DETECTION OF BRAIN TUMOUR USING CLASSIFICATION ALGORITHM

Amruta Chitari¹ and Mr V.V.BAG²
¹PG Research Scholar, N.K. Orchid College of Engineering and Technology, Solapur
²Associate Professor, N.K. Orchid College of Engineering and Technology, Solapur

ABSTRACT

Medical Imaging is a need for physicians. Physicians are relaying on the reports that are generated by radiologist. Lots of time and efforts are needed to be given by radiologist to prepare a report. Radiologist must read and observe the MRI image to detect and locate the lesion in MRI brain images. The research is helpful to reduce the effort of radiologist to detect the abnormalities in the MRI Image and classify them. It also help to increase the accuracy in detection of abnormalities. The main objective is to find the lesion present in MRI Image and to classify them. The brain tumor diagnostic procedure consists of three phases. The first phases involve image pre-processing. The feature extraction is the second phase and the classification using the k-nearest neighbor algorithm is the last phase.

Keywords: Magnetic Resonance Imaging, Feature Extraction, Gray scale coherence matrix, Classification-Nearest Neighbor

1. INTRODUCTION

In this era the medicine is important part of human life. Radiologist and Image diagnosis center generate millions of images daily. Medical images are obtained from biomedical devices. This medical image data indicates the presence or absence of different diseases. Mining this millions medical images is a challenge of today’s world. The radiologist manually observes the images and tries to detect and locate the abnormalities in the images. To diagnosis the abnormalities takes lot of efforts and require a lot of time. Analyzing manually make the radiologist tiresome.

Medical imaging techniques help the medical practitioners and radiologists to quickly diagnose the disease.

Brain cancer is one of the leading causes of death. The brain controls and coordinates the behavior of human body. MRI image can determine the anomalies accurately then CT_Scan, X-ray. The main reason behind brain tumor disease is abnormalities of genes or some chemical change in the structure due to working in chemical industries. Which is main cause of uncontrolled cell growth.

There are most common symptoms that vary depending on the size and around the brain (Headache, vomiting, nausea, personality changes, irritability, depression, drowsiness etc.). There are two main types of brain cancer. They are benign and malignant stage. Benign is starting stage of cancer which is curable by medicine. Secondary stage is known as malignant stage in this patient need to undergo surgery and therefore it is necessary to detect the brain tumor at beginning stage itself.

2. LITERATURE REVIEW

Lots of researches have been performed for the segmentation of normal and abnormal tissues in MRI brain images. Some of the recent related works regarding the brain tumor detection and location are reviewed in the following papers.
R. B. Dubey et al. [15] developed a semi-automatic method that was developed the segmentation of brain tumor from MR images. Replacing the constant propagation term by a Statistical force overcome many limitations and result in convergence to a stable solution.

Chunming Li et al. [4] have introduced a region-based approach for image segmentation, which has the potential to deal with intensity in homogeneities in the segmentation. Initially, the local clustering criterion function for the image intensities in neighborhood of each point has been defined based on the model of images with intensity in homogeneities. The proposed approach has been validated on synthetic images and real images of different modalities and obtained good performance in the presence of intensity in homogeneities. Experiments have revealed that the approach was robust to initialization, faster, and more accurate.

Medical image processing with data mining techniques are emerged as one of the most important tools to identify and diagnose various anomalies. This technique enables doctors to visualize and analyze the medical images for finding the abnormalities in internal structures of human organs e.g. brain [10]. Association rule mining technique is used to classify the MRI brain images. For this study three categories have been taken namely normal, benign and malign. Low level feature extracted from images and high level knowledge from the specialists is combined into the system [8]. Detection, segmentation & characterization of the Brain tumor is done by the novel approach that is by using graph mining technique. In this technique it exploits prior knowledge in the form of a sparse graph which represents the expected spatial positions of tumor classes [10].

3. MOTIVATION
In existing scenarios, the radiologist manually observes the images and tries to detect and locate the abnormalities in the images. To locate the abnormalities takes lot of efforts and require a lot of time. So, it is required to automate the process by developing a tool which will detect the brain tumor in the MRI image and stage accurately. This developed tool should strive toward improving the accuracy, precision, and computation speed of algorithms, while reducing the amount of manual interactions needed.

4. PROPOSED WORK
The proposed system presents brain tumor diagnosis, require a detailed histological analysis. In this system tumor diagnostic procedure is divided into the following phases. First phase is Image pre-processing, feature extraction is the second phase, and the last phase is Image classification.
4.1 IMAGE PRE PROCESSING

The medical image data which are obtained from multiple, heterogeneous sources and typically have huge size of image (often several gigabytes or more) due which they are highly susceptible to noisy, missing, and inconsistent data. Mining low-quality data will lead to low-quality results. It is required to improve image data by suppressing unwanted distortions and enhance image for further processing. Pre-processing is nothing but the removing of noise, bringing the image in the proper format i.e. size of the image.

The Image pre-processing [1], [7] is the technique of enhancing data images prior to computation processing. It is essential for reducing the complexity and computation time of the Computer aided tool. Image pre-processing enhances the data by normalizing the intensity of the images, resizing an image, colour to gray scale conversion and filter the image from noise using Median filter. A method of contrast adjustment using the images histogram in image processing is widely known as histogram equalization. When the useful information contained in the image is represented by close contrast values, histogram equalization generally increase the global contrast of images.

4.2 FEATURE EXTRACTION
Transforming the input data into set of features is called feature extraction. When the input data applied to an algorithm is very large to be processed and it is suspected to be redundant then the input data will be transformed into a reduced representative set of features which is called as features vector. Transforming the input data into a set of features extracted must be carefully selected so that the feature set will extract the relevant information from the input data in order to perform the desired task using these reduced feature vectors instead of the full size input pattern.

The purpose of feature extraction is to reduce the original data set by measuring certain properties, or features, that distinguish one input pattern from another. The extracted features provide the characteristics of the input type to the classifier by considering the description of the relevant properties of the image [42]. Features are extracted by using Grey level Coherence Matrix.

**Grey level- Coherence Matrix (GLCM)**
A Co-occurrence matrix also referred as a co-occurrence distribution is defined over an image to be the distribution of co-occurring values at a given offset. Represents the distance and angular spatial relationship over an image sub-region of specific size.GLCM is created from gray-scale image.GLCM is calculates how often a pixel with gray-level value i occurs either horizontally, vertically, or diagonally to adjacent pixels with the value j.

GLCM direction of Analysis
- Horizontal(0º)
- Vertical (90º)
- Diagonal
  a) Bottom left to top right (-45º)
  b) Top left to bottom right (-135º)

Contrast = \[ \sum | i - j |^2 p(i,j) \]
Correlation = \[ \sum (i - \mu_i)(j - \mu_j) \frac{p(i,j)}{\sigma_i \sigma_j} \]
Energy = \[ \sum p(i,j)^2 \]
Homogeneity = \[ \sum p(i,j) \]
\[ \frac{1}{1 + | i - j |} \]

where i , j are the grey level values in the image.
Table 1. Description Table

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contrast</td>
<td>Gives a measure of the intensity contrast between a pixel and its neighbor over the entire image. Range = [0(size(GLCM,1)-1^2] Contrast is 0 for a constant image</td>
</tr>
<tr>
<td>Correlation</td>
<td>Gives a measure of how correlated a pixel is to its neighbor over the entire image. Range = [-1 1] Correlation is 1 or -1 for a perfectly positively or negatively correlated image.</td>
</tr>
<tr>
<td>Energy</td>
<td>Returns the sum of squared elements in the GLCM. Range = [0 1] Energy is 1 for constant image.</td>
</tr>
<tr>
<td>Homogeneity</td>
<td>Returns a value that measures the closeness of the distribution of elements in GLCM to the GLCM diagonal. Range = [0 1] Homogeneity is 1 for a diagonal GLCM</td>
</tr>
</tbody>
</table>

4.3 IMAGE CLASSIFICATION

The fundamental objective for carrying out image classification in image mining is to acquire content information the users are interested in from the image group label associated with the image. Intelligently classifying image by content is an important way to mine valuable information from large image collection. The classification module in the mining system is usually called classifier. Recognizing the challenges that lies in grouping images into semantically meaningful categories based on low-level visual features. Currently, there are two major types of classifiers, the parametric classifier and non-parametric classifier develops a variety of classifiers to label the pixels in a Land set multispectral scanner image.

K -Nearest Neighbor Classification

A decision tree is a powerful method for classification and prediction and for facilitating decision making in sequential decision problems. Often the medical decision maker will be faced with a sequential decision problem involving decisions that lead to different outcomes depending on chance. If the decision process involves many sequential decisions, then the decision problem becomes difficult to visualize and to implement. They allow for intuitive understanding of the problem and can aid in decision making.
Decision tree induction is a very popular and practical approach for pattern classification. Decision tree induction is the learning of decision trees from class-labelled training tuples. KNN is a non parametric lazy learning algorithm. In other words, there is no explicit training phase or it is very minimal. This means the training phase is pretty fast.

Assumptions in KNN

KNN assumes that the data is in a feature space. More exactly, the data points are in a metric space. The data can be scalars or possibly even multidimensional vectors. Since the points are in feature space, they have a notion of distance. Each of the training data consists of a set of vectors and class label associated with each vector. In the simplest case, it will be either + or – (for positive or negative classes). But KNN, can work equally well with arbitrary number of classes. We are also given a single number "k. This number decides how many neighbors (where neighbors is defined based on the distance metric) influence the classification. This is usually an odd number if the number of classes is 2. If k=1, then the algorithm is simply called the nearest neighbor algorithm.

KNN algorithm

\[
\begin{align*}
\text{Input:} & \quad D \text{ //Training data (Feature extracted from the MRI Brain image)} \\
\text{Output:} & \quad T \text{ // (Classify the MRI brain image into 3 category that is Normal, Benign, Malignant)} \\
\text{KNN Algorithm:} & \\
& \text{// Simplistic algorithm to illustrate K nearest neighbour algorithm} \\
& \text{if } k=0 \text{ then} \\
& \quad T=\text{Normal}; \\
& \text{else if } k=1 \text{ then} \\
& \quad T=\text{Benign}; \\
& \quad \text{Apply the K –means Clustering Algorithm to } T; \\
& \text{else if } k=2 \text{ then} \\
& \quad T=\text{Malignant}; \\
& \quad \text{Apply the K-means Clustering Algorithm to } T; \\
\end{align*}
\]

4.4 EXPERIMENTAL WORK

1) Image Pre-processing

MRI brain images are obtained from the hospitals. These images may be affected by noise and may have inconsistent data. So, these images must undergo the image pre-processing process. The output window has shown below displays four buttons. The first button named by test image, clicking the button it displays the original image. Second button named as Denoising, clicking the button displays the filtered image. Third button named as
enhancement. Image is enhanced where in intensity of image is improved. Last button Submit clicking on the button this generate next the window.

![Image Process](image.png)

Fig. 2: Image Process

2) Image Classification

In the second window we browse the MRI brain image and then clicking on submit button image is classified. Image is classified into three categories, normal, benign, malignant stage.

4.5 Test results

<table>
<thead>
<tr>
<th>Test Image, Radiologist result, Output of developed tool</th>
<th>Test Image, Radiologist result, Output of developed tool</th>
<th>Test Image, Radiologist result, Output of developed tool</th>
</tr>
</thead>
</table>

Table 2: Test Results.
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1) NORMAL, NORMAL</td>
<td>3) MALIGNANT, MALIGNANT</td>
<td>3) BENIGN, BENIGN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4) NORMAL, NORMAL</td>
<td>5) BENIGN, MALIGNANT</td>
<td>6) BENIGN, BENIGN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The proposed model is more accurate than previous methods. I have applied this classification approach on 75 images out of which 73 images are classified correctly and accuracy of 97.3% is achieved with this method, which is far better than other classification techniques.

5. CONCLUSION

Image Mining is gaining its popularity in almost all applications of real world. Decision trees are so popular because they produce human readable classification rules and easy to interpret than other classification methods. Frequently used decision tree classifiers are studied and the experiments are conducted to find the best classifier for Medical Diagnosis. Using this approach, brains tumors could be efficiently classify into their types. It is also observed that K-Nearest Neighbor performs well for classification on medical data sets of increased size.

MRI brain tumor detection should strive toward improving the accuracy, precision, and computation speed of the algorithms and accurate location of the abnormalities. This is particularly important as MR imaging is becoming a routine diagnostic procedure in clinical practice. This application will also be helpful to the radiologist and neurologist for detection and locating abnormalities in brain images.
REFERENCES


