Computational Intelligence Based Model for Detection of Disease using Chest Radiographs

Hrishikesh Mane
Student (B.E.),
Department of Computer Engineering,
Modern Education Society's College of Engineering,
Pune, India
hrishi.mane26@gmail.com

Parag Ghorpade
Student (B.E.),
Department of Computer Engineering,
Modern Education Society's College of Engineering,
Pune, India
paragghorpade1998@gmail.com

Vedant Bahel
Student (B.E.),
Department of Information Technology,
GH Raisoni College of Engineering,
Nagpur, India
vbahel@ieee.org

***Abstract*— There are many diseases associated with lungs or thoracic cavity and the diagnosis of these diseases at once becomes difficult for any medical profusion. Although Chest Radiography is the most common way of screening done when the thoracic cavity comes into the picture. However, diagnosing multiple diseases from a single scan becomes difficult. This paper proposes an intelligent machine learning-based model which tries to detect 14 chest diseases out from a single radiograph with greater accuracy. This paper makes use of advanced deep learning techniques like neural networks, masking algorithms, etc. to assure higher performances.**

Keywords—Radiographs, Deep Neural Network, Imagining, Healthcare

I. INTRODUCTION

In recent years, the use of Machine Learning (ML) and Intelligent system has increased exponentially. That has, in turn, increased the demand for ML application in the field of healthcare. Machine Learning enables the extraction of the hidden patterns from the available data. In India, healthcare is very vast and so are the number of health-related problems. Medical Problems associated with thoracic cavities are pretty common these days. As per the report published by UNICEF in 2019, 14 children died every hour due to pneumonia in India in 2018. Lung cancer accounts to have 9.8% of total cancer-related deaths in India [8].

This paper demonstrates a virtual assistant model for accurately detecting 14 chest diseases/ elements namely: Atelectasis, Cardiomegaly, Infiltration, Effusion, Mass, Pneumonia, Pneumothorax, Nodule, Consolidation, Edema, Fibrosis, Emphysema, Pleural Thickening, and Hernia. The same concept can further extend for MRI, CT Scans, and various further images. Fig. 1 demonstrates a sample dataset for 8 out of the 14 diseases. In this paper, authors have also put forward a general comparison of performances of various algorithms over the same dataset and have tried to reach the peak performance metric for the model. The authors have tried to describe the structure of data before explaining the model.

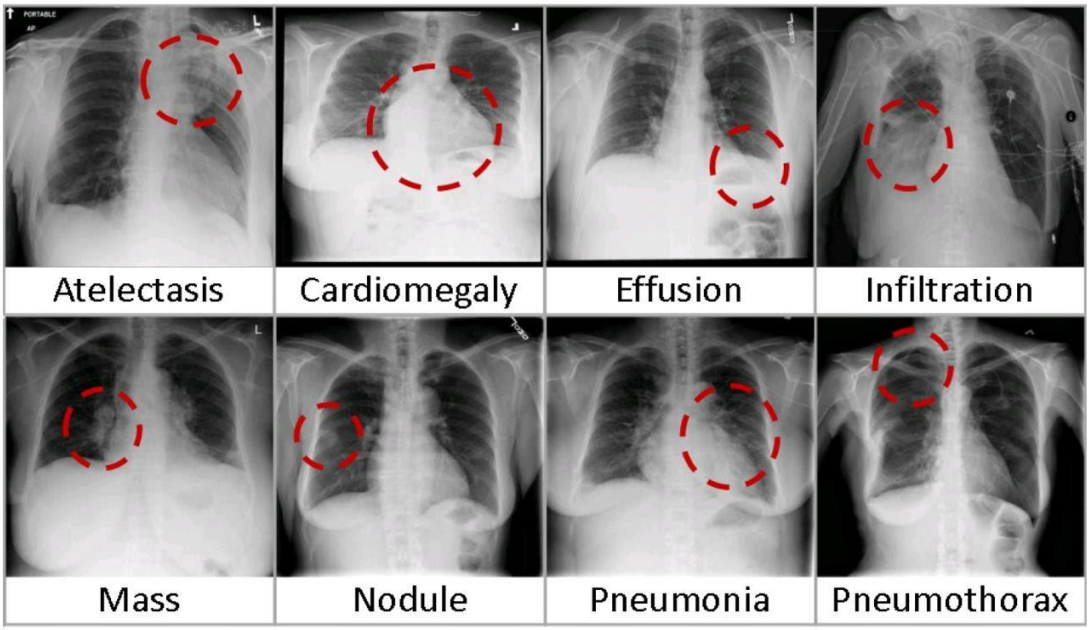


Fig. 1. Sample Image dataset for common thoracic diseases

II. DATASET

The Dataset used in this experiment was provided by the NIH Clinical Center. The Dataset consists of over 100,000 images of Chest X-ray. The dataset was compiled from the scans of more than 30,000 patients and is completely anonymous. Using this dataset, the author tries to modulate machine learning-based models using various algorithms to detect namely 14 diseases using chest radiographs. The radiograph available in the dataset is used as training data to formulate the ML-based models. The dataset helps in identifying the gradual changes over multiple x-rays which might go undetected. In [6], the performance of the proposed neural network model was narrowed due to a limited number of training data. The dataset used, NIH Chest X-ray 14 offers a better form with 6 more categories and new images of the dataset utilized in the recent work [1]. The dataset contains around 60% of all frontal chest x-rays in the hospital. Thus, it can expect for this dataset to precisely represent the real distribution of the patient population, which in turn will showcases realistic diagnosis of diseases. This size of the dataset will help better, train the neural networks and make accurate predictions. The Dataset comprises of 112,120 frontal-view X-ray images of 30,805 unique patients. Fourteen types of disease image labels are mined with the help of radiological reports using Natural Language Processing. Fourteen common thoracic pathologies include Atelectasis, Consolidation, Infiltration, Pneumothorax, Edema, Emphysema, Fibrosis, Effusion, Pneumonia, Pleural thickening, Cardiomegaly, Nodule, Mass and Hernia.[1] Table 1 demonstrates the structure of the number of images in the dataset for each class.

A. Dataset Configuration

More specifically the structure if the dataset can be summarised as follows:

- 112,120 frontal-view chest X-ray PNG labelled images having a resolution of (1024*1024) pixels
- Image Meta Data: Image Index, Finding Labels, Follow-up, Patient ID, Patient Age, Patient Gender, View Position, Original Image Size, and Original Image Pixel Spacing.
- Bounding boxes for around 1000 images: Image Index, Finding Label, Bounding Box [x, y, w, h]. [x y] are coordinates of each box's top-left corner. width and height of each box is given in [w h]
- For the purpose of training Neural Networks, the dataset is divided into two data split files (Training and Test Dataset). The images are divided into two sets by patient distribution. The details of the same patient might appear in either Testing or Training set.

TABLE I. DESCRIPTION OF THE DATASET (IMAGES/CLASS)

Class	Number of Images
Atelectasis	11,535
Cardiomegaly	2,772
Effusion	13,307
Infiltration	19,871
Mass	5,746
Nodule	6,323
Pneumonia	1,353
Pneumothorax	5,298
Consolidation	4,667
Edema	2,303
Emphysema	2,516
Fibrosis	1,686
Pleural Thickening	3,385
Hernia	227

III. PROPOSED MODEL

The thoracic detection classification task is a multiclass classification task, dataset having classification of 14 different types of diseases. The input to the model is an X-ray image of chest and the output is a vector which gives the confidence for each of the 14 classes.

***Model Architecture and Training*:** Authors implemented multiple models using transfer learning techniques for faster training to optimize the weights of various Deep Convolution Neural Network like VGG [2], ResNet [3] and DenseNet [4] on NIH Chest X-ray 14 dataset [1][5]. The final fully connected layer was replaced with an output layer having a sigmoid activation function. The initial weights of all the models are initialized to the pretrained ImageNet, trained entirely using Adam optimizer and standard parameters. The models are trained on the batch size of 32 and the initial learning rate of 0.001. Recently there have been increased availability of large and high-quality labelled datasets that are publicly available for experiments and research in the field of deep learning. Thanks to the availability of high-performance deep learning neural networks, performance and accuracy of the models have been increasing in medical imaging. Convolutional Neural Networks have ameliorated every year. Task of classification, detection and segmentation from convolutional neural networks has helped in successfully detecting tumours, cancer and pathology detection and many other such application.

***Disease Classification*:** Authors trained multi label classification models on 14 classes on different Deep Convolutional Neural Networks- VGG16, VGG19, ResNet50 and DenseNet121 with last convolutional layer as

block5_conv3, block5_conv4, activation_49 and bn respectively each having image input size as (224,224,3). Every model was trained on similar hyperparameters with slight changes in few, according to model architectures and GPU memory. The evaluation of the models was done on the basics of Area under curve (AUC) for each of the 14 classes. Authors could obtain the disease heatmap using the activations of the transition layer and the weights from the prediction layer of the model and then overwriting the heatmap to 0 its value is less than 0.2 and overlapped on the input image of the x-ray. Model could also plot the Blue bounding box (1,600 instances) on the same image to get localization heatmap and ground truth on the input image from the test dataset.

TABLE II. GENERAL COMPARISION OF YIELD OF VARIOUS MODEL ON CORRSPONDING CLASSES IN THE DATASET

Classes	VGG16	ResNet50	DenseNet121
Atelectasis	0.6161	0. 6831	0.8263
Cardiomegaly	0.7012	0. 8002	0.8732
Effusion	0.6250	0. 7215	0.8939
Infiltration	0.5966	0. 6173	0.7325
Mass	0.4899	0. 6343	0.8479
Nodule	0.6335	0. 6271	0.7633
Pneumonia	0.4980	0. 6261	0.7784
Pneumothorax	0.7426	0. 7891	0.9084
Consolidation	0.6012	0. 7136	0.7406
Edema	0.6751	0. 7823	0.8847
Emphysema	0.7477	0. 8172	0.9595
Fibrosis	0.5891	0. 7472	0.8003
Pleural Thickening	0. 5792	0. 6386	0.8092
Hernia	0. 7299	0. 8175	0.8295

As the Table II shows DenseNet121 yields more promising results by a great margin, followed by Resnet50 and then VGG16. This also proves that deeper the neural network the more it is precise. VGG is the smallest of the all then followed by ResNet, having 16 and 50 layers respectively. DenseNet has 121 layers and hence gives relatively better AUC than others. But as the neural network gets deeper the problem of vanishing gradient arises. Both ResNet and DenseNet have their own way of tackling this problem and thus the first layers cannot be updated through backpropagation of the error.

Data Augmentation on the dataset was used while training, which in turn used to increase the diversity of the data. Here only horizontal flip was used because of feature preservation. Resnet preserves the gradient by Projection shortcuts and identity matrix i.e. backpropagate through the identity function. The gradient would be multiplied by 1. ResNet proved that if the convolutional neural networks have shorter connection between layers close to the input and those close to the output the models can be build deeper, efficient and more accurate. DenseNet leverages this idea to connect each layer to every other layer. Each layer of Densenet the feature maps of all preceding layers are feed as inputs and its own feature maps are used as inputs to all subsequent the layers. Thus learning from highly complex features like that in medical imaging is much better in DenseNet than others because of feature reuse. Therefore DenseNet was used for this multiclass classification problem. The Model after training gives the mean AUROC of 0.8296.

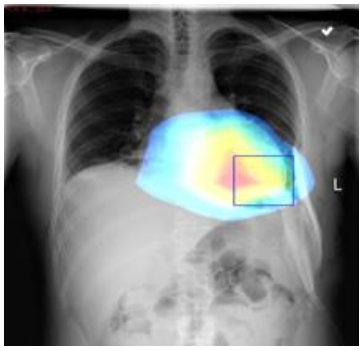
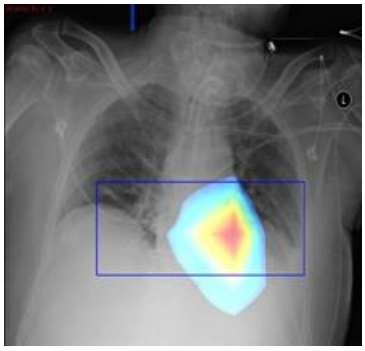
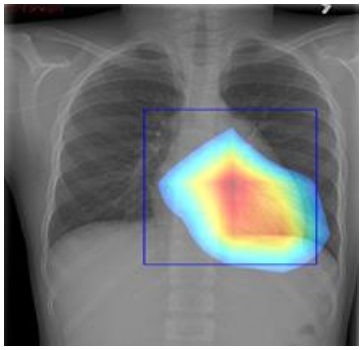
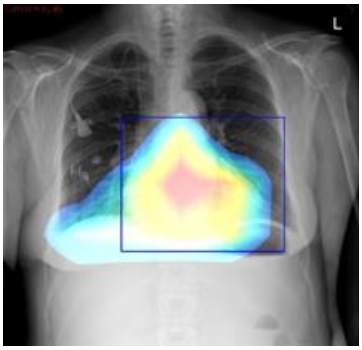
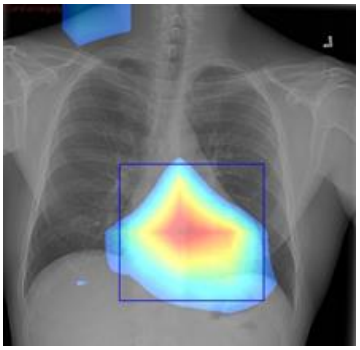
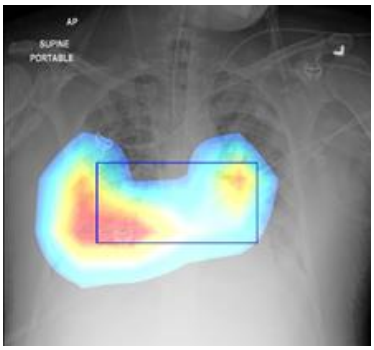
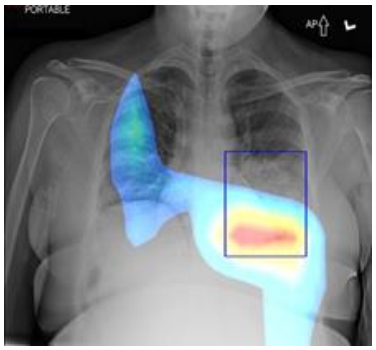
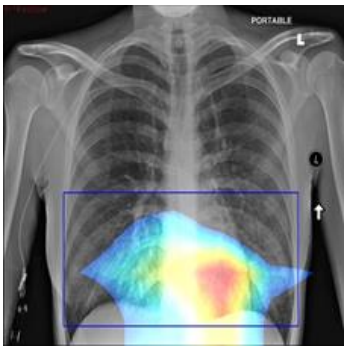
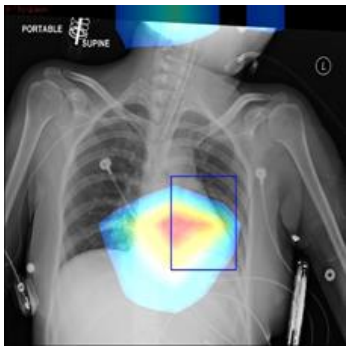
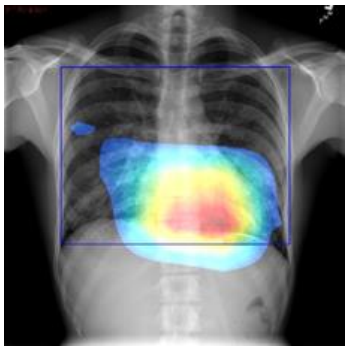
Table III contrasts the AUC Scores for the given 14 labels. It is evident from the table that the proposed Model beats the AUC scores of Wang et al [1] and Yao el al.[7] and closely completes with CheXNet[5]. Similar to CheXNet, the proposed model is also based on DenseNet 121 with few changes in hyperparameters and Data augmentation. Proposed model slightly outperforms in classifying certain diseases with higher AUC than the CheXNet, Namely: Atelectasis,Effusion,Pneumonia, Pneumothorax,Emphysema and Pleural Thickening. The Mean AUROC of our model is only 0.012 less to that the CheXNet (AUROC=0.841)[5]

TABLE III. COMPARISION OF PERFORMANCES OF VARIOUS RESEARCHERS ON THE DEFINED DATASET USINF VARIOUS ALGORITHMS & TECHNIQUES.

Classes	Wang et al. [1]	Yao et al. [7]	CheXNet [5]	Proposed Model
Atelectasis	0.716	0.772	0.8094	0.8263
Cardiomegaly	0.807	0.904	0.9248	0.8732
Effusion	0.784	0.859	0.8638	0.8939
Infiltration	0.609	0.695	0.7345	0.7325
Mass	0.706	0.792	0.8676	0.8479
Nodule	0.671	0.717	0.7802	0.7633
Pneumonia	0.633	0.713	0.7680	0.7784
Pneumothorax	0.806	0.841	0.8887	0.9084
Consolidation	0.708	0.788	0.7901	0.7406
Edema	0.835	0.882	0.8878	0.8847
Emphysema	0.815	0.829	0.9371	0.9595
Fibrosis	0.769	0.767	0.8047	0.8003
Pleural Thickening	0.708	0.765	0.8062	0.8092
Hernia	0.767	0.914	0.9164	0.8295

Table IV shows the classification and localisation of proposed model. Blue bounding box are the annotation given by the radiologists from Stanford i.e the ground truth on the test dataset and the heatmap is the localization map obtained from the designed model.

TABLE IV. VALIDATION OF MODEL OVER TEST DATA DEMONNSTRATING BOUNDING BOXES AND HEAT MAP FOR THE MODEL.

Disease	Model Testing Results		
Atelectasis			
Cardiomegaly			
Mass			
Infiltration			

IV. Conclusion

In this paper, authors implemented deep convolutional neural network which detects 14 different anomalies from frontal view of chest x-ray images and localize them on an input x-ray image. The comparison of the performance of the proposed solution to some others that previously built solution on the same chest X-ray dataset was also performed. The comparison clearly shows that proposed solution outperformed for most of the classes in comparison with previous solutions.

The solution is highly scalable and can be extended to other diseases as well. With such automated solutions, the healthcare access to other medical imaging expertise where resources are limited can be improved. The solution is reliable and fast and can be used for classification in large batch sizes as well.

References

[1] Wang, Xiaosong, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, and Ronald M. Summers. "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2097-2106. 2017.

[2] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).

[3] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778. 2016.

[4] Huang, Gao, Shichen Liu, Laurens Van der Maaten, and Kilian Q. Weinberger. "Condensenet: An efficient densenet using learned group convolutions." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2752-2761. 2018.

[5] Rajpurkar, Pranav, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding et al. "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning." *arXiv preprint arXiv:1711.05225* (2017).

[6] Shin, Hoo-Chang, Kirk Roberts, Le Lu, Dina Demner-Fushman, Jianhua Yao, and Ronald M. Summers. "Learning to read chest x-rays: Recurrent neural cascade model for automated image annotation." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2497-2506. 2016.

[7] Yao, Li, Eric Poblenz, Dmitry Dagunts, Ben Covington, Devon Bernard, and Kevin Lyman. "Learning to diagnose from scratch by exploiting dependencies among labels." *arXiv preprint arXiv:1710.10501* (2017).

[8] Maji, Kamal Jyoti, Anil Kumar Dikshit, and Ashok Deshpande. "Disability-adjusted life years and economic cost assessment of the health effects related to PM 2.5 and PM 10 pollution in Mumbai and Delhi, in India from 1991 to 2015." *Environmental Science and Pollution Research* 24, no. 5 (2017): 4709-4730.