RESEARCH ARTICLE OPEN ACCESS

Classification of Brain Tumor Using Support Vector Machine Classfiers

Dr.D. J. Pete[1], Mrs A.J. Nirmal[2], Mrs S.N. Vaidya[3]
[1](Prof. & Head, Department of Electronics Engineering, Mumbai University, D.M.C.E., Navi Mumbai)
[2,3](Assistant Professor, Department of Electronics Engineering, Mumbai University, D.M.C.E. Navi Mumbai)

Abstract
Magnetic resonance imagi ng (MRI) is an imaging technique that has played an important role in neuro science research for studying brain images. Classification is an important part in order to distinguish between normal patients and those who have the possibility of having abnormalities or tumor. The proposed method consists of two stages: feature extraction and classification. In first stage features are extracted from images using GLCM. In the next stage, extracted features are fed as input to Kernel-Based SVM classifier. It classifies the images between normal and abnormal along with Grade of tumor depending upon features. For Brain MRI images; features extracted with GLCM gives 98% accuracy with Kernel-Based SVM Classifiesr. Software used is MATLAB R2011a.
Keywords: GLCM, Feature Extraction, MRI, Kernel-Based SVM classifiers.

I. INTRODUCTION

Brain cancer can be counted among the most deadly and intractable diseases. Cancer’s location and ability to spread, makes treatment like fighting an enemy hiding out among minefields and caves. Brain cancer’s location and ability to spread quickly makes treatment with surgery or radiation like fighting an enemy hiding out among minefields and caves. The occurrences of brain tumors are on the rise and will be detected too late, after symptoms appear. Hence, computer-assisted, advanced image-guided technology have become increasingly useful for detection, planning and in Neuro surgery.

A brain tumors cancer is classified as Non-cancerous, and cancerous, means it spreads and invades the surrounding tissue. Malignant tumors are typically called brain cancer. These tumors are graded as **Grade I to Grade IV(**Less spread to Maximum spread)**.**

It is important for brain cancer diagnosis that system should be developed for detection of cancer from a given brain MRI image and recognizes the extracted data for *Classification*. The system will be useful in the field of biomedical cancer detection. The system will be efficiently used in the area of medical science such as Computer aided diagnosis & Mammography.

Image Detection and Classification, for Various Grades of tumors is real world problem from the domain of medical image analysis that requires efficient pattern recognition. Hence, the real motivation to this research work is to work on classification technique for classifying the objects into corresponding **Grades.**

Approaches used for classification falls into two categories. First category is supervised learning technique such as Artificial Neural Network (ANN), Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) Algorithm which are used for classification. Another category is unsupervised learning for data clustering such as K-means Clustering, Self Organizing Map (SOM). In this paper, SVM is used for classification as it gives better accuracy and performance [9]. SVM is a nonlinear classification algorithm based on kernel methods.In contrast to linear classification methods, the kernel methods map the original parameter vectors into a higher (possibly infinite) dimensional feature space through a nonlinear kernel function. High dimensional input spaces can be computationally difficult and time consuming for classifiers, e.g. weight adjustment of Artificial Neural Network. It is often required that the input dimension needs to be reduced. It is desired that with the limited resources (computer memory, computer speed, etc.) a SVM classifiers can solve the computation as fast as possible. Computational efficiency of SVM is high [3].

II. METHODOLOGY

In this paper we proposed a Brain Cancer Detection and Classification System. The system uses computer based procedures to detect tumor blocks and classify the type of tumor using Kernel-Based SVM in MRI images of different patients with Meningioma type of brain tumors.

The important steps in the implementation of the system are as follows:

1. Image Acquisition (Gray Scale MR images)
2. Image Pre-processing (Median Filtering, Histogram equalization and Thresholding)
3. Image Segmentation (Square Based Segmentation and Component Labeling)
4. Feature Extraction (Texture features using Gray Level Co-occurrence Matrix GLCM)
5. Classification (Kernel-Based SVM)

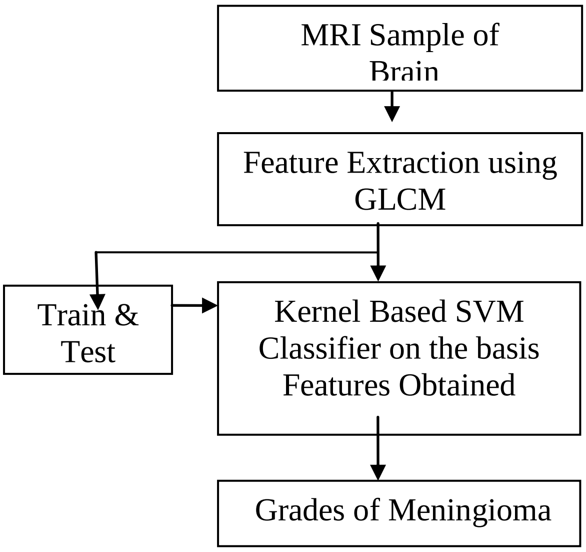

Figure2.1 Block diagram for the Proposed System.

III. FEATURE EXTRACTION

1. Maximum Probability:

f1= max P(i,j)

2. Contrast: A measure of difference moment and is defined as:

$$f2= \sum_{i,j=1}^{N} |i-j|^2 P(i,j)$$

3.Inverse Differencet Moment (Homogenity): A measure of local homogeneity that can be defined as

$$f3 = \frac{\sum_{i,j=1}^{N} P(i,j)}{1+(i-j)^2}$$

4. Entropy: A measure of non-uniformity in the image based on the probability of co-occurrence values and can be defined

$$f4 = \sum_{i,j=1}^{N} P(i,j)[-log(P(i,j))]$$

3. Angular second moment (Energy): A measure of homogeneity that can be defined as

$$f5 = \sum_{i,j=1}^{N} (P(i,j))^2$$

4. Correlation Coefficient: A measure of linear dependency of brightness and can be defined

$$f6= \frac{\sum_{i,j=1}^{N} ij \; P(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$

7. Dissimilarity

$$f7 = \sum_{i,j=1}^{N} P(i,j) \; |i-j|$$

8. Grey Level Co-occurrence Mean: It is an average value and measures the general brightness of an image

$$f8= \sum_{i,j=1}^{N} i \, P(i,j)$$

9. **Variance**

$$f9= \sum_{i,j=1}^{N} P(i,j) \; (i-\mu_i)^2$$

IV. SVM BASED CLASSIFICAYION

Classification is the procedure for classifying the input pattern into analogous classes. When the input data set is represented by its class membership, it is called supervised learning. It employs two phases of processing- training phase and testing phase. For training phase, characteristics properties of image features are isolated and a unique description of each classification category is created. In testing phase these features space partitions are used to classify image features [9].

SVM is a nonlinear classification algorithm based on kernel methods. In contrast to linear classification methods, the kernel methods map the original parameter vectors into a higher (possibly infinite) dimensional feature space through a nonlinear kernel function.

4.1 Linear SVM Classifier

SVM maps input vectors into a higher dimensional vector space where an optimal hyper plane is constructed. Among the many hyper planes available, there is only one hyper plane that maximizes the distance between itself and the nearest data vectors of each category. This hyper plane which maximizes the margin is called the optimal separating hyper plane and the margin is defined as the sum of distances of the hyper plane to the closest training vectors of each category The basic theme of SVM is to maximize the margins between two classes of the hyperplane [3][10]. The description is given below:

Step1: The simplest form of discriminating function is linear. Linear discriminating function f(x) is written as:

$$f(x)=w^T.x+b$$

Expression for hyper plane

$$f(x)=w^T.x+b = 0$$

x – Set of training vectors

w – vectors perpendicular to the separating hyper plane

b – offset parameter which allows the increase of the margin

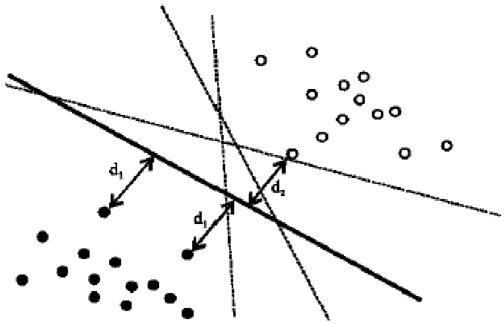

Figure 4.1: SVM Classifier

Margin is d_1+d_2

Step2: Maximal margin

Consider the class of hyperplanes $w^T x + b = 0$, $w \in R^n$, $b \in R$, correspondingto a decision function

$$f(x) = \text{sign}(w^T x + b)$$

A hyperplane is constructed which maximally separates the classes :**(maximum margin)**

$\max_{w,b} \min [\ \|x - x_k\|\ ;\ x \in R^n,\ w^T x + b=0,\ k=1,\dots N]$

To show how this hyperplane can be constructed in an efficient way, we need use definitions of Separability given by following equation :

A training set $D = \{(x_1, y_1), ..., (x_N, y_N) :\ x_k \in R^n,\ y_k \in \{-1,+1\}\}$ is called separable by a hyperplane $w^T x + b = 0$ if there exist both a unit vector w ($\|w\|= 1$) and a constant 'b' such that the following equalities hold:

$$w^T x_k + b \geq +1 \text{ for } y_k = +1 \quad \text{……(1)}$$

$$w^T x_k' + b \leq -1 \text{ for } y_k= -1 \quad \text{………(2)}$$

Step3: Optimal separating hyperplane or maximum margin hyperplane

The Optimal hyperplane of a training set D is defined by:

$$(w^*, b^*) = \arg\max D(w,b)$$

The unit vector w* and the constant b* which maximize the margin of the training set D (w, b) and also satisfy the condition (1) and (2).

Step4: The Euclidean Distance 'd' can be calculated.

$|f(x)|$ is a measure of Euclidean distance of the point 'x' from decision hyperplane. One side of the plane f(x) has positive values and on the other negative. In the special case b=0 the hyperplane passes through the origin.

Some criteria commonly used in classification are distance measure. In the following those criteria are explained:

• Distance measure is the simplest and most direct approach to classify data points. Basically, the idea is to classify a data point into the class closest to it. The Euclidean distance is the most common definition. Suppose we have 'K' classes with (μi, Si) as the known parameter set of class 'i', where μi is the reference vector of class 'i' and Si is the covariance. The Euclidean distance of an observation vector 'x' from class 'i' is given by the following equation [3].

$$d_i(x)=\sqrt{\|\ x - \mu_i\ \|^2}$$

4.2 Non-Linear SVM

The first section introduces the idea of maximal margin classification, optimal separating hyperplane, followed by kernel methods as the basis for the extension towards nonlinear classification as introduced by Vapnik.

Kernel function is used when decision function is not a linear function of the data and the data will be mapped from the input space through a non linear transformation rather than fitting non-linear curves to the vector space to separate the data. With an optimal kernel function implemented in SVM model, the classification task is able to scale high dimensional data relatively well, tradeoff between classifier complexity, and classification error can be controlled explicitly [14]. Various steps are given below:

Kernel Criteria

Steps involved:

Step1. Let $x \in D \in R^n$ denote a real valued random input vector, and $y \in \{-1,+1\}$ discrete real valued random output variable and let $\Omega \in R^{nH}$ denote a high dimensional feature space. The SVM method basically maps the input vector 'x' into the high dimensional feature space through some nonlinear mapping $\varphi : D \rightarrow \Omega$. In this feature space, one consider the linear function

$$f(x) = \text{sign}\ [w^T(x) + b]$$

This linear function is well in solving classification problems, however, it remains a problem to solve the calculation in the high dimensional feature space. Interestingly, no explicit construction of the nonlinear mapping 'φ(x)' is needed. This is motivated by the following result.

Step2. The inner product in the feature space '$\varphi(x_k)^T \cdot \varphi(x_l)$' can be replaced with the corresponding kernel $K(x_k, x_l)$ satisfying Mercer's condition.

Using Mercer's theorem to replace the inner product φ $(x_k)^T$φ (x_l) with its corresponding kernel $K(x_k, x_l)$ is often called the kernel trick. It enables us to work in a huge dimensional feature space without actually having to do explicit computations in this space. Computations are done in another space after applying this kernel trick.

In the case of support vector machines, one starts from a formulation in the primal weight space with a high dimensional feature space by applying transformations φ (·). The solution is calculated not in this primal weight space, but in the dual space of Lagrange multipliers after applying the kernel trick. In this way classification is done implicitly in a high dimensional feature space rather than in the original input space.

Step3. Non-Linear Conversion

With slight modification, for the nonlinear case we can write

$$w^T \varphi(x_k) + b \geq +1 \quad \text{for } y_k = +1$$
$$w^T \varphi(x_k) + b \leq -1 \quad \text{for } y_k = -1$$

In this quadratic form, the kernel trick is applied

$$K(x_k, x_l) = \varphi(x_k)^T \varphi(x_l) \quad \text{for } k = 1, \ldots, N.$$

Finally the nonlinear SVM classifier takes the form

$$y(x) = \text{sign}[\ \textstyle\sum_{k=1}^{N} \alpha k \alpha_k y_k yk \ \ K(x_k, x_l)+b]$$

Choice of kernel function:

Two common chioces of kernel functions are:

(1) $K(x,z) = \exp(-\|x - z\|^2/\sigma^2)$ **(Radial Basis Function)**

(2) $K(x,z) = (\tau + x^T z)^d$ **(Polynomial of Degree d)**

The kernel function most commonly used is polynomial function and Gaussian radial basis function.

V. RESULT

The implementation of the paper is carried out in MATLAB R20010a with Image Processing Toolbox and Neural Network Toolbox. In this proposed work, we have made used of Normal Brain MRI images and Meningioma type of brain tumor MRI images, which consists of four different grades of tumors as Grade I, Grade II, Grade III and Grade IV types. In segmentation method, each histogram equalized image is examined for the Region Of Interest (ROI). The ROI here in this work is, the tumor part. This segmentation should be stopped when tumor is able to be detected. In our work, we suggest a set of 12 GLCM based textural features which can be extracted from each of the gray tone spatial-dependence matrices. For classification we used SVM with 12 inputs, 6 hidden layers and 4 output layer. The 12 inputs to the neural networks are the 12 texture features that are extracted. The training in SVM is of supervised mode. The training is done for total of 45 images, out of which 5 are of normal images, 8 images of Grade I type, 8 images of Grade II type, 16 images of Grade III type and 8 images of Grade IV type. In the testing phase, a total of 15 images (3 x 5=15, 3 images from each set) are given as input to the SVM. Fig 3.1 shows the Result of Detection of tumor from the Brain MR Image. The result of individual step are shown as, Where (a) The Brain Tumor image (b)Enhanced Image after Filtering, Histogram Equalization, segmentation & Thresholding. (c) shows the result of classification in testing phase

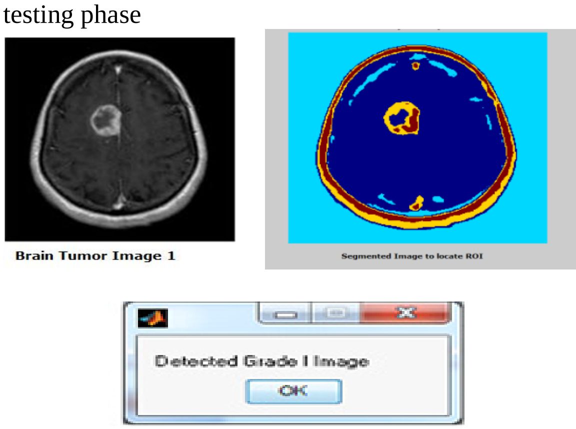

Fig 5.1 Detection & classification of the tumor from Brain MRI image

VI. ACCURACY OF CLASSIFICATION

The overall accuracy of the system is found to 94.8718%. Grade I, Grade II and Grade IV tumors have an accuracy of 100% , which means that all the input images are correctly being predicted by the developed system. And Grade III tumors have an accuracy of 92%. Table 6.1 gives system accuracy rate. Fig 6.2 shows the Accuracy of Classifcation of Brain Tumor for Linear, RBF & Quadratic- SVM.

Tumor Type	**No. of Input Images**	**Number of Correctly Predicted Images**	**Accuracy Rate**	**Spread of Tumor**
Class I	8	8	100%	Less Spread
Class II	8	8	100%	Moderate
Class III	16	15	92%	Moderate
Class IV	8	8	100%	Max Spread
Normal	5	4	80%	Nil

Total Accuracy = 94.8718%

Table 6.1: Accuracy Rate for All the Tumor Grades and Normal Images

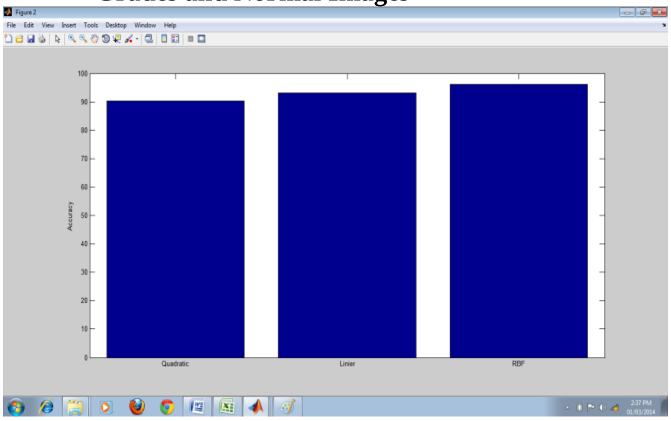

Figure6.2 Accuracy of Classification of Brain Tumor for Linear, RBF, Quadratic- SVM

VII. CONCLUSION

This Paper work presents an automated recognition system for the MRI image using the Kernel-Based SVM. It is observed that the system result in better classification during the recognition process. The considerable iteration time and the accuracy level is found to be about 50-60% improved in recognition compared to the existing neuro-classifier.

The system has been tested with the Meningioma type of brain cancer Images only. The system can be designed to classify other types of cancers as well with few modifications. Also, large patient data can be used to improve accuracy. More features can be added and the most discriminating features can be selected for training to increase the accuracy and to make the system robust .

REFERENCES

[1] Kadam D. B., Gade, S. S., M. D. Uplane and R. K. Prasad4 " **Neural Network Based Brain Tumor Detection Using MR Images",** International Journal Of Computer Science

[2] Sonali Patil, V.R. Udupi **"Computer Aided Diagnostic System For Classification Of Brain Tumors Using Texture Features & Probabilistic Neural Network"** International Journal Of Computer Science Engineering Vol. 3, Issue1, March 2013.

[3] Prof. J. Vandewalle, Prof. S. Van Huffel " **Least Squares Support Vector Machines Classification Applied To Brain Tumour Recognition".**

[4] Jayashri Joshi, Mrs.A.C.Phadk **"Feature Extraction and Texture Classification in MRI",** Special Issue of IJCCT Vol. 2 Issue 2, 3, 4; 2010, International Conference [ICCT-2010], 3rd-5th December 2010.

[5] Ming-Ni Wu, Chia-Chen Lin **"Brain Tumor Detection K-Means Clustering Segmentation".**

[6] T. Logeswari and M. Karnan **"An Improved Implementation Of Brain Tumor Detection Using Segmentation based on Soft Computing".**

[7] Rajeswari.S, Theiva Jeyaselvi.K, **"Support Vector Machine Classification For MRI**

[8] **Images".** International Journal of Electronics and

[9] Computer Science Engineering.

[10] Qurat-ul-ain, Ghazanfar Latif, Anwar M Mirza **"Classification &Segmentation of Brain Tumor using Texture Analysis"**

[11] Rosy Kumari **"SVM Classification an Approach on Detecting Abnormality in Brain MRI Images"** International Journal of Engineering Research and Applications.Vol. 3 Issue 4, July-Aug 2013.

[12] Robert M Haralick, K Shanmugam, And IT'Shak Dinstein **"Textural Features For Image"** , IEEE Transctions on systems, man & Cybernetics. Vol.SMC-3,No 6, Nov2000.

[13] Mamata S.Kalas **"An Artificial Neural Network for Detection of Biological EarlyBrain Cancer"** ©2010 International Journal of Computer Applications (0975 – 8887) Volume 1 – No. 6.

[14] S. Murugavalli, V. Rajamani, **"An Improved Implementation of Brain Tumor Detection Using Segmentation Based on Neuro Fuzzy Technique",** Journal of Computer Science 3

[15] Ashby LS, Troester MM, Shapiro**"Central Nervous System Tumors"** CancerTherapeutics, Vol.1, pp. 475-513, 2006.

[16] http://www.fp.ucalgary.ca/mhallbey/tutorial.htm Bayer Mryka Hall, 2007.GLCM Tutorial.

[17] **"Identification of brain tumors in 2D MRI using automatic seeded region growing method",** Journal of Education, ISSN: 2249-5517 & E-ISSN: 2249-5525, Volume 2, Issue 1, 2012, pp.-41-43, Swati Tiwari, Ashish Bansal, Rupali Sagar.