International Journal of Advance Research in Science and Engineering
Volume No.07, Special Issue No.04, April 2018
www.ijarse.com
IJARSE
ISSN: 2319-8354

Classification of Brain MRI using Watershed Segmentation and Support Vector Machine

Ashwini S. Kamble[1], V. S. Jadhav[2]

[1]*Department of Electronics and Telecommunication, MGM COE, Nanded-431 606, (India)*

[2]*Department of Electronics and Telecommunication, MGM COE, Nanded-431 606, (India)*

ABSTRACT

Magnetic resonance imaging is an imaging technique that acquires high quality image of the anatomical structures of the human body and provides rich information for medical analysis and biomedical research. Segmentation and classification technique for MR brain images are highly important for medical analysis and treatment selection. In this paper an automated watershed algorithm for segmentation and decision support system for classification is proposed. This method first employs watershed segmentation for MR image segmentation and then haar wavelet is used to extract the features from images then applies the principle component analysis (PCA) to reduce the dimensions of features. These features are given as input to the SVM classifier with three different types of SVM Kernel. The proposed system is efficient for the classification of brain tumor images in to Benign or Malignant with a high accuracy of 97.14 for Linear kernel, 92.10 for RBF kernel, and 97.50 for Polynomial kernel.

Keywords- LBP, Watershed Segmentation, Wavelet, PCA, SVM

I. INTRODUCTION

In the development of healthcare services, medical imaging acquires its importance with increase in the need of automated and efficient diagnosis of diseases in short period of time. Magnetic Resonance Imaging (MRI) is an imaging technique that is used in both clinical and surgical environment. Its characteristics include superior soft tissue differentiation, high spatial resolution and contrast. They provide rich information for disease diagnosis and biomedical research.

Magnetic resonance images are examined by radiologist to identify the presence of normal tissue through visual interpretation. Large volume of MRIs and shortage of radiologists have made the work labour intensive, cost expensive and often inaccurate. Hence there is a need for automated systems for analysis and classification of medical images. There exists a large research of work in analyzing the MR brain images. In [1] detection and quantification of brain tumor from MRI of brain are presented. This approach can be able to find the status of increase in the disease using quantitative analysis. In [2] a texture based tumor detection and automatic segmentation using seeded region growing method is presented. This method is region growing segmentation method for segmentation of brain tumor in MRI, in which it is possible to determine abnormality is present in the image or not.

In [3] hybrid techniques for classification of brain are presented. These included wavelet transform, principle component analysis and the supervised learning methods (FP-ANN and K-NN) that gave very promising results in classifying the healthy and brain patient. In [4] a classification system for brain MRIs was proposed by extraction features from the horizontal (LH) and vertical (HL) directions instead of the more common LL sub-band of the wavelet transform. In [5] a principle approach for investigating brain abnormalities were based on wavelet based feature extraction and PCA based feature selection. Classification was done with deep and extreme machine learning comparative to various others classification.

The contribution of this paper deals with watershed segmentation for MR image segmentation and then HAAR wavelet is used to extract the features from images then applies the principle component analysis (PCA) to reduce the dimensions of features. These features are given as input to the SVM classifier with three different types of SVM Kernel. This paper is organized as follows, Section II presents the Overview of the methods, Section III presents the proposed system, and Section IV presents the Experimental results, while the Conclusion is presented in Section V.

II. OVERVIEW

A. Segmentation

Watershed Segmentation is another segmentation technique used to extract the region of interest which in this case is a tumor from the smeared MRI image. This segmentation technique is particularly useful to segment objects when they are touching each other. This algorithm helps in finding the catchment basins and ridge lines in the image. In this case, the ridge line represents the height that separates the two catchment basins. For this, bwdist() method is used to compute the distance from every pixel to every nonzero pixel. Watershed algorithm is then implemented using watershed () function. This returns a labeled matrix which consists of positive integer values for different regions and 0 for ridge lines. This image is not very useful as there is only one catch basin spanning the entire image. Here catchment basins are the regions we want to identify. Therefore we take the compliment of our image and apply bwdist() on the complimented image after which we negate the distance in order to determine the bright catchment basins that represents individual regions. We then apply watershed () method which will return a labeled matrix consisting of positive values along with catchment basins. To display the image, we convert the labeled matrix into an rgb image and display the result. However before taking the compliment of the image to apply watershed algorithm, several pre-processing techniques are applied to enhance the region of interest we want to extract. In the project, Otsu's thresholding followed by morphological erosion of disk size 3 has been applied to enhance the contrast of the tumor region. This technique gives accurate results if the image is of reasonable contrast so that the watershed algorithm can extract the catchment basins efficiently.

B. Feature Extraction:

The proposed system uses the Discrete Wavelet Transform (DWT) coefficients as feature vector. The wavelet transform is a powerful mathematical tool for feature extraction, and has been used to extract the wavelet

 IJARSE
ISSN: 2319-8354

coefficient from MR images. Wavelet are localized basis functions, which are scaled and shifted versions of some fixed mother wavelets is that they provide localized frequency information about a function of a signal, which is particularly beneficial for classification [6]

The basic fundamental principle of DWT is introduced as follows [7]: The continuous wavelet transform of x(t) relative to a given wavelet (t) which is square-integral function is defined as:

$$\mathbf{W_{\psi}(a,b) = \int X(t) * \psi_{a,b}(t)dx} \qquad (1)$$

Where, Ψ (t) = $\frac{1}{\sqrt{|a|}}\ \Psi(\frac{t-a}{b})$, and the wavelet Ψ is computed from the mother Ψ wavelet by translation and dilation: a the dilation factor and b the translation parameter (both being real positive numbers).

The eq. (1) can be discredited by restraining a and b to a discrete lattice to give the discrete wavelet transform (DWT). DWT can be expressed as

$$\mathbf{DWT_{X(n)} \quad =d_{j,k}=x(n) \quad h_j*2_{j,k}, \qquad a_{j,k}=x(n) \quad g_j*2_{j,k}} \qquad \mathbf{(2)}$$

The coefficient $d_{j,k}$ refer to the detail components where as $a_{j,k}$ refer to the approximation components in the signal x(n) that corresponds to the wavelet function. In the equation, function h(n) and g(n) represents the coefficients of the high pass and low pass filters, respectively, while parameter j and k are wavelet scale and translation factors.

<table>
<tr><td>LL2</td><td>HL2</td><td rowspan="2">HL1</td></tr>
<tr><td>LH2</td><td>HH2</td></tr>
<tr><td colspan="2">LH1</td><td>HH1</td></tr>
</table>

Fig. 1 Two level 2D-DWT decomposition of an image

The main feature of DWT is multiscale representation of function. By using the wavelets, given function can be analyzed at various levels of resolution. Fig.1 illustrates DWT schematically. The original image is process along the x and y direction by h(n) and g(n) filters which, is the row representation of the original image. As a result of this transform there are 4 sub-band (LL, LH, HH, and HL) images at each scale. The LL sub-bands can be regarded as the approximation component of the image, while the LH, HL and HH sub-bands can be regarded as the detailed components of the image. In our algorithm, 3-level decomposition via Haar wavelet was utilized to extract features.

C. Feature Reduction:

Principal Component Analysis is an efficient tool to reduce the dimension of data set consisting of a large number of interrelated variables while retaining most of the variations. The PCA basis vectors are the eigenvector of the covariance matrix of the input data. This is useful for the exploratory data analysis of multivariations as the new dimensions called principal components PCs [8]. A reduced dimension is formed by choosing the PCs associated with highest eigenvalues.

Calculation of PCA for data set X(X: matrix of dimensions M × N) involves the following steps:

Step1: Calculate the empirical mean $\frac{1}{N}\sum X[m,n]$

Step2: Calculate the deviations from the mean and store the data in matrix B [M×N]: B=X - u×h, where h is a 1×N row vector of all 1s: h(n) =1 for n=1….. , N.

Step3: find the covariance matrix C: C= $\frac{1}{N}$ B×B.

Step4: Find the eigenvectors and eigenvalues of the covariance matrix $V^{-1}CV=D$: V the eigenvectors matrix; D the diagonal matrix of eigenvalues of C.

Step5: Rearrange the eigenvectors and eigenvalues.

Step6: Choosing components and forming a feature vector: save the first column of V

Step7: Selection Of Features: The eigenvector with the highest eigenvalues are projected into space. Number of Feature can be selected by calculating the classification accuracy.

D: SVM Classification:

Support vector machines become effective in high dimensional space. It can be effective in case be effective in cases where number of dimensions is greater than the number of samples [9]. SVM are memory efficient as it uses the subset of training points (also called as support vectors) in the decision function. They are versatile in the sense that different Kernel functions can be specified for the decision function. Supervised Learning is implemented in SVM by assigning the data set with known labels [10]. This indicates whether the system is performing in a right way or not. Validation of the accuracy of the system is calculated from the information gained from a desired response. Thus it helps the system to act correctly. A hyper plane may be used to divide the linear data separately. However if the data is far from linear distribution then the data sets are inseparable. Kernels can be used to non-linearly map the input data to a high –dimensional space. The new mapping now becomes linearly separable. The Kernel trick is used by SVM to form nonlinear boundaries.

The dot product of nonlinearly mapped data set can be expensive. The kernel trick will pick a suitable function that will correspond to the dot product of some nonlinear mapping data set. A particular kernel is only chosen by trial and error on the test data. Choosing the correct Kernel based on the application or problem would enhance SVM performance.

RBF is the most popular kernel function used in the classification of data set by support vector machine. Consider X and X' be the two feature vectors in input space. It is defined as:

$$K(X,X') = \exp\left(\frac{\|x-x'\|^2}{2\sigma^2}\right) \quad (3)$$

A polynomial function uses a polynomial mapping. It is defined as:

$$K(X,X') = ((X,X') + 1)^d \quad (4)$$

A linear Kernel function is defined as:

$$K(X - X') = (X,X') \quad (5)$$

III. PROPOSED SYSTEM

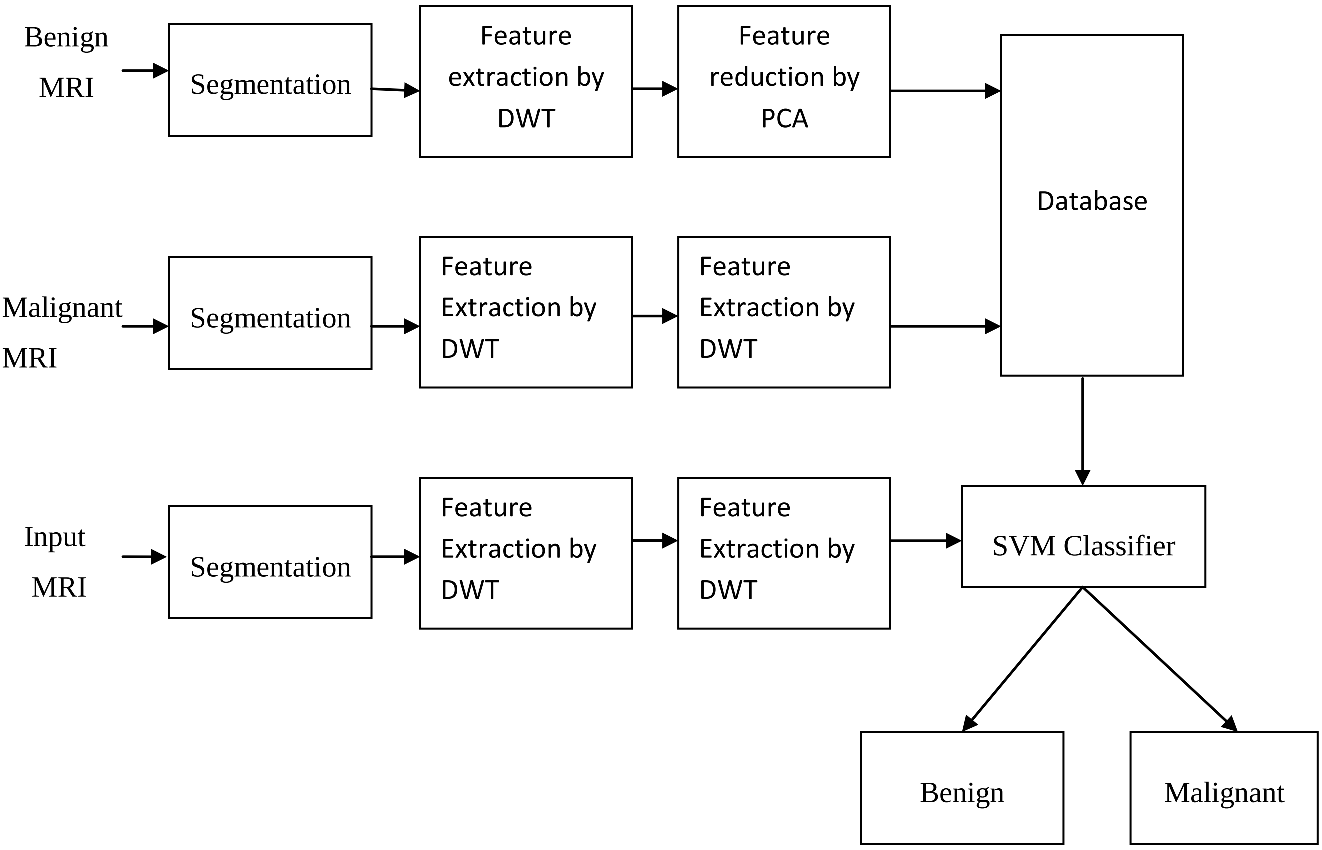
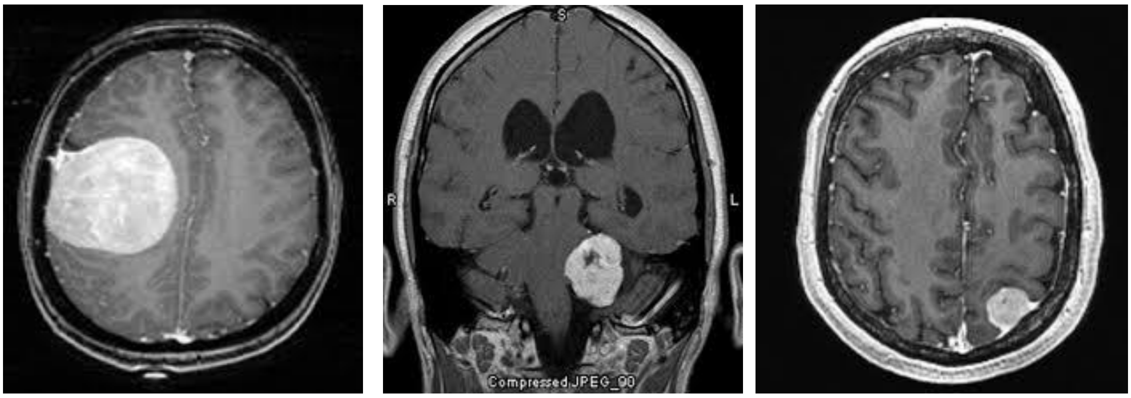


Fig. 2 Proposed system

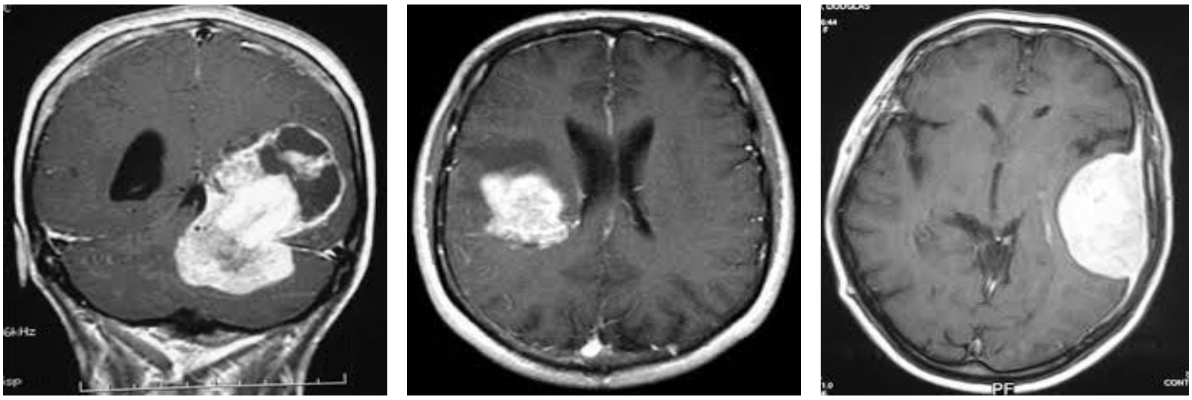
Figure 2 shows the flow diagram for the proposed system. The proposed techniques are based on the following techniques: Watershed segmentation, Discrete Wavelet Transform (DWT), Principal Component Analysis (PCA) and classification performed by SVM with different kernels. Watershed algorithm is used for the image

IJARSE
ISSN: 2319-8354

segmentation is to find the watershed line. After segmentation feature extraction stage and a feature reduction stage. Wavelet transform is an effective tool for feature extraction from MR brain images, due to its multi resolution analytic property. This helps to analyses the images at various levels of resolution [11]. However, this technique requires large storage and is computationally expensive. In order to reduce the feature vector dimensions and increase the discriminative power, the principal component analysis (PCA) is used [12]. PCA is appealing since it effectively reduces the dimensionality of the data and therefore reduces the computational cost of analyzing new data. The classification of images under two categories, either Benign or Malignant. Classification is possible via supervised classification methods such as support vector machine (SVM) [13].



(a)



(b)

Fig. 3 (a) Benign Images (b) Malignant Images

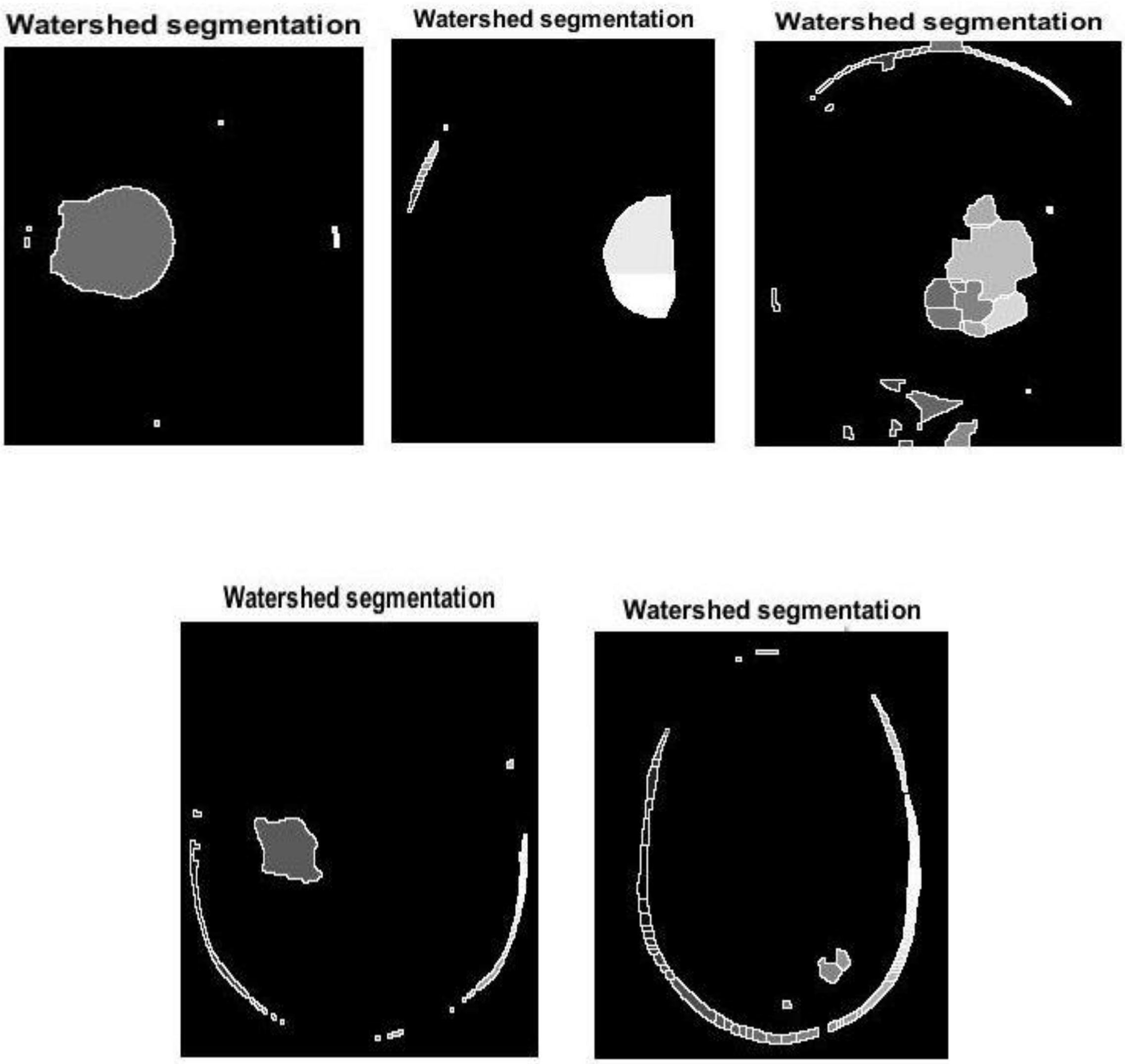


Fig. 4 Segmented Images

IV. EXPERIMENTAL RESULTS

The datasets consists of 50(25 images are benign and 25 images are malignant) MR brain images in axial plane and 256×256 in plane resolution, which were downloading from the Harvard Medical School website (http://med.harvard.edu/AANLIB/). Table 1 shows the classification accuracy with levels of Watershed Segmentation + Haar + PCA with linear, rbf, polynomial kernel. Classification accuracy is improved by adding segmentation method. Figure 3 shows the sample of benign and malignant T2- weighted images.

$$\text{Classification Accuracy}=\frac{\text{Correctly } \mathit{Classified\ images}}{\mathit{Test\ Images}}$$

Table 1. Classification accuracy results for watershed segmentation based KSVM

Kernel SVM	Test Image	Correctly Classified Image	Incorrectly Classified Image	Classification accuracy
RBF	38	35	3	92.10
Polynomial Kernel	40	39	1	97.50
Linear kernel	35	34	1	97.14

V. CONCLUSION

Classification of brain MRI as Benign and Malignant is a main problem faced physicians. In order to get an actual assessment of disease a medically accurate support system is required. System performance can be improved by good segmentation, good feature extraction, feature selection and accurate classification. The proposed system is designed using an automated watershed algorithm for segmentation and decision support system for classification is proposed. This method first employs watershed segmentation for MR image segmentation and then haar wavelet is used to extract the features from images then applies the principle component analysis (PCA) to reduce the dimensions of features. These features are given as input to the SVM classifier with three different types of SVM Kernel. The proposed system is efficient for the classification of brain tumor images in to Benign or Malignant with a high accuracy of 97.14 for Linear kernel, 92.10 for RBF kernel, and 97.50 for Polynomial kernel.

REFERENCES

[1] S. Roy and S. K. Bandyopadhyay, “Detection and Quantification of Brain Tumor from MRI of Brain and its Symmetric Analysis”, International Journal of Information and Communication Technology Research, K Y, USA, June 2012.

[2] M. Kumar and K.K. Mehta, “A Texture based tumor detection and automatic Segmentation using Seeded Region Growing Method”, International Journal of Computer Technology and Application, August 2011.

[3] E.-S.A El-Dahshan, T. Hosny, and A.-B.M. Salem, ”Hybrid Intelligent Techniques for MRI Brain Images Classification,” Digital Signal Processing, vol 20,pp: 433441, 2010.

[4] Salim Lahmiri., Mounir Boukadoum,” Classification of brain MRI using the LH and HL wavelet transform sub-bands”, In proceeding of: International Symposium on Circuits and Systems (ISCAS 2011) , Rio de Janeiro, Brazil,1025-1028,2011.

[5] Lavneet Singh, Girija Chetty, ”Detecting the Brain Abnormalities From Mri Structural Images Using Machine Learning And Pattern Recognition Tools”, IJREISS Volume 2, Issue 11 ISSN: 2250-0588, 15-30,2012.

[6] M. Kocionek, A. Materka, M. Strzelecki, P. Szczypin ski,” Discrete wavelet transform derived features for digital image texture analysis”, in: Proc. Of International Conference on Signals and Electronic Systems, Lodz, Poland, 1821 September, pp. 163168, 2001.

[7] References Burrus CS, Gopinath RA, Guo H. ”Introduction to wavelets and wavelet transforms”. Englewoood Cliffs, NJ: Prentice Hall, 1997.

[8] A. Sengur,” An expert system based on principal component analysis, artificial immune system and fuzzy k-NN for diagnosis of valvular heart diseases”.

[9] N. Cristianini and J. Shawe-Taylor, An Introduction to Support Vector Machines, Cambridge, U.K, Cambridge Univ. Press, 2000.

[10] M. OFarrell, E. Lewis, C. Flanagan, N. Jackman, “Comparison of k-NN and neural network methods in the classification of spectral data from an optical fibre-based sensor system used for quality control in the food industry”, Sens. Actuators B: Chemical 111112C ,354362,2005.

[11] S. Kara, F. Dirgenali, “A system to diagnose atherosclerosis via wavelet transforms, principal component analysis and artificial neural networks”, ExpertSyst. Appl. 32 , 632640,2007.

[12] V.N. Vapnik, ”Statistical Learning Theory”, Wiley, New York, 1998.

[13] S. Chaplot, L.M. Patnaik, N.R. Jagannathan,” Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network”, Biomed. Signal Process. Control 1, 8692,2006.

[14] Harvard Medical School, Web, data available (http://med.harvard.edu/AANLIB/)