International Journal of Electrical and Electronics Engineers
Vol. No. 9, Issue No. 01, January-June 2017
ISSN (O) 2321-2055
ISSN (P) 2321-2045

DETECTION AND CLASSIFICATION OF ABNORMALITIES IN BRAIN MR IMAGES USING SUPPORT VECTOR MACHINE WITH DIFFERENT KERNELS

Vishnumurthy T D[1], H S Mohana[2], Vaibhav A Meshram[3]
Pramod Kammar[4]

[1]*Research Scholar,Dept. of E&CE, Jain University Bengaluru, (India)*
[2,4]*Dept. of E& I, Malnad College of Engineering, Hassan, (India)*
[3]*Dept. of ECE, Dayananda Sagar University, Bengaluru, (India)*

Abstract

Magnetic resonance imaging (MRI) is a popular, powerful and versatile imaging technique that has been employed for studying brain images in neuroscience research. Classification is an important part in order to distinguish between normal patients and those who have the possibility of abnormality or tumor. The proposed method consists of two stages: feature extraction and classification. In the former stage, features are extracted from GLCM technique. In the latter stage, extracted features are used as inputs to SVM classifier. The SVM classifies the brain images as normal or abnormal. The SVM technique is applied along with various kernels and the performance of each kernel is evaluated for standard parameters. The classification accuracy ~97% is obtained using SVM with kernel functions Gaussian Radial basic function and Multilayer Perceptron kernel for the selected database.

Keywords: GLCM, Feature Extraction, MRI, SVM classifier, SVM Kernel.

I. INTRODUCTION

The MRI is the most powerful imaging technique that uses magnetic fields, RF pulses and computer to capture images of the human brain. In the recent years, there has been a significant development in the field of MRI technology which contributed significantly for clinical neuroimaging and analysis. MRI is used mainly to determine different brain pathologies and is preferred over Computed Tomography (CT) scan for diagnosing cancers and tumors. MRI detects signals emitted from normal and abnormal tissue, providing clear images of most tumors. It is very effective in identifying stroke, generative diseases like Alzheimer's, infectious diseases, and dementia due to the contrast scans obtained. The classifications of brain MRI data as normal and abnormal are important to prune the normal patient and to consider only those who have the possibility of having abnormalities or tumor.

In the past few years, different brain tumor detection algorithms have been developed and used. The automatic segmentation is very difficult and challenging technique that needs to be solved fully and satisfactorily. The main aim of the automated segmentation is to detect and identify the tumor from brain MR images. It takes into account the statistical features and/or structural features of the brain structure to represent it by significant feature points. In the literature review, there are many techniques and algorithms which were developed and implemented for image segmentation. They

International Journal of Electrical and Electronics Engineers
Vol. No. 9, Issue No. 01, January-June 2017
ISSN (O) 2321-2055
ISSN (P) 2321-2045

are histogram based, edge-based, artificial neural network based methods, region-based methods (region growing, splitting and merging), physical model based approaches, and clustering methods (K-means clustering, Fuzzy C-means clustering, Mean Shift and Expectation Maximization) [1]-[4], in spite of it, still there is a necessity to identify and develop an efficient and fast technique for medical image segmentation, because all these techniques have their own advantages and limitations with reference to their suitability, applicability, performance and computational time for the selected database. In prior to the classification, Morphological segmentation technique is implemented for detection of tumor [5].

The approach 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 etc. Another category is unsupervised learning for data clustering such as K-means Clustering, Self Organizing Map (SOM). In this paper, SVM classification is used with different kernel function for better accuracy and performance over other supervised learning techniques. ANN, PNN and KNN have their own limitations and advantages. The factors such as size of the database is increased, number of attributes and the presence of the artifacts, high computing cost which consumes high CPU and physical memory usage [6].

II. METHODOLOGY

The proposed method is described in the flowchart shown in figure1. In the present work, emphasis is given on feature extraction and SVM classifier.

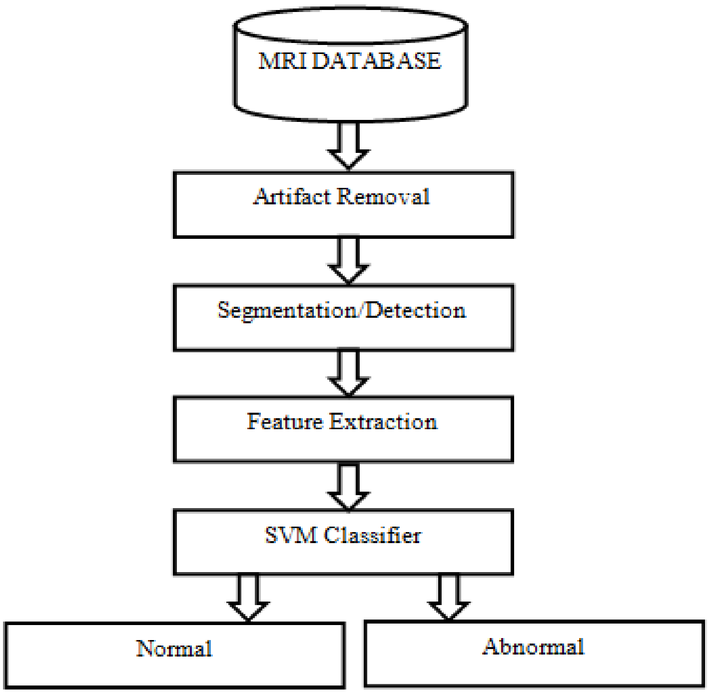

Fig 1. System overview.

Grey-Level Co-occurrence matrix (GLCM) is used for feature extraction and Support Vector Machine (SVM) for classification with different kernel functions. SVM is used to classify the extracted features into further two classes. For medical images; it classifies between normal and abnormal images. 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 automatically by SVMclassifier.

International Journal of Electrical and Electronics Engineers
Vol. No. 9, Issue No. 01, January-June 2017
ISSN (O) 2321-2055
ISSN (P) 2321-2045

2.1 MRI Database : Brain MR Image database is created by collecting the brain images from the radiologists of medical colleges and research centers as well as from open source. Database used for brain MR images consists of total 299 images; out of which 142 are normal and 157 are abnormal brain images. All MR brain images are taken with different views but with same resolution. The sample MR brain images with artifact taken in different view for testing are shown in figure 2.

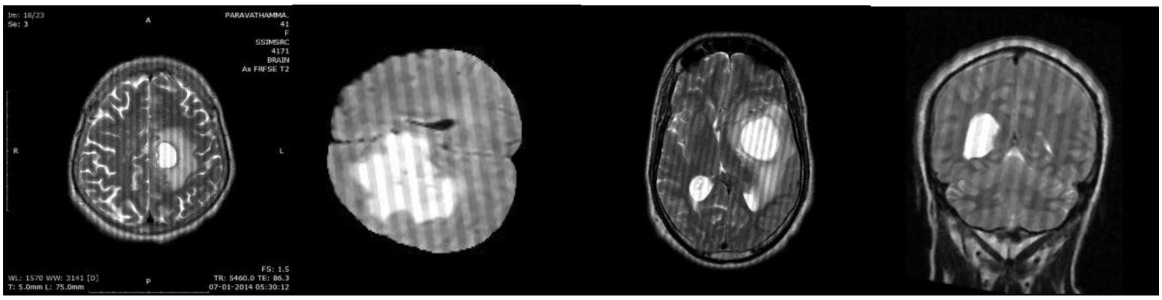

Fig 2. Sample MR Brain images.

2.2 Artifact Removal : Artifacts are seldom appear in MRI. The presence of artifacts can degrade the image quality, may confuse the pathology and reduce the accuracy of image segmentation and subsequently classification. Hence skull removal and herringbone artifact removal are considered in this work.. Herringbone artifact removal techniques such as combined FFT and Canny Edge Detector [7] and Combined Wavelet and FFT Based Filtering Technique are applied [8].

2.3 Tumor Detection : Automatic segmentation and detection in image processing is referred as the process of isolating an image into mutually exclusive regions. In processing of any medical images, detection is considered as an essential and crucial one because of the diverse image content, artifacts, and disordered objects; non-uniform object texture and other issues. Detection by morphological operations is implemented and result is compared with other segmentation techniques like Expectation maximization and Fuzzy C-Means with reference to performance measures and processing time. The performance measures such as Jaccard Distance, Dice coefficient, False positive ratio, and False negative ratio are used for comparison [5] [10].

2.4 Feature extraction : The feature is known as a function of one or more measurements, each specifies some quantifiable property of an object, and hence computed features quantify some meaningful characteristics of the object. All features can be classified into low-level features and high-level features. Low-level features are extracted directly from the original images; however high-level feature extraction must be based on low-level features. Texture is a surface property. It is characterized by the spatial distribution of gray levels in a neighbourhood. Since texture shows its characteristics both by pixel co-ordinates and pixel values, there are many approaches used for texture classification. The image texture depends on the scale or resolution at which it is displayed. A texture with specific characteristics in a sufficiently small scale could become a uniform texture if it is displayed at a larger scale.

Gray-Level Co-occurrence Matrix (GLCM) is a statistical method used for feature selection. The GLCM tabulation is given by Haralick, Shanmugam and Dinstein who familiarised the gray level co-occurrence matrix [9]. A gray-level co-occurrence matrix is fundamentally a two-dimensional histogram. The GLCM method studies the spatial connection between pixels of dissimilar gray levels.

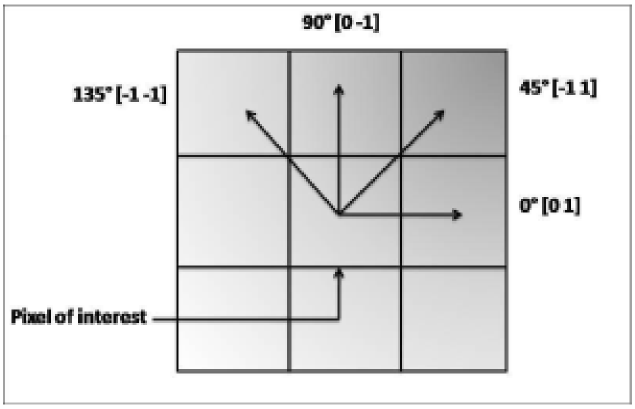

Fig 3. Direction for generation of GLCM.

This system uses MRI images which can be decomposed into patterns with systematic textures. So co-occurrence matrices represent these regular textures. To do so, the co-occurrence matrices in different phases of 0°, 45°, 90°, and 135° are utilised.

2.5 Texture Features

Texture is a feature utilised in the analysis and description of images and is described by a set of local statistical properties of pixel intensities. When the GLCM matrix is produced, the textures feature could be calculated. The second order statistics of an image can be obtained from GLCM which accounts for the spatial inter-dependency or co-occurrence of two pixels at specific relative positions. Co-occurrence matrices are calculated for the directions of 0, 45, 90, and 135 degrees. For each matrix, the fourteen features like Angular Second Moment or Energy, Contrast, Correlation, Sum of Squares or Variance, Inverse Difference Moment, Sum Average, Sum Variance, Sum Entropy, Entropy, Difference Variance, Difference Entropy, Information Measure of Correlation and Cluster Tendency are obtained for the synthesized placenta image. The homogeneity, contrast, entropy and energy are sensitive to the choice of the direction. The homogeneity and entropy supplies the indication on the dominancy values of the main diagonal on the basis of the frequencies of the problem. The energy supplies the information on the randomness of the spatial distribution.

Second order statistics account for the spatial or co-occurrence of two pixels at specific relative positions. Following nine features are employed in this work:

a) Contrast : Contrast calibrates the quantity of local changes in the picture. It reflects the sensitivity of the textures in relation to changes in the intensity. It returns the amount of intensity contrast between a pixel and its surroundings. For a constant image contrast value is zero. It is the amount of local variation present in an image. If local variation is large, the contrast feature also has consistently higher values comparatively. If the gray scale difference occurs continually, the texture becomes coarse and the contrast becomes large. If contrast value is small then textures becomes acute. Contrast can be calibrated as

$$\text{Contrast} = \sum_{n=0}^{Ng-1} n^2 \sum_{|i-j|=n} P_d(i,j) \qquad (1)$$

$\forall\, i \text{ and } j$, N_g is the total number of pixels, n is the each pixel value.

b) Correlation : This feature measures how correlated a pixel is to its neighbourhood. It is the measure of gray tone linear dependencies in the image. Feature values range from -1 to 1, these extremes indicating perfect negative and positive correlation respectively. μ_i and μ_j are the mean, σ_i σ_j are the standard deviations of $P_d(i)$ and $P_d(j)$, respectively. If the image has horizontal textures the correlation in the direction of 0° degree is often larger than those in other directions. It can be calculated as

$$\text{Correlation} = \frac{\sum_{i=1}^{Ng}\sum_{j=1}^{Ng}(i-\mu_i)(j-\mu_j)}{\sigma_i \sigma_j} \quad (2)$$

$\forall\, i \text{ and } j$.

c) Homogeneity: Homogeneity calibrates the similarity between pixels. Homogeneity is 'one' for diagonal gray level co-occurrence matrix also for constant image. Homogeneity turns large if local textures only have minimal changes. Homogeneity is given by

$$\text{Homogeneity} = \sum_{i=1}^{Ng}\sum_{j=1}^{Ng}\frac{P_d(i,j)}{1+|i-j|} \quad (3) \quad \forall\, i \text{ and } j.$$

d) Energy: Energy also named uniformity or angular second moment (ASM). The more homogeneous the image is, the larger the value. If the image is constant then energy will be 'one'. Energy is calculated as

$$\text{Energy} = \sum_{i=1}^{Ng}\sum_{j=1}^{Ng}P_d^2(i,j) \quad (4) \quad \forall\, i \text{ and } j.$$

e) Entropy: Randomness of intensity image is measured by entropy. Entropy is measured by

$$\text{Entropy} = \sum_{i=1}^{Ng}\sum_{j=1}^{Ng}P_d(i,j) * (-\log(P_d(i,j))) \quad (5) \quad \forall\, i \text{ and } j.$$

f) Dissimilarity: Dissimilarity measures the dissimilar between the pixels values and it is given by

$$\text{Dissimilarity} = \sum_{i=1}^{Ng}\sum_{j=1}^{Ng}P_d(i,j) * (|i-j|) \quad (6)$$

$\forall\, i \text{ and } j$.

g) Maximum probability: Maximum probability gives the maximum probability value amongst all the calculated probabilities. It is given by

$$\text{Maximum Probability} = max_{i,j}P_d(i,j) \quad (7) \quad \forall\, i \text{ and } j.$$

h) Variance: It measures the gray tone variance. It is given by equation

$$\text{Var} = \sum_{r=0}^{Ng-1}(r-m_w) * p(r) \quad (8)$$

where m_w is mean of pixel value $p(r)$.

i) Standard deviation: Measures the deviation of values from the mean value. It is given equation by

$$\text{STD} = sqrt\left(\sum_{r=0}^{Ng-1}(r-m_w) * p(r)\right) \quad (9)$$

F. SVM Classifier: 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. The Support Vector Machine (SVM) was first proposed by Vapnik. The SVM's generally are capable of delivering higher performance in terms of classification accuracy than other data classification algorithms [11]-[14].

- **Separable data:** It is a binary classifier based on supervised learning which gives better performance and classifies between two classes by constructing a best hyper plane in high-dimensional feature space. The best hyper plane for an SVM means the one with the largest margin between the two classes. Margin means the maximal width of the slab parallel to the hyper plane that has no interior data points. The support vectors are the data points that are closest to the separating hyper plane; these points are on the boundary of the slab. The figure 4 illustrates these definitions, with + indicating data points of type 1, and – indicating data points of type –1.

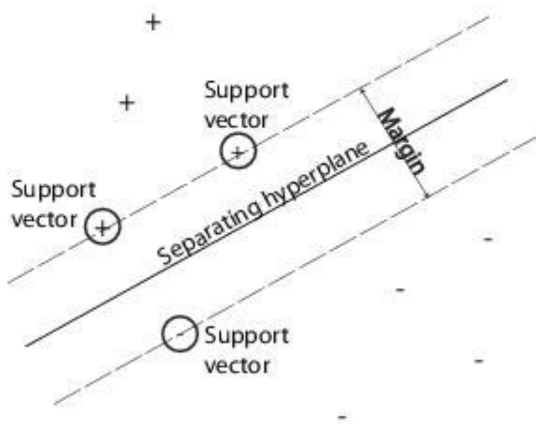

Fig:4. Support Vector Machine

The data for training is a set of points (vectors) x_j along with their categories y_j. For some dimension d, the $x_j \in \mathrm{R}^d$, and the $y_j = \pm 1$. The equation of a hyper plane is

$$f(x) = \mathrm{x}^T \beta + b = 0 \qquad (10)$$

where $\beta \in \mathrm{R}^d$ and b is a real number.

The following problem defines the best separating hyper plane (i.e., the decision boundary). Find β and b that minimize $\|\beta\|$ such that for all data points (x_j, y_j), $y_j f(x_j) \geq 1$.

The support vectors are the x_j on the boundary, those for which $y_j f(x_j) = 1$.

For mathematical convenience, the problem is usually given as the equivalent problem of minimizing $\|\beta\|$. This is a quadratic programming problem. The optimal solution ($\hat{\beta}, \hat{b}$) enables classification of a vector z as follows:

$$class(z) = \mathrm{sign}(\mathrm{z}^T \hat{\beta}, \hat{b}) = sign(\hat{f}(z)). \qquad (11)$$

$\hat{f}(z)$ is the classification score and represents the distance z is from the decision boundary.

- **Non separable Data:** Data might not allow for a separating hyper plane. In that case, SVM can use a soft margin, meaning a hyper plane that separates many, but not all data points.

- **Nonlinear transformation with kernels:** In some cases data points or data sets are always not separated by drawing a straight line between two classes. In that case, Kernel functions are used with SVM classifier. Kernel function provides

International Journal of Electrical and Electronics Engineers
Vol. No. 9, Issue No. 01, January-June 2017
ISSN (O) 2321-2055
ISSN (P) 2321-2045

the bridge between 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:

- **Linear kernel:** The linear kernel is the simplest kernel. It is defined by

$$K(x_i, x_j) = x_i^T x_i \quad (12)$$

- **Polynomial kernel:** The polynomial kernel is suited for problems with normalized training data. It is given by

$$K(x_i, x_j) = (\alpha x_i^T x_i + C)^d \quad (13)$$

where α is the slope, C is a constant and d is the polynomial degree.

- **Gaussian Radial Based Function kernel:** The GRBF kernel is defined as

$$K(x_i, x_j) = exp\left(\frac{\|x_i, x_j\|^2}{2\sigma^2}\right) \quad (14)$$

where σ is an adjustable parameter. It is a non-linear kernel and is very sensitive to noise.

- **Exponential Radial Based Function kernel:** The ERBF kernel is close to GRBF with only the square of the norm removed. It is defined as

$$K(x_i, x_j) = exp\left(\frac{\|x_i, x_j\|}{2\sigma^2}\right) \quad (15)$$

- **Multilayer Perceptron kernel:** Multilayer Perceptron kernel is also called as Hyperbolic Tangent kernel or sigmoid kernel. It is defined as

$$K(x_i, x_j) = tanh(\gamma x_i^T x_i + r) \quad (16)$$

where γ is the slope and r is the intercept constant.

III. CLASSIFIER PERFORMANCE PARAMETERS

All classification results could have an error rate and on occasion will either fail to identify an abnormality, or identify an abnormality which is not present. It is common to describe this error rate by confusion matrix or error matrix. This will contain information about actual predicted classifications done by a classification system. The table 1 shows the confusion matrix or error matrix for a two class classifier.

Table1. Confusion Matrix or Error Matrix

		Predicted	
		Negative	Positive
Actual	Negative	TN	FP
	Positive	FN	TP

True Positive (TP): the classification result is positive in the presence of the clinical abnormality.

True Negative (TN): the classification result is negative in the absence of the clinical abnormality.

False Positive (FP): the classification result is positive in the absence of the clinical abnormality.

False Negative (FN): the classification result is negative in the presence of the clinical abnormality.

Confusion matrix or error matrix is used to visualize the performance of the classification. Several standard terminology and derivations from a confusion matrix are used for more detailed analysis than mere proportion of correct guesses (accuracy) [12], [15].

a) *Sensitivity, recall, hit rate, or true positive rate (TPR or SE):* Measures the proportion of positives that are correctly identified as such (i.e. the percentage of sick people who are correctly identified as having the condition). It is calculated by

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} * 100 \tag{17}$$

b) *Specificity or true negative rate (TNR or SP):* Measures the proportion of negatives that are correctly identified as such (i.e., the percentage of healthy people who are correctly identified as not having the condition). It is calculated by

$$\mathrm{TNR} = \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}} * 100 \tag{18}$$

c) *Accuracy (ACC):* is the proportion of the total number of predictions that were correct. It is determined by

$$\mathrm{ACC} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{FP}+\mathrm{TN}+\mathrm{FN}} * 100 \tag{19}$$

d) *Precision or positive predictive value (PPV):* Is the proportions of the positive results in classification that are true positive. It is given by

$$\mathrm{PPV} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} * 100 \tag{20}$$

e) *Negative predictive value (NPV):* Is the proportions of the negative results in classification that are true negative. It is given by

$$\mathrm{NPV} = \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FN}} * 100 \tag{21}$$

f) *Miss rate or false negative rate (FNR):* Erroneous outcome of the result of classification which actually present. It is given by

$$\mathrm{FNR} = \frac{\mathrm{FN}}{\mathrm{FN}+\mathrm{TP}} * 100 = 1 - \mathrm{TPR} \tag{22}$$

g) *Fall-out or false positive rate (FPR):* Erroneous outcome of the result of classification which actually not present. It is given by

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP}+\mathrm{TN}} * 100 = 1 - \mathrm{TNR} \tag{23}$$

h) *False discovery rate (FDR):* Is the proportions of the positive results in classification that are false positive. It is given by

$$\mathrm{FDR} = \frac{\mathrm{FP}}{\mathrm{FP}+\mathrm{TP}} * 100 = 1 - \mathrm{PPV} \tag{24}$$

International Journal of Electrical and Electronics Engineers
Vol. No. 9, Issue No. 01, January-June 2017
ISSN (O) 2321-2055
ISSN (P) 2321-2045

i) *False Omission rate (FOR):*

$$FOR = 1 - NPV \tag{25}$$

j) *F1 score:* is the harmonic mean of precision and sensitivity. It is given by

$$F1 = 2 * \frac{PPV+TPR}{PPV+TPR} * 100 = \frac{2*TP}{2*TP+FP+FN} * 100 \tag{26}$$

k) *Matthews's correlation coefficient (MCC):* is the measure of the quality of the classifications.

$$MCC = \frac{TP*TN - FP*FN}{sqrt\,((TP+FP)(TP+FN)(TN+FP)(TN+FN))} * 100 \tag{27}$$

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The input to the feature extraction algorithm is the MR brain images. The pattern vectors (features) extracted from the images is given as input to the SVM classifier. Large database are required for the classifier to perform the classification. In this system samples of 299 MR brain images are collected, 157 images are normal brain images and remaining 142 images are abnormal brain images. Features are classified by SVM with other kernel functions.

In predictive analysis, a confusion matrix or error matrix is a table with two rows and two columns that reports the number of false positives (FP), false negatives (FN), true positives (TP), and true negatives (TN). Using this table detailed analysis can be done instead of doing just a proportion of correct guesses that is accuracy. Accuracy is not a reliable metric for the real performance of a classifier, because it will yield misleading results if the data set is unbalanced (that is, when the number of samples in different classes vary greatly). Table 2 shows the confusion matrix or error matrix generated when different kernel functions is used in SVM classifier.

Table 2. Generated confusion matrix of various kernel function in SVM Classifier.

Kernels used	Total No of Images	Performance Metric			
		TP	FP	FN	TN
Linear	299	135	10	7	147
Quadratic	299	135	8	7	149
Polynomial	299	137	7	5	150
GRBF	299	139	3	7	150
MLP	299	136	4	6	153

Commonly used evaluation measures are Sensitivity, Specificity and Accuracy. Fallout or False Positive Rate (FPR) are the proportion of Real Negatives that occur as Predicted Positive (ring-ins); Miss Rate or False Negative Rate (FNR) are the proportion of Real Positives that are Predicted Negatives (false-drops).

Table 3. Performance analysis of various kernel functions in SVM Classifier.

Performance Parameters	Linear	Quadratic	Polynomial	GRBF	MLP
SE	95.08	95.08	96.48	**97.88**	95.77
SP	93.63	94.90	95.54	95.54	**97.45**
ACC	94.31	94.98	95.98	**96.65**	**96.65**

International Journal of Electrical and Electronics Engineers
Vol. No. 9, Issue No. 01, January-June 2017
ISSN (O) 2321-2055
ISSN (P) 2321-2045

PPV	93.11	94.41	95.14	95.21	**97.15**
NPV	95.45	95.51	96.77	**98.03**	96.22
FNR	4.92	4.92	3.52	**2.12**	4.23
FPR	6.37	5.1	4.46	4.46	**2.55**
FDR	6.89	5.59	4.86	4.79	**2.85**
FOR	4.55	4.49	3.23	**1.97**	3.78
F1-Score	94.07	94.73	95.8	**96.52**	**96.45**
MCC	88.62	89.94	91.96	**93.33**	**93.29**

Classifier parameters are evaluated for all SVM kernel functions. The linear SVM classifier yielded low results when compared with all other kernel functions. Sensitivity says that, at what rate abnormal MR brain images are correctly identified as abnormal images. If sensitivity is taken into account then SVM with GRBF and polynomial kernel function have better result. Specificity says that, at what rate normal MR brain images are correctly classified as normal images. It is observed in the table 3 that SVM with MLP is better result than all other. Accuracy is overall classification of the classifier. GRBF and MLP have same accuracy but they differ in classifying the normal and abnormal MR brain images. Classifier how precisely classifies the positive results. PPV is more in MLP classifier. Negative predictive value tells how precisely the classifier classifies the negative results. FNR is erroneous outcome of the negative result that actually should have been present. FPR is erroneous outcome of the positive result that actually should have been the absent. FNR value will be zero if sensitivity senses the presence of abnormal brain images to 100% and FPR value will be zero if specificity specifies absence of the normal brain images. FNR is approaching zero in GRBF and FPR approaching zero in MLP. FDR is false discovery rate that will tell the proportionality of the positive results in classification that are false positive. It is very less in MLP which means classifier is good at classifying the normal brain images. FOR is false omission rate will be proportional to the negative results in classification that are false negative. It is very less in GRBF which means GRBF is good at classifying the abnormal brain images. F1-socre is harmonic mean of the precision and sensitivity and better result is obtained in GRBF. Quality of the classification is measured using MCC and better result is obtained in the GRBF.

V. CONCLUSION

The SVM classifier for classification of MR brain images varies its classification efficiency and accuracy with different kernel functions. The sensitivity, specificity and accuracy are also improved. 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. As it is said earlier, performance of the classification not merely relies on the accuracy but also depends on the other aspects and parameters. In this paper, SVM classifier with GRBF and MLP kernel functions gives better results for the selected image database. GRBF gives better classification result in classifying the abnormal brain images whereas MLP gives better classification result in classifying the normal brain images.

REFERENCES

[1]. El-Sayed A. El-Dahshan, Heba M. Mohsen, Kenneth Revett, Abdel-Badeeh M. Salem,“Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm”, ELSEVIER, Expert Systems with Applications, 41 pp5526–5545, 2014.

[2]. Orlando J.Tobias, and Rui Scara,“Image Segmentation by Histogram Thresholding Using Fuzzy sets” IEEE Transactions on Image Processing, Vol. 11, No. 12, December 2002.

[3]. Natarajan P, Krishnan N, Natasha Sandeep Kenkre, Shraiya Nancy, Bhuvanesh Pratap Singh, “Tumor Detection using threshold operation in MRI Brain images”, 978-1-4673-1344-5/12$31.00@2012 IEEE.

[4]. Y. Zhang, M. Brady, and S. Smith, “Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm,” IEEE Transactions on Medical Imaging, vol. 20, no. 1, pp. 45–57, 2001.

[5]. Vishnumurthy T D, Mohana H S and Vaibhav A Meshram, “Automatic Segmentation of Brain MRI images and Tumor Detection using Morphological Techniques” , 2016 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT). 2016.

[6]. Hari Babu Nandpuru, Dr. S. S. Salankara and Prof. V. R. Bora, “MRI Brain Cancer Classification Using Support Vector Machine”, 2014 IEEE Students' Conference on Electrical, Electronics and Computer Science, 978-1-4799-2526-1/14/$31.00 ©2014.

[7]. Vishnumurthy T D, Mohana H S, Vaibhav A Meshram, Pramod Kammar, “A Novel Algorithm for Removal of Herringbone Artifact in Brain MR Images Using FFT and Canny Edge Detector”, American Journal of Engineering Research, Vol.5, Issue.05, pp. 184-189, June 2016, e-ISSN: 2320-0847

[8]. Vishnumurthy T D, Mohana H S, Vaibhav A Meshram, Pramod Kammar, “Suppression of Herringbone Artifact in MR Images of Brain Using Combined Wavelet and FFT Based Filtering Technique” International Journal of Computer Sciences and Engineering, Vol.4, Issue-02, pp.66-71, February 2016, E-ISSN:2347-2693

[9]. Haralick, R., Shanmugam, K., Dinstein, I,”Textural features for image classification”,IEEE Trans. on Systems Man and Cybern 3(6), 610–621 (1973).

[10]. Zhang, Y., Gerbrands, J., “Segmentation Evaluation Using Ultimate Measurament Accuracy. Image Processing Algorithms and Techniques”, III 1657, 449–460, 1996.

[11]. Xiangdong Liu,Meilin Zhu,Zhaoqian Chen,Shifu Chen,“ Support Vector Machine and its Applications in Pattern Recognition”, [J]. Computer Science, 2003, 30(6):113-117.

[12]. Ting, Kai Ming,“Encyclopedia of machine learning”. Springer. ISBN 978-0-387-30164-8. (2011)

[13]. Hastie, T., R. Tibshirani, and J. Friedman. The Elements of Statistical Learning, second edition. Springer, New York, 2008.

[14]. Christianini, N., and J. Shawe-Taylor. An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press, Cambridge, UK, 2000.

[15]. Powers, David M W (2011). "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation" (PDF). Journal of Machine Learning Technologies. 2 (1): 37–63.