International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-1 Issue-6, August 2012

Classification of Abnormalities in Brain MRI Images Using GLCM, PCA and SVM

Daljit Singh, Kamaljeet Kaur

Abstract: Accurate automatic detection and classification of images is very challenging task whether they are medical images or other natural images. This paper presents a hybrid technique for automatic classification of MRI images as well as natural images. The proposed method consists of two stages: feature extraction and classification. In first stage, features are extracted from images using PCA and GLCM. In the next stage, extracted features are fed as input to SVM classifier. It classifies the images between normal and abnormal along with type of disease depending upon features. Also it classifies between natural images. For Brain MRI images; features extracted with GLCM gives 100% accuracy with SVM -RBF kernel function. Similarly for natural images; features extracted by GLCM gives 91.67% accuracy with SVM-RBF kernel function. Software used is MATLAB R2011a. Main focus is given on Brain MRI images classification as it deals with precious human life.
Index Terms: Feature, GLCM, Kernel, MRI, PCA, SVM.

I. INTRODUCTION

Now-a-days, Computer technology covers a wide range of medical area; as cancer research, heart diseases and brain diseases. [1] MRI is commonly used imaging technique uses magnetic fields and radio waves to produce high-quality two or three dimensional images of body, thus it is a non-aggressive, non-radioactive and pain-free technique for visualizing and detecting the brain tumors without any human involvement. It gives the detailed information regarding normal and abnormal tissue. Accurate brain diagnosis can be done automatically with more accuracy in feature extraction and classification of disease. 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) which are used for classification purposes.

Another category is unsupervised learning for data clustering such as K-means Clustering, Self Organizing Map (SOM). In this paper, [3] SVM is used for classification as it gives better accuracy and performance than other classifiers.

II. METHODOLOGY

The proposed method as described in fig. 5 (a) and fig. 5(b) is based on following discussed techniques: Grey-Level Co-occurrence matrix (GLCM), Principal Component Analysis (PCA) and Support Vector Machine (SVM). This method consists of two stages: Feature Extraction and Feature Classification. As SVM is binary classifier; it is used to classify the extracted features into further two classes. For medical images; it classifies between normal and abnormal images along with type of abnormality exist and for natural images; it classifies whether it is author's image or any other natural image in train dataset. Two classes have been defined i.e. class 0 and class 1. In Brain MRI images; class 0 is defined for normal images and class 1 is defined for abnormal images. Images are classified by specialist and SVM. For natural images; class 1 is defined for author's images and class 0 is defined for other natural images.

III. DATABASE

As discussed above; this work consists of two databases: for medical images and other is for natural images. Database used for Brain MRI images consists of total 59 images; out of which 11are normal and 48 are abnormal which are of bleed, clot, Acute-infarct, tumor and trauma. All the images are T2-Weighted MRI images with different views but with same resolution. Dataset used for natural images classification consists of total 28 images; out of which 11 are author's images and remaining 17 are other natural images. This dataset consists of author's images and other natural images. Only few images of Brain MRI are shown here in below figures.

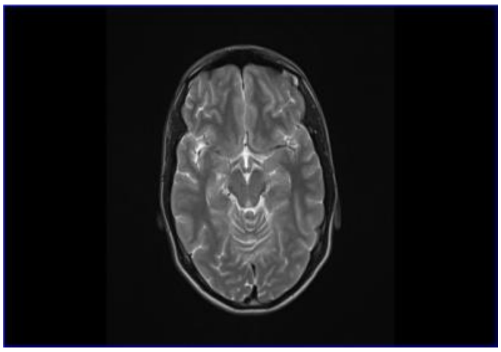

Figure 1: Normal Brain MRI image

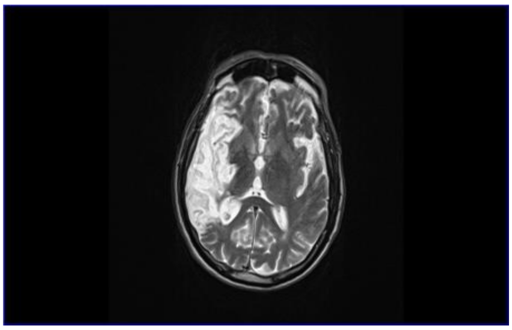

Figure 2: Brain MRI image infected by clot

Manuscript published on 30 August 2012.
***** Correspondence Author (s)
Daljit Singh*, Electronics and Communication Engineering, Ludhiana College of Engineering and Technology, Malerkotla, India,
Kamaljeet Kaur, Electronics and Communication Engineering, Ludhiana College of Engineering and Technology, Ludhiana, India.

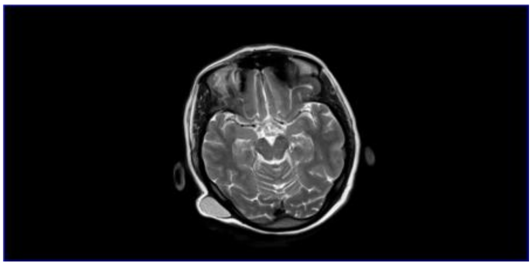

Figure 3: Brain MRI image infected by tumor

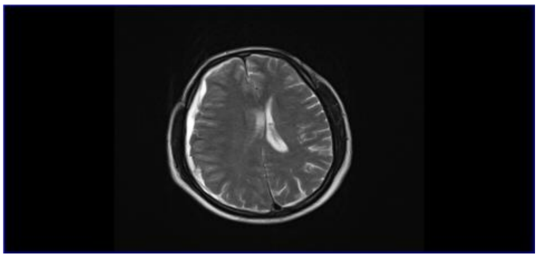

Figure 4: Brain MRI image infected by bleed

IV. FEATURE EXTRACTION

Features are said to be properties that describes the whole image. It can also refer as an important piece of information which is relevant for solving the computational task related to specific application. The purpose of feature extraction is to reduce the original dataset by measuring certain features. The extracted features acts as input to classifier by considering the description of relevant properties of image into feature space.

Authors have based their feature extraction on PCA and GLCM. PCA is a transformation that converts the set of correlated variables into set of uncorrelated variables. First principal component has maximum variance. GLCM calculates the co-occurrence matrix of an image by computing how often a pixel with a certain intensity 'i' occurs in relation with other pixel 'j' at a certain distance d' and orientation. In our paper, statistical features based on image intensity and features from gray level co-occurrence matrix are used to distinguish between normal and abnormal patient.

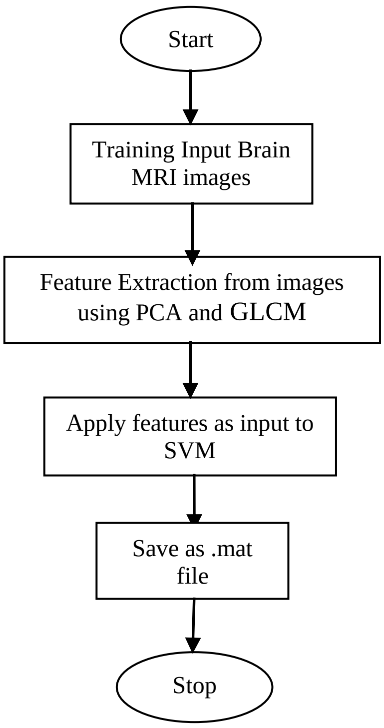

Figure 5(a): Flow chart to train SVM classifier

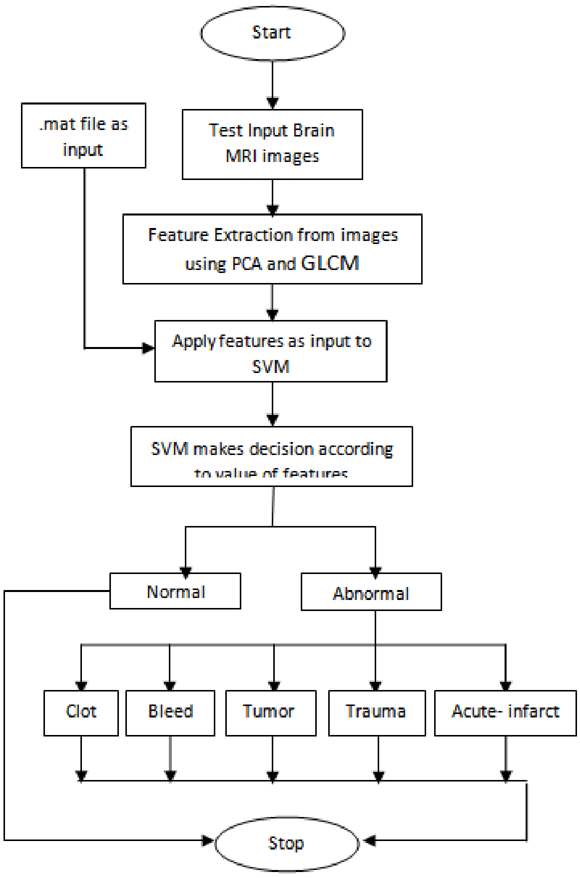

Figure 5(b): Flow chart to test SVM classifier

Above fig. 5(a) & fig. 5(b) shows the methodology to train and test the SVM classifier; for classification between normal and abnormal images. Once they are classified as abnormal; they are further classified between types of an existing abnormality as shown in fig.5 (b).

A. Extracted features are listed below:

Mean: It is an average value and measures the general brightness of an image.

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

Entropy: It measures the non-uniformity in the image based on probabilities of co-occurrence values.

$$\sum_{i,j=0}^{N-1} P(i,j)[-ln(P(i,j))]$$

Contrast: It is a measure of intensity contrast between a pixel and its neighbor pixel over a whole image. It is zero for constant image.

$$\sum_{i,j=0}^{N-1} |i-j|^2 P(i,j)$$

Energy: It returns the sum of squared elements in GLCM. It is 1 for constant image.

$$\sum_{i,j=0}^{N-1} P(i,j)^2$$

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

Energy is also known as uniformity, Angular Second Moment.

V. CLASSIFICATION

Classification analyses the numerical properties of image features and organize the data into different categories. It employs two phases of processing- training phase and testing phase. In training phase, characteristic 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.

VI. SUPPORT VECTOR MACHINE (SVM)

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-

$$w.x + b = 0 \qquad (1)$$

w is weight vector and normal to hyperplane.

b is bias or threshold.

A. Linear Separable Binary Classifier

Consider N training points, where each input x_i has 'A' attributes and is in one of two classes.[4] $y_i = +1$ or -1, i.e. training data is of the form: (x_i, y_i), i=1, 2, 3…N; $y_i \in \{+1, -1\}$.

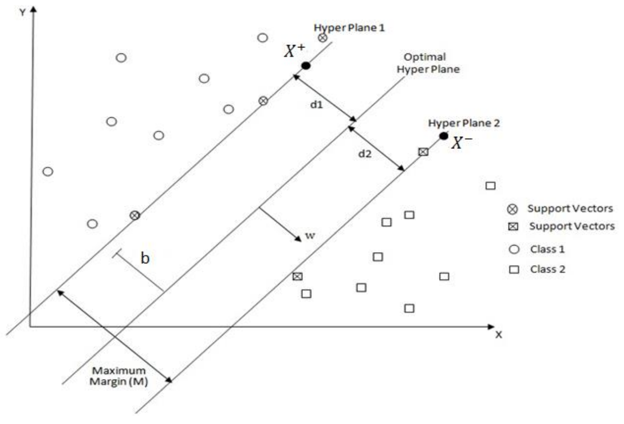

Figure 6: Linear separable binary classification.

The main purpose of using Support Vector Machine (SVM) is to orientate hyperplane in such a way as to be as far as possible from the closest members of both classes. Training data can be described by:

$$(w.\ x_i + b) \geq 1 \ for\ y_i = +1 \qquad (2)$$

$$(w.\ x_i + b) \leq 1 \ for\ y_i = -1 \qquad (3)$$

Equation (2) and (3) can also be written as:

$$y_i\ (w.\ x_i + b) - 1 \geq 0 \ \ \forall i \qquad (4)$$

Consider the points as shown in fig.6 that lie closest to the separating hyperplane, i.e. the Support Vectors; then the two planes; hyperplane 1 and hyper plane 2 lie on points which can be described as:

For Hyper Plane 1, $w.x_i + b = +1$ (5)

For Hyper Plane 2, $w.x_i + b = -1$ (6)

Mathematics geometry analysis defined a distance from point P (m, n) to a line $ax + by + c = 0$ as:

$$\frac{(am + bn + c)}{\sqrt{a^2 + b^2}}$$

In the same way, the distance from a point X^+ on HyperPlane1 to Optimal HyperPlane is given by $d1$ as

$$d1 = \frac{1}{||w||} \qquad (7)$$

Similarly from a point X^- on HyperPlane2 to Optimal HyperPlane is given by $d2$ as:

$$d2 = \frac{1}{||w||} \qquad (8)$$

By adding (7) and (8) we will get margin width

$$M = \frac{2}{||w||} \qquad (9)$$

Simple vector geometry in equation (9) shows that the margin is equal to $2/||w||$ and maximizing it subject to the constraint in (4) is equivalent to finding:

$$min\ ||w||$$

$$Subject\ to\ \ y_i\ (w.\ x_i + b) - 1 \geq 0 \ \ \forall i \qquad (10)$$

Minimizing $||w||$ is equivalent to minimizing $\frac{1}{2}||w||^2$(the factor $\frac{1}{2}$ being used for mathematical convenience) and the use of this term makes it possible to perform Quadratic Programming (QP) optimization later on. We therefore need to find:

$$min\ \frac{1}{2}||w||^2$$

$$st\ \ y_i\ (w.x_i + b) - 1 \geq 0 \quad \forall i \qquad (11)$$

In order to cater for constraints in this minimization, we need to allocate them Lagrange's multiplier (LM)α, where $\alpha_i \geq 0\ \forall i$; the primal form of the function is:

$$L_{P=}\ \frac{1}{2}||w||^2 - \alpha[\,y_i\ (w.x_i + b) - 1] \ \ \forall i \qquad (12)$$

To get the optimal hyper-plane, the resulting classifier and objective functions are:

$$f_\alpha(x) = sgn[\textstyle\sum_{i=1}^{N} y_i \alpha_i\ (x_i.x) + b] \qquad (13)$$

The equation (12) gives the new formulation which being dependent on α, thus need to maximize it:

$$L_D = \textstyle\sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i,j} \alpha_i \alpha_j y_i\ y_j\ \ x_i\ x_j \qquad (14)$$

$$Subject\ to \quad \alpha_i \geq 0 \quad \forall i$$

This is new formulation referred to as the dual form of the primary L_P.

Having moves from minimizing L_P to maximizing L_D, find:

$$\max_{\alpha}[\textstyle\sum_{i=1}^{N} \alpha_i - \frac{1}{2}\alpha^T H\ \alpha] \qquad (15)$$

$$st \quad \alpha_i \geq 0 \quad \forall i\ , \quad \textstyle\sum_{i=1}^{N} \alpha_i\ y_i = 0$$

This is convex quadratic optimization problem thus need to run a quadratic programming solver which will return α and by differentiating equation (12); it gives 'w':

$$w = \textstyle\sum_{i=1}^{N} y_i \alpha_i x_i$$

B. Linear non-separable binary classifier:

When the training dataset is not linearly separable as shown in fig.7; optimal separating hyperplane is found by solving an optimization problem relaxed. by introducing a set of slack variable ξ_i and a penalization for cases that are misclassified or inside the margin.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

The task for finding the optimal hyper plane is to minimize the following objective function,

$$\min \frac{1}{2}||w||^2 + C\sum_{i=1}^{N}\xi_i, \qquad i=1, 2, 3\ldots\ldots N$$

$$\text{subject to} \qquad y_i\,(w.\,x_i+b) \geq 1-\xi_i$$

$$\xi_i \geq 0 \; i=1, 2\ldots\ldots N$$

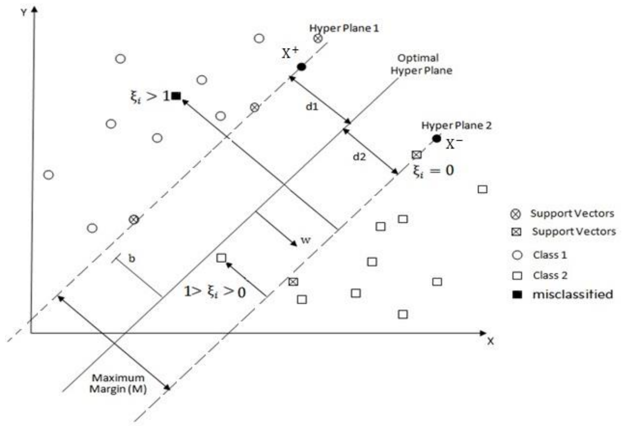

Figure 7: Linearly non-separable binary classification.

C is regularization parameter and controls the trade-off between the slack variable penalty and the size of the margin.

1. Small C allows constraints to be easily ignored i.e. wide margin. It allows a lot of samples that are not in ideal position.
2. Large C allows constraints hard to ignore i.e. narrow margin. It allows very few samples that are not in ideal position.
3. C=∞ enforces all constraints i.e. hard margin.

C. Non-Linear SVM:

In both above discussed cases of SVM classifier also shown in fig.6 and fig.7, 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 fig.8 can't be separable by using both SVM's discussed cases.

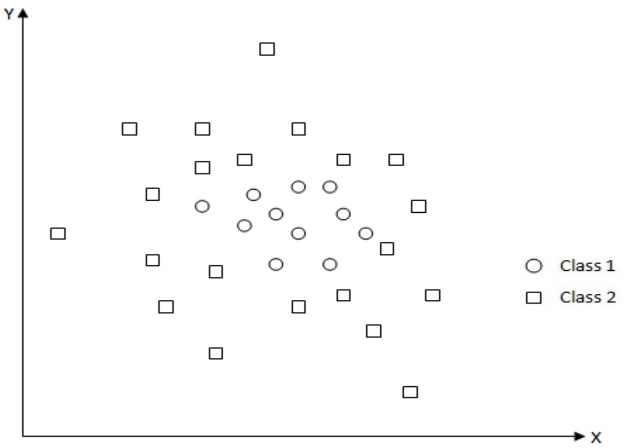

Figure 8: Non-linear data points

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

VI. RESULTS

After classification; it has been concluded that value of contrast by PCA for author's image lie 0.5626 to 0.6879. Similarly, value of contrast GLCM for author's image lies 0.0000 to 0.0015. Images having contrast value outside this range are considered as other natural images in train dataset.

In Brain MRI images; normal images have contrast value varies from 0.3593 to 0.4215 for PCA and contrast value varies from 0.7000 to 0.7550 for GLCM. Images outside this range are considered as abnormal images.

As GLCM gives maximum classification accuracy; images are further categorized into type of existing abnormalities. If mean value lie 14.0000 to 14.5000 then patient is suffering from bleed. For clot; mean value lie 26.1500 to 27.8200 and for tumor mean value lie 24.6500 to 26.1200.

Features are classified by SVM with other kernel functions also but not discussed here. From the above discussed results; it has been concluded that for natural images, out of 12 images only 9 are correctly classified whereas 3 are misclassified with features by PCA classified by SVM with RBF kernel function. Similarly features extracted by GLCM when classified by SVM with RBF kernel function; out of 12 images, 11 are correctly classified and 1 is misclassified.

For Brain MRI images; out of 26 images, only 15 images are correctly classified with features by PCA classified by SVM with RBF kernel function. Similarly for GLCM; all images are correctly classified without any misclassification. Only those kernel functions are discussed in this paper which gives better classification performance. Others are not discussed here.

A. Graphical representation:

In this section, experimental results of the proposed technique are shown in graphical form for both natural and Brain MRI images using different kernel functions in SVM. Comparison between algorithms is made on the basis of accuracy and execution time.

For graphs of brain MRI images; red points denotes data points for normal images data points, green points denotes data points for abnormal images and blue points denotes data points for images in test dataset. Black circles are support vectors and lines drawn are hyperplanes.

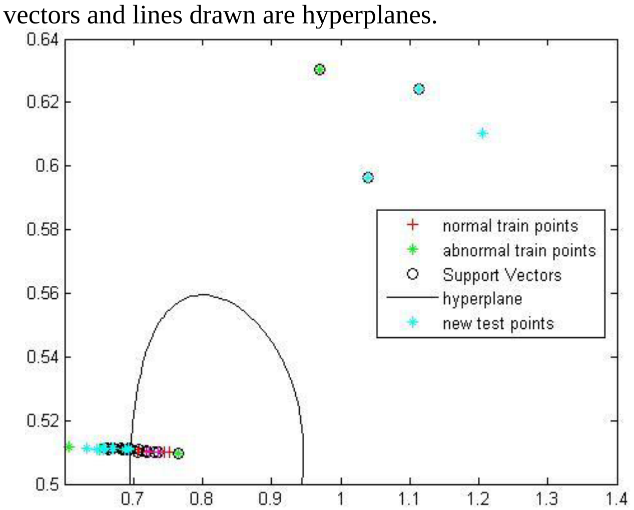

Figure 9: GLCM with SVM-RBF Kernel function for Brain MRI Images

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

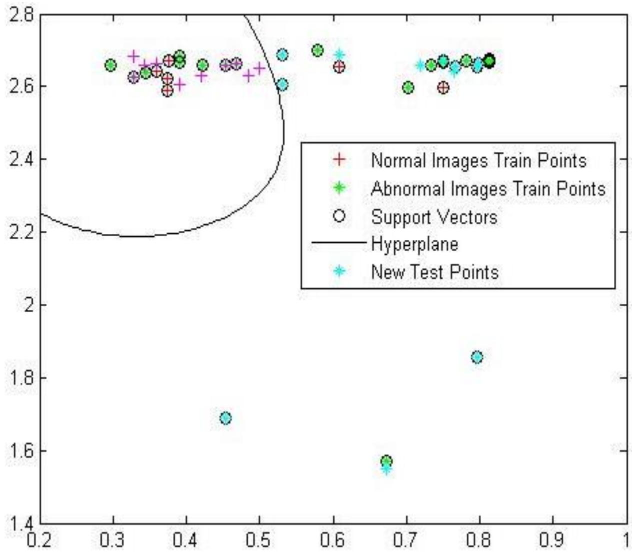

Figure 10: PCA with SVM-RBF Kernel function for Brain MRI images

For graphs of natural images; red points denotes data points for author's images, green points denotes data points for other natural images and blue points denotes data points for test images. Black circles are support vectors and lines drawn are hyperplanes.

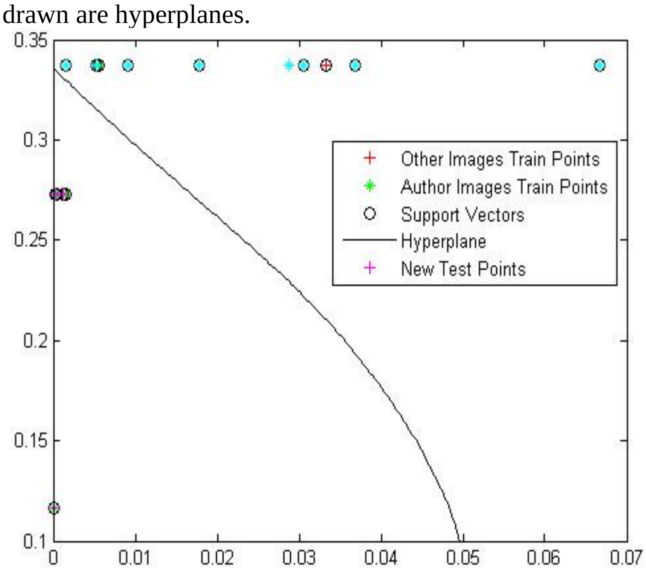

Figure 11: GLCM with SVM-RBF Kernel function for natural images

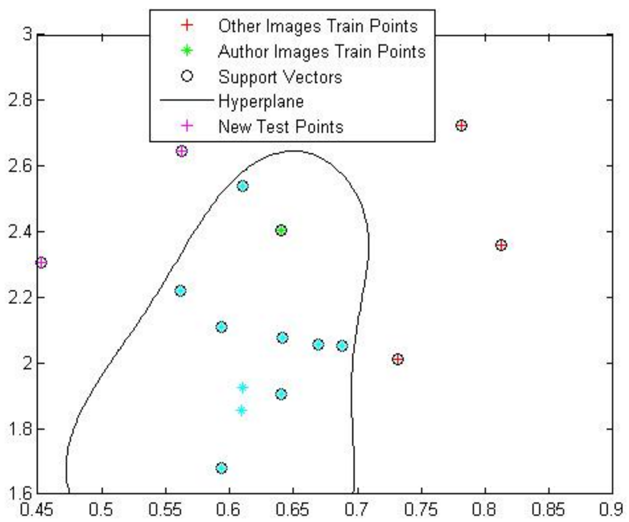

Figure 12: PCA with SVM-RBF Kernel function for natural images

Table I: Combined results for Natural Images using PCA and GLCM

Feature Extraction Technique	Kernel Function	Accuracy (%)	Execution Time (Seconds)
PCA	RBF	75	16.5058
	Linear	58.33	17.4874
	Quadratic	83.33	16.6581
GLCM	RBF	91.667	4.3987
	Linear	91.667	4.4024
	Quadratic	91.667	4.3992

Table II: Combined results for Brain MRI Images using PCA and GLCM

Feature Extraction Technique	Kernel Function	Accuracy (%)	Execution Time (Seconds)
PCA	RBF	57.69	30.4749
	Linear	50	31.7292
	Quadratic	73.07	31.0520
GLCM	RBF	100	10.3172
	Linear	96.15	10.4848
	Quadratic	100	10.3837

VII. CONCLUSION

In this paper, automatic classification technique is developed to classify between normal and abnormal images as well as to classify between author's images and other natural images. Features are extracted from images using feature extraction algorithms GLCM and PCA and classified further using SVM classifier. For brain MRI images, features extracted by GLCM classified with SVM using RBF kernel function gives 100% accuracy along with low execution time of 10.3172 seconds whereas with PCA with same kernel function gives accuracy of 57.69%. GLCM using SVM for linear and quadratic kernel functions gives classification accuracy of 96.15% and 100% with execution time of 10.4848 seconds and 10.3837 seconds respectively. Such maximum accuracy is not possible with features by PCA. This method helps to reduce the burden of medical practitioners when large amount of data is available. But the final decision is made after consultation with specialist.

Similarly for natural images; features extracted by GLCM gives maximum accuracy of 91.67% with all kernel functions but there is difference in their execution times. Execution time for SVM-RBF is 4.3987 seconds. PCA with same kernel function gives maximum accuracy of 75%.

REFERENCES

[1] EL- Sayed, EL- Dahshan, Abdul- Badeeh M. Salem, Tamer H.Yousin, "A Hybrid Technique for Automatic MRI Brain Images Classification". Studia Univ. Babes.Bolyai, Volume LIV, Number 1, 2009.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

[2] SaharJafarpour, Zahra Sedhgi, Mehdi ChehelAmirani, "A Robust Brain MRI Classification with GLCM Features", International Journal of Computer Applications (0975-8887) Volume 37- No.12, January 2012.
[3] Jin Huang, Jingjing Lu, Charles X.Ling, "Comparing Naive Bayes, Decision Trees, and SVM with AUC and Accuracy" IEEE International Conference on Data Mining, 2003.
[4] C.Cortes, V.Vapnik, in Machine Learning, PP. 273-297 (1995).
[5] El-Sayed Abdal Wahed, Abrahim Al Emam, Ami Badr, "Feature selection for Cancer Classification: An SVM based approach", International Journal of Computer applications (0975-8887), Volume 46- No.8, May 2012-07-23.
[6] F. Keyvanfard, M.A. Shoorehdler, M. Teshndab, "Feature Selection and Classification of Breast Cancer on Dynamic Magnetic Resonance Imaging using ANN and SVM" American Journal of Biomedical Engineering: 2011 ;1(1): 20-25.
[7] N.Christiani, J.Shawe-Taylor, "An Introduction to Support Vector Machines; and other kernel-based learning methods. Cambridge university Press, New York, NY, USA (2000).
[8] Noramalina Abdulla uni Kalthum Ngah, Shalihatun Azlin Aziz, "Image classification of Brain MRI using Support Vector Machine", 978-1-61284-896-9/11/$26.00,2011 IEEE.
[9] S.N. Deepa, B. Aruna Devi, "Second order Sequential Minimal Optimization for Brain tumor classification" , European Journal of scientific research, ISSN 1450-216X Vol. 64 No. 3 (2011), pp. 377-386.
[10] Ahmad Kharbat, Karim Gasmi, Mohamed Ben Messaond, Nacera Benamrane , Mohamed Abid , "A Hybrid Approach for Automatic Classification of Brain MRI using Genetic Algorithm and Support Vector Machine " , Journal of Leonardo Sciences , Issue 17, July-December 2012, P 71-82, ISSN 1583-0233.

Daljit Singh is pursuing his M-Tech (regular) thesis in Biomedical Image Processing. He had completed his Bachelor of technology in Electronics and Communication Engineering in 2010. He had published 3 national papers and 1 international journal. He has attended several conferences on microcontrollers and image processing. He had completed two projects using PIC18f microcontroller. His areas of interest are Digital Image Processing, Signal Processing and Microcontrollers.

Kamaljeet Kaur is pursuing her M-Tech (regular) thesis in Biomedical Image Processing. She had completed her B-Tech in Electronics and Communication Engineering in 2009.She has attended 2 national conferences on image processing. She had published 3 national papers and 1 international journal. Her areas of interest are Digital Image Processing and Microcontrollers.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)