Grenze International Journal of Engineering and Technology, July 2017

An MRI Brain Image Segmentation and Tumor Detection using SOM-Clustering and PSVM Classifier

D. Rammurthy [1] and Dr. Mahesh P. K [2]
[1]Navodaya Institute of Technology/Dept. of ECE, Raichur, India
Email: rammurthy.nit.ece@gmail.com
[2]Don Bosco Institute of Technology/Dept. of ECE, Bangalore, India
Email: mahesh24pk@gmail.com

***Abstract*—In recent days, image processing is widely used in diagnosis of disease such as brain tumor, Cancer, Diabetes etc. Brain tumor is one such dangerous disease and currently moreover 600,000 people have this type of disease. Image segmentation is an important technique highly used to extract the suspicious parts from medical images such as MRI, CT scan, and Mammography etc. With this motivation in this work, SOM clustering is proposed for MRI brain image segmentation. Before the segmentation the Histogram Equalization is utilized for feature extraction which will improve the segmentation accuracy. After the segmentation process, the feature extraction using Gray Level Co-occurrence Matrix is utilized which avoids the formation of misclustered regions. The Principle Component Analysis (PCA) method is used for the feature selection to improve the classifier accuracy. An effective classifier Proximal Support Vector Machines (PSVM) is used to automatically detect the tumor from MRI brain image. This method is faster and computationally more efficient than the existing method SVM.**

***Index Terms*— Image segmentation, MRI, Proximal Support Vector Machines, Self-organizing maps, Feature extraction.**

I. INTRODUCTION

The Magnetic Resonance Imaging (MRI) is a widely used medical imaging technique which provides detailed information of the internal tissue constitutions of the image. In disease diagnosis, the high resolution and non-invasive MR images have a vital control thus the image segmentation has more importance. The segmentation process used to detach the soft brain tissues like White Matter (WM), Gray Matter (GM) and Cerebral Spinal Fluid (CSF) etc in the form of neuroanatomical structures surrounded by the medical images and these regions are considered as pathological tissues. There are two types of segmentation techniques existing such as manual and automatic segmentation. Though the manual segmentation technique depends on experience or expert knowledge of human, it utilizes more time for tiring practice which significantly reduces the computational efficiency. On the other side, the automatic segmentation highly utilizes the histogram based which is fully based on the intensity of pixels. But, the results are not based on the spatial information and it is not effective when the noise occurs. Clustering is widely used for segmentation process by

Grenze ID: 01.GIJET.3.3.3

supervised and unsupervised learning [1] and is most popular too. The supervised learning based segmentation is about the apriori knowledge and the features of image are extracted using unsupervised learning based technique such as K-means, Fuzzy C-Means (FCM) and ANN. K-means algorithm is a hard segmentation method which may allocates a pixel to a class or does not [2]. On the other hand, FCM utilizes a membership function with the intention that a pixel can belongs to several clusters having different degree [3]. ANN can modifies the replies in consistent with the environmental states and study from the knowledge. SOM is an unsupervised ANN that uses competitive learning algorithm. The segmentation of MR images is a hard task because the image encloses strength in homogeneity, noise and Partial Volume Effect (PVE). This made more intended and involuntary movement of the patients and gears. MR brain image segmentation separated into three major soft tissues. For example: WM, GM and Cerebrospinal Fluid (CSF). The accurate segmentation of brain tissues assists to examine the tissue volume, detection of tumors and measuring tumor volumes. Clustering techniques executes as a main role in image segmentation field. SOM [4, 5] is a clustering technique which is used for unsupervised segmentation of MR brain images. SOM maps high dimensional data to a low dimensional discrete lattice of neurons. For continuous segmentation of MR brain images SOM and knowledge based expert system is exercised in [6].

In this study urbanized two fully unsupervised segmentation methods for MR image segmentation using SOM. Initially, one is a fast and efficient method for segmentation that depends upon the histogram developed from the image. Next one depends on the inherent features extracted from the image and it is a robust scheme under noisy and bad intensity image conditions. The proposed work consists of two phase for effective MR image segmentation and classification. The First phase utilizes the information from the volume image histogram to construct feature vectors which is to be classified using SOM. MRI image is segmented by using SOM clustering which uses each SOM unit as a cluster. The Second phase takes the input as segmented images and computes feature extraction using GLCM. PSVM are trained by using vectors as an input which comprises the selected features. The paper is organized as follows: Section 2: explains the materials and methods used in this work where describes the image databases and pre-processing stage which is common to the two segmentation approaches and also presents a faster implementation of the method which uses information extracted from the image histogram for segmentation of the whole volume and; Section 3: contains the feature extraction and classification to detect the tumor. Section 4: depicts the experimental results obtained from the evaluation of the proposed methods using the IBSR database and discusses the main questions derived from them. Section 5: finally, conclusions are drawn in.

II. AN AUTOMATED MRI BRAIN IMAGE SEGMENTATION AND TUMOR DETECTION

This section comprises of three subsections which summarize the segmentation methods and the image databases which is used in the work to evaluate the proposed methods.The overall architecture diagram is given in Fig.1.

A. Databases

In order to examine the performance of our image segmentation approach in comparison with other existing methods, manual segmentation labelling of the processed databases is required. Internet Brain Segmentation Repository (IBSR) from the Massachusetts General Hospital [8] is suitable for this purpose, as it provides manually guided expert segmentation results along with a set of magnetic resonance images. Meanwhile, IBSR 1.0 provides 20 T1-weighted volumetric images and IBSR 2.0 set provides 18 T1- weighted volumetric images that have been spatially normalized and processed by the Center for Morphometric Analysis (CMA) at the Massachusetts General Hospital with the biasfield correction routines which have been applied already. On the other hand, the overlap comparison metric is provided for each volume while comparing different segmentation methods. Consequently, images from the IBSR 1.0 database were used to compute the average overlap metric. Moreover, an image set consisting of high resolution MR images from the Nuclear Medicine Service of the "Virgen de las Nieves" Hospital, Granada, Spain (VNH) are also used to evaluate the proposed segmentation algorithms.

B. Image preprocessing

As soon as MR image has been acquired, a pre-processing is performed to remove noise and clean-up the image background. Several algorithms have been constructed for the purpose of Brain tissue extraction from the undesired structure such as Brain Surface Extractor (BSE), Brain Extraction tool (BET), Minneapolis

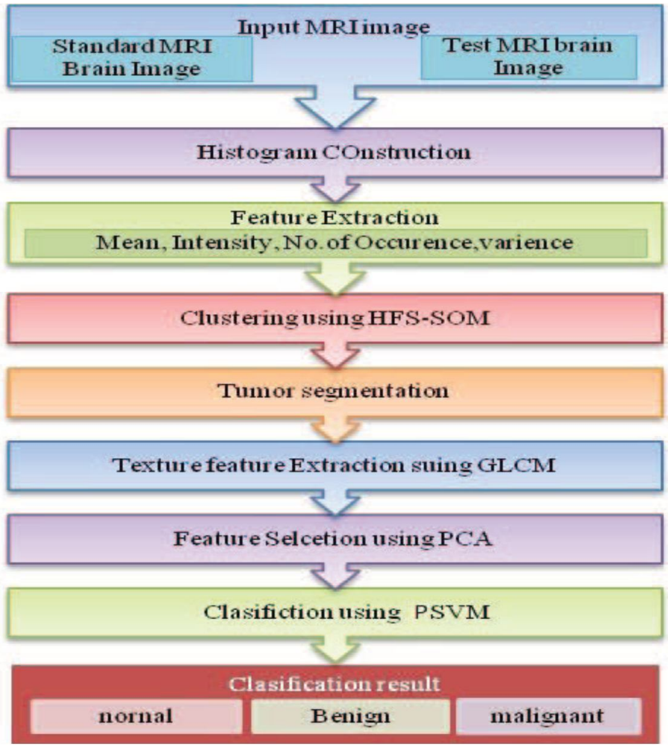


Fig. 1. Overall Block Diagram of Brain Tumor Diagnosis System

Consensus Strip (McStrip) or Hybrid Watershed Algorithm (HWA) [9]. These structures are already removed from the ISBR 1.0 database. However, images are provided by IBSR 2.0 which is distributed without the scalp/skull already removed. In the latter database, the brain has been extracted in the pre-processing stage using BET. In order to remove background noise uses a binary mask that built by detecting the greatest contiguous object in the image. After multiplying the binary mask (which contains 0 at the background voxels and 1 otherwise) by the original image, gets the background in black.

C. An Automated MRI Brain Image Segmentation

Histogram Equalization for Feature Extraction

Histogram equalization technique is used to increase the dynamic range of the histogram of an image. This technique [10] assigns the input image with their intensity values of pixels. In such a way, there is uniform distribution of intensities in the output image. This improves the contrast of an image. It is known that the good feature set will increase the classification accuracy result and it is very difficult too. As the tissues exists in brain are tedious to classify using only shape features or texture features or shape which defines the intensity level of information. Most of the works done in this area is utilized only the texture features or the shape and texture combination feature for MRI bran classification. By considering this fact and to improve the performance of the system color, texture and shape features which have been extracted in this work and considered for diagnosis. To achieve this goal Mean, intensity, number of occurrences and variance values are calculated to each MRI brain images. The best mean, intensity, variance and number of occurrences values are considered as extracted features results for input MRI image segmentation results. Finally, the gray scale image is converted to either black or white called binary image is done based some threshold value.

Clustering using SOM

The first step in the system is presented for isolating the tumor from the image. Since the tumor appears dark on the image, the detection of the edge of the tumor becomes confusing. Histogram Equalization is used to overcome this problem. The fast volume segmentation algorithm (HFSSOM) method is used for effective segmentation of brain image. The method is based on image histogram and the features are generated from the computed histogram. Tumor regions are effectively segmented by SOM clustering algorithms and thus the tumor portion from MRI image is detected. Fig.1 shows the segmentation process.

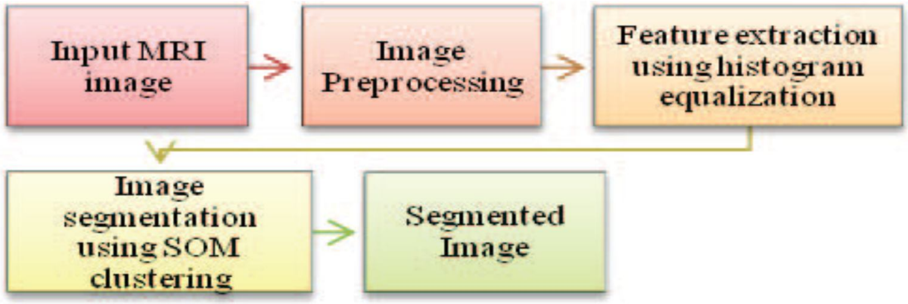


Fig. 2. Segmentation Method Process

D. An Automated MRI Brain Image Classification for Tumor Detection

Feature Extraction

In this module, Texture feature is defined by using Gray Level Co-occurrence Matrix (GLCM) [11]. Grayscale image from the segmentation phase is obtained from the color image, and then the image co-occurrence matrix is generated. As already known the features are the unique characteristics of in an image or object. To extract these features, various feature extraction techniques is proposed in such a way that the within-class similarity is maximized and between-class similarity is minimized. In this work, the GLCM feature extraction is utilized. The work involves extraction of the important features for brain tumor recognition. The features extracted gives the property of the texture, and are stored in knowledge base and further compared with the features of unknown sample image for classification. Thus, texture features are used to distinguish between normal and abnormal brain tumors. The important texture features are Autocorrelation, Contrast, Correlation, Cluster Prominence, Cluster shade, Dissimilarity, Energy, Entropy, Homogeneity, Maximum probability , Sum of squares, Sum average, Sum variance, Sum entropy, Difference variance, Difference entropy, Information measure of correlation, Inverse difference moment.

Contrast: It returns a measure of the intensity contrast between a pixel and its neighbor over the whole image. Contrast is 0 for a constant image. **Angular Second Moment (ASM):** It provides a strong measure of homogeneity. **Homogeneity (HOM):** It returns a value that measures the closeness of the distribution of elements in the Gray Level Cooccurrence Matrix (GLCM) to the GLCM diagonal. **Inverse Difference Moment (IDM):** It is the measure of local homogeneity.

Energy (E): Returns the sum of squared elements in the GLCM. Energy is 1 for a constant image.

Entropy (EN): It is a measure of randomness.

Variance (VAR): It is the measure that tells about how much the gray level are varying from the mean.

Feature Reduction using PCA

In feature reduction stage, which have been applied Principal Component Analysis (PCA) in order to reduce dimensionality of data to get most favorable features from entire data set [12]. PCA converts input feature space to high dimensional feature space wherever they are linearly distinguishable. The reduced Principal Components are then sorted in ascending order. The reduced matrix of PCA features has been arranged as in (1);

$$PC1 \geq PC2 \geq PC3 \ldots\ldots\ldots \qquad (1)$$

In (1), PC stands for principle component and N is the number of features for an image. Left side of the matrix contains most significant features and right side of the matrix has least significant features after PCA calculation. Least important features are ineffective which hold very fewer information and have no impact on precision. Features on the left side hold more important information because left side of the matrix having very high variation. The aim of the research is to reduce classifiers calculation and to make classification proficient, which has to choose smallest amount of features that can give best precision. Finally, PCA features have selected first L columns of matrix M. Although, we have chosen first few columns of PCA reduced feature that have high variations.

$$PC1 \geq PC2 \geq PC3 \dots\dots\dots\dots \geq PCL \qquad (2)$$

Where L is the number of columns in above equation. Detail of PCA method is given in [12]. Finally, PSVM classification is performed to identify true candidates. III.

III. Experimental Results and Discussions

In this section the classification for normal and abnormal detection with segmentation results are discussed with real MRI brain images from IBSR database. The proposed method is implemented in MATLAB language. Image preprocessing followed by segmentation using SOM clustering produces an effective result than the manual segmentation and other existing clustering methods such as K-means, FCM etc. The segmentation result for MRI brain image is shown in Fig.3. The performance of the proposed HFS-SOM method is evaluated through using Tanimoto's index. This metric is widely used to evaluate the segmentation performance of themethods which is defined as where is the segmentation set and is the ground truth.

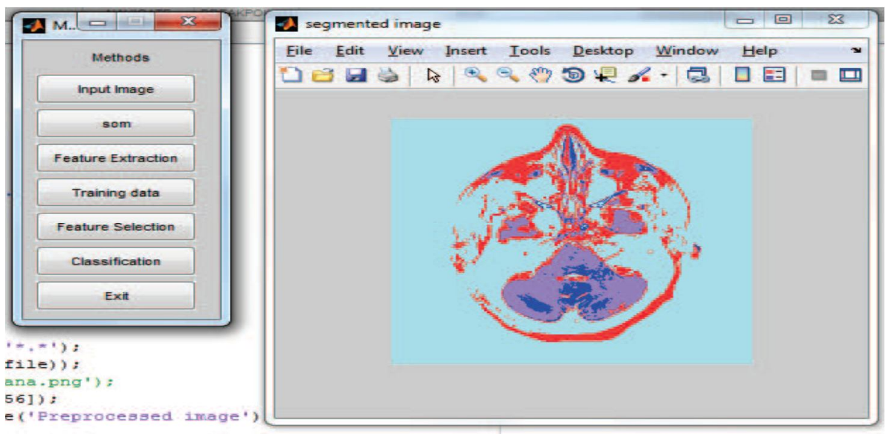
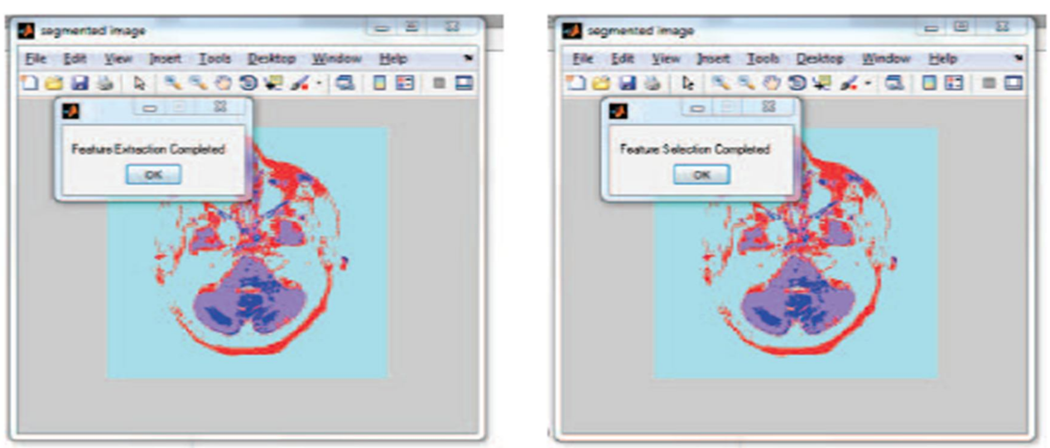


Fig. 3. Segmented Image

In the case of the tumor detection phase, GLCM is utilized as feature extraction method which can determined to be larger sufficient to extract textural features with no loss in resolution and the feature selection is done using Principle Component analysis. The feature extraction and selection results are shown in Fig.4.



(a) Feature Extraction (b) Feature Selection

Fig. 4. Feature Extraction and Selection Process

For the normal and abnormal detection stage, PSVM utilized as a classifier for brain images classification using MATLAB 7.8 with GLCM and PCA based features selection to improve the detection accuracy. The final result of classification process is shown in Fig.5.

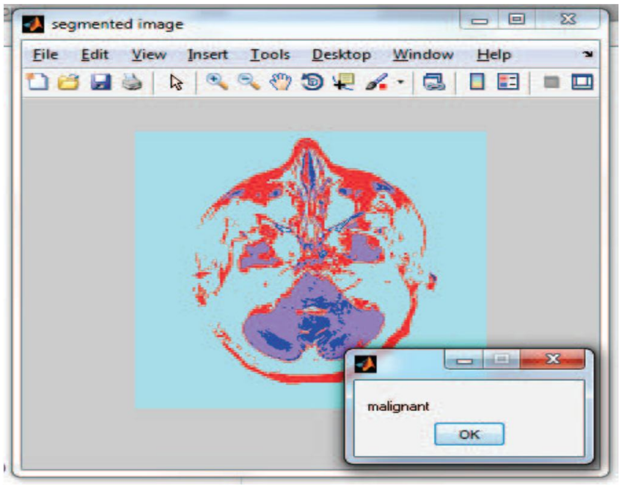


Fig. 5. The classification result of MRI brain image

A. Efficiency of the System

The efficiency or accuracy of the classifiers is evaluated based on the error rate results and is given as follows

$$FP = false\ positive\ pixels\ number\ /tumor\ s \quad (3)$$

$$FN = false\ negative\ pixel\ number\ /\ tumor\ s \quad (4)$$

$$Correct\ rate = FP + \quad (5)$$

The overall accuracy percentage details are shown in fig 6. And the comparative analysis is shown in Table 1.

TABLE I. ACCURACY RESULTS COMPARISON WITH OTHER TECHNIQUES

Techniques	Accuracy%
SVM-SOM	89
SVM-HFS-SOM	92
SVM- HFS-SOM-GLCM-PCA	94
PSVM- HFS-SOM-GLCM-PCA	97

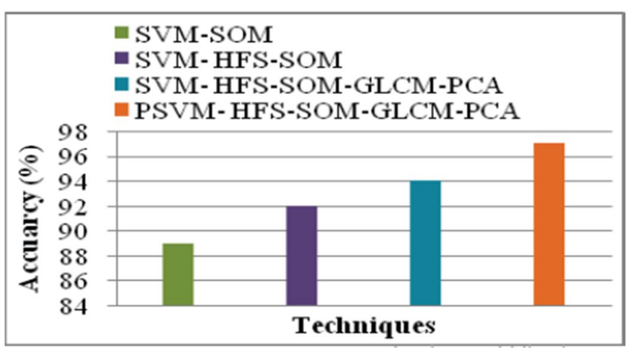


Fig. 6. Overall Accuracy Percentage Comparative Result

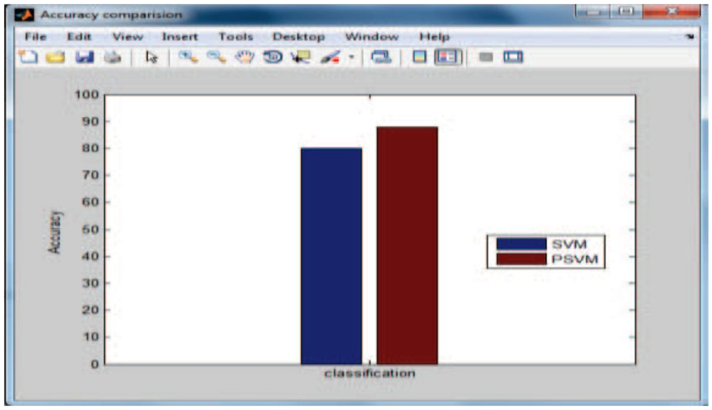


Fig. 7. Accuracy Comparison

The Fig.7 shows the accuracy comparison result of existing SVM and proposed PSVM algorithm. From the Fig.7, it is well known that the proposed system works better than existing SVM system with the high accuracy result of 92%. The values are tabulated in Table.II.

TABLE II. ACCURACY RESULTS COMPARISON

Technique	Accuracy%
SVM	82
PSVM	92

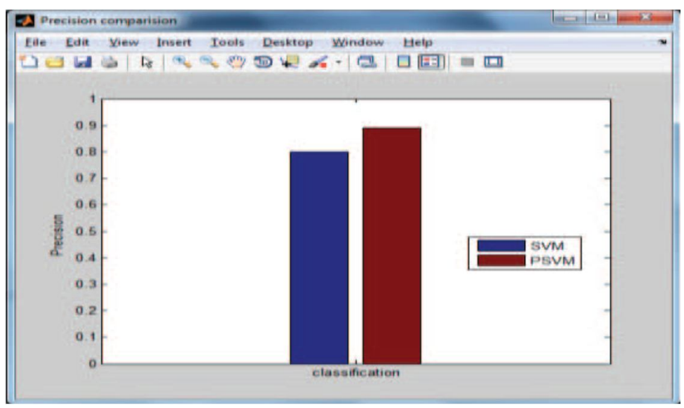


Fig. 8. Precision Comparison

The Fig.8 shows the precision comparison result of existing SVM and proposed PSVM algorithm. From the Fig.7, it is obvious that the proposed system has high precision rate of 0.93 which is 0.9 higher than the existing SVM algorithm. The reason is that the proposed system has less computational complexity than the SVM algorithm. The values are tabulated in Table.III.

TABLE III. PRECISION RESULTS COMPARISON

Technique	Precision Rate
SVM	0.82
PSVM	0.93

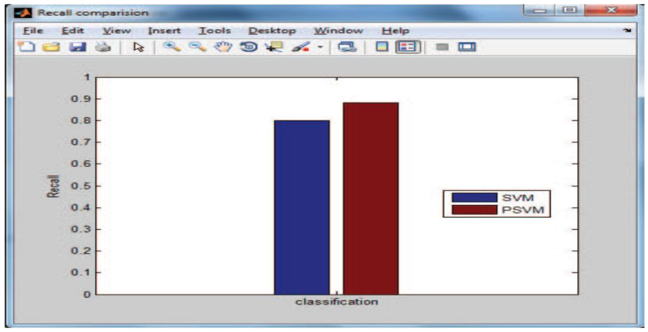


Fig. 9. Recall Comparison

The Fig.9 shows the recall comparison result of existing SVM and proposed PSVM algorithm. From the Fig.7, it is obvious that the proposed system has high recall rate as it has less execution time than the SVM algorithm. The values are tabulated in Table.IV.

TABLE IV. RECALL RESULTS COMPARISON

Technique	Recall Rate
SVM	0.82
PSVM	0.94

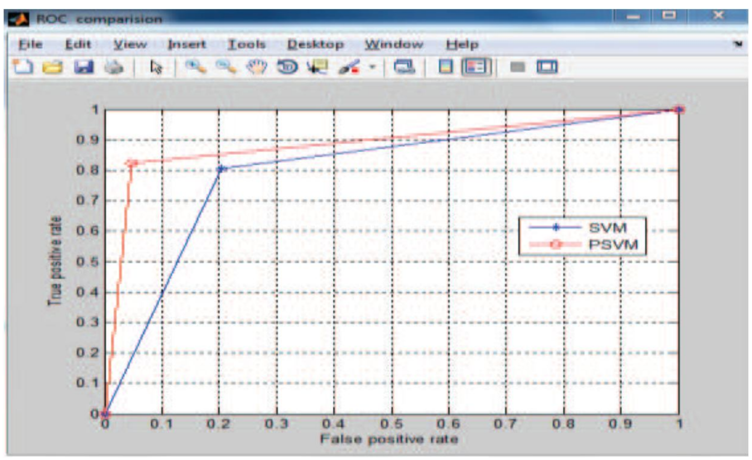


Fig. 10. ROC Comparison

The Fig.10 shows the ROC comparison result of existing SVM and proposed PSVM algorithm.

TABLE V. ROC RESULTS COMPARISON

False Positive Rate	ROC Rate	
	SVM	PSVM
0.2	0.79	0.81
0.4	0.82	0.85
0.6	0.85	0.87
0.8	0.89	0.94
1	0.95	1

IV. CONCLUSION

In this proposed work, effective segmentation and classification is proposed using HFS-SOM and PSVM. After segmentation, the resultant image is given as input to the PSVM classifier followed by feature extraction and selection using GLCM-PCA. At the training phase of PSVM, the texture features are utilized which can reduce the computation complexity of PSVM classifier. The experimental result shows that the proposed system shows a high accuracy rate and less error rate. The proposed system is highly effective for classification to classify normal or abnormal brain with high sensitivity, specificity and accuracy rate. In future the system can be improved to support other types of cancer images with few modification either in segmentation and classification stage. It is necessary to support large number of input and should improve the accuracy rate. To achieve this more number of features can be added with the utilization swarm based feature selection to improve the tumor detection and the classification result.

REFERENCES

[1] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magnetooptical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740-741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982]. A. K Jain, M. N. Murty,P. J Flynn., (1999), "Data Clustering – A Review, ACM Computing Survey", vol.31, no. 3, pp. 265 322.

[2] J. B. MacQueen., (1967), "Some Methods for Classification and Analysis of Multivariate Observations", Proceedings of 5-th Berkeley symposium on Mathematical Statistics and Probability, vol. 1, pp.

[3] J. C. Bezdek, (1981), "Pattern Recognition with Fuzzy Objective Function Algorithms", Plenum Press, New YorkT. Kohonen, Selforganizing maps, Springer,2001

[4] N. Duta and M. Sonka, "Segmentation and Interpretation of MR Brain Images: An Improved Active Shape Model,"IEEE Transactions On Medical Imaging, Vol. 17, No. 6, December 1998, pp.1049:1062.

[5] Guler I., Demirhan A. and Karakis R., Interpretation of MR Images using Self Organizing Maps and Knowledge based Expert systems, Digital Signal Processing 19 66866, 2009.

[6] K. Tasdemir, E. Merényi, Exploiting data topology in visualization and clustering of self-organizing maps, IEEE Transactions on Neural Networks 20 (April (4))(2009) 549–562.

[7] J. Alirezaie, M. Jernigan, C. Nahmias, Automatic segmentation of cerebral MR images using artificial neural networks, IEEE Transactions on Nuclear Science 45 (4) (1998) 2174–2182.

[8] T. Logeswari, M. Karnan, Hybrid self-organizing map for improved implementation of brain MRI segmentation, in: International Conference on Signal Acquisition and Processing, 2010.

[9] Bagade, Sapana Shrikrishna, and Vijaya K. Shandilya. "USE OF HISTOGRAM EQUALIZATION IN IMAGE PROCESSING FOR IMAGE ENHANCEMENT." International Journal of Software Engineering Research and Practices 1, no. 2 (2011): 6-10.

[10] Jian, Y. U. "Texture Image Segmentation Based on Gaussian Mixture Models and Gray Level Co- occurrence Matrix." In Information Science and Engineering (ISISE), 2010 International Symposium on, pp. 149-152. IEEE, 2010.

[11] Mubashir Ahmad, Mahmood ul-Hassan, Imran Shafi, Abdelrahman Osman, Classification of Tumors in Human Brain MRI using Wavelet and Support Vector Machine, IOSR Journal of Computer Engineering (IOSRJCE), Volume 8, Issue 2 (Nov. - Dec. 2012), PP 25-31

[12] Fung, Glenn, and Olvi L. Mangasarian. "Proximal support vector machine classifiers." In Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 77-86. ACM, 2001.[13] Vapnik, Vladimir. The nature of statistical learning theory. springer, 2000.