.

5.3.3 Pre-trained CNN Models and Their Modifications

Except training new CNN models we investigated the usage of pre-trained CNN models for image classification that available with the Keras packet, in particular the VGG [28], MobileNet [29], DenseNet [30], and ResNet [31] pre-trained CNN models. The pre-trained CNN models, the architectures of which were modified and used for pigmented skin lesions classification, were originally trained using ImageNet [32], which is a large images dataset consisting of 1000 classes as employing a total of 1.2 million images for the training, and 50,000 ones for the validation.

The VGG [28] network was constructed by the Visual Geometry Group, University of Oxford for the needs of ILSVRC-2014 competition. The VGG architecture is

a sequential network which has 3 × 3 convolution layers with stride equal to 1 and 2 × 2 max-pooling layers with stride equal to 2. The original VGG network was trained by the ImageNet which has 1000 classes with the input image size being equal to 224 × 224. In this paper we have used the VGG network with a depth of 16 layers (noted as VGG16) to perform as a classification model. In the modified VGG16 network, the last max-pooling layer was truncated, and the output was connected to the fully connected layer with softmax activation function for implementing the classification task as presented in Table 5.3.

The MobileNet [29] is introduced by Google Inc. for mobile and embedded vision applications. The MobileNet uses the concept of depthwise separable convolutions in its light-weight architecture and to keep the channel of the network, the 1 × 1 filter is used as a pointwise convolution. The advantage of this architecture is the reduced number of computations needed, both during training the CNN model and during online testing. Table 5.4 shows the modified architecture of the MobileNet in which

Table 5.3 VGG16 [28] architecture and its modifications for skin lesion classification

Layers	Original VGG16		Modified VGG16	
	Output size	Layer's structure	Output size	Layer's structure
Convolution	224 × 224	3 × 3 conv, stride 1	64 × 64	3 × 3 conv, stride 1
Convolution	224 × 224	3 × 3 conv, stride 1	64 × 64	3 × 3 conv, stride 1
Pooling	112 × 112	2 × 2 max pool, stride 2	32 × 32	2 × 2 max pool, stride 2
Convolution	112 × 112	3 × 3 conv, stride 1	32 × 32	3 × 3 conv, stride 1
Convolution	112 × 112	3 × 3 conv, stride 1	32 × 32	3 × 3 conv, stride 1
Pooling	56 × 56	2 × 2 max pool, stride 2	16 × 16	2 × 2 max pool, stride 2
Convolution	56 × 56	3 × 3 conv, stride 1	16 × 16	3 × 3 conv, stride 1
Convolution	56 × 56	3 × 3 conv, stride 1	16 × 16	3 × 3 conv, stride 1
Convolution	56 × 56	3 × 3 conv, stride 1	16 × 16	3 × 3 conv, stride 1
Pooling	28 × 28	2 × 2 max pool, stride 2	8 × 8	2 × 2 max pool, stride 2
Convolution	28 × 28	3 × 3 conv, stride 1	8 × 8	3 × 3 conv, stride 1
Convolution	28 × 28	3 × 3 conv, stride 1	8 × 8	3 × 3 conv, stride 1
Convolution	28 × 28	3 × 3 conv, stride 1	8 × 8	3 × 3 conv, stride 1
Pooling	14 × 14	2 × 2 max pool, stride 2	4 × 4	2 × 2 max pool, stride 2
Convolution	14 × 14	3 × 3 conv, stride 1	4 × 4	3 × 3 conv, stride 1
Convolution	14 × 14	3 × 3 conv, stride 1	4 × 4	3 × 3 conv, stride 1
Convolution	14 × 14	3 × 3 conv, stride 1	4 × 4	3 × 3 conv, stride 1
Pooling	7 × 7	2 × 2 max pool, stride 2	2 × 2	2 × 2 max pool, stride 2
Flatten	25,088		2048	
Dense	4096			
Dense	4096			
Classification	1000 fully-connected, softmax		7 fully-connected, softmax	

Table 5.4 MobileNet [29] architecture and its modifications for skin lesion classification

Layers	Original MobileNet		Modified MobileNet	
	Output size	Layer's structure	Output size	Layer's structure
Convolution	112 × 112	3 × 3 × 32 conv, stride 2	32 × 32	3 × 3 × 32 conv, stride 2
Convolution	112 × 112	3 × 3 depthwise conv, stride 1	32 × 32	3 × 3 depthwise conv, stride 1
Convolution	112 × 112	1 × 1 × 64 conv, stride 1	32 × 32	1 × 1 × 64 conv, stride 1
Convolution	56 × 56	3 × 3 depthwise conv, stride 2	16 × 16	3 × 3 depthwise conv, stride 2
Convolution	56 × 56	1 × 1 × 128 conv, stride 1	16 × 16	1 × 1 × 128 conv, stride 1
Convolution	56 × 56	3 × 3 depthwise conv, stride 1	16 × 16	3 × 3 depthwise conv, stride 1
Convolution	56 × 56	1 × 1 × 128 conv, stride 1	16 × 16	1 × 1 × 128 conv, stride 1
Convolution	28 × 28	3 × 3 depthwise conv, stride 2	8 × 8	3 × 3 depthwise conv, stride 2
Convolution	28 × 28	1 × 1 × 256 conv, stride 1	8 × 8	1 × 1 × 256 conv, stride 1
Convolution	28 × 28	3 × 3 depthwise conv, stride 1	8 × 8	3 × 3 depthwise conv, stride 1
Convolution	28 × 28	1 × 1 × 256 conv, stride 1	8 × 8	1 × 1 × 256 conv, stride 1
Convolution	14 × 14	3 × 3 depthwise conv, stride 2	4 × 4	3 × 3 depthwise conv, stride 2
Convolution	14 × 14	1 × 1 × 512 conv, stride 1	4 × 4	1 × 1 × 512 conv, stride 1
5 × Convolution	14 × 14	3 × 3 depthwise conv, stride 1	4 × 4	3 × 3 depthwise conv, stride 1
5 × Convolution	14 × 14	1 × 1 × 512 conv, stride 1	4 × 4	1 × 1 × 512 conv, stride 1
Convolution	7 × 7	3 × 3 depthwise conv, stride 2	2 × 2	3 × 3 depthwise conv, stride 2
Convolution	7 × 7	1 × 1 × 1024 conv, stride 1	2 × 2	1 × 1 × 1024 conv, stride 1
Convolution	7 × 7	3 × 3 depthwise conv, stride 2	2 × 2	3 × 3 depthwise conv, stride 2
Convolution	7 × 7	1 × 1 × 1024 conv, stride 1	2 × 2	1 × 1 × 1024 conv, stride 1
Classification	1024	7 × 7 global average pool	4096	flatten
	1000 fully-connected, softmax		7 fully-connected, softmax	

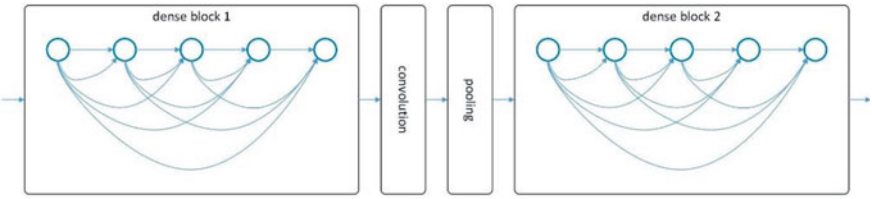


Fig. 5.2 DenseNet [30] with two dense blocks and transition layer between the dense blocks consisting of a convolution and pooling layer

the input size has changed from 224 × 224 to 64 × 64. All convolutional layers are then supported by batch normalization, and the ReLU for activation function.

The Dense Convolutional Network (DenseNet) [30] connects every layer to all next layers, while a conventional network connects only between previous layer to next layer. The block of the connections which connects each layer to other layers in feed-forward architecture noted as "dense block" as shown in Fig. 5.2.

The DenseNet has a 3 × 3 filter size at the convolution layer, and between dense blocks a 1 × 1 convolution layer supported by a 2 × 2 average-pooling layer are used as transition layer. Finally, softmax classifier is attached at the end of the last convolution layer. DenseNet has four dense blocks, but different number of convolution layers inside each block. In this paper, we used the DenseNet-121, DenseNet-169, and DenseNet-201 CNNs where the number after DenseNet name is indicating the total number of convolution layers, and also the fully-connected layers (as seen in Tables 5.5, 5.6 and 5.7, respectively).

ResNet [31] was introduced by Microsoft Research team based on the residual learning framework. In residual networks shortcut connections between stacks of convolution layers are inserted, as illustrated in Fig. 5.3. The ResNet has different number of convolution layers as shown in Tables 5.8, 5.9 and 5.10, respectively. For this evaluation we consider the ResNet-50, ResNet-101, and ResNet-152.

All pre-trained CNN models of Keras packet were originally trained using ImageNet dataset, and modified for the needs of the present evaluation by changing the input size from 224 × 224 to 64 × 64 and connecting the last convolution layer with a flatten layer supported by a fully-connected softmax layer. Afterwards, the modified pre-trained models were re-trained using the HAM10000 dataset for skin lesions classification. The re-training phase-stage of the pre-trained CNN models was done using 50 epochs and batch normalization size equal to 16 by using Adam optimizer and the cross-entropy (categorical) as loss function.

5.4 Evaluation Results

The related architecture-model over convolutional neural networks for classifying skin lesions (as presented in Sect. 5.2) was evaluated according to the experimental

Table 5.5 DenseNet-121 [30] architecture and its modifications for skin lesion classification

Layers	Original DenseNet-121		Modified DenseNet-121	
	Output size	Layer's structure	Output size	Layer's structure
Convolution	112 × 112	7 × 7 conv, stride 2	32 × 32	7 × 7 conv, stride 2
Pooling	56 × 56	3 × 3 max pool, stride 2	16 × 16	3 × 3 max pool, stride 2
Dense Block (1)	56 × 56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	16 × 16	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition (1)	56 × 56	1 × 1 conv	16 × 16	1 × 1 conv
Transition (1)	28 × 28	2 × 2 average pool, stride 2	8 × 8	2 × 2 average pool, stride 2
Dense Block (2)	28 × 28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	8 × 8	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition (2)	28 × 28	1 × 1 conv	8 × 8	1 × 1 conv
Transition (2)	14 × 14	2 × 2 average pool, stride 2	4 × 4	2 × 2 average pool, stride 2
Dense Block (3)	14 × 14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	4 × 4	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$
Transition (3)	14 × 14	1 × 1 conv	4 × 4	1 × 1 conv
Transition (3)	7 × 7	2 × 2 average pool, stride 2	2 × 2	2 × 2 average pool, stride 2
Dense Block (4)	7 × 7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	2 × 2	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$
Classification	1024	7 × 7 global average pool	4096	flatten
	1000 fully-connected, softmax		7 fully-connected, softmax	

protocol described in Sect. 5.3. At this point, general performance of the evaluated CNN models was measured in terms of classification accuracy, i.e.

$$c\text{Accuracy} = \frac{TrP + TrN}{TrP + TrN + FaP + FaN} \tag{5.3}$$

where TrP corresponds to the number of true positives, TrN means the number of true negatives, FaP is the number of false positives, and finally, FaN is the number of false negatives of the classified dermatoscopic images. For avoiding overlap between the training and testing subsets of evaluated data a cross validation evaluation setup using 10 folds was employed.

The evaluation results of identification of skin lesion types from dermatoscopic images using new CNN models and re-trained models of modifications of pre-existing

Table 5.6 DenseNet-169 [30] architecture and its modifications for skin lesion classification

Layers	Original DenseNet-169		Modified DenseNet-169	
	Output size	Layer's structure	Output size	Layer's structure
Convolution	112 × 112	7 × 7 conv, stride 2	32 × 32	7 × 7 conv, stride 2
Pooling	56 × 56	3 × 3 max pool, stride 2	16 × 16	3 × 3 max pool, stride 2
Dense Block (1)	56 × 56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	16 × 16	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition (1)	56 × 56	1 × 1 conv	16 × 16	1 × 1 conv
Transition (1)	28 × 28	2 × 2 average pool, stride 2	8 × 8	2 × 2 average pool, stride 2
Dense Block (2)	28 × 28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	8 × 8	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition (2)	28 × 28	1 × 1 conv	8 × 8	1 × 1 conv
Transition (2)	14 × 14	2 × 2 average pool, stride 2	4 × 4	2 × 2 average pool, stride 2
Dense Block (3)	14 × 14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	4 × 4	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$
Transition (3)	14 × 14	1 × 1 conv	4 × 4	1 × 1 conv
Transition (3)	7 × 7	2 × 2 average pool, stride 2	2 × 2	2 × 2 average pool, stride 2
Dense Block (4)	7 × 7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	2 × 2	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$
Classification	1664	7 × 7 global average pool	6656	flatten
	1000 fully-connected, softmax		7 fully-connected, softmax	

CNN models initially trained from big data are presented in Table 5.11. The best classification accuracy of the new CNN models and the best classification accuracy of the modified pre-existing CNN models are shown in bold-text.

As seen in Table 5.11 the top performing CNN model trained only with HAM10000 data ('New' CNN models) was CNN model three with classification accuracy equal to 76.83%. Among the evaluated pre-existing and modified CNN models the best classification accuracy was observed by the ResNet-50 with the accuracy equal to 93.89%, followed by the ResNet-152, the MobileNet, the DenseNet-169 and the DenseNet-121 CNN models with classification accuracies equal to 93.52%, 91.32%, 91.30% and 91.21%, respectively. The remaining big data-based CNN models achieved classification accuracy lower than 90%. All CNN models pre-trained from big datasets for general purpose image classification outperformed the four evaluated CNN models trained explicitly from the evaluation dataset for skin lesion

Table 5.7 DenseNet-201 [30] architecture and its modifications for skin lesion classification

Layers	Original DenseNet-201		Modified DenseNet-201	
	Output size	Layer's structure	Output size	Layer's structure
Convolution	112 × 112	7 × 7 conv, stride 2	32 × 32	7 × 7 conv, stride 2
Pooling	56 × 56	3 × 3 max pool, stride 2	16 × 16	3 × 3 max pool, stride 2
Dense Block (1)	56 × 56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	16 × 16	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition (1)	56 × 56	1 × 1 conv	16 × 16	1 × 1 conv
Transition (1)	28 × 28	2 × 2 average pool, stride 2	8 × 8	2 × 2 average pool, stride 2
Dense Block (2)	28 × 28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	8 × 8	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition (2)	28 × 28	1 × 1 conv	8 × 8	1 × 1 conv
Transition (2)	14 × 14	2 × 2 average pool, stride 2	4 × 4	2 × 2 average pool, stride 2
Dense Block (3)	14 × 14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	4 × 4	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
Transition (3)	14 × 14	1 × 1 conv	4 × 4	1 × 1 conv
Transition (3)	7 × 7	2 × 2 average pool, stride 2	2 × 2	2 × 2 average pool, stride 2
Dense Block (4)	7 × 7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	2 × 2	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$
Classification	1920	7 × 7 global average pool	7680	Flatten
	1000 fully-connected, softmax		7 fully-connected, softmax	

classification, thus indicating that the availability of large data collections is essential in training robust skin lesion classification models as well as that within the application of skin lesion classification CNN transferability is possible, i.e. pre-trained convolutional neural networks can successfully be modified and re-trained on smaller datasets of dematoscopic images and offer competitive classification accuracy.

As the second step, we analyzed the classification accuracy on skin pathology category level for the best performing CNN classification model, i.e. ResNet-50. The confusion matrix of the skin lesion classification (using ResNet-50) is shown in Table 5.12.

As it can be observed within Table 5.12, the highest classification accuracy was achieved for the actinic keratosis (akiec) skin pathology with accuracy equal to 97.90%, followed by nevus (nv) pathology with accuracy 90.14% and pigmented

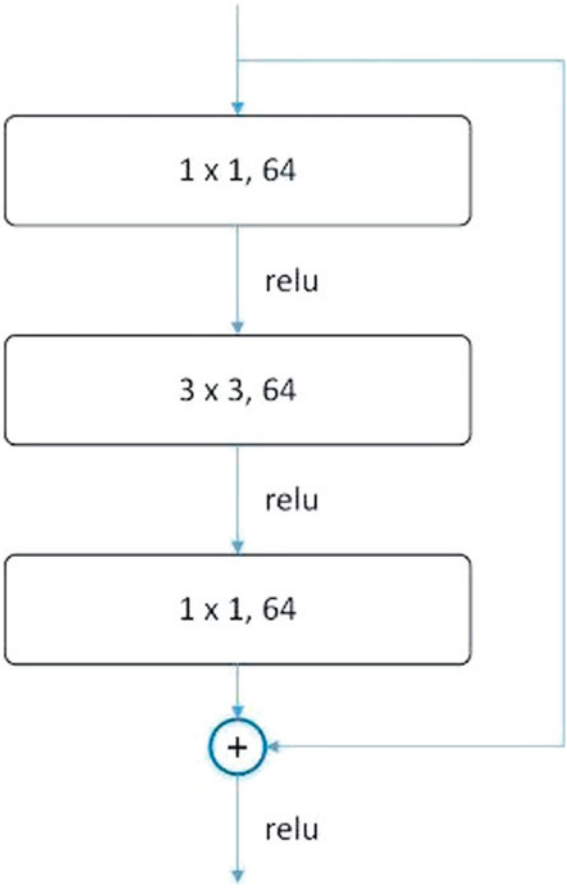


Fig. 5.3 Shortcut connection between a stack of three convolution layers of a residual network

benign keratosis (bkl) skin pathology with the accuracy 87.33%. The skin pathologies which were found the most difficult among the evaluated ones to be correctly classified were melanoma (mel) and basal cell carcinoma (bcc) with classification accuracies equal to 79.13 and 82.88%, respectively. The experimental results indicate the potential of computer-aided diagnosis regarding the skin lesions, by employing image processing and machine learning technology. Moving from the achieved classification accuracy, both in average (93.89%) and for the case of melanoma (79.13%), we deem that with the development of larger dermatoscopic image datasets and the progress in designing new CNN architectures, new and more accurate classification models are feasible in support of the diagnosis of dermatologists.

Table 5.8 ResNet-50 [31] architectures and their modifications for skin lesion classification

Layers	Original ResNet-50		Modified ResNet-50	
	Output size	Layer's structure	Output size	Layer's structure
Convolution	112×112	7×7 conv, stride 2	32×32	7×7 conv, stride 2
Pooling	56×56	3×3 max pool, stride 2	16×16	3×3 max pool, stride 2
Convolution	56×56	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	16×16	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
Convolution	28×28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	8×8	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
Convolution	14×14	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	4×4	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
Convolution	7×7	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 215 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	2×2	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 215 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
Classification	2048	7×7 global average pool	8192	flatten
	1000 fully-connected, softmax		7 fully-connected, softmax	

5.5 Conclusion

In this chapter, an architecture-model over convolutional neural networks, for classifying skin lesions was presented. The CNN-based architecture is based on preprocessing of dermatoscopic images followed by segmentation to extract the lesion area and subsequently processing of the corresponding segment of the image by a convolutional neural network for classification and labeling to a set of skin pathology types. In the evaluation new convolutional neural networks were trained via python packets of Keras and also TensorFlow (CUDA supported). In addition, we modified and re-trained well known publicly available CNN models for image classification which have been pre-trained using big data collections. The experimental results showed that all evaluated modified and re-trained pre-existing CNN models outperformed the CNN models trained explicitly from the dermatoscopic image dataset

Table 5.9 ResNet-101 [31] architectures and their modifications for skin lesion classification

Layers	Original ResNet-101		Modified ResNet-101	
	Output size	Layer's structure	Output size	Layer's structure
Convolution	112 × 112	7 × 7 conv, stride 2	32 × 32	7 × 7 conv, stride 2
Pooling	56 × 56	3 × 3 max pool, stride 2	16 × 16	3 × 3 max pool, stride 2
Convolution	56 × 56	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix} \times 3$	16 × 16	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix} \times 3$
Convolution	28 × 28	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix} \times 4$	8 × 8	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix} \times 4$
Convolution	14 × 14	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix} \times 23$	4 × 4	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix} \times 23$
Convolution	7 × 7	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 215 \\ 1\times1, 2048 \end{bmatrix} \times 3$	2 × 2	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 215 \\ 1\times1, 2048 \end{bmatrix} \times 3$
Classification	2048	7 × 7 global average pool	8192	flatten
	1000 fully-connected, soft-max		7 fully-connected, softmax	

used in the evaluation, with the best performing classification model being ResNet-50 with the accuracy equal to 93.89%. Analysis on skin lesion pathology type showed that melanoma and basal cell carcinoma were able to be correctly classified 79.13 and 82.88% of the times, respectively, while other types of pathologies like actinic keratosis and nevus were more easily correctly classified with accuracies 97.90 and 90.14%, respectively. The evaluation results show the potential of developing software tools for accurate computer-aided diagnosing the skin diseases, which can be used as objective supportive tools for dermatologists.

Table 5.10 ResNet-152 [31] architectures and their modifications for skin lesion classification

Layers	Original ResNet-152		Modified ResNet-152	
	Output size	Layer's structure	Output size	Layer's structure
Convolution	112×112	7×7 conv, stride 2	32×32	7×7 conv, stride 2
Pooling	56×56	3×3 max pool, stride 2	16×16	3×3 max pool, stride 2
Convolution	56×56	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	16×16	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
Convolution	28×28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$	8×8	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
Convolution	14×14	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$	4×4	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
Convolution	7×7	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 215 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	2×2	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 215 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
Classification	2048	7×7 global average pool	8192	flatten
	1000 fully-connected, soft-max		7 fully-connected, softmax	

Table 5.11 Accuracy regarding the dermatoscopic image classification (in percentages), as for different CNN model architectures

CNN Model		Accuracy (%)
New	CNN Model 1	72.68
New	CNN Model 2	76.38
New	CNN Model 3	**76.83**
New	CNN Model 4	75.29
Re-trained	VGG16 (modified)	85.19
Re-trained	MobileNet (modified)	91.32
Re-trained	DenseNet-121 (modified)	91.21
Re-trained	DenseNet-169 (modified)	91.30
Re-trained	DenseNet-201 (modified)	89.56
Re-trained	ResNet-50 (modified)	**93.89**
Re-trained	ResNet-101 (modified)	90.93
Re-trained	ResNet-152 (modified)	93.52

Table 5.12 Confusion matrix for skin lesion classification using ResNet-50 (modified) CNN model

	akiec	bcc	bkl	df	mel	nv	vasc
akiec	**97.90**	0.24	0.97	0.63	0.03	0.07	0.16
bcc	8.17	**82.88**	1.36	2.53	1.17	0.97	2.92
bkl	7.82	0.45	**87.33**	3.05	0.27	0.27	0.81
df	8.92	1.00	1.91	**86.17**	0.09	0.27	1.64
mel	5.22	8.70	2.61	2.61	**79.13**	0.00	1.74
nv	4.93	1.41	1.41	0.70	0.70	**90.14**	0.70
vasc	5.50	3.06	2.45	3.67	0.92	0.31	**84.10**

References

1. A.F. Jerant, J.T. Johnson, C. Demastes Sheridan, T.J. Caffrey, Early detection and treatment of skin cancer. Am. Fam. Physician. **62**(2), (2000)
2. R. Sumithra, M. Suhil, D. Guru, Segmentation and classification of skin lesions for disease diagnosis. Procedia. Comput. Sci. **45**, 76–85 (2015)
3. X. Li, J. Wu, H. Jiang, E. Z. Chen, X. Dong, R. Rong, Skin lesion classification via combining deep learning features and clinical criteria representations. bioRxiv, 382010, (2018)
4. M.A. Sheha, A. Sharwy, M.S. Mabrouk, Automated imaging system for pigmented skin lesion diagnosis. Int. J. Adv. Comput. Sci. Appl. **7**(10), (2016)
5. US Emerging Melanoma Therapeutics Market, "a090-52," Tech. Rep., 2001
6. S.A. Mahdiraji, Y. Baleghi, S.M. Sakhaei, Skin lesion images classification using new color pigmented boundary descriptors. In *2017 3rd International Conference on Pattern Recognition and Image Analysis* (*IPRIA*) (IEEE, 2017), pp. 102–107
7. F. Nachbar et al., The ABCD rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions. J. Am. Acad. Dermatol. **30**(4), 551–559 (1994)
8. M. Binder et al., Epiluminescence microscopy: a useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists. Arch. Dermatol. **131**(3), 286–291 (1995)
9. C.R. Moffatt, A.C. Green, D.C. Whiteman, Diagnostic accuracy in skin cancer clinics: the Australian experience. Int. J. Dermatol. **45**(6), 656–660 (2006)
10. J.F. Alcón et al., Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis. IEEE J. Sel. Top. Sign. Proces. **3**(1), 14–25 (2009)
11. W.-Y. Chang et al., Computer-aided diagnosis of skin lesions using conventional digital photography: a reliability and feasibility study. PLoS ONE **8**(11), e76212 (2013)
12. P. Rubegni et al., Automated diagnosis of pigmented skin lesions. Int. J. Cancer **101**(6), 576–580 (2002)
13. R. Refianti, A.B. Mutiara, R.P. Priyandini, Classification of melanoma skin cancer using convolutional neural network. IJACSA **10**(3), 409–417 (2019)
14. S. Prigent et al., Classification of skin hyper-pigmentation lesions with multi-spectral images. RR-8105, INRIA. 2012. ffhal-00745367
15. J. Kawahara, A. BenTaieb, G. Hamarneh, Deep features to classify skin lesions, in *2016 IEEE 13th International Symposium on Biomedical Imaging* (*ISBI*) (IEEE, 2016), pp. 1397–1400
16. G. Capdehourat, A. Corez, A. Bazzano, P. Musé, Pigmented skin lesions classification using dermatoscopic images, in *Iberoamerican Congress on Pattern Recognition* (Springer, 2009), pp. 537–544
17. C. Lehman, M. Halicek, G. AlRegib, Skin lesion classification: a transformation based approach to convolutional neural networks, in *Submitted to IEEE International Workshop on Machine Learning and Signal Processing* (*MLSP2017*), 2017

18. R. Mishra, O. Daescu, Deep learning for skin lesion segmentation, in *2017 IEEE International Conference on Bioinformatics and Biomedicine* (*BIBM*), (IEEE, 2017), pp. 1189–1194
19. M.H. Jafari et al., Skin lesion segmentation in clinical images using deep learning, in *2016 23rd International conference on pattern recognition* (*ICPR*), (IEEE, 2016), pp. 337–342
20. M. ur Rehman, S.H. Khan, S.D. Rizvi, Z. Abbas, A. Zafar, Classification of skin lesion by interference of segmentation and convolotion neural network, in *2018 2nd International Conference on Engineering Innovation* (*ICEI*), (IEEE, 2018), pp. 81–85
21. S. Kaymak, P. Esmaili, A. Serener, Deep learning for two-step classification of malignant pigmented skin lesions, in *2018 14th Symposium on Neural Networks and Applications* (*NEUREL*) (IEEE, 2018), pp. 1–6
22. T. Lee, V. Ng, R. Gallagher, A. Coldman, D. McLean, Dullrazor®: A software approach to hair removal from images. Comput. Biol. Med. **27**(6), 533–543 (1997)
23. J.A. Salido, C. Ruiz JR, Hair artifact removal and skin lesion segmentation of dermoscopy images. Asian. J. Phar. Clin. Res. **11**(3) (2018)
24. G.V. Tcheslavski, Morphological image processing: gray-scale morphology. ELEN **4304**, 5365 (2010)
25. T.F. Chan, L.A. Vese, Active contours without edges. IEEE Trans. Image Process. **10**(2), 266–277 (2001)
26. N.C. Codella, D. Gutman, M.E. Celebi, B. Helba, M.A. Marchetti, S.W. Dusza, A. Kalloo, K. Liopyris, N. Mishra, H. Kittler. Skin lesion analysis toward melanoma detection: a challenge at the 2017 international symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC). In:*2018 IEEE 15th International Symposium on Biomedical Imaging* (ISBI 2018) (IEEE, 2018), pp. 168–172
27. P. Tschandl, C. Rosendahl, H. Kittler, The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Scientific Data **5**, 180161 (2018)
28. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556 (2014)
29. A.G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, Mobilenets: efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:170404861 (2017)
30. G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708 (2017)
31. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778 (2016)
32. J. Deng, W. Dong, R. Socher, L-J. Li, K. Li, L. Fei-Fei, Imagenet: a large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (IEEE, 2009), pp. 248–255

Chapter 6
Combined Radiology and Pathology Based Classification of Tumor Types

N. Ravitha Rajalakshmi, B. Sangeetha, R. Vidhyapriya, and Nikhil Ramesh

Abstract Computer Aided Detection plays a crucial role in the early detection of deadly diseases such as cancer (or) tumor. Pathology and radiology images form the core of tumor diagnosis. Pathology images provide clinical information about the tissues whereas the radiology images can be used for locating the lesions. This work aims at proposing a classification model which categorizes the tumor as oligodendroglioma (benign tumors) (or) astrocytoma (Malignant tumors) using features of both the radiology and pathology images. Dataset from MICCAI Computational Precision Medicine Challenge is used for model building. The proposed model uses dedicated workflows for processing the pathology and radiology images. The feature descriptors of the images are obtained using pre-trained Inception v3 model. The resulting vectors are then used as input to the linear SVM (Support Vector Machine) classification model. The SVM model provided an accuracy of 75% on the blind folded test dataset provided in the competition.

6.1 Introduction

Brain tumor is a collection of mass or abnormal tissues in the brain. Early detection of tumor [10] can result in better treatment and speedy recovery of the patients. With the increased availability of digitized images, computer aided detection systems (CAD) are performing better at tasks like locating the tumor regions and also categorizing the grade of tumor. Amongst the various tumor grades, detection of lower grade tumors is still found to be a challenging task for CAD systems as the tumor cell structure is similar to the normal cells. Radiology Images are used at primary screening level wherein doctors look out for any suspicious regions (or) accumulation of white

N. Ravitha Rajalakshmi (✉) · B. Sangeetha · R. Vidhyapriya
Department of Information Technology, PSG College of Technology, Coimbatore, Tamil Nadu, India
e-mail: nrr.it@psgtech.ac.in

N. Ramesh
PSG College of Technology, Coimbatore, Tamil Nadu, India

U. Kose and J. Alzubi (eds.), *Deep Learning for Cancer Diagnosis*,
Studies in Computational Intelligence 908,
https://doi.org/10.1007/978-981-15-6321-8_6

mass. Then based on the report of radiologist, pathology of the suspicious regions is studied to identify the grade and malignancy of tumor. Further in the prognosis stage, information from both pathology and radiology images are correlated to decide on the further course of action. Sometimes, the pathology may reveal the abnormal region to be normal. In particular, sensitivity of models in identifying low-grade tumors at initial screening level is so less. Hence, the challenge was an initiative to integrate various levels of information from both the radiology and pathology images.

6.2 Methodology

In this work, an independent image processing pipeline is used for processing pathology and radiology images. The images are gleaned and necessary structures suitable for the task are extracted. It is then provided as input to pretrained Inception v3 for extraction of distinct high-level complex feature descriptors. The descriptors of pathology and radiology images are then concatenated. In linear SVM hypothesis space, the model for segregating the descriptors of the two classes are then identified. Figure 6.1 illustrates the overall workflow of the proposed system for tumor classification.

6.2.1 Dataset Description

The dataset for the work is obtained from 2018 Computational Precision Medicine Challenge Organized by MICCAI. The dataset contains both the Radiology and Pathology images obtained from the same patients. Each case corresponds to a single patient. For each case, the following set of images are provided.

- One whole slide tissue image.
- A set of Magnetic Resonance (MR) images which comprises of T1 (pre and post contrast), FLAIR and T2.

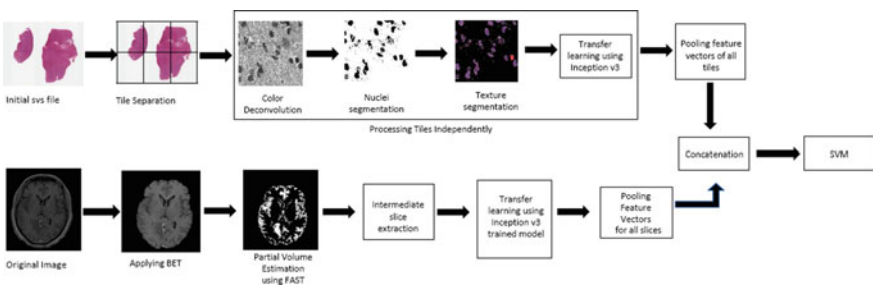


Fig. 6.1 Workflow for tumor classification

Table 6.1 Distribution of cases across two different classes in training set

Class	Number of cases
Oligodendroglioma	16
Astrocytoma	16

The Radiology data are provided in **NIFTI format**. The pathology images are **Hematoxylin and Eosin** (H&E) stained whole slide tissue images in **SVS format**. SVS files are used by a number of medical/microscope scanners such as Aperio, scanscope (AxioVision), and others. It actually contains the images acquired at different resolution. Table 6.1 shows the distribution of cases corresponding to each class in Training set.

6.2.2 Processing Pathology Images

Whole slide images will contain diagnostic images of varying resolution which are arranged as pyramidal structure starting from an image with lowest resolution to highest resolution. These microscopic images contain large number of pixels in the order of 100,000 * 100,000 which is difficult to analyze because of the constrained memory. Hence, it is usually divided into smaller tiled images of specific size [3, 14]. In this work, only the baseline image (i.e.) image with highest resolution is used for further processing. Image is divided into patches of size 256 * 256. Only the patches that contained tissue occupying 90% of the patch are considered.

Usually, histopathology images are stained using dyes and then it is captured. Due to the differences in the application of dyes, preparation of the tissue specimen and the time spent in the staining process before capturing the image, color variations are introduced in the resulting images [15]. To address the problem, color normalization of histopathological images is generally carried out before conducting further analysis. In this work, histogram based Color Normalization [8] known as stretch is applied to balance the number of white pixels. The resultant image after applying stretch to a patch image (Fig. 6.2(a)) is shown in Fig. 6.2(b).

6.2.2.1 Color Deconvolution

Nuclei of benign tumor's (oligodendroglioma) appear round shaped and exhibit smooth texture whereas the nuclei of malignant tumor's (astrocytoma) appear elongated with irregular texture [7, 14]. Hence, to differentiate between the different types of tumor, morphology of cell nuclei is considered as an important attribute. To facilitate segmentation of nuclei, Color Deconvolution is performed. This separates the stain components from the image.

H&E images are stained using two components Hematoxylin and eosin respectively. Hematoxylin stains the nuclei blue and Eosin stains the cellular structures

pink. Separation of the components will enable the separation of nuclei. Stain separation is carried out in the optical density space. Beer-lambert law defines the optical density space as

$$\mathrm{OD} = -\log_{10}(I)$$

where I denote input RGB image wherein pixels are normalized to a value between [0, 1]. The Saturation of stains can be extracted from the optical density image using the simple color deconvolution scheme given below.

$$\mathrm{OD} = v \cdot S;\ S = v^{-1}(\mathrm{OD})$$

where v and S denote stain matrix and the saturation of stains onto pixels respectively. In the work, we have utilized a standard v matrix for H&E as provided in [9].

$$v = \begin{bmatrix} 0.65 & 0.70 & 0.29 \\ 0.07 & 0.99 & 0.11 \\ 0.27 & 0.57 & 0.78 \end{bmatrix}$$

Rows in the stain matrix corresponds to haematoxylin, eosin and DAB space respectively. Multiplying the inverse of v with optical density image results in the saturation image whose channels correspond to the concentration of each of the stains. Only the channel corresponding to haematoxylin is used. The resulting channel is subjected to segmentation to recover darker objects (nuclei) from the rest of the image as shown in Fig. 6.2(c).

6.2.2.2 Nuclei Segmentation

Many techniques are available to segment nuclei from histopathology image [1, 5]. In this work, Local Maximum Clustering is used for segmentation of the nuclei. It utilizes the response of LoG filter which smoothens the image using Gaussian filter and identifies the pixel corresponding to abrupt intensity changes using second order derivative Laplacian operator. The LoG records maximum response for the pixels corresponding to the centre of the nuclei. Clustering algorithm then uses pixels with maximum value in a defined neighbourhood of the filter response as seed points and identifies regions corresponding to the nuclei through region growing strategy. The pixels are then labelled using the labels corresponding to its nearest local maxima whereby the size of the local neighbourhood is set to a value of 10. The nuclei detected for sample image is shown in Fig. 6.2(d). Using the mask generated by the algorithm, the pixel intensity of nuclear regions is preserved in the original image characterizing the nuclei texture [7]. Figure 6.2(e) shows the image whereby texture of the segmented nuclei is recovered.

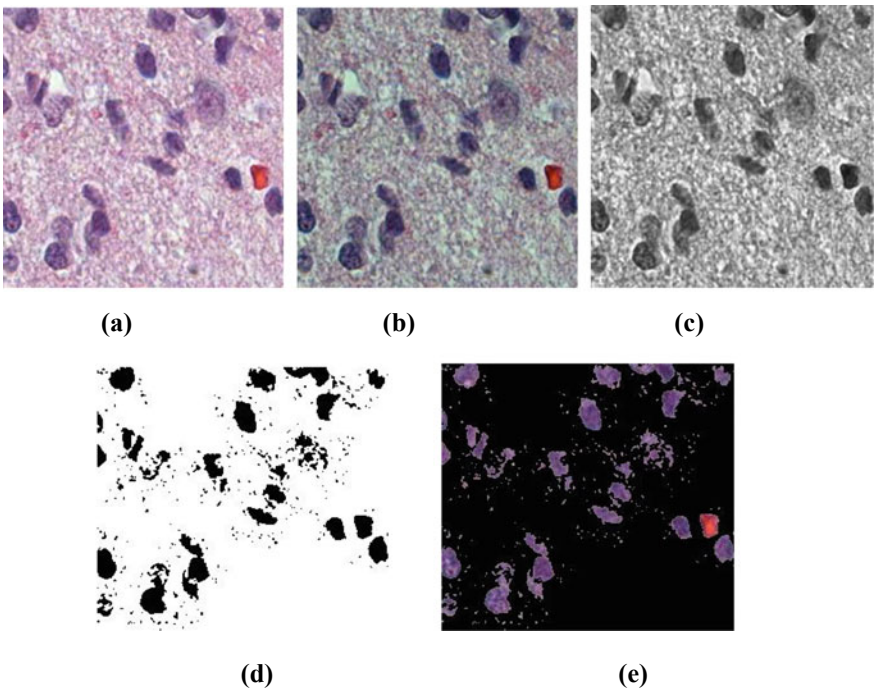


Fig. 6.2 **a** Sample patch extracted from Hematoxylin and Eosin stained whole slide image (WSI). **b** Resultant Image after applying stretch algorithm. **c** Hematoxylin constituent in the image extracted using color deconvolution. **d** Detected nucleus in the image using local maximum clustering. **e** Retaining the intensity of the nuclei from the original image

6.2.2.3 Feature Extraction

Deep Convolutional Neural Networks are considered as rich feature extractors characterizing both shape and texture of objects in an image. In H&E stained WSI images, processing is carried out using patches. Hence, multiple feature descriptors will be extracted from a single WSI each corresponding to single patch of WSI slide [3]. Pooling is then used as an aggregator to find the global descriptor for the entire WSI slide. In this work, inception v3 [11] model pretrained on ImageNet dataset is used as feature extractor. Inception v3 was selected as it has the ability to extract features of varying scale by applying convolutions in parallel to an input feature map and concatenating its output responses. The resulting feature maps has rich information about the input data. The patches were fed as input to Inception v3 and the vector generated after the average pooling layer is used as the feature vectors. Figure 6.3 illustrates the computation of global descriptor. It is computed as the maximum value recorded for each feature across the feature descriptors of individual patches.

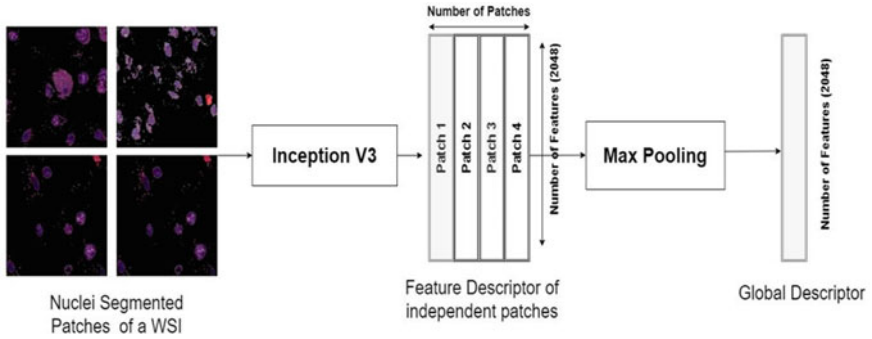


Fig. 6.3 Global feature descriptor extraction for whole slide image (WSI) using pretrained deep convolutional network

6.2.3 Processing Radiology Images

The two major non-invasive imaging modalities involved in the diagnosis of tumor in brain include: Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) [12]. The most common sequences of MRI images of high spatial resolution with good contrast in the tissues are provided in the dataset. Table 6.2 summarizes the various MRI sequence images provided, the intensity of various tissue types in these images and their TE and TR times. The MRI images are generally captured by introducing a magnetic field using an external varying levels of RF energy. Repetition Time (TR) denotes the time lapse between the successive pulses of RF signals to the same slice and Time to Echo (TE) indicates the time difference between the delivery of the RF signal and the echo signal. They capture images from a different anatomical plane and represents the prominent tissue types: Gray Matter (GM), White Matter (WM), Cerebrospinal Fluid (CSF) using varying intensities.

Table 6.2 Characteristics of MRI Images with varying weights

Weighting	Anatomical plane	Tissue types			Repetition time (TR)	Time to echo (TE)
		Gray matter (GM)	White matter (WM)	Cerebra spinal fluid (CSF)		
T1-weighted	Sagittal, axial, coronal	Dark	Light	Dark	500	14
T2-Weighted	Axial	Light gray	Dark gray	Bright	4000	90
T1-weighted Contrast	Sagittal	Gray	Light	Dark	500	14
Flair	Axial	Light gray	Dark gray	Bright	9000	114

Anatomical plane is primarily classified into three: axial plane; slice perpendicular to the *z*-axis, sagittal or median plane; the plane that divides the brain into left and right parts is known as sagittal or median plane and Coronal or Frontal plane; plane that divides the brain into posterior and anterior parts. Not all the four images were available for every case. In this work T1C and T2 weighted images are used as these images were available for every case and T2 weighted image enables clear separation of tumor cells.

The T2-weighted image are co-registered with T1 weighted contrast image and are pre-processed to improve the visibility of tumor. Preprocessing includes bias field correction, skull striping and partial volume estimation which can be done using popular open source tools FSL, SPM, FreeSurfer etc [4]. In this work, FSL Library [6] is used.

6.2.3.1 Image Registration

Aligning the multiple volume images into the same geometric space is called Co-registration. T2 weighted image is co-registered using T1C as reference image as suggested by physicians. FLIRT command of FSL is used for this purpose.

6.2.3.2 Bias Field Correction

It is used to ensure piecewise constant property across all the images. Piecewise constant property states that the intensities of different tissue classes should be the same across the entire image. In other words, the intensity of gray matter appearing in any part of the image should be the same. Moreover, the intensity of the gray matter should not be larger than the intensity of the white matter tissue class. These bias in the intensities are otherwise called as intensity inhomogeneity or shading which is caused during the image acquisition process. The inhomogeneity in the images can be handled by retrospective methods or prospective methods. The retrospective approach focuses on removing the shading after the image is captured by applying appropriate mask. In this work, the bias correction is done with the help of FMRIB Software Library (FSL). The FSL package uses segmentation-based methods for performing dual task of segmenting different tissue classes from the image and also handling inhomogeneity. The bias corrected T2 weighted image is shown in Fig. 6.4(a).

6.2.3.3 Skull Stripping

The MRI images acquired may contain extra-cranial or non-brain tissues like skin, fat, muscles, scalp, eyes etc. The Skull striping [2] is a pre-processing technique

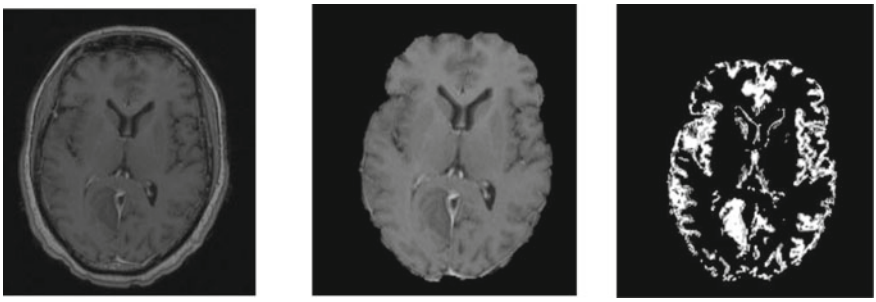


Fig. 6.4 Processing of radiology images: slice of T2-weighted MRI image. **a** Bias corrected image. **b** Skull Stripping using BET. **c** Partial volume map of the fourth class obtained using FAST algorithm

applied to segregate the non-brain tissues from the brain tissues in an MRI image. The non-brain tissues have to be stripped off because of its overlapping intensities with other tissue classes. Various approaches available for skull scripting include morphology-based approaches, intensity-based methods, deformable-surface based method and hybrid-based method.

In this work, Brain Extraction Tool (BET) of the FSL library which employs the deformable-surface based approach for skull stripping. Deformable-surface method uses an active contour region for identifying the boundary of the objects in a given image. Active contour region starts with a closed curve and either expands or shrinks based on the gradient value. The contour defines the region of interest in the images (in this case the skull boundary of the image) as a boundary of points. These boundary points are interpolated using the linear, polynomial or spline function to define the contour of the image in identifying the shape. The skull stripped image of Fig. 6.4(a) is shown in Fig. 6.4(b).

6.2.3.4 Partial Volume Estimation

FAST tool of FSL is used for Partial Volume Estimation. It extracts tissues of varying types specified in Table 6.2 using Hidden Markov Random Field with Expectation Maximization. It generates partial volume maps for four classes gray matter, white matter, Cerebra Spinal Fluid and other tissues (tumor). The partial volume map of tumor corresponding to T2 weighted image (Fig. 6.4(c)) is used for feature extraction.

6.2.3.5 Feature Extraction

As in the case of pathology images, global descriptor for entire voxel is obtained. It utilizes 10 slices in the middle region of the brain which is provided as input for

Inception v3. The feature descriptors for every slice is extracted and then max pooled to generate a global descriptor.

6.2.4 Support Vector Machine

Support Vector Machine (Cortes C, 1995) is a statistical machine learning algorithm which finds an optimal hyperplane to segregate the data points belonging to two different classes. The algorithm formulates the hyperplane detection as a constrained optimization problem whereby It selects the hyperplane that maximizes the margin (i.e.) the distance from the hyperplane to the nearest data points of both the classes. In this work, feature descriptors of radiology and pathology images are used to construct a linear SVM to find the optimal set of parameters $\{W, b\}$ for the hyperplane in the input space which can provide better generalization accuracy.

6.3 Results

Since the number of samples in the input dataset is 32, reliable accuracy estimate can be obtained either by using the technique of cross validation or bootstrapping. In this work, model is evaluated using cross validation. The dataset is divided into three folds and model is trained on two of three folds and the remaining one-fold is used for evaluation [13]. Receiver Operating Characteristic curve which measures the goodness of the model is shown in Fig. 6.5. ROC is a plot of false positive rate vs true positive rate at varying score levels. The graph indicate that Radiology data is not highly discriminative as it records only an average AUC of 0.51 (Fig. 6.5(a)) whereas pathology data provides an AUC of 0.62 (Fig. 6.5(b)). Using Combined model (i.e.) concatenating the vectors of Radiology and Pathology offered a better average AUC of 0.72 as shown in Fig. 6.5(c). Non-linear SVM when tried as an option resulted in poor accuracy even with hyperparameter tuning with Grid Search CV. The top performing model for combined pathology with AUC of 0.95 is used for the blind folded test set provided in the competition and it recorded an accuracy of 75% as per the competition scoreboard. Table 6.3 summarizes the performance of the proposed model on Training and Test dataset. The code for the work can be found at https://github.com/Ravitha/Combined-Pathology-and-Radiology.

6.4 Conclusions

A unified framework for combining radiology and pathology images is explored in the work. To facilitate extraction of distinct features, region of interest in both pathology images and radiology images are segregated and then processed using pretrained

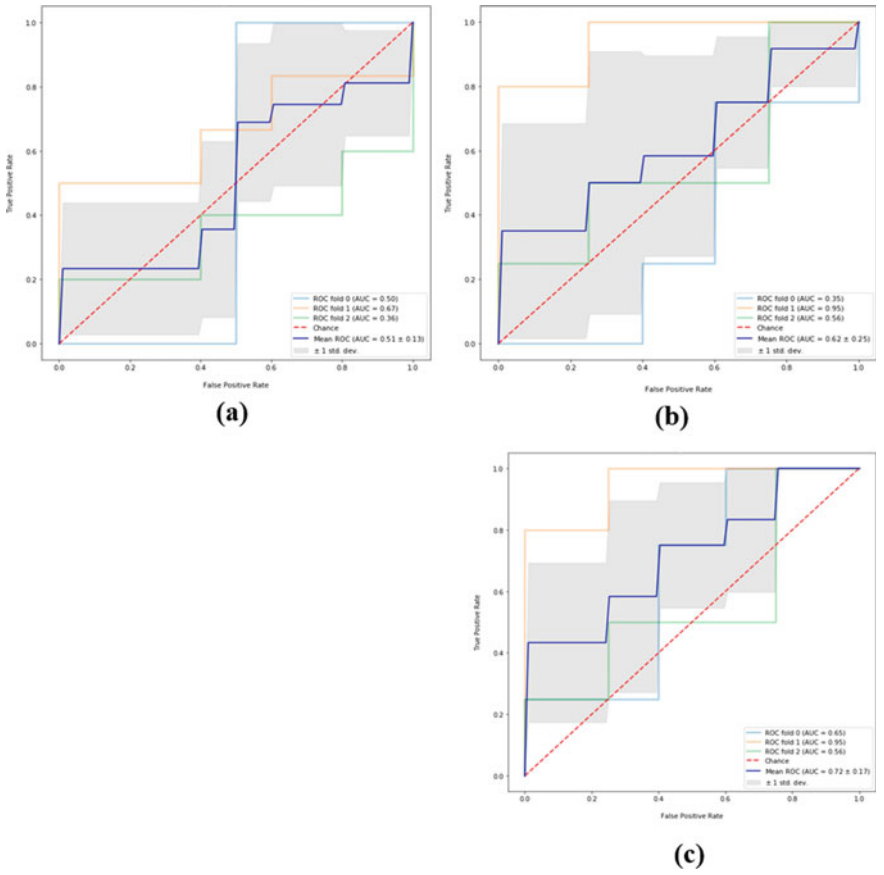


Fig. 6.5 ROC curve of the SVM classifier trained using the cross validation, false positive rate is plotted on the *x*-axis, and the true positive rate is plotted on the *y*-axis **a** Radiology data. **b** Pathology data. **c** Combining radiology and pathology data

Table 6.3 Accuracy of the model evaluated for train and test set using 3-fold cross validation

Dataset	Input data	AUC	Accuracy (%)
Train (32)	Radiology	0.51	50
	Pathology	0.62	61
	Combined	0.72	70
Test (10)	**Combined**	–	**75**

deep convolutional network. The resulting descriptors are found to be effective in segregating the classes. The model can be enhanced by eliminating correlated features using dimensionality reduction techniques or by utilizing a neural network as meta learner to combine the decisions of individual SVM classifiers trained on the pathology and radiology images separately.

References

1. Y. Al-Kofahi, W. Lassoued, W. Lee, B. Roysam, Improved automatic detection and segmentation of cell nuclei in histopathology images. IEEE Trans. Biomed. Eng. **57**(4), 841–852 (2010). https://doi.org/10.1109/TBME.2009.2035102
2. A.G.R. Balan, A.J.M. Traina, M.X. Ribeiro, P.M.A. Marques, C. Traina, Smart histogram analysis applied to the skull-stripping problem in T1-weighted MRI. Comput. Biol. Med. **42**(5), 509–522 (2012). https://doi.org/10.1016/j.compbiomed.2012.01.004
3. P. Courtiol, E.W. Tramel, M. Sanselme, G. Wainrib, *Classification and Disease Localization in Histopathology Using Only Global Labels: A Weakly-Supervised Approach* (2018), pp. 1–13. http://arxiv.org/abs/1802.02212
4. Z. Cui, C. Zhao, G. Gong, Parallel workflow tools to facilitate human brain MRI post-processing. Frontiers Neurosci. **9**(APR), 1–7 (2015). https://doi.org/10.3389/fnins.2015.00171
5. D. Forsberg, N. Monsef, Evaluating Cell Nuclei Segmentation for Use on Whole-Slide Images in Lung Cytology, in *2014 22nd International Conference on Pattern Recognition* (2014), pp. 3380–3385. https://doi.org/10.1109/ICPR.2014.582
6. M. Jenkinson, C.F. Beckmann, T.E.J. Behrens, M.W. Woolrich, S.M. Smith, Review FSL. NeuroImage **62**, 782–790 (2012). https://doi.org/10.1016/j.neuroimage.2011.09.015
7. A.D. Lee, J.K. Cooper et al., An integrative approach for in silico glioma research. IEEE Trans. Biomed. Eng. **23**(1), 1–7 (2011). https://doi.org/10.1161/CIRCULATIONAHA.110.956839
8. D. Nikitenko, M. Wirth, K. Trudel, Applicability of white-balancing algorithms to restoring faded colour slides: An empirical evaluation. J. Multimedia **3**(5), 9–18 (2008). https://doi.org/10.4304/jmm.3.5.9-18
9. A.C. Ruifrok, D.A. Johnston, Quantification of histochemical staining by color deconvolution. Anal. Quant. Cytol. Histol. **23**(4), 291–299 (2001)
10. J.D. Schiffman, P.G. Fisher, P. Gibbs, Early detection of cancer: past, present, and future. Am. Soc. Clin. Oncol. Educ. Book, **35**, 57–65 (2015). https://doi.org/10.14694/edbook_am.2015.35.57
11. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the Inception Architecture for Computer Vision, in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016-Decem* (2016) pp. 2818–2826. https://doi.org/10.1109/CVPR.2016.308
12. K. Usman, K. Rajpoot, Brain tumor classification from multi-modality MRI using wavelets and machine learning. Pattern Anal. Appl. **20**(3), 871–881 (2017). https://doi.org/10.1007/s10044-017-0597-8
13. G. Varoquaux, L. Buitinck, G. Louppe, O. Grisel, F. Pedregosa, A. Mueller, Scikit-learn. GetMobile Mobile Comput. Commu. **19**(1), 29–33 (2015). https://doi.org/10.1145/2786984.2786995
14. F. Wang, T.W. Oh, C. Vergara-Niedermayr, T. Kurc, J. Saltz, Managing and Querying Whole Slide Images, in *Medical Imaging 2012: Advanced PACS-Based Imaging Informatics and Therapeutic Applications*, vol. 8319 (2012), p. 83190J. https://doi.org/10.1117/12.912388
15. Y. Zheng, Z. Jiang, H. Zhang, F. Xie, J. Shi, C. Xue, Adaptive color deconvolution for histological WSI normalization. Comput. Methods Programs Biomed. **170**, 107–120 (2019). https://doi.org/10.1016/j.cmpb.2019.01.008

Chapter 7
Improved Deep Learning Techniques for Better Cancer Diagnosis

K. R. Sekar, R. Parameshwaran, Rizwan Patan, R. Manikandan, and Ambeshwar Kumar

Abstract Over the past several decades, Computer-Aided Diagnosis (CAD) for diagnosis of medical images has prospered due to the advancements in the digital world, advancements in software, hardware and precise and fine-tune images acquired from sensors. With the advancement in the field of medical and applications of Artificial Intelligence scaling to the height of improvement, modern state-of-the-art applications of Deep Learning for better cancer diagnosis have been incepted in recent years. CAD and computerized algorithms and solutions in diagnosing cancer obtained from different modalities, i.e., MRI, CT scans, OCT and so on plays an immense impact on disease diagnosis. Learning model based on transfer mechanisms that stored knowledge for one aspect and using it for another aspect with Deep Convolutional Neural Network paved the way for automatic diagnosis. Recently, improved deep learning algorithm has resulted in great success resulting in robust image characteristics, involving higher dimensions. Analysis of bi-cubic interpolation preprocessing technique paves way for robust obtaining of a region of interest. For an inflexible object with a higher amount of dissimilarity, a comprehensive form for detecting the region of interest and determination of actual positioning may not be robust. Robust perception and localization schemes are analyzed. By integrating Deep Learning with Neighborhood Position Search unseen cases are said

K. R. Sekar · R. Manikandan (✉) · A. Kumar
School of Computing, SASTRA Deemed University, Thanjavur, India
e-mail: srmanimt75@gmail.com

K. R. Sekar
e-mail: sekar1971kr@gmail.com

A. Kumar
e-mail: ambeshwar.kumar@gmail.com

R. Parameshwaran
Department of ECE, National Institute of Technology, Tiruchirappalli, India
e-mail: paramu32@gmail.com

R. Patan
Department of Computer Science and Engineering, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, AndhraPradesh, India
e-mail: prizwan5@gmail.com

U. Kose and J. Alzubi (eds.), *Deep Learning for Cancer Diagnosis*,
Studies in Computational Intelligence 908,
https://doi.org/10.1007/978-981-15-6321-8_7

to be identified and segmented accordingly via Maximum Likelihood decision rule, forming robust segmentation. The favorable result of an a better cancer diagnosis is indeed contingent on the cancer diagnosis however, an anticipating prediction should consider certain factors more than a straight forward diagnostic decision. Besides the application of different medical data analyses and image processing techniques used in the study of cancer diagnosis deeper insights of the relevant solutions in the light of higher collections of deep learning techniques are found to be vital. Hence, certain factors to be analyzed are the forecasting of risk involved, forecasting of cancer frequency and the forecasting of cancer survival. These factors are analyzed according to the diagnosis criterion, sensitivity, specificity, and accuracy.

Keywords Computer-Aided diagnosis · Artificial intelligence · Deep convolutional neural network · Supervised deep learning · Recurrent neural network

7.1 Computer Analysis Interpretation for Medical Visual Representation

Computer Analysis Interpretation refers to the system that aids the doctors while interpreting images pertaining to medical data. With the assistance of certain imaging process like, high energy electromagnetic radiation, and diagnosing ultrasound images, on the other hand, requires voluminous amount of information to be handled from the side of doctor involving comprehensive analysis in a minimum interval. Digital images processed by computer aided systems are potentially utilized to identify the presence or absence of disease at an early stage.

One of the prominent causes of extinction of life globally is cancer. Both eminent research personalities and physicians are facing threats of fighting cancer. Based on the research conducted by the American cancer society in 2019, the occurrence of death rate due to skin cancer is 96,480, death rate due to lung cancer is 142,670, mortality due to breast cancer is 42,260, mortality rate due to prostate cancer is 31,620 and death rate due to brain cancer is 17,760 [1].

Therefore, the topmost priority for saving lives lies in the cancer discovery at the preliminary stage. Different visual examination and manual procedures are carried out to discover and diagnosis cancer at the preliminary stage. However, this human intervention of medical images necessitates a higher amount of time consumption and results in different types of faults. For this reason, diagnosis of disease using computer systems were brought to helps the medical experts and specialists to enhance the effectiveness of medical image interpretation.

7.1.1 Medical Imaging Types and Modalities

In recent years, disease diagnosis and patient treatment are said to be performed with minimal side effects. Medical imaging involves the materials and methods of producing representations of interior visual images for clinical medicine analysis and interposition accordingly. In other words, it is defined as the visual representation of the function of certain organs or tissues. With the aid of medical imaging helps in disclosing the concealed by the skin and also utilized in the diagnosing and treatment of the corresponding disease. Using medical imaging, abnormalities arising are compared with the normal anatomy and physiology.

With the visual representation, this objective is said to be attained where the functioning of organs inside the body is said to be perceived without the requirement for surgery or supplementary measures. Medical imaging refers to the utilization of imaging modalities to acquire pictures of the human body, track ongoing issues and finally assist in treatment. Hence, medical imaging is said to be of use for both disease diagnosis and healing purposes.

Different types of medical imaging techniques are said to exist, each possessing its own advantages and disadvantages. The following section gives an introduction to the most common imaging techniques, their working, advantages, and limitations that will pave the way for accurate imaging. Some of the common medical imaging types for better cancer diagnosis are shown in Fig. 7.1.

As depicted in the above figure, different medical imaging types along with their working, advantages, and disadvantages are given below.

7.1.1.1 Ultrasound

One of the safest forms of medical imaging is ultrasound. This is because of the reason that there arises no harmful hazard while utilizing the ultrasound due to the sound waves and Doppler used instead of the ionizing radiation. As in the case of sound waves, a conducting gel is utilized as a probe, with the aid of those waves, an

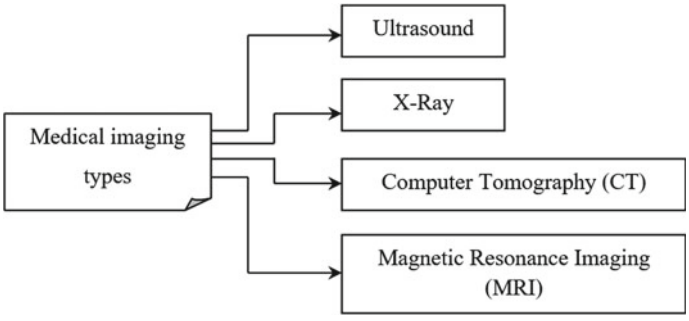


Fig. 7.1 Block diagram of medical imaging types

image is crated for the purpose of diagnosis. On the other hand, in case of Doppler, sound waves are utilized as diagnosis tool where the blood flow is performed via arteries and veins.

Working and Uses

- The device includes a probe that continuously emits high-frequency sound waves.
- The device correspondingly bounces off of several body fragments, resulting in echoes.
- Contrarily, when these bounce back to the probe, it can also detect them.
- This, in turn, results in the live image on another scanner, lasting from 15 to 45 min, performed either externally internally or endoscopic ally.

Advantages

- Minimal risk involved
- Cost-effective medical imaging.
- The safest form of medical imaging.

Limitations

- Found to be highly allergic.
- Causes sore throat or bloating.
- Extreme cases result in internal bleeding.

7.1.1.2 X-Ray Imaging

The second form of imaging used for cancer diagnosis is the most customarily utilized imaging types and this is called the X-Ray imaging. Here, an X-ray machine is utilized to obtain anatomy images most common bone cancer diagnosis. The working of X-ray is based on the wavelength and frequency that are found to be unable to be seen with the naked human eye but pierce into the skin to produce an impression of what's going on underneath. This type of imaging is specifically utilized for cancer diagnosis via mammography.

Working and Uses

- One of the types of electromagnetic radiation is x-rays, found to be invisible to humans, easily penetrate through the body.
- The patient possessing a detector on the other end will see the absorbent rate, with which an image is said to be generated.

- In certain cases, even a contrast agent is said to be injected to the patient with the purpose of seeing the soft tissues more easily on the image.
- As the process is said to be very fast, the entire procedure is said to take place within a few minutes.

Advantages

- Low cost incurred
- Quick and relatively ease of patient.

Limitations

- Risk due to radiation injection.
- Radiation-induced cancer.

7.1.1.3 CT Scan

CT scan also referred to as Computer Tomography or Computed Axial Tomography (CAT) utilizes X-rays to generate a cross-sectional body image. It consists of a large circular opening provided with a source, a detector that rotates automatically around the patient resulting in a snapshot. These snapshots are then merged into one or any number of images which are then further used for cancer diagnosis.

Working and Uses

- During the performing of a CT scan, a table is used where the patient lies down.
- The table here moves via a doughnut-like ring called a gantry.
- The gantry on the other end possesses an X-ray tube.
- Rotating of this X-ray tube is conducted on all sides of the patient. At this juncture, narrow beams of X-rays are shoot via the body.
- Finally, the digital detectors pick up the X-rays that are found to be directly opposite the source.

Advantages

- Find tumors and detail picturesque after a conventional scan.
- Monitoring is also said to be an important advantage of CT scan allow progress to be kept of any developing conditions, i.e., cancer.
- Greater clarity than conventional X-rays.
- Prevents the requirement for exploratory surgery.

Limitations

- Risk of cancer.
- Harm to the unborn child.
- Possibility of an allergic reaction.

7.1.1.4 MRI (Magnetic Resonance Imaging)

A vibrant magnetic field is utilized by the so called MRI that in turn generates images of body portions usually invisible through x radiation or cross sectional images. Even structures of tissue found to be very tender in nature are also measured via MRI, or identify tumors within the body. With this, even a joint or ligament are said to be observed accurately rather than just outside the view. Customarily, MRI is utilized in diagnosing strokes, tumors and brain functionalities.

Working and Uses

- Powerful magnets are employed in MRIs.
- The magnets seen powerful in nature tend to push the protons in the body and accordingly align with that field.
- On the other hand, upon turning off of the radio frequency, the energy released at this juncture are seen to be detected by the MRI sensors during realignment of a proton with the magnetic field.

Advantages

- No side effects.
- MRI is painless and found to be highly safe.
- They do not possess any radiation exposure as found in the case of X-ray.

Limitations

- With the involvement of a strong magnet, can result in a hazard.
- Loud noise requires ear protection.

7.2 Review of Cancer Diagnosis via Novel Deep Learning

Feature extraction is considered to be the preliminary steps in machine learning. Several feature extraction methods and mechanisms for discrete cancer types are

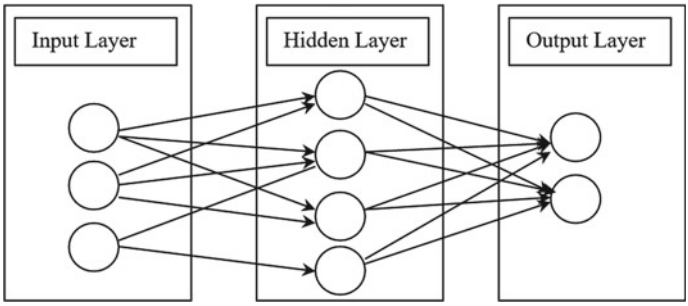


Fig. 7.2 A depiction of ANN

explored in recent years. Despite the effectiveness and efficiency explored, these feature extraction materials and methods possess certain deficiencies. To control and enhance the performance, several mechanisms in learning have been presented and investigated in the recent past.

High-level feature depiction is being developed by learning the deep features involved in it [2]. For example, cancer detection with favorable performance has been detected using convolutional neural networks. Due to nonlinear nature of processing involved in neurons, neural networks or improved deep learning possess the potentiality of performing complicated computation. With the higher amount of prediction being said to be attained using an artificial neural network, it is said to be higher in use as far as the medical images are concerned.

To be more specific, in an artificial neural network or improved deep learning, test images considered for experimentation are provided as input to the neurons for training. Followed by which the backpropagation algorithm is utilized to train neurons. Then, the resultant output images are said to be compared to those of the input images, where the error signal is said to be obtained if a proper match is not said to occur. Figure 7.2 given below shows the sample depiction of ANN.

As illustrated in the above figure, three layers are present in the ANN structure. They are:

- Input layer.
- Hidden layer.
- Output layer.

To start with, in the input layer, the elementary component of computation in ANN is the neuron or the node present in the input layer. It also received input from certain other nodes also and with the aid of the hidden layer where the actual processing is performed measures an output.

Each input possesses the corresponding weight factor that is said to be allocated based on the relative significance to other input nodes or neurons. Corresponding, the node or the neuron is said to be applied with a certain task to the corresponding weighted input sum, resulting in the disease diagnosis in case of medical imaging. In other words, the input layer provides information from the outside to the network

and hence, in this layer no computation is said to be performed in any of the input nodes, where only passing of the information to the hidden nodes is said to exist.

In the hidden layer where the hidden nodes are present does not communicate with each other due to the absence of direct communication hence it is referred to as the hidden layer. However, arithmetical calculations and information transfer are said to be performed between input and output layer. Therefore, the hidden layer comprises of group of hidden nodes.

Finally, the actual computation and transformation of information are said to be conducted in the output layer where the output nodes are present and here only the contact with the output world is said to take place here. In the activation function, the output from the neuron is measured. Here, the non-linearity is said to be established into the neuron output. This is because of the reason that most medical imaging is said to be non-linear in nature and neurons have to be learned for these corresponding non-linear representations. Every non-linear representation obtains a single number and conducts a definite mathematical operation. Some of the activation functions included is:

- Sigmoid.
- Tanh.
- ReLU (Rectified Linear Unit).

To minimize the error, different types of weights are prevalent and according to the type of cancer to be diagnosed, weights are adjusted accordingly. This processing is said to proceed until the error is said to be nullified. In the design of a neural network, a layered structure is present with different numbers of nodes found to be connected with each other and an activation function.

Some of the non-linear transformation model used for better cancer diagnosis is threshold function, tangent hyperbolic function sigmoid function and so on. Depending upon the type of cancer diagnosis, the usage of the activation function also differs. Input patterns (i.e. image structure) are provided as input to the network in input layer, which is then connected to the hidden layer and this hidden layer is in turn association with the output layer.

7.2.1 Deep Convolutional Transfer Learning-Based Neural Network

In recent years, convolutional neural networks (CNNs) have been proven to efficiently differentiate between benign and malignant lesions. Compared with conventional methods, CNNs minimizes the process involved in the design of the image feature extraction. To be more specific, they provide image or feature characteristics directly into the network that in turn identifies the selective features in an automatic manner.

The convolutional neural network (CNN) being specialized feed-forward neural network processes multi-dimensional data. The architecture includes layers to be

convoluted, layers to be pooled and finally, forming a completed set of layers in connected form. Figure 7.3 illustrates the architecture of CNN used in diagnosis of cancer. When compared to layers associated in a fully connected form, a single element of layer to be convoluted is specifically associated with a trivial open field of its input.

On the other hand, the filter bank is associated with the connection weight and the functions to be convolved are used to slide the filter across the input. This is turn forms activations at the corresponding receptive field that combines simultaneously resulting in a feature map.

A CNN is a type of Deep Learning algorithm that acquires an input image, allocates weights and bases to several images features so that efficient differentiation is said to take place between them. As far as pre-processing is concerned, the convolution network is said to be performed in a negligible pattern upon comparison with the other classification types. While in elementary models, filters are said to be performed manually with enough form of training, CNN has the potentiality to learn these filters.

With the above CNN architecture, the classification of tumors of a cancer diagnosis is performed. The convolution operation as depicted in the above figure comprises passing a kernel over an input volume representing an image. During this pass, the image value that matches kernel size is performed with a matrix multiplication to provide the value of single-cell on the output volume. The kernel here then slides in the correct direction. This is performed in an iterative manner until the iteration reaches the final position. If the sliding is performed by 1, it is called stride of one with the total positions being 4.

On the other hand, if the stride is increased by 2, it is called stride of two with the total positions being 2. The next element included in deep CNN is padding. Padding is a technique that includes zeros to the image margin to enhance the size. The padding required to achieve a similar volume on both the convolution side is called the Same Padding. Figure 7.4 show the tumor classification with the resultant values to be either of malignant or not. The dashed images represent misclassification of tumor type via CNN.

As far as transfer learning-based neural networks with deep convolution is concerned, it transfers the knowledge according to the input pattern being considered

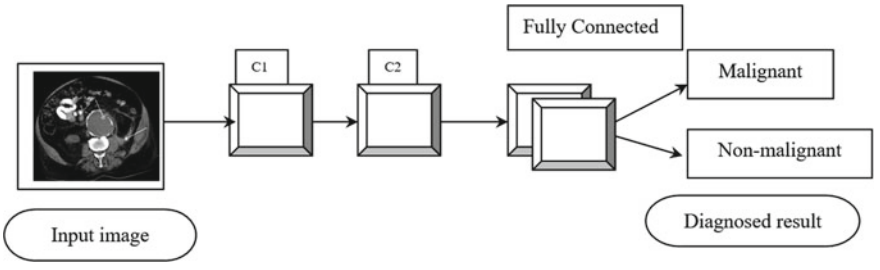


Fig. 7.3 CNN architecture

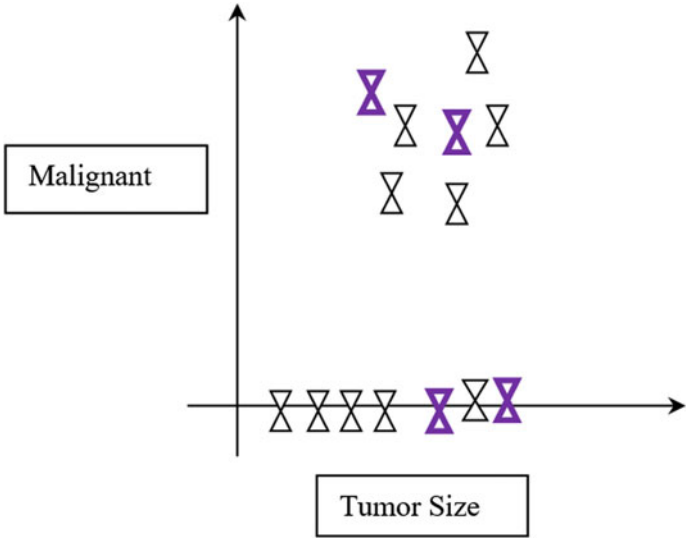


Fig. 7.4 Cancer diagnosis using CNN

and is customarily advantageous with confined annotated data, to name a few are image cytometry. In these cases, manual annotations are found to be time-consuming to be acquired, necessitating a higher level of expertise to make. Besides, the biomedical images for cancer diagnosis said to be trained using CNNs that are extracted with specific experimental settings also results in substandard detection or suggestions.

To eliminate the above-said disadvantages, large annotated datasets are said to be first pre-trained using the new CNNs. Followed by which the transferred parameter values generating better initial values are said to be regulated to the pertinent chosen data. In addition, transfer learning also permits the attachment of intense networks according to hardly any task distinct annotated images. Besides speedy concurrence is also said to be attained contributing to a higher level of classification performance and findings. Therefore, by integrating both transfer learning and deep CNNs, results in a higher and accurate classification rate.

7.2.2 Data Augmented Convolved Neural Network

Data augmentation [3] in machine learning refers to the methods that artificially enhance a dataset by administering a modification on the existing samples, hence augmenting the quantity of accessible training data. Despite the presence of new data points that are not self-sufficient and uniformly dispersed, absolutely establish the models and enhance findings, as confirmed by statistical learning. For the past few years, data augmentation has been in the long utilized in machine learning and it has

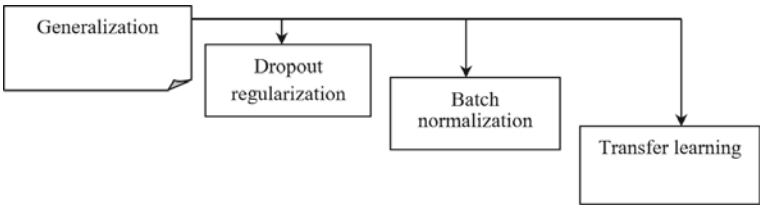


Fig. 7.5 Schematic generalization views

been recognized as an analytical element of several models. Framing a high-level context of data augmentation for CNN includes the following.

Figure 7.5 given below shows the schematic view of generalization performance with data augmented CNN.

As depicted in the above figure, the three generalization view for data augmentation address over fitting from the root problem, where the design is preceded with the assumption that more information is said to be extracted via augmentation, therefore contributing efficient disease diagnosis, specifically, a cancer diagnosis.

7.2.2.1 Dropout Regularization Technique

Dropout involves a regularization technique that proceeds with the assumption of zeroing out all the activation values. This is performed by selecting the neurons in a random manner during simulation. With the aid of this threshold value, the network in turn is said to obtain features that are said to be more robust in nature. Besides, with the assistance of new type of dropout called, Spatial Dropout, instead of dropping individual neurons, the entire features were dropped.

7.2.2.2 Batch Normalization Regularization Technique

One of the sought out regularization technique that is designed based on the normalization of activation sets is the batch normalization. Here, the normalization is said to be performed in batch as it includes both subtraction and division. Here, two operations are performed. First, the subtraction operation is performed with the batch mean and activation. Followed by which, a division is performed via standard deviation. This batch normalization technique, in addition to the standardization technique, remains to be a standard technique while performing preprocessing during the image processing, followed by a better cancer diagnosis is said to be ensured with the elimination of irrelevant features or pixels.

7.2.2.3 Transfer Learning Regularization Technique

Transfer learning regularization technique helps in preventing over-fitting. Network training is said to be performed via Transfer Learning. Then, these weights are then used as initial weights during image categorization. In this transfer learning process, specifically, the weights in convolutional layers are obtained. This transfer learning regularization are very efficient as several medical databases possess low-level spatial features that are comparative learned in a finer manner with big data.

7.2.3 Unbalanced and Skew Supervised Deep Learning

As far as supervised learning models are concerned, labeling of training data is highly required and in case of the classification each image specimen pertains to a familiar group. Consider two different types of groups, called the majority group and minority group. While considering a dual categorization issue pertaining to image specimens from dual different classes, the group disparity is said to occur when single classification, the minority group, consists of an extensively lesser amount of specimens in comparative to the second classification, the majority group.

In several image classification and diagnosis issues, the minority group remains to be the categorization of significance, i.e., the positive class or simply the malicious cases correctly identified as malicious. On the other hand, the majority group remains to be to other classes of interest, i.e., the negative class or simply the normal cases incorrectly identified as malicious.

To name a few, a prominent class disparity [4] machine learning framework is the job of a cancer diagnosis. Here, detecting the disease remains to be the main factor to be analyzed with maximum number of patients found to be in the healthy stature. With the maximum number of patients to be found in the healthy condition are classified as in negative class section. With these disparity datasets, learning is said to be a highly cumbersome process, specifically when dealing with big data. Hence in these cases, certain non conventional model of learning techniques is applied to arrive at the expected results.

An in-depth investigation of the class disparity issue and the algorithms accessible for addressing the above-said class imbalance problem is indispensable, due to the reason that skewed data are said to prevail in several real-world applications. Therefore, whenever class imbalance is said to exist within training data, the over-classification of the majority group is said to occur due to its escalated antecedent likelihood. As a result, the specimens pertaining to the group in the minority section is said to be misclassified customarily than when compared to the major section.

7.2.4 Spatio Temporal Pattern-Based Recurrent Neural Network

One of the classifications of ANN is the Recurrent Neural Network that associates the nodes in such a manner resulting in a graph of directed nature both stretching in a temporal and spatial sequence. This, in turn, ensures both temporal dynamic behavior and exhibit spatial information. When compared to a feed-forward neural network, RNNs are said to be utilized with their internal memory with the purpose of processing sequences of inputs. This makes RNN to be applied to certain tasks such as object recognition, medical imaging or speed recognition and so on. The basic structure of RNN is given below (Fig. 7.6).

As depicted in the above figure, the basic RNN consists of three different layers. They are input, hidden and output layer. An RNN is able to successfully encapsulate the spatial and temporal factors present in medical images via the incorporation of relevant filters. The spatial-temporal factors with RNN conduct a better fitting to the medical images due to the reduction in the size of parameters involved and weight reusability. In other words, the RNN with the involvement of spatial and temporal factors perceives the image in a sophisticated manner.

The term spatiotemporal RNN refers to dual network classes. They are spatiotemporal finite impulse and spatiotemporal infinite impulse. The spatiotemporal finite and infinite networks exhibit temporal and spatial dynamic behavior. When compared to the spatiotemporal infinite impulse, being a directed cyclic graph in nature are said to be uncovered and restored with a feed-forward neural network.

The spatiotemporal infinite impulse RNN cannot be uncovered. Both spatiotemporal finite and infinite RNN possess an additional stored state. The storage is not performed with the neural network. Upon spatial or temporal constraints the storage is said to be restored via one more network if that causes delay in timing, also referred to as the Feedback Neural Network.

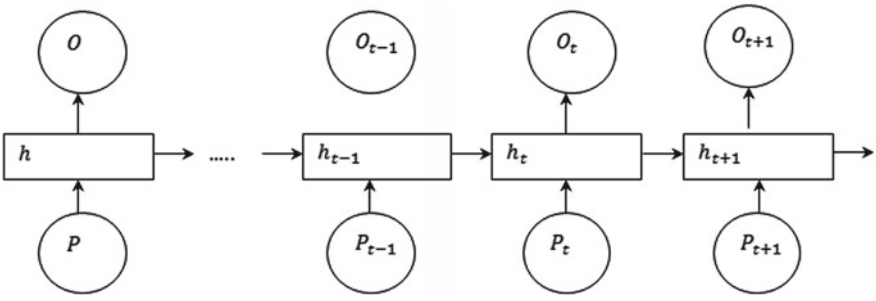


Fig. 7.6 Structure of basic RNN

7.3 Image Processing Techniques

The Image process is a method that is customarily utilized to enhance raw input images that are acquired from different resources. The Image processing technique transforms an image into a digital form and executes certain actions on it, with the purpose of obtaining an enhanced image. The purpose of image processing is split into different groups that are listed below (Fig. 7.7).

An in-depth investigation of image processing techniques is elaborated in the following sections.

7.3.1 Visualization or Image Acquisition

The first part of every visualization scheme is image acquisition or visualization. Image processing is utilized to recognize those objects that are not observable. When the image is obtained then several processes are applied to the image and the most prevalent method for image acquisition is real-time acquisition.

7.3.2 Image Edge Sharpening and Image Restoration

Medical image enhancement techniques have received many advantages in recent years and hence found to be invariably applied in the medical field. This is due to the reason that enhanced images are required by the doctors to assists in disease diagnosis that are affected by noise and other devices with which the images are acquired. In image processing, several materials and methods are applied to the raw picture with the purpose of obtaining a refined image.

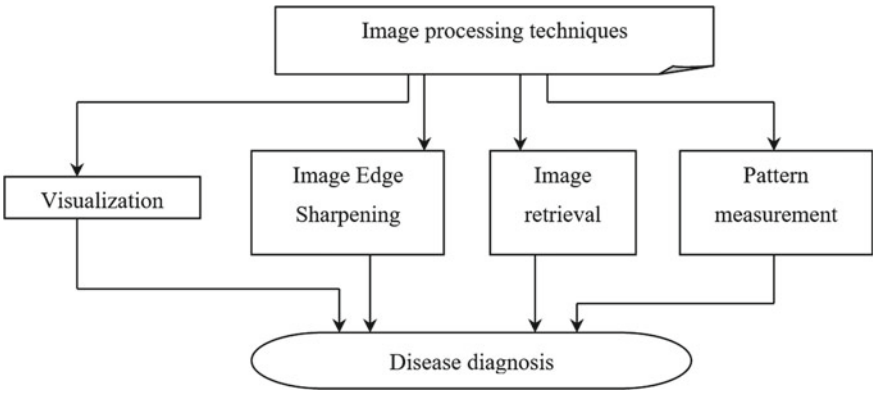


Fig. 7.7 Schematic views of image processing techniques

Image restoration involves a method with which a noisy image is processed in such a manner that a fine-tune image is obtained. Image restoration concentrates on the processing of medical images suitable for a specific application. This is due to the reason that image enhancement that works well for X-Ray does not fit for CT scans and so on. Hence, restoration reconstructs those images whose quality of an image is spoiled due to several aspects like, noise, contrast, error and so on.

7.3.3 Image Retrieval

With the extensive utilization of digital imaging data in hospitals, the medical image size repository where the large or huge image dataset is kept is progressing swiftly. This results in cumbersome management and querying resulting in the requirement of Content-Based Medical Image Retrieval (CBMIR) systems. With the aid of image processing, the user detects only those parts of the picture that is of high relevance to the user. Specifically, two distinct representations are used for image representation. They are

- Boundary representation.
- Region representation.

On one hand, boundary depiction exhibit central appearance of the picture that refers to the object shape, i.e., sharp edge, curve edge or rounded edge or any other shape. On the other hand, region representation refers to the internal characteristics.

7.3.4 Pattern Measurement

Various constituents in an image are investigated. The major aspects of pattern measurement rely on the capturing of pictures from several resources. While capturing the images, image quality is said to be compromised. Pattern measurement adjusts the elements of the pictures so that image clarity is said to be improved. Two different types of patterns are said to be measured. They are

- Local pattern measurement.
- Global pattern measurement.

Global patterns are said to be homogenous in nature and on the other hand, the local patterns are said to be droplets, black dots, vascular in appearance or blue-white surrounding. The rationale of these measurements remains in the evaluation of qualitative factors separately.

7.3.5 Image Recognition

Finally, with the above-mentioned steps, the actual image is said to be recognized or in other words, features in an image are identified. With the above-said image processing techniques, the process involved in the design for better cancer diagnosis is depicted in the figure given in 7.8.

The elaborate description of the four different processes is described below in the following sections.

7.3.6 Bi-cubic Interpolated Pre-processing

Raw medical images comprise of a certain amount of noise in it, hence the first and foremost step in image processing for better cancer diagnosis is preprocessing, i.e., improving image standard via irrelevant image discarding. If not properly addressed several levels of inaccuracies are said to persist. Several types of filtering mechanisms are therefore applied to the raw images for the elimination of statistical noise, salt and pepper noise, Shot noise, and besides certain filtering mechanisms, like, nonlinear digital image filtering, sliding window spatial filter, adaptive median filter, Gaussian smoothing and linear time invariant filter.

The correct blend of pre-processing tasks results in a higher amount of accuracy rate. Some of the preprocessing methods are contrast enhancement, color equalization, elimination of black frame, Karhunen–Loe've transform, blur and noise removal, pseudo-random filter and so on. Preprocessing in cancer diagnosis specifically comprises of delineation of tumors, performing mammogram labeling and orientation using preprocessing.

However, with low-resolution images, one of the addendum of interpolation between lists of images is bi-cubic interpolation. Here, a regular grid type structure in the form of two dimensional structures is utilized to interpolate the data points. Here, the interpolated plane is found to be even than corresponding surfaces generated via bilinear interpolation [5]. It is said to be accomplished by means of interpolating in-determinates, piecewise third order polynomials, or edge-directed image scaling algorithm.

While processing images for diagnosing cancerous portions, bi-cubic interpolation is specifically utilized than the bilinear interpolation. On the other hand, when compared to interpolation in the form of bilinear patterns, that only utilizes 4 pixels, 16 pixels are considered while applying bi-cubic interpolation. Therefore, re-sampled images via bi-cubic interpolation are found to be highly smoother and have fewer noises with low contrast also.

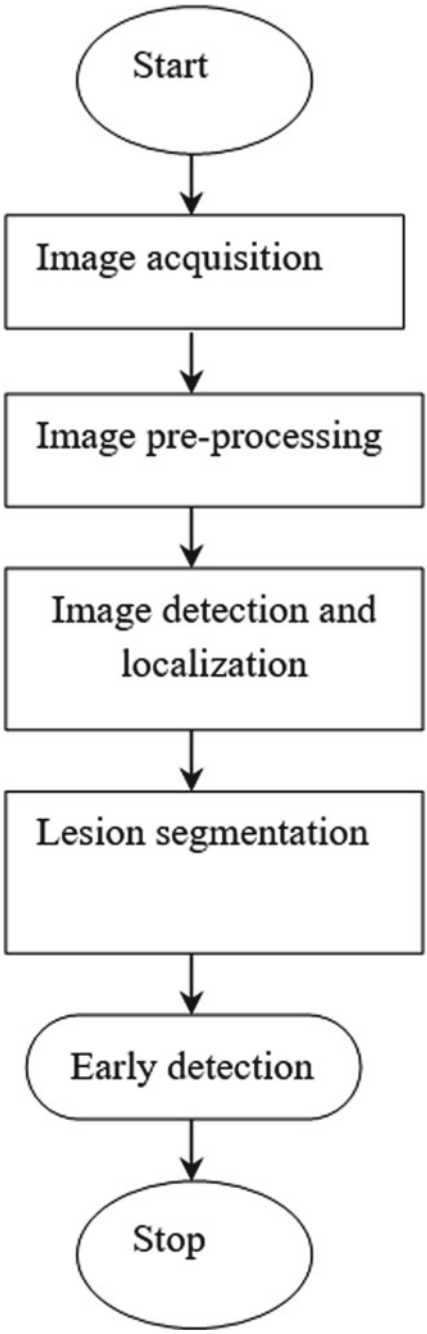


Fig. 7.8 Image processing flow diagram

7.3.7 Robust Computer-Aided Detection and Localization

For the last two decades, CAD has been found to be a major attraction with several researches conducted in this area. However, at the same time, the fullest advantage is said to be only attained by only improving the overall system quality of service. This therefore necessitates a fine assessment model that requires optimization of the algorithms in the part of the CAD designers.

In order to assess the CAD performance, Free Response-receiver Operating Characteristic Curves (FROC) is extensively utilized to measure the CAD performance. But, a fine assessment hardly ever interests beyond. Recent studies conducted in the area of medical image processing have manifested that low dose computed tomography (LDCT) can prove to be a significant screening mechanism to minimize the mortality rate involved in cancer. In such cases, Computer-Aided Detection (CAD) would be an advantageous one for doctors and radiologists.

Studies have illustrated [6] that though Iterative Reconstructions (IR) results in the improvement of cancer diagnosis quality, however, degrade the CAD performance or is said to increase the false positive rate. With the purpose of enhancing the CAD performance, retraining with robust detection mechanism and localization as a standard preprocessing technique remains to be the solutions.

7.3.8 Maximum Likelihood Lesion Segmentation

Segmentation refers to the process of splitting the image into several regions or parts. As far as image segmentation is concerned, an image is split into subparts based on the requirement of the user or the issue to be addressed with respect to a certain problem, where the image is split into pixels.

In other words, image segmentation splits the image in such a manner that the required image portions to be processed again is called as the segmentation and is found to be highly accurate. On the other hand, the outcome of image segmentation lies in a set of sections or groups that altogether shields the entire image discarded from the image. Hence, the objective of image segmentation lies in simplifying the illustration of picturesque in such a form that it is found to be highly notable and simple to measure, resulting in finer image appearance.

Segmentation is primarily the dissociation of a locality to be obtained from the image backdrop, when the locality to be obtained remains the major portion that is said to be utilized for processing. In diagnosis of cancer, the lesion part is said to be extracted from the diseased part. However, the largest probability assessment has different competent attributes.

The largest probability assessment is said to be only efficient, when the assessment factor is unbiased in nature and also evolves to a minimum bound on the variance. The purpose of applying the largest probability assessment function is to arrive at the resultant values that improves the potentiality of the existence of the evaluated

average signal forces at all the calculated values. Hence, the probability assessment factor necessitates a precise and stringent probability formulation of the measured average signal force.

7.3.9 Early Detection and Assessment

Early diagnosis [7] is referred to as the early cancer identification in patients those who have disease symptoms, on the other hand, cancer screening involves the process of obtaining the cancerous portions that are invisible to the naked eye or pre-cancerous lesions in an evidently healthy chosen population. However, both screening and cancer early diagnosis are said to be highly important, but the fundamental variation refers to the resource required for diagnosing the disease, requirements in terms of infrastructure and monetary cost involved in the diagnosis play a major factor.

The concentration of diagnosing cancer in an early stage is cancer early diagnosis is patient those with the indication and signs as that of a patient with cancer. The purpose of early detection and assessment remains in identifying the symptoms as early as possible and associating the patient for remedy without any time delay. When done at a prompt stage, cancer is said to be detected at a potentially curable stage, therefore enhancing both the survival rate and also the life cycle. For early diagnosis some of the three steps involved are:

- Consciousness of cancer indications and acquiring attention.
- Clinical assessment, recognition and directing.
- Acquiring to therapy.

From the above analysis, certain key elements or early cancer diagnosis are

- Volume of participants.
- Test.
- Health system requirements.
- Awareness and assessing care.
- Training and human resource needs.
- Clinical evaluation.
- Public awareness.
- Follow up.
- Potential advantages.
- Potential limitations.

7.4 Survey of Improved Deep Learning Applications in Cancer

Different research works and state-of-the-art methods providing deep learning techniques, methods, principles, ancient surveys, and applying the same in different research areas have been presented. This section manifests an appraisal of improved deep learning applications in cancer using different microscopic image evaluation that is considered to be a hot topic of research.

Definitely, in short, it presents the amateur improved deep neural networks in the specific area, narrates the forecast of cancer susceptibility, recurrence and survival rate in detail. With the application of improved deep learning in AI can enhance the process of updating deep learning from conventional toward medical images and enhance the overall system performance.

- Prediction of cancer vulnerability.
- Prediction of cancer frequency.
- Prediction of cancer continuity.

7.4.1 Prediction of Cancer Vulnerability

The major findings in designing answerable tools that differentiate between cancerous and non-cancerous images while diagnosing cancer are risk stratification. According to the researchers who have conducted the field of research in this area, conventional works based on the utilization of computer-aided diagnosis methods have used specific machine learning techniques such as deep learning to evaluate the danger of the patient surviving from cancer.

Several researchers have applied deep learning with the objective of designing a prediction method that efficiently and effectively classifies cancerous mammographic images from non-cancerous mammographic images. The method was constructed utilizing a different numbers of hidden layers that generalized comparatively in a better manner than a smaller number of hidden nodes. The mammographic records obtained by the radiologists were also reviewed by obtaining the reading information.

Followed by which the performance was measured via ten-fold cross-validation even by discarding over fitting. Here, the network error was said to be controlled during the training and if in case of over fitting the training was stopped. Besides, two fascinating features are the assessment of two main elements of accuracy, specifically, distinction and assessment. In the case of distinction, it is used to evaluate with the purpose of separating benign abnormalities from malignant images. On the other hand, assessment is used as a measure wherein the patients are classified into high or low-risk classes. With this, a Receiver Operating Characteristics is plotted to measure the discerning capability of their model.

7.4.2 Prediction of Cancer Frequency

In this section, the prediction of cancer frequency is discussed. In a study conducted on recurrence prediction of Oral Squamous Cell Carcinoma (OSCC) [9] a multi-parametric decision making activities was designed to evaluate the OSCC evolvement via heterogeneous data and hence a subsequent recurrence.

Besides, accuracy, true positive rate and false positive rate were evaluated and used for differentiation between the classification methods being used. Besides, more than one classification technique or ensemble of classification was also utilized to obtain robust results. A Support Vector Machine based model for analyzing the recurrence of breast cancer was also performed. Here, with the categorization of patients suffering from cancer into risky or non-risky, treatment and follow-up planning were made in an effective manner. Besides, machine learning techniques involving hyperplane classification, neural networks and regression models were also utilized for obtaining the optimal results. Here, too, accuracy, true positive rate and false positive rate were measured for estimating the cancerous and non-cancerous patient in a more reliable fashion.

7.4.3 Prediction of Cancer Continuity

For the forecast of cancer, survival robustness plays a major role under the model's parameter disparity. Survivability here refers to both types of patients who have not survived and who have survived after being diagnosed at the early stage using improved deep learning techniques.

Amongst the most revealing characteristics are the tumor size, numbers and age of tumor during the diagnosis stage. Five-turn hybrid-recognition was utilized for measuring the predictive performance models. Based on the finding results, the Semi-Supervised Learning model was found to be the best model for analyzing the survival rate, where the most informative features were obtained without separate pre-processing steps.

7.4.4 Measuring Sensitivity, Specificity, and Accuracy

In several tests, including cancer diagnosis, sensitivity refers to the magnitude to which genuine positives cases are not ignored and in this case, the false-negative rates are said to be few. On the other hand, specificity [8] refers to the magnitude to which genuine negative cases are categorized as such and in this case false positives are said to be few.

Therefore, a highly sensitive test hardly ever fails to observe a genuine positive (for scenario, presenting tumor nonexistence in spite of certain tumor being observed).

On the other hand, as far as a genuine positive (for scenario, presenting nonexistence tumor in spite of certain tumor being observed) is concerned, a highly specific test seldom gives a positive categorization for any images however, this may not be the case for testing.

7.4.4.1 Sensitivity

Sensitivity denotes the test's potentiality to precise identify patients those who are ill who do possess the cardinality or who are actually ill. For example, of a medical test used to diagnose cancer, the sensitivity also referred to as the detection rate refers to the percentage ratio of the patient testing to be positive for the disease scenario disease among those patients who actually possess the disease. The sensitivity is formulated as given below.

$$\text{Sensitivity} = \frac{\text{No of true positives}}{\text{No of true positives} + \text{No of false negatives}}$$
$$\frac{\text{No of true positives}}{\text{Total no of cancer diagnosied patients in samples}}$$

7.4.4.2 Specificity

Specificity refers to the test's potentiality to precisely reject non-tumor patients without a condition. For example, a medical test for diagnosing cancer, the specificity of a test is the ratio of a healthy or non-tumor patient known not to have the disease, testing negative for it. Mathematically, this can also be written as

$$\text{Specificity} = \frac{\text{No of true negatives}}{\text{No of true negatives} + \text{No of false positives}}$$
$$= \frac{\text{No of true negatives}}{\text{Total no of non cancer patients in samples}}.$$

7.4.4.3 Accuracy

Let us consider a scenario measuring new hypothetical testing that identifies patient for a cancer diagnosis. Each patient those who are subject to test either is said to be a cancerous patient or referred to as non-cancerous patient. The results of the test may be positive, i.e., diagnosing the patient with cancer or the results may be negative i.e., diagnosing the patient without cancer. However, there also arises situation where the results of the test for each patient may not match with the patient's actual status. In

that setting:

$$\text{TP(True positive)} = \text{Correctly identified}$$
$$\text{FP(False positive)} = \text{Incorrectly identified}$$

From the above settings, the accuracy is used to measure the resultant value and it involves classification of results in a dual state with the correct identification of the said condition. On the other hand, accuracy refers to the percentage ratio of true results with respect to the total patients considered for experimentation. The formula for quantifying binary accuracy is written as given below.

$$\text{Acc} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{FP} + \text{TN} + \text{FN})}$$

More commonly as the definition is given above, accuracy refers to a description of systematic errors or refers to the proximity of the measurements to a determined value.

References

1. Cancer Facts and Figures 2019, American Cancer Society, 2019. Available online https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annualcancerfacts-and-figures/2019/cancer-facts-and-figures-2019.pdf
2. K. Kourou, T.P. Exarchos, K.P. Exarchos, M.V. Karamouzis, D.I. Fotiadis (2014) Machine learning applications in cancer prognosis and prediction. Comput. Struct. Biotechnol. J.
3. Hernández-García A, König P, *Further Advantages of Data Augmentation on Convolutional Neural Networks* (Springer, 2018)
4. J.M. Johnson, T.M. Khoshgoftaar, Survey on deep learning with class imbalance. J. Big Data (2019)
5. S.E. El-Khamy, M.M. Hadhoud, M.I. Dessouky, B.M. Salam, F.E. Abd El-Samie, A new approach for regularized image interpolation. J. Braz. Comput. Soc. **11**(3), 65–79 (2005)
6. L. Raghupathi, P.R. Devarakota, M. Wolf, Learning-based image preprocessing for robust computer-aided detection, *Medical Imaging 2013: Computer-Aided Diagnosis*, vol. 8670, International Society for Optics and Photonics, p. 867034 (2013)
7. Guide to cancer early diagnosis, World Health Organization (2017)
8. https://en.wikipedia.org/wiki/Sensitivityandspecificity
9. K.P. Exarchos, Y. Goletsis, D.I. Fotiadis, Multiparametric decision support system for the prediction of oral cancer reoccurrence. IEEE Trans. Inf Technol. Biomed. **16**, 1127–1134 (2012)

Chapter 8
Using Deep Learning Techniques in Detecting Lung Cancer

Osamah Khaled Musleh Salman, Bekir Aksoy, and Koray Özsoy

Abstract Today, with the rapid rise in the number of illnesses, there is a significant increase in the number of people who died due to these diseases. Nowadays, cancer diseases, in particular, are one of the important types of diseases that cause fatal outcomes. The World Health Organization stated that approximately 9.6 million people died from cancer worldwide in 2018. According to the World Health Organization, among these cancer types, approximately 1.8 million people pass away from cancer. Lung cancer has been identified by the World Health Organization as the deadliest cancer type among all cancer types. For this reason, the early diagnosis of lung cancer is very important for human health. Computed Tomography (CT) images are frequently utilized in the detection of lung cancer. In this book section, academic studies on the diagnosis of lung cancer are examined.

Keywords Deep learning · CT images · Lung cancer

8.1 Definition of Cancer

Cancer is a tumor that occurs when cells in organs or tissues in any part of the body divide and multiply irregularly. As a consequence of the uncontrolled increase of cells, the normal working order of the body is disrupted. According to statistics, the number of people with cancer in the world is about 12.7 million per year, but this number is estimated to rise to 21 million by 2030 [1]. When the causes of death are

O. K. M. Salman · B. Aksoy (✉)
Faculty of Technology Mechatronics Engineering, Isparta University of Applied Sciences, Isparta, Turkey
e-mail: bekiraksoy@isparta.edu.tr

O. K. M. Salman
e-mail: yl1930654007@stud.sdu.edu.tr

K. Özsoy
Senirkent Vocational School Electric and Energy Department, Isparta University of Applied Sciences, Isparta, Turkey
e-mail: korayozsoy@isparta.edu.tr

U. Kose and J. Alzubi (eds.), *Deep Learning for Cancer Diagnosis*,
Studies in Computational Intelligence 908,
https://doi.org/10.1007/978-981-15-6321-8_8

examined, it is seen that one in every five men and one in every six women in the world get cancer [2]. Therefore, early detection and treatment planning process is very important for cancer to increase the probability of survival.

8.1.1 Definition of Lung Cancer

The lung cancer is among the leading types of cancer that causes death, considering the mortality rates (due to cancer) in both men and women [3, 4]. That cancer type is caused by the formation of a tumor inside the lung as a result of uncontrolled cell proliferation. Lung cancer first occurs in the lungs, but in the later stages it can cause damage to the surrounding tissues. The majority of patients in lung cancer are Non-Small-Cell-Lung-Cancer (NSCLC), while Lung-Adenocarcinoma (LUAD) and Lung-Squamous-Cell-Carcinoma (LUSC) are the most common subtypes [5]. The main cause of lung cancer is genetic and epigenetic damage caused by active or passive exposure to tobacco smoke [6].

8.2 CT Imaging

Computed-Tomography (CT) imaging, which is one of the medical imaging methods, is an important radiological diagnostic method used in the diagnosis and treatment of lung diseases [7]. CT allows to create a cross-sectional image of a region using X-rays on the desired area of the body. By combining the images of the region taken from different angles, a three-dimensional view of the inner structure of that region can be obtained [8]. The creation of the CT image takes place in three stages as (1) scanning phase, (2) reconstruction phase, and (3) digital-to-analog conversion phase. In the scanning phase where X-ray tube and detectors are used, the projection of X-rays passing over the body section is taken. After the scanning phase is completed, the data obtained from the detectors that collect the rays are transferred to the computer digitally and processed with different algorithms. CT images consist of many points called pixels, and each of these pixels has a different volume according to the section thickness. At the last stage, digital data is converted to analog data by using electronic components to make the image grayscale [9, 10].

8.3 Deep Learning

Deep learning is briefly a method of machine learning, as a sub-branch of the artificial intelligence that started to spread in 2000s [11]. Deep learning allows to estimate the output using the given dataset and to train artificial intelligence according to this result [12]. Multi-layered neural networks were created by Hinton in 2006 to process

data such as images, sound, text, and output, and were then used frequently in deep learning practices [13]. The reason why deep learning is becoming more widespread nowadays is that the amount of data required for education is sufficient and there is an infrastructure to process this data [14]. Deep learning generally compromises of an input layer, three or more concealed layers and an output layer [15]. Since the deep learning consists of more than one layer, the learning process takes place more successfully [16]. In deep learning, the learning process can be carried out as supervised, unsupervised and semi-supervised [17]. In deep learning architecture, the feature learned in each layer is the input of the other layer [18]. Deep learning is used in a wide range of fields such as object recognition [19], speech recognition [20], and natural language processing [21].

8.3.1 Convolutional Neural Network (CNN)

CNN is an algorithm that is inspired by the visual perception mechanism of living things and is frequently used in deep learning [22–24]. The CNN algorithm has gathered attention with its remarkable results in ImageNet competitions, and has been widely used especially in image processing techniques [25, 26]. CNNs are two-dimensional neural networks and are frequently used in image analysis, image classification, clustering, object recognition, medical image analysis, and natural language processing [27, 28]. In CNN, weights are learned by the system and features can be determined by filtering (convolution) techniques [29]. As CNN structure, Fig. 8.1. As seen, it is formed by adding convolution, pooling (sampling) and fully connected layers to the basic neural network layers [30]. In the convolution layer, feature extraction is performed by using a convolution process with a dot product between the input and the kernel. The pooling layer is generally located between two convolution layers. The purpose of the pooling layer is to reduce the calculation load on the system by reducing the size of the feature maps and to perform a secondary feature extraction [31, 32]. Some types of pooling used in this layer are max pooling, stochastic pooling, mixed pooling and average pooling [32, 33]. In the last layer, which is called fully connected layer, neurons are directly connected to all previous neurons as in classical neural networks and they act as classifiers in the CNN model.

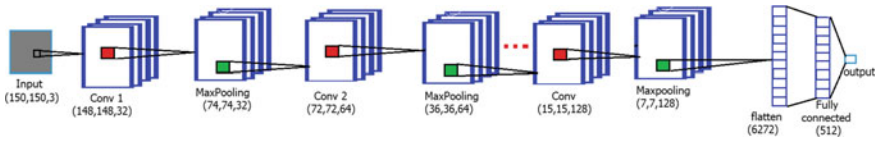


Fig. 8.1 Convolutional neural network layers

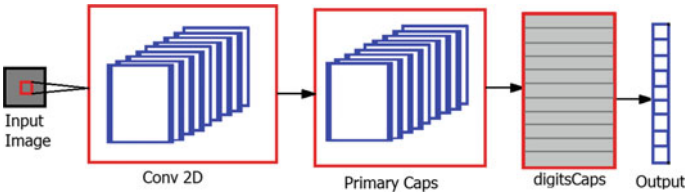


Fig. 8.2 CapsNet architecture

8.3.2 Capsule Neural Network (CapsNet)

Another deep learning algorithm used in the analysis of medical images is Capsule Networks (CapsNet). Although successful results of CNN architectures are noteworthy within the scope of deep learning algorithms, the fact that pooling processes in this algorithm cause loss of information led to the development of alternative algorithms. In this context, the CapsNet algorithm, which provides better learning through capsule layers consisting of a group of neurons, has been introduced to the literature in order to prevent the loss of knowledge that take place during changes in input parameters [34, 35]. The depth of a typical CapsNet architecture is created not only with layers, but also with the help of capsules in one layer, thus, better feature extraction and learning can be obtained. In addition to the ReLU function in CNN architectures, the Squash function, whose inputs and outputs are vectors, is also used in CapsNet [36, 37]. Although these improvements show that the results obtained with CapsNet are better than CNN-based architectures, it is also seen that using CapsNet capsules also requires extra computing power [38]. General components of CapsNet architecture can be seen in Fig. 8.2.

8.3.3 The Importance of Deep Learning Techniques in Detecting Lung Cancer

Detecting the lung cancer earlier can save lives. Because of that, it is very important to correctly classify the nodules in CT images. It is observed that the use of deep learning methods gives fast and accurate results and helps physicians in the stage of cancer prognosis [39]. In addition, early start of the treatment process with early detection of cancerous cells using deep learning techniques is effective for reducing the rates of mortality due to lung cancer. Recently, many datasets have been prepared for the detection of lung cancer. Common datasets used in the detection of lung cancer are given in Chap. 4.

8.4 Commonly Used Datasets for Lung Cancer Detection

In order to apply deep learning methods in the detection of lung cancer, large amounts of data and computer computing power to analyze this data are needed [40]. Some of the datasets frequently used in the detection of lung cancer are.

8.4.1 LIDC/IDRI Dataset

The dataset, called LIDC (Lung-Image-Database-Consortium) and IDRI (Image-Database-Res.-Initiative), are among the frequently used datasets for performing detection regarding the lung cancer. The LIDC/IDRI dataset was prepared in collaboration done by eight medical imaging companies, and also seven academic centers, as including 1018 cases. In cases, there are clinical thoracic CT-scan images and the XML files providing information for the results of two-phase image annotation process done by four experienced thoracic radiologists. Each data contains CT images and an XML report file created by an expert [41].

8.4.2 LUNA-16 Dataset

That dataset is briefly a sub-set within the LIDC/IDRI dataset. The Luna 16 dataset comprises of 888 pectoral CT scan images, marked by at least three of the four medical specialists for each lesion condition. In this data set, only nodules smaller than 3 mm in diameter are accepted as the positive samples, while all rest of the lesions are called negative samples [42]. The Luna 16 dataset consists of three-dimensional images and a CSV file with annotations [43].

8.4.3 NLST (National-Lung-Screening-Trial) Dataset

The NLST dataset consists of data from 53,454 people with a high risk of lung cancer in 33 medical centers in the USA between August 2002 and April 2004. The data set was created by randomly assigning 53,454 volunteers to undergo three annual screenings with either low dose CT (26,722 participants) or single-view posteroanterior breast radiography (26,732 participants) [44]. NLST data set has been used in studies since 2009 [40].

8.4.4 ELCAP (Early-Lung-Cancer-Action-Program) Dataset

ELCAP is a dataset created by the collaboration between the ELCAP and VIA research groups. This data set was created to develop a common dataset that can be used to assess the performance of difference computer-aided detection systems. The ELCAP dataset was first released in December 2003. ELCAP dataset is a prototype for web-based image data archives. The ELCAP dataset consists of 50 low dose CT scan images with a cross section thickness of 1.25 mm. In the ELCAP data set, the locations of the nodules detected by expert radiologists are also given [40, 45].

8.4.5 Data Science Bowl 2017 (KDSB17) Dataset at the Kaggle

The KDSB17 dataset was obtained from the axial CT scan image of the chest cavity of 2101 patients [46]. Each CT scan image in the KDSB17 data set is labeled as 'with cancer' if the patient is diagnosed within 1 year, otherwise it is labeled as 'without cancer' [47]. In the KDSB17 data set, the nodules are not labeled according to their location and size.

8.5 Literature Research

When the academic studies for the detection of lung cancer with deep learning are examined;

Anthimopoulos et al. proposed the CNN model, which is the sub-branch of deep learning, for the classification of lung disease samples. The proposed model was evaluated on 12,096 CT scan images of 120 patients and consists of 5 convolutional layers, an average pooling layer and 3 dense layers. The model in the study uses Leaky ReLU activation function and Adam optimizer as an optimizer. As a result, they achieved classification performance of approximately 85.5% [48].

Bayraktar et al. used deep learning methods to determine whether the cancerous cell is benign or malignant. A dataset of 669 samples with 10 different features was used to diagnose cancer caused by irregular division of cells. In order to get the best results, 2 and 4 layer models were developed, epoch numbers, activation functions and batch-size were changed. In the first model of the application developed using the Python programming language; first layer is composed of 70 nodes, second layer is composed of 60 nodes, the number of epochs is taken as 100, the Exponential function is selected as the activation function and Adadelta selected as an optimizer. In the second model; first layer is composed of 70 nodes, the other three layers consist of 60 nodes, the number of epochs is taken as 100, the ReLU function is selected as the activation function and Adamax selected as an optimizer. As a result, a success

rate of 99.05% was determined for both models. It has been found that Batch size has no effect on the model and increasing the number of epochs has a negative effect on the success of the model [49].

Cheng et al. conducted studies on the computer-based diagnosis (CADx) system to determine whether lesions and tumors are benign or malignant. In their work, they used the Stacked Denoising-Autoencoder (SDAE) architecture, which is a more developed version of the Stacked-Autoencoder (SAE) deep learning model. This structure is a model based on input reconstruction by making use of the unknowns in the input data, but it also has the feature to automatically discover different patterns. This feature allows the changes in the appearance of the lesions to be more effectively removed and compared. In conclusion, in this study, SDAE architecture was compared with CURVE, RANK and MORPH on lung CT and breast ultrasound images and SDAE architecture has been shown to perform better compared to the other three algorithms [50].

Sun et al. used CNN, Deep Belief Network (DBN) and Stacked Denoising-Autoencoder (SDAE) algorithms for the diagnosis of computer-aided lung cancer. In their study, they used 52 × 52 pixels, 174,412 sample datasets on CADx system. They obtained accuracy ratios of 0.7976, 0.8119 and 0.7940 for CNN, DBN and SDAE algorithms, respectively [51].

Coudray et al. used CNN algorithm (inception v3) to classify the common lung cancer types, LUAD, LUSC, and the gene mutations of cancer. In the study, it was demonstrated that the presented method successfully detected lung cancer types with 0.97 area under the curve (AUC) using dataset containing 512 × 512 pixel tiles, 1176 tumor tissue and 459 non-tumor tissue images [52].

Ciompi et al. used CNN to classify nodules with abnormal tissue growth in lung cancer screening. The model was trained with Italian MILD screening trial data containing 943 patients and 1352 nodule images, and tested with Danish DLCST screening trial data containing 468 patients and 639 nodule images. As a result of the study, when the analyzes made by human observers and computer were compared, it was found that 3 scale CNN had 79.5% accuracy rate [53].

Hua et al. used DBN and CNN for the categorization of lung nodules in CT images. In the categorization of pulmonary nodules, it was found that deep learning methods performed better than traditional computer-aided diagnosis (CAD). The performance sensitivity and specificity of DBN and CNN models were shown to be 73.4%, 82.2% for DBN, 73.3% and 78.7% for CNN, respectively [54].

Ardila et al. used deep learning algorithms in low-dose computed tomography (LDCT) to detect the risk of lung cancer. In the study, a 3-dimensional (3D) CNN model was created. With this model, analysis of LDCT data for tumor detection is performed for each patient. Presented model achieved 94.4% AUC on 6716 National Lung Cancer Screening Trial cases [55].

Song et al. used deep learning methods to classify lung pulmonary nodules on CT images. They used CNN, DNN and SAE deep learning algorithms in their studies. The study was carried out using the LIDC-IDRI database containing 1010 cases and 244,527 images. As a result, it was found that the CNN algorithm performed best

with an accuracy of 84.15% in the classification performed to determine whether the nodules are benign or malignant [56].

Lakshmanaprabu et al. performed analysis on CT images for lung cancer detection using Optimal Deep Neural Network (ODNN) and Linear Discriminate Analysis (LDA). In order to classify the lung nodules as benign and malignant, the dimensionality of the images was extracted using feature Latent Dimensionality Reduction (LDR). After feature reduction, LDA was used to reduce computation time. The model is optimized using the modified gravitational search algorithm (MGSA). With the ODNN model proposed in the study, 96.2% accuracy, 95.26% sensitivity and 96.2% specificity performance were obtained [39].

Hosny et al. used deep learning for lung cancer prediction (prognosis) in their studies. In NSCLC, it was seen that many factors affect the prognosis and the same tumor stage can give different results in each patient. In this regard, analysis was carried out using the CNN model on 1194 NSCLC patients with a median age of 68.3 who were treated with radiotherapy or surgery. In the study, it has been revealed that deep learning methods perform better in surgical patients compared to current methods [57].

Çevik and Dandıl conducted a study on the classification of lung nodules using CNN. In the study, 23 patients and 1218 CT images were used and 426 regions of interest (ROI) were tagged. The dataset used in the study consists of a total of 214 lung nodule images, 99 of which are benign and 115 are malignant. In the study, 75% of dataset was used as education dataset and 25% was used as test dataset. With the proposed CNN-based computer-assisted detection (CAD) 75% performance was obtained according to the AlexNet architecture result [58].

Kumar et al. proposed a CAD system using autoencoder (AE) that can assist radiologists in the diagnosis of lung cancer. In their study, they used the LIDC dataset containing 1010 patients' CT images to classify lung nodules as benign or malignant. As a result of the study, they obtained 85.25% sensitivity and 75.01% accuracy with the proposed system [59].

Wang et al. used the deep learning method to classify g mediastinal lymph node metastasis of NSCLC. In their study, they used CNN and compared CNN to four machine learning methods, including random forest, adaptive boosting and support vector machines (SVM). These methods were assessed using 1397 lymph node PET/CT scan images collected from 168 patients, and comparison was conducted using 10 times ten fold cross validation. In the study, 81–85% ACC (accuracy) and 0.87–0.92 AUC values were obtained as a result of classical machine methods. With CNN, 84 sensitivity, 88 specificity, 86 ACC and 0.91 AUC values were obtained, while 73 sensitivities, 90 specificities and 82 ACC values were obtained with human doctors. As a result, no significant difference was observed between the best machine learning method and CNN, and these methods were found to have higher sensitivity but lower specificity than physicians [60].

Rosetto and Zhou used CNN in their studies to classify lung cancer CT scan images. In the study, the lung cancer imaging database supplied by the Kaggle Data Science Bowl 2017 was used as a data set. The images in the dataset are used both raw and smoothed with Gaussian filter. With the data labeled positive and negative

for cancer, the training and testing phase of the model was carried out. As a result of the study, 97.5% accuracy was obtained and it was revealed that the percentage of false positives was lower than 10% [61].

Shimizu et al. used SAE for the preliminary diagnosis of lung cancer. In the study, human urine was converted to three-dimentional data (ionic strength, retention time, mass-to-charge ratio) with Gas Chromatography Mass Spectrometer (GC-MS). The neural network used gives two different outputs as a result of calculations: the patient has lung cancer, or not. With this method, 90% accuracy was obtained in determining whether the patient had cancer or not [61].

References

1. T. Tarhan, Kanser Araştırmalarına Yönelik Manyetik O-Karboksimetil Kitosan Nanopartiküllerin Sentezlenmesi, Karakterizasyonu, İrinotekan yüklenmesi ve Glioblastoma Multiforme (Beyin Tümörü) Hücre Hatları Üzerine Sitotok-sisite Değerlendirilmesi (2020)
2. E. Kara, *Validity and Reliability of the Turkish Cancer Loneliness Scale and the Cancer-Associated Negative Social Expectations Scale* (Pamukkale University Institute of Health Sciences, 2019)
3. L.A. Torre, R.L. Siegel, A. Jemal, Lung cancer statistics, in *Lung Cancer and Personalized Medicine*, ed. by A. Ahmad, S. Gadgeel. Advances in Experimental Medicine and Biology, vol. 893 (Springer, Cham, 2016)
4. A. Alberg, J. Samet, Epidemiology of lung cancer. Chest **123**(1), 21S–49S (2003). https://doi.org/10.1378/chest.123.1_suppl.21s
5. R. Herbst, D. Morgensztern, C. Boshoff, The biology and management of non-small cell lung cancer. Nature **553**(7689), 446–454 (2018). https://doi.org/10.1038/nature25183
6. N. Yanaihara, N. Caplen, E. Bowman, M. Seike, K. Kumamoto, M. Yi et al., Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. Cancer Cell **9**(3), 189–198 (2006). https://doi.org/10.1016/j.ccr.2006.01.025
7. A. Aydoğdu, Y. Aydoğdu, Z. Yakıncı, Recognition of basic radiological investigation methods. J. Inonu Univ. Health Serv. Vocat. Sch. **5**(2) (2017)
8. M. Kahraman, *Segmentation and Nodule Detection with Matlab of Lung CT Images* (Yeni Yüzyıl University Faculty of Engineering and Architecture) (2017)
9. Z. Işık, H. Selçuk, S. Albayram, Bilgisayarlı Tomografi ve Radyasyon. Klinik Gelişim **23**, 16–18 (2010)
10. B. Arslan, *Medical Imaging Methods with Computerized Tomography* (Istanbul Technical University, Institude of Science and Technology, 2005)
11. Z. Seyitoğlu, *Changing of Consumer Experience in Digital Public Relations in Turkey: Chatbot Applications* (Istanbul Kültür University, 2019)
12. N. Şimşek, Derin Öğrenme (Deep Learning) Nedir ve Nasıl Çalışır? (2019). Retrieved 26 February 2020, from https://medium.com/@nyilmazsimsek/derin-%C3%B6%C4%9Frenme-deep-learning-nedir-ve-nas%C4%B1l-%C3%A7al%C4%B1%C5%9F%C4%B1r-2d7f5850782
13. S. Çalışkan, S. Yazıcıoğlu, U. Demirci, Z. Kuş, Yapay Sinir Ağlari, Kelime Vektörleri Ve Derin Öğrenme Uygulamalari. Retrieved from http://acikerisim.fsm.edu.tr:8080/xmlui/bitstream/handle/11352/2702/%c3%87al%c4%b1%c5%9fkan.pdf?sequence=1&isAllowed=y
14. O. Inik, E. Ülker, Deep learning and deep learning models used in image analysis. J. Gaziosmanpasa Sci. Res. **6**(3), 85–104 (2017). Retrieved from https://dergipark.org.tr/en/pub/gbad/issue/31228/330663

15. Ç. Uyulan, T. Ergüzel, N. Tarhan, Elektroensefalografi tabanli sinyallerin analizinde derin ogrenme algoritmalarinin kullanilmasi. J. Neurobehav. Sci. 1 (2019). https://doi.org/10.5455/jnbs.1553607558
16. Derin Öğrenme (Deep Learning) Nedir? (2019). Retrieved 26 February 2019, from https://www.beyaz.net/tr/yazilim/makaleler/derin_ogrenme_deep_learning_nedir.html
17. H.İ. Çelenli, Application of paragraph vectors to news and tweet data, in 2018 *26th Signal Processing and Communications Applications Conference (SIU)*, Izmir (2018), pp. 1–4
18. G. Işık, H. Artuner, Recognition of radio signals with deep learning neural networks, in *2016 24th Signal Processing and Communication Application Conference (SIU)*, Zonguldak (2016), pp. 837–840
19. R. Daş, B. Polat, G. Tuna, Derin Öğrenme ile Resim ve Videolarda Nesnelerin Tanınması ve Takibi. Fırat Üniversitesi Mühendislik Bilimleri Dergisi, **31**(2), 571–581
20. G. Işık, *Identification of Turkish Dialects Using Deep Learning Techniques* (Hacettepe University Department of Computer Engineering, 2019)
21. G. Nergız, Y. Safali, E. Avaroğlu, S. Erdoğan, Classification of Turkish news content by deep learning based LSTM using fast text model, in *2019 International Artificial Intelligence and Data Processing Symposium (IDAP)*, Malatya, Turkey (2019), pp. 1–6
22. L. Zhong, L. Hu, H. Zhou, Deep learning based multi-temporal crop classification. Remote Sens. Environ. **221**, 430–443 (2019). https://doi.org/10.1016/j.rse.2018.11.032
23. Q. Zhang, L. Yang, Z. Chen, P. Li, A survey on deep learning for big data. Inform. Fus. **42**, 146–157 (2018). https://doi.org/10.1016/j.inffus.2017.10.006
24. J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai et al., Recent advances in convolutional neural networks. Pattern Recogn. **77**, 354–377 (2018). https://doi.org/10.1016/j.patcog.2017.10.013
25. J. Ou, Y. Li, Vector-kernel convolutional neural networks. Neurocomputing **330**, 253–258 (2019). https://doi.org/10.1016/j.neucom.2018.11.028
26. M. Sajjad, S. Khan, K. Muhammad, W. Wu, A. Ullah, S. Baik, Multi-grade brain tumor classification using deep CNN with extensive data augmentation. J. Comput. Sci. **30**, 174–182 (2019). https://doi.org/10.1016/j.jocs.2018.12.003
27. K. Hanbay, Hyperspectral image classification using convolutional neural network and two dimensional complex Gabor transform. J. Fac. Eng. Archit. Gazi Univ. **35**(1), 443–456 (2020). https://doi.org/10.17341/gazimmfd.479086
28. Evrişimsel Sinir Ağları (2020). Retrieved 26 February 2019, from https://tr.wikipedia.org/wiki/Evri%C5%9Fimsel_Sinir_A%C4%9Flar%C4%B1
29. G. Polat, Y.S. Dogrusöz, U. Halici, Effect of input size on the classification of lung nodules using convolutional neural networks, in *2018 26th Signal Processing and Communications Applications Conference (SIU)*, Izmir (2018), pp. 1–4
30. R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, R. Gao, Deep learning and its applications to machine health monitoring. Mech. Syst. Sign. Process. **115**, 213–237 (2019). https://doi.org/10.1016/j.ymssp.2018.05.050
31. L. Huang, J. Li, H. Hao, X. Li, Micro-seismic event detection and location in underground mines by using convolutional neural networks (CNN) and deep learning. Tunn. Undergr. Space Technol. **81**, 265–276 (2018). https://doi.org/10.1016/j.tust.2018.07.006
32. T. Guo, J. Dong, H. Li, Y. Gao, Simple convolutional neural network on image classification, in *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)*, Beijing (2017), pp. 721–724
33. Q. Zhang, M. Zhang, T. Chen, Z. Sun, Y. Ma, B. Yu, Recent advances in convolutional neural network acceleration. Neurocomputing **323**, 37–51 (2019). https://doi.org/10.1016/j.neucom.2018.09.038
34. F. Beşer, M.A. Kizrak, B. Bolat, T. Yildirim, Recognition of sign language using capsule networks, in *2018 26th Signal Processing and Communications Applications Conference (SIU)*, Izmir (2018), pp. 1–4
35. R. Mukhometzianov, J. Carrillo, CapsNet comparative performance evaluation for image classification (2018). arXiv preprint arXiv:1805.11195

36. A. Körez, N. Barışç, Classification of objects in unmanned aerial vehicle (UAV) images using capsule networks, in *3rd International Symposium on Innovative Approaches in Scientific Studies*. Ankara, Turkey (2019)
37. H. Tampubolon, C. Yang, A. Chan, H. Sutrisno, K. Hua, Optimized CapsNet for traffic jam speed prediction using mobile sensor data under urban swarming transportation. Sensors **19**(23), 5277. https://doi.org/10.3390/s19235277
38. W. Zhang, P. Tang, L. Zhao, Remote sensing image scene classification using CNN-CapsNet. Remote Sens. **11**(5), 494 (2019). https://doi.org/10.3390/rs11050494
39. S.K. Lakshmanaprabu, S.N. Mohanty, K. Shankar, N. Arunkumar, G. Ramirez, Optimal deep learning model for classification of lung cancer on CT images. Fut. Gener. Comput. Syst. **92**, 374–382 (2019). https://doi.org/10.1016/j.future.2018.10.009
40. P. Monkam, S. Qi, H. Ma, W. Gao, Y. Yao, W. Qian, Detection and classification of pulmonary nodules using convolutional neural networks: a survey. IEEE Access **7**, 78075–78091 (2019)
41. S. Armato, G. McLennan, L. Bidaut, M. McNitt-Gray, C. Meyer, A. Reeves et al., The lung image database consortium (LIDC) and image data-base resource initiative (IDRI): a completed reference database of lung nodules on CT scans. Med. Phys. **38**(2), 915–931 (2011). https://doi.org/10.1118/1.3528204
42. A. Setio, A. Traverso, T. de Bel, M. Berens, C. Bogaard, P. Cerello et al., Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge. Med. Image Anal. **42**, 1–13 (2017). https://doi.org/10.1016/j.media.2017.06.015
43. J. Salamon, *Lung Cancer Detection Using Deep Convolutional Networks* (Dublin Institute of Technology, 2018)
44. National Lung Screening Trial Research Team, Reduced lung-cancer mortality with low-dose computed tomographic screening. New England J. Med. **365**(5), 395–409
45. Public Lung Image Database (2020). Retrieved 26 February 2019, from http://www.via.cornell.edu/databases/lungdb.html
46. K. Kuan, M. Ravaut, G. Manek, H. Chen, J. Lin, B. Nazir, V. Chandrasekhar et al., Deep learning for lung cancer detection: tackling the Kaggle data science bowl 2017 challenge (2017). arXiv preprint arXiv:1705.09435
47. M. Khan, S. Rubab, A. Kashif, M. Sharif, N. Muhammad, J. Shah et al., Lungs cancer classification from CT images: an integrated design of contrast based classical features fusion and selection. Pattern Recogn. Lett. **129**, 77–85 (2020). https://doi.org/10.1016/j.patrec.2019.11.014
48. M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, S. Mougiakakou, Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE Trans. Med. Imaging **35**(5), 1207–1216 (2016)
49. U. Bayraktar, Derin Öğrenme Tabanlı Kanserli Hücre Tespiti. Retrieved from https://www.researchgate.net/profile/Umut_Bayraktar2/publication/334151448_Derin_Ogrenme_Tabanli_Kanserli_Hucre_Tespiti/links/5d1a651192851cf4405c8806/Derin-Oegrenme-Tabanli-Kanserli-Huecre-Tespiti.pdf
50. J. Cheng, D. Ni, Y. Chou, J. Qin, C. Tiu, Y. Chang et al., Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans. Sci. Rep. **6**(1) (2016). https://doi.org/10.1038/srep24454
51. W. Sun, B. Zheng, W. Qian, Computer aided lung cancer diagnosis with deep learning algorithms, in *Medical Imaging 2016: Computer-Aided Diagnosis* (2016). https://doi.org/10.1117/12.2216307
52. N. Coudray, P. Ocampo, T. Sakellaropoulos, N. Narula, M. Snuderl, D. Fenyö et al., Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nat. Med. **24**(10), 1559–1567 (2018). https://doi.org/10.1038/s41591-018-0177-5
53. F. Ciompi, K. Chung, S. van Riel, A. Setio, P. Gerke, C. Jacobs et al., Towards automatic pulmonary nodule management in lung cancer screening with deep learning. Sci. Rep. **7**(1) (2017). https://doi.org/10.1038/srep46479

54. Y.-J. Chen, K. Hua, C. Hsu, W. Cheng, S. Hidayati, Computer-aided classification of lung nodules on computed tomography images via deep learning technique. Oncotargets Ther. (2015). https://doi.org/10.2147/ott.s80733
55. D. Ardila, A. Kiraly, S. Bharadwaj, B. Choi, J. Reicher, L. Peng et al., End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nat. Med. **25**(6), 954–961 (2019). https://doi.org/10.1038/s41591-019-0447-x
56. Q. Song, L Zhao, X. Luo, X. Dou, Using deep learning for classification of lung nodules on computed tomography images. J. Healthc. Eng. **2017**, 1–7 (2017). https://doi.org/10.1155/2017/8314740
57. A. Hosny, C. Parmar, T. Coroller, P. Grossmann, R. Zeleznik, A. Kumar et al., Deep learning for lung cancer prognostication: a retrospective multi-cohort radiomics study. PLOS Med. **15**(11), e1002711 (2018). https://doi.org/10.1371/journal.pmed.1002711
58. K. Çevik, E. Dandıl, Classification of lung nodules using convolutional neural networks on CT Images, in *2nd International Turkish World Engineering and Science Congress*. Antalya, Turkey (2019)
59. D. Kumar, A. Wong, D.A. Clausi, Lung nodule classification using deep features in CT images, in *2015 12th Conference on Computer and Robot Vision*, Halifax, NS (2015), pp. 133–138
60. H. Wang, Z. Zhou, Y. Li, Z. Chen, P. Lu, W. Wang et al., Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from 18F-FDG PET/CT images. EJNMMI Res. **7**(1) (2017). https://doi.org/10.1186/s13550-017-0260-9
61. A.M. Rossetto, W. Zhou, Deep learning for categorization of lung cancer CT images, in *2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, Philadelphia, PA (2017), pp. 272–273

Chapter 9
Effective Use of Deep Learning and Image Processing for Cancer Diagnosis

J. Prassanna, Robbi Rahim, K. Bagyalakshmi, R. Manikandan, and Rizwan Patan

Abstract The area of medical image processing obtains its significance with the requirement of precise and effective disease diagnosis over a short period. With manual processing becoming more complicated, stagnant and unfeasible with higher data size, there necessitates automatic processing that can transform contemporary medicine. Deep learning mechanisms can arrive at a higher rate of accuracy in processing and classifying images in comparison with human-level performance. Deep learning not only assist in selecting and extracting features but also possesses the potentiality of measuring predictive target audience and bestows prediction in a more action format to help doctors significantly. Unsupervised Deep Learning for cancer diagnosis is advantageous whenever the involvement of unlabeled data is huge. By bestowing unsupervised deep learning techniques to such unlabeled data, features of pixels that are superior compared to manually obtained features of pixels are said to be learned. Supervised Discriminating Deep Learning directly provides discriminating potentiality for cancer diagnosis purposes. Finally, hybrid deep learning for labeled and unlabeled data is specifically used for cancer diagnosis with a resource

J. Prassanna
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India
e-mail: jprassanna@gmail.com

R. Rahim
Department of Management, Sekolah Tinggi Ilmu Manajemen Sukma, Medan, Indonesia
e-mail: usurobbi85@zoho.com

K. Bagyalakshmi
Department of Electrical and Electronics Engineering, Sri Ranganathar Institute of Engineering and Technology, Coimbatore, India
e-mail: bagiraviram@gmail.com

R. Manikandan (✉)
School of Computing, SASTRA Deemed University, Thanjavur, India
e-mail: srmanimt75@gmail.com

R. Patan
Department of Computer Science and Engineering, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India
e-mail: prizwan5@gmail.com

U. Kose and J. Alzubi (eds.), *Deep Learning for Cancer Diagnosis*,
Studies in Computational Intelligence 908,
https://doi.org/10.1007/978-981-15-6321-8_9

or poor pixel representations and hence early detection and diagnosis performed via bank features. Deep Neural Network, as the name implies includes several layers, emphasizing the complex non-linear relationships between the features present in the images, therefore contributing to higher accuracy. Deep Belief Network used in both supervised and unsupervised deep learning adopting greedy mechanism, maximizing the likelihood nature of detection and diagnosis at an early stage. Sequential event analysis is said to be performed by Recurrent Neural Network with the weights being shared across all neurons, contributing diagnosis accuracy. Certain fine-tuned learning parameters of consideration for better and precise learning are Interaction and Non-linear Rectified Activation function, Circumventing over-fitting via Dropout and Optimal Epoch Batch Normalization. In the last section, challenges about the application of deep learning for cancer diagnosis are discussed.

Keywords Medical image processing · Deep learning · Artificial intelligence · Machine learning · Sparse auto encoders · Convolution neural network

9.1 Deep Learning Techniques for Cancer Diagnosis Using Medical Image Processing

The medical healthcare category is entirely different from other fabrication. As far as the medical healthcare sector is concerned, it is to be of the highest priority and people anticipate the highest care and services irrespective of cost. This is because of the reason that most of the medical data performances are being performed by the medical expert. The second reason is that in terms of execution of images by human expert, due to its characteristics it is found to be finite with high complicated images and due to the substantial dissimilarities across the images being acquired.

With the favorable result of deep learning in other real-world applications, exciting solutions are also said to be achieved with a higher rate of accuracy. In this section, the outline, threats and prospects of deep learning and image processing for cancer diagnosis are discussed.

- Outline
- Threats
- Future Prospects.

9.1.1 Outline

With the enormous advancement in image acquisition sensors, devices, that data is progressing towards big data that makes medical imaging a demanding and fascinating for image analysis and disease diagnosis. This extensive widening in medical images and modalities necessitates considerable and cumbersome endeavors by

doctors. However, it is found to be highly susceptible to human error or fatalities, resulting in variations across several experts.

One of the alternative models is the utilization of machine learning algorithms that in turn automates the process of disease diagnosis, however, the conventional types of machine learning algorithms are not efficient in dealing with the complicated issues related to medical imaging.

Integration of high-level computing performance with machine learning promotes in dealing with big medical imaging process for the accurate and efficient diagnosis, specifically, cancer. With the aid of deep learning not only the selection and extraction of features are performed effectively, but also the newer features are said to be constructed. Besides, not only the disease diagnosis is said to be performed but also measure prediction in a significant manner and assists physicians to take action accordingly.

Hence, specifically, techniques of Machine Learning (ML) and Artificial Intelligence (AI) have contributed in the medical field like medical image processing, computer-aided diagnosis, image interpretation and so on. With these techniques, doctors and physicians are assisted in disease diagnosis at an early stage and also measure for prevention.

9.1.2 Threats

Since the inception of digital imaging, deep learning techniques have been the research topic for disease diagnosis. Most researchers in the field of medical imaging have started thinking that deep learning-based applications will conquer humans and due to this most of the diagnosis will be said to be performed using intelligent machines but will also assists the doctors in prescribing the medicine and guiding in treatment accordingly.

Even though, a large amount of investment and interest have bestowed in the field of medical imaging, deep learning future in medical imaging is not that near when compared to the conventional type of imaging applications due to the complications involved in the medical imaging process. The idea of applying deep learning-based techniques to medical imaging is an interesting and budding research area. However, several threats decelerate their progress [1]. These threats are

- Dataset unavailability
- Privacy and legacy related issues
- Dedicated medical experts
- Nonstandard data machine learning techniques.

9.1.3 Future Prospects

Three movements that manage the deep learning mutiny and resulting in prospects are:

- Big data availability
- Current deep learning techniques designed on the human brain
- Power of the processing machine
- Power of processing power.

Though deep learning prospective advantages are exceedingly noteworthy and so are the inceptive endeavors and prices. Several big tycoons like, Google Deep Mind, research labs in addition to the most significant hospitals are joining hands and working toward the optimal solution of big medical imaging. Similarly, research lab to name a few like, Google, IBM are also putting money into effective imaging application delivery.

9.1.4 Deep Learning Over Machine Learning

Correct disease diagnosis depends upon the images being acquired and the interpretation being made for the corresponding disease. Several image acquisition devices are said to be in existence in the market and with this, with the higher amount of devices, radiological images like, X-Ray, CT and MRI scans etc. are used for obtaining higher resolution images.

- Amongst them, one of the best machine learning applications is computer vision.
- This is because of the reason that conventional machine learning techniques depend on features being crafted utilizing expert analysis.

For example, lung tumor detection necessitates structure features to be extracted. Due to the significant differentiation being observed from patient to patient medical imaging data, conventional learning models are not found to be authentic. Hence, over the last few years, machine learning techniques have progressed by its potentiality to transfer via complicated and big data.

This deep learning mechanism has received a wide range of attention in all the fields and specifically in medical image analysis, i.e., cancer diagnosis. So deep learning mechanism for medical image analysis is considered to be the most efficient and effective supervised machine learning model. This supervised machine learning model utilizes the deep neural network, an extension of Neural Network (NN) utilizing the advance model when compared to the conventional NN.

- The fundamental computational unit in a NN model consists of a neuron.
- It is said to be a notion influenced by the human brain learning that obtains multiple types of signals as input, integrates the multiple inputs with the aid of weights.

- These weights are then linearly integrated with the aid of weight and then pass these integrated signals via nonlinear operations to produce the corresponding output signals.

Some of the key aspects of deep learning over machine learning are:

- Data Dependencies
- Hardware Dependency
- Problem Solving Perspective
- Feature Engineering
- Execution time.

9.1.4.1 Data Dependencies

The most significant distinction between deep learning and conventional machine learning is its execution as the size of data gets bigger. Therefore with larger data size in the image, deep learning algorithms perform well because using the conventional machine learning only handcrafter rules are said to prevail.

9.1.4.2 Hardware Dependencies

Compared to conventional machine learning methods, deep learning largely is influenced by high-end machines. This is because of the reason that deep learning methods encompass Graphics Processing Units (GPUs) that are an essential part of its functioning. Besides as deep learning methods involve a larger amount of matrix multiplication operations, efficient optimization of these operations is said to be done using GPU.

9.1.4.3 Problem Solving Perspective

When solving an issue using conventional machine learning method, it is normally suggested to split the issue down into different portions or segments, individually address the issue and integrate them to obtain the final result. This is in the case of conventional machine learning method and hence the complexity involved is found to be higher, whereas, deep learning method recommends to solve the issue of disease diagnosis for medical imaging, end-to-end.

9.1.4.4 Feature Engineering

Feature engineering involves the process of placing the domain representation into the formation of feature extractors to minimize the data complexity or image complexity and ensure patterns more apparent to learning algorithms to work. This procedure is

found to be cumbersome and involves higher computational costs both in terms of time and overhead.

- In the case of conventional machine learning, the features have to be recognized by a medical expert and then have to be handcrafted as per the requirement of diagnosis and the data type.
- On the other hand, in the case of deep learning, a higher level of features are said to be learned from the data itself.
- Hence, deep learning minimizes the task of developing a new feature for every problem case scenario.

9.1.4.5 Execution Time

During the training of medical images, when applied with a deep learning algorithm, the longer time is said to be consumed. This is due to the reason that a higher amount of parameters are required during the training of deep learning algorithms. In the case of a machine learning algorithm, much lesser time is said to be consumed during the training stage. However, this process is found to be opposite in the testing time.

During testing time, by applying deep learning algorithms, much lesser time is said to be consumed and however, a higher amount of time is said to be consumed during testing when applied with the conventional machine learning algorithm.

9.1.5 Unsupervised Deep Learning for Unlabeled Data

Unsupervised learning [2] is important in scenarios involves an enormous amount of unlabeled data or images. By employing unsupervised deep learning algorithms to such unlabeled data or images, features that are found to be comparatively better to the handwritten label of features is learned. Some of the unsupervised deep learning algorithms for unlabeled data for cancer diagnosis are:

- Auto Encoders
- Sparse Auto Encoders
- Stacked Sparse Auto Encoders.

9.1.5.1 Auto Encoders

An Auto Encoder (AE) for cancer diagnosis is utilized in unsupervised deep learning methods.

- An AE comprises an encoder, followed by a decoder as shown in Fig. 9.1.
- The task of the encoder remains in transforming the input image 'p' to intermediate image 'f' using the auto-encoder function '$f^e(*)$'.

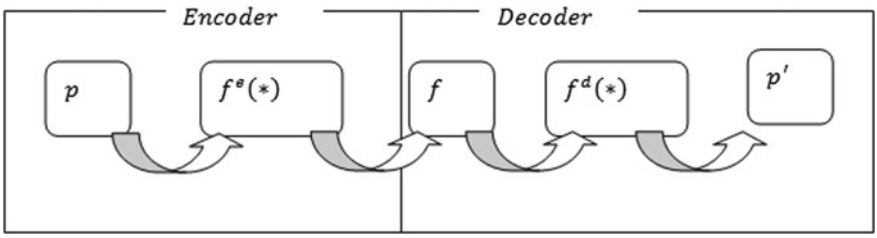


Fig. 9.1 Construction of auto encoder

- On the other hand, the decoder extracts the original image input 'p'' from 'f', forming a coarse-grained restoration of the original input 'p'.

On the other hand, the decoder extracts the original image input 'p'' from 'f', forming a coarse-grained restoration of the original input 'p'.

9.1.5.2 Sparse Auto Encoders (SAE)

A sparse auto-encoder is said to be formed using a sparsity regularizer utilizing a systematic auto-encoder. When almost all of the feature appearances are the exclusion of an only a certain portion of non-zero entries, this type of feature representation is referred to as the sparse representation. These sparse representations on the SAE are highly susceptible to noise and also found to be cost-effective. By including overall samples and sparsity compact encoding on the images, example-specific features for cancer diagnosis can be obtained.

9.1.5.3 Stacked Sparse Auto Encoders (SSAE)

SSAE is said to be formulated by stacking simultaneously multiple layers of fundamental Stacked Sparse Auto Encoders in such a manner that the output from the first layer is provided as the input to the second layer, the output of the second layer provided as input to the third layer and so on. Finally, a classifier is applied to the last output for early cancer diagnosis.

9.1.6 Supervised Discriminative Deep Learning for Labeled Data

Supervised Discriminative Deep Learning is deliberated to straightforwardly furnish discerning capability for cancer diagnosis purpose. It is specifically performed by distinguishing the rear disseminations of categories governed on the visible data.

Selected labeled data are specifically accessible in unambiguous or ambiguous forms for such types of supervised learning.

9.1.6.1 Convolutional Neural Network (CNN)

A CNN is dissimilar from the conventional Back Propagation Neural Network (BPN).This is because of the reason that the BPNN performs specifically on the extracted handcrafted image features. On the other hand, a CNN performs directly on the image provided as input where the necessary features are extracted forming room for efficient image classification and diagnosis. A CNN consists of different layers, namely,

- Convolutional Layers: apply different types of filtering mechanism to input images to detect relevant features of important or region of interest.
  - Hand crafted
  - Automated.
- Pooling Layers: down samples the features detected in the corresponding feature maps once the images are said to be convolved.
  - Minimum pooling
  - Maximum pooling
  - Average pooling.
- Fully Connected Layers: performs classification of images into cancerous or non-cancerous according to the resultant images obtained via pooling.
- Classification Layers.

9.1.7 Hybrid Deep Learning for Labeled and Unlabeled Data

Finally, hybrid deep learning for labeled and unlabeled data is said to be formed in conjunction with both the supervised and unsupervised learning, with both the labeled and unlabeled data.

9.1.8 Reinforced Deep Learning

Reinforcement Deep Learning (RDL) is an extent of Machine Learning (ML). RDL works on the principle of taking relevant action to increase the reward in a specific circumstance. This is said to be achieved by utilizing several software and machines to identify the best probable behavior it should consider for diagnosing cancer and differentiate between cancerous and non-cancerous cells.

- The differentiation between the RDP and the SL is that in case of SL, the training images have the patterns for diagnosis.
- Hence the model is trained with the correct answer.
- In the case of RDL, no patterns are said to be perceived, but the agent present in the RDL model decides upon the action to be taken and hence differentiate between cancer and non-cancerous cell based on the occurrences.

9.2 Labeled and Unlabeled Data Modeling

Labeled data refers to a category of illustrative that has been marked with one or more than one labels.

- The task of labeling the data commonly obtains a set of unlabeled data.
- Next, it makes bigger each piece of that unlabeled portion of that data with significant tags informative.
- On the other hand, unlabeled data modeling comprises of image samples involving natural form that are said to be acquired straightforwardly.

9.2.1 Cross Entropy Auto Encoder

Cross-Entropy [3] is specifically utilized in machine learning as a measure for loss function. It evaluates the difference between two probability distributions (i.e. between the measured value and actual value). With minimum cross-entropy, an Auto Encoder utilized back propagation to produce output value nearly samples as the input values with minimum noise or cross-entropy. Figure 9.2 shows the schematic view of Cross-Entropy Auto Encoder.

The working of cross entropy auto encoder is given below.

- The input is in the form of image including high dimensionality images and processes the neural network producing smaller representation with dual chief elements.
- The first element forms the encoder, representing a group of layer that acquires the input image, compressing it to minimum dimensionality.
- The second element forms the decoder, where the actual reconstruction of input image is said to be obtained with minimum cross entropy.

9.2.2 Sparse Specific Feature Auto Encoder

Sparse Specific Feature Auto Encoder (SSFAE) utilizing a sparsity enforcer enforces a single layer network that in turn evaluates code dictionary, therefore minimizing the

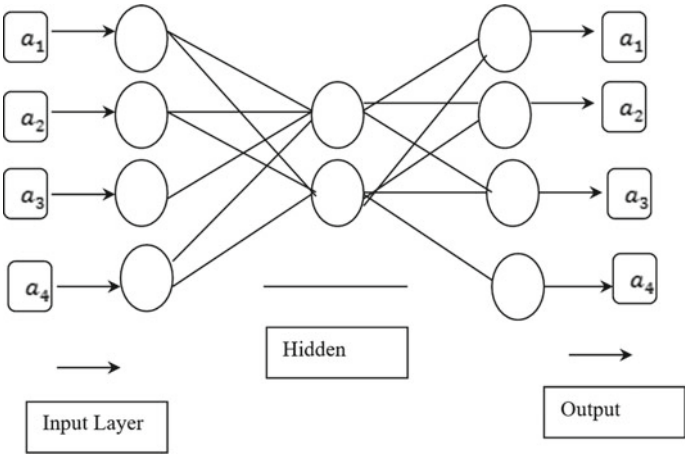


Fig. 9.2 Schematic view of cross entropy auto encoder

error and confining code words for reorganization. Figure 9.3 shows the schematic view of SSFAE.

As depicted in Fig. 9.3, SSFAE comprises of a single hidden layer '*f*', that is connected to the input vector '*a*' via weight factor '*w*' resulting in the encoding section. The hidden layer then outputs to reconstructed the image vector utilizing a weight factor '*w′*' resulting in the decoding section. Here, the decoded image forms either the cancerous or non-cancerous image.

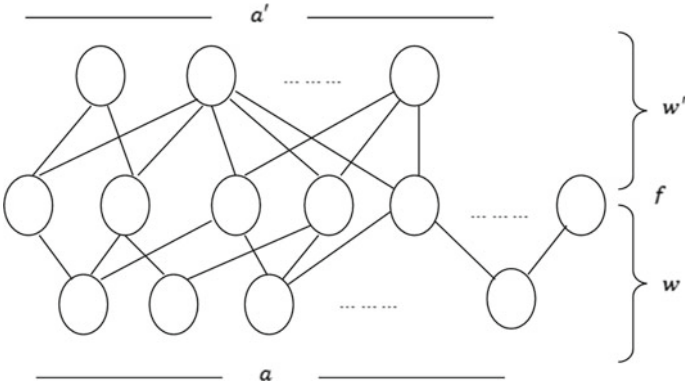


Fig. 9.3 Schematic view of SSFAE

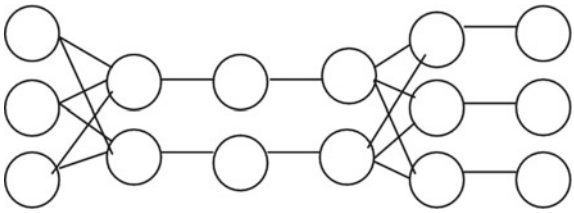


Fig. 9.4 Schematic view of SLSAE

9.2.3 Stacking Layer Sparse Auto Encoder

A Stacking Layer Sparse Auto Encoder (SLSAE) comprises of several layers of sparse auto encoders. Here, the output of each hidden layer is associated with the input of the consecutive hidden layer. The schematic view of SLSAE is given in Fig. 9.4.

As depicted in Fig. 9.4, the SLSAE comprises of three steps. They are:

- Training of auto encoder utilizing input image data and obtain the learned data.
- The learned data from the first layer is used as an input for the second layer, the learned data from the second layer is used as an input for the third layer and this process is said to be continued until the entire training process is said to be accomplished.
- Finally, upon the training of all the hidden nodes in the hidden layers, cost functions are minimized and the corresponding weights are updated to obtain fine-tuned parameters, contributing to accurate disease diagnosis.

9.2.4 Convolutional Neural Network

A Convolutional Neural Network (CNN) is another form of Deep Learning algorithm that obtains an input image, allocates weights and biases to several image objects to differentiate between each objects. One of the main advantages of this method is the minimum amount of pre-processing required when compared to other disease diagnosis classification algorithms.

9.3 Deep Learning Models

Under a moderate acceptance on the activation function, a two-layer neural network, with a limited number of hidden nodes in hidden layers coarse grains a portion of the consecutive function. Hence, deep learning models are also referred to as a ubiquitous estimator. Nevertheless, it is also probable to coarse grain complicated functions to similar accuracy factors utilizing a deep architecture with few neurons overall. In this manner, the weight functions are minimized, hence ensuring the training of images

with a comparatively small-sized image. Therefore, using deep learning models, with the aid of domain expert and engineering skills, good features are said to be discovered in a hierarchically automatic manner.

9.3.1 Deep Neural Network

A Deep Neural Network (DNN) corresponds to a neural network involving a certain amount of complexity, including more than two layers.

- For achieving deeper insight, advanced mathematical modeling is utilized so that processing of data or features are said to be cone in a complicated fashion.
- A neural network synonyms to the process of human brain activity performs the task of pattern recognition, specifically, patterns for classifying into cancerous and non-cancerous cell and the progress of input via different layers of simulated neural associations
- Each layer as mentioned above includes an input layer, an output layer and at-most one hidden layer in the middle.
- With each layer performing particular types of activities, one of the main purposes of using this network is dealing with unlabeled data.

By utilizing the characteristics of artificial intelligence diagnose the risk of cancer by classifying them into cancerous or non-cancerous cell beyond simple input/output postulates.

- Deep Boltzman Machine
- Deep Belief Network
- Recurrent Neural Network.

9.3.2 Deep Boltzman Machine (DBM)

Deep Boltzmann Machine (DBM) [4] is a representation of a DNN produced from several layers of neurons using nonlinear activation functions. With the aid of DBM, the most sophistical and complexity relationships between features are said to be made ensures advanced performance in learning of high-level feature representation when compared to the conventional ANNs (Fig. 9.5).

The advantages of using DBM are:

- First, DBM's have the prospects of learning internal depictions that become progressively complicated and hence considered to be a favorable means of solving object and cancer diagnosis problems.
- Second, high-level depictions are said to be constructed from a considerable size of unlabeled feature inputs and very confined labeled are then said to be utilized to moderately adjusting the model for a fixed task at hand.

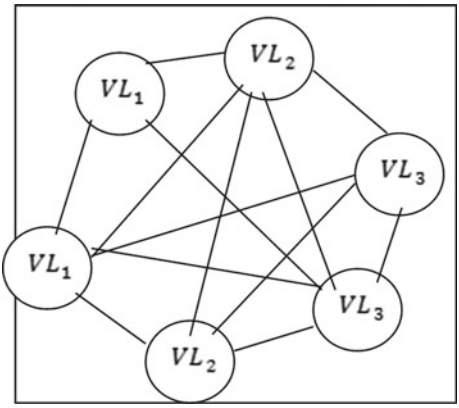


Fig. 9.5 Schematic view of deep Boltzmann machine

- Finally, DBM better disseminates unpredictability and hence deal with more robust and ambiguous input images.

9.3.3 Deep Belief Network

By stacking multiple DBMs results in a single probabilistic model called, Deep Belief Network (DBN). In other words, the DBN comprises one visible layer 'VL' and a series of hidden layers 'HL^1, HL^2, HL^3, . . . , HL^n' as shown in Fig. 9.6.

As depicted in Fig. 9.6, while the top two layers form undirected generative models, whereas the lower layers form directed generative models. The generation structure of DBN is given below.

- Train the first layer.

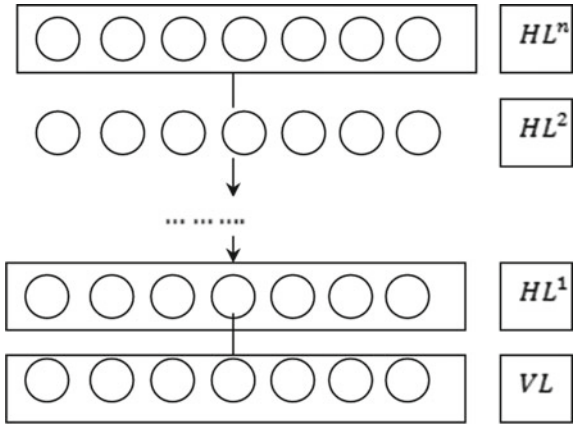


Fig. 9.6 Schematic view of deep belief network

- Make use of the first layer to acquire a depiction of the input image that is said to be utilized as the monitoring for the second layer, i.e., either using the mean activations or samples drawn.
- Train the second layer as DBM, using the transformed samples mean activations as training examples.
- Iterate steps 2 and step 3 for chosen number of layers, each time communicating ascending either samples or mean activations.

9.3.4 Recurrent Neural Network

A Recurrent Neural Network (RNN) [5] is special type of ANN where the connection between nodes results in a directed graph in the form of a temporal sequence, therefore contributing to temporal energetic behavioral patterns. RNN utilizes their internal memory to exercise the input sequence of images. This makes them applicable to cancer diagnosis involving medical imaging traits. Figure 9.7 shows the block diagram of RNN.

As depicted in Fig. 9.7, the RNN refers to two classes of networks, finite impulse and infinite impulse.

- Both classes of networks show temporal dynamic aspects.
- As far as the finite impulse is concerned, it is a directed acyclic and the other infinite impulse is concerned, it is a directed cyclic graph.
- Both finite and infinite impulse RNNs possesses an additional stored state.
- It is said to be under direct control by a NN.

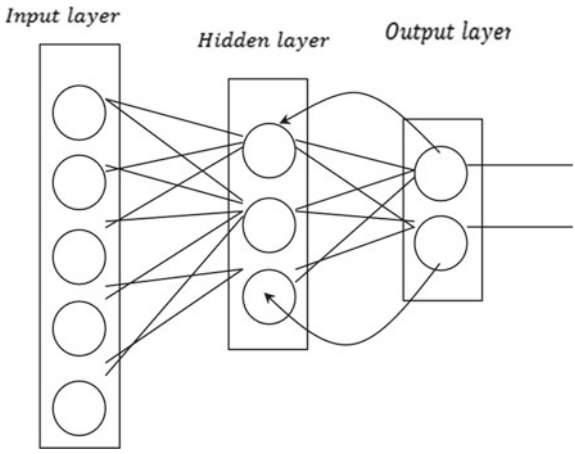


Fig. 9.7 Block diagram of RNN

9.4 Fine-Tuned Learning Parameters

When training of images of any modalities or types is performed with CNN, the foremost task in deep learning tasks is to train the network from scratch. This is because of the reason that an enormous amount of parameters are involved in deep neural networks, usually in the range of millions. Hence, fine-tuning is said to be required whenever involving deep learning.

- First training a CNN on a smaller dataset involving any medical images specifically smaller than the number of required parameters greatly hurts on the CNN, usually resulting in overfitting.
- Therefore, if the dataset of any images if not different from the pre-trained model is trained on, fine-tuning is said to be undertaken.
- A Pre-trained network on a large dataset extracts most unique features like curves, edges, contrast that are said to be highly relevant and useful to most of the disease diagnosis problems.
- In cases of images involving specific domain, medical images, with a cancer diagnosis with the existence of no pre-trained networks, training the network from scratch is said to be accomplished.
- Truncate final layer.
- Apply smaller learning rate.
- Freeze weight on first few layers.

9.4.1 Interaction and Non-Linear Rectified Activation Function

For several computer vision applications, involving medical imaging for disease diagnosis, in recent years deep networks have been used for early diagnosis. The advantage of using DNNs consisting of multiple layers transforming input images into output involves the learning process with higher-level features. For this purpose, deep learning heavily depends on learning different levels of hierarchical data representations [6, 7].

- Due to the involvement of the hierarchical structure, the DNN parameters are said to be tuned to approximate the target functions more effectively than the parameters involved in a shallow format.
- Hence, activation functions are utilized in neural networks (NN) that transforms the weighted sum of the provided input data and biases.
- In this manner, the decision regarding the neuron to be fired or not fired for classification is said to be made.

Some of the customarily utilized activation functions are:

- Sigmoid
- Hyperbolic Tangent (tanh)

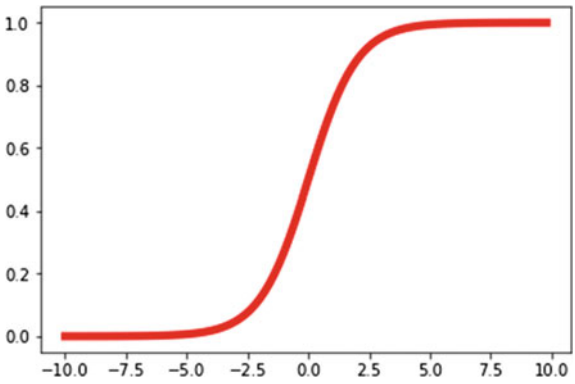


Fig. 9.8 Sigmoid curve

- Rectified Linear Unit (ReLU).

9.4.1.1 Sigmoid Activation Function

In neural networks, the widely used functions are sigmoid functions [8]. With the presence of sigmoid activation function, differentiation between the sigmoid neuron and logistic neuron is made. Different types of perceptions are

- Perceptron outputs discrete 0 or 1 value
- Sigmoid neuron outputs a continuous range of values between 0 and 1

The sigmoid curve given above is in the shape of S. It has a finite limit of '0' as 'x' approaches '$-\infty$' and '1' as 'x' approaches '$+\infty$' respectively (Fig. 9.8).

9.4.1.2 Hyperbolic Tangent

In neural networks, a hyperbolic tangent function is considered as an alternative to sigmoid function. While performing back-propagate, a portion of activation function is involved as derivative that is said to be indulged while measuring error effects on weights. In these cases, derivative of a hyperbolic tangent function is said to be used. Figure 9.9 given illustrates the hyperbolic tangent curve.

As depicted in Fig. 9.9, the hyperbolic tangent function generates output in the form of '$[-1, +1]$' and case of medical imaging, i.e., cancerous or non-cancerous. Besides, it is said to be in the form of continuous function. Therefore, the hyperbolic tangent function in Fig. 9.9 is said to produce output for every feature value for disease diagnosis.

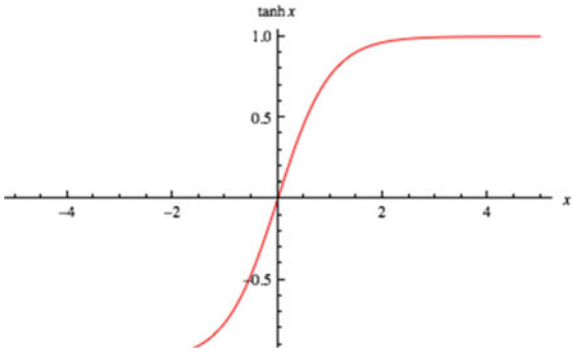


Fig. 9.9 Hyperbolic tangent curve

9.4.1.3 Rectified Linear Unit

The use of ReLU enabled supervised training of state-of-the-art DNNs. Upon comparison to the other conventional activation functions, ReLU is found to be symmetric. Figure 9.10 given illustrates the ReLU.

As depicted in Fig. 9.10, the characteristic of ReLU provides the advantage for the hidden units to be sparse and hence more biologically feasible. This property encourages the hidden units to be sparse, and thus more biologically plausible. Some of the features of ReLU included are given below.

- To utilize stochastic gradient descent with back-propagation of errors for training in deep neural networks, an activation function is found to be highly necessary.
- Though it acts as linear function, non-linearity is said to be ensured using a rectified activation function, therefore ensuring complex relationships.
- As far as the neural network is concerned, the responsibility of activation function remains in transforming the summed weighted input from the node into the activation of the node or output for that input.

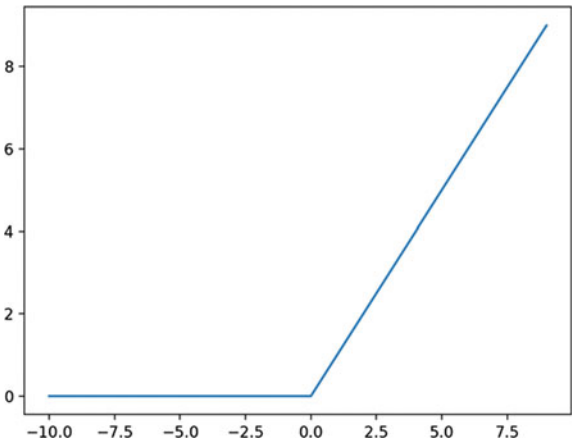


Fig. 9.10 Rectified linear unit activation function

- Besides, the rectified linear activation function is a piecewise linear function, where the rectified linear activation function output the input image in a direct manner if it is positive. On the other hand, it will output zero. With the aid of rectified linear activation, better performance is said to be achieved due to ease in training.

  Some of the domains being in used are

- Facial recognition
- Facial verification
- Medical fraud analysis and detection
- Sentimental analysis and evaluation
- Sound detection
- Speech and visual pattern recognition.

9.4.2 Circumventing Over-Fitting via Dropout

9.4.2.1 Over-Fitting

Whenever training images of any nature is said to be involved in a statistical model, over-fitting is said to occur. Hence, whenever, a model is said to be trained with an enormous amount of information in the images, learning is said to be initiated from the noise and inaccurate feature entries in our data set. Due to this reason, the model does not properly correctly classify the data due to the presence of higher amount of noise.

- The reasons for over fitting [8] are due to the involvement of non-parametric and non-linear methods.
- These types of machine learning methods have more individuality in constructing the model based on dataset of images and hence they are said to be unrealistic models.
- Solutions to address over-fitting are either application of linear algorithms in case of linear data or application of decision trees in case of maximal depth.

  Some of the mechanisms to avoid over-fitting are

- Cross-validation
- Early stopping
- Pruning
- Regularization.

Cross-Validation

One of the most fundamental means of avoiding over-fitting is the application of cross-validation. Here, prediction error is determined based on the out of sample value using fivefold cross validation.

Early Stopping

By applying the principle of early stopping, over fitting is said to be avoided. In other words, the number of iterations required to be run is determined, therefore early stopping is said to be achieved.

Pruning

Whenever related models have said to be constructed, pruning is applied. In other words, with the application of pruning, the nodes or features that possess little prediction power is removed.

Regularization

Here, a cost function is introduced for bringing more features based on the objective function. With this regularization concept, several variables are pushed to zero, therefore minimizing the cost term, contributing to over-fitting.

- Over-fitting is said to be reduced through dropout [9].
- This is performed to minimize the computational overhead so that regularization of deep neural network is said to be attained via dropout.
- This model works on the principle of probabilistic model, where the dropping is said to be performed in a probabilistic manner by either eliminating or discarding or removing certain irrelevant features, not required in the diagnosis stage while performing with medical images.
- It works in such a manner that simulation is conducted on a comparatively larger number of features with very different feature structure and hence making more robust to the input features.

9.4.2.2 Dropout the Regularization Mode

Some of the dropout regularization modes are

- Use with all feature modalities
- Rate of dropout
- Large network usage

- Grid search parameters
- Utilization of weight constraint
- Utilization of smaller dataset.

9.4.3 Optimal Epoch Batch Normalization

9.4.3.1 Batch Normalization

Batch normalization works on the principle of normalizing the input layer by making alterations and activation scaling. For example, consider a scenario where the features for medical images for diagnosing cancer include between 0 and 1 and certain non cancerous cell between 1 and 100 [10].

- To normalize the cancerous and non-cancerous cell, the learning process is said to be initially learnt, for speeding up the process.
- If by making the changes, advantage in the input layer is said to take place, then changes are also said to be made in the hidden layers.
- So training speed are said to be changed in an overall manner by 10 times.

In this manner, employing batch normalization, the amount by which the hidden values is said to be shift around is said to be improved via covariance shift. For example, let us consider a scenario on cancer diagnosis. The training of medical images is said to be performed on only cancerous cell image.

So, when applied this network to medical image on non-cancerous call image, it is obvious, the process is not said to be doing well. The training set and the prediction set are both medical images but they differ a little bit. In other words, if an algorithm learned some a to b mapping and if the distribution of a changes, then the learning algorithm has to be retrained in such a manner to change the distribution of a with the distribution of b.

9.4.3.2 Batch Normalization Working

Figure 9.11 shows the batch normalization block diagram. To minimize the problem of covariate shift as mentioned above, batch normalization is added in between the hidden layers to normalize the function. Here normalization is said to be performed in a separate manner for each neuron and hence it is referred to as 'batch' normalization.

The working of batch normalization is said to be proceeded as follows:

- To improve and enhance the neural network stability and focus, batch normalization performs the process of normalizing the output of a preceding activation layer by deducting the batch mean and obtaining the ratio through batch standard deviation.
- With the shifting operation of activation outputs by certain amount of random parameters, the weights in the succeeding layer are not found to be optimal.

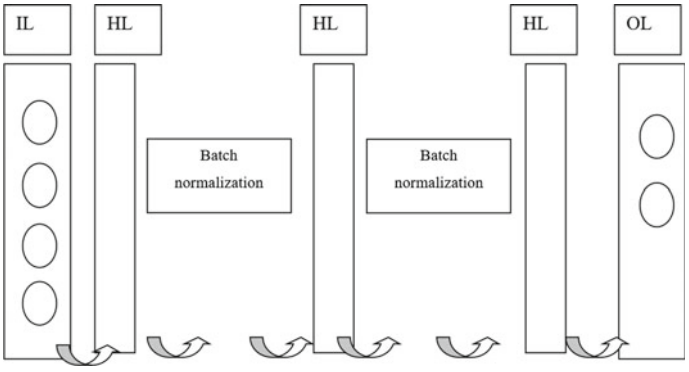


Fig. 9.11 Batch normalization. IL—Input layer, HL—Hidden layer, OL—Output

- To ensure optimality, Stochastic Gradient Descent function is applied to undergo normalization, so that the loss function is said to be minimized.

9.4.3.3 Optimal Epoch Batch Normalization

Improved Deep neural networks (DNN) have resulted in an exponential improvement in the recent few years due to its varied applications in several domains like image processing, object detection, disease diagnosis and so on. However, due to certain levels of problems found in the training stage, thousands or millions of images are said to be processed to obtain fine-tuned images fitting to a specific domain.

In this section, to achieve fine tune learned features, optimal epoch according to the domain area and applications are said to be applied using batch normalization to improve the generalization capability of DNN. By enhancing the statistical parameters involved in disease diagnosis from the input image to the disease diagnosis in all batch normalization using optimal epoch across network, deep adaptation in the field of medical imaging for cancer diagnosis can be attained.

- Batch normalization is specifically said to be implemented as a supplementary layer in NN to both speed up the processing rate and time, therefore contributing for medical imaging, specifically cancer diagnosis.
- For each iteration during the cancer diagnosis, of the training process, the optimal epoch is said to be arrived at by normalizing the batch with its input using both mean and the variance of each cancerous and non-cancerous image to produce output possessing zero mean and unit variance.
- In this case, both the average and dispersion value on the input medical images is probably simulated in such a manner that it should be similar to the average and dispersion over the overall image.
- During the testing stage, the layer performs the actual normalization process utilizing the save average and dispersion, that are running averages and dispersion measured during the training process.

- With optimal epoch, the input at the batch normalization layer is normalized in such a manner that the next layer in the network extracts the input images that are simpler to be trained, therefore improving the training speed.

References

1. M.I. Razzak, S. Naz, A. Zaib, Deep learning for medical image processing: overview, challenges and future, in *Classification in BioApps* (Springer, 2017)
2. D. Selvathi, A.A. Poornila, Deep learning techniques for breast cancer detection using medical image analysis, in *Techniques For Image Processing Applications.* Lecture Notes in Computational Vision and Biomechanics (Springer, 2018)
3. B. Yan, G. Han, Effective feature extraction via stacked sparse auto encoder to improve intrusion detection system. IEEE Access (2018)
4. A. Taherkhani, G. Cosma, T.M. McGinnity, Deep-FS: a feature selection algorithm for Deep Boltzmann Machines. Neuro Comput. (2018)
5. https://en.wikipedia.org/wiki/Recurrent_neural_network
6. G. Castaneda, P. Morris, T.M. Khoshgoftaar, Evaluation of maxout activations in deep learning across several big data domains. J. Big Data (2019)
7. H.H. Aghdam, E.J. Heravi, D. Puig, Recognizing traffic signs using a practical deep neural network, in *Second Iberian Robotics Conference* (Springer, 2016)
8. B.N. Saha, G. Kunapuli, N. Ray, J.A. Maldjian, S. Natarajan, *AR-Boost: Reducing Overfitting by a Robust Data-Driven Regularization Strategy* (Springer, 2013)
9. T.M. Khoshgoftaar, E.B. Allen, Controlling overfitting in classification-tree models of software quality. Empirical Softw. Eng. (2001)
10. J. Bjorck, C. Gomes, B. Selman, K.Q. Weinberger, in *Understanding Batch Normalization*, 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, Canada

Chapter 10
A Deep Learning Architecture for Identification of Breast Cancer on Mammography by Learning Various Representations of Cancerous Mass

Gokhan Altan

Abstract Deep Learning (DL) is a high capable machine learning algorithm with the detailed analysis abilities on images. Although DL models achieve very high classification performances, the applications are trending on using and fine-tuning pre-trained DL models by transfer learning due to the dependence on the number of data, long train time, employments in modeling the most meaningful architecture. In this chapter, we proposed own pruned and simple DL architectures on ROIs extracted from mammography to classify cancer-normal using Convolutional Neural Network (CNN) and Deep Autoencoder (Deep AE) models, which are the most popular DL algorithms. Breast Cancer, which occurs as a result of developing of normal breast tissue to a tumour, is one of the deadliest diseases according to WHO reports. The detection of cancerous mass at early stages is one of the decisive step to start the treatment process. Mammography images are the most effective and simplest way of the diagnosis of breast cancer. Whereas early diagnosis of breast cancer is a hard process due to characteristics of mammography, the computer-assisted diagnosis systems have ability to perform a detailed analysis for a complete assessment. The aim of this study is proposing a robust cancer diagnosis model with a light-weighted DL architecture and comparing the efficiency of the dense layer with the Deep AE kernels against CNN. The ROIs from mammography images were fed into the DL algorithms and the achievements were evaluated. The proposed Deep AE architecture reached the classification performance rates of 95.17%, 96.81%, 93.65%, 93.38%, 96.95%, and 0.972 for overall accuracy, sensitivity, specificity, precision, NPV, and AUROC, respectively.

Keywords Deep learning · ROI · Breast cancer · Deep autoencoders · Mammography · Deep AE · DDSM · ConvNet

G. Altan (✉)
Iskenderun Technical University, Computer Engineering Department, Hatay, Turkey
e-mail: gokhan.altan@iste.edu.tr

U. Kose and J. Alzubi (eds.), *Deep Learning for Cancer Diagnosis*,
Studies in Computational Intelligence 908,
https://doi.org/10.1007/978-981-15-6321-8_10

10.1 Introduction

Breast cancer is the most prevalent and one of the deadliest diseases in the women population all over the world. It causes more than half a million deaths per year [1]. Even though there are many studies on the identification of breast cancer, where cancer conditions increase with a rising momentum, it still continues to be one of the most deadly diseases. Asymptotic phases of breast cancer can be detected directly from mammography images. Mammography images, which use low dose X-ray, are the most effective and primary diagnostic tool to identify and screen the cancerous tissue [2]. The cancerous cell in the tissue is particularly prominent for advanced cancer conditions since its density is high in mammography. Therefore, mammography images still stand out as the primary diagnostic tool in monitoring and diagnosing the disease. Detection of cancerous lesions in the early stages is more difficult to detect in pathology and color densities since mammography is a gray-scale image [3, 4]. Early diagnosis supports a key role in stopping and even keeping the progression of disease under control. Due to this situation, clinicians double-reading, or processes that require detailed analysis of at least two different radiologists, or expensive diagnostic systems are frequently used for early detection. However, this procedure incurs conditions such as additional clinical tests, workload, and the rate of physicians' interest in new patients [5]. These disadvantages can be detected in a short time by using computer-aided diagnosis (CADx) systems, which can detect smaller pathologies that can be missed by even two radiologists, and speed up the patient-specific treatment process.

In the machine learning algorithms used to develop successful CADx models, database and feature extraction take an important role. The homogenization of the distribution in the dataset, preventing overfitting by increasing the amount of data, preparing the train and test sets independently based on the problem are among the parameters to be considered during modeling [5–7]. Conventional CADx systems on mammography should be trained to include the characteristics of cancer lesions and should be determined in the learning process for responsible areas to achieve high discrimination capability [8]. Excluding the redundant regions from the analysis is one of the simplest methods. It is used in optimizing the model, especially the pattern, histogram, hand-crafted and shape-based features of pathological regions [9]. This allows disciplines such as machine learning, image processing to maintain its popularity and to stand out as more specific areas. Whereas most of the CADx systems use hand-crafted features obtained by statistical analysis [2, 10], transformation methods and expert guidance [11, 12], novel CADx systems perform high accurate identification performances using automatized analysis, feature learning and transfer learning. One of these leading and novel methods is Deep Learning (DL), which has proven its effectiveness on the image in many disciplines.

DL is a machine learning algorithm that minimizes the use of image processing and hand-crafted features, which is more distinctive with the multi-layered structures used in supervised learning [13, 14]. The most important advantage of the DL is that it extracts the low-, middle-, and high-level features of the images by

transferring them between adjacent layers and feeds the size-independent characteristics into the supervised classifier models [10, 14]. In this way, the shape, pattern, color, and density information of the images can be analyzed with different feature maps. The basis of advanced image analysis methods is the main approach of the feature learning phase in DL. DL achieves successful results in many fields such as image processing, natural language processing, computer vision, time-series analysis, and provides enhancing the ability of models with superior analysis capability in many different disciplines [15]. Owing to the transfer learning phase in DL models, the weights obtained by training with different datasets. It provides very successful segmentation, regression and classification performances and called pre-trained DL architectures [12]. In addition to the aforementioned disciplines, DL has become one of the most popular methods in many areas including medical image processing. Different variations were obtained by transferring different optimization and unsupervised models to feature learning [10]. The main advantages of DL are detailed avoiding overfitting, analyzing specific feature maps, and using many hidden layers in supervised learning procedures.

There are various studies that applied pre-trained DL algorithms which are based on convolutional neural network (CNN) framework on different mammography databases [4, 6, 8, 15, 16]. They reported well-enough classification and mass detection performances. In particular, in addition to mammography images, histology and histopathology images, which are different cancer imaging techniques, were also analyzed with DL and achieved remarkable segmentation and classification results. However, the pre-trained CNN architectures including AlexNet, DenseNet, VGGNet, GoogleNet, DenseNet, and more have a high number of optimization parameters about 12–132 M [15–17]. This case increases the requirements including proper computational capability and training time. This study aims at proposing a robust cancer diagnosis model with a light-weighted DL architecture and comparing the efficiency of the dense layer with the Deep AE kernels.

10.1.1 Literature Overview

Breast Cancer analysis is a discipline that is frequently studied due to the disease is fatal and has a high prevalence [1, 4]. While mass detection and diagnostic analysis are performed using hand-crafted features [18], especially before the DL approach, an indispensable tool in the segmentation of pathological patterns, mass detection, mass bounding, and classification of the lesions with the flexibility, feature learning capability and adaptability specifications of DL has been highlighted for medical images [9, 19]. Histology and histopathology images are popular as well as mammography; however, they contain very different patterns and color-based characteristics compared to mammography images [1]. Moreover, these diagnostic tools are expensive medical devices. Therefore, mammography images were preferred in this study since they are not applicable to every healthcare institution. While most studies in the literature focus directly on cancer-normal, malignant-benign, mass-calcification,

malignant-benign-normal classification from mammography images [20, 21], some studies have carried out mass detection and localization of the mass pathologies and calculating region of interest (ROI) performance for predicted regions to evaluate the responsibility of the models [22–24].

Pardamean et al. separated the mammography images into normal and cancer classes using transfer learning on ChexNet model weights. They augmented the DDSM database using 5 crop augmentation techniques due to the necessity of a large database on DL. Their model performed various experiments to define the impact of the learning rate, dropout factor and L2 regularization on pre-trained DenseNet architecture with 121 layers. They achieved accuracy rates of 98.02 and 90.38% for training and testing fold [21].

Zeng et al. used a smoothing technique on convolution layers and the ROI pooling layer for mass detection. They applied faster regions with CNN (R-CNN) by combining rectangular mass tissues. They identified the benign and malignant masses on mammography with a sensitivity rate of 93.60% and an average false positive value of 0.651 per image. Their model enabled extracting multi-level features and transferring the specified characteristics for the supervised learning stage [3].

Yoon and Kim extracted hand-crafted features including mass shape, mass margin, calcification type, density, assessment to separate mammography with benign from mammography with malignant. They used 5-fold cross-validation on multiple support vector machines (SVM) with a recursive feature elimination model. They modified the radial basis function kernel and Ada-boost classifier for the SVM for the classification system. They reported an area under the receiver operating characteristic curve value of about 0.9473 for training. They suggested the descriptor features as deterministic mass characteristics [18].

Ertosun and Rubin proposed a regional probabilistic approach for the classification of mammography with and without mass. They also detected the mass area for mammography for the pathological patterns. They performed a data augmentation by cropping, translation, rotation, flipping, and scale-based methods. They utilized adapted pre-trained CNN architectures and reported the success of GoogleNet for mass detection [25].

Suzuki et al. analyzed the mammography images to classify them into mass and normal by a CADx system. They extracted ROI images from the DDSM database. They trained AlexNet architecture using transfer learning on 1656 ROI images with an equal number of images for each class. They tested the model using a separate testing fold (198 images) [26].

Zhu et al. adapted the adversarial model to segment the mass tissue for DL architecture. They applied pixel position-dependent normalization on the DDSM database. They augmented the mammography images by flipping horizontally and vertically. They proposed an adaptive level set segmentation on breast cancer segmentation. They iterated the conditional random field algorithm, adversarial model, and CNN architecture using hybrid methods. They reported the dice index for the mass detection with the rates of 90.21%, 90.97%, and 91.30% for CNN, CNN with conditional random field algorithm, and CNN with Adversarial and conditional random field

algorithm, respectively. They commented on the efficiency of the hybrid methods on the identification of mass tissue using mammography [13].

Xi et al. used class activation maps on pre-trained CNN architectures to classify calcification and mass mammography. They applied 5-fold cross-validation to the training phase of the classifier architectures. They achieved mass detection performance rates of 91.23%, 92.53%, 91.80%, and 91.10% for AlexNet, VGGNet, GoogleNet, and ResNet, respectively. They commented on the feature extraction efficiency of VGGNet on mammography images for localizing pathological tissue by class activation maps [19].

Sarosa et al. cropped ROI images of mass from the DDSM mammography database. They classified malignant and benign masses for the identification of breast cancer. They extracted contrast sharpening and histogram equalization algorithms. They analyzed gray-level co-occurrence matrix features using the SVM classifier. They utilized 3-fold cross-validation on the training of their proposed model and reported an accuracy rate of 63.03% and a specificity rate of 89.01% for binary classification [11].

Nguyen and Lim integrated Gabor filter into the CNN architecture. They applied thresholding and median filtering as the pre-processing of their proposal. They combined two databases (DDSM and MIAS) for defining the robustness of the model. They classified malignant, benign and normal mammography images with a sensitivity rate of 96%. They chose to feed the classifier with ROI images of the abnormalities instead of getting the whole mammography image as input. Even though they achieved a high sensitivity rate for separation of three lesions, using different databases without standardization has possibilities to lead into error in the application and medical use [24].

Touahr et al. have proposed a CADx system to perform automated breast tumor detection. They analyzed malignant and benign tissues using local binary patterns. They applied batch normalization on the CNN architecture. They compared the classification performance of local binary patterns and directly CNN using mammography. They reported CNN/CNN with local binary patterns classification performance rates of 93.24/96.32%, 92.91/96.81%, 91.92/95.83%, and 97/97% for accuracy, sensitivity, specificity, and AUROC, respectively [27].

10.2 Materials and Methods

10.2.1 Deep Autoencoders

Autoencoder (AE) is an unsupervised algorithm to generate different presentations using encoding and decoding procedures. The adaptable dimensional characteristic of AE on encoding procedure makes it a popular module for unsupervised stages of the DL algorithms [28]. AE constructs a set of encoding weights using input data into an intermediate representation and then generates another set of decoding weights to

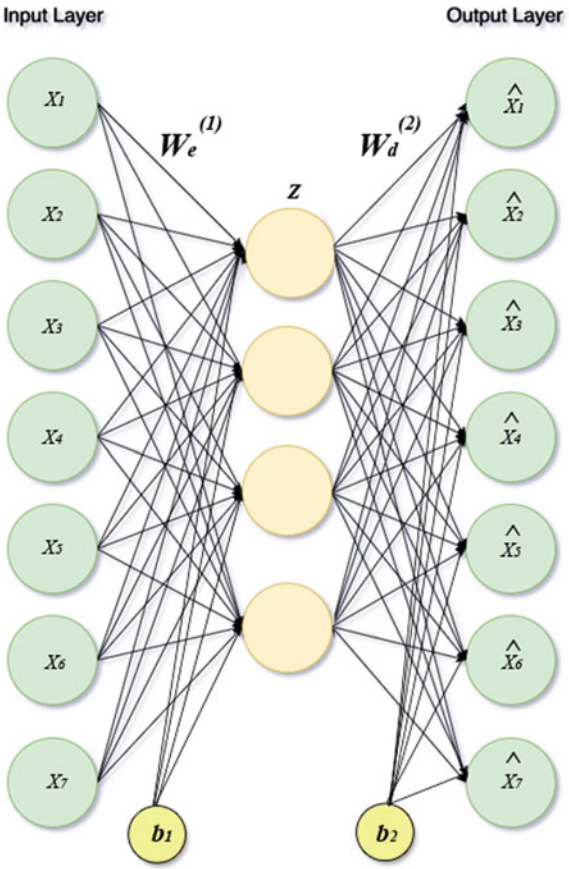


Fig. 10.1 Autoencoder structure

convert intermediate representation into an approximate representation of the input data [7].

Whereas AE models generate different presentations of the input image, it is also a method used for feature dimensionality reduction and data enhancement. They usually produce compressed representation and ensure that responsible and significant features are transferred to the next layers [28, 29]. The structure of AE is depicted in Fig. 10.1.

The encoding process is a mapping function that constructs input $x \in R^{dx}$ into $x \in R^{d^{(1)}}$ using non-linearity. d_x and $d^{(1)}$ are dimensions of the input and encoded representation, respectively.

$$z^{(1)} = h^{(1)}\left(W^{(1)}x + b^{(1)}\right) \tag{10.1}$$

where $h^{(1)}$ represents a non-linear transfer function for the encoding process. $W^{(1)}$ and $b^{(1)}$ represent encode weight matrix and encoder bias, respectively. The decoding

process is a mapping function that reconstructs input using encoded representation.

$$\hat{x} = h^{(2)}\left(W^{(2)}z^{(1)} + b^{(2)}\right) \tag{10.2}$$

where $h^{(2)}$ represents a non-linear transfer function for the decoding process. $W^{(2)}$ and $b^{(2)}$ represent decode weight matrix and decoder bias, respectively. Typically the dimension of under complete AE code is smaller than the input dimension due to provide learning most characteristic information of the training. The learning procedure aims at minimizing the cost function which states the error between input data and reconstructed output ($\hat{x}$) [29].

$$\hat{J}(\psi, b) = \left[\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{2}\|z_i^{(l-1)} - \hat{x_i}^{(l)}\|^2\right)\right] + \frac{\vartheta}{2}\sum_{j=1}^{u_l-1}\sum_{i=1}^{u_l}\left(\psi_{ij}^{l-1}\right)^2 \tag{10.3}$$

where N, ϑ, and u_l represent the number of samples, weight decay parameter, and the number of units in the 1st layer, respectively ($z_i^1 = x_i$). The second term denotes the weight decay for avoiding overfitting. Sparse AE is a simple AE approach that adds a regularization parameter to the cost function to obtain useful representations. The sparsity regularization parameter is a function that denotes the average output activation of a neuron [30]. The cost function of sparse AE is formulated as follows:

$$\hat{J}(\psi, b) = J(\psi, b) + \gamma.\Omega_{sparsity} \tag{10.4}$$

where Ω represents sparsity penalty term and γ defines the responsibility of sparsity term.

Deep AE models are encoded by handling layer by layer (see Fig. 10.2). The Deep AE models have a higher learning capacity than shallow and linear AE models [7]. Compressed representations have the ability to be decoded for many layers using Deep AE approaches [28, 29]. Hereby, the efficiency and robustness of CNN and Deep AE with dense layers were compared on the same supervised learning stages on mammography images with cancer and non-cancer tissues.

10.2.2 Digital Database for Screening Mammography (DDSM)

DDSM, one of the largest mammography databases with online access, was used for the proposed model [31]. DDSM is a wide breast cancer database that has a total number of 2620 cases. The DDSM was gathered and opened for analysis for the researchers within the scope of the project by Massachusetts General Hospital, Wake Forest University School of Medicine, Sacred Heart Hospital and Washington University in St. Louis. Having mammography images from different devices is one of its most important features to analyze various image resolutions and bits.

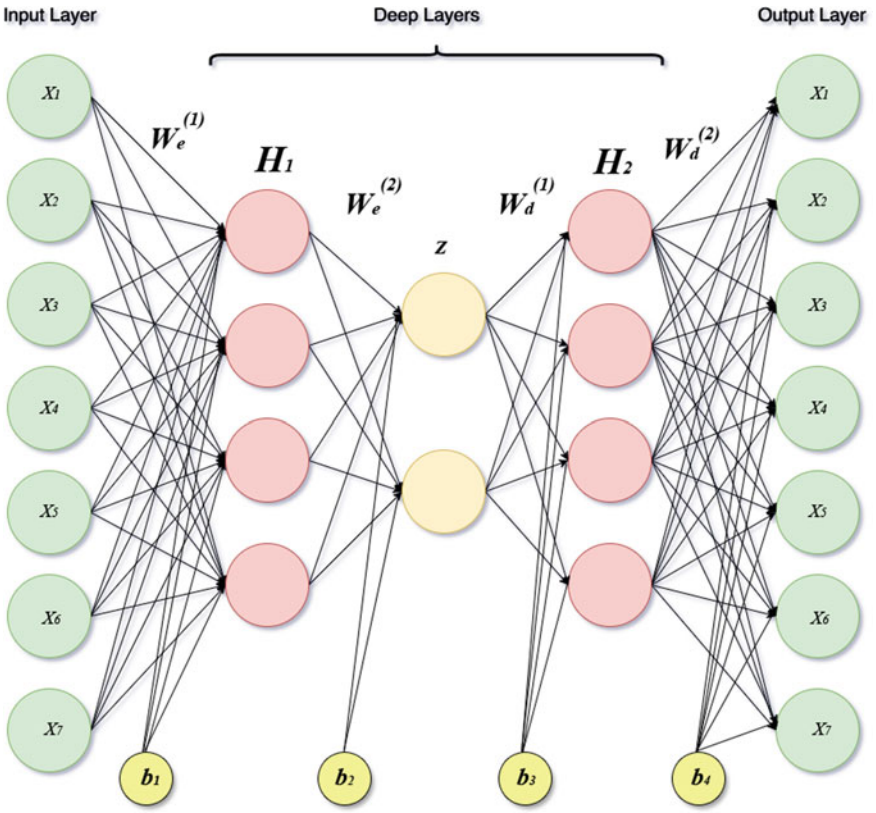


Fig. 10.2 Deep Autoencoder structure with one hidden layer in decoding and encoding

Among them, we excluded the benign case sets. We collected normal and cancer cases from all devices since the proposed Deep AE models need a big number of data and avoiding overfitting. The analyzed mammogram images in this paper comprise the complete set of normal and cancer cases available in the DDSM. Table 10.1 presents the mammography and ROI distribution according to the scanners in DDSM. DDSM also has ROI segmentation, general locations of lesions, and bounding boxes

Table 10.1 The number of normal mammography, cancer mammography an ROI images in the analysis

	Mammography		ROIs	
	Cancer	Normal	Cancer	Normal
DBA	97	430	117	457
Howtek	424	183	435	317
Lumisys	393	82	421	278
Total	914	695	973	1052

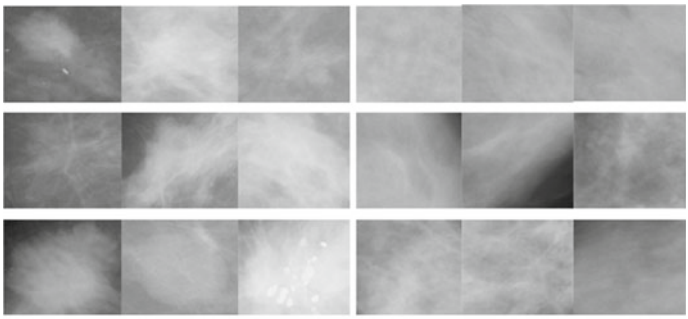


Fig. 10.3 Randomly selected ROIs with cancer (**a**) and normal (**b**)

for cancer mammogram images for the training of the machine learning algorithms. The sample analyzed ROIs are indicated in Fig. 10.3. We augmented the ROIs using cropping and flipping techniques.

10.2.3 Performance Evaluation

Dataset was tested using 5-fold cross-validation. In this algorithm, the dataset is divided into 5 folds with a homogeneous distribution of classes. In other words, the number of samples from different classes in each fold must be equal. One fold is used for testing the trained model, whereas remaining folds are used for the training of the classifier [32]. This process is implemented for each fold and the average of the test results for each fold is used to evaluate the overall performance of the model [7].

The predicted labels obtained in the test results are compared with the actual labels to obtain the confusion matrix. Performance evaluation metrics can be calculated using True Positive (TP), False Positive (FP), True Negative (TN), False Negative (FN), which are the distribution in the confusion matrix [32]. TP is the number of ROIs that are correctly classified as cancer; FN is the number of ROIs that

are actual cancer, are predicted as normal; FP is the number of ROIs that are actually normal, falsely predicted as cancer; TN is the number of ROIs that are correctly classified as normal. Utilizing the output of the proposed Deep AE models, we calculated accuracy, sensitivity, specificity, precision, negative predictive value (NPV), F1 Score, and AUROC. The formulation of the performance evaluation metrics are as follows:

$$Accuracy : \frac{(TP + TN)}{(TP + FP + TN + FN)} \tag{10.5}$$

$$Sensitivity : \frac{TP}{(TP + FN)} \tag{10.6}$$

$$Specificity : \frac{TN}{(FP + TN)} \tag{10.7}$$

$$Precision : \frac{TP}{(TP + FP)} \tag{10.8}$$

$$NPV : \frac{TN}{(FN + TN)} \tag{10.9}$$

$$F1Score : \frac{2 * TP}{(2 * TP + FP + FN)} \tag{10.10}$$

10.3 Experimental Evaluation

In this section, we shared the experimental setup, model architectures, classification parameters and tuned ranges for Deep AE and CNN models. The best results for both classifiers were presented and compared with the existing literature. A series of experiments with a standardized classifier model for both classifiers at a limited variety is carried out on ROI images with cancer lesions and normal tissue from mammography. The efficiency of the representation capabilities for Deep AE and convolution layers are compared with each other.

10.3.1 Experimental Setup

Cancer and normal ROIs segmented using mass locations from mammography images were used for the analysis of breast cancer on DL models. Hereby, experiments were carried out based on two different DL models. In the first proposed DL model, conventional CNN models with fully connected layers were trained using the extracted ROI images for cancer-normal mammography classification. In the sec-

ond DL model, the Deep AE outputs were fed into the fully connected layer. The classification parameters at the supervised stage including the number of the fully connected layer and neuron size were iterated at a specific region for a complete comparison of Deep AE and convolutional layers performances. The same classification parameters including learning rate, dropout parameter, and Adam optimization were set for each model at the supervised training. In the proposed models, the classifier performances on CONV and Deep AE were evaluated for the same fully-connected layers. In the proposed models, the convolution filter size, max pooling, dropout factor, stride, layer sizes, which affect the classifier performance, are increased at certain intervals and their contribution to the generalization performance and achievements were improved. The depth for convolution layers and Deep AE layers is fixed at 3. The number of nodes for each layer in Deep AE models is iterated at the range of 10–250 with a step-size of 10, whereas the filter size for Convolutional layers is set at a range of 3–10 with an increasing 1 by 1. The filter size of the max-pooling layer was set to 2×2 for each layer after convolution (CONV). CONV kernels sifted with a stride length of 1 (non-stride) and no-padding until the entire ROIs were traversed. Rectified Linear Unit (RELU) was utilized for eliminating negative values after each max-pooling. Meanwhile, the sparsity regularization parameter at each AE layer is fixed at 10^{-3}. Sparsity penalty term (γ) variance is set over $\{0.1, 0.75\}$ in the experiments. Fully-connected layers were tuned over two layers. The number of nodes at each fully connected layer was tuned at a range of 50–200 with a step-size of 10 neurons. The dropout factor was set as 0.5 for each layer. The training of fully-connected layers was stopped at the 200th epoch. The output of the fully-connected layer was fed to the softmax classifier as input. The softmax classifier predicts the output as a probability function that defines the level of relation to cancer or normal.

10.3.2 Data Augmentation

Unlike other machine learning approaches, deep learning algorithms have capabilities to train the models that increase generalization performance and classification performance for large-scale databases. Therefore, data augmentation is frequently used in DL-based image analysis. Data augmentation in medical images enables generating additional data using simple transformations that solve the rarity problems of Deep AE training.

Using larger training datasets, DL models have the ability to learn more robust and characteristic information. Moreover, it prevents the classifier models from overfitting, which is the main and compulsive issue for machine learning algorithms. However, it is an important issue to have a homogenous distribution by classes.

The most popular data augmentation techniques including cropping, flipping, rotating, and more are not only simple but also effective for deep learning algorithms. Filtering is not well chosen for data augmentation, because of the convolution layers are already a kind of filtering in images. Considering these characteristics in Deep Learning models, we augmented the ROIs to enhance the performance of the

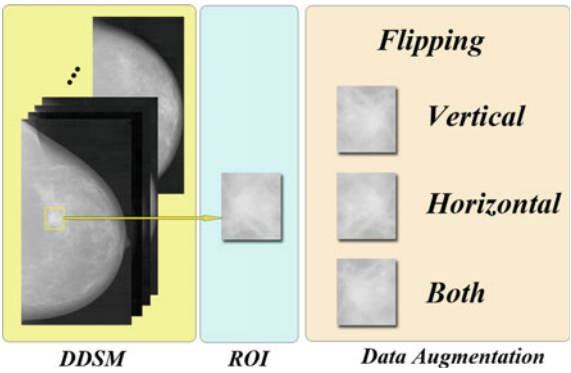


Fig. 10.4 Data augmentation progress steps for ROIs

proposed models using flipping in different directions (vertically, horizontally and both). The ROI mammography dataset was augmented by 4 times (1 original ROI, 3 flipped ones). The details of the data augmentation process on DDSM is depicted in Fig. 10.4. Using this augmentation procedure, we attained 3892 and 4208 ROIs from mammography for cancer lesions and normal tissue, respectively. Each ROI image was resized to 224 × 224 for Deep AE and CNN analysis.

10.3.3 Results and Analysis

In this subsection, the analysis results proposed Deep AE and CNN with fully-connected layers were evaluated considering the various AE layers and convolutional layers. The experimental results and the highest classification performances for both Deep AE and CNN were compared with the existing literature on the identification performance of cancer and normal tissues on mammography images. We want to highlight that original and augmented ROI images were fed into the Deep AE and CNN architectures to classify the ROIs and to evaluate the generating significant representations performances. The best achievements are presented in Tables 10.2 and 10.3 for Deep AE and CNN, respectively. It is evident that the proposed Deep AE with two fully-connected layers achieved better performance on all sparsity penalty terms. However, vying performances were achieved for conventional CNN architectures as well.

Experimental iterations with different Deep AE models have shown that a three-layer Deep AE model has the ability to classify cancer-normal mammography images with high classification performance achievements. Figure 10.5 depicts the classification performance achievements for proposals. Especially, Deep AE models with low sparsity penalty value had higher generalization capacities than a high sparsity penalty value. The best classification performances with low sparsity penalty term ($\gamma = 0.1$) were reached to the rates of 95.17%, 96.81%, 93.65%, 93.38%, and 96.95% for overall accuracy, sensitivity, specificity, precision, and NPV, respectively.

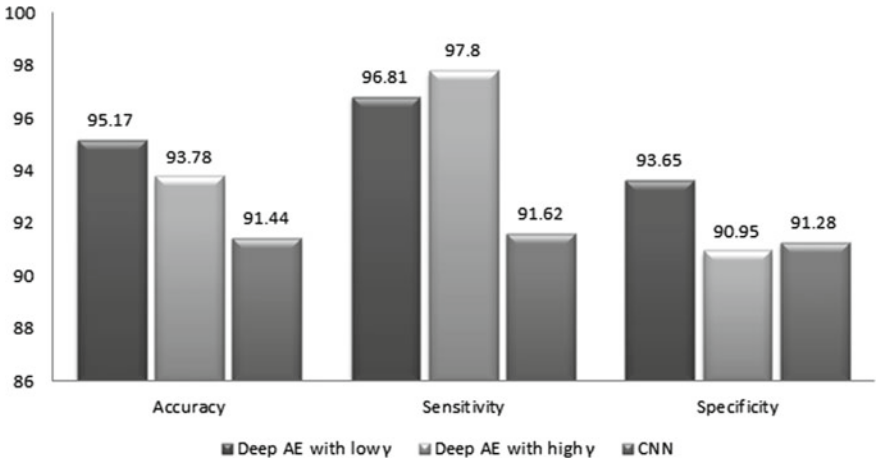


Fig. 10.5 Comparison of proposed Deep Learning architectures considering system performances (%)

The highest classification performance was achieved with a low sparsity penalty term; additionally, successful models were generally achieved for its high value, as well. The highest classification performances with high sparsity penalty term ($\gamma = 0.75$) were reached to the rates of 93.78%, 97.80%, 90.95%, 88.36%, and 98.33% for overall accuracy, sensitivity, specificity, precision, and NPV, respectively. It is experienced that most of the accurate Deep AE models consisted of funnel-shaped compression architecture.

Table 10.2 The best five classification performances (%) for Deep AE models with fully-connected layers

Deep AE model	FC layers	Γ	Acc	Sen	Spe	Pre	NPV	F1 Score
180-120-90	130–210	0.1	95.17	96.81	93.65	93.38	96.95	0.9507
120-100-60	90–120		94.48	94.96	94.04	93.64	95.28	0.9430
200-190-110	70–160		93.74	95.30	92.30	91.97	95.50	0.9360
90-90-130	80–120		93.65	98.00	89.64	89.74	97.97	0.9411
170-140-60	90–70		93.40	97.69	89.42	89.52	97.66	0.9343
110-90-50	160–220	0.75	93.78	97.80	90.95	88.36	98.33	0.9284
150-130-90	130–110		92.62	89.18	95.79	95.15	90.54	0.9207
130-140-130	110–170		92.59	88.67	96.22	95.60	90.18	0.9201
100-40-30	230–40		91.07	97.53	85.10	85.82	97.39	0.9130
210-30-30	100–240		90.68	96.33	85.46	85.97	96.18	0.9085

Acc accuracy, *Sen* sensitivity, *Spe* specificity, *Pre* precision, *FC* fully-connected

Table 10.3 The best five classification performances (%) for CNN models with fully-connected layers

CNN model	FC layers	Acc	Sen	Spe	Pre	NPV	F1 Score
CONV(5) @96	210–220	91.44	91.62	91.28	90.67	92.18	0.9114
CONV(8) @64							
CONV(9) @32							
CONV(5) @96	200–200	90.21	90.65	89.81	89.16	91.20	0.8990
CONV(8) @64							
CONV(9) @32							
CONV(5) @96	190–230	90.02	89.44	90.57	89.76	90.27	0.8960
CONV(6) @64							
CONV(7) @32							
CONV(7) @96	170–180	89.59	89.13	90.02	89.20	89.95	0.8917
CONV(8) @64							
CONV(3) @32							
CONV(3) @96	240–180	89.21	93.47	85.27	85.44	93.39	0.8928
CONV(4) @64							
CONV(3)@32							

Acc accuracy, *Sen* sensitivity, *Spe* specificity, *Pre* precision, *FC* fully-connected

10.4 Discussion

The CADx systems are frequently brought to the fore with various approaches on breast cancer, especially as novel models that provide fast and detailed analysis of all image types. The fact that DL algorithms enable more detailed analysis opportunities for CADx systems is DL has an advantage for the development of these models. Accordingly, low-, middle-, high-level feature extraction capabilities of CNN architectures provide tuning successful diagnostic models in the analysis of mammography images. Besides being a model that has proven itself in many disciplines; using different supervised techniques as a feature learning approach is a highlighted proposal in recent years, as well. In this chapter, we explored the popular DL algorithms including Deep AE and CNN with the same classification parameters to evaluate the efficiency of Deep AE kernels on mammography images to identify cancer tissue against CNN. Moreover, we focused on answering the question that "Do deep AE kernels have the ability to extract significant characteristics from mammography to learn cancer lesions?". To tackle this question, we adopted Deep AE as a feature learning stage to abstract informative patterns from ROIs with cancer and normal tissue and to learn complex mapping relations in ROIs. Afterward, we fed the compressed representations and flattened CONV features to the fully-connected layer-based supervised learning stage with the same classification parameters to evaluate the performances.

As is clearly seen in Table 10.4, most of the studies focused on the analysis of breast cancer using Deep Learning algorithms. Especially, the pre-trained CNN archi-

Table 10.4 The state-of-art for classification of mammography for cancer-normal identification on DDSM

Papers	Methods	Classifier	Acc	Sen	Spe	AUROC
Pardamean et al. [21]	Transfer learning from ChexNet	CNN (DenseNet)	90.38	–	–	–
Ertosun and Rubin [25]	Regional probabilistic approach	CNN (GoogleNet)	85.00	85.00	–	–
Suzuki et al. [26]	Transfer learning from ImageNet	CNN (AlexNet)	85.35	89.90	–	–
Nasir Khan et al. [20]	Multi-View Feature Fusion	CNN (VGGNet)	94.45	98.07	88.13	0.918
Al-antari et al. [9]	First-Higher Order Statistical Features	DBN	92.86	–	–	–
Swiderski et al. [22]	Non-negative matrix factorization	CNN	85.82	82.38	86.59	0.919
Agarwal et al. [23]	Transfer learning from ImageNet	CNN (ResNet-50)	84.00	–	–	0.92
This study	ROI extraction	CNN	91.44	91.62	91.28	0.931
		Deep AE	95.17	96.81	93.65	0.972

Acc accuracy, *Sen* sensitivity, *Spe* specificity, *Pre* precision

tectures are the main and the most popular approaches for cancer detection due to the transfer learning advantage and adaptability of CNN models.

Pardamean et al. transferred the weights of the pre-trained ChexNet model for learning mammography images. They experimented on fully-connected layer size, learning rate, dropout factor and regularization parameters at the supervised stage of their CNN model. They adapted 121 layers of DenseNet architecture and optimization of supervised learning. They achieved an accuracy rate of 90.38% using their proposal [21]. Ertosun and Rubin utilized CNN classifiers with popular pre-trained models. They achieved classification accuracy rates of 84.00%, 82.00%, and 85.00% for AlexNet, VGGNet, and GoogleNet, respectively. They reported the highest sensitivity rate of 85.00% and an average false positive value of 0.9 using GoogleNet architecture for each mammogram for ROI localization [25]. Suzuki et al. proposed a deep neural network structure with the transfer learning capability of ImageNet weights. They used AlexNet architecture and achieved the classification performance rates of 85.35%, 89.90%, and 19.20% for accuracy, sensitivity, and average false positive, respectively [26]. Nasir Khan et al. proposed a multi-view feature fusion model. They used four mammography images for each breast. They focused on different classification problems including cancer-normal, mass-calcification, and malignant-benign. They used bilateral filter and contrast enhancement techniques as a pre-processing stage. They adapted pre-trained CNN architectures including VGGNet, GoogleNet, and ResNet. They reported VGGNet as the best CNN archi-

tecture for distinguishing cancer and normal mammography images by training their own weights. They reached classification performance rates of 94.45%, 98.07%, and 88.13% for accuracy, sensitivity, and specificity, respectively [20]. Al-antari et al. proposed a CADx system using representational learning architecture Deep Belief Networks which is similar to Deep AE on a limited number of mammography images. Their proposal used restricted Boltzmann machines as the supervised stage and attained pre-trained weights for supervised learning. They extracted first- and high-order statistical features including entropy, histogram, asymmetry, skewness, kurtosis, moments from various orders, dissimilarity rates, and more from the ROIs from mammography images with cancer and normal. They fed the statistical feature set into multiple machine learning algorithms such as linear discriminant analysis, quadratic discriminant analysis, SVM, and Deep Belief Networks. They achieved the highest classification performance with an accuracy rate of 92.86% using Deep belief networks [9]. Agarwal et al. proposed a cancer mass detection model using transfer learning from ImageNet weights. They adapted the pre-trained ResNet-50 which is a popular CNN architecture on augmented mass and normal mammography images. They achieved a classification accuracy rate of 84.00% and an AUROC value of 0.92 for ImageNet weights on DDSM. They fine-tuned the ResNet-50 model using the InBreast database by training a novel architecture. They reached an accuracy rate of 92.00% and an AUROC value of 0.98 [23]. Swiderski et al. applied the non-negative matrix factorization method that enables generating a different presentation in the factors form of the Hilbert matrix by a vector. They proposed a novel light-weighted CNN architecture instead of using pre-trained CNN models. They ensured a model with extracting middle- and high-level features from mammography. They tested the CNN model using 10-fold cross-validation. They classified cancer and normal mammography images with classification performance rates of 85.82%, 82.38%, and 86.59% for accuracy, sensitivity, and specificity, respectively. Their training reaches an AUROC value of 0.919 [22].

Whereas the related works focused on using pre-trained CNN models, our proposal analyzed the efficiency of Deep AE as a feature learning stage. The proposed model has also trained a novel CNN architecture with the capability of extracting significant information from mammography images. The achieved results at iterated classification parameter ranges show that Deep AE has a higher and more responsible feature extraction ability against CONV layers on ROI images. The constructing of Deep AE models needs to define the layer sizes and number of neurons at each layer for generating a responsible representation. The proposed Deep AE model classified cancer and normal ROIs with high classification performance rates of 95.17%, 96.81%, 93.65%, 93.38%, and 96.95% for accuracy, sensitivity, specificity, precision, and NPV, respectively. The proposed CNN architecture classified cancer and normal ROIs with the performance rates of 91.44%, 91.62%, 91.28%, 90.67%, and 92.18% for accuracy, sensitivity, specificity, precision, and NPV, respectively. The Deep AE model had a better generating responsible representation capability against the CONV model at experimented classification parameter ranges. The results show that using a compression model with a funnel-shaped Deep AE model has a better generalization capability for mass classification. Using lower sparsity penalty

term value has advantages of carrying significant characteristics for ROIs on mammography images. Using a big number of nodes at fully-connected layers provided a better classification performance for Deep AE compressed representations. This case enables performing more detailed learning procedures for DL algorithms.

Considering the conventional CNN models among the proposed DL models, the highest classification performance was achieved using the CNN architecture that is composed of CONV(5) with 96 filters, CONV(8) with 64 filters, CONV(9) with 32 filters, FC1(210), and FC2(220), respectively. We want to note that each CONV layer is followed by a max-pooling layer with a filter size of 2×2 and a RELU layer. The proposed CNN architecture has separated the ROIs into cancer and normal with classification performance rates of 91.44%, 91.62%, 91.28%, 90.67%, and 92.18% for overall accuracy, sensitivity, specificity, precision, and NPV, respectively.

10.5 Conclusion

We explored the CNN and Deep autoencoder algorithms on mammography considering feature extraction stages; in particular, how deep AE kernels separated ROIs with cancer and normal tissue from mammography images without impairing generalization capability and classification performance of popular DL algorithms. This study demonstrates how the Deep AE algorithm is effective not only using dense layers with many optimization parameters but also using AE kernels on simplified feature learning stages of DL architectures.

There are many existing studies focused on Deep AE kernels [10, 14]. However, these Deep AE models rarely show how they can be integrated into the DL classifiers as feature learning stages. This chapter compared the efficiency of CONV and Deep AE performances on the identification of mass lesions for breast cancer.

This section presented an approach to combine the Deep AE into various machine learning algorithms that perform unsupervised training for Deep AE on generating representations of detailed pattern analysis. The result of proposed simplified Deep AE architecture as feature learning is an adaptable algorithm to improve medical imaging applications, made available to a wider community of academics and practitioners, and, in addition, is easy to perform for even DL models with many hidden layers using Deep AE kernels.

References

1. P.C. Gøtzsche, Screening for breast cancer with mammography. Lancet **358**, 2167–2168 (2001). https://doi.org/10.1016/S0140-6736(01)07198-7
2. A. Jalalian, S.B.T. Mashohor, H.R. Mahmud, M.I.B. Saripan, A.R.B. Ramli, B. Karasfi, Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: a review. Clin. Imaging **37**(3), 420–426 (2013). https://doi.org/10.1016/j.clinimag.2012.09.024

3. Q. Zeng, H. Jiang, L. Ma, Learning multi-level features for breast mass detection, in *ACM International Conference Proceeding Series* (2018). https://doi.org/10.1145/3285996.3286000
4. H.P. Chan, R.K. Samala, L.M. Hadjiiski, C. Zhou, Deep learning in medical image analysis. Adv. Exp. Med. Biol. (2020). https://doi.org/10.1007/978-3-030-33128-3
5. J. Shiraishi, Q. Li, D. Appelbaum, K. Doi, Computer-aided diagnosis and artificial intelligence in clinical imaging. Semin. Nucl. Med. (2011). https://doi.org/10.1053/j.semnuclmed.2011.06.004
6. D. Abdelhafiz, C. Yang, R. Ammar, S. Nabavi, Deep convolutional neural networks for mammography: Advances, challenges and applications. BMC Bioinformatics (2019). https://doi.org/10.1186/s12859-019-2823-4
7. G. Altan, Y. Kutlu, Generative Autoencoder Kernels on Deep Learning for Brain Activity Analysis. Natural and Engineering Sciences (2018). https://doi.org/10.28978/nesciences.468978
8. K. Mendel, H. Li, D. Sheth, M. Giger, Transfer learning from convolutional neural networks for computer-aided diagnosis: a comparison of digital breast tomosynthesis and full-field digital mammography. Acad. Radiol. **26**(6), 735–743 (2019). https://doi.org/10.1016/j.acra.2018.06.019
9. M.A. Al-antari, M.A. Al-masni, S.U. Park, J.H. Park, M.K. Metwally, Y.M. Kadah, S.M. Han, T.S. Kim, An automatic computer-aided diagnosis system for breast cancer in digital mammograms via deep belief network. J. Med. Biol. Eng. (2018). https://doi.org/10.1007/s40846-017-0321-6
10. G. Altan, Y. Kutlu, A.Ö. Pekmezci, A. Yayık, Diagnosis of chronic obstructive pulmonary disease using deep extreme learning machines with LU autoencoder kernel, in *7th International Conference on Advanced Technologies (ICAT'18)* (2018)
11. S.J.A. Sarosa, F. Utaminingrum, F.A. Bachtiar, Mammogram breast cancer classification using gray-level co-occurrence matrix and support vector machine, in *3rd International Conference on Sustainable Information Engineering and Technology, SIET 2018—Proceedings* (2018). https://doi.org/10.1109/SIET.2018.8693146
12. R. Wang, Y. Ma, W. Sun, Y. Guo, W. Wang, Y. Qi, X. Gong, Multi-level nested pyramid network for mass segmentation in mammograms. Neurocomputing **363**, 313–320 (2019). https://doi.org/10.1016/j.neucom.2019.06.045
13. W. Zhu, X. Xiang, T.D. Tran, G.D. Hager, X. Xie, Adversarial deep structured nets for mass segmentation from mammograms, in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* (2018), pp. 847–850. https://doi.org/10.1109/ISBI.2018.8363704
14. G. Altan, Y. Kutlu, Hessenberg Elm autoencoder kernel for deep learning. J. Eng. Technol. Appl. Sci. **3**(2), 141–151 (2018). https://doi.org/10.30931/jetas.450252
15. J. Arevalo, A. Cruz-Roa, F.A. González, *Hybrid Image Representation Learning Model with Invariant Features for Basal Cell Carcinoma Detection*, ed. by J. Brieva, B. Escalante-Ramírez (2013), pp. 89220M-6). https://doi.org/10.1117/12.2035530
16. J. Arevalo, F.A. Gonzalez, R. Ramos-Pollan, J.L. Oliveira, M.A.G. Lopez et al., DeepMammo breast mass classification using deep convolutional neural networks. Comput. Methods Progr. Biomed. (2018). https://doi.org/10.1016/j.acra.2018.06.019
17. S. Pan, J. Zhang, T. Wang, L. Kong, X-ray mammary image segmentation based on convolutional neural network, in *2019 IEEE 4th International Conference on Image, Vision and Computing (ICIVC)* (2019), pp. 105–108. https://doi.org/10.1109/ICIVC47709.2019.8981350
18. S. Yoon, S. Kim, AdaBoost-based multiple SVM-RFE for classification of mammograms in DDSM. BMC Med. Inform. Decis. Mak. (2009). https://doi.org/10.1186/1472-6947-9-S1-S1
19. P. Xi, C. Shu, R. Goubran, Abnormality detection in mammography using deep convolutional neural networks, in *IEEE International Symposium on Medical Measurements and Applications (MeMeA)* (2018), pp. 1–6. https://doi.org/10.1109/MeMeA.2018.8438639
20. H. Nasir Khan, A.R. Shahid, B. Raza, A.H. Dar, H. Alquhayz, Multi-view feature fusion based four views model for mammogram classification using convolutional neural network. IEEE Access (2019). https://doi.org/10.1109/ACCESS.2019.2953318
21. B. Pardamean, T.W. Cenggoro, R. Rahutomo, A. Budiarto, E.K. Karuppiah, Transfer learning from chest X-ray pre-trained convolutional neural network for learning mammogram data. Procedia Comput. Sci. **135**, 400–407 (2018). https://doi.org/10.1016/j.procs.2018.08.190

22. B. Swiderski, J. Kurek, S. Osowski, M. Kruk, W. Barhoumi, Deep learning and non-negative matrix factorization in recognition of mammograms, in *Eighth International Conference on Graphic and Image Processing (ICGIP 2016)* (2017). https://doi.org/10.1117/12.2266335
23. R. Agarwal, O. Diaz, R. Marti, X. Llado, Mass detection in mammograms using pre-trained deep learning models, in *14th International Workshop on Breast Imaging (IWBI 2018)*, ed. by E.A. Krupinski (SPIE, 2018), p. 12. https://doi.org/10.1117/12.2317681
24. V.D. Nguyen, K. Lim, M.D. Le, N. Dung Bui, Combination of Gabor filter and convolutional neural network for suspicious mass classification, in *2018 22nd International Computer Science and Engineering Conference (ICSEC)* (2018), pp. 1–4. https://doi.org/10.1109/ICSEC.2018.8712796
25. M.G. Ertosun, D.L. Rubin, Probabilistic visual search for masses within mammography images using deep learning, in *2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (2015), pp. 1310–1315. https://doi.org/10.1109/BIBM.2015.7359868
26. S. Suzuki, X. Zhang, N. Homma, K. Ichiji, N. Sugita, Y. Kawasumi, T. Ishibashi, M. Yoshizawa, Mass detection using deep convolutional neural network for mammographic computer-aided diagnosis, in *2016 55th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)* (2016), pp. 1382–1386. https://doi.org/10.1109/SICE.2016.7749265
27. R. Touahri, N. AzizI, N.E. Hammami, M. Aldwairi, F. Benaida, Automated breast tumor diagnosis using local binary patterns (LBP) based on deep learning classification, in *International Conference on Computer and Information Sciences (ICCIS)* (2019), pp. 1–5. https://doi.org/10.1109/ICCISci.2019.8716428
28. P. Vincent, H. Larochelle, Y. Bengio, P.-A. Manzagol, Extracting and composing robust features with denoising autoencoders (2008). https://doi.org/10.1145/1390156.1390294
29. J. Wang, X. Zhang, Q. Gao, H. Yue, H. Wang, Device-free wireless localization and activity recognition: a deep learning approach. IEEE Trans. Vehic. Technol. **66**(7), 6258–6267 (2017). https://doi.org/10.1109/TVT.2016.2635161
30. B.A. Olshausen, D.J. Field, Sparse coding with an overcomplete basis set: a strategy employed by V1? Vis. Res. (1997). https://doi.org/10.1016/S0042-6989(97)00169-7
31. M. Heath, K. Bowyer, D. Kopans, R. Moore, W. Philip Kegelmeyer, The digital database for screening mammography, in *The Fifth International Workshop on Digital Mammography*, ed. by M.J. Yaffe (Medical Physics Publishing, 2001). ISBN 1-930524-00-5
32. R. Duda, P. Hart, D. Stork, *Patterns Classification* (Wiley, New York, 2012). ISBN: 9781118586006

Chapter 11
Deep Learning for Brain Tumor Segmentation

Khushboo Munir, Fabrizio Frezza, and Antonello Rizzi

Abstract Brain tumors are considered to be one of the most lethal types of tumor. Accurate segmentation of brain MRI is an important task for the analysis of neurological diseases. The mortality rate of brain tumors is increasing according to World Health Organization. Detection at early stages of brain tumors can increase the expectation of the patients' survival. Concerning artificial intelligence approaches for clinical diagnosis of brain tumors, there is an increasing interest in segmentation approaches based on deep learning because of its ability of self-learning over large amounts of data. Deep learning is nowadays a very promising approach to develop effective solution for clinical diagnosis. This chapter provides at first some basic concepts and techniques behind brain tumor segmentation. Then the imaging techniques used for brain tumor visualization are described. Later on, the dataset and segmentation methods are discussed.

Keywords Deep learning · Convolutional neural network · Brain tumor segmentation · Artificial intelligence

11.1 Introduction

Machine learning (ML) [1, 2] is a fundamental approach for designing intelligent systems and it is considered a basic ingredient in artificial consciousness [3]. Artificial intelligence systems, mostly based on ML, are nowadays applied in many fields such as agriculture [4, 5], computer vision [6, 7], medical diagnosis [8, 9] and

K. Munir (✉) · F. Frezza · A. Rizzi
Department of Information Engineering, Electronics and Telecommunications (DIET), Sapienza University of Rome, Rome, Italy
e-mail: khushboo.munir@uniroma1.it

F. Frezza
e-mail: fabrizio.frezza@uniroma1.it

A. Rizzi
e-mail: antonello.rizzi@uniroma1.it

U. Kose and J. Alzubi (eds.), *Deep Learning for Cancer Diagnosis*,
Studies in Computational Intelligence 908,
https://doi.org/10.1007/978-981-15-6321-8_11

scene analysis [5, 10]. Neural networks based advancements [11, 12] introduced deep learning (DL) into the market which nowadays is considered to be the most powerful approach. It can discover deep feature representations meaningful for a particular application, difficult to be found by the specialist. The methodologies used to determine those features are known as automatic strategies of feature acquisition [13, 14]. Deep learning has been widely used in the field of medicine and has proven to be the best among all other clinical methods.

Among all others, brain is the most important part of the body, whereas brain tumor is caused by a disordered proliferation of cells. Central Nervous System tumors can develop either in the spinal cord or in the brain [15], while the prime symptoms are headaches, neural coordination issues, mood swings, and memory loss. Brain tumors are categorized based on their growth origin, nature, and progression rate [16, 17]. There are benign tumors and malignant tumors. Benign tumors have low progression rate, they don't contain cancerous cells and are non-invasive, whereas the malignant tumors have a high progression rate, they consist of cancerous cells and are invasive in nature [18]. Brain Tumors are classified into four grades by brain tumor associations [19] in which grade I and II are benign, whereas III and IV are malignant. If the benign tumor is not treated carefully it can turn itself into malignant. In the invasive case, incision is done for collecting tumor samples which, is also known as biopsy. Non-invasive diagnostic approaches makes use of images of different types, for example, positron emission tomography (PET), computed tomography (CT) and magnetic resonance imaging (MRI) [20, 21]. MRI is useful in providing human cerebrum information and is more useful because of its non-ionizing radiation and non-intrusive nature.

Deep learning is used for investigating various types of brain tumors. The convolutional neural network is the most used deep learning structure outperforming other structures, especially for cerebrum tumor segmentation and categorization. The idea is to convolve the images through learned Finite Impulse Response (FIR) filters, organized in many layers in order to achieve a chain of feature activations. Basic steps of computer vision system design, at training stage, are image acquisition, preprocessing, segmentation, feature extraction, feature reduction, and model synthesis. Feature extraction has the purpose of computing a set of meaningful characteristics, while feature reduction (selection) is in charge of selecting the most related subset of characteristics, given the problem at hand. This chapter basically provides the recent trends for segmenting brain tumors digital images, used as basic features for classification by using deep-learning techniques.

11.2 Brain Image Compression Techniques

Image compression is basically a process of reducing the size in bytes of images deprived of demeaning the superiority and quality of the image to an objectionable level. The decrease in image size permits further images to be stored in a specified quantity of disk or memory space. In order to represent a part of the human body in

digital form, CT images are used. There exists a variety of work that has been done in dealing with the medical imaging process. Image compression is another way of handling such images. A huge amount of work has been done in this prospect. After knowing what exactly image compression is, we will now overview various methods and procedures presented in this prospect.

11.2.1 Computed Tomography (CT)

CT images are the combination of X-rays processed images taken from different angles of the brain to produce images centered on the degree of alteration. CT process basically consists of the subject of examination which is being controlled. In Table 11.1, some of the methods of brain compression using CT techniques are given.

11.2.2 Electroencephalography (EEG)

In EEG method electrodes are attached to the scalp in order to record the electrical signals and activities of the brain. EEG enables us to analyze distinct brain zones by providing with the strength, power, and position of electrical deeds. Table 11.2 discusses some of the works using EEG compression.

Table 11.1 Comparison of compression techniques of CT brain images

CT compression	CT compression pros	CT compression cons	CT compression results
[22] (Rhodes et al., 1985)	Heavy image compression	Expensive in terms of computation	5% Compression achieved
[23] (Lee et al., 1991)	3D image compression in addition to 2D data	Expensive in terms of computation and system stops compressing in the worst case	Efficient and effective method
[24] (Hashimoto et al., 2004)	Obtained image quality is adequate	Expensive in respect of computation	Improved eminence and pixel quality
[25] (Li et al., 2006)	Image quality is preserved	Only for limited types of images	A portion of choice is compressed
[26] (Signoroni et al., 2009)	Coding efficient with increased reliability	The system is slow	Effective result
[27] (Ju and Seghouane, 2009)	No effect on the image quality	Expensive in terms of computation	96% sensitivity for polyps achieved

Table 11.2 Comparison of compression techniques of EEG brain images

EEG compression	EEG compression pros	EEG compression cons	EEG compression results
[28] (Cinkler et al., 1997)	Accurate and fast	Limitation of memory	Acceptable compression
[29] (Memon et al., 1999)	Error rate less than 1%	Expensive in terms of computation	achieved a high compression rate
[30] (Aviyente and Selin 2007)	small data used	affects image quality	Acceptable compression rate
[31] (Higgins et al., 2010b)	Power is minimized	Expensive in terms of computation	The method is better than other EEG compression methods

11.2.3 Magnetic Resonance Imaging (MRI)

MRI basically measures the brain activity by using magnetic grounds and radio waves hence producing enriched superior signals in 3D or 2D spatial dimensions. Some of the works using MRI compression are given in Table 11.3 and 3D MRI are given in Table 11.4.

Table 11.3 Comparison of compression techniques of MRI brain images

MRI compression	MRI compression pros	MRI compression cons	MRI compression results
[32] (Cavaro-Menard et al., 1999)	Can be used for segmentation application	Compression ratio is limited	The achievable compression ratio is 67:1
[33] (Raghavan et al., 2002)	If even area are in MRI than substantial compression is achievable	Increase in image storage	Mask image compression achieved to large value with large even areas sum
[34] (Badawy et al., 2002)	Quality of image is preserved	Large systems and complex systems	It provides an acceptable compression ratio
[35] (Gornale et al., 2007)	Handles variation in images, its intensity and frequency	Compromised results for higher frequencies	The ratio of compression is highly dependent on the image
[36] (Karras et al., 2009)	Independent of parameters defined by user	Obstruction possession removal for divided borders	In addition to image particular protection method acquired effective outcomes

Table 11.4 Comparison of compression techniques of 3D-MRI brain images

3D-MRI compression	3D-MRI compression pros	3D-MRI compression cons	3D-MRI compression results
[37] (Yodchanan et al., 2008)	Decreases the complexity of computation	Only skull region of brain MRI can be handled	40% of enactment by dent is achieved
[38] (Corvetto et al., 2010)	3D images are handled	–	–
[39] (Dhouib et al., 2009)	3D images are handled	High storage requirements for images	An effective method with a large storage requirements

11.3 Image Segmentation Using Multi-models

Analysis of the medical images becomes a challenging problem because of variation of the shape, size and target issue location. Although many different algorithms are proposed for the segmentation, still their comparison is not possible due to the fact that their evaluation was done on different types of datasets with different metrics. So to face this issue public health challenges were created, where the same data are used in order to perform a fair and useful comparison, such as Ischemic Stroke Lesion Segmentation (ISLES), Brain tumor Segmentation (BraTS) [40], MR Brain Image Segmentation (MRBrainS) [41], Combined (CT-MR) Healthy Abdominal Organ Segmentation (CHAOS), Neonatal Brain Segmentation (NeoBrainS) [42], Automatic intervertebral disc localization and 6-month infant brain MRI Segmentation (Iseg-2017) [43] and segmentation from 3D Multi-modality MR (M3) Images (IVDM3Seg).

Table 11.5 describes the detailed dataset information mentioned above.

Table 11.5 Summary of multi-model image segmentation dataset

Data-set provider	Task	Training	Testing	Images size	Modalities
BraTS (2012)	Brain tumor	35	15	$160 \times 216 \times 176$ $176 \times 176 \times 216$	T1, T1C, T2, flair
BraTS (2013)	Brain tumor	35	25	$160 \times 216 \times 176$ $176 \times 176 \times 216$	T1, T1C, T2, flair
BraTS (2014)	Brain tumor	200	38	$160 \times 216 \times 176$ $176 \times 176 \times 216$	T1, T1C, T2, flair
BraTS (2015)	Brain tumor	200	53	$240 \times 240 \times 155$	T1, T1C, T2, flair
BraTS (2016)	Brain tumor	200	191	$240 \times 240 \times 155$	T1, T1C, T2, flair
BraTS (2017)	Brain tumor	285	146	$240 \times 240 \times 155$	T1, T1C, T2, flair
BraTS (2018)	Brain tumor	285	191	$240 \times 240 \times 155$	T1, T1C, T2, flair
BraTS (2019)	Brain tumor	285	191	$240 \times 240 \times 155$	T1, T1C, T2, flair

11.4 DataSet

Before the implementation of any algorithm designed for brain cancer diagnosis problem, the first step is to analyze dataset at hand. MICCAI BraTS provides free access to a large amount of MRI volumes for the research of brain tumors. These MRI volumes consist of four modalities, T1-weighted, T2-weighted, T1CE-weighted and FLAIR. Each modality is explained briefly in this section.

Tl-weighted is basically used for differentiating among healthy tissues from diseased ones and these scans provide gray and white matter contrast (is shown in Fig. 11.1).

T2-weighted are well suited for brain diseases in which water accumulates inside the tissues of the brain because of the sensitivity of this modality to water content. This modality delineates the edema region hence resulting in the production of a bright signal on the image. A colorless fluid present in the spinal cord and brain known as cerebrospinal fluid (CSF) can be successfully differentiated by using T1-, while T2-weighted images. This CSF looks bright in T2-weighted modality images and in T1-weighted images it looks dark (Fig. 11.2).

T1-weighted MRI with gadolinium contrast enhancement (T1-Gd) is the fourth type of MRI sequence (Fig. 11.3). In this modality accumulated contrast agent such as gadolinium ions are used in the active cell region of the tumor tissues to create a bright signal which enables to distinguish between the border of the tumor. As the necrotic cells don't interact with the used contrast agent, they are differentiated as a hypointense part of the tumor core, therefore easing the segmentation of hypointense part from active cell region. Fluid Attenuated Inversion Recovery (FLAIR) except for its acquisition protocol is similar to T2-weighted images (Fig. 11.4). Here suppression of water molecule is carried out, which in turn helps to differentiate between edema from Cerebrospinal Fluid (CSF). FLAIR can suppress the water signals and because of which the hyperintense periventricular lesion is visible clearly.

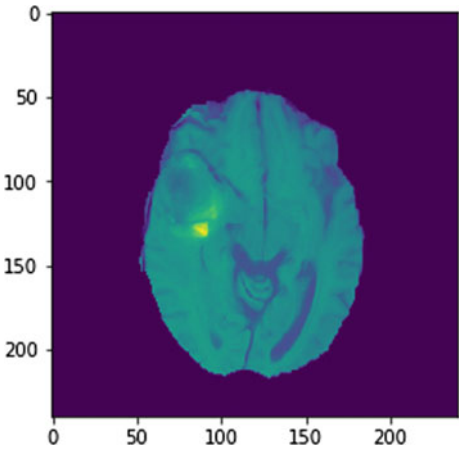


Fig. 11.1 MRI T1-weighted volume used for brain tumors

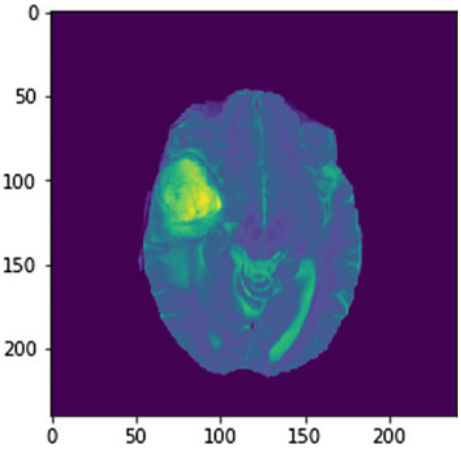


Fig. 11.2 MRI T1CE-weighted volume used for brain tumors

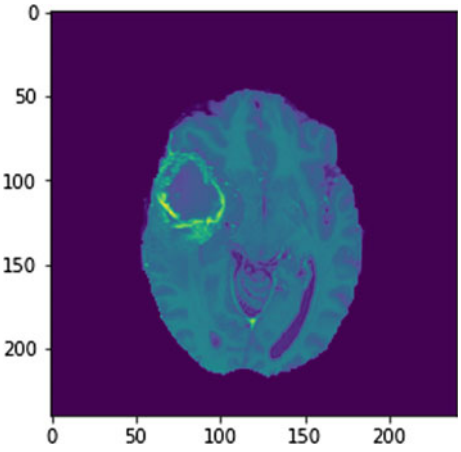


Fig. 11.3 MRI flair volume used for brain tumors

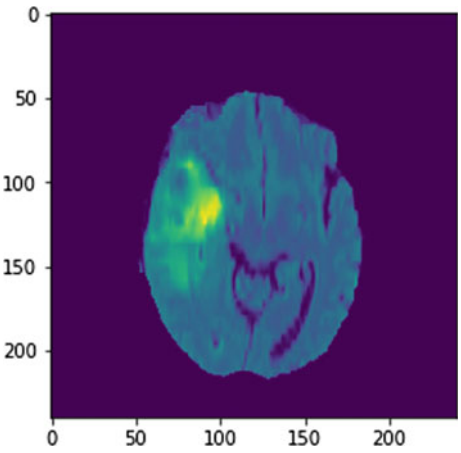


Fig. 11.4 MRI T2-weighted volume used for brain tumors

11.5 Data Processing

Data preprocessing is a fundamental step, consisting in selection of dimensions of the data, pre-processing of the images, and data-augmentation. These steps form the basis of deep learning segmentation.

11.5.1 Dimension of Data

There are 3D images used for the segmentation of medical images. There are 2D-models which breaks these 3D images into slices for the model training [44–47]. 3D-models are created which directly makes use of these images for the training of the model [48–50]. 3D Models inputs 3D images with 3D convolution kernels in order to extract spatial features of the images. The 3D approach is computationally expensive. A solution to this computational problem would be to break the 3D images into small patches in order to train the model. One of the examples of these 3D patch training approach is the algorithm proposed by Kamnitsas et al. [51] who randomly extracted 10 k 3D patches to train its model for brain tumor segmentation. On the other hand, there is a 2D approach in which 3D images are converted into 2D slices and by using 2D convolutional kernel spatial information is extracted. This approach helps in reducing the computational cost but it loses the spatial information of z-direction. Mlynarski et al. [52] extracted 2D features using a CNN-based model from the images of axial, coronal and sagittal views, feeding a 3D-CNN based model. This method achieved the median dice score for whole tumor 0.918, 0.883 for the tumor core and 0.854 for enhancing core. This method can learn features directly in 3D representations.

11.5.2 Data Pre-processing

Images with multiple models shows variations in contrast, intensity and noise in the images; therefore pre-processing has a vital role in order to make the images smooth for training. Among the most popularly used pre-processing techniques are the intensity normalization, bias field correction and image registration. BraTS dataset undergoes through the image registration process before it is made public [53, 54]. MRI data distortion is corrected by applying the N4ITK method [45, 51, 55]. Images are then normalized by subtracting mean and dividing by standard deviation of the brain region.

11.5.3 Data-Augmentation

A labeled data with a large number of patterns for training is not practically available because this labelling requires a large amount of time of experts. Conversely, if the model is trained with limited data then overfitting may occur [56]. In order to increase the data set and reduce overfitting, a technique of data-augmentation is introduced. This technique applies operations such as rotation, scaling, translation, flipping, addition of Gaussian noise and distortion addition over the dataset images and creates new images. These newly created images along with the original dataset images are fed into the neural network. Overfitting problem was reduced by application of data augmentation techniques such as random scaling, random rotation, gamma correction, random elastic deformations and mirroring on the dataset was used by Isensee et al. [48].

11.6 Segmentation of Tumor

The most famous type of brain tumors are gliomas [55]. Patients with low-grade tumors have the survival chance of many years as they are less destructive in nature. On the contrary high-grade tumors decrease the survival expectancy of the patient to two years [57]. Classification of gliomas and glioblastomas is really difficult as they are combined with edema with poor contrast and undefined structures. Segmentation of tumor means to localize the tumor region. There are three types of tumor tissues dynamic, necrotic and edema.

11.7 Deep Learning Strategies for Brain Tumor Segmentation and Classification

There are various forms of segmentation blocks used for the segmentation of brain tumors nowadays, such as Convolutional neural networks (CNNs), Deep convolutional neural networks (DCNNs), Auto Encoders (AEs), Recurrent Neural Networks (RNNs), Generative Adversarial Networks (GANs), Deep Neural Networks (DNNs) and Long Short Term Memory (LTSM). Deep learning methodologies are able to determine automatically useful features for the final classification stage. The general strategy is to pass an image through the trained pipeline of deep learning building blocks and input image segmentation is performed depending on the deep features.

Classification basically depends on the classifier's methodology. In general, at first image features are extracted and fed to the classifier. Most of the classification research focuses on CNNs. Many researchers used fuzzy c-mean and discrete wavelet transform (DWT) for reduced feature extraction, which are then passed to a deep neural network for prediction. Features obtained from the CNN network are

fused with handcrafted features for tumor classification by Saba et al. [58]. Capsule networks are used by Afshar et al. [59] for brain tumor classification. These Capsule networks are the newly formed AI structure introduced to remove the weaknesses of CNNs. Nie et al. [60] used a 3D CNN for feature extraction and the support vector is trained on these features for the classification of tumors. Xu et al. [61] used AlexNet for deep feature extraction for further use of tumor classification. Multiclass predictions using CNN for tumor classification has been studied by Sultan et al. [62].

11.8 Discussion and Conclusion

In clinics the general practice for brain tumor analysis is performed by the medical experts, which is a challenging task as the brain images are varying and complex. Therefore these manual methods are tedious whereas the automated segmentation methods make this task to be performed with ease. Computer-aided systems have given a positive contribution in the medical field of diagnostics because of its fast speed and accuracy. Deep learning provides automatic feature acquisition which is time-efficient as compared to manual methods. In addition Graphics Processing Units (GPUs) makes it even fast. Training data is another factor to increase the performance of the system. There are some drawbacks too in addition to the advantages of DL and GPUs which causes limitations of the DL approaches such as the cost of GPU. This review helps the readers to evaluate what recent DL algorithms have been developed for the analysis of brain tumors, pointing out further research problems in existing DL approaches.

Quite a bunch of work has already been performed in recent years on brain tumor MRI segmentation and classification using the deep learning methods but the complexity of MRI images makes it a challenging technical area that provides a big amount of room to researchers for further study in this area. Classification and segmentation both provide medical experts to refer themselves for a second opinion which is purely based on the results provided by the automated results. At the same time, there is an issue of robustness in accuracy which makes this domain suffer.

References

1. J. Amin, M. Sharif, M. Raza, M. Yasmin, Detection of brain tumor based on features fusion and machine learning. J. Ambient Intell. Hum. Comput. 1–17 (2018)
2. J. Amin, M. Sharif, M. Raza, T. Saba, M.A. Anjum, Brain tumor detection using statistical and machine learning method. Comput. Methods Progr. Biomed. **177**, 69–79 (2019)
3. J. Liu, Y. Pan, M. Li, L. Ziyue Chen, C.L. Tang, J. Wang, Applications of deep learning to MRI images: a survey. Big Data Min. Anal. **1**(1), 1–18 (2018)

4. M. Sharif, M.A. Khan, Z. Iqbal, M.F. Azam, M.I. Ullah Lali, M.Y. Javed, Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection. Comput. Electron. Agric. **150**, 220–234 (2018)
5. R.J. Martis, V.P. Gurupur, H. Lin, A. Islam, S.L. Fernandes, Recent advances in big data analytics. Internet Things Mach. Learn. (2018)
6. S.M. Naqi, M. Sharif, M. Yasmin, Multistage segmentation model and SVM-ensemble for precise lung nodule detection. Int. J. Comput. Assist. Radiol. Surg. **13**(7), 1083–1095 (2018)
7. M. Sharif, M.A. Khan, M. Faisal, M. Yasmin, S.L. Fernandes, A framework for offline signature verification system: best features selection approach. Pattern Recognit. Lett. (2018)
8. S. Masood, M. Sharif, A. Masood, M. Yasmin, M. Raza, A survey on medical image segmentation. Curr. Med. Imaging **11**(1), 3–14 (2015)
9. S. Masood, M. Sharif, M. Yasmin, M. Raza, S. Mohsin, Brain image compression: a brief survey. Res. J. Appl. Sci. Eng. Technol. **5**(1), 49–59 (2013)
10. G.J. Ansari, J.H. Shah, M. Yasmin, M. Sharif, S.L. Fernandes, A novel machine learning approach for scene text extraction. Future Gener. Comput. Syst. **87**, 328–340 (2018)
11. M. Yasmin, M. Sharif, Sajjad Mohsin, Neural networks in medical imaging applications: a survey. World Appl. Sci. J. **22**(1), 85–96 (2013)
12. M.A. Khan, T. Akram, M. Sharif, M.Y. Javed, N. Muhammad, M. Yasmin, An implementation of optimized framework for action classification using multilayers neural network on selected fused features. Pattern Anal. Appl. **22**(4), 1377–1397 (2019)
13. M. Fayyaz, M. Yasmin, M. Sharif, J.H. Shah, M. Raza, T. Iqbal, Person re-identification with features-based clustering and deep features. Neural Comput. Appl. 1–22 (2019)
14. M. Sharif, U. Tanvir, E.U. Munir, M.A. Khan, M. Yasmin, Brain tumor segmentation and classification by improved binomial thresholding and multi-features selection. J. Ambient Intell. Hum. Comput. 1–20 (2018)
15. Brain tumor basics
16. American Cancer Society website
17. Brain tumor diagnosis
18. V. Rajinikanth, S.C. Satapathy, S.L. Fernandes, S. Nachiappan, Entropy based segmentation of tumor from brain MR images—a study with teaching learning based optimization. Pattern Recognit. Lett. **94**, 87–95 (2017)
19. L.M. DeAngelis, Brain tumors. N. Engl. J. Med. **344**(2), 114–123 (2001)
20. U.R. Acharya, S.L. Fernandes, J. En WeiKoh, E.J. Ciaccio, M.K. Mohd Fabell, U.J. Tanik, V. Rajinikanth, C.H. Yeong, Automated detection of Alzheimer' disease using brain MRI images—a study with various feature extraction techniques. J. Med. Syst. **43**(9), 302 (2019)
21. S.L. Fernandes, U.J. Tanik, V. Rajinikanth, K.A. Karthik, A reliable framework for accurate brain image examination and treatment planning based on early diagnosis support for clinicians. Neural Comput. Appl. 1–12 (2019)
22. M.L. Rhodes, J.F. Quinn, J. Silvester, Locally optimal run-length compression applied to CT images. IEEE Trans. Med. Imaging **4**(2), 84–90 (1985)
23. H. Lee, Y. Kim, A.H. Rowberg, E.A. Riskin, 3-d image compression for X-ray CT images using displacement estimation, in *1991 Proceedings. Data Compression Conference* (IEEE, 1991), p. 453
24. M. Hashimoto, K. Matsuo, A. Koike, H. Hayashi, T. Shimono, Ct image compression with level of interest, in *2004 International Conference on Image Processing, 2004. ICIP'04*, vol. 5 (IEEE, 2004), pp. 3185–3188
25. L. Guoli, Z. Jian, W. Qunjing, H. Cungang, D. Na, L, Jianping, Application of region selective embedded zerotree wavelet coder in ct image compression, in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference* (IEEE, 2006), pp. 6591–6594
26. A. Signoroni, S. Masneri, A. Riccardi, I. Castiglioni, Enabling solutions for an efficient compression of pet-ct datasets, in *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)* (IEEE, 2009), pp. 2747–2751
27. J.L. Ong, A.-K. Seghouane, False positive reduction in ct colonography using spectral compression and curvature tensor smoothing of surface geometry, in *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (IEEE, 2009), pp. 89–92

28. J. Cinkler, X. Kong, N. Memon, Lossless and near-lossless compression of EEG signals, in *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems and Computers (Cat. No. 97CB36136)*, vol. 2 (IEEE, 1997), pp. 1432–1436
29. N. Memon, X. Kong, J. Cinkler, Context-based lossless and near-lossless compression of EEG signals. IEEE Trans. Inf. Technol. Biomed. **3**(3), 231–238 (1999)
30. S. Aviyente, Compressed sensing framework for EEG compression, in *2007 IEEE/SP 14th Workshop on Statistical Signal Processing* (IEEE, 2007), pp. 181–184
31. G. Higgins, S. Faul, R.P. McEvoy, B. McGinley, M. Glavin, W.P. Marnane, E. Jones, EEG compression using jpeg2000: how much loss is too much?, in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology* (IEEE, 2010), pp. 614–617
32. C. Cavaro-Ménard, A. Le Duff, P. Balzer, B. Denizot, O. Morel, P. Jallet, J.-J. Le Jeune, Quality assessment of compressed cardiac MRI. Effect of lossy compression on computerized physiological parameters, in *Proceedings 10th International Conference on Image Analysis and Processing* (IEEE, 1999), pp. 1034–1037
33. S. Raghavan, S. Chatterjee, M.B. Waldron, Image compression applied to MRI images, in *Images of the Twenty-First Century. Proceedings of the Annual International Engineering in Medicine and Biology Society* (IEEE, 1989), pp. 526–527
34. W. Badawy, M. Weeks, G. Zhang, M. Talley, M.A. Bayoumi, MRI data compression using a 3-d discrete wavelet transform. IEEE Eng. Med. Biol. Mag. **21**(4), 95–103 (2002)
35. S.S. Gornale, V.T. Humbe, S.S. Jambhorkar, P. Yannawar, R.R. Manza, K.V. Kale, Multi-resolution system for MRI (magnetic resonance imaging) image compression: a heterogeneous wavelet filters bank approach, in *Computer Graphics, Imaging and Visualisation (CGIV 2007)* (IEEE, 2007), pp. 495–500
36. D.A. Karras, Compression of MRI images using the discrete wavelet transform and improved parameter free Bayesian restoration techniques, in *2009 IEEE International Workshop on Imaging Systems and Techniques* (IEEE, 2009), pp. 173–178
37. W. Yodchanan, Lossless compression for 3-D MRI data using reversible KLT, in *2008 International Conference on Audio, Language and Image Processing* (IEEE, 2008), pp. 1560–1564
38. A. Corvetto, A. Ruedin, D. Acevedo, Robust detection and lossless compression of the foreground in magnetic resonance images, in *2010 Data Compression Conference* (IEEE, 2010), pp. 529–529
39. D. Dhouib, A. Nait-Ali, C. Olivier, M.S. Naceur, Comparison of wavelet based coders applied to 3D brain tumor MRI images, in *2009 6th International Multi-Conference on Systems, Signals and Devices* (IEEE, 2009), pp. 1–6
40. B.H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, et al., The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans. Med. Imaging **34**(10), 1993–2024 (2014)
41. A.M. Mendrik, K.L. Vincken, H.J. Kuijf, M. Breeuwer, W.H. Bouvy, J. De Bresser, A. Alansary, M. De Bruijne, A. Carass, A. El-Baz, et al., MRBrainS challenge: online evaluation framework for brain image segmentation in 3T MRI scans. Comput. Intell. Neurosci. **2015** (2015)
42. I. Išgum, M.J.N.L. Benders, B. Avants, M.J. Cardoso, S.J. Counsell, E.F. Gomez, L. Gui, P.S. Hűppi, K.J. Kersbergen, A. Makropoulos, et al., Evaluation of automatic neonatal brain segmentation algorithms: the NeoBrainS12 challenge. Med. Image Anal. **20**(1), 135–151 (2015)
43. L. Wang, D. Nie, G. Li, É. Puybareau, J. Dolz, Q. Zhang, F. Wang, J. Xia, W. Zhengwang, Jia-Wei Chen et al., Benchmark on automatic six-month-old infant brain segmentation algorithms: the iSeg-2017 challenge. IEEE Trans. Med. Imaging **38**(9), 2219–2230 (2019)
44. S. Pereira, A. Pinto, V. Alves, C.A. Silva, Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Trans. Med. Imaging **35**(5), 1240–1251 (2016)
45. S. Cui, L. Mao, J. Jiang, C. Liu, S. Xiong, Automatic semantic segmentation of brain gliomas from MRI images using a deep cascaded neural network. J. Healthc. Eng. **2018** (2018)
46. G. Wang, W. Li, S. Ourselin, T. Vercauteren, Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks, in *International MICCAI Brainlesion Eorkshop* (Springer, 2017), pp. 178–190

47. D. Nie, L. Wang, Y. Gao, D. Shen, Fully convolutional networks for multi-modality isointense infant brain image segmentation, in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)* (IEEE, 2016), pp. 1342–1345
48. F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, K.H. Maier-Hein, Brain tumor segmentation and radiomics survival prediction: contribution to the brats 2017 challenge, in *International MICCAI Brainlesion Workshop* (Springer, 2017), pp. 287–297
49. Y. Qin, K. Kamnitsas, S. Ancha, J. Nanavati, G. Cottrell, A. Criminisi, A. Nori, Autofocus layer for semantic segmentation, in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer, 2018), pp. 603–611
50. J. Dolz, K. Gopinath, J. Yuan, H. Lombaert, C. Desrosiers, I.B. Ayed, Hyperdense-net: a hyper-densely connected CNN for multi-modal image segmentation. IEEE Trans. Med. Imaging **38**(5), 1116–1126 (2018)
51. K. Kamnitsas, C. Ledig, V.F.J. Newcombe, J.P. Simpson, A.D. Kane, D.K. Menon, D. Rueckert, B. Glocker, Efficient multi-scale 3d CNN with fully connected CRF for accurate brain lesion segmentation. Med. Image Anal. **36**, 61–78 (2017)
52. P. Mlynarski, H. Delingette, A. Criminisi, N. Ayache, 3d convolutional neural networks for tumor segmentation using long-range 2d context. Comput. Med. Imaging Graph. **73**, 60–72 (2019)
53. X. Zhao, W. Yihong, G. Song, Z. Li, Y. Zhang, Y. Fan, A deep learning model integrating FCNNs and CRFs for brain tumor segmentation. Med. Image Anal. **43**, 98–111 (2018)
54. K. Kamnitsas, W. Bai, E. Ferrante, S. McDonagh, M. Sinclair, N. Pawlowski, M. Rajchl, M. Lee, B. Kainz, D. Rueckert, et al., Ensembles of multiple models and architectures for robust brain tumour segmentation, in *International MICCAI Brainlesion Workshop* (Springer, 2017), pp. 450–462
55. M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, H. Larochelle, Brain tumor segmentation with deep neural networks. Med. Image Anal. **35**, 18–31 (2017)
56. L. Perez, J. Wang, The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint* arXiv:1712.04621 (2017)
57. S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, R.T. Shinohara, C. Berger, S.M. Ha, M. Rozycki, et al., Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge. *arXiv preprint* arXiv:1811.02629 (2018)
58. T. Saba, M.A. Khan, A. Rehman, S.L. Marie-Sainte, Region extraction and classification of skin cancer: a heterogeneous framework of deep CNN features fusion and reduction. J. Med. Syst. **43**(9), 289 (2019)
59. P. Afshar, K.N. Plataniotis, A. Mohammadi, Capsule networks for brain tumor classification based on MRI images and coarse tumor boundaries, in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, 2019), pp. 1368–1372
60. D. Nie, H. Zhang, E. Adeli, L. Liu, D. Shen, 3d deep learning for multi-modal imaging-guided survival time prediction of brain tumor patients, in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer, 2016), pp. 212–220
61. Y. Xu, Z. Jia, L.-B. Wang, Y. Ai, F. Zhang, M. Lai, I. Eric, C. Chang, Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. BMC Bioinform. **18**(1), 281 (2017)
62. H.H. Sultan, N.M. Salem, W. Al-Atabany, Multi-classification of brain tumor images using deep neural networks. IEEE Access **7**, 69215–69225 (2019)

Chapter 12
Convolutional Neural Network Approach for the Detection of Lung Cancers in Chest X-Ray Images

D. A. A. Deepal and T. G. I. Fernando

Abstract Chest X-rays are considered to be the most widely used technique within the health industry for the detection of lung cancers. Nevertheless, it is very difficult to identify lung nodules using raw chest X-ray images and analysis of such medical images has become a very complicated and tedious task. This study mainly concerned on convolutional neural network approach to identify whether a suspicious area is a nodule or a non-nodule. The JSRT digital images of the chest X-ray database developed by the Japanese Society of Radiological Technology (JSRT) is used to train and test these models. Further, support vector machines and multilayer perceptrons are used for comparison with convolutional neural network model. "Pylearn2" research library is used to build the convolutional neural network model and multilayer perceptron model. "scikit-learn" Python library is used to build the support vector machine models. "MATLAB" is used to extract nodule and non-nodule locations from the original images and other image processing parts. Under support vector machine models, three functions (linear, polynomial and radial) are used and the linear function showed the highest accuracy rate (92%). Comparing of these three approaches, the convolutional neural network approach showed the highest accuracy rate (94%).

Keywords Convolutional neural network (CNN) · Graphics processing unit (GPU) · Gray level co-occurrence matrix (GLCM) · Japanese society of radiological technology (JSRT) · Multilayer perceptron (MLP) · Support vector machine (SVM)

12.1 Introduction

Human is considered as the brainiest animal in the earth. Their inventions and discoveries have made the world developed. Other than the physical, technical and process developments in their urge to study has extended to study their own body. Even

D. A. A. Deepal · T. G. I. Fernando (✉)
Department of Computer Science, Faculty of Applied Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda 10250, Sri Lanka
e-mail: tgi@sjp.ac.lk

U. Kose and J. Alzubi (eds.), *Deep Learning for Cancer Diagnosis*,
Studies in Computational Intelligence 908,
https://doi.org/10.1007/978-981-15-6321-8_12

though the study of the body can be very enlightening, it is rather mysterious. There are functions that the body performs unconsciously, such as the heart beats, respiration and digestion. Other than the amazing aspects, the body of human suffers from different diseases. The cancer is among the most dangerous diseases of human life. Bladder, breast, leukemia, colon and rectal, endometrial, kidney, lung, melanoma, non-Hodgkin lymphoma, prostate, pancreatic and thyroid cancers are the generic types of cancers in the human body. Other than the said types, many people suffer and die from lung cancers than any other cancer [1–3].

The government of Sri Lanka has taken steps to control non-communicable diseases and had declared year 2013 as the year of non-communicable disease prevention. Cancer is a non-communicable disease where the lifetime risk of developing any type of cancer, is one in every 13 people. It is one in every 12 for males and one in every 13 for females.

According to the cancer registry issued in 2007 relating to Sri Lanka has recorded that the cancers in trachea, bronchus and lungs are showing an increasing trend in the age level of thirty and above. Lung cancer is a disease that occurs due to uncontrolled cell growth in tissues of the lung [4]. This can cause metastasis, by affecting adjacent tissue and infiltration beyond the lungs.

It has been revealed that the survival rate of lung cancer patient is only about 14%. If a patient can identify lung cancer symptoms in an early stage, the survival rate can be increased up to 50% [3]. The survival rate is notably improved, but there is a need to increase this survival rate more than the current rate [1]. If lung cancer nodules can be recognized exactly at an early stage, the patients' survival rate can be increased by a significant percentage [5]. There are several methods to use to take an image of the inside of the human body. They are like CT scans, X-rays, MRIs, etc. The CT scan is the most suggested method which produces 3D images of the lungs [1].

The chest X-rays are considered to be the most widely used technique within the health industry for the detection of lung cancer. But it is very difficult to identify lung nodules using raw chest X-ray images and analysis of such medical images has become a very complicated and tedious task [5].

Death of a human cannot be borne by his/her beloveds. The life of a human is momentous. Cancer can destroy the ambitions, love to be lived and the feelings of a person. Not only the cancer patient but also his/her family members and others should suffer. Therefore, this study will be useful to detect lung cancers, so that speedy treatments can be implemented thereafter.

12.2 Literature Review

Udeshani et al. [5] proposed a method to detect a lung cancer in chest X-ray images, using a "Statistical Feature-based Neural Network Approach." Chest X-ray images have been used by them in the form of a novel approach in order to detect lung cancers. They have used a pipeline of image processing routines at the initial stage. This is in

support of removing the noise and segment of the lung out of the other anatomical structures in the X-ray of the chest. Also, regions that exhibit shape characteristics of lung nodules can be extracted by using this image processing routines. Inputs to the neural network are the first and second order statistical texture features. It verifies whether a region extracted in the initial stage a nodule or not and this approach detects nodules in the diseased area of the lung by using the pixel-based technique and the feature-based technique.

Zhou et al. [6] proposed an automatic pathological diagnosis procedure named Neural Ensemble based Detection (NED) to identify lung cancer cells in Chest X-ray images. The ensemble comes with two-level ensemble architecture. Each individual network has only two outputs which can be identified as normal or cancer cell. First level is used to evaluate whether a cell is normal with high confidence. Cells that are evaluated as cancer cells by the first-level deal with the second level and each individual network has five outputs (adenocarcinoma, squamous cell carcinoma, small cell carcinoma, large cell carcinoma and normal). Amongst them the previous four are different types of lung cancer cells. The predictions of those individual networks are combined with an existing method. Neural ensemble-based detection achieved high rates of overall identification with low rates of false negative identification.

Ozekes et al. [7] proposed a method for nodule detection by using the density value of each pixel in CT images. Thereafter, rule-based lung region segmentation has been performed. 8-directional search is used to extract the Regions of Interests (ROIs). After that preliminary classification is executed using Location Change Measurement (LCM). The later nodules are checked using a trained Genetic Algorithm from the images of ROIs. The system showed not only 93.4% sensitivity, but also 0.594 false positive rate.

Ozekes et al. [8] proposed lung cancer nodules detection in four steps using Genetic Cellular Neural Networks (GCNN) and 3D template matching with fuzzy rule-based thresholding. The regions of interests (ROIs) are thereafter extracted using 8-directional search. Convolution-based filters have been used to detect the nodules by searching through 3D images with a 3D template. Finally, fuzzy rule-based thresholding is applied and the ROIs are found.

Abdullah and Shaharum [9] proposed a Lung Cancer Cell Classification Method (LCCCM) using an Artificial Neural Network. In this method, image processing procedures such as image re-processing, lung nodule detection, lung field segmentation, and feature extraction have been used and additionally it has been used an artificial neural network for the classification process.

Shriwas et al. [10] proposed a lung cancer detection and prediction method by using a neural network in CT scan images. Preprocessing of images, image enhancements, image segmentation techniques, feature extraction, GLCM (Gray Level Co-Occurrence Method), binarization approach and neural network classifiers have been used for lung cancer detection.

Kaur and Ada [11] proposed a lung cancer detection algorithm by using a neural network in CT scan images. Preprocessing of image, morphological operators, feature extraction, GLCM (Gray Level Co-Occurrence Method), binarization

approach, PCA (Principal Component Analysis) and neural network classifiers have used for lung cancer detection.

Nancy and Kaur [12] proposed to automate the classification process for the early detection of lung cancer in CT scan images. Lung cancer detection techniques such as preprocessing, training and testing of samples, feature extraction, and classification algorithm (i.e. neural network) and optimization (Genetic Algorithm) have been used in the above study.

With the use of the CT scan images in Dicom (DCM) format, Shettar and Naresh [13] attempted to diagnose lung cancers at its initial stage. This input image has been converted to a gray scaled image and then, to remove Gaussian white noise, a non-local mean filter has been used. The lung part is segmented by using Otsu's threshold from lung CT image and, after that its textural and structural features are extracted from the processed image. Three classifiers (SVM, ANN, and kNN) were applied for the detection of lung cancers and the severity of disease (stage I or stage II). Then, it has been compared with ANN, and kNN classifier according to different quality attributes, namely accuracy, sensitivity (recall), precision and specificity. It revealed that SVM scores higher accuracy of 95.12% when ANN achieves 92.68% accuracy in the analysis of the given dataset and kNN is of least accuracy of 85.37%. SVM algorithm which achieved 95.12% accuracy helps patients to take remedial action on time and reduces mortality rate from this deadly disease.

Retico et al. [14] recommended that the pleural region could be identified with the help of the morphological opening and directional-gradient concentration (DGC). The regions of interest are taken from the segmented pleura region. The features are extracted and candidate nodules are classified using a feedforward neural network. Each nodule candidate is characterized by 12 morphological and textural features, which are analyzed by a rule-based filtering and a neural network classifier.

Maeda et al. [15] emphasized the usage of temporal subtraction of consecutive CT images in order to detect candidate nodules. These features of candidate nodules are calculated thereafter. Afterwards, these are refined under rule-based feature analysis. The feature space is thereafter lowering by referring to PCA and artificial neural network for nodule classification.

Tan et al. [16] introduced the isotropic resampling of a CT image to change the resolution of the image. The lung region is segmented and the nodule center is estimated using the divergence of normalized gradient. The multi-scale nodule and vessel enhancement filtering has been used to segment nodule clusters. The invariant shape and regional descriptors are calculated thereafter. The mix of ANN, GA (FD-NEAT) and SVM have been used for the feature selection and nodule classification.

Kiran et al. [17] proposed an ANN based approach to identify lung cancers from raw chest X-ray images. It mainly consists of three stages. They are pre-processing, feature extraction and classification. The pre-processing involves resizing which enables the further processing is easier. Feature extraction involves extracting the features. The classification stage involves an artificial neural network.

12.3 Research Methodology

This section describes the research methods used in this study in detail. This study proposes two approaches to detect lung cancer nodules using chest X-rays. The JSRT digital image database was used in this study. It contains 154 chest X-rays of lung nodules (100 malignant cases, 54 benign cases), and 93 chest X-rays of non-nodules.

154 cancer nodules and 154 non-nodules were used to train and test the above two approaches. More than one location from some non-cancer X-rays were used to create 154 non-nodule locations from the 93 chest X-ray images and cancer nodule locations of 154 lung cancer X-ray images were used to create 154 lung cancer nodules.

As the initial stage, histogram equalization was performed on all images to improve their intensities.

The methodology used in the first approach is based on MLP and CNN. The rotation and shifting processes were carried out on all images with the aim of increasing the datasets. Then, an area of size128*128 area extracted and the extracted area was then sharpened. The pixel values were used as the training and testing datasets to the MLP and CNN.

The second approach is based on SVM methodology. From the histogram equalized X-ray images, an area of size 128*128 was extracted. The extracted area was then sharpened. The calculated values of the 10 selected statistical values were used as the training and testing datasets for the SVM.

12.3.1 First Order Statistical Features

First order statistical features are based on the statistical moments.

The nth moment is defined as,

$$\mu_n = \sum_{i=0}^{L-1} (z_i - m)^n p(z_i)$$

where μ_n—nth moment, m—average gray level, z_i—possible value of intensity, $p(z_i)$—percentage of pixels with intensity value z_i, and L—Possible intensity levels.

The First order statistical features are listed below (Table 12.1).

12.3.2 Second Order Statistical Features

Gray Level Co-Occurrence Matrix (GLMC) was used to measure the spatial relationship between pixels of examining texture. It is a statistical method, also known as the gray-level spatial dependence matrix. The GLCM functions describe the texture

Table 12.1 First order statistical features

Texture features	Expression	Measure of texture
Average gray level	$m = \sum_{i=0}^{L-1} z_i p(z_i)$	A measure of average intensity
Standard deviation	$\sigma = \sqrt{\mu_2(z)}$ $\mu_2(z)$ is second moment	A measure of average contrast
Smoothness	$R = 1 - \frac{1}{(1+\sigma^2)}$	Measures the relative smoothness of the intensity in a region. This is 0 for a region of constant intensity and 1 for regions with large excursions in the value of its intensity level
Skewness (Third moments)	$\mu_3 = \sum_{i=0}^{L-1} (z_i - m)^3 p(z_i)$	Measures the skewness of a histogram. This measure is 0 for symmetric histograms. Positive for histograms skewed to the right and negative for the left
Uniformity	$U = \sum_{i=0}^{L-1} p^2(z_i)$	Measures uniformity. This measure has its maximum when all gray levels are equal (maximally uninform) and decreases from that point
Entropy	$e = -\sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i)$	A measure of randomness

of an image. It calculates how often pairs of pixels with specific values and in a specified spatial relationship occur in an image. The pixel of interest and the pixel to its immediate right (horizontally adjacent) is defined as the spatial relationship. Statistical measures can be extracted from this matrix. These statistics provide information about the texture of an image. The following Table 12.2 lists the second order statistics.

Table 12.2 Second order statistical features

Statistic	Description
Contrast	Measures the local variations in the gray-level co-occurrence matrix
Correlation	Measures the joint probability occurrence of the specified pixel pairs
Energy	Provides the sum of squared elements in the GLCM. Also known as uniformity or the angular second moment
Homogeneity	Measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal

12.3.3 System Architecture

Figure 12.1 shows the system architectures of the two approaches used in this research.

12.3.3.1 Histogram Equalization

Histogram equalization method was applied to X-ray images. It helps to adjust image intensities to enhance the contrast of X-ray images (Fig. 12.2).

12.3.3.2 Approach 1

The first approach is based on Convolutional neural network (CNN) and Multilayer perceptron (MLP). The pixels values were used as the training and testing datasets to MLP and CNN.

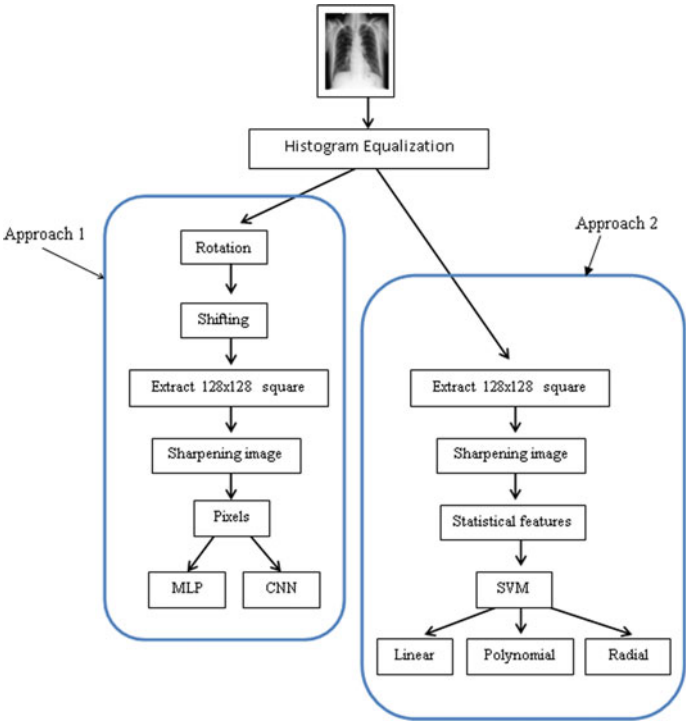


Fig. 12.1 System architecture

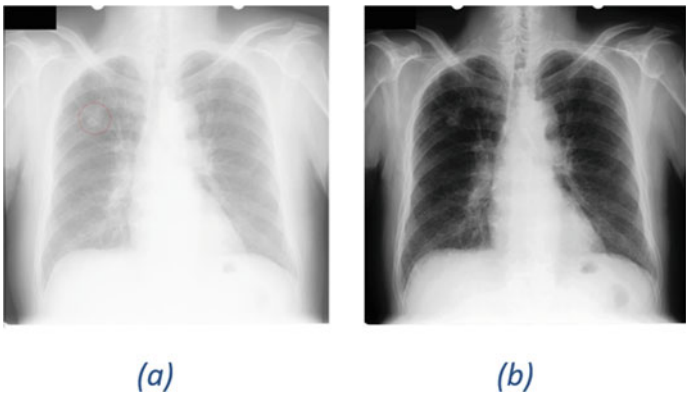


Fig. 12.2 Histogram equalization. **a** Before histogram equalization. **b** After histogram equalization

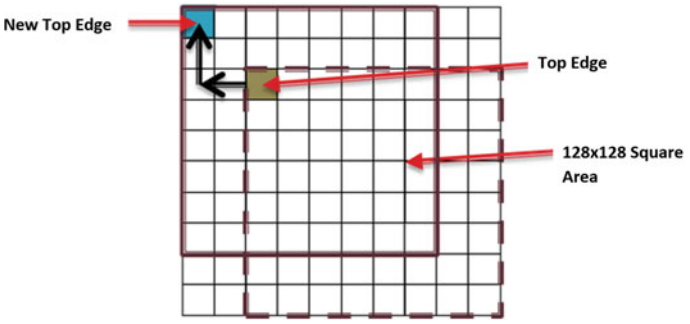


Fig. 12.3 Step 1

To increase the number of images in each category, the following augmenting methods were used.

Rotation

In this method the original image is rotated by 0°, 5°, 10°, 15°, 20°, −5°, −10°, −15° and −20°.

Pixels Shifting

Step 1: Shift the top edge of the marked 128 × 128 area by two pixels left and two pixels top (Fig. 12.3).

Step 2: Extract an area of size 128 × 128 pixels from the image and shift the new top edge by one pixel right (Fig. 12.4).

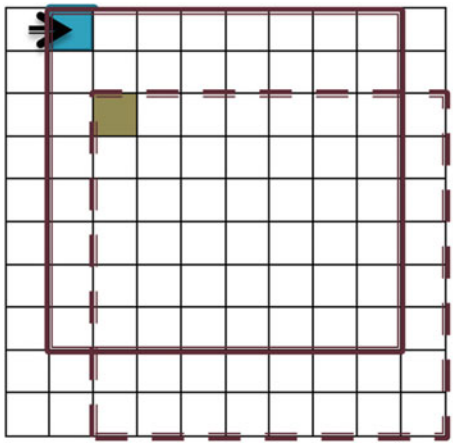


Fig. 12.4 Step 2

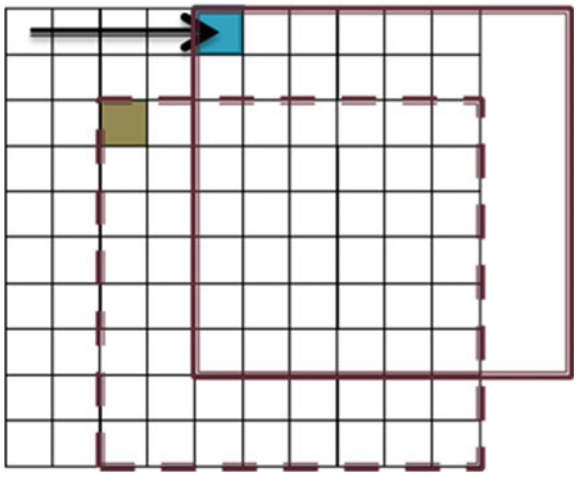


Fig. 12.5 Step 3

Step 3: Repeat the step 2 until total right shifted pixel count is equal to 5 (Fig. 12.5).

Step 4: Shift the new top edge by 5 pixels left and shift down by one pixel (Fig. 12.6).

Step 5: Repeat step 3 and step 4 until total down shifted pixel count is equal to 5 (Fig. 12.7).

25 images were created from one image using the pixel shifting method.

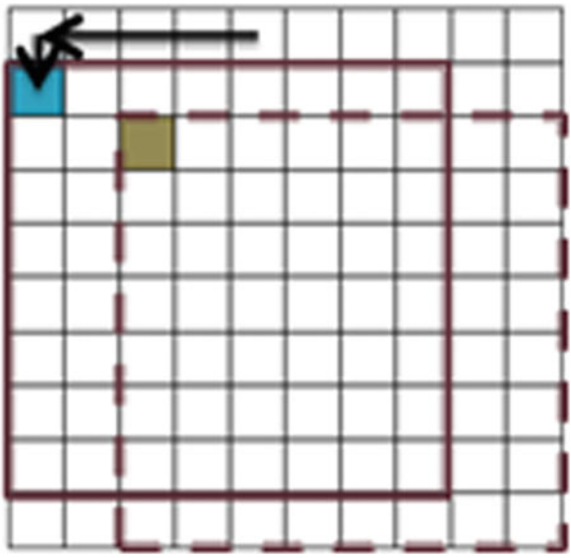


Fig. 12.6 Step 4

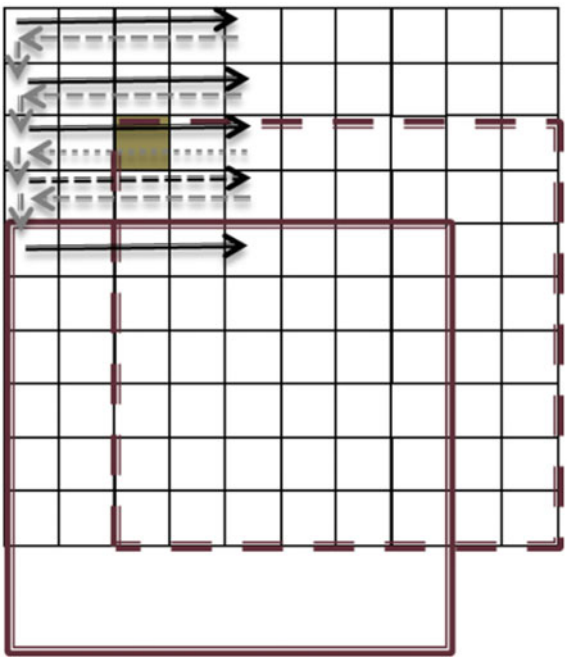


Fig. 12.7 Step 5

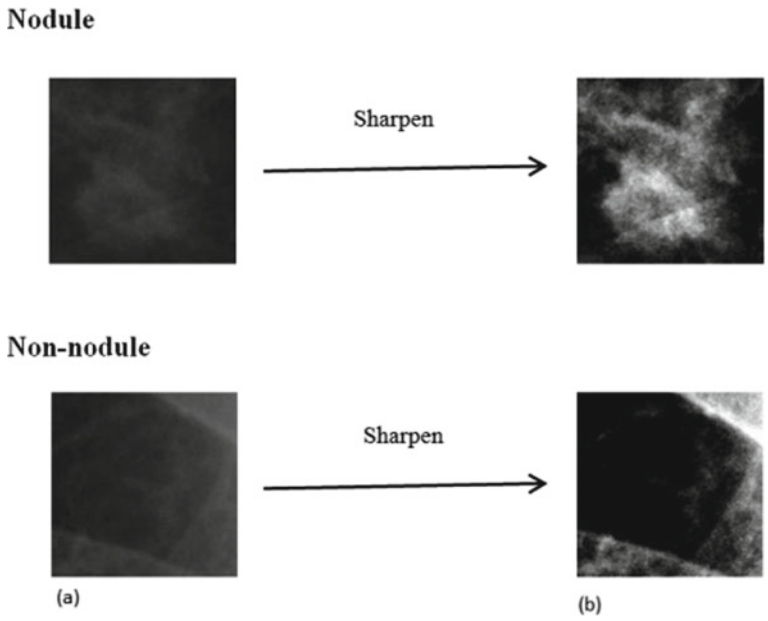


Fig. 12.8 Sharpening the image. **a** Original images. **b** Sharpen images

Extract 128 × 128 square areas from X-Ray images

128*128 areas of pixels were extracted from the images generated by applying the previously mentioned (histogram equalization, rotation and pixels shifting) methods to original images.

Sharpening the Image

It enhances the definition of edges in the extracted area (Fig. 12.8).

Datasets for MLP and CNN

The pixel values were used as the input to MLP and CNN. These neural networks require a large dataset. Rotating and pixel shifting were used to expand the dataset. The original image is rotated by 0°, 5°, 10°, 15°, 20°, −5°, −10°, −15° and −20° and extracted the above 25 shifted images (128 × 128) from each rotated image.

Two hundred and twenty five (225) images were created from one image using the above two methods. These images were used to train MLP and CNN networks.

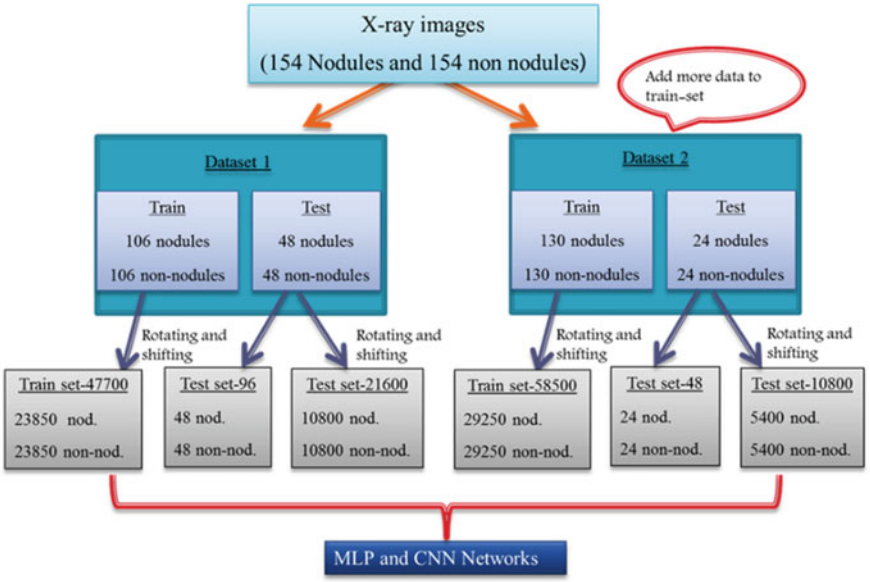


Fig. 12.9 Training and testing datasets for MLP and CNN

Training and Testing Datasets for MLP and CNN

In this approach 308 images (nodule and non-nodule) were used and they were divided into training and testing datasets (Fig. 12.9). Then two pairs of training/testing datasets were created using augmented methods (rotating and shifting).

First Training/Testing Datasets

The first training dataset contains 212 images (106 nodules and 106 non-nodules) and these images were expanded to 47,700 images using rotating and pixels shifting methods. Other 96 images (48 nodules and 48 non-nodules) were used as a testing dataset and these 96 images also expanded to 21,600 images using rotating and pixels shifting methods.

These expanded 47,700 images were used to train MLP and CNN networks and above 96 images and expanded 96 images (21,600 images) were used to test these trained networks.

Second Training/Testing Datasets

In the second training dataset, more 48 images (24 nodules and 24 non-nodules) were added from the testing dataset to the training dataset. The second training dataset contains 260 images (130 nodules and 130 non-nodules) and these images were also expanded to 58,500 images using rotating and pixels shifting methods. Other 48

images (24 nodules and 24 non-nodules) were used as the testing dataset and these 48 images were also expanded to 10,800 images using rotating and pixels shifting methods.

These expanded 58,500 images were used to train MLP and CNN networks and above 48 images and expanded 48 images (10,800 images) were used to test these trained networks.

MLP Approach

In the first method, first training dataset (47,700 images) was used to train the MLP network and this trained MLP network was tested with 96 images and expanded 96 images (21,600 images).

In the second method, second training dataset (58,500 images) was used to train the MLP network and this trained MLP network was tested with 48 images and expanded 48 images (10,800 images).

Parameters of the MLP Approach

The batch size of the MLP training model which was developed in this study is 1,000 and the learning rate is 0.0001. The coefficient of the model is 0.0005 and include probability of the dropouts is 0.8.

MLP Network Architecture

This MLP network has input layer, 5 hidden layers (4 fully-connected layers with 4,000 nodes and one fully-connected layer with 1,024 nodes) and a softmax layer. Dropouts were used to reduce the overfitting. The MLP network is shown in Fig. 12.10.

CNN Approach

In the first method, first training dataset (47,700 images) was used to train the CNN network and this trained CNN network was tested with 96 images and expanded 96 images (21,600 images).

In the second method, second training dataset (58,500 images) was used to train the CNN network and this trained CNN network was tested with 48 images and expanded 48 images (10,800 images).

Parameters of the CNN Approach

The batch size of the CNN training model developed in this study is 100 and the learning rate is 0.0001. The coefficient is 0.0005 and the probability of the dropouts is 0.8.

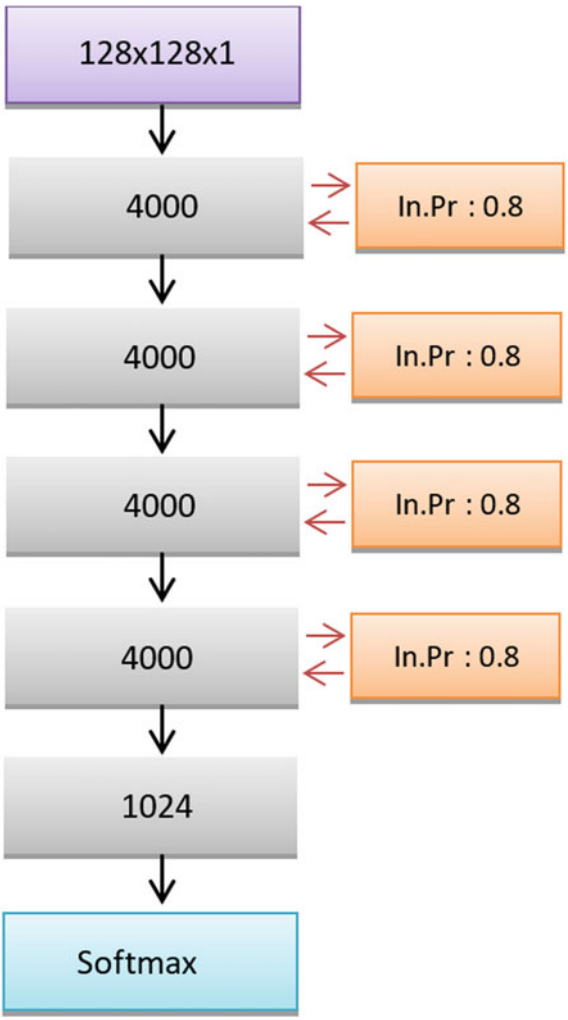


Fig. 12.10 MLP network architecture

CNN Network Architecture

This CNN network has 5 Convolutional layers, 3 Max-pooling layers and 4 fully-connected layers and 1 softmax layer. Dropouts were used to reduce the overfitting and apply dropout to fully-connected layers. The MLP network is shown in Fig. 12.11.

12.3.3.3 Approach 2

This approach is based on SVM methodology. The calculated values of the 10 selected statistical features were used as the training and testing datasets for the SVM.

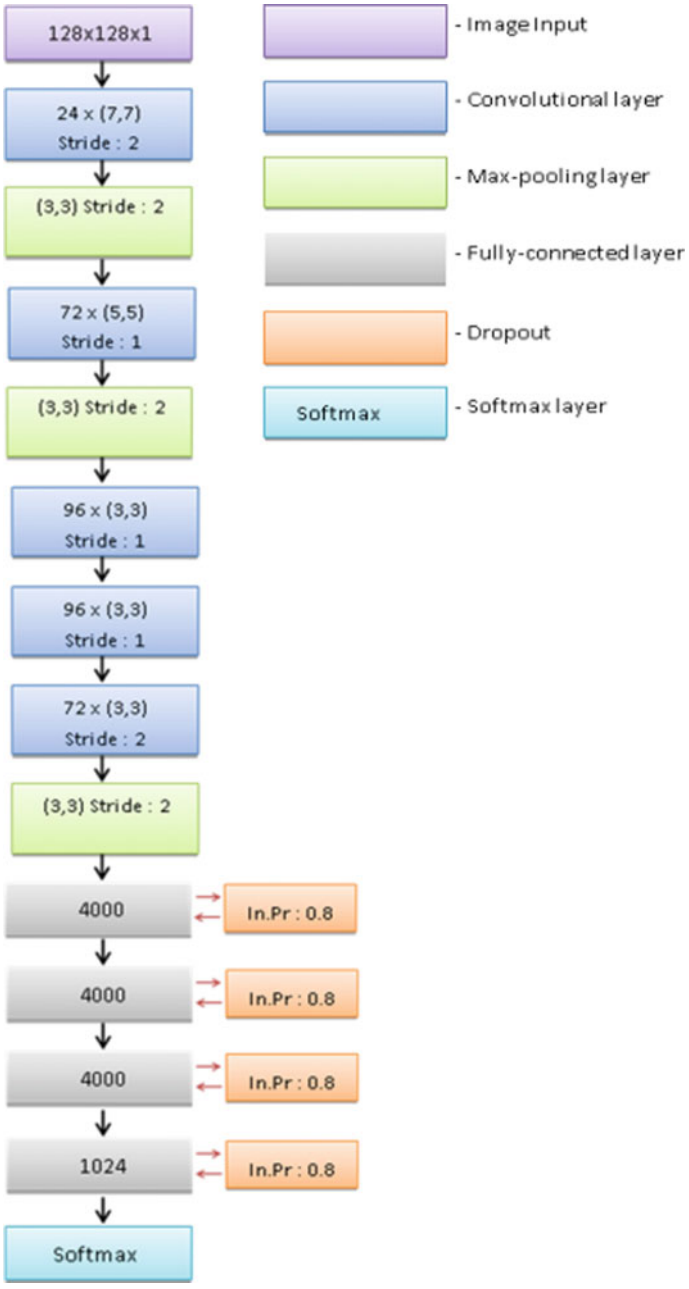


Fig. 12.11 CNN network architecture

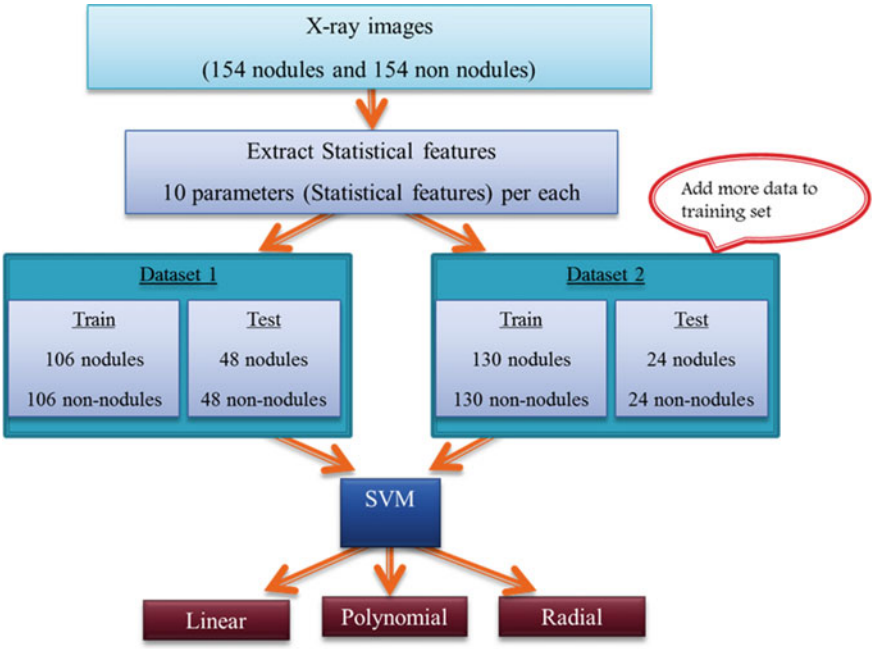


Fig. 12.12 Training and testing datasets for SVM

Dataset for SVM

From the histogram equalized X-ray images, 128*128 area was extracted. The extracted area was then sharpened. Ten statistical features were extracted from the above 154 nodule locations and 154 non-nodule locations. These statistical features were divided into training and testing sets.

In the first training dataset, these 308 images were divided into two samples for training (212 images) and testing (96 images) and the second training dataset, another 48 images were added from the earlier testing dataset to training set. There were 260 images for training and 48 images for testing (Fig. 12.12).

SVM Approach

In this approach, linear, polynomial and radial based SVMs were used to identify lung cancer nodules and statistical features were used as the input to the above SVMs. In first method, first training dataset (212 images) and testing dataset (96 images) were used to train and test the above SVMs. In second method, second training dataset (260 images) and testing dataset (48 images) were used to train and test the above SVMs (Table 12.3).

Table 12.3 Parameters of the SVM approach

SVM model	Penalty parameter of the error term (C)	Degree	Gamma
Linear	0.151		
Radial	0.5	2	0.01
Polynomial	0.5	2	0.009

12.4 Results and Discussion

This section presents the results obtained for the two main approaches carried out to achieve our research objectives. In the first approach, the extracted pixel data from the X-ray images were given to MLP and CNN to identify the lung cancer nodules in the images and in the second approach, extracted data were tested using the SVM methodology. Under SVM method, we carried out the experiments based on linear, polynomial and radial kernel functions.

12.4.1 Results

12.4.1.1 Test Results of MLP Network

In the first method, the first training dataset (47,700 images) was used to train the MLP network. Batch size of the model was 1,000. Learning rate of the model was 0.0001 and the coefficient of the model was 0.0005.

This trained MLP network was tested with 96 images. The accuracy rate of the identifying an actual nodule as a nodule was 91.67% (Table 12.5) and the accuracy rate of the identifying an actual non-nodule as a non-nodule was 91.67% (Table 12.5). The error rate of the identifying an actual nodule as a non-nodule was 08.33% (Table 12.5) and the error rate of the identifying an actual non-nodule as a nodule was 08.33% (Table 12.5). The overall accuracy of this model was 91.667% (Table 12.4).

Next, this trained MLP network was tested with expanded 96 images (21,600 images). The accuracy rate of the identifying an actual nodule as a nodule was decreased to 88.64% (Table 12.5) and the error rate of the identifying an actual

Table 12.4 Test results of the CNN and MLP

Method	Overall accuracy (%)			
	Train images—47,700		Train images—58,500	
	Test images 96	Test images 21,600	Test images 48	Test images 10,800
MLP	91.667	89.417	91.667	89.945
CNN	89.001	88.593	94.001	94.538

Table 12.5 Image identification rates of the CNN and MLP for 47,700 training images

Method	Image type	Train images—47,700			
		Test images—96		Test images—21,600	
		Identify as a nodule (%)	Identify as a non-nodule (%)	Identify as a nodule (%)	Identify as a non-nodule (%)
MLP	Nodule	91.67	8.33	88.64	11.36
	Non-nodule	8.33	91.67	9.75	90.25
CNN	Nodule	89.58	10.42	88.54	11.46
	Non-nodule	12.50	87.50	11.35	88.65

nodule as a non-nodule was increased to 11.36% (Table 12.5). Also, the accuracy rate of the identifying an actual non-nodule as a non-nodule was decreased to 90.25% (Table 12.5) and the error rate of the identifying a non-nodule as a nodule was increased to 9.75% (Table 12.5). The overall accuracy of the model was decrease to 89.417% (Table 12.4).

In the second method, the second training dataset (58,500 images) was used to train the above MLP network and this trained MLP network was tested with 48 images and expanded 48 images (10,800 images). Batch size of the model was 1,000. Learning rate of the model was 0.0001 and the coefficient of the model was 0.0005.

This trained MLP network was tested with 48 images. This MLP model remains with same identification rates and error rates compared to its earlier trained model tested with 96 images (Table 12.6). The overall accuracy of this model was not changed and it was 91.667% (Table 12.4).

Thereafter this trained MLP network was tested with expanded 48 images (10,800 images). The accuracy rate of the identifying an actual nodule as a nodule was relatively not changed and it was 91.69% (Table 12.5). The accuracy rate of the identifying an actual non-nodule as a non-nodule was decreased to 88.20% (Table 12.6). The overall accuracy of the model was decreased to 89.945% (Table 12.4).

Table 12.6 Image identification rates of the CNN and MLP for 58,500 training images

Method	Image type	Train images—58,500			
		Test images—48		Test images—10,800	
		Identify as a nodule (%)	Identify as a non-nodule (%)	Identify as a nodule (%)	Identify as a non-nodule (%)
MLP	Nodule	91.67	8.33	91.69	8.31
	Non-nodule	8.33	91.67	11.80	88.20
CNN	Nodule	95.83	4.17	94.13	5.87
	Non-nodule	8.33	91.67	5.06	94.94

12.4.1.2 Test Results of CNN Network

In the first method, first training dataset (47,700 images) was used to train the CNN network. Batch size of the model was 1,000. Learning rate of the model was 0.0001 and the coefficient of the model was 0. 00005.

This trained CNN network was tested with 96 images. The accuracy rate of the identifying an actual nodule as a nodule was 89.58% (Table 12.5) and the accuracy rate of the identifying an actual non-nodule as a non-nodule was 87.50% (Table 12.5). The error rate of the identifying an actual nodule as a non-nodule was 10.42% (Table 12.5) and the error rate of the identifying an actual non-nodule as a nodule was 12.50% (Table 12.5). The overall accuracy of this model was 89.001% (Table 12.4).

In the second method, second training dataset (58,500 images) was used to train the above CNN network and this trained CNN network was tested with 48 images and expanded 48 images (10,800 images). Batch size of the model was 1,000. Learning rate of the model was 0.0001 and the coefficient of the model was 0. 00005.

This trained CNN network was tested with 48 images. This CNN model was succeeded in identifying an actual nodule as a nodule image with a rate of 95.83% (Table 12.6). The error rate of the identifying a nodule as a non-nodule was decreased to 4.17% (Table 12.6). Also, the accuracy rate of the identifying an actual non-nodule as a non-nodule was increased to 91.67% (Table 12.6) and the error rate of the identifying an actual non-nodule as a nodule was decreased to 8.33% (Table 12.6). The overall accuracy of this model was increased to 94.001% (Table 12.4).

Next, this trained CNN network was tested with expanded 48 images (10,800 images). The accuracy rate of the identifying an actual nodule as a nodule was increased to 94.13% (Table 12.6) and the error rate of the identifying an actual nodule as a non-nodule was decreased to 05.87% (Table 12.6). The accuracy rate of the identifying an actual non-nodule as a non-nodule was increased to 94.94% (Table 12.6) and the error rate of the identifying a non-nodule as a nodule was decreased to 05.06% (Table 12.6). The accuracy of the model was increased to 94.538% (Table 12.4).

12.4.1.3 Test Results of SVM

Linear SVM Model

In this method, linear kernel function was used to fit the linear model. Penalty parameter of the error term (C) was set to 0.151.

First dataset containing 212 images was used to train the model. It gave 92.71% overall accuracy (Table 12.7) when it was tested using the testing dataset of 96 images. The accuracy rate of the identifying an actual nodule as a nodule was 87.50% (Table 12.8) and the accuracy rate of the identifying an actual non-nodule as a non-nodule was 97.92% (Table 12.8). The error rate of the identifying an actual nodule as a non-nodule was 12.50% (Table 12.8) and the error rate of the identifying an actual non nodule as a nodule was 02.08% (Table 12.8).

Table 12.7 Test results of the SVMs

Method	Overall accuracy (%)	
	Train images—212 Test images—96	Train images—260 Test images—48
Linear	92.708	87.500
Polynomial	83.333	79.167
Radial	84.375	77.083

Table 12.8 Image identification rates of the SVMs for 260 training images

Method	Image type	Train images—212 Test images—96	
		Identify as a nodule (%)	Identify as a non nodule (%)
Linear	Nodule	87.50	12.50
	Non nodule	2.08	97.92
Polynomial	Nodule	81.25	18.75
	Non nodule	14.58	85.42
RBF	Nodule	89.58	10.42
	Non nodule	20.83	79.17

Table 12.9 Image identification rates of the SVMs for 212 training images

Method	Image type	Train images—260 Test images—48	
		Identify as a nodule (%)	Identify as a non nodule (%)
Linear	Nodule	83.33	16.67
	Non nodule	8.33	91.67
Polynomial	Nodule	79.17	20.83
	Non nodule	20.83	79.17
RBF	Nodule	83.33	16.67
	Non nodule	29.17	70.83

Second dataset with 260 images was used to train the above linear model and tested this model using 48 images. The accuracy rate of the identifying an actual nodule as a nodule was decreased to 83.33% (Table 12.9) and the error rate of the identifying an actual nodule as a non-nodule was increased to 16.67% (Table 12.9). The accuracy rate of the identifying an actual non-nodule as a non-nodule was decreased to 91.67% (Table 12.9) and the error rate of the identifying a non-nodule as a nodule was increased to 08.33% (Table 12.9). The overall accuracy of this linear model was 87.50% (Table 12.7).

Polynomial SVM Model

In this method, polynomial kernel function with degree 2 was used to fit the SVM model. Penalty parameter of the error term (C) was 0.5 and kernel coefficient (gamma) of this model was 0.009.

First dataset with 212 images was used to train this polynomial model and tested this model using 96 images. The accuracy rate of the identifying an actual nodule as a nodule was 81.25% (Table 12.8) and the error rate of the identifying an actual nodule as a non-nodule was 18.75% (Table 12.8). The accuracy rate of the identifying an actual non-nodule as a non-nodule was 85.42% (Table 12.8) and the error rate of the identifying an actual non-nodule as a nodule was 14.58% (Table 12.8). The overall accuracy of this polynomial model was 83.33% (Table 12.7).

Second dataset (260 images) was used to train the above polynomial model and tested this model using 48 images. The accuracy rates of the identifying an actual nodule as a nodule and actual non-nodule as a non-nodule were decreased to 79.17% (Table 12.9) and the error rates of the identifying a nodule as a non-nodule and actual non-nodule as a nodule were increased to 20.83% (Table 12.9). The overall accuracy of this polynomial model was 79.17% (Table 12.7).

Radial SVM Model

In this model, radial kernel function was used to fit the SVM model. The degree of this radial model was 2 and penalty parameter of the error term (C) was 0.5. Kernel coefficient (gamma) of this model was 0.01.

First dataset with 212 images was used to train this radial model and tested this model using 96 images. The accuracy rate of the identifying an actual nodule as a nodule was 89.58% (Table 12.8) and the accuracy rate of the identifying an actual non-nodule as a non-nodule was 79.17% (Table 12.8). The error rate of the identifying an actual nodule as a non-nodule was 10.42% (Table 12.8) and the error rate of the identifying an actual non-nodule as a nodule was 20.83% (Table 12.8). The overall accuracy of this radial model was 84.375% (Table 12.7).

Second dataset with 260 images was used to train the above radial model and tested this model using 48 images. The accuracy rate of the identifying an actual nodule as a nodule was decreased to 83.33% (Table 12.9) and the error rate of the identifying an actual nodule as a non-nodule was increased to 18.75% (Table 12.9). Also, the accuracy rate of the identifying an actual non-nodule as a non-nodule was decreased to 70.83% (Table 12.9) and the error rate of the identifying an actual non-nodule as a nodule was increased to 29.17% (Table 12.9). The overall accuracy of this radial model was decreased to 77.083% (Table 12.9).

12.4.2 Discussion

When compared with the original test images, MLP has succeeded in identifying nodule images and non-nodule images correctly with a ratio of 91.67% while CNN method identifying nodule images correctly with 89.58% and non-nodule images correctly with 87.5% for 47,700 training images. When compare with overall accuracies, MLP shows a higher overall accuracy of 91.667% over CNN method for 47,700 training images.

But, MLP fails to increase the accuracy when the training dataset is increased to 58,500 images. It remains with same accuracy rates compared to its earlier results of 47,000 training images. On the other hand, CNN has succeeded in increasing its accuracy rate of identifying nodule images correctly to 95.83% and non-nodule images to 91.67%. It also gives the lowest rate of 4.17% in identifying nodule images as non-nodule images.

When compared with overall accuracies, CNN has succeeded to increase its overall accuracy when the training dataset is increased (89.00% $\rightarrow$ 94.00%).

When compared with the expanded test images, the accuracy rates of both MLP and CNN models trained with 47,000 images, does not show significant difference for identifying nodule images correctly and identifying non- nodule images correctly. But both MLP and CNN models have been able to increase their accuracy rates when the training dataset increased. Again, CNN method gives the highest accuracy rate in identifying nodule images correctly and lower rates of 5.87% and 5.06% in incorrectly identifying nodule images as non-nodule images and vice versa. When compare with overall accuracies, MLP shows a lower accuracy rate compared to its previous results. But CNN has again succeeded in increasing its accuracy rate when the training dataset is increased (88.59% $\rightarrow$ 94.54%).

Therefore, it can be concluded that CNN shows higher overall performance when compared to MLP.

Among the three SVM approaches, linear approach shows the highest accuracy rate. Radial based approach gives higher accuracy than polynomial based approaches. It is also notable that in all three approaches have failed to increase their accuracy rates when the training dataset is increased.

Therefore, we can conclude that linear based approach is capable of identifying lung cancer nodules than other two SVM approaches.

Comparing of these SVM, MLP and CNN approaches, the CNN network approach showed the highest accuracy and it was reported as 94.54%.

12.5 Conclusions

The main objective of this research was to develop a convolutional neural network approach for the detection of lung cancers in chest X-ray images. SVM and MLP

approaches were used to compare the results of the above CNN network. Feature-based technique was used to train the SVM model and pixel-based technique was used to train MLP and CNN networks.

Among the three SVM approaches, the linear based approach shows the highest accuracy. Therefore, it can be concluded that linear based approach is capable of identifying lung cancer nodules than other two SVM approaches.

When compared with deep learning approaches, CNN shows the highest accuracy rate compared to MLP. Therefore, it can come to the conclusion of saying CNN approaches are capable of identifying lung cancer nodules than MLP approach.

Comparing of these SVM, MLP and CNN approaches, the CNN network approach shows the highest accuracy and therefore it can be concluded that CNN approaches are capable of identifying lung cancer nodules than other two approaches.

References

1. S. Akram, A. Khanum, U. Qamar, A. Hassan, M.Y. Javed, Artificial Neural Network Based Classification Of Lungs Nodule Using Hybrid Features From Computerized Tomographic Images. Appl. Math. Inf. Sci. **9**(1), 183–195 (2015)
2. R.T. Greenlee, T. Murray, S. Bolden, P.A. Wingo, Cancer statistics. CA Cancer J. Clin. **50**(1), 7–33 (2000)
3. K.W. Jung, Y.J. Won, S. Park, H.J. Kong, J. Sung, H.R. Shin, E.C. Park, J.S. Lee, Cancer statistics in Korea: incidence, mortality and survival in 2005. J. Korean Med. Sci. **43**, 1–11 (2011)
4. A. Vishwa, Pre-diagnosis of lung cancer using feed forward neural network and back propagation algorithm. Int. J. Comput. Sci. Eng. (IJCSE), **3**(9) (2011). ISSN: -0975-3397
5. K.A.G. Udeshani, T.G.I. Fernando, R.G.N. Meegama, Statistical feature-based neural network approach for the detection of lung cancer in chest X-ray images. Int. J. Image Process. (IJIP) **5**(4), 425–434 (2011)
6. Z.H. Zhou, S.F. Chen, Y.B. Yang, Y. Jiang, Lung cancer cell identification based on artificial neural network ensembles, in *Artificial Intelligence in Medicine*, vol. 24, pp. 25–36 (2002). ISSN 0933-3657
7. S. Ozekes, Rule-based Lung region segmentation and nodule detection via Genetic Algorithm trained template matching. Istanbul Commer. Univ. J. Sci. **6**, 17–30 (2007)
8. S. Ozekes, O.N. Ucan, O. Osman, Nodule detection in a lung region that's segmented with using genetic cellular neural networks and 3D template matching with fuzzy rule-based thresholding. Korean J. Radiol. **9**, 1–9 (2008)
9. A. Abdullah, S.M. Shaharum, Lung cancer cell classification method using artificial neural network. Inf. Eng. Lett. 2(1). ISSN: 2160-4114
10. R.S. Shriwas, D. Akshay, M.E. Dikondawar, J.D.I.E.T. Yavatmal, Lung cancer detection and prediction by using neural network. IPASJ Int. J. Electron. Commun. (IIJEC) 3(1) (2015). ISSN 2321-5984
11. R.K. Ada, Early detection and prediction of lung cancer survival using neural network classifier. Int. J. Appl. Innov. Eng. Manage. (IJAIEN) 2(6) (2013). ISSN 2319 – 4847
12. P.K. Nancy, Identifying lung cancer in its early stage using neural networks and GA algorithm. Int. J. Adv. Res. Comput. Eng. Technol. (IJARCET) 4, 341, ISSN: 2278 – 1323
13. R. Shettar, P. Naresh, Early detection of lung cancer using neural network techniques. Int. J. Eng. Res. Appl. 4(8), 78–83 (2014). ISSN: 2248-9622

14. A. Retico, M.E. Fantacc, I. Gori, P. Kasae, B. Golosio, A. Piccioli, P. Cerello, G.D. Nunzio, S. Tangaro, Pleural nodule identification in low-dose and thin-slice lung computed tomography. Comput. Biol. Med. **39**, 1137–1144 (2009)
15. S. Maeda, Y. Tomiyama, H. Kim, N. Miyake, Y. Itai, J.K. Tan, S. Ishikawa, A. Yamamoto, Detection of lung nodules in thoracic MDCT images based on temporal changes from previous and current images. Adv. Comput. Intell. Intell. Inf. **15**, 707–713 (2011)
16. M. Tan, R. Deklerck, B. Jansen, M. Bister, J. Cornelis, A novel computer-aided lung nodule detection system for CT images. MedPhys **38**, 5630–5645 (2011)
17. S. Kiran, *An efficient lung cancer detection based on artificial neural network* (Int. J. Eng. Sci. Manage. Res., ISSN, 2015), pp. 2349–6193

Chapter 13
Future of Deep Learning for Cancer Diagnosis

Pinar Koc and Cihan Yalcin

Abstract Long-term treatments observed in cancer-derived diseases have a higher risk of death than other diseases. The effectiveness of the developing technology in human life is manifested in the treatment and diagnosis of diseases. '*Early diagnosis saves lives*', which is frequently heard in all diseases, comes to life in this part. The first of the most important points in cancer and derivative diseases is early detection of the disease. Artificial intelligence is used to simulate and simplify the human life offered by developing technology. This study focuses on the methods of deep learning, which is one of the subfields of artificial intelligence. The aim of this study is to emphasize the deep learning methods used in cancer diagnosis. As a result of emphasizing the methods, the present and future potential of the literature in terms of cancer diagnosis has been revealed. It is thought that the study will be a current reference for the researchers who will conduct research within the scope of the subject.

Keywords Deep learning · Artificial intelligence · Deep learning in cancer diagnosis · Deep learning in bioinformatics · Machine learning

13.1 Introduction

Technological tools, machines and innovations have occurred in the life of human beings since the existence and continue to come. Although these developments and innovations were sometimes out of curiosity, sometimes in line with the needs. One of the most important issues affecting the development of technology and science, in line

P. Koc
Institute of Education Sciences, Computer and Instructional Technologies Education, Suleyman Demirel University, Isparta, Turkey
e-mail: pinarkoc1995@gmail.com

C. Yalcin (✉)
Institute of Education Sciences, Computer Engineering Department, Suleyman Demirel University, Isparta, Turkey
e-mail: cihanyalcin@sdu.edu.tr

U. Kose and J. Alzubi (eds.), *Deep Learning for Cancer Diagnosis*,
Studies in Computational Intelligence 908,
https://doi.org/10.1007/978-981-15-6321-8_13

with the needs is the diseases that affect human life. Although some of the diseases can be easily diagnosed and treated, a long and difficult process may be required in the diagnosis and treatment of some diseases. Cancer, which requires a long and challenging process and has difficulties in its diagnosis and treatment, includes cancer. Considering the definition of cancer is the formation of bad neoplasms with the irregular division and reproduction of the cells in the tissue and organ [4]. According to another definition, cancer is a cluster of more than 100 diseases that contain groups of cells in the body [14]. If a general definition of cancer is to be made from the given definitions, it is possible to define "cancer, in the organs and in the cells inside the organs, to have different diseases in the uncontrolled and different form of normal growth." Diagnosis and treatment of cancer disease, requires both a long process and high costs. At the same time, cancer disease takes the second place in diseases that result in death [4]. Therefore, WHO (World Health Organization) emphasizes the methods of protection from them in order not to catch diseases [13]. WHO stated that approximately 9.6 million people died from various cancer types in 2018, and that even by 2030, approximately 13.1 million people could die due to cancer disease [29]. The most important issue in preventing these deaths caused by cancer is the most important diagnosis (early diagnosis). When current data are analyzed, it is concluded that 90% of cancer patients are treated by early diagnosis [25].

Artificial intelligence is a learning system that will imitate human intelligence [13]. The purpose of artificial intelligence technology is to create intelligent systems through to digital systems to solve real world problems [16–19]. It is the learning of machinery to the branch of science that makes people's lives easier and allows them to think and imitate them as people [13]. Machine learning, which is one of the most important sub-branches of artificial intelligence, forms the basis of the system called intelligent systems that learn the existing situation, develop it and adapt to this situation [16–19]. In fact, machine learning is the backbone of artificial intelligence [16–19]. Machine learning is a system that includes learning, application and testing processes.

The machine learning technique completes the training of the algorithm thanks to the learning time and checks whether there is any problem in the learning part of the system with a test process before the actual application [16–19]. In this way, the application of the algorithm that the system has been trained and in a healthy way can be observed. Machine learning, which constitutes the building blocks of the current technology, emerged in the 1980s [3]. It is possible to divide the learning approaches of artificial intelligence and machine learning into 4. These approaches are; Consultant learning, Unconsultant learning, Reinforced learning and Semi-Consultant learning. The educational data in consulting learning method and the result to be used in the method/system to be used are obvious. In consultant-free learning, there are again educational data, but the outcome is uncertain. In reinforced learning, the existing problem situation is externally directed with feedbacks such as yes-no, true-false, and the applied method is continued with experience. Semi-Consultant learning is the method in which systems with and without the results will be used together [16–19]. Although machine learning is used in many fields such as industry, material production and security systems, it is also used in medical fields such as diagnosis

of diseases and making forward predictions [8, 20], Ayturan [3]. Machine learning is one of the most used systems in the early diagnosis of diseases [13].

It is thought that any disruption experienced in cancer diagnosis and treatment that continues to increase causes tremendous problems. Within the scope of the study, emphasis is placed on diagnostic applications using deep learning methods in cancer diagnosis. In this study, the information about which deep learning method is used in studies according to cancer diagnosis is given. The study content in the section one; what is deep learning, what are their types and deep learning architectures included. In section two, which deep learning methods used in the field of health and case studies. Especially, the intensity of deep learning information used in diagnostics is given. In the third part; processes of cancer diagnosis which has used deep learning methods and expectations are given as information. In addition, how to benefit from deep learning can be used in cancer treatment and the expectations/predictions about what direction the studies should be. Discussion and conclusion part, using deep learning technologies advantages of commonly used methods for cancer diagnosis perhaps a remedy solution and the way of direction to how future studies will provide a given direction during the event is referred.

13.2 Deep Learning

In cancer diagnosis, deep learning in the classes of machine learning is used. In deep learning, it is the area studied by the analogy method from the neural networks in the brain. Deep learning is the area used to simulate multilayered and mixed problems of machines produced by comparing the neural networks of the brain [3]. According to a different definition, deep learning is a system that processes raw data and automatically extracts the desired features [31]. The term deep learning is also expressed in the form of the Deep Neural Network [6]. The biggest differences between machine learning and deep learning are the ability to easily process larger data with the machine learning method and the way deep learning solves this data [3]. One way of deep learning does not learn the rules, coded differently from classical machine learning systems, instead it is used for sound, video, picture etc. It automatically learns files such as icons [13]. It has an algorithm structure in deep learning and these algorithms can be supervised/unsupervised [24]. The system that allows multi-layer network structures to be used together is deep learning [26]. In the system in consecutive layer structures, data that is output on one layer is taken as input on the other layer [24]. It is possible to talk about different architectures (methods) in the field of deep learning. These architectures are; Deep Neural Networks (DNN), Evolutionary (Convolutional) Neural Networks (ENN), Repetitive (Recursive) Neural Networks (RNN), Long-Short Term Memory Networks (LSTM), Deep Boltzman Machine (DBM), Deep Faith Networks (DFN) and Deep Auto-Encoders (DAE) It can be examined in 7 subgroups [31].

Deep Neural Networks (DNN), having more than two layers, this network structure is generally used for classification and analysis. It enables the identification of complex data structures [13].

Evolutionary (Convolutional) Neural Networks (ENN) is seen as the most basic architecture among deep learning architectures [10]. These are structures that are used in multi-layered sensors and generally in 2D structures such as visual elements [24]. The most frequently used method of deep learning methods used in cancer diagnosis is Evolutionary (Convolutional) Neural Networks. The architectures of convolutional neural networks have an important place in cancer diagnosis. CNN models are being developed for cancer diagnosis. The methods used for cancer diagnosis of the developed CNN models; AlexNet, GoogleNet, VGGNet, ResNet are encountered [31].

Repetitive Neural Networks (RNN) also includes the ability to analyze the data flow of RNN, which has the ability to model the time dependence and memorization ability of the sequential events that occur [13]. The basic system used in RNN uses sequential data and RNN creates a directed loop of existing connections between units within the system [24].

Deep Boltzman Machine (DBM) are random neural networks that learn probability distribution logic in input data. It is at the forefront as fast learning algorithm [24]. Undirected connections occur between layers in the network [13].

Long-Short-Term Memory Networks (LSTM), this system used before the discovery of RNN architectures is a knowledge-using based system [24]. LSTM idea was found to eliminate the problems caused by the gaps between time contexts in RNN. There are memory cells in LSTM and it is the network structure used to calculate hidden situations in RNN [7].

Deep Faith Networks (DFN) is defined as the data stack/storage of long-short term memory networks [24]. Only the first two layers have non-directional connectivity and therefore the network can be trained both supervised and unsupervised [13]. These are graphical structures that will represent the data set in the network structure consisting of hierarchical binding of the layers in the restricted boltzman machine [7].

Deep Auto-Encoders (DAE), which is used for unsupervised learning purposes, is also called Diabolo [24]. Auto-encoders used for educational purposes are used in different types to represent a healthy representation (Fig. 13.1). Some developed auto-encoders; They appear with the names of sparse (intermittent) auto-encoders, noise-clearing auto-encoders, auto-encoders with pulling power, and convective auto-encoders [13]. DAE, which has been in neural networks for a long time, started to be mentioned in this field with the formation of deep learning architectures [7].

As frequently used models in deep learning field, it emerges us; AlexNet, ZF Net, GoogleNet and Microsoft RestNet. AlexNet is used for pattern and visual identification purposes and is a 11% healthier and more reliable model in terms of reducing the error rate from 26 to 15% [7]. With this computerized object identification system, architecture can classify 100 objects. The matrix structure is 11×11 [10]. The ZF Net algorithm was obtained with the development of AlexNet architectural structure. ZF Net reduces the error rate by 11.2%, and a visualization technique called

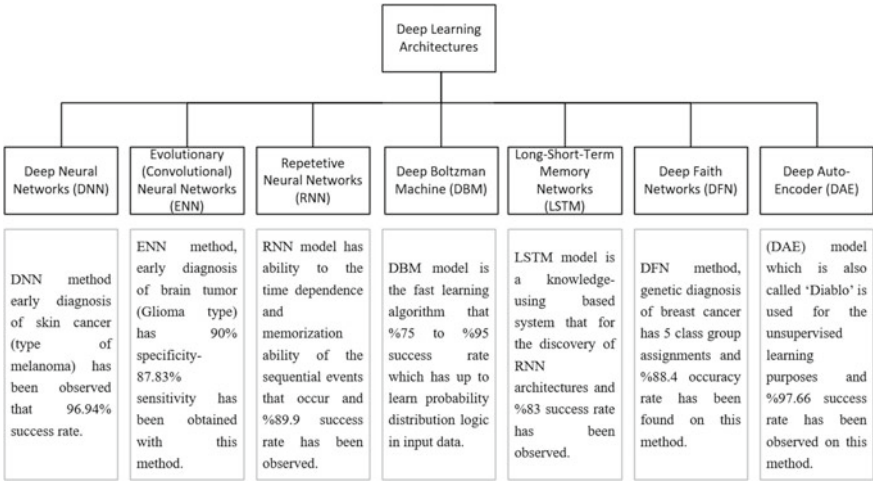


Fig. 13.1 Deep learning methods for diagnosis

Deconvolutional Network has been developed and a different dimension has been developed in this way [7, 32]. The 11 × 11 matrix structure on AlexNet has been made 7X7. With this change, the small filter size in the convolution layer provides protection of the original pixel information contained in the input size [10].

GoogleNet has a complex structure, the error rate of this system is 5.7%. The 144-layer structure filters different dimensions through Inception module. Thanks to this filtering, size reduction can be made [7]. In this structure Microsoft RestNet with an error rate of 3.6%, the number of layers is higher than other structures. There are two RELUs, a linear layer and a one-time residual value-enhancing block in this structure [7]. Microsoft RestNet architecture has a structure consisting of Residual blocks [10].

13.3 Deep Learning in Cancer Diagnosis

Deep learning methods can be used for different purposes in the field of health. For these purposes, gene diversity, drug design, interaction of proteins, interactions of proteins with DNA and RNA, cancer diagnosis, etc. Such situations can be counted [13]. Deep learning architectures in systems generally used in bioinformatics vary according to cancer types. The reason for the changes observed in the architectures used are the data sets used and the reliability rates obtained from the methods used. The studies and the methods used in the studies below are included.

Xu et al. [30] uses multi-model deep learning methods in their study for the diagnosis of Cervical Dysplasia. They use the convolutional (neural networks) method to convert video and non-image data into combinable feature vector. The multi-model

structure proposed and used in the study contains a large dataset in the diagnosis of cervical dysplasia. Compared to other methods, ENN method has been found to be more successful and 90% specificity-87.83% sensitivity has been obtained with this method.

Nie et al. [22] used three-dimensional ENN method for early diagnosis of brain tumor (Glioma type). They also propose a new network architecture that works with consulting learning and large datasets. In the study, a training is given for the patient's survival time. In this method 89.90% accuracy value was determined.

Ibrahim et al. [9] provided 9% more success in the diagnosis of liver cancer, 6% in lung cancer and 10% in breast cancer (in F1 measurements) as a result of the experimental studies they performed instead of the classical methods (HCC).

Khademi and Nedialkov [15] conducted a study for the genetic diagnosis of breast cancer in their study. Within the scope of this study, they use Deep Faith Networks (DFN) which are deep learning methods to complete the missing features and extract the desired features from microarrayed data. They applied the structure learning algorithm of clinical data and as a result, the clinical model can be automatically extracted. The resulting model was integrated with softmax nodes and achieved more promising results.

Ahmed et al. [1] aim to early diagnosis and treatment of breast cancer in their study for the diagnosis of breast cancer. Within the scope of the study, they develop methodology based on deep convolutional neural systems. As a result of the methodology, they made 5-class group assignments and reached an accuracy rate of 88.4%.

Yildiz [31] proposes a system that enables early diagnosis of skin cancer (type of melanoma) and deep diagnosis and machine learning methods to be treated in his study titled "Detection of melanoma from dermoscopy images with deep learning methods". In the experimental study, which is used the deep neural networks C4Net model, it was observed that 96.94% success rate in the diagnosing of skin cancer.

Kadiroglu et al. [12], types of AlexNet and VGG16 trained before the convolutional neural networks, were used for early diagnosis of skin cancer (type of lesion), which is frequently encountered using deep attributes and colors. A total of 200 images were used within the scope of the study and these images were tested in AlexNet fc6, AlexNet fc7, VGG16 fc6 and VGG16 fc7 structures. Accuracy rates are 92%, 94%, 94%, 94%, respectively. In the total evaluation of the data, the accuracy rate was found 98%.

Varol and Iseri [27] used 135 digital images in their study for the diagnosis of lymph cancer by making use of the image processing structures of deep learning. In an experimental study with pre-trained algorithms, three types of lymph cancer (CLL, FL and MCL) structures were examined. As a result of experiments using though image processing structures, an accuracy rate of 89.72% was succeed.

Ismael et al. [11] used the deep learning method for brain cancer diagnosis. In their study, using a network system in the deep learning, three types of brain tumors (meningiomas, Gliomas and Pituitary), was evaluated. Compared to other studies, better results were obtained with the data set that the study was conducted. In this study, the used model has reached the highest accuracy rate of 99%.

Zhang et al. [33] proposed that image-based processing method in order to identify early skin cancer. In their study, they use the Convolution Neural Network (CNN) architecture. CNN method was tried to be optimized with the thrived "whale optimization". To achieve their goals, they used two different methods on two data sets. As a result of the simulations, the proposed method has obtained better results in percentage compared to other methods.

Nayak et al. [21] conducted a study for the brain abnormalities (discrepancy in abilities and skills) and diagnosis of breast cancer using the auto-coding architecture of deep learning. In this study they proposed "Random Vector Functional Link (RVFL) based Autoencoder (SRVFL-AE)" neural network structures. SRVFL-AE which they used for RVFL structure, forms the basis of the system. The reason for the opt this structure, customary auto-encoder based structure methods, allow comparisons to be made and increase the speed of learning. To compare proposal SRVFL-A method, two brain data sets (MD-1 and MD-2) were examined. Analysis showed that, MD-1 96.67% and MD-2 95% accuracy rate were observed. The same method, was also tested in breast cancer dataset. As a result of the tests, shorter training period and better success were achieved thanks to the proposed method.

Arefan et al. [2] in their study on breast cancer diagnosis, they used digital mammograms of the scanned traditional scans as datasets. After these scans, aimed at speeding up the diagnosis of breast cancer with deep learning methods, they compared the two structures in deep learning (GoogleNet and GoogleNet-LDA). In the study it concluded that, GoogleNet-LDA perform better results than GoogleNet. It should be emphasized that larger data studies should be conducted by using deep learning method for breast cancer risk assessment.

Coccia [5] in his study, emphasizes of using deep learning method in cancer imaging. Thanks to deep learning technology, pathologist's highlights that "helps to detect and classify the stage of cancer". Diagnosing of cancer type with deep learning methods it is emphasized that the correct detection rate is better. At the same time deep learning technology for cancer imaging, shows the resource spread point (the source of the cancer). Beside this advantage deep learning, is emphasized that it has socioeconomic benefits and supports filling the current gap in the health sector.

13.4 Cancer Diagnostication and Treatment Proposal with Deep Learning Methods in Future

In the previous sections, cancer diagnosis and use of deep learning methods in cancer diagnosis are mentioned. This section includes that using deep learning for cancer diagnosis, treatment methods and the way of direction to works that can be made, forecasts in the future. Information which will be described in this section, to guide the deep learning studies in the future and cancer diagnosis/treatment expectations are included.

Working in the field of health with deep learning provides great convenience. It is especially important in the diagnosis of cancer, which is tackle within the scope of the study. However many difficulties are encountered when using deep learning in the field of health. Difficulties Encountered [23];

- "*Data shortage*" Data access due to ethical concerns and various rules/regulations.
- "*Information sharing*" Cooperation between relevant institutions is time consuming or difficult to obtain the necessary data for the studies.
- "*Computing skills*" Failure to create a team of artificial intelligence (data scientist, developer and solution architect) experts to fill the gap created by other software with deep learning study.
- "*Deep learning skills*"; The team, which will be formed by artificial intelligence experts, must be knowledgeable both in AI and in the field of health. Because, information such as error status, internal/external factors, and/or insufficiency of the conditions in the data obtained as a result of the operations made can be obtained from experts.

The situation observed in the studies conducted that cancer types can be diagnosed with deep learning. In our study it includes that which cancer types is determined with the deep learning architecture by processing datasets. The studies to be carried out in future, expectations from deep learning for cancer diagnosis and treatment suggestions are higher than other study methods. Beyond the cancer diagnosing, deep learning should now be included in the treatment.

Today's world health and medical sectors are concentrated on visualization and scanning technologies. The number of people whose waiting for diagnosis with an increase in the patient profile is directly proportional to the order of cancer scanning. Modern imaging devices such as NMRI (Nuclear Magnetic Resonance Imaging), BT (Computed Tomography), PET (Positron Emission Tomography) have harmful ɣ Gamma radiation and radiation values that are reflected in the body compared to the old conventional imaging devices. Transition from single detector conventional imaging devices (Fig. 13.2) to complex modern imaging devices with a minimum of 8- to 256 detectors;

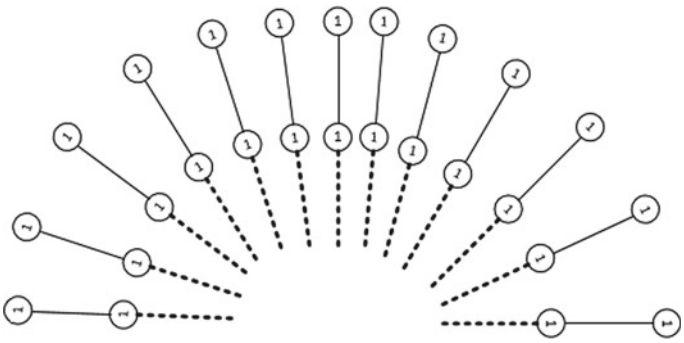


Fig. 13.2 Conventional imaging devices

- Close-up time to approximately 1 h to 45 s.
- Reduced high radiation levels to the amount of radiation exposed during a long-range flight (Istanbul-London).
- Increase high electricity consumption levels to a large volume electric heater level.
- Patient movement originated shooting errors are minimized.
- Hardware sustainability and disruption.

By minimizing those rates, it has led to a major leap forward in its intended production target.

Besides going to the numerical increase of imaging devices and technological innovations, most soft tissue cancers and some derivatives, the conventional diagnostic methods for the popularity continues. In line with humanitarian needs such as intense work tempo, sleeplessness and morale, the data obtained from scanning device with diagnose interpreted by human fatigue and this all carried out by the doctors.

In the interpretation and acquisition of data section, a combination of information of radiologists and Artificial Intelligence is observed. Radiologists take the necessary shots through modern imaging devices (Fig. 13.3) and imaging results are sent to the relevant doctors computer [28].

In new studies, after diagnosing with deep learning; it is expected to include information about the treatment methods and/or medicine schedule that can be cure to the patient. After the necessary tests are made for the individual and the diagnosis is confirmed; it should provide a report either suggestion on the path to be followed. This report will include information such as the type of treatment recommendations (irradiation, chemotherapy, medication, etc.).

The system which is expected to emerge, thanks to the relationship between deep learning and cancer diagnosis/treatment, stages from diagnosis to treatment of the disease are as follows:

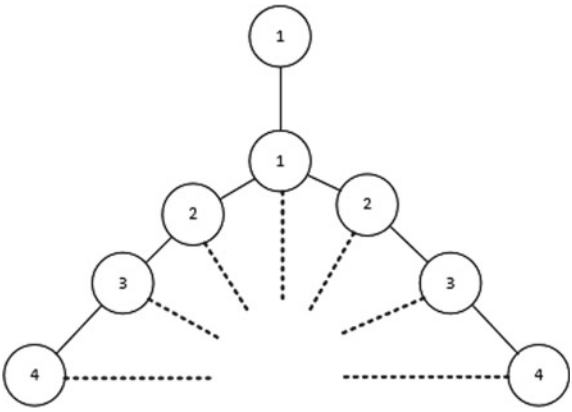


Fig. 13.3 Modern imaging devices

1. The dataset which is obtained from the individuals, whose disease has been diagnosed should be analyzed in deep learning architectures, and the diagnosis of cancer and the type of cancer should be learned. The part up to this point is still carried out today.
2. Diagnostic proses and cancer derivatives, passed through again from the deep learning architectures which can be implemented method of treatment (radiation therapy, chemotherapy, medication etc.) should be listed. At the same time, the report which includes deep learning and treatment method recommendations, it should also contain the steps to be follow during the treatment. The observation of result, the drug to be given the chosen treatment which is chemotherapy "How much CC should be applied to patient, how often (Treatment Schedule) and which steps to follow on the phases" that report, should have offer these information's to aimed the expectations If radiotherapy is to be applied the patient, it is also expected to provide a suggestion for the apply region and duration to be irradiated.

If these studies can be carried out, it is possible to list the advantages of this situation as follows.

1. Diagnostics for the decision for treatment, which saves time will be lost.
2. Saves time lost in the decision-making process on which method will be implemented.
3. In current situation, in case of negative result of the chosen treatment method; the patient's body to unnecessary drug/radiation installation's disadvantage to get free from.
4. Receiving peer interpretations for the treatment to be applied and preventing time loss in the procedural process.

The use of deep learning method in cancer diagnosis and classification, will prevent loss of time until passed to the treatment process consequently it is estimated that it will lead to the treatment of the disease much faster.

13.5 Conclusions

Cancer is a disease that has a huge impact on human life. It is known that the diagnosis and treatment of cancer is a long and difficult process. Studies are carried out to shorten the long and difficult process experienced in the diagnostic process as much as possible. Artificial intelligence are machines that mimic human behaviors. In the cancer diagnosis process, deep learning is used from the sub-fields of artificial intelligence. Deep learning are machines that enable the solving of many and complex data problems of artificial neural networks produced by comparing them to the neural networks of the brain. Although deep learning is used with many fields such as industry also it has been frequently used in medical field since 2014. Deep learning has been used in cancer diagnosis in recent years.

The use of deep learning methods varies according to cancer types. When the data in the previous section is examined, it is seen that different deep learning methods

are used for each type of cancer. Thanks to the studies conducted, the accuracy rates are also emphasized in order to know which method should be used in the diagnosis of the cancer type. In this study, deep learning methods, architectures and deep learning methods used in diagnosis of cancer types are emphasized. In this study, Recent studies are highlighted and current dates are evaluated. It is thought that the study will lead future studies in terms of the current informations and the in-depth examination of the field. It is also important to work with the feature of directing the work to be done.

References

1. S.R.A. Ahmed, O.N. Uçan, A.D. Duru, O. ve Bayat, Breast cancer detection and image evaluation using augmented deep convolutional neural networks. Aurum Mühendislik Sistemleri Ve Mimarlık Dergisi **2**(2), 121–129 (2018)
2. D. Arefan, A.A. Mohamed, W.A. Berg, M.L. Zuley, J.H. Sumkin, S. Wu, Deep learning modeling using normal mammograms for predicting breast cancer risk. Med. Phys. **47**(1), 110–118 (2020)
3. Y.A. Ayturan, Airborne Particulate Matter Concentration Estimation with Deep Learning (In Turkish). (*Master Thesis*). Institute of Science. Konya Karatay University. Konya (2019)
4. U. Bayraktar, *Cancer Cell Detection Based on Deep Learning* (2018). (In Turkish). https://www.researchgate.net/publication/334151448
5. M. Coccia, Deep learning technology for improving cancer care in society: new directions in cancer imaging driven by artificial intelligence. Technol. Soc. **60**, 101198 (2020)
6. Ö. Deperlioğlu, Diagnosis of Hepatitis Disease with Deep Neural Networks, in *ECSAC'19 IVth European Conference on Science, Art Culture.* (2019), pp. 467–473. (In Turkish)
7. F. Doğan, İ. ve Türkoğlu, A compilation of deep learning models and application areas. DÜMF Eng. J. **10**(2), 409–445 (2019). (In Turkish)
8. L.B. Holder, M.M. Haque, M.K. Skinner, Machine learning for epigenetics and future medical applications. Epigenetics **12**(7), 505–514 (2017)
9. R. Ibrahim, N.A. Yousri, M.A. Ismail, N.M. El-Makky, Multi-level gene/MiRNA Feature Selection Using Deep Belief Nets and Active Learning, in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.* IEEE (2014), pp. 3957–3960
10. Ö. Inik, E. ve Ülker, Deep learning models used in deep learning and image analysis. Gaziosmanpaşa Sci. Res. J. **6**(3), 85–104 (2017). (In Turkish)
11. S.A.A. Ismael, A. Mohammed, H. Hefny, An enhanced deep learning approach for brain cancer MRI images classification using residual networks. Artif. Intell. Med. **102**, 101779 (2020)
12. Z. Kadiroğlu, B.N. Akılotu, A. Şengür, M. ve Kayaoğlu, Classification of dermoscopic images using deep features and color (2019). (In Turkish)
13. U. Kaya, A. Yılmaz, Y. ve Dikmen, Deep learning methods used in health. Eur. J. Sci. Technol. **16**, 792–808 (2019). (In Turkish)
14. İ. Keser, K. Özdemir, B. Ertürk, M. Haspolat, T. Duman, M. ve Esmer, Analysis of the services offered at the oncological rehabilitation unit for cancer patients. Gazi J. Health Sci. **2**(1), 18–27 (2017). (In Turkish)
15. M. Khademi, N.S. Nedialkov, Probabilistic graphical models and deep belief networks for prognosis of breast cancer, in *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA).* IEEE (2015), pp. 727–732
16. U. Kose, Developing a human controlled model for safe artificial intelligence systems. J. Eng. Sci. Des. **6**(1), 93–107 (2018). (In Turkish)

17. U. Kose, Artificial intelligence and cyber wars of the future. TÜBİTAK Sci. Tech. Maga. **52**(618), 76–84 (2018). (In Turkish)
18. U. Kose, Artificial intelligence and the future: should we worry? Sci. Utopia. **24**(284), 39–44 (2018). (In Turkish)
19. U. Kose, Artificial intelligence: paradoxes in the science of the future. Popular Sci. Mag. **25**(261), 12–21 (2018)
20. K. Kourou, T.P. Exarchos, K.P. Exarchos, M.V. Karamouzis, D.I. Fotiadis, Machine learning applications in cancer prognosis and prediction. Comput. Struct. Biotechnol. J. **13**, 8–17 (2015)
21. D.R. Nayak, R. Dash, B. Majhi, R.B. Pachori, Y. Zhang, A deep stacked random vector functional link network autoencoder for diagnosis of brain abnormalities and breast cancer. Biomed. Signal Process. Control **58**, 101860 (2020)
22. D. Nie, H. Zhang, E. Adeli, L. Liu, D. Shen, 3D Deep Learning for Multi-Modal Imaging-Guided Survival Time Prediction of Brain Tumor Patients, in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer, Cham, 2016), pp. 212–220
23. M.K. Saraf, P. Mike, *Deep Learning for Cancer Diagnosis: A Bright Future* (2018). Online https://software.intel.com/en-us/articles/deep-learning-for-cancer-diagnosis-a-bright-futureadresinden. Retrieved 22 Mar 2020
24. A. Şeker, B. Diri, H.H. ve Balık, A study on deep learning methods and applications. Gazi J. Eng. Sci. **3**(3), 47–64 (2017). (In Turkish)
25. Sencard.com.tr, Importance of early diagnosis in cancer (2019). (In Turkish). Online https://www.sencard.com.tr/kanserde-erken-teshisadresinden. Retrieved 19 Dec 2019
26. M. Toğaçar, B. ve Ergen, Comparison of existing methods with deep learning in biomedical images. Fırat Univ. J. Eng. Sci. **31**(1), 109–121 (2019). (In Turkish)
27. A.B. Varol, İ. İşeri, Classification of pathology images of lymph cancer by machine learning methods. Eur. J. Sci. Technol. 404–410 (2019). (In Turkish)
28. J. Vincent, Why cancer-spotting AI needs to be handled with care (2020). Online https://www.theverge.com/2020/1/27/21080253/ai-cancer-diagnosis-dangers-mammography-google-paper-accuracy. Retrieved 22 Mar 2020
29. WHO. Cancer (2019). Online https://www.who.int/health-topics/cancer#tab=tab_1. Retrieved 18 Dec 2019
30. T. Xu, H. Zhang, X., Huang S. Zhang, D.N. Metaxas, Multimodal deep learning for cervical dysplasia diagnosis, in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer, Cham, 2016), pp. 115–123
31. O. Yildiz, Detection of melanoma from dermoscopy images with deep learning methods: comprehensive study. Gazi Univ. Faculty Eng. Architecture J. **34**(4), 2241–2260 (2019). (In Turkish)
32. M.D. Zeiler, R. Fergus, Stochastic pooling for regularization of deep convolutional neural networks (2013). arXiv preprint arXiv:1301.3557
33. N. Zhang, Y.X. Cai, Y.Y. Wang, Y.T. Tian, X.L. Wang, B. Badami, Skin cancer diagnosis based on optimized convolutional neural network. Artif. Intell. Med. **102**, 101756 (2020)

Chapter 14
Brain Tumor Segmentation Using 2D-UNET Convolutional Neural Network

Khushboo Munir, Fabrizio Frezza, and Antonello Rizzi

Abstract Gliomas are considered as the most aggressive and commonly found type among brain tumors. This leads to the shortage of lives of oncological patients. These tumors are mostly by magnetic resonance imaging (MRI) from which the segmentation becomes a big problem because of the large structural and spatial variability. In this study, we propose a 2D-UNET model based on convolutional neural networks (CNN). The model is trained, validated and tested on BRATS 2019 dataset. The average dice coefficient achieved is 0.9694.

Keywords Deep learning · Deep UNET · Brain tumor segmentation · Artificial intelligence · Convolutional neural network

14.1 Introduction

Brain tumors are one of the most lethal types of cancers. Studies have shown that magnetic resonance imaging can be effectively used for the detection of the brain tumors [1–3]. Multi-modal MRI protocols have been used for the evaluation of brain tumor vascularity, cellularity, and blood-brain barrier integrity. This is because of the crucial complementary information provided by the different contrasts of these multimodal images. The most generally used MRI protocols are T1-weighted, T2-Weighted, T1-weighted with gadolinium enhancement (T1CE) and Fluid-Attenuated Inversion Recovery (FLAIR) resulting in valuable diagnosis [4]. One of the most wanted and critical problems in brain tumor diagnosis is its segmentation because of the fact

K. Munir (✉) · F. Frezza · A. Rizzi
Department of Information Engineering, Electronics and Telecommunications (DIET), Sapienza University of Rome, Rome, Italy
e-mail: khushboo.munir@uniroma1.it

F. Frezza
e-mail: fabrizio.frezza@uniroma1.it

A. Rizzi
e-mail: antonello.rizzi@uniroma1.it

U. Kose and J. Alzubi (eds.), *Deep Learning for Cancer Diagnosis*,
Studies in Computational Intelligence 908,
https://doi.org/10.1007/978-981-15-6321-8_14

then correct segmentation can eliminate the confounding structures from tissues of the brain and can help in the classification of different types of tumors. The main challenges of brain tumor are due to high variations in shape, size, location, regularity and heterogeneous appearance [5]. Studies related to brain tumor segmentation are categorized into supervised learning based [6–8], and unsupervised learning based [9–11]. Combining region-growing and fuzzy clustering for brain tumor scans by T1-weighted and T2-weighted was done by Hsieh et al. [12] and achieved the accuracy of segmentation up to 73%. Brain tumor segmentation using multi-stage fuzzy c-means was implemented by using multi-model MRI scans but it was tested on a very limited dataset achieving promising results. Under the conditional random field framework, superpixel features were extracted by Wu et al. [13] but the obtained results were varied from patient to patient while under performing for the images of LGG.

Supervised deep convolutional neural networks are used to extract complex features directly from data [14]. Using BraTS benchmark deep learning got itself on the top of all other methods of brain tumor segmentation [15, 16]. Apart from the developments in deep learning, the main challenges relating to brain tumors are first the brain tumor segmentation which is considered as an abnormal region detection problem and is relatively more difficult than the classification task. Secondly if the network can give high performance for HGG it can still perform poor for the LGG. Thirdly, comparing the segmentation of complete tumor the delineation of enhanced infiltrative regions and core tumor regions is still not well performed, particularly when considering other important features of the trained model, such as memory footprint and processing efficiency. This study presents a 2D Unet convolutional neural network for the segmentation of brain tumor using multi-modal 3D MRI volumes form BraTS 2019 archive [17–20]. In order to increase the performance of the system data augmentation technique is used. The proposed method has performed well on both HGG and LGG. The results are comparable to the state-of-the-art methods for the segmentation of brain tumors.

14.2 Method

14.2.1 Convolution Neural Networks

Convolutional neural networks are considered to be the current state of the art for the problem of image recognition in the machine learning domain [21, 22], in particular when dealing with biomedical MRI [23]. They are formed by a sequence of typical layers, convolution, pooling and flatten ones. In the convolution layer filter size and number of filters are defined. A moving dot product is performed between a fixed defined $K \times K$ filter against subpictures of the original image. The output of this dot product is finally feeded into a nonlinear activation function. The most famously used activation functions are rectified linear unit and sigmoid function. Both activation

functions are differentiable and can fit during the training into the framework of standard gradient descent learning techniques for neural networks.

When a $k \times k$ kernel is convolved with a sub-picture of $p \times p$ the resultant image will be of size $(p - k + 1) \times (p - k + 1)$. Usually, a convolutional layer is followed by a pooling layer in order to reduce the dimensionality to make the training of the CNN network a little bit easier.

14.2.2 Convolutional U-Network

After the application of series of a convolutional filters, the dimension of the final layer is usually a way much smaller than that of the original image provided to the network. In order to tackle the problem of checking whether the selected pixel of the input image is a part of a lesion or not, the output of the network must be of the same dimensions as the input was. Ciresan et al. [24] tried to solve this dimensionality problem by considering each input image pixel and a small localized region around it instead of the entire image into the neural network. Whereas Ronneberger et al. [17] presented a powerful solution of a U-Net CNN. It basically solves the problem by deconvolving the layer performing an upsampling procedure to increase image dimensionality.

14.2.3 Data Acquisition and Pre-Processing

The method proposed in this paper is trained, validated and predicted on BraTs 2019 dataset [25] which contains High-Grade Gliomas (HGG) and Low-Grade Gliomas (LGG) patient scans. The dataset consists of four types of modalities: these are T1-weighted, T2-weighted, T1-weighted with gadolinium-enhancing contrast (T1CE) and FLAIR. T1, T2, and FLAIR were co-registered into T1CE dataset which was the resampled and interpolated into $1 \times 1 \times 1\text{mm}^3$, corresponding to image size in pixels equal to image as $240 \times 240 \times 155$. Data is normalized by subtracting each MRI modality with its mean and dividing it by its standard deviation. The images dataset provided by the BraTS are 3D volumes, each represented with four modalities. This dataset is divided into six sets for HGG and two sets of LGG. Each set consists of 40 volumes. The volumes are shown in Fig. 14.1.

The first step in the preprocessing is the conversion of these 3D volumes into 2D slices. Among these slices there exist the slices which do not contain any useful information thereby causing overfitting problem. In order to reduce the overfitting, the slices containing low or no information are discarded. On the remaining slices, data augmentation will be performed. The basic purpose of performing data augmentation is to increase the training data in order to improve network performance. In doing so rotation, flipping and shifting are applied to the slices obtained in the previous step.

14.2.4 Proposed CNN Based on UNET Architecture

Biomedical data images contain detailed patterns of the brain tumor with variable edge boundary definitions of tumor objects. Long et al. [26] proposed an architecture based on skip connections which have the basic function of combining decoding layer high-level representation with the encoding layer shallow appearance representation. The method provided good results on natural images and can be applied to the biomedical images [26, 27]. Combining this skip-architecture with the unet-architecture Ronneberger et al. [17] solved the problem of cell tracking. The architecture which we are proposing consists of two distinct sub-structures, an encoding part (downsampling) and decoding one (upsampling).

The encoding part consists of 5 convolutional blocks where each block contains two 2D convolutional layers with kernel size of 3×3, with same padding and rectified linear unit activation. In the third layer batch normalization is applied. The fourth layer is composed of 2D max pooling. After the first two convolutional blocks a drop-out layer of 0.2 is applied. The last encoding block consists of only one convolutional layer. Decoding part of each block consists of a deconvolutional layer with kernels of size 3×3, stride (2, 2), same padding and relu activation. This first layer of upsampling is concatenated with the batch normalization layer of encoding block therefore named as a merge block. The third layer performs 2D convolution, therefore, a decoding block consists of these three layers. The four decoding blocks follows the same pattern except the fact that merging layer of the first decoding block concatenates with normalization layer of fourth encoding block, merging layer of the second encoding block concatenates with the normalization layer of third encoding block, merging layer of third decoding block concatenates with the normalization layer of the second encoding block and the merging layer of fourth decoding block concatenates with normalization layer of first encoding layer. Drop out of 0.2 is applied after the first two decoding blocks. After the application of four decoding blocks, a convolutional layer is applied with 3×3 kernel, same padding, and relu activation. The final layer is an output layer consisting of 2D convolution with softmax activation.

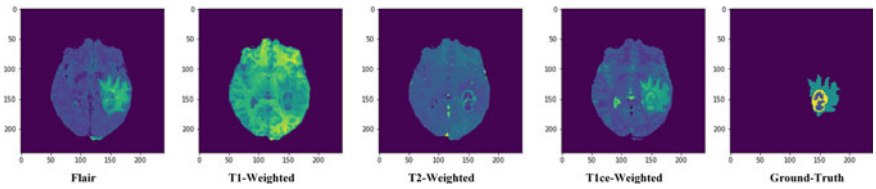


Fig. 14.1 Examples of MRI images in the considered dataset

14.2.5 Network Training

For the training, the used cost function is soft dice metric as given in [28] rather than using a quadratic cost function or cross-entropy. Stochastic gradient-based optimization is required for the training of a deep neural network, therefore, minimizing the cost function with respect to network parameters. To estimate the parameters adaptive moment estimator known as Adam is used [29]. Generally, first and second gradient moments are used by the adam to update and correct current gradients moving average. Learning rate of Adam optimizer is set to $1e^{(-5)}$, the batch size is set to 100 and the number of epochs are 50. As the dataset was divided into six sets to reduce the computational complexity and cost reduction the network is allowed to train on each preprocessed set one by one.

14.2.6 Experiment, Results and Discussion

The evaluation has been done using a set by set approach for both HGG and LGG data respectively. Dice coefficient is used for the evaluation of the tumor region. This study proposed a U-Net convolutional neural network for the segmentation of brain tumor. Essentially the problem of tumor detection and segmentation belongs to semantic segmentation. A comprehensive technique of data augmentation is used. Dice coefficient and the loss tested for sets are given in the Fig. 14.2. The average dice coefficient achieved by the proposed network is 0.9694 which is a very much promising result as compared to the pre-made neural networks. Figure 14.3 shows the segmentation results comparison with the ground truth for HGG and Fig. 14.4 shows the segmentation results for LGG. The network is implemented by using the TensoFlow and Keras libraries. The network is processed by using the Tesla K80 google collab Graphic Processing unit with 25 GB memory. The basic problem in training the network with 2D images or 3D volumes is because of the memory limits as a 2D network can be processed with full slices at once, whereas only one patch

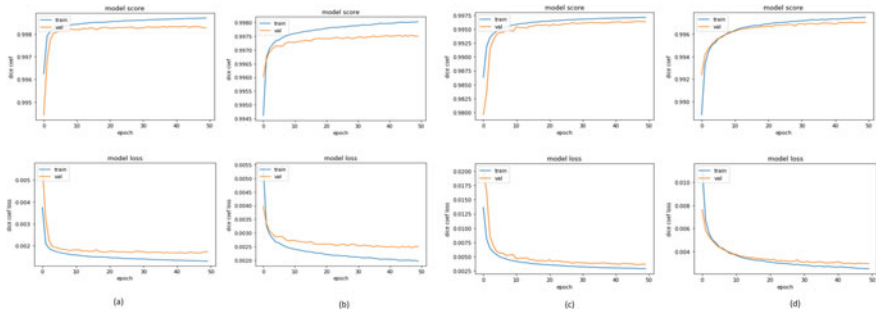


Fig. 14.2 Dice coefficient and dice coefficient loss

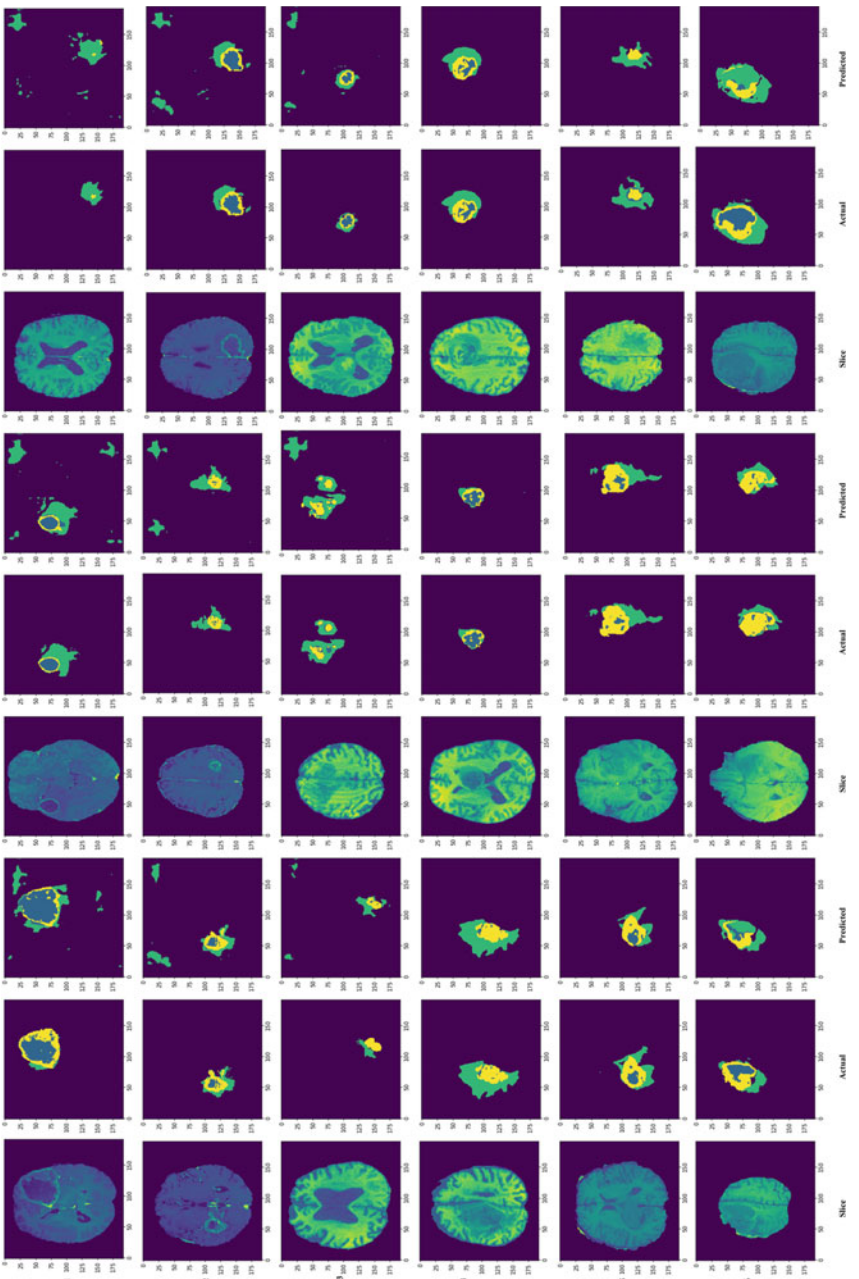


Fig. 14.3 Segmentation results for the sets of HGG

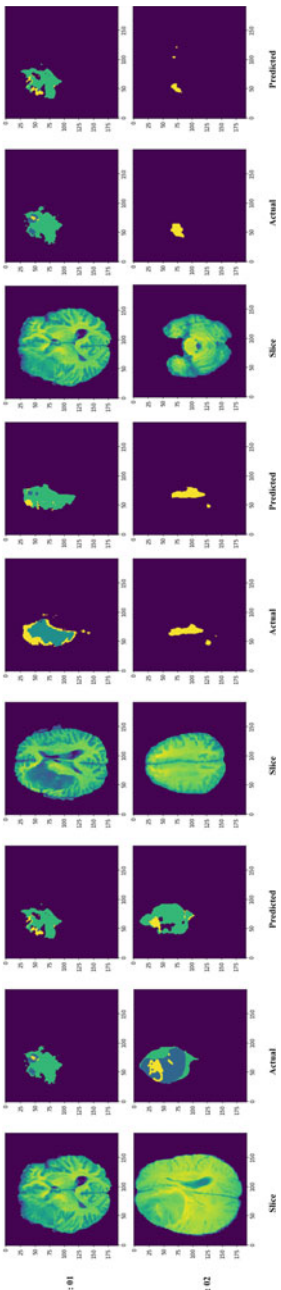


Fig. 14.4 Segmentation results for the sets of LGG

of 3D volume can be processed by the 3D system at the same time. The processing time of 1 epoch is 12 seconds which is more time-efficient as compared to previous studies such as [16] which took 30 seconds, [15] which consumed approximately from 25 s to 3 min and [14] which took 8 min for prediction of one tumor study. There are some limitations of the network we proposed, since there is a possibility to perform evaluation through cross-validation scheme by running separate different datasets that may provide different evaluation results. Secondly, some parameters in the network have to be set carefully. Apart form these limitations our network gave very efficient results for brain tumor segmentation. Further study by taking different scenarios into account can be done and this network can be made more efficient.

14.3 Conclusion and Future Work

A fully automatic method of brain tumor segmentation is presented in this paper using a 2D U-Net based deep convolutional network. The network was trained and tested on the well-known BraTS 2019 benchmark containing both HGG and LGG patients scans. Our method provided efficient and robust segmentation which was comparable to the delineated ground truth. Comparing to other state of the art methods our algorithm achieved quite promising results. This network can also be tested for different datasets in the future and can be further improved by the introduction of different other data augmentation methods and noise addition. However, this proposed method makes it possible to segment the tumor region without a manual interface therefore allowing to perform automatically the diagnostic task, treatment planning, and patient monitoring.

References

1. G.P. Mazzara, R.P. Velthuizen, J.L. Pearlman, H.M. Greenberg, H. Wagner, Brain tumor target volume determination for radiation treatment planning through automated MRI segmentation. Int. J. Radiat. Oncol. Biol. Phys **59**(1), 300–312 (2004)
2. T. Yamahara, Y. Numa, T. Oishi, T. Kawaguchi, T. Seno, A. Asai, Keiji Kawamoto, Morphological and flow cytometric analysis of cell infiltration in glioblastoma: a comparison of autopsy brain and neuroimaging. Brain Tumor Pathol. **27**(2), 81–87 (2010)
3. S. Bauer, R. Wiest, L.-P. Nolte, M. Reyes, A survey of mri-based medical image analysis for brain tumor studies. Phys. Med. Biol. **58**(13), R97 (2013)
4. T.L. Jones, T.J. Byrnes, G. Yang, F.A. Howe, B.A. Bell, T.R. Barrick, Brain tumor classification using the diffusion tensor image segmentation (D-SEG) technique. Neuro-oncology **17**(3), 466–476 (2015)
5. M. Soltaninejad, G. Yang, T. Lambrou, N. Allinson, T.L. Jones, T.R. Barrick, F.A. Howe, X. Ye, Automated brain tumour detection and segmentation using superpixel-based extremely randomized trees in FLAIR MRI. Int. J. Comput. Assist. Radiol. Surg. **12**(2), 183–203 (2017)
6. P.A. Mei, C.C. de Carvalho, S.J. Fraser, L.L. Min, F. Reis, Analysis of neoplastic lesions in magnetic resonance imaging using self-organizing maps. J. Neurol. Sci. **359**(1–2), 78–83 (2015)

7. J. Juan-Albarracin, E. Fuster-Garcia, J.V. Manjon, M. Robles, F. Aparici, L. Martí-Bonmatí, J.M. Garcia-Gomez, Automated glioblastoma segmentation based on a multiparametric structured unsupervised classification. PLoS One **10**(5), e0125143 (2015)
8. A. Rajendran, R. Dhanasekaran, Fuzzy clustering and deformable model for tumor segmentation on mri brain image: a combined approach. Procedia Eng. **30**, 327–333 (2012)
9. M. Jafari, S. Kasaei, Automatic brain tissue detection in mri images using seeded region growing segmentation and neural network classification. Aust. J. Bas. Appl. Sci. **5**(8), 1066–1079 (2011)
10. M. Goetz, C. Weber, J. Bloecher, B. Stieltjes, H.-P. Meinzer, K. Maier-Hein (2014) Extremely randomized trees based brain tumor segmentation, in *Proceeding of BRATS challenge-MICCAI*, pp. 006–011
11. N. Subbanna, D. Precup, T. Arbel, Iterative multilevel MRF leveraging context and voxel information for brain tumour segmentation in MRI, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2014), pp. 400–405
12. T.M. Hsieh, Y.-M. Liu, C.-C. Liao, F. Xiao, I.-J. Chiang, J.-M. Wong, Automatic segmentation of meningioma from non-contrasted brain MRI integrating fuzzy clustering and region growing. BMC Med. Inf. Decis. Making **11**(1), 54 (2011)
13. W. Wu, A.Y. Chen, L. Zhao, J.J. Corso, Brain tumor detection and segmentation in a CRF (conditional random fields) framework with pixel-pairwise affinity and superpixel-level features. Int. J. Comput. Assist. Radiol. Surg. **9**(2), 241–53 (2014)
14. S. Pereira, A. Pinto, V. Alves, C.A. Silva, Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Trans. Med. Imag. **35**(5), 1240–1251 (2016)
15. M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, H. Larochelle, Brain tumor segmentation with deep neural networks. Med. Image Anal. **35**, 18–31 (2017)
16. K. Kamnitsas, C. Ledig, V.F. Newcombe, J.P. Simpson, A.D. Kane, D.K. Menon, D. Rueckert, B. Glocker, Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med. Image Anal. **1**(36), 61–78 (2017)
17. O. Ronneberger, P. Fischer, T. Brox. U-net: convolutional networks for biomedical image segmentation, in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer, 2015), pp. 234–241
18. B.H. Menze et al., The multimodal brain tumor image segmentation benchmark (brats). IEEE Trans. Med. Imag. **34**(10), 1993–2024 (2015)
19. H. Akbari, A. Sotiras, S. Bakas et al., Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features, in *ci Data 4, 170117* (2017). https://doi.org/10.1038/sdata.2017.117
20. S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, R.T. Shinohara, C. Berger, S.M. Ha, M. Rozycki, M. Prastawa et al., Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge. arXiv preprint arXiv:1811.02629 (2018)
21. A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in *Advances in Neural Information Processing Systems* (2012), pp. 1097–1105
22. Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition. Proc. IEEE **86**(11), 2278–2324 (1998)
23. J. Bernal, K. Kushibar, D.S. Asfaw, S. Valverde, A. Oliver, R. Martí, X. Lladó, Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review. Artif. Intell. Med. **95**, 64–81 (2019)
24. D. Ciresan, A. Giusti, L.M. Gambardella, J. Schmidhuber, Deep neural networks segment neuronal membranes in electron microscopy images, in *Advances in Neural Information Processing Systems* (2012), pp. 2843–2851
25. B.H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, L. Lanczi, The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans. Med. Imag. **34**(10), 1993–2024 (2014)

26. E. Shelhamer, J. Long, T. Darrell, Fully convolutional networks for semantic segmentation. IEEE Ann. Hist. Comput. **04**, 640–651 (2017)
27. M. Drozdzal, E. Vorontsov, G. Chartrand, S. Kadoury, C. Pal, The importance of skip connections in biomedical image segmentation, *Deep Learning and Data Labeling for Medical Applications* (Springer, Cham, 2016), pp. 179–187
28. F. Milletari, N. Navab, S.-A. Ahmadi, V-net: fully convolutional neural networks for volumetric medical image segmentation, in *2016 Fourth International Conference on 3D Vision (3DV)* (IEEE, 2016), pp. 565–571
29. D.P. Kingma, J. Ba, Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)

Chapter 15
Use of Deep Learning Approaches in Cancer Diagnosis

M. Hanefi Calp

Abstract Cancer is among serious health problems with an uncertain and complex structure that causes fatal results. Cancer is a disease that consists of uncontrolled proliferation of cells in different organs, whose clinical appearance, treatment and approach are different from each other and that should be controlled in the early stages. The cancer burden should be estimated in order to determine priorities for cancer control. In this context, there are many studies on diagnosis and treatment methods and a rapid development is observed in this regard. The aim is to increase the survival rate of people with cancer. In order to achieve this goal effectively, early and accurate diagnosis is especially important in the treatment of cancer, as it causes fatal results. It is known that cancer is very difficult to diagnose in the early stages and accurately with traditional diagnostic methods. At this point, the artificial intelligence, a new or current approach, comes to the agenda. Developments in this area offer very important opportunities in cancer diagnosis as in many areas. Therefore, in this study, deep learning approaches which are an artificial intelligence technique in the literature for the diagnosis of cancer disease are examined, and the applications in the literature on how these approaches are used are included. Since the subject of the study is up to date, it is considered that the study will be a guide for people or institutions working in this field.

Keywords Artificial intelligence · Deep learning · Cancer diagnosis · Treatment

15.1 Introduction

Cancer is a type of disease that threatens human life and has fatal consequences. This disease is expected to be among most serious diseases that cause deaths in the coming years. According to the data of the World Health Org. (WHO) [1], cancer corresponds to 13% of all deaths in the world in 2018. While the number of people

M. Hanefi Calp (✉)
Department of Management Information Systems, Karadeniz Technical University, Trabzon, Turkey
e-mail: mhcalp@ktu.edu.tr

U. Kose and J. Alzubi (eds.), *Deep Learning for Cancer Diagnosis*,
Studies in Computational Intelligence 908,
https://doi.org/10.1007/978-981-15-6321-8_15

dying from cancer will increase significantly in the future, a predicted 12 million people will die from cancer in 2030. Scientists, researchers and doctors are working on new techniques to effectively fight cancer, however, this struggle is known to be quite difficult [2–4].

Over the years, manual, automatic or controlled methods have been tried to detect cancer disease or cancerous region and many different solutions have been offered by applying these methods in a single or hybrid structure. However, these techniques are not very effective in detecting cancer if sufficient data could not be obtained. Therefore, in recent years, computer-aided diagnostic detection systems and similar techniques have attracted the attention of researchers, scientists and radiologists [2, 5]. At this point, various options for cancer treatment are still being sought and developed. In particular, artificial intelligence-based forecasting methods are among these techniques. Improvements in the predictive accuracy helps doctors plan patient treatments and eliminate physical and mental difficulties caused by the disease. Improvement in the percentage of diagnosis using the most advanced artificial intelligence technology is a support factor for clinical researchers. Technical and theoretical advances in computer software and statistics have allow of computer or software engineers and health scientists to achieve successful results in diagnosis using traditional logistic regression, multi-factor analysis and artificial intelligence-aided analysis. The accuracy of such estimates are significantly higher than the experimental estimates. With the application of artificial intelligence, scientists have recently begun to create new models to estimate and diagnose cancer. These models are very important in improving the accuracy of cancer sensibilty and survival estimates [4].

However, the realization of this diagnosis in the early stages of the disease is as important as the diagnosis and treatment of cancer. Early diagnosis of cancer is the top priority to save many people's lives. Normally, visual inspection and manual approaches are utilized for types of this cancer diagnosis. That explication (manual) of medical images requires a lot of time and is highly prone to errors [3]. Diagnosis of cancer in the early stages is a very difficult process due to uncertain symptoms, uncertainties in mammograms and screening, and there is a possibility that it will reappear after treatment. Therefore, better predictive models should be developed using multivariate data and high-resolution diagnostic/treatment tools in clinical cancer studies. When the studies on the diagnosis and treatment of cancer are examined in the literature, it is seen that the use of artificial intelligence technique is increasing. In addition, it has been revealed that traditional analysis solutions like the statistical and multivariate analysis are not as successful as artificial intelligence technique. Among the artificial intelligence techniques, especially the deep learning approach gives very effective results [6–9]. Deep learning is a kind of neural network with many hidden layers. In recent years, deep learning has been used in a wide variety of fields. In particular, it has provided high performance-success in application types like as speech recognition, and also image recognition and classification within high-tech devices such as autonomous vehicles and drones [10–12]. In addition, basic classifications like healthy and cancerous tissue diagnosis are performed and traditional machine learning solutions are employed in the developed models. However, artificial intelligence-based deep neural networks are a more effective method of generating

classification models with data matrices. These models are useful for diagnosing cancer, seeing and predicting its progress, and thus applying timely and optimal cancer treatment [13].

In this study, deep learning approaches used in cancer diagnosis and treatment are examined. The aim of the study is to show with the support of the literature the extent to which effective results are obtained with a deep learning approach, which is one of the artificial intelligence techniques of a disease such as cancer, which methods and techniques are used, and how these methods are used. In the second part of the study, artificial intelligence and deep learning technique are explained; in the third part, deep learning approaches in cancer diagnosis and treatment processes and examples in the literature; In the fourth part, the results obtained from the study and the suggestions about the study are included.

15.2 Artificial Intelligence and Deep Learning

15.2.1 Artificial Intelligence

From the past to the present day, human intelligence has been explored through science or various ways, and it has been tried to be created by imitating or modeling. The studies in this field have come to an advanced level every day. Intelligence can be improved by working, teaching, training, and acquired knowledge and experience, and skills based on experience. The ability to adapt, understand, learn and analyze an event encountered for the first time can be realized with intelligence. Intelligence can be imitated by software or integrated chips, and this is called artificial intelligence. Since human intelligence is accepted as the most complex structure in the world, the concept of artificial intelligence did not go beyond representing the effort to imitate the human brain. While human intelligence can perform a simple numerical operation in a few minutes, a very short time is required for a perception related process. It is insufficient on issues the application of the knowledge gained through understanding and applying, while the computer analyzes very complex numerical operations instantly [14, 15].

Artificial Intelligence is a branch of science that has the abilities of machines to such as reasoning, use of from past knowledge, planning, learning, communication and perception. Systems ensuring certain human behaviors with artificial intelligence and that run human thinking process related to certain specialties such as computation and medical diagnosis can be established [16]. Artificial intelligence is a technique that provides solutions to real world problems such as classification, optimization, prediction, prediction, pattern and image recognition, control, diagnosis and interpretation. Therefore, it is used by all disciplines in almost every field [17–21]. Artificial intelligence is based on imitating people's ways of thinking and behavior, however, in time, any kind of dynamism, which can be put into mathematical, and also logical patterns in the sense of the frame of intelligent functions within

nature can be expressed. Briefly, it is a multidisciplinary field in which it is aimed to automate the activities that take place with human intelligence [22].

According to Stuart J. Russell and Peter Norvig, Artificial Intelligence can be examined in the context of four different categories: systems that think like people, systems that act like people, systems that think rationally and systems that act rationally. They also stated that artificial intelligence evaluates human decisions, that it have a human characteristic, such as thinking, and that it has logical decision-making capabilities [23]. Essential objective of the artificial intelligence is to create machines smarter, understand about what intelligence is, and also ensure machines more useful. In addition, imitating human intelligence by means of a computer means in this sense to give a certain degree of learning ability to computers. Thus, artificial intelligence often corresponds to methods, which try to model human thinking, the working model of the brain, or the evolution (biological) of the nature [24]. In summary, artificial intelligence is the equipping of machines by modeling with human intelligence, physiological and neurological structure. It is computer system that think and act like a human [25, 26].

15.2.2 Deep Learning

In recent years, deep learning, which is a very popular subject and has a high success in almost every field, is an effective technique used in the perception and understanding of the world of machines. Deep learning methods are used in areas such as image classification, speech recognition, video analysis, and the natural language learning. Analysis can be made without the need for an attribute extractor based on a mathematical model previously created with deep learning. One of the most important advantages of deep learning methods is that it can be generalized. A learned neural network approach can be used for other applications and data structures. If the data set is insufficient, the performance of deep learning will also be low [27, 28].

Deep learning is known as a machine learning sub-field, which follows advantages of nonlinear processing unit layers. The next layer receives the output of the previous layer as the input. In the deep learning technique, the data are based on learning from the representation of the data by learning more than one feature level. Top-level features are derived from low-level features, thereby creating a hierarchical representation. Deep learning methods have been developed as oriented on artificial neural networks generally, but have more hidden neurons and layers. Deep learning methods have yielded very successful results in the processing of many types of data such as video, audio, text [29–33].

Deep learning methods are used in the application types like [30, 34]:

1. Natural language processing/language modeling
2. Speech and audio processing
3. Information retrieval

4. Computer vision/object recognition
5. Multi-modal and multi-task learning.

15.3 Deep Learning Within Cancer Diagnosis

Doctors often rely on their personal knowledge, skills and experience when diagnosing the disease. However, no matter how skilled he is, a doctor cannot claim that the diagnosis of the disease is absolutely correct, and he definitely makes misdiagnoses. Therefore, at this point, artificial intelligence technologies come to the agenda. Because artificial intelligence has the ability to analyze large amounts of data, solve complex problems and predict with high accuracy. Deep neural network, one of the most up-to-date artificial intelligence techniques of today, expresses a series of computer models that are used effectively in obtaining information from images. Deep learning algorithms have been applied for activities in many medical specialties (most commonly, radiology and pathology). In addition, high performance has been get in the sense of running deep learning solutions in the areas of cancer biology, as like clinical imaging of several species [4, 35]. In Fig. 15.1, artificial intelligence techniques and their relationship are given in estimating cancer diagnosis.

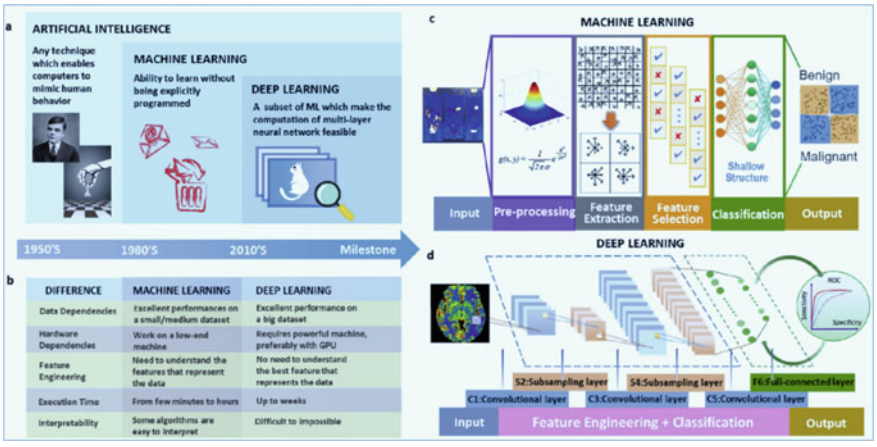


Fig. 15.1 The techniques of AI, ML, and DL, and their relationships [4]. **a**: Use of AI technology such as ML and DL as a decision support tool, **b**: The differences exist between ML and DL, **c**: The example for ML, **d**: Involves more operations than DL

15.3.1 Steps Regarding Cancer Diagnosis

15.3.1.1 Pre-process Step

Since the raw images contain noise, the first step in the detection process is pretreatment. Preprocessing is to improve the quality of the considered image that will be used more by removing image information, which is unwanted and called image sounds. If this problem is not corrected, incorrect classification can be made. The right combination of preprocessing tasks is required to increase the accuracy rate [3].

15.3.1.2 Image Segmentation

Segmentation is essentially the partition or division of a zone. It is divided into four part: threshold-based, pixel-based, region-based, and also model-based segmentation. There are also other techniques such as histogram threshold, adaptive threshold, distributed and localized region definition, gradient flow vector, clustering and statistical region growth, active contours, boot learning, supervised learning, edge detection. These methods are also used in hybrid forms [3, 36, 37].

15.3.1.3 Post-processing Step

Next to the pre-process and the image segmentation steps, the post-processing process begins. The basic steps of the three titles described so far are summarized in Table 15.1 [3].

Table 15.1 Steps regarding cancer diagnosis

Pre-processing step	Image segmentation	Post-processing step
Adjustment of the contrast	Histogram thresholding	Opening and closing operations
Remove of the vignetting effect	Distributed and localized	Island removal
	Region identification	
Correction of the color	Clustering and active contours	Region merging
Image smoothing	Supervised learning	Border expansion
Hair removal	Edge detection and Fuzzy logic	Smoothing
Normalization and the localization	Probabilistic modeling and graph theory	

15.3.1.4 The Rule of ABCD

The ABCD characters in this rule corresponds to A-asymmetry, B-border, C-color, and the D-diameter of the image of lesion, respectively and each is defined as follows [38]:

(A) for the Asymmetry: The related input image is divided into a vertical axis, giving the lowest possible value of the asymmetry score.
(B) for the Border: The related image is divided into eight and checks for sharp and sudden changes.
(C) for the Color: Shades of color are used for cancer detection.
(D) for the Diameter: A ceraful check for the diameter of the lesion is done.

15.3.1.5 7-Point Checklist Method

It has two types of criteria, large and small, and the main criteria have three scores, and each of the small criteria has 4 points with a value of 1 point [39].

15.3.1.6 Menzies Method

There are several positive and negative features. The asset of negative features means that melanoma is malignant. That will be benign if one or more positives are true and both negatives are absent [39].

15.3.1.7 Pattern Analysis

A method oriented on determining patterns that are global or local can be run. Main objective of that method is to perform a qualitative evaluation of dermoscopic criteria separately [3].

15.3.2 Deep Learning Methods and Use in Cancer Treatment

Deep learning methods are widely used because they have multi-layered neural networks and perform successfully especially in areas such as computer vision and pattern recognition. The most important feature of this method is that the feature extraction process is accepted as a part of the learning process [40, 41].

The deep learning methods in the literature are listed below and each of them is briefly explained and several examples in practice are included.

15.3.2.1 Deep Neural Networks (DNN)

A simple neural network has an input layer and is directly linked to the output. DNNs that are effective in solving complex problems have multiple hidden layers and the weight of each layer has been adjusted with the delta learning method. By adding more hidden layers to deep neural networks, complex nonlinear relationships can be detected. DNNs are used in supervised and unsupervised learning problems, but learning processes are very slow. However, it is generally used for classification and regression purposes, and high performance results can be obtained [32, 34].

In Fig. 15.2, an example DNN model is given.

Ohmori et al. detected endoscopic detection and differentiation of esophageal lesions using deep neural network. It was revealed that there was no significant difference between the artificial intelligence system and experienced endoscopists in terms of diagnostic performance. The neural network system has high sensitive and they developed has shown high accuracy in differentiating non-cancerous lesions [43].

Schwyzer et al. investigated whether deep neural networks can be used to detect lung cancer in the environment of low positron emission tomography/computed tomography scans. They found that the deep neural network algorithm achieved high performance for lung cancer detection. They showed that the further development of this technology increased the success of lung cancer screening efforts [44].

Tan and colleagues designed a model on adaptive melanoma diagnosis, by employing clustering, ensemble and DNNs. They used the hybrid Particle Swarm Optimization algorithm for skin lesion segmentation and also classification. They used the Firefly algorithm to diversify and so that improve accordingly the algorithm. It has been used in conjunction with the clustering algorithm of K-Means, in order to increase segmentation of the lesion. In addition to these techniques, the lesion classification was performed with deep neural networks using the skin lesion data set, and the results showed higher performance than other techniques [45].

Skourt et al. conducted a research on lung computed tomography image segmentation using the deep neural network approach. At this point, U-net architecture, one

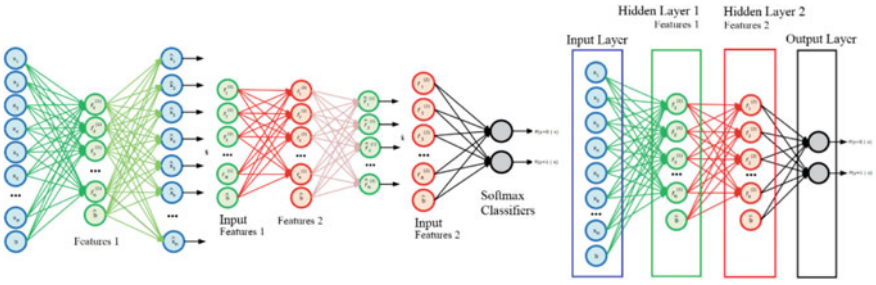


Fig. 15.2 An example for DNN model [42]

of the architectures commonly used in deep learning technique for image segmentation, was used. Because the deep neural network can be trained from very few images and end to end, they have achieved better performance than other methods [46].

Pomponiu and colleagues have proposed a model that classifies a skin mole lesion with a pre-trained deep neural network. They found that experimental tests on a clinical data set performed better than other state-of-the-art classification performance using deep neural network-based features. That is, experimental results have shown that the transferred properties are a suitable solution for skin mole characterization and achieve better classification accuracy [47].

Jiang et al. have introduced a model for recognition or segmentation of basal cell carcinoma in digital histopathology images captured by smartphones with a deep neural network. With the model they developed, they were able to divide all the images they obtained with high accuracy. As a result, they were able to recognize basal cell carcinoma pathology with high sensitivity and accuracy by using deep learning method via smartphone [48].

15.3.2.2 Convolutional Neural Networks (CNN)

CNN is a type of multilayer sensors. CNN algorithm, a forward neural network, has been developed by inspiring the visual characteristics of animals. The mathematical convolution process here can be considered as a neuron's response to stimuli from its impulse domain. CNN consists of one or more fully bound layers [29, 33].

In the convolution process, the symmetry of the filter, which will be applied to two dimensional data, is taken according to the x and y axis. As the filter is moved on the image depending on the step length, the overlapping values are multiplied by element on each step and the sum of all values is recorded as the relevant element of the output matrix. The process performed when the filter is not symmetrical is called cross-correlation. While the input data is single channel, this operation can be done simply. The number of channels of the exit sign is also equal to the number of filters applied [28, 49, 50] (Fig. 15.3).

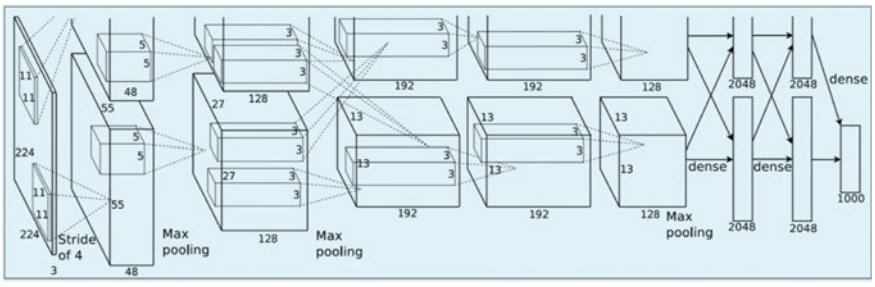


Fig. 15.3 An example for CNN model [51]

Paul et al. used Deep Neural Network and Traditional Image Features approaches to increase survival estimation accuracy in lung cancer patients. They reported that deep feature extraction via (pre-trained) CNN is successful for some image areas. The researchers applied a pre-trained evolutionary neural network method for obtaining deep features from computed tomography images (lung cancer), and then classified them to predict long-term and also short-term survivors. Using the technique in question, they achieved high accuracy and performance results [52].

Ghoneim et al. have developed a system for cervical cancer classification using CNNs and extreme learning machines. First of all, cell images were transmitted to convolutional neural networks and input images were classified with extreme learning machines. The test results were carried out using the created database. Researchers have revealed that the proposed system can be used for high accuracy for detection and classification problems [53].

Rasti et al. have studied breast cancer diagnosis to distinguish benign and malignant breasts using CNNs. In the proposed system, the masses in the chest are classified using the data. Experimental results revealed that the proposed model can be used as an effective tool for radiologists to analyze breast images [54].

Arevalo et al. performed a representative learning for mammography mass lesion classification using CNNs. In their work, they proposed an innovative model that automatically learned discriminatory features by avoiding the design of custom handmade image-based feature detectors for breast cancer diagnosis in mammography. To improve image details, they implemented two phases: preprocessing and supervised training to learn the classifier of breast imaging lesions. Compared to the latest technology image identifiers of the method they developed, they achieved successful results in terms of performance. As a result, they suggested that the deep learning based convolutional neural networks approach is an effective method to automatically handle the classification of breast mass lesions in mammography [55].

Zhang et al. have developed a new optimized image processing-based model using convolutional neural networks for early diagnosis of skin cancer. They used the whale optimization algorithm to optimize the CNNs part of the model. They compared with different methods in the different data clusters to evaluate the proposed method. The experimental results showed that the proposed method is superior to other methods compared [56].

15.3.2.3 Recurrent Neural Networks (RNN)

RNN is a type of artificial neural network where connections between units ensure a loop, which is directed. Unlike feed forward neural networks, RNNs use sequential information to process their input memory random strings of inputs. The reason why RNN architecture is called duplicate is that it performs the same task for each element of an array based on previous outputs. RNN is often preferred because of its serial processing feature in the field of natural language processing. It produces successful results in sequential flow data such as speech, text, music, and video [29, 49, 57] (Fig. 15.4).

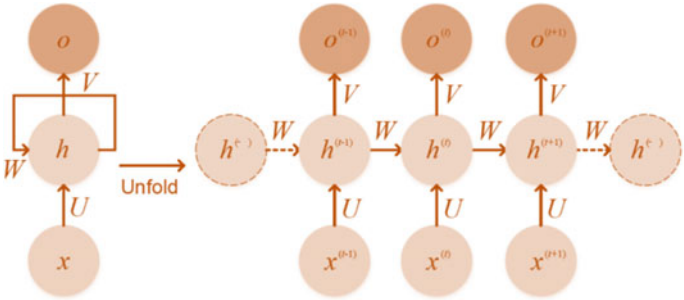


Fig. 15.4 An example for RNN model [58]

Azizi et al. conducted a study on deep recurrent neural networks for the detection of prostate cancer. In the study, they used deep repetitive neural networks to model temporal information. They achieved a high sensitivity and accuracy rate that distinguishes cancer in the prostate from benign tissue. They have proven to significantly improve cancer detection accuracy, based on previously presented studies using RNN [59].

Zanjani et al. conducted a study aimed at detecting the cancerous region on mass spectrometry imaging data using repetitive neural networks. In the proposed RNN model, lung and bladder cancer datasets, mass spectra have a fast training time and have been able to classify with high accuracy and better performance [60].

Chen et al. proposed a model using the incremental combination of repetitive neural networks for estimation of malignancy in breast cancer, which is very common in women. The researchers stated that the early detection of malignancy of breast cancer and a treatment suitable for diagnosis greatly increased the survival rate of patients. As a result, they have achieved highly effective and successful results using a repetitive neural network layer to classify the benign and malignant features of breast cancer [61].

Amirkhan et al. Employed a recurrent neural network model for predicting colorectal cancer among patients. In the study, they applied repetitive neural networks to see whether these networks can learn and produce accurate prediction models. The experimental results showed that equal performance is achieved with the latest technology algorithms. They also reported that such models could be a source of support for issues such as cancer detection and could be used as a resource for future research [62].

15.3.2.4 Deep Boltzmann Machine (DBM)

DBM is a productive random artificial neural network model, which has the ability of learning probability distribution over the input set. They consist of two-part graphs

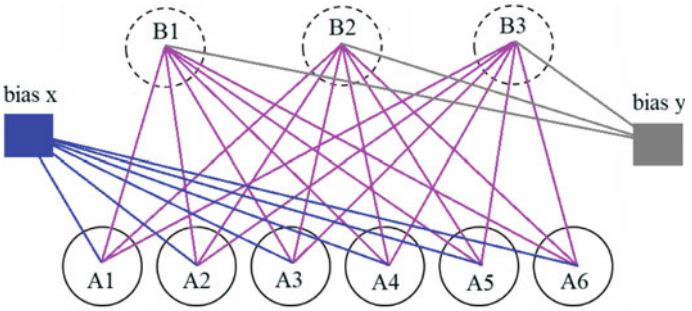


Fig. 15.5 An example for DBM model [64]

with symmetrical connections between them, that are visible and hidden [29, 63] (Fig. 15.5).

Nivaashini and Soundariya have proposed a model for deep boltzmann machine-related breast cancer risk assessment for health systems. In the proposed system, it is aimed to find a particularly effective feature set and to divide the cancer status into two groups as good and malignant. Outputs from the study were evaluated by using specificity, sensitivity and accuracy with the classifiers like support vector machine (SVM), unified neural network (CNN), probabilistic neural network (PNN), multilayer perceptual neural network (MLPNN), repetitive neural network (RNN). Naïve Bayes (NB), SMO and C4.5 techniques were used to classify breast tumors. A high detection rate was obtained in the estimation and classification of breast tumors using the proposed medical diagnostic decision-making system; that is, the authors have shown very good results and performance [65].

Jeyaraj and Nadar have provided the classification of the cancerous region in order to provide a better health system by using the DBM algorithm. At this point, computer-aided diagnostic performance is enhanced by confirming the region in the hyperspectral image with the pre- and post-cancer region classification. A three-layer unsupervised network with back propagation structure was used for implementation. Image fragments were collected from the presented dataset. It has been revealed that the proposed model produces solutions with high precision and accuracy, and according to the classification, accuracy and success rates achieved, DBM better classifies complex images compared to the traditional convolution network [66].

15.3.2.5 Deep Belief Networks (DBN)

DBN's are defined as a stack of DBMs. Each DBM layer is linked with both previous and next layers. However, the nodes of any layer do not have horizontal communication with each other. It has the ability to classify as the last layer or cluster for unsupervised learning. DBN architectures are implemented in image recognition and production [29, 67] (Fig. 15.6).

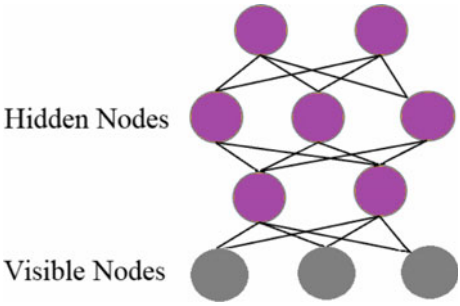


Fig. 15.6 An example for DBN model [68]

Zhao et al. designed a new deep learning and clustering model called DBN and Fuzzy C-Means (FCM) to cluster lung cancer patients. In the proposed model, the images after preprocessing are first coded into multiple layers of hidden variables to achieve hierarchical properties and feature distribution and to create high-level representations. Experimental results reveal that it performs better than unsupervised classification methods and also show that the proposed model enables practical implementations in lung cancer pathogenesis studies and is a useful guide for cancer treatment [69].

Shahweli realized an exemplary study the use of deep learning technique in predicting diseases, classifying and analyzing the causes of diseases such as cancer. In the study, he used the DBNs method to predict Predisposition to Lung Cancer in TP53 Gene. The data set is trained with the back propagation neural network; The training is divided into three parts to represent the set of validation and testing. The experimental results revealed that the neural network model can highly predict lung cancer [70].

Renukadevi and Karunakaran have proposed a method to optimize deep belief network parameters using the grasshopper algorithm for liver disease classification. Initially, image quality is enhanced by preprocessing techniques, and then features such as texture, color, and shape are obtained. In the study, DBN parameters were optimized to recognize liver disease. The proposed method has yielded successful results with a high accuracy, precision and precision rates compared to existing techniques in the simulation method [71].

15.3.2.6 Deep Autoencoders (DA)

DAs is a machine learning method with input and output layers having the same properties, consisting of clustering multiple auto-encoders on top of each other to represent multi-dimensional data. The purpose of this method is to automatically extract the attribute set or reduce the number of data sizes. That is, to capture the more accurate ratio with less attributes. In this method, unsupervised learning method is used; that is, there is no need for a marked data set for learning. However, the method

requires a pre-learning phase to find suitable weights. DAs are frequently used in feature extraction and data compression processes [32, 34, 67, 72, 73].

Mallick et al. have proposed a technique for cancer detection that enables image classification using a deep wavelet auto-encoder based DNN. In the study, they blended the image decoupling feature of the wavelet transform together with the main feature reduction feature of the automatic encoder. A brain image was taken as a data set and the data was classified with the proposed image classifier. The performance criterion of the model has been compared with other existing classifiers such as automatic encoder-DNN or DNN, and it has been stated that the proposed method outstrips existing methods [75].

Kucharski et al. proposed a semi-controlled segmentation approach, thanks to convective autoencoders so that it was possible to solve the problem of segmentation tasks with few ground-reality images. They used deep network architecture to differentiate benign and malignant skin lesions of cancerous areas in histopathological images of skin samples. They have developed a deep learning tool with computer vision support that can effectively ensure segmentation of slots based on the autocoder with two learning steps. With their experimental results, they confirmed the effectiveness-success of the proposed approach and the high similarity coefficient, sensitivity and discrimination rate [76] (Fig. 15.7).

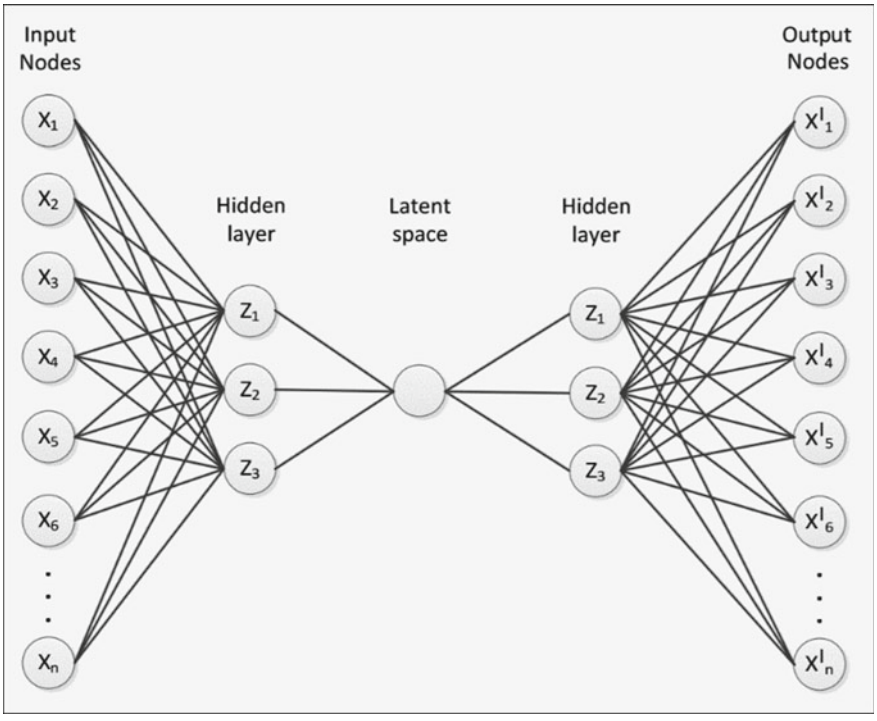


Fig. 15.7 An example for DA model [74]

Pyman et al. investigated the microRNA regulation of cancer with a context-sensitive deep cancer class. At this point, they proposed a new class of miRNA-based deep cancer that includes genomic and hierarchical tissue detailing, which can accurately predict the presence of cancer in a wide range of human tissues. They have created a deep learning architecture by leveraging stacked auto-codecs and a multi-layered sensor. Four-fold cross-validation is used to evaluate the performance of this model. As a result, DA networks have been found to be a powerful tool for modeling complex miRNA-phenotype relationships in cancer and increase classification accuracy [77].

15.4 Conclusion and Suggestions

In this study, deep learning techniques used for cancer diagnosis are examined. It has been observed that these techniques primarily lead to a method change in cancer diagnosis, are an auxiliary tool for doctors, increase efficiency, standardize quality processes and achieve highly effective results with successful predictions. To date, deep learning technology applications have been tested and substantially validated in subjects such as reproducibility and generalizability in cancer treatment. Because, in order to reveal the effect or contribution of this technique, versatile feasibility studies have been conducted and applied in different clinical settings, thus increasing the efficiency levels of doctors. So much so that, by using these techniques, cancer diagnosis was made in the early stages and contributions such as survival of many patients or prolonging their life were provided. In particular, in the field of medicine, it can be said that the technological change based on deep learning technology has started to improve the regional and national health sector. For this reason, it is beneficial to reach more theoretical knowledge by applying deep learning technology in cancer diagnosis and medicine in general. In addition, deep learning techniques should be supported by hybrid approaches, using different methods such as statistical methods to achieve higher levels of performance.

Although these techniques or methods have serious contributions to cancer diagnosis, these developments are still at an early stage. More effective and advanced deep learning approaches will be used in the near future and important steps will be taken in the diagnosis and treatment of a vital disease such as cancer disease. As a result, at least until today, deep learning technology has been used in the field of national and international medicine and it has been found that it provides very successful improvements in a vital issue such as cancer treatment. At this point, another important issue is to ensure that deep learning techniques used in the diagnosis and treatment of cancer are more involved in theoretical and academic studies and that this issue is developed.

References

1. WHO Cancer Fact Sheets (2020), [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cancer
2. J. Tang, R.M. Rangayyan, J. Xu, I. El Naqa, Y. Yang, Computer-aided detection and diagnosis of breast cancer with mammography: recent advances. IEEE Trans. Inf Technol. Biomed. **13**(2), 236–251 (2009)
3. K. Munir, H. Elahi, A. Ayub, F. Frezza, A. Rizzi, Cancer diagnosis using deep learning: a bibliographic review. Cancers **11**(9), 1235 (2019)
4. S. Huang, J. Yang, S. Fong, Q. Zhao, Artificial intelligence in cancer diagnosis and prognosis: opportunities and challenges. Cancer Lett. **28**(471), 61–71 (2020)
5. R. Fakoor, F. Ladhak, A. Nazi, M. Huber, Using deep learning to enhance cancer diagnosis and classification, in *Proceedings of the International Conference on Machine Learning*, vol. 28 (ACM, New York, USA, 2013)
6. Z. Obermeyer, E.J. Emanuel, Predicting the future-big data, machine learning, and clinical medicine. The New England journal of medicine **375**(13), 1216 (2016)
7. A. Allahyar, J. Ubels, J. De Ridder, A data-driven interactome of synergistic genes improves network-based cancer outcome prediction. PLoS Comput. Biol. **15**(2), e1006657 (2019)
8. M.J. Mitchell, R.K. Jain, R. Langer, Engineering and physical sciences in oncology: challenges and opportunities. Nat. Rev. Cancer **17**(11), 659 (2017)
9. A. Hosny, C. Parmar, J. Quackenbush, L.H. Schwartz, H.J. Aerts, Artificial intelligence in radiology. Nat. Rev. Cancer **18**(8), 500–510 (2018)
10. J. Jeong, Deep learning for cancer screening in medical imaging. Hanyang Med. Rev. **37**(2), 71–76 (2017)
11. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 770–778)
12. A. Graves, A.R. Mohamed, G. Hinton, Speech recognition with deep recurrent neural networks, in *2013 IEEE international conference on acoustics, speech and signal processing* (IEEE, 2013), pp. 6645–6649
13. A. Park, S. Nam, Deep learning for stage prediction in neuroblastoma using gene expression data. Genomics Inf. **17**(3), e30 (2019)
14. İ.Y. Kazu, O. Özdemir, Öğrencilerin bireysel özelliklerinin yapay zeka ile belirlenmesi (Bulanık mantık örneği). Akademik Bilişim (2009), pp. 11–13
15. Ç. Elmas, Bulanık Mantık Denetleyiciler-Kuram. *Uygulama, Sinirsel Bulanık Mantık, Seçkin Yayıncılık: Ankara* (2003)
16. V. Nabiyev, *Yapay Zeka, Birinci Baskı* (Seçkin Yayıncılık Sanayi ve Ticaret, AŞ, Ankara, 2003)
17. A.I. Strong Applications of artificial intelligence & associated technologies. Science, 5(6) (2016)
18. S. Makridakis, The forthcoming artificial intelligence (AI) revolution: its impact on society and firms. Futures **90**, 46–60 (2017)
19. M.H. Calp, The role of artificial intelligence within the scope of digital transformation in enterprises, in *Advanced MIS and Digital Transformation for Increased Creativity and Innovation in Business* (IGI Global, 2020), pp. 122–146
20. I.M. Cockburn, R. Henderson, S. Stern, The impact of artificial intelligence on innovation (No. w24449) (National Bureau of Economic Research, 2018). https://doi.org/10.3386/w24449
21. H.M. Calp, Medical diagnosis with a novel SVM-CoDOA based hybrid approach. BRAIN Broad Res. Artif. Intell. Neurosci. **9**(4), 6–16 (2018)
22. C. Williams, A brief introduction to artificial intelligence, in *OCEANS'83, Proceedings* (IEEE, 1983), pp. 94–99
23. S. Russell, P. Norvig, A. Intelligence, Knowledge and reasoning: a modern approach, in *Artificial Intelligence* (Prentice-Hall, Egnlewood Cliffs, 1995), p. 25, 27
24. M. Tektaş, A. Akbaş, V. Topuz, Yapay zeka tekniklerinin trafik kontrolünde kullanilmasi üzerine bir inceleme (2002)

25. M.E. Balaban, E. Kartal, *Veri Madenciliği ve Makine Öğrenmesi Temel Algoritmaları ve R Dili ile Uygulamaları* (Çağlayan Kitabevi, İstanbul, 2015)
26. M. Atalay, E. Çelik, Büyük Veri Analizinde Yapay Zekâ Ve Makine Öğrenmesi Uygulamalari-Artificial Intelligence and Machine Learning Applications in Big Data Analysis. Mehmet Akif Ersoy Üniversitesi Sosyal Bilimler Enstitüsü Dergisi **9**(22), 155–172 (2017)
27. Ş. Karahan, Y.S. Akgül, Eye detection by using deep learning, in *2016 24th Signal Processing and Communication Application Conference (SIU)* (IEEE, 2016), pp. 2145–2148
28. İ. Özkan, E. Ülker, Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri. Gaziosmanpaşa Bilimsel Araştırma Dergisi **6**(3), 85–104 (2017)
29. A. Şeker, B. Diri, H.H. Balık, Derin öğrenme yöntemleri ve uygulamaları hakkında bir inceleme. Gazi Mühendislik Bilimleri Dergisi (GMBD) **3**(3), 47–64 (2017)
30. L. Deng, D. Yu, Deep learning: methods and applications. Found. Trends® Sign. Process. **7**(3–4), 197–387 (2014)
31. Y. Bengio, Learning deep architectures for AI. Found.® Trends Mach. Learn. **2**(1), 1–127 (2009)
32. D. Ravì, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, G.Z. Yang, Deep learning for health informatics. IEEE J. Biomed. Health Inf. **21**(1), 4–21 (2016)
33. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. Nature **521**(7553), 436–444 (2015)
34. D. Küçük, N. Arıcı, Doğal Dil İşlemede Derin Öğrenme Uygulamaları Üzerine Bir Literatür Çalışması. Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi **2**(2), 76–86 (2018)
35. N. Sompairac, P.V. Nazarov, U. Czerwinska, L. Cantini, A. Biton, A. Molkenov, U. Kairov, Independent component analysis for unraveling the complexity of cancer omics datasets. Int. J. Mol. Sci. **20**(18), 4414 (2019)
36. M. Emre Celebi, H.A. Kingravi, H. Iyatomi, Y. Alp Aslandogan, W.V. Stoecker, R.H. Moss, J.M. Malters, J.M. Grichnik, A.A. Marghoob, H.S. Rabinovitz, S.W. Menzies, Border detection in dermoscopy images using statistical region merging. Skin Res. Technol. **14**(3), 347–353 (2008)
37. N. Tong, H. Lu, X. Ruan, M.H. Yang, Salient object detection via bootstrap learning, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2015), pp. 1884–1892
38. M. Garbaj, A.S. Deshpande, Detection and analysis of skin cancer in skin lesions by using segmentation. IJARCCE (2015)
39. R.H. Johr, Dermoscopy: alternative melanocytic algorithms—the ABCD rule of dermatoscopy, menzies scoring method, and 7-point checklist. Clin. Dermatol. **20**(3), 240–247 (2002)
40. A.A Cruz-Roa, J.E.A. Ovalle, A. Madabhushi, F.A.G. Osorio, A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection, in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer, Berlin, Heidelberg, 2013), pp. 403–410
41. Y. Bengio, A. Courville, P. Vincent, Representation learning: a review and new perspectives. IEEE Trans. Pattern Anal. Mach. Intell. **35**(8), 1798–1828 (2013)
42. A. Baştürk, M.E. Yüksel, A. Çalışkan, H. Badem, Deep neural network classifier for hand movement prediction, in *2017 25th Signal Processing and Communications Applications Conference (SIU)* (IEEE, 2017), pp. 1–
43. M. Ohmori, R. Ishihara, K. Aoyama, K. Nakagawa, H. Iwagami, N. Matsuura, S. Shichijo, K. Yamamoto, K. Nagaike, M. Nakahara, T. Inoue, Endoscopic detection and differentiation of esophageal lesions using a deep neural network. Gastrointest. Endosc. **91**(2), 301–309 (2020)
44. M. Schwyzer, D.A. Ferraro, U.J. Muehlematter, A. Curioni-Fontecedro, M.W. Huellner, G.K. von Schulthess, P.A. Kaufmann, I.A. Burger, M. Messerli, Automated detection of lung cancer at ultralow dose PET/CT by deep neural networks–initial results. Lung Cancer **1**(126), 170–173 (2018)
45. T.Y. Tan, L. Zhang, C.P. Lim, Adaptive melanoma diagnosis using evolving clustering, ensemble and deep neural networks. Knowl.-Based Syst. **187**, 104807 (2020)
46. B.A. Skourt, A. El Hassani, A. Majda, Lung CT image segmentation using deep neural networks. Procedia Comput. Sci. **127**, 109–113 (2018)

47. V. Pomponiu, H. Nejati, N.M. Cheung, Deepmole: deep neural networks for skin mole lesion classification, in *2016 IEEE International Conference on Image Processing (ICIP)* (IEEE, 2016), pp. 2623–2627
48. Y.Q. Jiang, J.H. Xiong, H.Y. Li, X.H. Yang, W.T. Yu, M. Gao, X. Zhao, Y.P. Ma, W. Zhang, Y.F. Guan, H. Gu, Recognizing basal cell carcinoma on smartphone-captured digital histopathology images with a deep neural network. Br. J. Dermatol. **182**(3), 754–762 (2020)
49. M.A. Kizrak, B. Bolat, Derin öğrenme ile kalabalık analizi üzerine detaylı bir araştırma. Bilişim Teknolojileri Dergisi **11**(3), 263–286 (2018)
50. F. Doğan, İ. Türkoğlu, Derin öğrenme algoritmalarının yaprak sınıflandırma başarımlarının karşılaştırılması. Sakarya University Journal of Computer and Information Sciences **1**(1), 10–21 (2018)
51. A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in *Advances in Neural Information Processing Systems* (2012), pp. 1097–1105
52. R. Paul, S.H. Hawkins, L.O. Hall, D.B. Goldgof, R.J. Gillies, Combining deep neural network and traditional image features to improve survival prediction accuracy for lung cancer patients from diagnostic CT, in *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (IEEE, 2016), pp. 002570–002575
53. A. Ghoneim, G. Muhammad, M.S. Hossain, Cervical cancer classification using convolutional neural networks and extreme learning machines. Future Gener. Comput. Syst. **102**, 643–649 (2020)
54. R. Rasti, M. Teshnehlab, S.L. Phung, Breast cancer diagnosis in DCE-MRI using mixture ensemble of convolutional neural networks. Pattern Recogn. **72**, 381–390 (2017)
55. J. Arevalo, F.A. González, R. Ramos-Pollán, J.L. Oliveira, M.A.G. Lopez, Representation learning for mammography mass lesion classification with convolutional neural networks. Comput. Methods Programs Biomed. **127**, 248–257 (2016)
56. N. Zhang, Y.X. Cai, Y.Y. Wang, Y.T. Tian, X.L. Wang, B. Badami, Skin cancer diagnosis based on optimized convolutional neural network. Artif. Intell. Med. **102**, 101756 (2020)
57. T. Mikolov, M. Karafiát, L. Burget, J. Černocký, S. Khudanpur, Recurrent neural network based language model, in *Eleventh Annual Conference of the International Speech Communication Association* (2010)
58. W. Feng, N. Guan, Y. Li, X. Zhang, Z. Luo, Audio visual speech recognition with multimodal recurrent neural networks, in *2017 International Joint Conference on Neural Networks (IJCNN)* (IEEE, 2017), pp. 681–688
59. S. Azizi, S. Bayat, P. Yan, A. Tahmasebi, J.T. Kwak, S. Xu, B. Turkbey, P. Choyke, P. Pinto, B. Wood, P. Mousavi, Deep recurrent neural networks for prostate cancer detection: analysis of temporal enhanced ultrasound. IEEE Trans. Med. Imag. **37**(12), 2695–2703 (2018)
60. F.G. Zanjani, A. Panteli, S. Zinger, F. van der Sommen, T. Tan, B. Balluff, D.R. Vos, S.R. llis, R.M. Heeren, M. Lucas, H.A. Marquering, Cancer detection in mass spectrometry imaging data by recurrent neural networks, in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)* (IEEE, 2019), pp. 674–678
61. D. Chen, G. Qian, C. Shi, Q. Pan, Breast cancer malignancy prediction using incremental combination of multiple recurrent neural networks, in *International Conference on Neural Information Processing* (Springer, Cham, 2017), pp. 43–52
62. R. Amirkhan, M. Hoogendoorn, M.E. Numans, L. Moons, Using recurrent neural networks to predict colorectal cancer among patients, in *2017 IEEE Symposium Series on Computational Intelligence (SSCI)* (IEEE, 2017), pp. 1–8
63. R. Salakhutdinov, G. Hinton, Deep boltzmann machines, in *Artificial Intelligence and Statistics* (2009), pp. 448–455
64. A. Oppermann, Deep learning meets physics: restricted Boltzmann machines part I. Towar. Data Sci., (2018). https://towardsdatascience.com/deep-learning-meets-physics-restricted-boltzmann-machines-part-i-6df5c4918c15
65. M. Nivaashini, R.S. Soundariya, Deep boltzmann machine based breast cancer risk detection for healthcare systems. Int. J. Pure Appl. Math **119**, 581–590 (2018)

66. P.R. Jeyaraj, E.R.S. Nadar, Deep Boltzmann machine algorithm for accurate medical image analysis for classification of cancerous region. Cogn. Comput. Syst. **1**(3), 85–90 (2019)
67. G.E. Hinton, R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks. Science **313**(5786), 504–507 (2006)
68. Manish Nayak, Feature Detection Using Deep Belief Networks (DBN), https://mc.ai/feature-detection-using-deep-belief-networksdbn/
69. Z. Zhao, J. Zhao, K. Song, A. Hussain, Q. Du, Y. Dong, J. Liu, X. Yang, Joint DBN and Fuzzy C-Means unsupervised deep clustering for lung cancer patient stratification. Engineering Applications of Artificial Intelligence. **1**(91), 103571 (2020)
70. Z.N. Shahweli, Deep belief network for predicting the predisposition to lung cancer in TP53 gene. Iraqi J. Sci. **61**(1), 171–177 (2020)
71. T. Renukadevi, S. Karunakaran, Optimizing deep belief network parameters using grasshopper algorithm for liver disease classification. Int. J. Imag. Syst. Technol. **30**(1), 168–184 (2020)
72. M. Uçar, E. Uçar, Derin Otomatik Kodlayıcı Tabanlı Özellik Çıkarımı İle Android Kötücül Yazılım Uygulamalarının Tespiti. Yönetim Bilişim Sistemleri Dergisi **5**(2), 21–28 (2019)
73. P. Baldi, Autoencoders, unsupervised learning, and deep architectures, in *Proceedings of ICML Workshop on Unsupervised and Transfer Learning* (2012), pp. 37–49
74. T. Amarbayasgalan, B. Jargalsaikhan, K.H. Ryu, Unsupervised novelty detection using deep autoencoders with density based clustering. Appl. Sci. **8**(9), 1468 (2018)
75. P.K. Mallick, S.H. Ryu, S.K. Satapathy, S. Mishra, G.N. Nguyen, P. Tiwari, Brain MRI image classification for cancer detection using deep wavelet autoencoder-based deep neural network. IEEE Access **7**, 46278–46287 (2019)
76. D. Kucharski, P. Kleczek, J. Jaworek-Korjakowska, G. Dyduch, M. Gorgon, Semi-supervised nests of melanocytes segmentation method using convolutional autoencoders. Sensors **20**(6), 1546 (2020)
77. B. Pyman, A. Sedghi, S. Azizi, K. Tyryshkin, N. Renwick, P. Mousavi, Exploring microRNA regulation of cancer with context-aware deep cancer classifier, in *PSB* (2019), pp. 160–171

Chapter 16
Deep Learning for Magnetic Resonance Images of Gliomas

John J. Healy, Kathleen M. Curran, and Amira Serifovic Trbalic

Abstract Gliomas are tumors that arise in the glial cells of the brain and spine. Gliomas are one of the most common brain cancers, and comprise 80% of all malignant brain tumours. Gliomas are classified by cell type, grade, and location. Prognosis of patients presenting with high grade gliomas remains poor. The gold standard for grading of gliomas remains histopathology, but a working radiological diagnosis can be established from a magnetic resonance imaging (MRI) scan. MRI is typically used throughout the patient pathway because routine structural imaging provides detailed anatomical and pathological information. In addition, advanced techniques can provide additional physiological detail. Traditionally, MRIs were read exclusively by radiologists, but improvements in machine learning has sparked considerable interest in its application to enhanced and automated diagnostic tools. Machine learning approaches are also of interest in monitoring the progression of low grade gliomas, and in monitoring of patients undergoing treatment. Convolutional neural networks (CNNs), especially when trained using transfer learning, have been shown to grade gliomas with up to 94% accuracy. Given an MR image of a brain tumour, we wish to manually segment the various tissues to aid diagnoisis and other assessments. This manual process is difficult and laborious; hence there is demand for automatic image segmentation of brain tumors. Public datasets and the BRATS benchmark have enabled clearer and more objective comparison of segmentation techniques. State-of-the art automated segmentation of gliomas is currently represented by deep learning methods. Deep learning is also now on the rise for prediction of molecular biomarkers. Novel approaches to explainable artificial intelligence are now required to aid the extraction of novel useful features from machine learning approaches. Finally,

J. J. Healy (✉)
School of Electrical and Electronic Engineering, University College Dublin, Dublin, Ireland
e-mail: john.healy@ucd.ie

K. M. Curran
School of Medicine, University College Dublin, Dublin, Ireland
e-mail: kathleen.curran@ucd.ie

A. Serifovic Trbalic
Faculty of Electrical Engineering, University of Tuzla, Tuzla , Bosnia and Herzegovina
e-mail: amira.serifovic@untz.ba

U. Kose and J. Alzubi (eds.), *Deep Learning for Cancer Diagnosis*,
Studies in Computational Intelligence 908,
https://doi.org/10.1007/978-981-15-6321-8_16

CNNs have been developed to make predictions of patient survival times. There are many exciting new directions for this field, from novel CNN architectures, to the integration of information from advanced MRI and complementary imaging modalities and spectroscopic techniques. In time, this may lead to clinically acceptable automation of a variety of radiological determinations, with positive consequences for patient outcomes. In this chapter, we will give an overview for engineers and computer scientists of the deep learning applications used in glioma detection, characterization/grading and overall survival prognosis of the patients. We will highlight the limitations and challenges of deep learning techniques as well as the potential future of these methods in prognosis, clinical diagnostics, and decision making.

16.1 Introduction

Gliomas are tumors that arise in neuroglial stem or progenitor cells—the glial cells of the brain and spine. Gliomas are one of the most common brain cancers. They comprise 80% of all malignant brain tumours, and cost patients more lost years of life than any other tumor. Gliomas may be classified by location, grade, and cell type. Prognosis of patients presenting with high grade gliomas remains poor. Risk factors are not well established for gliomas, with just two – high dosage exposure to ionizing radiation, and inherited mutations that are associated with certain rare syndromes. Therefore gliomas are a problem that can strike almost equally in male vs. female. The gold standard for the grading of gliomas remains histopathology, but a working radiological diagnosis can be established from a magnetic resonance imaging (MRI) scan. MR imaging is generally used through the entire patient pathway because routine structural imaging provides detailed information about the anatomy and pathology of the disease. In addition, advanced techniques can provide additional physiological detail. Traditionally, MRIs were read exclusively by radiologists, but improvements in machine learning has sparked considerable interest in its application to enhanced and automated diagnostic tools. Machine learning approaches are also of interest in monitoring the progression of low grade gliomas, and in monitoring of patients undergoing treatment. Radiotherapy, surgery, and alkylating agent chemotherapy remain the main treatments, but personalised strategies may ultimately improve outcome [91], and image processing has a role to play in helping rapidly and repeatably categorising tumours.

The chapter is structured as follows. We begin in Sect. 16.2 with the acquisition and processing of a MR images. In addition to image formation from the raw data collected, this traditionally includes such techniques as stripping the images of the skull, normalization and correction of intensity bias, and image denoising. Deep learning techniques may replace traditional approaches in many of these topics, and indeed may finally unify them into a single pre-processing step for computational efficiency. In Sect. 16.3, we consider processing problems related to the interpretation of medical images, including medical image registration and segmentation. Medical image registration finds an optimal spatial transformation that best aligns the underlying

anatomical structures in two or more images, and thus enables population analyses as well as comparison of structures across different scans and modalities. Medical image segmentation partitions an image, separating it into multiple regions based on similarity of different attributes. This enables the localization and quantification of organs or lesions. We also consider content-based image retrieval, which attempts to make features of medical images, such as anatomic regions or similar pathological category, searchable in order to assist clinicians in making decisions. In Sect. 16.4, we consider diagnosis and prediction. MRI is routinely used in the management of patients with glioma in the diagnosis, in the prediction of ultimate patient outcome, and in the assessment of the patient's response to treatment. MRI allows us to extract different kinds of information about a glioblastoma, including information about its structure and physiology, as well as functional information. This permits MRI multidimensional, in-vivo characterization of the tumour. The limitation of these noninvasive descriptors is that they are molecularly unspecific. Radiogenomics, that is the prediction of genomic properties of tissue using radiological imaging, has recently been established to study the relationships between the imaging features we can extract noninvasively and the corresponding molecular characteristics. We consider machine learning approaches to all of these problems, architectures, results, and challenges.

16.2 Image Acquisition and Processing

In this section, we will discuss image pre-processing steps, from skull stripping [34], to normalisation and the correction of intensity bias [89], to denoising of the MR images [15]. These defects directly affect the quality of segmentation results. It seems likely that deep learning techniques will replace traditional approaches in many of these topics, and indeed may finally unify them into a single pre-processing step for computational efficiency. Image registration is the transformation of different datasets into a common, unified coordinate system. If we want to directly and quantitatively compare or synthesise (fuse) the data obtained from different measurements, we turn to image registration techniques. As image registration is a particularly important topic, and as it is feature-specific, we will address it separately in Sect. 16.3.

We will then turn to artefact detection, focussing mainly on Gibbs ringing. Ultimately, that discussion is entangled with the question of super-resolution, which is the name for a broad category of techniques for obtaining an improved resolution image.

16.2.1 Pre-processing

The presence of the skull, along with other non-brain tissues such as the eyes and the dura, complicates registration and normalisation steps. As such, the crucial first step for many neuroimaging workflows is the removal of those non-brain tissues. This is referred to as skull stripping. Skull stripping can be integrated into segmentation algorithms, but is usually performed separately by specialised brain extraction algorithms such as Statistical Parametric Mapping, Brain Extraction Tool, Brain Surface Extractor, or Robex. The gold standard remains manual extraction, but this is time-consuming and (as in so many applications) it is desirable to automate it in order to (1) reduce inter-observer variance, (2) to facilitate analysis and reproducibility of large clinical studies. Automated extraction of the brain tissue is complicated by contrast and intensity variations arising from imaging protocol, imaging artifacts, and scanner platform variations. Age and other factors affect anatomical variability [23]. Tumours and changes arising from treatment can also alter the brain structure in a fashion that confounds algorithms. The border of the brain can be defined histologically as the transition of myelination from oligodendroglia to Schwann cells, which unfortunately is not visible on present-day MR scans. Finally, ground truth labels may themselves be ill defined for a variety of reasons. Kleesiek et al. (2016) proposed a CNN-based approach for skull stripping. They claim that their approach matches or betters state-of-the-art performance, which they demonstrate on a trio of public data sets [45].

Inconsistencies in the grayscale distribution of MRIs arising from variations in hardware and acquisition protocol complicates training of deep learing algorithms. Therefore, a pre-processing stage that normalises all images to have similar intensity distributions is a useful step before training or using a deep learning network. It is important that every modality to be weighted equally in training, so we want them all to have the same range. We refer to this step as intensity bias correction and normalization. Havaei et al. reported that the most popular pre-processing approaches reported by researchers in this area are as follows [32]:

- Application of N4/N3 bias field correction. (Alternatively, normalization of intensity using the mean cerebral spinal fluid value.)
- Removing outliers by truncating data in the lowest 1 or 0.1% of the histogram.
- Histogram normalization.
- Subtracting the mean, and normalising to have unit variance on all modalities, or at least on the selected training patches.

Bias correction and intensity normalization may be replaced with some success by the z-score computation [41, 88], but Pereira et al. have shown improved results when applying normalization prior to deep learning based segmentation procedures [67].

MR image acquisition corrupts the raw MRI data with statistical noise. The most suitable noise model depends on how many coils are in the system [29]. If the system is single-coil, the measured noise presents as a Rician distribution, whereas

multi-coil systems give rise to a non-central chi-distribution noise term. The k-space representation of the MRI data is a complex function of spatial frequency coordinates x_f and y_f, $D(x_f, y_f)$. $d(x, y)$, the inverse DFT of $D(x_f, y_f)$ contains real and imaginary components.

$$\Re(d) = s(x, y)\cos(\theta(x, y)) + n_{\text{Re}}(x, y) \tag{16.1}$$

$$\Im(d) = s(x, y)\sin(\theta(x, y)) + n_{\text{Im}}(x, y) \tag{16.2}$$

where the signal is $s(x, y)$, and $n_{\text{Re}}(x, y)$ and $n_{\text{Im}}(x, y)$ are the additive white Gaussian noise (AGWN) present in the real and imaginary components of the signal respectively. Finally, we see that the MR image is given by the absolute value of $d(x, y)$.

$$|d(x, y)| = \sqrt{\Re(d)^2 + \Im(d)^2} \tag{16.3}$$

Such a function, and hence any MR image, is corrupted by the presence of Rician noise. It is important to use the correct noise model here. An old technique estimated the noise power from the σ of the pixel intensity in a region of the image that did not contain any features, and that retain cause one to underestimate the true power of the noise by as much as 60% [29]. The Rician noise distribution approximates a Rayleigh distribution in low intensity regions, and a Gaussian distribution in high intensity regions. The noise sources in MRI acquisition include thermal noise in the receiving coils and inductive losses in the sample (depending on B0 and the sample volume size), the dimensions of the voxels, the bandwidth of the receiver and how many averages were used in the acquisition of the image.

Tumor classification is less accurate in the presence of noise. Hence, denoising of MR images has been studied extensively. Many authors have attempted denoising directly using deep learning techniques, e.g. Jain et al. proposed image denoising using CNNs and claimed to obtain performance similar to or improving on state-of-the-art approaches based on wavelets and Markov random fields even if the set of training images is small [37]. Latif et al. proposed a novel technique to denoise glioma MRIs with a deep CNN and anisotropic diffusion, and demonstrated improved segmentation based on a watershed transform [51].

Classical approaches to denoising are also still under investigation, not least because of the positive impact denoising can have on learning, e.g. the wavelet transform [10].

16.2.2 Artefact Detection and Superresolution

An MRI scan produces a series of samples of the $k-$space representation of the MR image, related to the image by means of the Fourier transform. That Fourier relationship results inevitably in an artefact called Gibbs ringing. Gibbs ringing

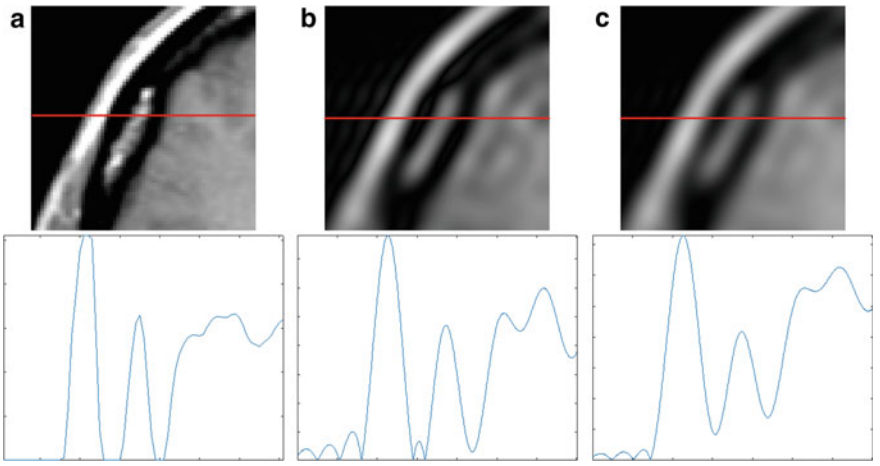


Fig. 16.1 Suppression of Gibbs ringing. We began with an image from "Axial post-contrast T1 (top) and T2 (bottom) weighted MRI showing an IDH1 mutant frontal lobe glioblastoma with sparse enhancement despite large size", by Jpoozler, licensed under CC BY-SA 4.0 (https://creativecommons.org/licenses/by-sa/4.0/). We cropped it to produce **a**. Truncation (which causes Gibbs ringing and loss of resolution) and zero padding in the Fourier domain (which causes upsampling) produces **b**. Image **c** is made using filtered Fourier reconstruction to suppress Gibbs ringing. The red lines indicate the rows of pixels which are then plotted on the corresponding lower row of plots. Note the additional lobes in **b**, which are suppressed by the filtered Fourier reconstruction. Deep learning approaches have been shown to improve on filtered Fourier reconstruction in this problem

arises from a limitation in the physics of image formation in MRI, and is exacerbated by the truncation of the $k-$space representation that inevitably arises from a finite series of measurements. Gibbs ringing is an oscillation that occurs around sharp edges—tissue boundaries in this case—in a function reconstructed from a truncated Fourier series. The ringing is a mathematical property of the Fourier transform. It reflects the inherent difficulty in approximating a discontinuous function by a finite sum of continuous functions (Fig.16.1).

Traditionally, Gibbs ringing is suppressed by a mixture of filtering methods and reprojection methods such as the Gegenbauer polynomials [77]. Hybrid methods have also shown promise, retaining the speed of the filtering methods and the accuracy of the reprojection methods [24]. Older machine learning methods, including autoregressive modeling and multilayer neural networks, have been used to reduce the ringing artifact. That reduction is achieved by estimating the missing high frequency components from the low frequency data acquired by the MRI [4, 80, 98]. Deep learning approaches have been investigated extensively over the past four years for suppression of Gibbs ringing [62, 84, 90, 100, 103].

Wang et al. is the earliest paper on this topic that we are aware of [90]. The authors reduce the k-space of an MRI by 70% in one direction only, replacing the values with zeros. Signals reconstructed from this reduced k-space exhibit Gibbs

ringing. The authors compare a deep learning approach with two filtered Fourier approaches, using Hamming and Tukey window functions. They assess the quality of the reconstructions using peak SNR, the high-frequency error norm, and structural similarity. Their results show that they gain a few dB improvement in PSNR over the filtered Fourier results.

Muckley et al. examine a deep learning approach for suppressing Gibbs ringing and other noise in a diffusion weighted MRI [62]. The authors observe that diffusion MRI is a particularly suitable test for Gibbs ringing suppression as it can have drastic effects on diffusion parameter maps. The calculation of higher order diffusion parameters can require high b-values, "which, coupled with strong diffusion gradients and long echo times, leads to low signal-to-noise ratio (SNR)" and significant errors in kurtosis maps. The proposed model is a modified U-net with over 31 million parameters, incorporating wavelet transforms for downsampling and up sampling. The authors train their model entirely on simulations of MRI acquisitions with images from the ImageNet data set, claiming that this mitigates the risk of overfitting. They compare the output of their network with the MP-PCA method for denoising followed by subvoxel shifting, and with other CNNs of their own devising. They test their networks on rotations of an edge spread function (knife edge), using the full width half maximum of the linespread function at various contrast-to-noise ratios. They further tested the network on 40,000 images from the imagenet test data using the power spectral ratio, and on a measurement of the diffusion weighted MRI of a volunteer's in-vivo head. In spite of this commendable array of tests, the results are largely qualitative. The details of and motivations behind the noise model used are also unclear.

Zhang et al. [100] proposed a 4-layer convolutional network for Gibbs suppression. Their training data consisted of over seventeen thousand T2-weighted images of the brains of 136 healthy adult subjects, which were collected as part of The Human Connectome Project. The test data included T2-weighted images of normal and pathological brains, diffusion-weighted brain images, and T2-weighted images of knees. The final image with suppressed ringing is found by subtracting the output of the network (which estimates the error) from the input. The image is then Fourier transformed, and the low-frequency data are replaced by the Fourier transform of the input. Both of these ideas are recurring in this field. Praxis has shown that training on the error introduced by Gibbs ringing is easier than training on the desired image. Several authors have also treated the low-frequency data as strictly correct, which better preserves mid-band frequencies but calls into question whether they are suppressing Gibbs ringing or simply upsampling the image. Zhang et al. use a number of metrics to evaluate their results, including RMSE, peak SNR, and the structural similarity index. They compare the images resulting from their proposed method with the images produced using a filtered Fourier reconstruction using a Hamming filter. They also compare with the LSS method. Their results show PSNR gains of just a few dB.

Zhao et al. [103] propose an end-to-end model which they derive from the EDSR model. EDSR was originally designed for super-resolution of a single natural image. They used transfer learning to train on 600 images from the div2k dataset before

turning to the IXI dataset for MRI images, which has 577 subjects scanned to produce T1- and T2-weighted and proton density images. They argue that if you compare medical images with natural images, the former are simpler in terms of structure and details of texture. Hence transfer learning should be effective in this context. They simulated Gibbs-ringing artifacts by means of a truncating the image in k-space by 75%. We can compare this idea with the results of Zhang et al. [100], who trained on T2-weighted images only, but successfully validated on several other types. Muckley et al. [62] also trained on natural images, arguing it limits the risk of overfitting. Zhao et al. compared the results of their network with bilateral filtering, non-local MRI upsampling, and the CNN architecture of Wang et al. [90], using average PSNR and the structural similarity index. Zhao et al. report only an ≈ 3 dB improvement over filtering, although the results in their Fig. 16.2 seem promising. This only serves to highlight the difficulty in comparing different results in this field. Authors use different metrics on very different test data.

In general, authors test and train their networks for Gibbs ringing on truncated k-space data. Many authors imply that somehow the problem is simply *truncation*, e.g.

- "In Magnetic Resonance Imaging (MRI), the K-space data is often under-sampled and truncated to shorten the scan time. However, the truncation of K-space also

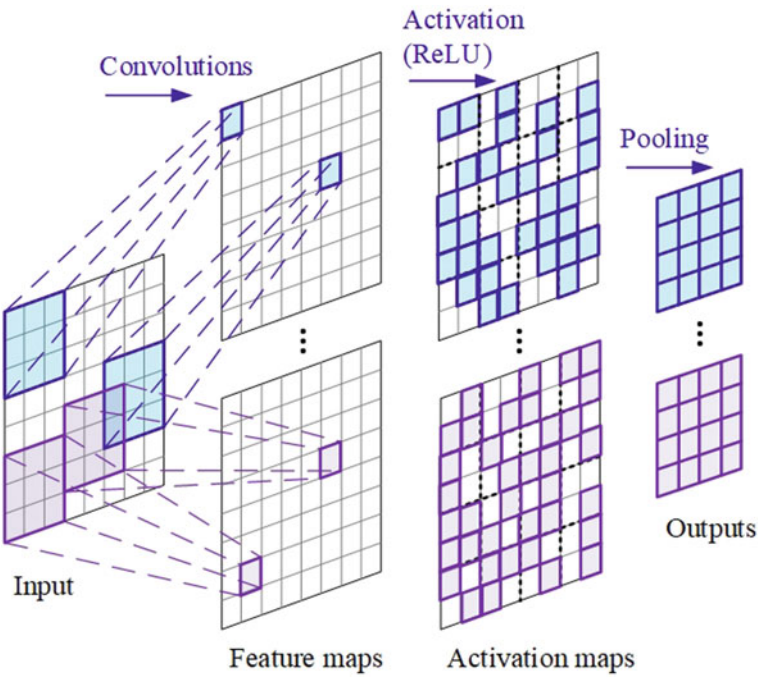


Fig. 16.2 The structure of a CNN

causes Gibbs ringing artifacts in the image, which seriously deteriorates the image quality" [90].
- "Gibbs- ringing artifact in MR images is caused by insufficient sampling of the high frequency data" [100].

Gibbs ringing is inevitable. No matter how many measurements we take in k-space, a Fourier reconstruction does not converge correctly at a discontinuity, exhibiting an $\approx 9\%$ error in the limit. Does a truncated k-space truly simulate Gibbs ringing? To what degree are we simply seeing a deep learning based superresolution rather than true suppression of Gibbs ringing? Zhao et al. [103] specifically use a model from the super-resolution literature, further blurring these lines. Some authors implicitly assume a fixed amount of truncation in their images, e.g. [90, 103]. Zhang et al. [100] noted that, in clinical practice, it may not be possible to establish the real degree of truncation and the extent of Gibbs ringing in MRIs. They trained and successfully validated a model using mixed data with truncation levels from 30 to 50%. This only solves part of the problem, but it does appear to be an important consideration.

Modern MRI systems practice super-resolution by means of compressive sensing. A random or pseudo-random subset of the k-space measurements is fed into a compressive sensing algorithm that then produces an MR image of normal resolution. The advantage of this process is a large reduction in how many measurements are required, which reduces scan times, improving patient throughput, reducing costs, reducing motion blur, and improving patient comfort. To date, none of the investigations of deep learning for Gibbs' ringing suppression in MRI has investigated the effects of compressive sensing. Muckley et al. claimed that "with CNN processing, high-quality parameter maps can be calculated from abbreviated acquisitions with a little more than half of the encoding time required for traditional full acquisitions" [62], but this must be demonstrated to outperform compressive sensing if we are to consider deep learning a truly cutting edge tool for reconstruction of MRI images. Recent developments in dynamic MRI reconstruction may lead the way in this regard [73].

16.3 Contextual Processing Prior to Diagnosis and Prediction

16.3.1 Image Registration

Medical image registration is a basic technique in many medical imaging problems. The goal is to find a best spatial transformation to align the underlying anatomical structures, and in doing so to simplify tasks like comparing different scans of the same structure, or to analyse similar structures in a large pool of patient data. The former kind of analysis helps us to understand how the brain anatomy of a patient evolves over time or to overlay multiple modalities simultaneously or to monitor tumour

growth, and the latter helps us understand inter-patient variability by normalizing the brain scan to a brain atlas [81].

Several studies have applied deep learning (DL) techniques to improve medical image registration. In this section, we will summarize cutting edge developments and challenges in registration of brain MRIs. In general, there are two prevalent strategies are in the literature: either using DL networks to estimate a similarity measure for two images and hence to drive iterative optimization strategies, or to use the networks to directly predict transformation parameters.

16.3.1.1 Deep Learning Similarity Based Methods

Some DL-based approaches to registration of brain MRIs attempted to replace the traditional image similarity measures with deep learning based similarity measures, reporting promising results. These similarity measures are inserted into a classical deformable registration framework with a defined interpolation strategy or output from transformation model, and optimization algorithm. The multimodal 3D case was considered in [78], the authors of which proposed a suitable similarity metric based on a CNN. The network could be trained without transfer learning or output from just a handful of aligned image pairs. They then incorporated this metric into first-order continuous optimization frameworks. The network output replaced mutual information in a deformable image registration for brain T1-T2 registration.

As with so many other applications of deep learning, the quality of the training data is critical to supervised learning. Accurate image alignment of training images is necessary, but very difficult to obtain.

Therefore, some researchers applied the DL in an unsupervised manner. For example, the authors of [93] presented an unsupervised DL approach. They combined convolutional layers with independent subspace analysis to extract features from various image patches in the brain MRI, and hence they learnt basis filters adaptive to local and global features. Rather than handcrafted features, their feature maps were used in the HAMMER registration algorithm. Stacked autoencoders (SAE) can be used similarly to learn feature maps without supervision [94]. Low-dimensional feature maps from the SAE are then used for calculating the similarity of images, and the learnt representations drive the HAMMER and Demons deformable image registration algorithms.

16.3.1.2 Deep Learning Transformation Estimation

In the previously described approaches, DL-based similarity metrics have shown potential to outperform traditional metrics for the same task, but the estimation of the transformation in these methods is still iterative; this slows the process of registration (particularly in high-dimensional cases, e.g. when the registration may be deformable). The optimization algorithms also require a smooth derivative, which can be tricky to difficult to guarantee. This observation prompted the exploration

of new deep learning approaches that estimate the parameters of the registration transformation in a single step. In [99], a patch-based deep encoder-decoder network was designed that learns the necessary mapping, using the large deformation diffeomorphic metric mapping registration methodology. In [6], a patch-based CNN regression model is designed to learn the mapping. A strategy called equalized active-points guided sampling mitigates the problems associated with a limited training set. Auxiliary contextual cues guide the learning more directly. A CNN was used also for implementation of parametrized registration function learnt from a collection of volumes in a unsupervised way [3]. A non-rigid image registration algorithm that uses fully convolutional networks is presented in [52]. Authors reported that their method estimated the registration transformation directly by maximizing an image-wise similarity metric between fixed and deformed moving images, which is similar to conventional image registration algorithms.

16.3.1.3 Challenges

Image registration of MR images of gliomas, as a pre-processing step, has close associations with the accuracy of the consequent quantification steps such as MR image segmentation of the subregions within brain tumors, computer-aided diagnosis and longitudinal studies for treatment response follow-up which are highly dependent on it.

Also, it is an essential step in analyzing brain tumor MR imaging data from multiple images, because it ensures the spatial correspondence of anatomy across different modalities for diagnosis and treatment. Many commercially available brain MR image analysis software have some implementation of image registration, and description and evaluation of such techniques can be found in the literature [46].

Supervised methods have allowed the robust registration for several applications. For supervised methods, there is a problem of lack of training datasets with known transformations for network training. The quality of the registrations that can be achieved using this kind of framework is ultimately dependent on the quality of the ground truth registrations. Some published works propose, as a solution, an application of data augmentation techniques for artificial transformations generation. Compared to supervised transformation estimation, unsupervised methods do not suffer from this lack of suitable training datasets, but it remains challenging to define a suitable loss function for the CNN without ground truth transformations. Therefore, a variety of regularization terms have been discussed in the literature. The problems that may occur in brain tumor image registration are much more observable when effects of surgery have to be considered, for example the registration of pre-operative images to intra-operative or post-resection tumor images.

In contrast to brain tumor classification and segmentation from MR images, there are few papers on DL techniques in brain MRI registration and there is little sign of convergence towards a best approach. We foresee more and more research publication on this topic before long.

16.3.2 Image Segmentation

Segmentation is the technique of isolating the regions of an image that contain distinct parts, anatomical or pathological regions as an organ, legion, or tumour. This step localises and quantifies the segmented objects from the background, and is a key step in many algorithms for further processing. This is a very well studied problem, but even here we see deep learning approaches beginning to displace traditional methods.

Our main focus in this section is on segmenting gliomas in brain MR images, a key step for initial treatment planning and determining disease progression. There are a number of standard MRI acquisition protocols used in brain tumour patients; these include T1-weighted (with or without gadolinium enhancement), T2-weighted, and T2-weighted with fluid attenuated inversion recovery (FLAIR) sequences. This particular segmentation problem is very challenging: the tissue within the region is heterogeneous, the intensity values of MRIs are nonuniform (especially across protocols), and the borders of the tumors are unclear, diffuse, and discontinuous. Brain tumor MRI data are inherently complex 3D data the characteristics of which can vary greatly between patients. Effective segmentation requires the use of multiple acquisition protocols, which introduces registration error to the problem.

The range of different brain MRI segmentation methods is extensive; comprehensive reviews have been carried out [19, 35, 39, 81]. The methods can be cagtegorised as: manual, semi-automatic, or fully automatic, depending on the amount of interaction required from the operator. Manual segmentation of the many MR images generated in the standard clinical routine is tedious and time consuming, and subject the user to intra- and inter-observer variability. Hence, automatic methods are desirable.

Semi-automatic methods can require user interaction for an initialization in which (s)he defines a region of interest;, or they may require an intervention or feedback response, in which the operator steers the automated algorithms towards the desired results. While generally less time consuming than the manual methods, the semi-automatic methods are also prone to intra- and inter-observer variability.

Therefore, current segmentation research in this field is intensely focused on the third class of methods: fully automatic methods. In these methods, there is no need for user interaction, instead, a prior knowledge and AI are used in combination to segment the images. Recently deep learning techniques have shown a great deal of promise and have become a leading light in the area.

16.3.2.1 Deep Learning Network Architecture

Many deep learning approaches are used by scientists for segmenting brain tumours. The main focus of the literature on this topic is on network architecture design to perform feature extraction; less attention is paid to image processing approaches that can work in tandem with the neural networks to improve performance. We can identify three approaches applied specifically to brain tumour segmentation:

patch-wise, semantic-wise, or cascaded architectures. We will now introduce the most significant network architectures used for glioma segmentation, and discuss the advantages and disadvantages of each one.

Convolutional Neural Networks (CNN): In medical image processing, the interest in DL mostly revolves around CNNs. CNNs are a type of neural network that are able to extract and learn typical complex features from training data. They consist of a sequence of layers, each of which performs a specific operation; common examples include convolution, and pooling. Each layer takes the output of the previous layer as its own input, and in turn its own output becomes the input to the subsequent layer. An input layer will hold the pixel values of the image. This is followed by a set of convolutional layers composed of a certain number of filters, also called kernels, that are convolved with the input data to obtain feature maps. Convolutional layers are typically followed by pooling layers to introduce nonlinearity to the activations. There are many possible activation functions; among the most common are the rectified linear unit (ReLu) and derived types such as Leaky ReLU. ReLU is very simple and computationally inexpensive, in addition to being highly non-linear: all negative inputs are zeroed. It depends on the design, but convolutional layers are frequently followed by a pooling layer. The aim of a pooling layer is to gradually reduce the dimensionality of the network while retaining the most salient feature information. There are two popular approaches to pooling: averaging and max-pooling. The last part of the network consists of fully connected layers, which endeavour to extract high-level abstractions of the data. The weights and biases of the network are optimized by the training phase. The structure shown in Fig. 16.2 is an example of a CNN. A diverse collection of approaches presented in the literature utilize CNNs for brain segmentation. Many of the currently popular CNNs used in the segmentation problems take a patch-wise approach.

In [104], a standard CNN implementation based on multi-channel 2D convolutions has been used, and it was adapted such that it operates on multichannel 3D data usually available for the brain tumor segmentation task. Another 2D CNN architecture was presented in [67], exploring the use of small (3×3) convolutional kernels. This approach mitigated against overfitting by assigning fewer weights in the network. They also examined patch intensity normalization as a preprocessing step together with the training data augmetation. A similar approach, using a 2D patch-wise CNN architecture was presented in [20, 101]. Later, Haveaei et al. proposed a two-pathway-based segmentation method [31]. The two pathways consist of a small convolution filter for local segmentation and large filter for global segmentation. Later, the the two pathways are reunited to produce the segmented image. They found that they could achieve better results if they cascading their pathways, feeding the output probabilities of one CNN as additional inputs to a second network. Training is split into two phases in order to deal with difficulties arising from imbalanced tumor labels (i.e., the scans contain far more 'normal' voxels than tumor voxels). One of the first 3D CNN models was introduced in [86]. MRI is inherently 3D data, and it is natural to consider that a 3D model may be advantageous. Their idea was followed by the DeepMedic model, a multiscale, dual-path 3D CNN, presented in [40]. The second pathway received patches from a decimated version of the image. 3-D CNNs

are computationally expensive, having a very large number of parameters to tune. Milletari et al. proposed an approach that combines a CNN with Hough voting on feature extracted by the network [60].

Fully Convolutional Network (FCN): In semantic-wise approaches, entire images are divided into regions; the outputs of different regions are then stitched together using additional convolutional filters. One example of such a structure is the fully convolutional network (FCN) [54], which replaces the later fully connected layers with a convolutional layer. This permits the network to arrive at a dense prediction, pixel-wise, and it produces excellent results. Higher resolution activation maps are used to achieve better localization by means of pixel-rather than patch-wise predictions. FCN has been used for brain tumor segmentation in research presented in [9].

U-Net: U-Net is a widely known and applied architecture for medical image segmentation (Fig. 16.3). It builds on FCN, employing some modifications avoid having to trade context off for localization. Big patches require more pooling which reduces localization accuracy, but small patches lack context. U-Nets therefore consist of two paths, called analysis and synthesis. The analysis path (left side of the network) captures context; the synthesis path (right side of the network) enables precision in the localization. The contracting path is fairly comparable to a regular CNN and the expansion path consists of upsampling and deconvolution layers. In both pathways, convolutional layers with dropout are followed by ReLU layers and pooling. The two pathways are linked by skip-connections. These skip-connections are necessary to provide high resolution features to the deconvolution layers.

Dong et al. [17] developed an automatic segmentation technique for brain tumours based on the U-Net, and validated it with five-fold cross-validation. This architecture was also successfully applied in [43], where a Jaccard loss function was used to handle imbalances in the training data. In [36], a modified U-Net was implemented which uses a Dice loss function for the same purpose, along with extensive data augmentation to prevent overfitting. A 3D U-Net architecture for both brain extraction and tumor segmentation was presented in [7]. Also, a U-Net-like network was proposed that introduces a new loss function, generalizing binary cross-entropy [59]. This handles label uncertainty well.

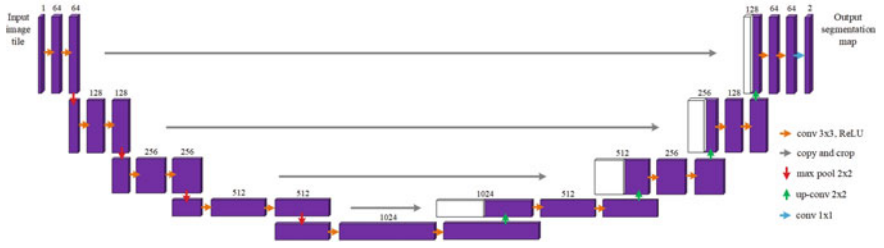


Fig. 16.3 The structure of a U-Net. Violet boxes represent feature maps. The number of channels is denoted above each feature map

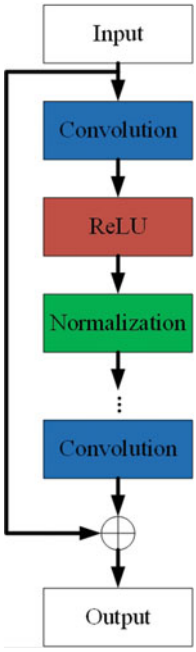


Fig. 16.4 A residual block of a RN

Residual Networks (RN): Residual networks are characterized by the use of batch normalization and the introduction of residual blocks periodically stacked throughout the network. A typical RN consists of multiple residual blocks and batch normalization between every convolutional layer and ReLU layer. These stacked residual blocks greatly improve training efficiency and largely resolve the degradation problem by employing batch normalization. The RN architecture and residual blocks are shown in Fig. 16.4.The 2D RNs have been demonstrated convinvingly in many 2D segmentation tasks, but so far insufficient studies have applied residual learning on volumetric data. Among the few is the voxelwise residual network (VoxResNet), proposed in [12], a 3D deep residual network. Residual learning reduces the degradation problems that arise when training a deep network; hence increasing the network depth can be used without qualm. VoxResNet has no fewer than 25 layers, and so it can generate feature maps of great generality. The network inputs are multi-modal to exploit as much data as possible.

Recurrent Neural Networks (RNN): The RNNs are networks empowered with recurrent connections which enables the network processing of sequences with long term dependencies. Since the ROIs in MR images are usually distributed over multiple adjacent slices, there are correlations between successive slices. Accordingly, RNNs can extract context by treating the slices as sequential data. The structure of an RNN extracts intra-slice information in two sections (which are any type of CNN), and the RNN itself which combines that information. In [28], the BraTS-17 dataset was selected to test the architecture on high grade glioma segmentation.

Long short-term memory networks (LSTM) are the most famous type of RNN. LSTM is an upgrade on the RNN, designed for sequential data. This is clearly a disadvantage for medical image segmentation as it discards the spatial information, which is undesirable, so LSTMs were modified to perform convolution in place of multiplication. Pyramidal Multi-Dimensional LSTM (PyraMiD-LSTM) systems [82] use a remarkable topology for tumour segmentation. The strategy is easily parallelizable, and less computationally intense overall [96].

Generative Adversarial Networks (GAN): GAN is a fascinating variation on the CNN that produces great results even with limited training data. It models a min-max game between which two neural networks, the generator and the discriminator. The discriminator is trained to classify inputs into two classes: ground-truth, and synthetic examples generated by the generator network. The generator is trained to produce new images from a noise input. Both networks are trained in an alternating manner; both gradually improve at their tasks. In [26, 97] a novel pair of adversarial network architectures called SegAN and SegAN-CAT were proposed. One is a segmentation network and the other is a discriminator network. The segmentation network takes as an input an MRI slice and generates a probability label map. The discriminator network takes as an input an MRI slice and its probability label map, that can be either the one computed by the segmentation network or the ground-truth one.

Auto Encoders (AE): An autoencoder is a simple neural network which can be divided into two parts: an encoder and a decoder. The encoder is used to generate a reduced representation of the image, and the decoder reconstructs the input from the encoder's output. Different AE variants have been used by researchers to segment brain tumours [1, 63]. Stacked denoising autoencoders using limited training data are deployed in [87]. Other works that feature autoencoders include: [48, 95].

Ensemble of Deep Learning Methods: Ensemble DL methods combine several baseline DL models, and use some combination of their prediction to segment the brain tumor from MRIs. Examples of ensemble methods can be found in numerous research papers, such as [16, 42, 102].

16.3.2.2 Network Training Techniques

Training of the deep learning networks can be: supervised, transfer learning or unsupervised training. Supervised training uses an expert-annotated dataset to allow the network to learn how to map input data based on a pixel- or voxel-level annotation. The training data consists of examples for which there is a pair: input data and a desired corresponding label or labels. Supervised training can be done by just doing the supervision at the output layer, or by deep supervision, where the direct supervision of the hidden layers is provided and propagated to lower layers. As discussed elsewhere in this chapter, annotated datasets have associated problems, and are difficult to come by for many problems.

Transfer learning is an approach to address the lack of available annotated medical image data. In transfer learning, a system reuses an previously trained model for a related task as the starting point for training on its task. These pretrained models reduce training time, but also provide a first approximation to boost performance

when the training data is limited. Naturally, transfer learning works better the more alike the two tasks are, but some research has shown that transferring the knowledge from far distant tasks is better than the random initialization. There are many pretrained models provided by major AI companies (e.g., GoogLeNet and AlexNet) which can be used to develop new AI applications.

Unsupervised training algorithms are those algorithms that can learn image features from an unlabelled dataset. They are an important tool for addressing the limitations of medical datasets. In recent years, generative models such as GANs and AEs have become dominant techniques for unsupervised deep learning.

16.3.2.3 Challenges

DL methods use a great deal of training data for segmentation of gliomas and other brain tumors on MR images. It is non-trivial to acquire suitable data for training DL models. Therefore, one of the major limitations in medical image segmentation is data scarcity. Data scarcity can lead to over-fitting, when a model has a good performance on training data but cannot perform well on unseen test data. Usually, a sufficiently large number of annotated medical images for is not available for training segmentation algorithms for multiple reasons which have been discussed elsewhere in this chapter. In addition, sharing patient medical data has many implications, legally, and in terms of privacy, data-ownership and technical issues. It is even more complicated at the international level. One can shrink the network in terms of complexity, but the preferred solution is data augmentation, which is used to produce new synthetic images by minor transformations of the data and the corresponding ground truth. Examples of data augmentation approaches include rotations, translations, horizontal and vertical flips, scaling, brightness variation, and elastic deformations [65]. The complexity of the network architecture is directly related to the computational cost of DL networks.

Another challenge in medical image segmentation is imbalanced data. We have seen that this arises for the case of gliomas too, as the pathological tissue is much smaller than the healthy tissue. Training with the imbalanced data can result in an unstable network, biased towards the larger class. A potential solution for this problem is to apply weighting to samples during training, where a higher weight will be assigned to the less extensive ROI.

Different DL techniques, presented in this section, can segment tumours from edema and from normal tissue. These techniques are promising with at least moderate accuracy.

The performance of many segmentation algorithms for MRI brain tumour is compared with the ground truth images of MIC-CAI Brain Tumor Segmentation Challenge (BRATS) dataset 2012–2018. In this dataset, four types of tumor are annotated by radiologists. The metric of choice to define accuracy of algorithm for automatic image segmentation is the Dice coefficient, that measures the overlap between the manual delineated brain tumor regions and the segmentation results of an automatic method. Table 16.1 gives the summary of some deep learning approaches used to segment MR image with glioma.

Table 16.1 A summary of some deep learning architectures and approaches used to segment glioma MR images

Publication	Approach	Database	Results (dice scores)
Zikic et al. [104]	CNN in a sliding-window fashion in the 3D space	MICCAI-BRATS 2013 dataset	0.84 whole tumor, 0.74 core, 0.69 enhancing tumor
Urban et al. [86]	3D CNN with 3D convolutional kernels	MICCAI-BRATS 2013 dataset	0.87 whole tumor, 0.77 core, 0.73 enhancing tumor
Pereira et al. [67]	CNN with small 3×3 convolutional kernels	MICCAI-BRATS 2013 dataset	0.88 whole tumor, 0.83 core, 0.77 enhancing tumor
Havaei et al. [31]	A cascade neural network architecture in which the output of a basic CNN has been treated as an additional source of information for a subsequent CNN	MICCAI-BRATS 2013 dataset	0.88 whole tumor, 0.79 core, 0.73 enhancing tumor
Kamnitsas et al. [40]	The DeepMedic model, a multiscale, dual-path 3D CNN	MICCAI-BRATS 2015 dataset	0.85 whole tumor, 0.67 core, 0.63 enhancing tumor
Cui et al. [16]	Tumor localization network (TLN)—a fully convolutional network (FCN) in conjunction with the transfer learning technology	MICCAI-BRATS 2015 dataset	0.90 whole tumor, 0.81 core, 0.81 enhancing tumor
Myronenko et al. [63]	Semantic segmentation network for tumor subregion segmentation from 3D MRIs based on encoder-decoder architecture	MICCAI-BRATS 2018 dataset	0.88 whole tumor, 0.81 core, 0.76 enhancing tumor
Isensee et al. [36]	A modified U-Net was implemented which uses a Dice loss function to cope with class imbalances and extensive data augmentation to address overfitting	MICCAI-BRATS 2017 dataset	0.9 whole tumor, 0.8 core, 0.73 enhancing tumor

Much research has been carried out in this particular segmentation field, but application of the segmentation methods in the clinics is still limited. In many cases, a manual tumor delineations are preferred clinically, probably due to the lack of communication between the relevant researchers and clinicians [81], and to the lack of interpretability of results which fails to inspire confidence. There are many factors that stand between this research and its widespread clinical adoption, including user-friendliness, demonstrably robust and interpretable results, and yet higher accuracy.

16.3.3 Content-Based Image Retrieval

With widespread use of digital imaging and Picture Archiving and Communication Systems (PACS) in the hospitals, the size of medical image collections are increasing rapidly. PACS is used to manage a vast library of medical image data compatible with the DICOM file format. However, the search capabilities provided by PACS are limited to the textual information stored in the DICOM header such as a textual keywords, including patient name, identifiers, date, modality, body parts examined and image device. Therefore, there is a need to develop an effective medical image retrieval system to aid the clinicians in browsing and analysing these large image repositories. The goal of content-based image retrieval (CBIR) in a medical setting is to retrieve similar medical cases given an image. This might involve retrieving the same anatomic regions, or retrieving lesions of clinical relevance (e.g., lesions with the same pathology) in order to assist clinicians in the decision-making process. The performance of a CBIR system depends largely on the representation of images in terms of their features and selected similarity measures.

A number of studies have developed CBIR systems in contexts relevant to this chapter, e.g. [18, 33, 49, 61]. Application of CBIR for brain tumors in MRI is difficult because of the complex appearances of tumors. Defining the similarity between two images of brain tumors is an ill-posed and very challenging goal, because there are multiple characteristics that must be taken into account, like type, shape and localization in the brain. Brain tumors of the same type may present very different appearances in different patients, and indeed different types of brain tumours may have share visual similarity. In addition, there is a problem of the "semantic gap", i.e. the information we lose when we convert a low level representation such as an image into a high-level representation such as its visual features.

In order to overcome these problems, some medical CBIRs found in the recent literature are implemented by DL algorithms. For example, in [69], a CNN-based DL CBIR algorithm was developed to infer the contrast of MRIs based on the image intensity of a number of slices. DL methods have recently also been used to generate radiology reports incorporating LSTM models. In [83], a CBIR system for brain tumors on T1-weighted CE-MRI was presented. To overcome the semantic gap between the MR image and the high-level information perceived by the medical expert, a deep CNN VGG19-based novel feature extraction framework was proposed to quantify the similarity between the query and database images. Experiments were

performed on a publicly available CE-MRI dataset consisting of glioma, meningioma, and pituitary tumours; the results they obtained are promising. In [68], they added multitask learning, and implemented a corresponding tumor image retrieval system. Validation of their solution was carried out on annotated brain tumor data.

16.3.3.1 Challenges

Content-based medical image retrieval is a very difficult task, with a large potential in the medical field. It has a significant role in a domains of diagnostics, teaching and research. To be used as an aid in diagnosis, CBIR algorithms for brain tumor MR images need to prove their performance and they need to be integrated into daily clinical practice. One of the inherent problems in CBIR is the semantic gap due to the inconsistency between the features extracted and the user interpretation of an medical image. But, as a CBIR application becomes more specialized for a limited domain, the semantic gap can be made shrink using expert domain knowledge. An additional problem with the CBIR application is the annotation is inherently subjective, depending largely on the annotator. Annotations vary not only between people, but also in time for the same person depending on their specific task at hand. User interaction and use of high-dimensional data are further important issues. To obtain best possible results, all types of interaction need to be exploited.

16.4 Diagnosis and Prediction

16.4.1 The Role of Radiogenomic Imaging in Diagnosing and Predicting Glioma

MRI is routinely used in the management of patients with glioma in diagnosis, outcome prediction and assessment of treatment response. We can provide a rich characterization of a glioblastoma non-invasively and in-vivo using MRI, as it can extract physiological, structural, and functional information. Unfortunately, that characterisation remains unspecific at the molecular level. Radiogenomics, that is the prediction of the genomic properties of given tissue using only the non-invasive modalities of radiological imaging, has recently developed. Radiogenomics studies the relationships between imaging features and molecular characteristics [25, 44]. Studies have demonstrated the value of using these molecular biomarkers: they allow us to group gliomas with similar clinical behaviours, similar therapy responses, and even similar patient outcomes. If we can identify these biomarkers, it may enable us to target specific treatments to each patient, to their benefit [38].

We use the term “molecular” instead of “genomic” because many important findings are not technically genomic, but they may be epigenomic (such as methylation of some portion of a genome, which then typically acts to deactivate the gene) or molec-

ular level or chromosomal, without representing any true alteration of the gene. The World Health Organisation announced new classifications in 2016 [55], in which they recognized five molecular subtypes of diffuse gliomas. These subtypes were based on isocitrate dehydrogenase (IDH) and 1p/19q genotypes as well as histologic phenotypes. Traditionally, the job of radiological imaging has been to image a patient's phenotype, i.e., to assess quantitatively the product of the genotype, the impacts of environment (such as trauma) or habit (including diet and exercise) on that particular individual.

16.4.1.1 Stratification of Molecular Sub-types in Glioma

There is an urgent clinical need to find out whether clinical MRI is capable of stratifing these molecular sub-types to support the diagnosis, monitoring, and treatment monitoring of brain tumours. Several studies to date have show that there are associations between observable imaging features and molecular characteristics. These include the location of the tumor [21], its volume [64], enhancement [72], invasiveness [14], edema [105], and diffusion restriction [70]). There are studies that suggest that simple volumetrics may be able to predict when low grade glioma (LGG) are likely to convert to glioblastoma multiforme (GBM) [105]. We should note that this result is controversial [71], but more sophisticated quantitative/DL methods may improve our knowledge about of this problem. Improvements in molecular classification of tumors have the potential to make personalized medicine and targeted therapies a clinical reality.

In order to predict the prognosis of a patient, or to make preoperative treatment plans, it is critically important to successfully grade glioma. In particular, it is of the utmost importance that we reliably differentiate between LGG (grades II and III) and higher grade glioma (HGG, grade IV) [53, 92].

16.4.2 Traditional Machine Learning Methods

Researchers have applied many machine learning methods, mostly supervised learning methods, to MRI data of brain tumors. As discussed earlier in the chapter, supervised learning occurs when there exists a known ground truth to each data point, which is typically labeled by experts. We split the data into training, validation, and testing datasets. We use the training set to develop our models, testing the accuracy by cross-validation on an internal subset of the data. We hold-out the testing set in order to be able to ultimately evaluate if the the model we have trained generalises to unseen data. E.g., if a project were to classify gliomas into IDH wild-type versus IDH mutated, the label would be the specific genomic subtype, the IDH mutation status. The training algorithm attempts to find those network weights and biases that minimize the error using some cost function between the machine learning algorithm output and the ground truth of genomic sub-type.

Tian et al. [85] use a Support Vector Machine (SVM) model, texture and histogram MRI features to attempt to classify grade III and grade IV gliomas. They report high accuracy of 96.8% for classifying low-grade from high-grade gliomas and 98.1% for classifying grades III from IV. Crucially, evidence suggests that histologic grade may not be a good predictor on its own of patient outcome, as outcomes are more closely tied to molecular sub-types rather than histologic grade [66]. Furthermore, imaging features have been found that relate to these molecular sub-types better than they relate to histologic grade [75].

K-means clustering is an example of an unsupervised machine learning method. Tumors in the training data are unlabeled. The algorithm discovers clusters of tumors it deems similar, taking into consideration, e.g., shape, volume, and texture features. Rathore et al. (2018) [74] applied k-means clustering and found three distinct sub-types of IDH-1 wild-type glioblastomas with unique imaging, molecular, and clinical, characteristics.

16.4.3 Machine Learning Based Radiomics for Gliomas

Feature discovery and extraction is a critical task in radiomics, the field that extracts quantitative features from radiographic images. To date, most radiomics studies on brain tumors extracted features typically including tumor shape, location, intensity, texture, and wavelet features. In the study of gliomas, various radiomics features have been discovered to be useful for genotype classification [74], grading [5], and outcome prediction [75].

Hand-engineered features are used extensively in machine learning-based radiomics. This is justified because those features are based on prior research that may mechanistically point toward specific pathophysiology; they also limit the feature space. The majority of studies use small data-sets of approximately 100 image volumes. In these situations, overfitting is potentially a problem. Unlike segmentation problem we discussed above, in machine learning radiomics, each image volume is associated with just one classification, which acts to limit the number of available training inputs.

16.4.4 Ground-Truth Data-Sets

In medical imaging, curating high-quality, ground truth data-sets that are representative of the population is challenging. We have already discussed the issue of overfitting in a previous section, and will not repeat it here.

16.4.4.1 Benchmark Data-Sets for Glioma

There are several benchmark open-source glioma data-sets available. The Medical Image Computing and Computer Assisted Intervention Society 2017 Brain Tumor

Segmentation challenge training data-set contains the annotated brain MRI scans of 210 patients who have HGGs and 75 patients who have LGGs. There are a further 46 separate scans from other patients which act as the validation set. Multiple neuroradiologists have manually segmented all of the cases.

A 2nd public data-set is a multi-institutional glioma collection of The Cancer Genome Atlas, which is available through The Cancer Imaging Archive [13]. It contains preoperative MRIs of 135 patients with glioblastoma and 108 patients who have LGGs [30]. Additionally, a panel of radiomic features of morphologic, intensity, histogram-based, volumetric, and textural parameters, spatial information, and parameters extracted from glioma growth models, are also included [30].

The limited brain tumour datasets with known molecular markers continues to impede research. A typical DL architecture consists of very many parameters, which demands a larger data-set than traditional ML methods to prevent overfitting. For DL methods, there are several approaches to deal with smaller data-sets, including drop-out [22], data augmentation [57], patch-based histogram analyses [58], and generative adversarial networks [2]. The reader is referred generally to the discussion in the section of this chapter on segmentation.

16.4.5 Deep Learning as a Tool for Molecular Biomarker Prediction

The emergence of novel machine learning methods (DL) to predict molecular properties of tissues has the potential to be a game changer for medical imaging in that the technology has the potential to achieve a performance exceeding that of humans in identifying the content of images.

DL can learn intricate and highly abstract patterns from medical imaging data that may not be apparent to clinicians. As discussed in previous section large data-sets can be difficult to curate.

CNNs are particularly useful in medical image analysis because the central idea—convolutions—preserves the spatial relationships of data. Compared with traditional machine learning methods, there is no need to compute features as a first step. The CNN finds the critical features as a natural part of its own search process. Figure 16.2 depicts a simple CNN architecture.

16.4.5.1 Deep Learning-Based Radiomics for Gliomas

In glioma, DL has shown promise in the task of predicting important molecular markers such as 1p19q codeletion and MGMT promoter methylation from MRI images alone [47]. Recent studies have attempted to assess deeper- and higher-order features using DL that may improve on the predictive performance of radiomics models for brain tumors.

Li et al. [53] used DL-based radiomics to extract characteristics from multi-modal MRI to predict IDH-1 mutation status in 151 patients with low-grade glioma. Their results show an AUC for IDH-1 estimation of 0.95, which compares with 0.86 for features extracted using traditional machine learning methods [74].

DL-based radiomics models for brain tumors have been proposed for molecular characteristics classification [8] and for prediction of patient survival [50]. Lao et al. [50] developed a CNN-based radiomics model using MRI to estimate overall survival in patients with glioblastoma. After selection of six deep features from pre-trained CNN, the established model predicted overall survival better than conventional models that relied on hand-engineered features, with a C-index of 0.71 (representing a good fit of a logistic regression model), that increased to 0.739 when it was combined with clinical risk factors [50].

Though the reported accuracies, C-indexes, and AUCs of some of these studies were favorable, many used small numbers of training datasets, with an imbalance of classes, and used validation datasets without separate datasets for testing [8, 50, 53], potentially limiting generalizability.

Machine learning-based radiomics have been reported as insufficient for reliable clinical usage [44], but advances in CNN architectures for tumor segmentation are now paving the way toward progress in radiomics [8, 50]. Previously, segmentation was often a significant bottleneck affecting radiomics [27]. DL algorithms can be stacked to achieve end-to-end training from image segmentation to classification to outcome prediction.

In addition, specialized regularization methods, which include adversarial training, extensive data augmentation, combining classification with segmentation during training, and semi-supervised learning methods via combining autoencoder and classifier architectures together, may further improve the quality of the learned hierarchic classifier features [2].

16.4.6 Explainable Artificial Intelligence

Early DL efforts were often criticised users could not understand the basis the algorithm used for a particular prediction. The field of "explainable AI" has arisen in direct response to those criticism, and attempts to address them. If we can understand the basis for a CNN's prediction, here are potentially several benefits for us: (1) it may increase clinician, patient, and public confidence in, and acceptance of, the decision or prediction; (2) it may allow us to better understand of the underlying physical basis for the decision or prediction; and (3) it may allow us to satisfy other statutory requirements to establish the basis for decision [e.g. General Data Protection Regulations (GDPR)]. For example, in the molecular prediction of biomarkers, it is vitally important to know if the algorithm's decision is based on areas of the image with enhancement or hyper-intensities, or if it makes a decision based on the background of the MRI.

Recent work in visualization of deep learning models is helping to bridge the gap in visualizing these biomarkers. While research in this area is at an early stage and has not been objectively and fully evaluated there is sufficient evidence to suggest that emerging explainable AI methods could pave the way to discovering entirely new ways of evaluating images.

A number of methods have been proposed to aid the interpretation of an algorithm's predictions. Local interpretable model-agnostic explanations (LIME), proposed by Ribeiroet al. [76], interprets model predictions based on locally approximating the model around a given prediction. SHAP (SHapley Additive exPlanations), proposed by Lundberg et al. [56], assigns each feature an importance value for a particular prediction. Their method introduces three tools, GradientExplainer, DeepExplainer (DEEP SHAP), and KernelExplainer (Kernel SHAP) for DL interpretation.

Visual explanations such as GradCAM [11] and saliency maps [79] have also been proposed as ways to understand and justify DL network decisions. The prediction of molecular markers is highly valued in clinical practice. If we can capturing the decision with high accuracy and also visualise the areas of the image that contributed to this decision, we can build trust and understanding with clinicians.

It may be that more comprehensive datasets will allow us to more accurately predict the treatment or treatments that will be most effective for a given patient, and to predicat the prognosis for patients from imaging alone. This could make MR imaging more valuable in the care of patients with brain tumours.

16.5 Conclusion

We have seen that every pre-processing step for MRI, from skull stripping, to denoising, to bias correction and normalisation, indeed the whole reconstruction pipeline, has seen deep learning approaches applied to good effect. The suppression of Gibbs and other artefacts, and image superresolution generally, has also been investigated using deep learning approaches. These problems may be unified into a single neural network in future work. Furthermore, we have seen that deep learning approaches to in brain MRI registration is a new field, with no consensus yet on the best approaches. Segmentation is better studied, but suffers from lack of sufficient training data, and imbalanced data. In spite of that, we have seen that deep learning approaches can segment different brain tumor tissues from edema and from healthy brain tissue and these approaches show promise, with accuracy beginning to approach that of human experts. Lack of interpretability remains a challenge for clinical acceptance. Content-based image retrieval has huge potential for diagnostics, teaching, and research. We have outlined the extensive challenges involved in this topic. Radiomics suffer from small datasets even more so than segmentation, as each image volume corresponds with a single output or classification; this acts to limit the number of available training inputs. We have outlined the publicly available datasets for this field, and the approaches to mitigating the effects of small training datasets. Radiomics is not ready for clinical use yet, but recent advances show the promise of deep learning

approaches. Finally, we discussed the field of explainable AI, which endeavours to justify how deep learning and other AI tools arrive at their conclusions, which is critical for confidence in, and acceptance of, the decision or prediction, establishing new research direction via new knowledge of the underlying physiognomy of the disease, and satisfaction of all necessary legal and/or regulatory requirements providing the basis for a decision to the patient or for later scrutiny.

Deep learning approaches have touched every aspect of MRI reconstruction and interpretation for gliomas, and have generally proven competitive with the state of the art, though research on this topic is too immature and on too limited data sets to establish deep learning as the leading approach in these fields yet. That is likely to change quickly.

Acknowledgements All three authors would like to acknowledge the support of the EU COST Action 18206.

References

1. J. Amin, M. Sharif, N. Gul, M. Raza, M.A. Anjum, M.W. Nisar, S.A.C. Bukhari, Brain tumor detection by using stacked autoencoders in deep learning. J. Med. Syst. **44**(2), 32 (2020). https://doi.org/10.1007/s10916-019-1483-2
2. S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J.S. Kirby, J.B. Freymann, K. Farahani, C. Davatzikos, Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. Sci. Data **4**, 170117 (2017)
3. G. Balakrishnan, A. Zhao, M.R. Sabuncu, A.V. Dalca, J. Guttag, An unsupervised learning model for deformable medical image registration, in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2018), pp. 9252–9260. https://doi.org/10.1109/CVPR.2018.00964
4. P. Barone, G. Sebastiani, A new method of magnetic resonance image reconstruction with short acquisition time and truncation artifact reduction. IEEE Trans. Med. Imag. **11**(2), 250–259 (1992)
5. S. Bisdas, H. Shen, S. Thust, V. Katsaros, G. Stranjalis, C. Boskos, S. Brandner, J. Zhang, Texture analysis-and support vector machine-assisted diffusional kurtosis imaging may allow in vivo gliomas grading and IDH-mutation status prediction: a preliminary study. Sci. Rep. **8**(1), 1–9 (2018)
6. X. Cao, J. Yang, J. Zhang, D. Nie, M. Kim, Q. Wang, D. Shen, Deformable image registration based on similarity-steered CNN regression, in *Medical image computing and computer-assisted intervention : MICCAI—International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol. **10433** (2017), pp. 300–308
7. K. Chang, A.L. Beers, H.X. Bai, J.M. Brown, K.I. Ly, X. Li, J.T. Senders, V.K. Kavouridis, A. Boaro, C. Su, W.L. Bi, O. Rapalino, W. Liao, Q. Shen, H. Zhou, B. Xiao, Y. Wang, P.J. Zhang, M.C. Pinho, P.Y. Wen, T.T. Batchelor, J.L. Boxerman, O. Arnaout, B.R. Rosen, E.R. Gerstner, L. Yang, R.Y. Huang, J. Kalpathy-Cramer, Automatic assessment of glioma burden: a deep learning algorithm for fully automated volumetric and bidimensional measurement. Neuro-Oncology **21**(11), 1412–1422 (2019). https://doi.org/10.1093/neuonc/noz106
8. P. Chang, J. Grinband, B. Weinberg, M. Bardis, M. Khy, G. Cadena, M.Y. Su, S. Cha, C. Filippi, D. Bota et al., Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas. Am. J. Neuroradiol. **39**(7), 1201–1207 (2018)
9. P.D. Chang, Fully convolutional deep residual neural networks for brain tumor segmentation, in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, ed. by

A. Crimi, B. Menze, O. Maier, M. Reyes, S. Winzeck, H. Handels (Springer International Publishing, Berlin, 2016)
10. L. Chato, E. Chow, S. Latifi, Wavelet transform to improve accuracy of a prediction model for overall survival time of brain tumor patients based on MRI images, in *2018 IEEE International Conference on Healthcare Informatics (ICHI)* (IEEE, 2018), pp. 441–442
11. A. Chattopadhay, A. Sarkar, P. Howlader, V.N. Balasubramanian, Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks, in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)* (IEEE, 2018), pp. 839–847
12. H. Chen, Q. Dou, L. Yu, J. Qin, P.A. Heng, Voxresnet: deep voxelwise residual networks for brain segmentation from 3d MR images. NeuroImage **170**, 446–455 (2017). https://doi.org/10.1016/j.neuroimage.2017.04.041
13. K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle et al., The cancer imaging archive (TCIA): maintaining and operating a public information repository. J. Digit. Imag. **26**(6), 1045–1057 (2013)
14. R.R. Colen, M. Vangel, J. Wang, D.A. Gutman, S.N. Hwang, M. Wintermark, R. Jain, M. Jilwan-Nicolas, J.Y. Chen, P. Raghavan et al., Imaging genomic mapping of an invasive MRI phenotype predicts patient outcome and metabolic dysfunction: a TCGA glioma phenotype research group project. BMC Med. Genomics **7**(1), 30 (2014)
15. P. Coupé, P. Yger, S. Prima, P. Hellier, C. Kervrann, C. Barillot, An optimized blockwise nonlocal means denoising filter for 3-d magnetic resonance images. IEEE Trans. Med. Imag. **27**(4), 425–441 (2008)
16. S. Cui, L. Mao, J. Jiang, C. Liu, S. Xiong, Automatic semantic segmentation of brain gliomas from MRI images using a deep cascaded neural network. J. Healthc. Eng. **2018**, 1–14 (2018). https://doi.org/10.1155/2018/4940593
17. H. Dong, G. Yang, F. Liu, Y. Mo, Y. Guo, Automatic brain tumor detection and segmentation using u-net based fully convolutional networks, in *Medical Image Understanding and Analysis*, ed. by M. Valdés-Hernández, V. González-Castro (Springer International Publishing, Cham, 2017), pp. 506–517
18. S. Dube, S. El-Saden, T.F. Cloughesy, U. Sinha, Content based image retrieval for MR image studies of brain tumors, in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society* (2006), pp. 3337–3340. https://doi.org/10.1109/IEMBS.2006.260262
19. C. Dupont, N. Betrouni, N. Reyns, M. Vermandel, On image segmentation methods applied to glioblastoma: state of art and new trends. IRBM **37**(3), 131–143 (2016). https://doi.org/10.1016/j.irbm.2015.12.004. http://www.sciencedirect.com/science/article/pii/S195903181500144X
20. P. Dvorak, B. Menze, Structured prediction with convolutional neural networks for multimodal brain tumor segmentation, in *Proceeding of the Multimodal Brain Tumor Image Segmentation Challenge* (2015), pp. 13–24
21. B. Ellingson, A. Lai, R. Harris, J. Selfridge, W. Yong, K. Das, W. Pope, P. Nghiemphu, H. Vinters, L. Liau et al., Probabilistic radiographic atlas of glioblastoma phenotypes. Am. J. Neuroradiol. **34**(3), 533–540 (2013)
22. B.M. Ellingson, M. Bendszus, J. Boxerman, D. Barboriak, B.J. Erickson, M. Smits, S.J. Nelson, E. Gerstner, B. Alexander, G. Goldmacher et al., Consensus recommendations for a standardized brain tumor imaging protocol in clinical trials. Neuro-oncology **17**(9), 1188–1198 (2015)
23. C. Fennema-Notestine, I.B. Ozyurt, C.P. Clark, S. Morris, A. Bischoff-Grethe, M.W. Bondi, T.L. Jernigan, B. Fischl, F. Segonne, D.W. Shattuck et al., Quantitative evaluation of automated skull-stripping methods applied to contemporary and legacy images: effects of diagnosis, bias correction, and slice location. Hum. Brain Mapp. **27**(2), 99–113 (2006)
24. A. Gelb, A hybrid approach to spectral reconstruction of piecewise smooth functions. J. Sci. Comput. **15**(3), 293–322 (2000)
25. O. Gevaert, L. Mitchell, A. Achrol et al., Glioblastoma multiforme: exploratory radiogenomic analysis by using quantitative image features. Radiology **273**(1), 168–174 (2014)

26. E. Giacomello, D. Loiacono, L. Mainardi, Brain MRI tumor segmentation with adversarial networks (2019)
27. R.J. Gillies, P.E. Kinahan, H. Hricak, Radiomics: images are more than pictures, they are data. Radiology **278**(2), 563–577 (2016)
28. S. Grivalsky, M. Tamajka, W. Benesova, Segmentation of gliomas in magnetic resonance images using recurrent neural networks, in *2019 42nd International Conference on Telecommunications and Signal Processing (TSP)* (2019), pp. 539–542. https://doi.org/10.1109/TSP.2019.8769056
29. H. Gudbjartsson, S. Patz, The Rician distribution of noisy MRI data. Magn. Reson. Med. **34**(6), 910–914 (1995)
30. D.A. Gutman, L.A. Cooper, S.N. Hwang, C.A. Holder, J. Gao, T.D. Aurora, W.D. Dunn Jr., L. Scarpace, T. Mikkelsen, R. Jain et al., Mr imaging predictors of molecular profile and survival: multi-institutional study of the tcga glioblastoma data set. Radiology **267**(2), 560–569 (2013)
31. M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.M. Jodoin, H. Larochelle, Brain tumor segmentation with deep neural networks. Med. Image Anal. **35**, 18–31 (2017). https://doi.org/10.1016/j.media.2016.05.004
32. M. Havaei, N. Guizard, H. Larochelle, P.M. Jodoin, *Deep Learning Trends for Focal Brain Pathology Segmentation in MRI* (Springer International Publishing, Cham, 2016), pp. 125–148. https://doi.org/10.1007/978-3-319-50478-0_6
33. M. Huang, W. Yang, Y. Wu, J. Jiang, Y. Gao, Y. Chen, Q. Feng, W. Chen, Z. Lu, Content-based image retrieval using spatial layout information in brain tumor t1-weighted contrast-enhanced MR images. PloS one **9**(7), e102754 (2014)
34. J.E. Iglesias, C.Y. Liu, P.M. Thompson, Z. Tu, Robust brain extraction across datasets and comparison with publicly available methods. IEEE Trans. Med. Imag. **30**(9), 1617–1634 (2011)
35. A. Işln, C. Direkoğlu, M.Şah, Review of MRI-based brain tumor image segmentation using deep learning methods. Procedia Comput. Sci. **102**, 317–324 (2016), in 12th International Conference on Application of Fuzzy Systems and Soft Computing, ICAFS, 29-30 Aug 2016 (Austria, Vienna, 2016). https://doi.org/10.1016/j.procs.2016.09.407. URL http://www.sciencedirect.com/science/article/pii/S187705091632587X
36. F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, K.H. Maier-Hein, Brain tumor segmentation and radiomics survival prediction: contribution to the brats 2017 challenge, in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, ed. by A. Crimi, S. Bakas, H. Kuijf, B. Menze, M. Reyes (Springer International Publishing, Cham, 2018), pp. 287–297
37. Jain, V., Seung, S.: Natural image denoising with convolutional networks, in *Advances in neural information processing systems* (2009), pp. 769–776
38. J.E. Eckel-Passow, D.H. Lachance, A.M. Molinaro, K.M. Walsh, P.A. Decker, H. Sicotte, M. Pekmezci, T. Rice, M.L. Kosel, I.V. Smirnov, G. Sarkar, Glioma groups based on 1p/19q, IDH, and TERT promoter mutations in tumors. New Engl. J. Med. **372**(26), 2499–508 (2015)
39. L. Jin, L. Min, J. Wang, F. Wu, T. Liu, Y. Pan, A survey of MRI-based brain tumor segmentation methods. Tsinghua Sci. Technol. **19**(6), 578–595 (2014). https://doi.org/10.1109/TST.2014.6961028
40. K. Kamnitsas, E. Ferrante, S. Parisot, C. Ledig, A.V. Nori, A. Criminisi, D. Rueckert, B. Glockern, Deepmedic for brain tumor segmentation, in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, ed. by A. Crimi, B. Menze, O. Maier, M. Reyes, S. Winzeck, H. Handels (Springer International Publishing, Cham, 2016), pp. 138–149
41. K. Kamnitsas, C. Ledig, V.F. Newcombe, J.P. Simpson, A.D. Kane, D.K. Menon, D. Rueckert, B. Glocker, Efficient multi-scale 3d CNN with fully connected CRF for accurate brain lesion segmentation. Med. Image Anal. **36**, 61–78 (2017)
42. P.Y. Kao, F. Shailja, J. Jiang, A. Zhang, A. Khan, J. Chen, B. Manjunath, Improving patch-based convolutional neural networks for mri brain tumor segmentation by leveraging location information. Front. Neurosci. **13**, 1449 (2020). https://doi.org/10.3389/fnins.2019.01449
43. B. Kayalibay, G. Jensen, P. van der Smagt, CNN-based Segmentation of Medical Imaging Data (2017). arXiv:1701.03056

44. P. Kickingereder, D. Bonekamp, M. Nowosielski et al., Radiogenomics of glioblastoma: machine learning-based classification of molecular characteristics by using multiparametric and multiregional mr imaging features. Radiology **281**(3), 907–919 (2016)
45. J. Kleesiek, G. Urban, A. Hubert, D. Schwarz, K. Maier-Hein, M. Bendszus, A. Biller, Deep MRI brain extraction: a 3d convolutional neural network for skull stripping. NeuroImage **129**, 460–469 (2016). https://doi.org/10.1016/j.neuroimage.2016.01.024. http://www.sciencedirect.com/science/article/pii/S1053811916000306
46. A. Klein, J. Andersson, B.A. Ardekani, J. Ashburner, B. Avants, M.C. Chiang, G.E. Christensen, D.L. Collins, J. Gee, P. Hellier et al., Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. Neuroimage **46**(3), 786–802 (2009)
47. P. Korfiatis, B. Erickson, Deep learning can see the unseeable: predicting molecular markers from mri of brain gliomas. Clin. Radiol. **74**(1), 367–373 (2019)
48. P. Korfiatis, T. Kline, B. Erickson, Automated segmentation of hyperintense regions in flair mri using deep learning. Tomogr. J. Imag. Res. **2**, 334–340 (2016). https://doi.org/10.18383/j.tom.2016.00166
49. A. Kumar, J. Kim, T.W. Cai, M.J. Fulham, D.D. Feng, Content-based medical image retrieval: a survey of applications to multidimensional and multimodality data. J. Digit. Imag. **26**, 1025–1039 (2013)
50. J. Lao, Y. Chen, Z.C. Li, Q. Li, J. Zhang, J. Liu, G. Zhai, A deep learning-based radiomics model for prediction of survival in glioblastoma multiforme. Sci. Rep. **7**(1), 1–8 (2017)
51. G. Latif, D. Iskandar, J. Alghazo, M. Butt, A.H. Khan, Deep CNN based MR image denoising for tumor segmentation using watershed transform. Int. J. Eng. Technol. **7**(2.3), 37–42 (2018)
52. H. Li, Y. Fan, Non-rigid image registration using self-supervised fully convolutional networks without training data, in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* (2018), pp. 1075–1078. https://doi.org/10.1109/ISBI.2018.8363757
53. Z. Li, Y. Wang, J. Yu, Y. Guo, W. Cao, Deep learning based radiomics (dlr) and its usage in noninvasive idh1 prediction for low grade glioma. Sci. Rep. **7**(1), 1–11 (2017)
54. J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2015), pp. 3431–3440. https://doi.org/10.1109/CVPR.2015.7298965
55. D. Louis, A. Perry, G. Reifenberger et al., The 2016 world health organization classification of tumours of the central nervous system: a summary. Acta Neuropathol. **131**, 803–820 (2016)
56. S.M. Lundberg, S.I. Lee, A unified approach to interpreting model predictions, in *Advances in Neural Information Processing Systems* (2017), pp. 4765–4774
57. A. Mahajan, A. Moiyadi, R. Jalali, E. Sridhar, Radiogenomics of glioblastoma: a window into its imaging and molecular variability. Cancer Imag. **15**(1), P14 (2015)
58. M.A. Mazurowski, Radiogenomics: what it is and why it is important. J. Am. Coll. Radiol. **12**(8), 862–866 (2015)
59. R. McKinley, R. Meier, R. Wiest, Ensembles of densely-connected cnns with label-uncertainty for brain tumor segmentation, in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, ed. by A. Crimi, S. Bakas, H. Kuijf, F. Keyvan, M. Reyes, T. van Walsum (Springer International Publishing, Cham, 2019), pp. 456–465
60. F. Milletari, S.A. Ahmadi, C. Kroll, A. Plate, V. Rozanski, J. Maiostre, J. Levin, O. Dietrich, B. Ertl-Wagner, K. Bötzel et al., Hough-CNN: deep learning for segmentation of deep brain regions in MRI and ultrasound. Comput. Vis. Image Underst. **164**, 92–102 (2017)
61. J. Moustakas, K. Marias, S. Dimitriadis, S.C. Orphanoudakis, A two-level CBIR platform with application to brain MRI retrieval, in *2005 IEEE International Conference on Multimedia and Expo* (2005), pp. 1278–1281
62. M.J. Muckley, B. des-Aron, A. Papaioannou, G. Lemberskiy, E. Solomon, Y.W. Lui, D.K. Sodickson, E. Fieremans, D.S. Novikov, F. Knoll, *Training a Neural Network for Gibbs and Noise Removal in Diffusion MRI* (2019)
63. A. Myronenko, 3D MRI brain tumor segmentation using autoencoder regularization, in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, ed. by A. Crimi, S. Bakas, H. Kuijf, F. Keyvan, M. Reyes, T. van Walsum (Springer International Publishing, Cham, 2019), pp. 311–320

64. K.M. Naeini, W.B. Pope, T.F. Cloughesy, R.J. Harris, A. Lai, A. Eskin, R. Chowdhury, H.S. Phillips, P.L. Nghiemphu, Y. Behbahanian et al., Identifying the mesenchymal molecular subtype of glioblastoma using quantitative volumetric analysis of anatomic magnetic resonance images. Neuro-oncology **15**(5), 626–634 (2013)
65. J. Nalepa, M. Marcinkiewicz, M. Kawulok, Data augmentation for brain-tumor segmentation: a review. Front. Comput. Neurosci. **13**, 83 (2019)
66. C.G.A.R. Network, Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas. New Engl. J. Med. **372**(26), 2481–2498 (2015)
67. S. Pereira, A. Pinto, V. Alves, C.A. Silva, Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Trans. Med. Imag. **35**(5), 1240–1251 (2016)
68. M. Pisov, G. Makarchuk, V. Kostjuchenko, A. Dalechina, A. Golanov, M. Belyaev, *Brain Tumor Image Retrieval via Multitask Learning* (2018). arXiv:1810.09369
69. R. Pizarro, H.E. Assemlal, D.D. Nigris, C. Elliott, S. Antel, D.L. Arnold, A. Shmuel, Using deep learning algorithms to automatically identify the brain MRI contrast: Implications for managing large databases. Neuroinformatics **17**(1), 115–30 (2018). https://doi.org/10.1007/s12021-018-9387-8
70. W. Pope, L. Mirsadraei, A. Lai, A. Eskin, J. Qiao, H. Kim, B. Ellingson, P. Nghiemphu, S. Kharbanda, R. Soriano et al., Differential gene expression in glioblastoma defined by adc histogram analysis: relationship to extracellular matrix molecules and survival. Am. J. Neuroradiol. **33**(6), 1059–1064 (2012)
71. W.B. Pope, Genomics of brain tumor imaging. Neuroimaging Clin. **25**(1), 105–119 (2015)
72. W.B. Pope, J.H. Chen, J. Dong, M.R. Carlson, A. Perlina, T.F. Cloughesy, L.M. Liau, P.S. Mischel, P. Nghiemphu, A. Lai et al., Relationship between gene expression and enhancement in glioblastoma multiforme: exploratory dna microarray analysis. Radiology **249**(1), 268–277 (2008)
73. C. Qin, J. Schlemper, J. Caballero, A.N. Price, J.V. Hajnal, D. Rueckert, Convolutional recurrent neural networks for dynamic MR image reconstruction. IEEE Trans. Med. Imag. **38**(1), 280–290 (2018)
74. S. Rathore, H. Akbari, J. Doshi, G. Shukla, M. Rozycki, M. Bilello, R.A. Lustig, C.A. Davatzikos, Radiomic signature of infiltration in peritumoral edema predicts subsequent recurrence in glioblastoma: implications for personalized radiotherapy planning. J. Med. Imag. **5**(2), 021219 (2018)
75. S. Rathore, M.A. Iftikhar, Z. Mourelatos, *Prediction of Overall Survival and Molecular Markers in Gliomas via Analysis of Digital Pathology Images using Deep Learning* (2019). arXiv preprint arXiv:1909.09124
76. M.T. Ribeiro, S. Singh, C. Guestrin, "Why should i trust you?" explaining the predictions of any classifier, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2016), pp. 1135–1144
77. B. Roberts, M. Wan, S.P. Kelly, J.J. Healy, Quantitative comparison of Gegenbauer, filtered Fourier, and Fourier reconstruction for MRI, in *Multimodal Biomedical Imaging XV*, vol. 11232 (International Society for Optics and Photonics, 2020), p. 112320L
78. M. Simonovsky, B. Gutiérrez-Becker, D. Mateus, N. Navab, N. Komodakis, A deep metric for multimodal registration, in *MICCAI* (2016)
79. K. Simonyan, A. Vedaldi, A. Zisserman, *Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps* (2013). arXiv preprint arXiv:1312.6034
80. M. Smith, S. Nichols, R. Henkelman, M. Wood, Application of autoregressive modelling in magnetic resonance imaging to remove noise and truncation artifacts. Magn. Reson. Imag. **4**(3), 257–261 (1986)
81. B. Stefan, W. Roland, N. Lutz-P, R. Mauricio, A survey of MRI-based medical image analysis for brain tumor studies. Phys. Med. Biol. **58**, R97–R129 (2013). https://doi.org/10.1088/0031-9155/58/13/R97
82. M.F. Stollenga, W. Byeon, M. Liwicki, J. Schmidhuber, Parallel multi-dimensional lstm, with application to fast biomedical volumetric image segmentation, in*Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2, NIPS'15* (MIT Press, Cambridge, MA, USA, 2015), pp. 2998–3006

83. Z.N.K. Swati, Q. Zhao, M. Kabir, F. Ali, Z. Ali, S. Ahmed, J. Lu, Content-based brain tumor retrieval for MR images using transfer learning. IEEE Access **7**, 17809–17822 (2019)
84. D. Tamada, *Review: Noise and Artifact Reduction for MRI Using Deep Learning* (2020)
85. Q. Tian, L.F. Yan, X. Zhang, X. Zhang, Y.C. Hu, Y. Han, Z.C. Liu, H.Y. Nan, Q. Sun, Y.Z. Sun et al., Radiomics strategy for glioma grading using texture features from multiparametric MRI. J. Magn. Reson. Imag. **48**(6), 1518–1528 (2018)
86. G. Urban, M. Bendszus, F.A. Hamprecht, J. Kleesiek, Multi-modal brain tumor segmentation using deep convolutional neuralnetworks, in *MICCAI BraTS (Brain Tumor Segmentation) Challenge. Proceedings, winningcontribution* (2014), pp. 31–35
87. A. Varghese, K. Vaidhya, S. Thirunavukkarasu, C. Kesavdas, G. Krishnamurthi, Semisupervised learning using denoising autoencoders for brain lesion detection and segmentation. J. Med. Imag. **4**, 041311 (2017)
88. A.V. Vasilakos, Y. Tang, Y. Yao et al., Neural networks for computer-aided diagnosis in medicine: a review. Neurocomputing **216**, 700–708 (2016)
89. U. Vovk, F. Pernus, B. Likar, A review of methods for correction of intensity inhomogeneity in MRI. IEEE Trans. Med. Imag. **26**(3), 405–421 (2007)
90. Y. Wang, Y. Song, H. Xie, W. Li, B. Hu, G. Yang, Reduction of Gibbs artifacts in magnetic resonance imaging based on convolutional neural network, in *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)* (IEEE, 2017), pp. 1–5
91. M. Weller, W. Wick, K. Aldape, M. Brada, M. Berger, S.M. Pfister, R. Nishikawa, M. Rosenthal, P.Y. Wen, R. Stupp et al., Glioma. Nat. Rev. Dis. Primers **1**(1), 1–18 (2015)
92. P.Y. Wen, J.T. Huse, 2016 world health organization classification of central nervous system tumors. Continuum Lifelong Learn. Neurol. **23**(6), 1531–1547 (2017)
93. G. Wu, M. Kim, Q. Wang, Y. Gao, S. Liao, D. Shen, *Unsupervised Deep Feature Learning for Deformable Registration of MR Brain Images* (2013), pp. 649–656. https://doi.org/10.1007/978-3-642-40763-5_80
94. G. Wu, M. Kim, Q. Wang, B.C. Munsell, D. Shen, Scalable high-performance image registration framework by unsupervised deep feature representations learning. IEEE Trans. Biomed. Eng. **63**(7), 1505–1516 (2016). https://doi.org/10.1109/TBME.2015.2496253
95. Z. Xiao, R. Huang, Y. Ding, T., Lan, R. Dong, Z. Qin, X. Zhang, W. Wang, A deep learning-based segmentation method for brain tumor in MR images, in *2016 IEEE 6th International Conference on Computational Advances in Bio and Medical Sciences (ICCABS)* (2016), pp. 1–6. https://doi.org/10.1109/ICCABS.2016.7802771
96. K. Xie, Y. Wen, Lstm-ma: a lstm method with multi-modality and adjacency constraint for brain image segmentation, in *2019 IEEE International Conference on Image Processing (ICIP)* (2019), pp. 240–244. https://doi.org/10.1109/ICIP.2019.8802959
97. Y. Xue, T. Xu, H. Zhang, R. Long, X. Huang, Segan: Adversarial network with multi-scale l_1 loss for medical image segmentation. Neuroinformatics **16**, 383–392 (2017). https://doi.org/10.1007/s12021-018-9377-x
98. H. Yan, J. Mao, Data truncation artifact reduction in mr imaging using a multilayer neural network. IEEE Trans. Med. Imag. **12**(1), 73–77 (1993)
99. X. Yang, R. Kwitt, M. Niethammer, *Fast Predictive Image Registration* (2016), pp. 48–57. https://doi.org/10.1007/978-3-319-46976-8_6
100. Q. Zhang, G. Ruan, W. Yang, Y. Liu, K. Zhao, Q. Feng, W. Chen, E.X. Wu, Y. Feng, MRI Gibbs-ringing artifact reduction by means of machine learning using convolutional neural networks. Magn. Reson. Med. **82**(6), 2133–2145 (2019)
101. L. Zhao, K. Jia, Deep feature learning with discrimination mechanism for brain tumor segmentation and diagnosis, in *2015 International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP)* (2015), pp. 306–309. https://doi.org/10.1109/IIH-MSP.2015.41
102. X. Zhao, Y. Wu, G. Song, Z. Li, Y. Zhang, Y. Fan, A deep learning model integrating fcnns and crfs for brain tumor segmentation. Med. Image Anal. **43**, 98–111 (2018)

103. X. Zhao, H. Zhang, Y. Zhou, W. Bian, T. Zhang, X. Zou, Gibbs-ringing artifact suppression with knowledge transfer from natural images to MR images, in *Multimedia Tools and Applications* (2019), pp. 1–23
104. D. Zikic, Y. Ioannou, M. Brown, A. Criminisi, *Segmentation of Brain Tumor Tissues with Convolutional Neural Networks* (2014)
105. P.O. Zinn, B. Majadan, P. Sathyan, S.K. Singh, S. Majumder, F.A. Jolesz, R.R. Colen, Radiogenomic mapping of edema/cellular invasion MRI-phenotypes in glioblastoma multiforme. PloS One **6**(10), e25451 (2011)