ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 4, April 2015

Image Processing Techniques for Automatic Detection of Tumor in Human Brain Using SVM

Mr. Ajaj Khan[1], Ms. Nikhat Ali Syed[2]

M. Tech Scholar, Central India Institute of Technology, Indore, India[1]

M.E. Scholar, Jawaharlal Institute of Technology, Borawan, India[2]

Abstract: Human brain tumor creates problem in speaking, learning, loss in memories, hearing problem, problems in talking and understanding or gaining anything etc. Magnetic Resonance images are used to find the presence of brain tumor in brain. Magnetic resonance imaging (MRI) is an imaging technique that has played an important role in neuro science research for studying brain images. Classification of MRI images is an important part to differentiate between normal patients and a patient who has tumor in brain. The proposed method has two stages: First is feature extraction and second one is classification. First stage is used to extract the features from images using GLCM. In the second stage extracted features that are used as input to Support Vector machine (SVM) classifier.

Keywords: Brain tumor detection, Imaaging technique,brain tumor, medical image processing, MRI images.

I. INTRODUCTION

Image processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. It is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that image. Usually Image processing system includes treating images as two dimensional signals while applying already set signal processing methods to them. Brain tumor is caused by an abnormal growth of cell in brain. Normally brain tumor emerges from brain cells, blood vessels or nerves that are present in the brain.

Early detection of brain tumor is necessary as death rate is higher among humans having brain tumor. Most Research in developed countries proves that the numbers of people who have brain tumors were died due to the fact of inaccurate detection. The image processing techniques allows the detection of tumor tissue in brain with more accuracy as compare to manual detection. It also reduces the time for analysis of an image. At the end of the process the tumor is extracted from the MR image and its exact position and the shape also determined. The following figure embraces the fundamental steps in image processing system [2].

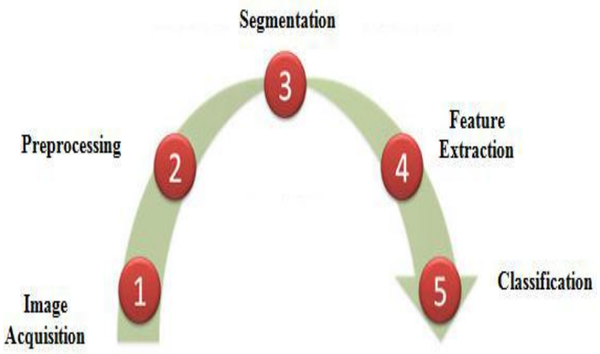

Fig. 1: Fundamentals steps in digital image processing

II. MAGNETIC RESONANCE IMAGE

Magnetic Resonance Imaging (MRI) is a scanning device that uses magnetic fields and computers to capture images of the brain on film. It does not use x-rays. It provides pictures from various planes, which permit doctors to create a three-dimensional image of the tumor. MRI detects signals emitted from normal and abnormal tissue, providing clear images of most tumors.

It has become a widely-used method of high quality medical imaging, especially in brain imaging where soft tissue contrast and non-invasiveness are clear advantages.MRI is examined by radiologists based on visual interpretation of the films to identify the presence of abnormal tissue. Brain images have been selected for the image reference for this research because the injuries to the brain tend to affect large areas of the organ [1].

The MRI machine produces the radio frequency (RF) pulse that specifically binds only to hydrogen. The system or machine sends the pulse to that particular part of the human body that needs to be examined. By the RF pulse, protons absorb the energy from that part of body to make them spin in a different direction. This is the resonance of MRI. The RF pulse makes the protons spin at the Larmour frequency, in a specific direction.

This frequency is found based on the particular tissue being imaged and the strength of the main magnetic field. MRI uses three electromagnetic fields:

Static field which is a very strong static magnetic field which polarizes the hydrogen nuclei; gradient field which is a weaker time-varying field used for spatial encoding; and a weak radio frequency field for manipulation of the hydrogen nuclei to produce measurable signals, which are collected through radio frequency antenna. The following figure shows the operation of MRI system [2].

Copyright to IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 4, April 2015

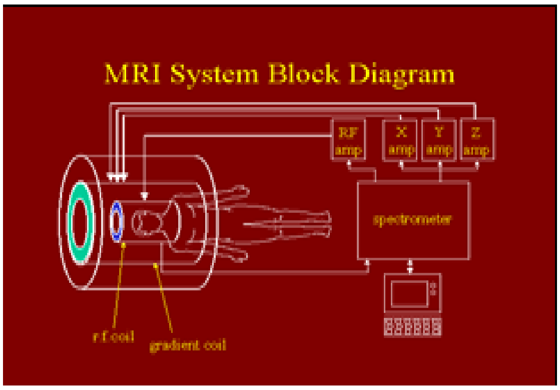

Fig. 2: Block Diagram for MRI System

III. METHODOLOGY

In the methodology part of this paper we describe our method which is used for detection of tumor in human brain using SVM. The proposed work mainly gives a review that what steps are performed throughout the entire process to detect tumor from MRI of brain. The frame work is mainly consists of two phases. In the first phase textural features are extracted from MRI and in the second phase tumor is classified as cancerous or non-cancerous. The measurements obtained from the study of textural feature are given as input to the SVM classifier for training in order to classify it.

The proposed method as described in the flowchart is based on following discussed techniques: First is Grey-Level Co-occurrence matrix (GLCM) and second is Support Vector Machine (SVM).

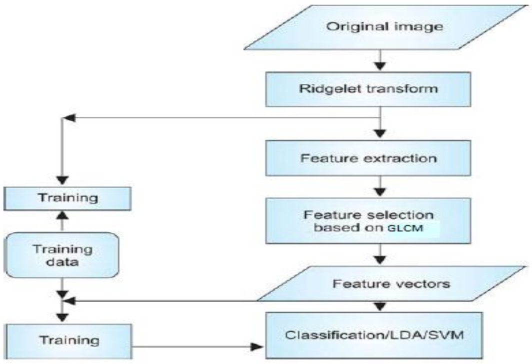

Fig: Flowchart

Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power or a classification algorithm which over fits the training sample and generalizes poorly to new samples.

IV. EXTRACTION OF GLCM

The Grey-Level Co-occurrence matrix (GLCM) method is a way of extracting second order statistical texture features. A GLCM is a matrix where the number of rows and columns is equal to the number of gray levels, G, in the image. The matrix element P (i, j | Δx, Δy) is the relative frequency with which two pixels, separated by a pixel distance (Δx, Δy), occur within a given neighborhood, one with intensity 'i' and the other with intensity 'j'. The matrix element P (i, j | d, ɵ) contains the second order statistical probability values for changes between gray levels 'i' and 'j' at a particular displacement distance d and at a particular angle (ɵ) [3].

Using a large number of intensity levels G implies storing a lot of temporary data, i.e. a G × G matrix for each combination of (Δx, Δy) or (d, ɵ). Due to their large dimensionality, the GLCM's are very sensitive to the size of the texture samples on which they are estimated. Thus, the number of gray levels is often reduced [3].In the classical paper [4], Haralick et. al introduced fourteen textural features from the GLCM and then in [5] stated that only six of the textural features are considered to be the most relevant. Those textural features are Energy, Entropy, contrast, Variance, Correlation and Inverse Difference Moment.

Energy is also called Angular Second Moment (ASM) where it measures textural uniformity [6]. If an image is completely homogeneous, its energy will be maximum. Entropy is a measure, which is inversely correlated to energy. It measures the disorder or randomness of an image [6]. Next, contrast is a measure of local gray level variation of an image. This parameter takes low value for a smooth image and high value for a coarse image. On the other hand, inverse difference moment is a measure that takes a high value for a low contrast image. Thus, the parameter is more sensitive to the presence of the GLCM elements, which are nearer to the symmetry line C(m,m) [6]. Variance as the fifth parameter is a measure that is similar to the first order statistical variables called standard deviation [7]. The last parameter, correlation, measures the linear dependency among neighboring pixels. It gives a measure of abrupt pixel transitions in the image [8].

V. SUPPORT VECTOR MACHINE

The Support Vector Machine (SVM) was first proposed by Vapnik and has since attracted a high degree of interest in the machine learning research community. Several recent studies have reported that the SVM (support vector machines) generally are capable of delivering higher performance in terms of classification accuracy than other data classification algorithms.

SVM is a binary classifier based on supervised learning which gives better performance than other classifiers. SVM classifies between two classes by constructing a hyperplane in high-dimensional feature space which can be used for classification. Hyperplane can be represented by equation-

$$w.x+b=0$$

w is weight vector and normal to hyperplane. b is bias or threshold.

Copyright to IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 4, April 2015

LINEAR SVM:
Begin with the simplest case, in which the training patterns are linearly separable. That is, there exists a linear function of the form f(x) = w T x + b (1) such that for each training example xi, the function yields ( ) ≥ 0 i f x for = +1, i y and f(xi)<0 for yi=-1.

In other words, training examples from the two different classes are separated by the hyperplane
f(x) = w T x + b =0,

where w is the unit vector and b is a constant.

For a given training set, while there may exist many hyperplanes that maximize the separating margin between the two classes, the SVM classifier is based on the hyperplane that maximizes the separating margin between the two classes (Figure 3.2).

In other words, SVM finds the hyperplane that causes the largest separation between the decision function values for the "borderline" examples from the two classes. In below Figure , SVM classification with a hyperplane that minimizes the separating margin between the two classes are indicated by data points marked by "X" s and "O"s. Support vectors are elements of the training set that lie on the boundary hyperplanes of the two classes.

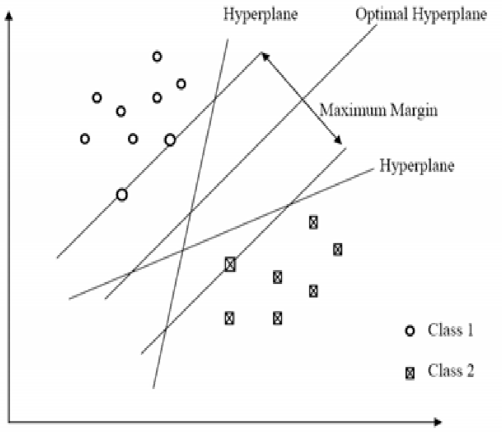

Fig: Linear SVM Classification

NON-LINEAR SVM:
In the above discussed cases of SVM classifier also shown in fig.straight line or hyperplane is used to distinguish between two classes. But datasets or data points are always not separated by drawing a straight line between two classes. For example the data points in the below fig.can't be separable by using above SVMs discussed .So, Kernel functions are used with SVM classifier.

Kernel function provides the bridge between from non-linear to linear. Basic idea behind using kernel function is to map the low dimensional data into the high dimensional feature space where data points are linearly separable. There are many types of kernel function but Kernel functions used in this research work are given below:

1. Radial basis function (RBF)
2. Linear
3. Quadratic

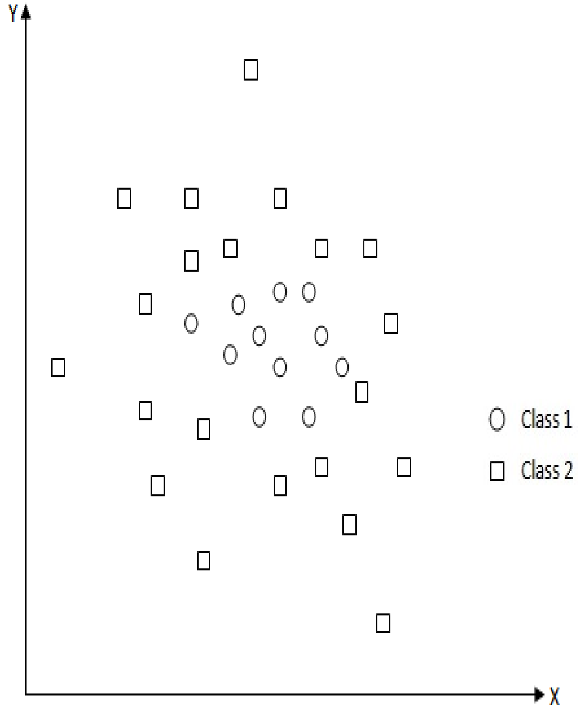

Figure: Non-linear data points

RBF is the main kernel function because of following reasons:
1. The RBF kernel nonlinearly maps samples into a higher dimensional space unlike to linear kernel.
2. The RBF kernel has less hyper parameters than the polynomial kernel.
3. The RBF kernel has less numerical difficulties.

MERITS AND DEMERITS OF SVM:
The advantages of support vector machines are:
- Effective in high dimensional spaces.
- Still effective in cases where number of dimensions is greater than the number of samples.
- Uses a subset of training points in the decision function (called support vectors), so it is also memory efficient.
- Versatile: different Kernel functions can be specified for the decision function. Common kernels are provided, but it is also possible to specify custom kernels.\

The disadvantages of support vector machines include:
- If the number of features is much greater than the number of samples, the method is likely to give poor performances.
- SVMs do not directly provide probability estimates, these are calculated using an expensive five-fold cross- validation.

The SVM Classifier Results Table:

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 4, April 2015

Feature extraction technique	Kernel Function	Sensitivity	Specificity	Accuracy	Executing on Time(Sec)
GLCM	RBF	93.67	96.64	98.69	10.321
	Linear	88.98	92.75	94.70	10.723
	Quadratic	91.47	96.21	97.89	10.468

VI. CONCLUSION

In this paper we proposed an approach of classification using Support Vector Machine Classifier which has very good working efficiency and produces the accurate results as compare to other classifiers. So that by the SVM classifier we can more accurately and effectively detect the tumor of human brain by the analysis of MRI images.

VI. FUTURE WORK

The proposed approach is computationally effective and yields good result. This automated analysis system could be further used for classification of images with different pathological condition, types and disease status. The future work is to improve the classification accuracy by extracting more features and increasing the training data set.

REFERENCES

[1]. Rosy Kumari,"SVM Classification an Approach on Detecting Abnormality in Brain MRI Images" International Journal of Engineering research and Applications, Vol. 3, Issue 4, Jul-Aug 2013, pp.1686-1690.
[2]. B. Chitradevi and R. Saranya, "A Review on Brain Tumor Detection and Classification System Based on Image Processing Techniques", International Journal for Scientific Research & Development,Vol. 2, Issue 12, 2015, ISSN (online): 2321-0613
[3]. P. Mohanaiah, P. Sathyanarayana, L. GuruKumar, "Image Texture Feature Extraction Using GLCM Approach", International Journal of Scientific and Research Publications, Volume 3, Issue 5, May 2013.
[4]. R. Haralick, K. Shanmugam, I. Dinstein, "Texture Features For Image Classification", IEEE Transaction,SMC-3(6). Pp. 610-621, 1973.
[5]. N. Otsu, "A Threshold Selection Method from Gray- Level Histogram", IEEE Trans. on System Man Cybernetics, vol. 9(I), pp. 62-66, 1979.
[6]. F. Argenti, L. Alparone, G. Benelli, "Fast algorithms for texture analysis using co-occurrence matrices", IEE Proc on Radar and Signal Processing, vol. 137(6), pp. 443-448, 1990.
[7]. P. Gong, J. D. Marceau, and P. J. Howarth, "Acomparison of spatial feature extraction algorithms for land-use classification with SPOT HRV data", Proc. Remote Sensing Environ, vol. 40, pp. 137-151, 1992.
[8]. A. Ukovich, G. Impoco, G. Ramponi, "A tool based on the GLCM to measure the performance of dynamic range reduction algorithms", IEEE Int. Workshop on Imaging Sys. & Techniques, pp. 36-41, 2005.