Textural Feature Extraction and Analysis for Brain Tumors using MRI

Anantha Padmanabha A G*, S Y Pattar**

*Student, M.Tech, ML, BMSCE, Bangalore, INDIA.

**Associate. Professor, ML, BMSCE, Bangalore, INDIA.

Abstract- The proposed research work is to perform textural analysis of the brain tumor on MRI images and this process aims by giving correct decisions towards medication and providing tools for automated extraction of the most discerning features of regions of interest in human brain. The detection of tumor in human brain (MRI) is performed through segmentation and for region characterization we use texture information. Based on extracted information the textural analysis of tumor is done. For the selection of region of interest we use FCM and Level set segmentation. After selecting region of interest we extract GLCM features from the segmented brain tumor image. The textural information is captured from the region of interest using Haralick textural descriptor features. To train and classify the tumor into benign or malignant we use SVM classifier. The ultimate aim is to develop an automated classification of region into one of the classes as benign or malignant. The digital medical imaging and clinical data associating on daily data are witnessed by the hospitals all over the world. The expertise will be able to analyze, inspect quickly the images.

Index Terms- MRI, Segmentation, FCM, Tumor

I. Introduction

To detect and to identify the pathogenic condition of the brain tumor we need to conduct an analysis on brain MR images. The analysis has been differentiated into three types' i.e shape analysis, intensity analysis and texture analysis. The recognition and identification of real-world objects is defined as Shape analysis. The subject that defines a particular model or intensity space is Intensity analysis. Texture analysis includes analyzing the pattern and intensity that are not even visible to human eye. A texture analysis is mainly used in longitudinal monitoring of recovery from particular disease.

Texture is a measure of intensity variation of a surface, quantifying properties such as smoothness, coarseness, and regularity. In image analyzing and computer vision it is used as region descriptors. For the analysis and characterization many different methods can be applied. To extract the textural features within medical images including fractal dimension, run-length encoding, discrete wavelet transform, and two-dimensional co-occurrence matrices can be used.

Of these, we have implemented both co-occurrence matrices and run-length matrices to classify the textures. The reason of choosing texture to characterize different types of regions resides in the fact that different organ tissues present different textures in the MR images, and thus, we expect the texture descriptors will have enough discrimination power to distinguish among different types of regions. We capture the texture information of the regions of interest using two second-degree statistical models: 1) The gray level co-occurrence matrix and 2) Gray level run length statistics.For each model a set of texture descriptors calculated.

Chest X-rays have been investigating a computer aided detection scheme to improve the sensitivity for providing assistances to the radiologists. The researcher although have worked from long years and still are working on improving the performance of CAD schemes, but still these schemes contain a more number of false positives (FPs). Due to this cause the radiologists are getting distracted and their efficiency is gradually reduced. In addition the radiologists will also lose their confidence with CAD scheme as an efficient tool. For a CAD scheme to be useful it should have a low false positive rate.

Compared to Chest X-Rays the imaging modality MRI gives more precise detailed information. This helps in taking a correct decisions regarding patient's condition of tumor growth as they give detailed textural features.

II.OBJECTIVES

Objectives of the Proposed System are to perform effective segmentation of the region of interest on organs like brain region within the human body. To extract and study the texture features of pathological regions. Analysis and classification of the pathological situation for a brain tumor. Determining whether the tumor is present and if it is benign or malignant.

Segmentation

The goal of segmentation is to simplify and change the representation of an image into something that is more meaningful and easier to analyze. MRI segmentation methods use either a single 2D or 3D image or a series of multispectral or multimodal images. Common segmentation approaches to MR images are thresholding, edge detecting, clustering, genetic algorithms, neural networks, and probabilistic techniques.

International Journal of Scientific and Research Publications, Volume 7, Issue 8, August 2017
ISSN 2250-3153

Types of Segmentation Methods.

Region Based
Clustering
Level set Method.

Region Growing:
Region-growing methods rely mainly on the assumption that the neighboring pixels within one region having similar values. The common procedure is to compare one pixel with its neighbors. If a similarity criterion is satisfied, the pixel can be set to belong to the cluster as one or more of its neighbors. The pixel with the smallest difference measured is assigned to the respective region. This process continues until all pixels are assigned to a region.
In Region Growing Segmentation, the algorithm specifies a pixel in the tumor part input image and after comparing it with the neighboring pixels, segments the tumor portion as seen in the output image.

Clustering Algorithms
The frequently used clustering algorithms are the K-Means and Fuzzy C-Means algorithm.
K-Means clustering algorithm:
K-Means is a well-known partitioning method. Objects are classified as belonging to one of k groups, k chosen a priori . Cluster membership is determined by calculating the centroid for each group and assigning each object to the group with the closest centroid. This approach minimizes the overall within-cluster dispersion by iterative reallocation of cluster members.

Fuzzy C-Means clustering algorithm:
In 1969, Ruspini has given the idea of using fuzzy set theory for clustering. The first specific formulation of Fuzzy C-Means (FCM) is credited to Dunn . But its generalization and current framing is designed by Bezdek .

Level Set Segmentation:
The level set method can be used to efficiently address the problem of curve/surface propagation in an implicit manner. The central idea is to represent the evolving contour using a signed function whose zero corresponds to the actual contour. Then, according to the motion equation of the contour, one can easily derive a similar flow for the implicit surface that when applied to the zero level will reflect the propagation of the contour. Level set methods have been shown to be versatile, robust, and efficient techniques for a wide class of problems in image processing . They work by embedding the propagating front as the zero level set of a higher dimensional function. Another advantage is that it is less sensitive to noise. The membership functions in Level set is given by

$$Mi\left(\emptyset_1(y), \ldots\ldots\ldots.\emptyset\kappa(\,y)\right) = \begin{cases} 1, y \in \Omega i \\ 0 \quad else \end{cases} \tag{1}$$

For N=3, $\emptyset_1$ and $\emptyset_2$ are level set functions.

$$\begin{aligned} M_1(\emptyset_1, \emptyset_2) &= H(\emptyset_1)\, H(\emptyset_2), \\ M_2(\emptyset_1, \emptyset2) &= H(\emptyset_1)\,(1- H(\emptyset_2)), \\ M_3(\emptyset_1, \emptyset_2) &= H(\emptyset_2)\,(1- H(\emptyset_1)), \end{aligned} \tag{2}$$

are the member functions for region. Data term energy using the level sets and member functions.

$$\in(\emptyset, c, b) = \int \sum_{i=1}^{N} e_i\,(\mathrm{x})\, M_i\,(\emptyset(\mathrm{x}))\, d\mathrm{x} \tag{3}$$

Energy in level set formulation including regularization term is given by

$$F(\emptyset, b, c) \triangleq \in(\emptyset, b, c) + R_p(\emptyset) \tag{4}$$

Minimization of energy is performed by solving the gradient flow equations.

International Journal of Scientific and Research Publications, Volume 7, Issue 8, August 2017
ISSN 2250-3153

$$\frac{\partial \emptyset_1}{\partial t} = -\sum_{i=1}^{N} \frac{\partial M_i(\emptyset)}{\partial \emptyset_1} e_i + \vartheta\delta(\emptyset_1) div\left(\frac{\nabla\emptyset_1}{|\nabla\emptyset_1|}\right) + \mu div\left(d_p(|\nabla\emptyset_1|)\right)\nabla\emptyset_1$$

$$\frac{\partial \emptyset_k}{\partial t} = -\sum_{i=1}^{N} \frac{\partial M_i(\emptyset)}{\partial \emptyset_k} e_i + \vartheta\delta(\emptyset_k) div\left(\frac{\nabla\emptyset_k}{|\nabla\emptyset k_1|}\right) + \mu div\left(d_p(|\nabla\emptyset_k|)\nabla\emptyset_k\right) \quad (5)$$

Gradient flow equation is iterated until contours are evolved to the proper shapes and sizes defining the segmentation boundaries for the regions.

Feature Extraction:
Two-dimensional co-occurrence matrices are generally used in texture analysis because they are able to capture the spatial dependence of gray-level values within an image. A 2D co-occurrence matrix, P, is an n x n matrix, where n is the number of gray-levels within an image. For reasons of computational efficiency, the number of gray levels can be reduced if one chooses to bin them, thus reducing the size of the co-occurrence matrix. The matrix acts as an accumulator so that P[i,j] counts the number of pixel pairs having the intensities i and j. Pixel pairs are defined by a distance and direction which can be represented by a displacement vector d = (dx,dy), where dx represents the number of pixels moved along the x-axis, and dy represents the number of pixels moved along the y-axis of the image slice. Run-length matrices capture the coarseness of texture in specified directions. The features extracted using both co-occurrence and run-length matrices provide valuable information about the MR images.

Feature extraction techniques:

1) SIFT - Scale-invariant feature transform (SIFT) is an algorithm in computer vision to detect and describe local features in images.
2) Gabor Filter – It is a linear filter used for edge detection. Frequency and orientation representations of Gabor filters are similar to those of the human visual system, and they have been found to be particularly appropriate for texture representation and discrimination.
3) GLCM - A statistical method of examining texture that considers the spatial relationship of pixels is the gray-level co-occurrence matrix (GLCM), also known as the gray-level spatial dependence matrix.

III.PROPOSED METHODOLOGY

The brain image is obtained from the MRI (Magnetic Resonance Imaging) scan. Images will be processed and the region of interest will be extracted. Texture analysis will be performed for the images. Identification of the disorder will be done using a classifier.

MR brain image → Brain tumor segmentation → Texture feature extraction → Texture feature analysis → Brain tumor classification

Figure 1: Block diagram of proposed methodology.

Image segmentation groups pixels into regions, and hence defines object regions. Segmentation uses the features extracted from the image.Fuzzy C Means and Level Set Segmentation techniques are studied and employed to effectively segment the region of interest.

GLCM features are extracted using Gray level Co-occurrence Matrix. The texture features such as contrast, homogeneity, energy and correlation may not contribute properly towards the effective classification. Hence GLCM features are also calculated and the tumors are classified. Classification is the process of arriving at a decision for a pathological situation which is normal or abnormal.

International Journal of Scientific and Research Publications, Volume 7, Issue 8, August 2017
ISSN 2250-3153

IV. PROPOSED ALGORITHM

Objective: To perform segmentation of Brain Tumor region.
Input: MRI scan image.
Expected Output: Segmented Brain Tumor region.
Procedure:
Step1: Initialize the following factors
- Number of clusters
- Assign centroid
- Number of iterations
- Termination parameters
- Fuzziness factor

Step2: Calculate / update membership values (µik)
- Calculate distance (dik)
- Membership values are to be calculated using calculated distances.

Step 3: Update centroids.
Step 4: Find Objective function (Jr)
Step 5: If Jr <> Jr-$_1$ then go to step 2 or stop.

Feature Extraction:
Feature Extraction is the transformation of input data into a set of features. It is a key stage in performing the task, which identifies sets of features that describe the visual texture of an image. MR Image segmentation is based on a set of measurable features which are extracted or computed from the images. Features themselves can be classified as pixel intensity-based features, calculated pixel intensity-based features and edge and texture-based features. Texture Features: An image texture is a set of metrics calculated in image processing designed to quantify the perceived texture of an image. Image texture gives us information about the spatial arrangement of intensities in an image or selected region of an image.

Classification:
In machine learning and statistics, classification is the problem of identifying a set of categories on the basis of a training set of data containing observations. Classification was performed by starting with the more discriminative features and gradually adding less discriminative features, until classification performance is no longer improved. In fuzzy set theory, every element in the universe belongs to a varying degree to all sets defined in the universe. But in fuzzy clustering objects are not classified as belonging to one and only cluster, but instead, they own a degree of membership with each of the clusters. FCM provides hyper spherically-shaped well separated clusters accurately.

V.RESULTS

a b c d

Figure2.a. Input brain image1 and b.Segmented Tumor. c .Input brain image2 and d. Segmented Tumor.

International Journal of Scientific and Research Publications, Volume 7, Issue 8, August 2017
ISSN 2250-3153

K-Fold Validation

Images	Classification Accuracy
Set 1	80.00%
Set 2	80.00%
Set 3	86.66%
Set 4	86.66%
Set 5	73.33%
Average	81.33%

Table1.Classification of benign and malignant tumors.

K-Fold Validation

K-fold validation is a model validation technique for assessing how the results of a statistical analysis will generalize to an independent data set. It is mainly used in settings where the goal is prediction, and one wants to estimate how accurately a predictive model will perform in practice. In a prediction problem, a model is usually given a dataset of known data on which training is run (training dataset), and a dataset of unknown data (or first seen data) against which the model is tested (testing dataset). The goal of cross validation is to define a dataset to "test" the model in the training phase (i.e., the validation data set).

VI.CONCLUSION & FUTURE SCOPE

The work aims to help the medical sector. Performing textural analysis of brain MRI scans, we aim to be able to detect tumors at the earliest possible stage for the patients. The Fuzzy C-means segmentation algorithm is giving effective segmentation of the tumor region. The algorithms and classifiers we used will be able to differentiate between a normal and a diseased brain and further if the tumor is benign or malignant. The normal features like contrast, homogeneity, energy and correlation are not good enough to distinguish the tumors; hence more GLCM features such as sum of variances, sum of entropy are used. The multifold validation technique is performed to validate the accuracy of the dataset being used. The average accuracy is over 80% for the proposed algorithm. Hospitals throughout the world are witnessing huge volumes of digital medical images and associated clinical data. The number of qualified personnel to inspect, analyze and make decisions is being outnumbered in relation to the number of images needing their expertise. Our work aims to make it easier for medical professionals in detecting brain tumors in patients and diagnosing them at the earliest possible stage.

The future scope of this work entails that this work can be continued by using variant Level set methods for efficient segmentation and different classifiers to further improve the classification accuracy we achieved in this work.

REFERENCES

[1] Chabat, Francis, Guang-Zhong Yang, and David M. Hansell. “Obstructive Lung Diseases: Texture Classification for Differentiation at CT1.” Chabat et al. Radiology 2003; Vol 228. pp. 871-877.

[2] Deerwester, S., Dumais, S. T., Landauer, T. K., Furnas, G. W., & Harshman, R. A.“Indexing By Latent Semantic Analysis”. Journal of the American Society For Information Science, 41, 391-407, 1990.

[3] Haralick, R.M. and L.G. Shapiro. “Computer and Robot Vision.” Addison-Wesley Publishing Co, 1992.

[4] Haralick, R. M, K. Shanmugam, and Itshak Dinstein. “Textural Features for Image Classification.” IEEE Transactions on Systems, Man, and Cybernetics, vol. Smc-3, no.6,Nov. 1973. pp. 610-621.

[5] Karkanis, S. A, G. D. Magoulas, D. Iakovidis, D. E. Maroulis, and M. O. Schurr. “On the Importance of Feature Descriptors for the Characterisation of Texture.” 4th World Multiconference on Systems, Cybernetics and Informatics (SCI 2000), vol. VI, Image Acoustic, Speech and Signal Processing, Part II. Orlando, Florida, 2000. pp. 96-101.

[6] Kass, Witkin and Terzopoulos. “Snakes: Active Contour Models”. International Journal of Computer Vision, 1988.

[7] Stan, D. and I. K. Sethi. Color Patterns for Pictorial Content Description. Proceedingsof the ACM Symposium on Applied Computing, Madrid, Spain, 2002.

International Journal of Scientific and Research Publications, Volume 7, Issue 8, August 2017
ISSN 2250-3153

[8] Tang, Xiaoou. “Texture Information in Run-Length Matrices.”IEEE Transactions on Image Processing, vol. 7, no. 11, Nov.1998.pp.1602-1609.

[9] Trefethen L. and Bau D. Numerical linear algebra. SIAM, 1997, 28-29.

[10] Xu and Prince. “Gradient Vector Flow: A New External Force for Snakes”. IEEE Proc.Conf. on Comp. Vis. Patt. Recog.,1997.

AUTHORS

First Author – **Anantha Padmanabha A G,** Student, M.Tech, ML, BMSCE, Bangalore, Email: anuram004@gmail.com
.
Second Author – S. Y. Pattar, Associate. Prof. M.Tech (PhD), ML, BMSCE, Bangalore, Email: syp.ml@bmsce.ac.in

Correspondence Author – S. Y. Pattar, Associate. Prof., ML, BMSCE, Bangalore, Email: syp.ml@bmsce.ac.in