

## Classification of Tumors in Human Brain MRI using Wavelet and Support Vector Machine

Mubashir Ahmad<sup>1</sup>, Mahmood ul-Hassan<sup>2</sup>, Imran Shafi<sup>3</sup>, Abdelrahman Osman<sup>4</sup>

<sup>1</sup>(Department of Computing and Technology, IQRA University Islamabad, Pakistan,

<sup>2</sup>(Department of Computer Science, Arts and Sciences college in Tabrjal/Al Jouf University, KSA,

<sup>3</sup>(Department of Computing and Technology, Abasyn University Islamabad, Pakistan,

<sup>4</sup>(Department of Computer Science, Arts and Sciences college in Tabrjal/Al Jouf University, KSA,

---

**Abstract :** The task of classification in recovery systems is to differentiate between normal and abnormal brain. In this paper feature extraction from MR Images is carried out by DAUB-4 Wavelet method. DAUB-4 is an efficient tool for feature extraction because it gives better contrast to an image. Due to better contrast it improves easily hanging signals of an image and reduces the overhead. PCA is used to select the best features for classification. These PCA selected features are given as an input to SVM for classification. In this work we are using two SVM kernel functions which are Linear Kernel and Radial Basis Kernel. Experimental results show that the proposed system have high classification accuracy of 98.7% with Radial Basis Kernel.

**Keywords:** Wavelets transform, Principle component analysis, Support vector machine, MRI, Brain tumor

---

### I. INTRODUCTION

Medical imaging is a warm research area in image processing with the development of efficient digital signal processing hardware. Medical diagnostics can easily provide image in digital formats. But with this provision researchers are now trying to automate the said diagnostics, give support to doctors and practitioners to extract information accurately and easily [1]. The mean of medical imaging is to give information about inner system of human body which can facilitate doctors to detect diseases and also explore such procedure that are non-persistent. It is normally used for investigation of different abnormalities like tumor, identifying bones fractures inside the body and vessels blockage. Medical images provides thorough detail examination regarding to the human organs and methods that are frequently used for identifying and investigating diseases consist of computed tomography (CT) scans, magnetic resonance imaging and magnetic resonance angiography. Two most widely used modalities among all above techniques are MRA and MRI. These methods are ideal due to their non-persistent natural history i.e. painless property and less exposure to radiation [2]. MRI is rapidly growing medical imaging technique and capture high resolution images of soft tissues [3]. Magnetic resonance imaging (MRI) is a non-invasive technique for classifying cells composed of tissues in human body [4]. In the soft tissues of the human body MRI provides thorough detail about abnormalities that may not be identified by X-rays and CT scan.

Many techniques have been developed for feature extraction from MRI but wavelet transform is the best method for feature extraction [5, 11]. Wavelet is a non statistical method which gives local frequency information and detail coefficients of the image at various levels. We have used Daub-4 Wavelet in our work which is the best feature extraction method. Daub-4 gives good contrast to an image. Due to good contrast Daub-4 extracted very low signals of MRI. Principle component analysis (PCA) reduces the dimensions and overcome the computational complexity is used for best feature selection. Recent work shows [11] that MRI classification is possible using supervised techniques like SOM and unsupervised techniques like ANN, KNN and SVM. SVM is the best method for MRI classifications because of its large margin in a feature space [13]. The aim of this research work is to select the best method for feature extraction and selection. Most important is to select a classifier which performs better for normal and abnormal human brain MRI classification.

### II. BACKGROUND WORK

R. Mishra [4] proposed tumor identification system using wavelet packet aspect and Artificial Neural Networks in MR images. Features were extracted from wavelet packet and classify features using Artificial Neural Network as normal and abnormal images in this system. The improvement of using wavelet packet is that it gives wealthy investigation by decomposing estimation as well as detail component every time as compared to wavelet transformation method. Artificial Neural network used inaccuracy propagation policies for its training. E. A. El-Dahshan and T. Hosney [5] proposed a hybrid system for tumor detection in MR images and categorize them using artificial neural networks (ANN) and k-nearest neighbor (KNN). During this method features were extracted using discrete wavelet transform (DWT) and then for selection of best extracted features, principle component analysis (PCA) method was used. These selected features were given as input to

classifiers as KNN and ANN. K-Nearest Neighbor and Artificial Neural Network involve two phases, training and testing. These two classifiers KNN and ANN categorize MR images as normal and abnormal images. The negative aspects of these systems are, every time new information is arrived needs new training and KNN precision can rapidly degrade when number of features grows. H. Selvaraj et.al [6] proposed a system for classification of Magnetic Resonance Images by means of wavelet features which were given as input to SVM and ANN. In proposed method neural network Self Organizing Maps (SOM) were used as classifiers for tumor recognition and it simply captures non linear computation and reached precision rate of 94% as evaluated to SVM which captures linear and non linear computation and precision rate attained was 98%. S. Chaplot and L. M. Patnaik [7] proposed brain tumor identification using wavelets transformation method and SVM. Generally two phases were considered. Initially, in processing phase noise was detached from the signal and through wavelet method features were extracted and then features were given as input to SVM for classification as normal and abnormal brain. A.E. Laskhari [10] proposed an a technique based on neural Network for brain tumor detection in MR images using geometric and Zernike moments. Magnetic Resonance images were used as input images. Feature extraction phase occupys statistics features collection by mean, median, entropy and standard deviation as well as a non-statistic feature by geometric moment's invariants. Selection of features was carried out by kernel F-score technique and given as input to Artificial Neural Network classifier. ANN classifies it into two classes as normal and cancer brain tissues. A. E. Laskhari [9] proposed a technique based on artificial neural network for brain tumor detection in MR images using Gabor Wavelets. This technique used PD, T1 and T2 weighted as input images. Preprocessing is performed for noise reduction to obtain excellent results. Feature extraction stage has two categories i.e. statistical features extraction by median, mean and variance wherever non-statistical feature extraction by Gabor wavelet technique. Feature selection stage is carried out by kernel F-score technique that is used to calculate the variation between two classes. Whilst features were chosen, these features were given as input to ANN for normal and abnormal brain tissues. A. Kharrat et.al [11] proposed a system for brain tumor classification using genetic algorithm and support vector machine (SVM). The features have been extracted by two methods. Firstly, extraction of features from both normal and abnormal images is carried out by spatial gray level dependence method (SGLDM). Secondly, the image is decomposed at second level by applying Daubechies wavelet transform. The optimal features were selected by genetic algorithm. Those optimal features were given as an input to classifiers SVM. The accuracy rate achieved by this technique was 96.36% to 97.59%.

### III. PROPOSED METHOD AND MATERIAL

This section explains the proposed system design and methodology which concerns its fundamental design and the features of each of three parts of proposed method. The proposed method consists of number of phases such as Feature extraction from MRI data set using Wavelets. Feature reduction using principal component analysis (PCA) to select most valuable features. This section also explains the classification stage in which support vector machine is used for classification.

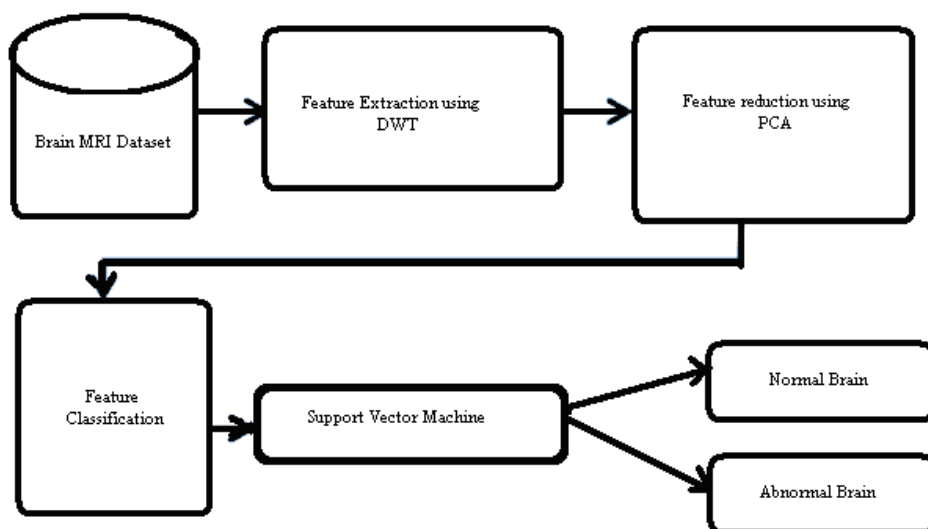


Figure 1: Proposed Methodology

#### 3.1 Brain MRI Dataset selection

Dataset we are using in this work consists of  $256 \times 256$  pixels, T2-Weighted MRI brain images. These images are taken from Ibrar MRI & CT Scan Rawalpindi, Pakistan of original patients. In this study dataset

having 75 images in which 65 images are abnormal having tumor and 10 images are normal. This database contains images of axial plane. The following figure shows Normal and Abnormal brain images.

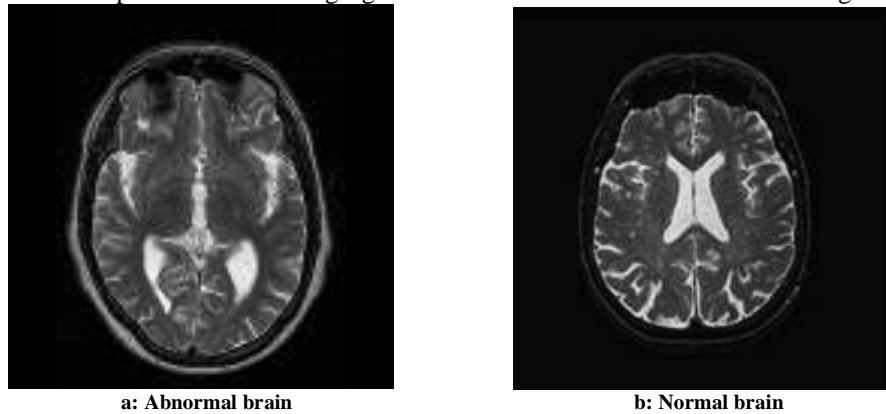


Fig.2: Normal and abnormal brain images

### 3.2 Feature Extraction by Wavelets

Feature extraction is the first stage of classification in which features of each image is separately extracted from MRI by Wavelets which considers the best method to extract most emphasizing pixels present in images to improve results. To decompose data into different frequency components Wavelets mathematical functions are used and then each component is studied having resolution matched to its degree. For analysis of complex datasets Wavelet is a new powerful mathematical tool. Fourier transform fails due to the redundant set of features and only give frequency information that is not spatially localized but wavelet provides time-frequency information which is localized in space and perfect tool for pattern recognition tasks.

#### 3.2.1 Discrete Wavelet Transform

DWT is an implementation method for wavelet obeying some defined rules and discrete set of wavelet translations [7]. It is necessary for practical computation to make wavelet discrete. The scale limit is then discredited with respect to the translation limit ( $\tau$ ). The following equations show the scale and translation of the wavelet.

$$s = 2^{-m}$$

$$\tau = n2^{-m}$$

Where  $m, n \in Z$ , so  $Z$  is the set of all integers. Now equation 1 represents the family of wavelet function.

$$\Psi_{m,n}(t) = 2^{m/2} (2^m t - n) \tag{1}$$

The equation no 2 and 3 shows DWT crumbles a signal  $x(t)$  into a family of synthesis wavelet are given below.

$$x(t) = \sum_m \sum_n C_{m,n} \Psi_{m,n}(t) \tag{2}$$

Wherever

$$C_{m,n} = (x(t), \Psi_{m,n}(t)) \tag{3}$$

$x[n]$  is the discrete time signal in the equation number 4, decomposition of wavelet on  $I$  is given below

$$x[n] = \sum_{i=1}^I c_i [n - 2^i k] + \sum_{k \in Z} d_{i,k} h_i [n - 2^i k] \tag{4}$$

Where wavelet coefficients are  $c_{i,k}$ ,  $i = 1 \dots \dots \dots I$ . and Scaling coefficients are  $d_{i,k}$

$i = 1 \dots \dots \dots I$ .

So wavelet is given in equation 5 and wavelet scaling coefficients are given in equation 6 below.

$$c_{i,k} = \sum_n x[n] g_i^* [n - 2^i k] \tag{5}$$

$$d_{i,k} = \sum_n x[n] h_i^* [n - 2^i k] \tag{6}$$

Where  $*$  specifies complex conjugate and  $g_i^* [n - 2^i k]$  represents discrete wavelets and  $h_i^* [n - 2^i k]$  represents scaling sequence.

#### 3.2.2 Two dimensional Discrete Wavelet Transform

DWT has been used to extract the wavelet coefficient from MRI at different directions and scales. Wavelet decomposition provides four sub bands [5] which are as follows.

- Approximated (LL)
- Horizontal (LH)
- Vertical (HL)
- Diagonal (HH).

Approximated sub band (LL) is normally further decomposed at different scales while three sub bands (LH, HL, HH) include characteristic of an image. DWT overcomes the drawback of Fourier transform by recognizing point discontinuities in image. Discrete Wavelet transform gives local frequency information. Fig.3 shows wavelet decomposition method.

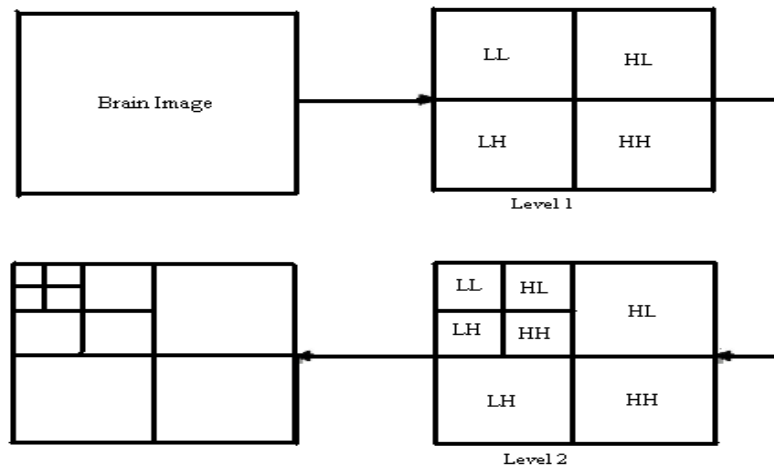


Fig.3: 2D Wavelet decomposition

LL is lower frequency component band. If LL band is further decomposed into next level then LL band will be distributed into further four parts and each part will have  $N/2 \times N/2$  coefficients and in second level  $N/2 \times N/2$  coefficients are further decomposed into four parts having  $N/4 \times N/4$  coefficients each and in the next level same procedure will be repeated.

**3.2.3 Choosing Mother Wavelet**

The origin of a wavelet transform that is restricted or localized is called the mother wavelet. The computationally very expensive form of wavelet is Daubachies (DAUB-4) wavelet [7, 13] has the benefit of superior resolution for efficiently varying signals. DAUB-4 wavelet is a family of orthogonal wavelets defining a discrete wavelet transforms and differentiated by a highest quantity of disappearance moments for some given support. DAUB-4 are the best tool for feature extraction, due to this reason we decide to extract the DAUB-4 wavelet coefficients of brain MRI and use these coefficients as feature vector for classification.

- DAUB-4 Wavelet improves the low signals which are neglected by other Wavelet transforms because DAUB-4 improves the contrast of an image [13].
- Fast tool for detection and classification with limited number of features.

**3.3 Feature Reduction using PCA**

In feature reduction stage, we have applied Principal Component Analysis (PCA) to reduce dimensionality of data to get most favorable features from entire data set [5]. PCA first 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

$$PC_1 \geq PC_2 \geq PC_3 \dots \dots \dots \geq PC_N \tag{8}$$

In equation 8, 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 can be useless as these 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 our research is to reduce classifiers calculation and to make classification proficient; so we have to choose smallest amount of features that can give best precision. At last, from PCA features we have selected first L columns of matrix M. Though, we have chosen first few columns of PCA reduced feature that have high variations.  $L \leq N$  in equation number 9.

$$PC_1 \geq PC_2 \geq PC_3 \dots \dots \dots \geq PC_L \tag{9}$$

L is the number of columns in above equation. Detail of PCA method is given in [5].

**3.4 Classification using Support Vector Machine**

Vapnik developed SVM which is a binary classification method to minimize structural risk. SVM is an attractive and systematic method for two class problems. In this research work we are classifying images into two separate classes such as normal and abnormal using Support Vector Machine. This classifier is usually used

in several research areas due to its out performance in pattern recognition and image processing tasks. Support vector machine (SVM) is extensively used in pattern recognition and image processing due to their originality of concepts using strong mathematical base and have strong realistic ability. SVM take intelligence from its training set to classify unknown data in testing phase. SVM suits most excellent for classification troubles with small training dataset and high dimensional feature space. Like neural networks, SVM also needs two preparation stages; training and testing stage. SVM trains itself by features given as an input to its learning algorithm. The aim of SVM is to choose the suitable margins between two classes during training. Features are labeled according to class associative with particular class. Artificial neural network has a few issues having local minima and number of neurons selection for each problem. In order to resolve this problem SVM occupies no local minima and overhead to neurons selection by initiating the idea of hyper planes.

### 3.4.1 Implementation using SVM Kernel functions

Support Vector Machine utilize kernel tricks; the impact of kernel is double fold by mapping the input to adequately large feature space; wherever patterns turns into linearly distinguishable that makes classification job easier. In this research work we have used two types of SVM kernels which are as follows.

#### a. Linear Kernel function:

This is the simplest case in which data is linearly separable. Linearly separable data is not performing well due to the small margin in feature space. The following equation shows the form of linear function

$$F(x) = W^T \chi + b, \tag{10}$$

Such that for every training example  $f(\chi_i) \geq 0$  for  $y_i = +1$ ,  $f(\chi_i) < 0$  for  $y_i = -1$ ; Training samples are separated by hyper plane using

$$F(x) = W^T \chi + b = 0; \tag{11}$$

So SVM separate the two classes due to their high margin.

#### b. Radial Basis Kernel function:

A decision rule of the radial basis function is in the following equation.

$$F(x) = \text{sign} \left( \sum_{i=1}^n a_i y_i k_\gamma(|x - x_i| - b) \right) \tag{12}$$

Equation 12 shows that N represents total number of support vectors, Kernel function having width parameter is  $\gamma$ . So  $k_\gamma(|x - x_i|)$  is depends on  $|x - x_i|$  that is the distance between two vectors.

$$K(x_i, x_j) = \exp \left[ -\frac{\|x_i - x_j\|^2}{2 \sigma^2} \right] \tag{13}$$

This function is called as Gaussian function and it is the activation function for Radial Basis Kernel [11] and basically it deal with data that have a class conditional probability distribution function approaching the Gaussian distribution. In equation 13  $(2 \sigma^2)^{-1} = \gamma$  where  $\gamma$  is the width of the kernel function. The advantages of the RBF kernel are that, all support vectors automatically obtained as a part of the training process. We have discussed two kernel functions in which radial basis function is very useful because vectors are nonlinearly mapped to a very high dimension feature space.

## IV. RESULTS AND DISCUSSION

We have taken the features of 75 T-2 Weighted images in which 65 images are abnormal and 10 images are normal. The size of each image is 256×256 having dicom format. Level-2 of wavelet transform gives 64 × 64 approximation coefficients and at level-3 32 × 32 coefficients are obtained but these coefficients are computationally very expensive and large set of approximation coefficients because we are using wavelet transform of Daubechies family. DAUB-4 is computationally expensive and gives better contrast. So due to better contrast it differentiates itself from other Wavelet families. By proper analysis of wavelet in MATLAB 7.8 we have decided that level 4 is the best choice using DAUB-4 wavelet. So the coefficients at level-4 are best fit to our problem. Only 16 × 16 features are extracted from LL band using DAUB-4 wavelet for each image. Level-2 and Level-3 are computationally very expensive. Level-4 gives best classification accuracy but also very manageable set of approximation coefficients. So feature extraction from DAUB-4 differentiate our work from other researches because it gives very good contrast. Due to good contrast DAUB-4 extract those pixels of an image which are discarded by other wavelet methods.

Wavelet features which are obtained from DAUB-4 wavelet are stored in a vector form. Reshape the coefficients in one row for each image and then total of 75 rows are stored in a single matrix having normal and tumor images. PCA gives the variation of each column from wavelet matrix. Now we store the variation of the matrix in ascending order. So the first five columns of high variations are selected for classification. The purpose of high variation features selection is very clear that we are selecting such type of features which differ for each image. In this work we have selected only 7 features for each image because our extraction approach is very efficient. Most emphasizing features are extracted from DAUB-4 wavelet method

**Table 1: Our System analysis with SVM Kernel functions**

| Our Hybrid Techniques | Total Images | TP | FN | TN | FP | Specificity% | Sensitivity% | Accuracy% |
|-----------------------|--------------|----|----|----|----|--------------|--------------|-----------|
| SVM-Linear            | 75           | 64 | 1  | 7  | 3  | 70           | 98.46        | 94.7      |
| SVM-RBF               | 75           | 65 | 0  | 9  | 1  | 90           | 100          | 98.7      |

TP= True Positive, FN= False Negative, TN= True Negative, FP= False Positive

In our work we are using SVM as a classifier for brain images classification using MATLAB 7.8 with DAUB-4 wavelet transform and PCA based features selection given as input to SVM-OSU MATLAB toolbox. This is 2-dimensional classification method. We are using two class pattern classification problem in this research work. SVM involves two stages which are training and testing with a specified Kernel functions. There are two types of SVM Kernel used in our work which are Linear and Radial Basis kernel. Training is very vital part of the image classification as it sets the parameters for a classifier because SVM thinking depends on its previous experience.

Linear SVM having cost of 3 is used for training. Linear kernel is not giving valuable results due to fewer margins in feature space. A radial basis function proves to be useful from linear kernel function, as the vectors are nonlinearly mapped to a very high dimension feature space. The  $\gamma$  and  $C$  are constants in radial basis kernel, where the kernel function has the width  $\gamma$  and error/trade-off parameter having cost  $C$  that adjusts the importance of the separation error in the creation of the separation surface. The classification for the MRI data set with  $(\gamma, C)$  varying along a network is carried out. RBF is used as a non linear function in our work with value of  $\gamma$  is 3 and  $C$  is 8.

#### 4.1 Efficiency of our System

Our System gives better results. The following reasons that our system using DWT-DAUB-4 and SVM-RBF having very high accuracy rate.

**Table 2: Results Comparison with other Techniques**

| Techniques  | Accuracy% |
|---|-----------|
| Wavelet (DAUB-4) + PCA + SVM-RBF (Our Technique)    | 98.7      |
| Wavelet (DAUB-4) + PCA + SVM-Linear (Our Technique) | 94.7      |
| Wavelet (Haar) + PCA + KNN [5]                      | 98.6      |
| Wavelet (Haar) + PCA + ANN [5]                      | 97        |

- i. We are using DAUB-4 wavelet transform which is expensive in computation but it gives better contrast to easily hanging signals of the image.
- ii. DAUB-4 wavelet is a fast tool for detection and classification with limited number of features. Our work proves that we have extracted very limited number of features from DAUB-4 wavelet at level-4.
- iii. We are using Support Vector Machine in our work. Artificial Neural Network has a few issues having local minima and number of neurons selection for each problem. In order to resolve this problem SVM occupies no local minima and overhead to neurons selection by initiating the idea of hyper planes.
- iv. SVM has good detection rate and very low false alarm because we are using a non linear kernel in SVM. Radial basis kernel (Non linear kernel) performs well due to its large margin in feature space.
- v. A radial basis function proves to be useful, as the vectors are nonlinearly mapped to a very high dimension feature space and prediction accuracy improved. ANN prediction accuracy can quickly degrade when number of attributes grows.
- vi. The combination of best methods like DAUB-4, PCA and SVM-RBF prove that our hybrid technique is outperforms and results show that it can also be performed well on the larger datasets.

## V. CONCLUSION

Classification of tumor using brain MRI is a main problem for doctors and practitioners to get valuable results. Previously proposed systems have certain problems that require crucial investigation. System performance is going to be degraded due to the training overhead, computational overhead, bad feature extraction methods, bad feature selection methods and complex classifiers. System performance can be improved by getting higher detection rate and reducing misclassification. This research work developed for a classification of brain MRI as a normal and abnormal class. The aim of this research work is to improve

performance and reduce false alarm. This hybrid technique is designed using Discrete Wavelet Transform (DWT) using DAUB-4 Wavelet, Principle Component Analysis (PCA) and Support Vector Machine (SVM). Proposed approach gives better results than earlier systems developed for the same purpose. This is possible due to DAUB-4 wavelet which gives better contrast and gives support to easily hanging signals of an image. SVM with radial basis kernel have large margin in feature space which reduces the misclassification. The proposed system is proficient for classification to classify normal or abnormal brain with high sensitivity, specificity and accuracy rate. We obtained best accuracy rate than previous systems.

#### REFERENCES

- [1] M. G. Mostafa and M. F. Tolba, Medical Image Segmentation Using Wavelet Based Multi resolution EM Algorithm, *IEEE trans on Industrial Electronics, Technology and automation, IETA*, 2001.
- [2] N. Passat, C. Ronse., J. Baruthio, J. P. Armspach., C. Maillot, *Magnetic resonance angiography: From anatomical knowledge modeling to vessel segmentation*, Medical Image Analysis, pp. 259-274, 2010.
- [3] J. H. Kazmi, K. Qureshi, H. Rasheed, *An Implementation of SAN Filter and Edge Sharpening Method for MRA images*, Malaysian Journal of Computer Science, 2007.
- [4] R. Mishra, *MRI based Brain Tumor Detection using Wavelet Packet Feature and Artificial Neural Networks*, Department of computer science India Ghandi Institute of Technology Indraprastha University, Delhi, India, 2010.
- [5] E. A. El-Dihshan, T. Hosney, A. B. M. Salem, *Hybrid intelligence techniques for MRI Brain images classification*, ELSEVIER, Digital Signal Processing, , Volume 20, pp.433-441, 2010
- [6] S. Chaplot, L. M. Patnaik, Classification of magnetic resonance brain images using wavelets as input to support vector machines and neural networks, *Biomedical Signal Processing and Control*, pp. 86-92, 2006.
- [7] S. Chaplot, L. M. Patnaik, Brain Tumor Diagnosis with wavelets and Support Vector Machine, *proceeding of 3<sup>rd</sup> international Conference on intelligent Systems and Knowledge Engineering*, 2008.
- [8] Q. Ain, I. Mehmood, S. M. Naqvi, M. A. Jaffar, *Bayesian classification using DCT features for Brain tumor detection*, National University of Computer & Emerging Sciences, FAST Islamabad, 2010.
- [9] A. E. Lashkari, *A Neural Network based Method for Brain Abnormality Detection in MR Images Using Gabor Wavelets*, International Journal of Computer Applications, vol. 4.no.7, 2010.
- [10] A. E. Lashkari, *A Neural Network based Method for Brain Abnormality Detection in MR Images Using Zernike Moments and Geometric moments*, International Journal of Computer Applications, vol. 4.no.7, 2010.
- [11] A. Kharrat, K. Gasmii, M. B. Messaoud, N. Benamrani, M. Abid, *A hybrid Approach for Automatic classification of Brain MRI using Genetic Algorithm and Support vector Machine*, Leonardo journal of sciences, ISSN: 1582-0233, pp.71-82, issue-17, 2010.
- [12] I. Ahmad, A. B. Abdullah, and A. S. Alghamdi, *Towards the Selection of Best Neural Network System for Intrusion Detection*, International Journal of the Physical Sciences, pp.1830-1839, 2010h.
- [13] M. F. B. Othman, N. B. Abdullah, N. F. B. Kamal, *MRI brain classification using Support vector machine*, *IEEE, Centre for Artificial Intelligence & Robotics (CAIRO), Universiti Teknologi Malaysia*, 2011.
- [14] S. Lahmiri and M. Boukadoum, Classification of Brain MRI using the LH and HL Wavelet Transform Sub-bands, *IEEE*, ISBN, 978-1-4244-9474-3, 2011.