A Neural Network based Method for Brain Abnormality Detection in MR Images Using Gabor Wavelets

AmirEhsan Lashkari
University of Tehran, School of
Electrical and Computer engineering
Campus No. 2, North Kargar Avenue
P.O. BOX 14395/515 Tehran. I.R. of
IRAN

ABSTRACT

Nowadays, automatic defects detection in *MR* images is very important in many diagnostic and therapeutic applications. This paper introduces a Novel automatic brain tumor detection method that uses *T1, T2*_weighted and *PD, MR* images to determine any abnormality in brain tissues. Here, has been tried to give clear description from brain tissues using *Gabor* wavelets, energy, entropy, contrast and some other statistic features such as mean, median, variance, correlation, values of maximum and minimum intensity .It is used from a feature selection method to reduce the feature space too. this method uses from neural network to do this classification. The purpose of this project is to classify the brain tissues to normal and abnormal classes automatically, that saves the radiologist time, increases accuracy and yield of diagnosis.

General Terms

Pattern Recognition, Image Processing, Soft Computation, Artificial Intelligent.

Keywords

Feature extraction, Kernel F-score feature selection, Gabor wavelets, artificial neural network, tumor detection, segmentation, MR images.

1. INTRODUCTION

Body is made of many cells. Each cell has specific duty. The cells growth in the body and are divided to reproduce other cells. These divisions are very vital for correct functions of the body. When each cell loses the ability of controlling its growth, these divisions is done with any limitation and tumor emerges. Tumors, their self, are divided to tow classes: benign and malignant. According to a statistical report published by the Central Brain Tumor Registry of the United States (*CBTRUS*), approximately 39,550 people were newly diagnosed with primary benign and primary malignant brain tumors in 2002 [1-3]. Furthermore, in 2000, more than 81,000 people, in the United States alone, were living with a primary malignant brain tumor and 267,000 were living with a primary benign brain tumor. The same report indicates that the incidence rate of primary brain tumors, whether benign or malignant, is 14 per 100,000, while median age at diagnosis is 57 years [3].

According to upon statistics, it is clear that medical imaging has strong function to increase the yield and accuracy of tumor diagnosis in short time. Medical imaging is divided to two classes of anatomical and physiological. The anatomical imaging contains *CT*, Ultrasound and *MRI*.

MR imaging technique, because of good ability in showing difference between soft tissues, high resolution, good contrast and non invasive technique for using no ionization rays is very appropriate. this paper uses images that are provided from this technique. Specially, main purpose of this work is, design, implement and evaluate strong system with complete pattern to extract diverse features that increases accuracy of tumor diagnosis with analyzing *T1*, *T2*_weighted and *PD* brain *MR* images. Moreover at the methods section, it has been explained about the principles of *MR* imaging technique and segmentation problem. Segmentation is the first and most important step in every work that related to image processing. In this section after a short explanation about the image data set and technique which is used to create input images, is explained about the preprocessing steps. Images usually contain one or more type of noise and artifact. In medical images, because of diagnostic and therapeutic applications, this issue is critical. Specially in *MR* images, inhomogeneous magnetic fields, Patient motions duration imaging times, thermal noise and exist of any metal things in imaging environment, are some reasons that can create noises and artifact, though in most of times, are not very important because of human studies on images ,but these are one of the main causes for computational errors in automatic or semi automatic image analyzing methods and so it is needed to be removed by preprocessing procedures before any analyzing. At next section, paper explains about the image processing procedures that itself divides to feature extraction, kernel F-score feature selection and neural network subsections. In each subsection it has been explained about details. At result section, it will be mentioned some results and show one of the output images has been provided from my project. the yield and accuracy of method is measured, using sufficiency indexes. At last, the conclusion section, tells the advantages and disadvantages of this work. Some other future works in this field has been purposed in this part too.

2. MATERIALS AND METHODS

2.1 Input Data Sets

MR images that have been used in this paper are *T1*_weighted, *T2*_weighted and *PD* images. *T1*_weighted *MR* images shows the hard tissue darker gray scale density than neighbor tissues and in *T2*_weighted and *PD, MR* images the hard tissue is brighter gray scale density than neighbor tissues but in all *MR* images modalities, the normal tissues are almost similarity gray scale density. Because of tumor tissues are harder than the normal tissues so abnormalities are seen darker than their neighbor tissues in *MR* images and are brighter than their neighbor tissues in *T2*_weighted and *PD, MR* images but in all types of *MR* images the normal tissues are almost similar. Most of the tumors are

heterogeneous tissues and their mean values of relaxation time are not at all sufficient to characterize the heterogeneity of the different tumor types [7, 8].

Providing an analyze from *(T2_T1)* and *(PD_T2)* images as input image can be used for accurate description from intensity values of abnormal tissues .so better contrast and good remove of normal tissue at brain *MR* image is obtained. It has been used from these images as input images for processing. The images have been obtained from 100 normal brain tissues and 100 abnormal brain tissues. In (Figure 1) one normal and abnormal input image has been showed.
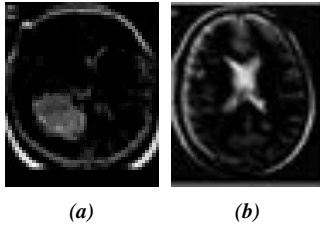

(a) *(b)*

Figure 1: input image   a) abnormal   b) normal

2.2 Preprocessing

Images usually contain one or more type of noise and artifact.

In medical images, because of diagnostic and therapeutic applications, this issue is critical. Specially in MR images, inhomogeneous magnetic fields, Patient motions duration imaging times, thermal noise and exist of any metal things in imaging environment, are some reasons that can create noises and artifact, though in most of times, are not very important because of human studies on images ,but these are one of the main causes for computational errors in automatic or semi automatic image analyzing methods and so it is needed to be removed by preprocessing procedures before any analyzing. Here, preprocessing is equal to remove seeds from images and increase contrast between normal and abnormal brain tissues. The procedure have been used here are Histogram equalization, using Median filter, using Un sharp mask, thresholding and using from Mean filter respectively for each image.In this step, two-dimensional discrete Fourier transform is computed for images. To reduce the noise a 3 by 3 pixel mean filter was implemented. This filtered averaged 9 points thus reducing the noise by 3. Because a single pass of this filter did not seem to provide sufficient noise reduction, the image was passed through the filter a second time (Figure 2).
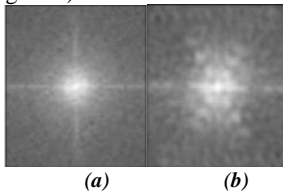

(a) *(b)*

Figure 2: Fourier transform   a) abnormal   b) normal

As it is obvious at images, the Fourier transform of normal tissue is intensive, whereas it is widespread and amorphous in abnormal images. This difference can be done as one feature named entropy that will be explained at the next sections.

2.3 Image Processing

2.3.1 Feature extraction

The purpose of feature extraction is to reduce the original data set by measuring certain properties, or features, that distinguish one input pattern from another. The extracted features provide the characteristics of the input type to the classifier by considering the description of the relevant properties of the image. [8].

The analyzing methods have been done until now used values of pixels intensities, pixels coordinates [8,18,19] and some other statistic features such as mean, variance or median which have much error in determination process and low accuracy and robustness in classification [8-19].

It has been explained about the features that have been used in this paper which are divide to two classes of statistic and non-statistic features. X (i,j) is the value of intensity for location of (i,j) at variation of $\Delta y, \Delta x$ between pixels.

2.3.1.1 Statistic features

Mean: The mean is defined as below:

$$\text{Mean (M)} = m: \text{xx} = \frac{1}{x+y}\sum_{i=1}^{x}\sum_{j=1}^{y}x(i,j) \tag{1}$$

Standard deviations: It is square of variance. The variance is defined as below:

$$\textit{Variance (V)} = \sigma^2 = \frac{1}{x+y}\sum_{i=1}^{x}\sum_{j=1}^{y}(x(i,j)-m)^2 \tag{2}$$

Entropy: A measure of non uniformity in the image based on the probability of co-occurrence values:

$$\textit{Entropy} = \sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{P(i,j)}{R}\right)\log\left(\frac{P(i,j)}{R}\right) \tag{3}$$

Where, N is the number of gray levels, equal to 256 for images in the present study. R is equal to the total number of pixel pairs used for the calculation of texture features in the specified angular direction.

Median: The value that multiplicity of upper values and lower values are equal.

Contrast: A measure of difference moment and is defined as below:

$$\text{Contrast} = \sum_{i=1}^{N}\sum_{j=1}^{N}(i-j)^2\left(\frac{P(i,j)}{R}\right) \tag{4}$$

Energy: A measure of homogeneity that can be defined as:

$$\text{Energy} = \sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{P(i,j)}{R}\right)^2 \tag{5}$$

Inverse Difference Moment: A measure of local homogeneity that can be defined as below:

$$I = \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\left[\frac{i\,P(i,j)}{R}\right]}{1+(i-j)^2} \quad (\text{for } i \neq j) \tag{6}$$

Correlation: A measure of linear dependency of brightness and can be defined as below:

$$\text{Correlation} = \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\left[\frac{i\,P(i,j)}{R}\right]-\mu x\mu y}{\sigma x\sigma y} \tag{7}$$

$$\mu x = \sum_{i=1}^{N}i\sum_{j=1}^{N}\left(\frac{P(i,j)}{R}\right), \quad \mu y = \sum_{i=1}^{N}j\sum_{j=1}^{N}\left(\frac{P(i,j)}{R}\right) \tag{8}$$

$$\sigma_x^2 = \sum_{i=1}^{N}(i - \mu x)^2 \sum_{j=1}^{N}\frac{P(i,j)}{R}, \quad \sigma_y^2 = \sum_{i=1}^{N}(j - \mu y)^2 \sum_{j=1}^{N}\frac{P(i,j)}{R}$$

In the above expressions, μx, μy, σx, σy are the mean and standard deviation values of GCM values accumulated in the x and y directions, respectively.

Line Plot: From the plots in (Figure 3) it can be seen that there is a local minimum and a local maximum in the normal images at a frequency between 50 and 150 mm-1. This would indicate that normal cells contain regular features which repeat, where as the abnormal cells do not contain this repeating nature.
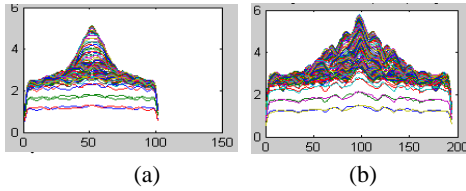

(a) (b)

Figure 3: line plot a) abnormal b) normal

2.3.1.2 Non statistic features
Gabor wavelets

Gabor wavelets transforms like other wavelets transforms have a good attribute in image processing and machine vision. These wavelets assemble tools for processing the image at frequency space and best advantage of them is unruffled variations in frequency space. if I(z) be the input image with gray level intensity values, image Gabor wavelets transforms is computed from convolution of one element from Gabor wavelets with input images ,such as below:

$$O_{\mu,v}(z)= I(z)*\Psi_{\mu,v}(z) \qquad (9)$$

Where * is convolution operator, O is transform result, ψ is the wavelet that has been applied in the transform with the side of μ and magnification of υ. Gabor wavelet equations actually express the simple wave with the certain frequency and side which has been trammeled under the Gaussian function. This equation can be defined with different forms depends on type of coordinate systems such as Cartesian and polar.The standard formula of Gabor wavelet is as below:

$$\Psi_{\mu,v}(z)=\frac{\|k_{\mu v}\|^2}{\delta^2} \exp\left(\frac{\|k_{\mu v}\|^2 \|z\|^2}{2\delta^2}\right)\left(e^{ik_{\mu v}z} - e^{\frac{\delta^2}{2}}\right) \qquad (10)$$

In this equation, k, express the length and side of wave and is computed from the equation which is as below:

$$k_{\mu v} = k_v e^{j\phi_\mu}, \quad k_v = \frac{k_{ma}}{f^v}, \quad \varphi_v = \mu\frac{\pi}{8} \qquad (11)$$

In this equation, μ has been multiplied by $\pi/8$ and have made the phase of k, so have integer values between 0 and 7. Upper values create waves with repetitive sides. Y can have values between 0 and 4 and so, we have 40 wavelets with different sides and sizes. Here has been showed one of the Gabor wavelet elements and $\delta = 2\pi$, k ma =π, f = 2, μ = 2, υ =3 has been showed at (Figure 4).
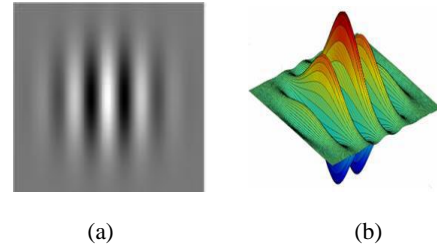

(a) (b)

Figure 4: Gabor wavelet a) 2D b) 3D

As showed in the (Fig. 4) Gabor wavelet is the wave with descent amplitude.

Here, has been explained the way of extraction feature from image. The Fourier transform of image windows that want to decide about them is multiplied at Gabor wavelet Fourier transforms and after that, computed inverse Fourier transforms merged and make the features of desired matrix. It is obvious that all these features are too much for each step and increase the detection process time and so should be reduce with the feature selection space algorithm which mentioned at next section.

2.3.2 Feature selection
Kernel F-score method

First, the *F-score* method is explained and after that, it is explained about the *kernel F-sco*re method. Readers can refer to [20] for more details.

F-score method is a basic and simple technique that measures the distinguishing between two classes with real values. In *F-score* method, *F-score* values of each feature in dataset are computed according to following equation (*Eq.* (12)) and then in order to select the features from whole dataset, threshold value is obtained by calculating the average value of *F-scores* of all features. If the *F-score* value of any feature is bigger than threshold value, that feature is added to feature space. Otherwise, that feature is removed from feature space. Given training vectors x k, $k = 1,. . . ,m$, if the number of positive and negative instances are n+ and n-, respectively, then the F-score of the *i-th* feature is explained as follows:

(12)

$$F(i) = \frac{(\bar{x}_i^{(+)} - \bar{x}_i)^2 + (\bar{x}_i^{(-)} - \bar{x}_i)^2}{\frac{1}{n_+ - 1}\sum_{k=1}^{n_+}(\bar{x}_{k,i}^{(+)} - \bar{x}_i)^2 + \frac{1}{n_- - 1}\sum_{k=1}^{n_-}(\bar{x}_{k,i}^{(-)} - \bar{x}_i)^2}$$

Where $\bar{x}_i$, $\bar{x}_i^{(+)}$, $\bar{x}_i^{(-)}$ is the average of the *i-th* feature of the whole, positive, and negative data sets respectively.$\bar{x}_{k,i}^{(+)}$, is the *i-th* feature of the *k-th* positive instance and $\bar{x}_{k,i}^{(-)}$ is the *i-th* feature of the *k-th* negative instance. The numerator shows the discrimination between positive and negative sets, and the denominator defines the one within each of the two sets. The larger *F-score* for one feature means this feature is more discriminative. But a disadvantage of *F-score* method does not take the mutual information between features into account. In the proposed feature selection method, *kernel F-score* feature selection method is provided both to transform from non-linearly

separable dataset to linearly separable dataset and to decrease the computation cost of classification algorithm. First of all, input spaces (features) of dataset have been mapped to kernel space using Linear (*Lin*) or Radial Basis Function (*RBF*) kernel functions. In this way, the dimensions of datasets have transformed to high dimensional feature space. After transforming from input space to kernel space, the F-score values of datasets with high dimensional feature space have been calculated using *F-score* formula. And then the mean value of calculated *F-scores* has been computed and also this value is selected as threshold value. If the *F-score* value of any feature in datasets is bigger than threshold value, that feature will be selected. Otherwise, that feature is removed from feature space. Thanks to *KFFS* method, the irrelevant or redundant features are removed from high dimensional input feature space. The cause of using *kernel* functions transforms from non-linearly separable medical dataset to a linearly separable feature space.

2.3.3 Purposed method
Proposed method contains 2 steps. In first step, defiant structures are estimated using a neural network and the statistic features that have been explained .These structures are thought be part of tumor or contains tumor.

In second step, finally it is decided about which defiant structures are really tumor or contains tumor using another neural network and non statistic feature (*Gabor* wavelets). At end, it has been done some post processing procedures with morphological procedures such as filling and connected components algorithms to connect the probable discrete points which are exist at image. Tumor location is determined by measuring the primary and extremity points coordinates and then compute lengh, width and height for measuring the volume of tumor. The parts below have explained more about each stage.

2.3.3.1 First stage: finding defiant structures

In this stage, each input feature vector contains all features that are mentioned at the statistic features. At each pixel such as x, the feature vector of F(x) formed from 9 line of vector. In other word $F(x) = [I(x) R(x) P(x) A(x) \ E(x) C(x) M(x) V(x) D(x)]$ where $I(x)، R(x)، A(x)، E(x)، C(x) ، M (x)، V(x)، D(x),$ are Inverse Difference Moment, Entropy, Energy, Contrast, Mean, Standard deviation, Median at the mask with the size of 3 with desired pixel center after passing through feature selection stage respectively. $P(x)$ is the number of the peaks at the line plot. Here, one *MLP* is used to explain the relationship between inputs and outputs. Different architectures have been tested and finally selected one neural network with 9 neuron(input features) at input layer,5 neuron at first hidden layer,3 neuron at second hidden layer, and one neuron at the output layer. The output features are distance function from tumor structures. This function is subset of defiant functions in description of structures and is positive at tumor regions, is zero on the boundary and negative at other areas. At the point that are Adjacent to boundary the absolute values decreases.

The neural network has been trained using back propagation algorithm and training process has been continued until the Mean Square Error (*MSE*) became constant. At this stage is not expected that *MSE* be zero, because have been used from statistic features and know some points that is determined by

network as tumor are not really tumor or contains no tumor. The training process lasts about 9 hours.

2.3.3.2 Second stage: accurate tumor determination

First stage outputs are diverse between different areas. The first stage neural network outputs are merged .this neural network, works such a classifier not such an estimator. The goal of this stage is classifying the image pixels to 2 classes of normal and abnormal.

The input feature vector in this stage contains the non statistic feature *(Gabor wavelets)* plus all first stage outputs. In other word, at each pixel such as x, the feature vector of F(x) formed from 3 line of vector. In other word, $H(x)$ can be defined $H(x) = [O(x),G(x),N(x)]$ where $O(x)$ is the vector that contains all outputs of first stage, $G(x)$ is the vector that contains Gabor wavelets after passing from feature selection stage and $N(x)$ is the vector that contains pixels coordinates in Cartesian coordinate system .the pixels coordinates helps the network to express better input-output relationship and filling the ruptures to create contiguous results.

 The neural network in this stage is *MLP* too. Here like first stage, different architectures tested, and finally one neural network with 23 neuron(input features) at input layer,13 neuron at first hidden layer,5 neuron at second hidden layer, and one neuron at the output layer has been selected. Activation functions are sigmoid at all layers. In this stage, training process lasts about 52 hours. Though this time is much but this process done for just one time and this time is lower in comparison of other similar methods in literatures.

2.3.4 Final processing
In this stage, small holes and unreal pixels is removed by image morphological algorithms. Here, has been used from two image processing morphological algorithms: image filling and connected-component algorithms. image filling is used for filling the holes between network outputs and connected-component used to remove the unreal pixels. In this stage, tumor location is determined by measuring the primary and extremity points coordinates and then compute lengh, width and height for measuring the volume of tumor.

3. RESULTS
During the classification process of the tumors types, a *MLP* has been used with two hidden layers only. In order to evaluate the classification efficiency, two metrics have been computed: (*a*) the training performance (i.e. the proportion of cases which are correctly classified in the training process) and (*b*) the testing performance (i.e. the proportion of cases which are correctly classified in the testing process). Basically, the testing performance provides the final check of the *NN* classification efficiency, and thus is interpreted as the diagnosis accuracy using the neural networks support. Recall that the testing performance, corresponding to the neural networks-based diagnosis accuracy, involves only cases with unknown diagnosis for the neural network classifier. This represents an alternative to traditional classification performance measures, such as sensitivity, specificity etc. and is directly related to this Artificial Intelligence technique.

Technically, 130cases are randomly selected used for training, 70 cases remaining for testing.

Firstly, notice that a testing performance of about 97.5% on average has been obtained (97.5% of cases are correctly classified in the testing process), together with a high training performance equaling about 98.5% on average (98.5% of cases are correctly classified in the training process). Thus, the neural networks-based diagnosis accuracy is about 98.15% on previously unknown cases, proving a good potential for this methodology. Usually, the training performance is higher than the testing performance, since the latter concerns unknown cases. In this case, the two measures are close enough (difference of about 1 percentage points only), indicating a low over-learning level, that is a balanced training/testing process. Moreover, the corresponding standard deviations equaling 2 and 3 percentage points respectively indicate a high stability of the model, especially in the training case. Note that, as training progresses, the training error naturally drops; it is desirable that the testing error should decrease as well as the training error as training progresses. If the difference between the two errors increases too much, this indicates that the network is starting to over-learn the data and thus it is applicable to other datasets anymore.

Secondly, the mean number of hidden processing units (*neurons*) in the network equals $13((5+13+5+3)/2=13)$, with a relative high standard deviation, equaling 3. This means that a two-hidden layer *MLP* with 10 neurons on average is able to provide 97.65% diagnosis accuracy. Moreover, it is possible to build a neural network model with 7 neurons only, that is a simple neural structure, and obtain a good enough accuracy. On the other hand, more than 16 neurons are not necessary to obtain a better classification. Thus, it can be concluded that it is possible to have a relative simple network structure (i.e. a small number of hidden neurons), that is a fast *NN* with a very good performance. Recall that an efficient neural computing solution to real-life problems implies the selection of the simplest *NN* architecture with high performance.

Performance measures

All classification result 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 the terms true and false positive and true and false negative as follows: [8, 29, 30]:

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.
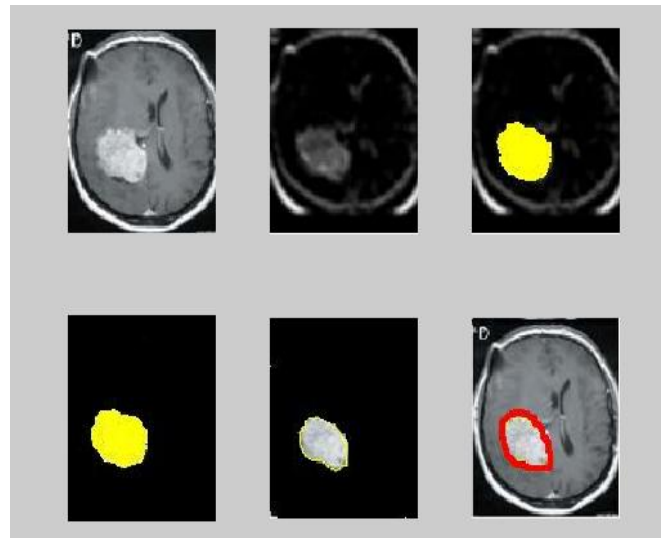
(a)
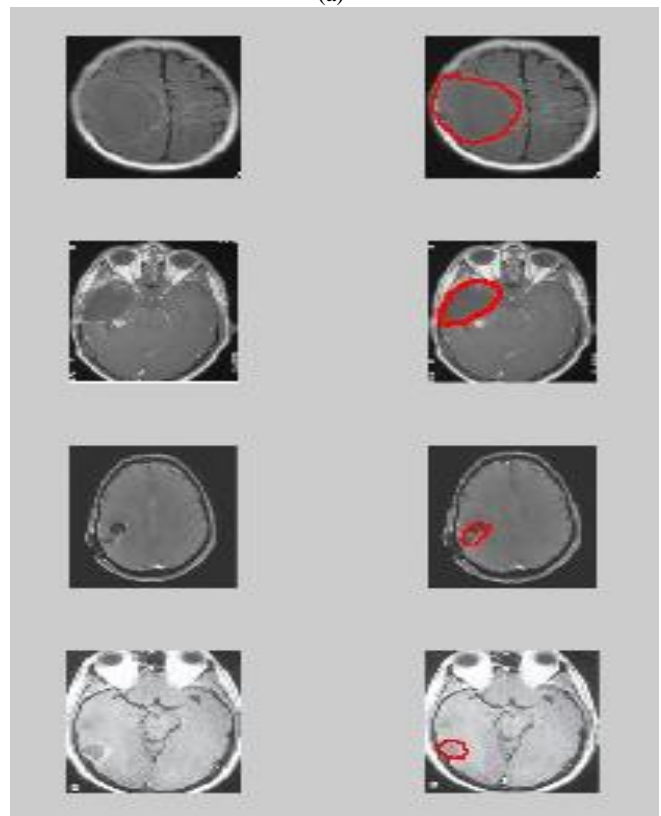
(b)

Figure 5: End results :a) from upper left to right : original image _difference image _after first stage _ image post processing _after second stage _output image   b) another images: left column is images with tumor, right column is output image by purposed method

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.

In this study because of measuring more real and robust results ,the method has been tested 10 times on all dataset and have computed the Performance measures as below:

Sensitivity = TP/ (TP+FN) *100%= 344/ (344+ 6) *100%=98.28%

Specificity = TN/ (TN+FP) *100%= 342/ (342+ 8) *100%=97.71%

Accuracy = (TP+TN)/ (TP+TN+FP+FN)*100 = (344+342)/ (344+342+ 8+ 6)*100 %= 98%

It has been presented qualitative comparison results of the segmented image too. The performance of the network is derived from the *MSE* value. No further improvement in the MSE is observed after epoch 100 as shown in (Figure 6), with a performance measure of 0.025.
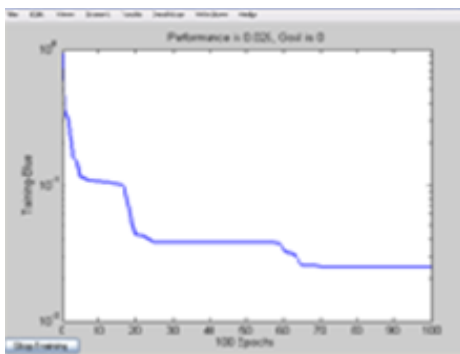

Figure 6: Method performance

4. CONCLUSION

Different types of elasticity imaging procedures have recently been described in the medical literature, with clinical applications already developed for the diagnosis of abnormalities. The methodology has been developed in this paper, is based on processing sample images of tumor and normal tissues, enables the exploration and analysis by automatic means of large quantities data from large number of patients. This provides a method which is an alternative to traditional human-based techniques, and optimally predicts the presence or absence of abnormality by using a noninvasive methodology.

This work has some limitations because of using all 3 modalities *T1 ،T2*_weighted and *PD ،MR* images. For working, though this can be justify by its high power, accuracy and yield in detecting each type of abnormalities.

The tasks that can derive from this work, include the integration of features derived from Fractal Analysis which describe Local Texture or Ruggedness in terms of an estimated value called *Hurst* Coefficient. These results are expected to be used in conjunction with Wavelet Multi resolution. Moreover, a classification performance analysis based on *ROC* curves is also needed to complete the study.

5. REFERENCES

[1] Ashby LS, Troester MM, Shapiro WR " Central Nervous System Tumors " Cancer Therapeutics, Vol.1, pp. 475-513, 2006.

[2] Cruickshank G, " Tumours of The Brain Surgery " Oxford, 22, pp.69-72 ,2004.

[3] Doolittle ND, " State of The Science in Brain Tumor Classification", Semin Oncol Nurs, vol.20, pp.224- 230, 2004.

[4] Chakeres DW, Schmalbrock P "Fundamentals of Magnetic Resonance Imaging " Williams & Wilkins, Baltimore ,1992.

[5] Buxton RB "Introduction to Functional Magnetic Resonance Imaging-Principles and Techniques". Cambridge University Press ,2002.

[6] Alan Wee, Chung Liew , Hong Yan, " Current Methods in the Automatic Tissue Segmentation of 3D Magnetic Resonance Brain Images," Current Medical Imaging Reviews, Vol.2, pp.1-13 ,2006.

[7] R. A. Lerski, K. Straughan, L. R. Schad, D.Boyce, S. Bluml, and I. Zuna, "MR ImageTexture Analysis- An approach To TissueCharacterization", Magnetic Resonance Imaging, Vol. 11, pp. 873-887,1993.

[8] H. Selvaraj1, S. Thamarai Selvi2, D. Selvathi3, L. Gewali1, " Brain MRI Slices Classification Using Least Squares Support Vector Machine " IC-MED, Vol.1, No. 1, Issue 1, Page 21 of 33 ,2007.

[9] Karteek Popuria, Dana Cobzasb, Martin Jagersandb, Sirish L. Shaha and Albert Murthac, "3D Variational Brain Tumor Segmentation on a Clustered Feature Set", SPIE,1-4 ,2009.

[10] Marcel Prastawa a, Elizabeth Bullitt c, Sean Ho a, Guido Gerig, "A Brain Tumor Segmentation Framework Based on Outlier Detection" Medical Image Analysis ,1-9 ,2004.

[11] Jiayin Zhou, Vincent Chong,Tuan-Kay Lim, Jing Huang, "MRI Tumor Segmentation for Nasopharyngeal Carcinoma Using Knowledge-based Fuzzy Clustering" International Journal of Information Technology Vol. 8, No. 2 ,September 2002.

[12] Jason J. Corso, Member, IEEE, Eitan Sharon, Shishir Dube, Suzie El-Saden, Usha Sinha, and Alan Yuille, "Efficient Multilevel Brain Tumor Segmentation with Integrated Bayesian Model Classification" IEEE Transactions on Medical Imaging,pp.1-7 ,2007.

[13] Dzung L. Pham, Chenyang XU, and Jerry L. Prince2 "Current methods in Medical Image Segmentation1" ,Biomed. Eng, vol.2 pp.315-337 ,2000.

[14] http://documents.wolfram.com

[15] Dzung L. Pham, Chenyang Xu, Jerry L. Prince;"A Survey of Current Methods in Medical Image Segmentation" Technical Report JHU / ECE 99-01, Department of Electrical and Computer Engineering. The Johns Hopkins University, Baltimore MD 21218 ,1998.

[16] 1Nathan Moon, 2Elizabeth Bullitt, 4Koen van Leemput, and 1;3Guido Gerig, " Automatic Brain and Tumor Segmentation", MICCAI2002,LNCS2488(I) pp. 372-379

[17] Rajeev Ratan A, Sanjay Sharma B, S. K. SharmaC, "Brain Tumor Detection Based on Multi-Parameter MRI Image Analysis" , ICGST-GVIP Journal, ISSN 1687-398X, Vol.9,Issue (III), June 2009.

[18] Carlos A. Parra, Khan Iftekharuddin and Robert Kozma, "Automated Brain Data Segmentation and Pattern Recognition Using ANN" , Computational Intelligence, Robotics and Autonomous Systems (CIRAS 03), December2003.

[19] Alan Wee-Chung Liew1 and Hong Yan, "Current Methods in the Automatic Tissue Segmentation of 3D Magnetic Resonance Brain Images" Current Medical Imaging Reviews, vol.2, 1-13 ,2006.

[20] Kemal Polat, Salih Gunes, "A New Feature Selection Method on Classification of Medical Datasets:Kernel F-score Feature Selection" , Expert Systems with Applications, vol.36 pp.10367–10373 ,2009.

[21] R.O. Duda, P.E. Hart, and D.G. Stork, "Pattern Classification", Wiley, 2nd edition ,2001.

[22] J. A. Freeman and D. M. Skapura, "Neural Networks, Algorithms, Applications and Programming Tech niques", Addison-Wesley Publishing Company,2002.

[23] C.M Bishop, "Neural Networks for Pattern Recognition", Oxford U.K: Clarendon ,1995.

[24] S. Haykin, "Neural networks: A comprehensive Foundation", 2nd ed. Englewood Cliffs, NJ: Prentice Hall ,1999

[25] F. M. Ham and I. Kostanic, "Principle of Neurocomputing for Science and Engineering" Tata McGraw-Hill publishing company Limited, New Delhi ,2002.

[26] F. Belloir, A. Fache, and A. Billat, "A General Approach to Construct RBF Net Based Classifier", Proceedings of the 7th European Symposium on Artificial Neural Network, pp. 399-404 ,1999.

[27] Y. S. Hwang and S. Y. Bang, "An Efficient Method to Construct a Radial Basis Function Neural Network Classifier", Neural Networks, vol. 10, no. 8, pp. 1495-1503 ,1997.

[28] C. J. C. Burges, "A tutorial on Support Vector Machines for Pattern Recognition", Data Mining and Knowledge Discovery,vol.2,no.2, pp.121–167 ,1998.

[29] R.M. Nishikawa, M.L. Giger, K. Doi, C.J. Vyborny, and R.A. Schmidt, "Computer Aided Detection of Clustered Microcalcifications in Digital Mammograms," Med. Biol. Eng. Comp. Vol. 33, pp. 174-178 ,1995.

[30] S. Theodoridis and K. Koutroumbas, "Pattern Recognition", Academic Press, San Diego, 1999.

[31] N.G. Raouf, Naguib, H. A. M. Sakim, M.S. Lakshmi, V. Wadehra, T. W. J. Lennard, J. Bhatavdekar, and G. V. Sherbat, "DNA Ploidy and Cell Cycle Distribution of Breast Cancer Aspirate Cells Measured by Image Cytometry and Analyzed by Artificial Neural Networks for their Prognostic Significance," IEEE Transaction on Information Technology in Biomedicine. Vol. 3, No. 1 ,1999.

[32] Alirezaie, J.; Jernagan, M.E.; Nahmias, C. " Automatic Segmentation of Cerebral MR Images Using Artificial Neural Networks" Nuclear Science Symposium (1996). Conference Record IEEE , Vol.3, pp.2-9 (1996). vol.3, pp. 1777 -1781 ,Nov. 1996.

[33] Alirezaie, J.; Jernagan, M.E.; Nahmias, C.; Neural Network Based Segmentation of Magnetic Resonance Images of the Brain" Nuclear Science Symposium and Medical Imaging Conference Record, (1995), IEEE , Vol.3 ,pp.21-28 (1995) vol.3 pp. 1397 -1401 ,Oct 1995.

[34] Sammouda, R.; Niki, N.; Nishitani, H. "Neural Networks Based Segmentation of Magnetic Resonance Images" Nuclear Science Symposium and Medical Imaging Conference(1994), IEEE Conference Record, Vol.4 (1994) vol.4, 1827 -1831 , 30 Oct , 5 Nov 1994.

[35] Porter, R.; Huckett, S.; Canagarajah, C.N "Optimal Feature Extraction for the Segmentation of Medical Image"s. Image Processing and Its Applications (1997)., Sixth International Conference on , Vol.2 , pp.14-17 vol.2 ,pp .814 -818 ,Jul 1997.

Amir Ehsan Lashkari is MSC Student of the Department of Bioelectrical Engineering, University of Tehran, Tehran, Iran,Asia. He received his BS from the Faculty of Bioelectric, Iran University of Isfahan in 2009 .His research interests are in: **D**igital Image Processing, **N**eural Networks, **P**attern Recognition, **I**ntelligent Systems, **M**achine Vision, **F**uzzy Logic, **C**haos and Fractal Phenomena, **D**igital Design, **A**rtificial Intelligent, **D**igital Signal Processing, **N**etwork and Path Planning, **M**ultiple Valued Function, **P**rogrammable Device.

He is top Student of university and he has published 3 books until now. The project he has done are: A Neural Network based Method for Brain Abnormality Detection in MR Images Using Gabor Wavelets, 2009/A Neural Network-Based Method for Brain Abnormality Detection in MR Images Using Zernike Moments and Geometric Moment Invariants, 2009/Full Automatic Micro Calcification Detection in Mammogram Images Using Artificial Neural Network and Gabor Wavelets, 2010/Automatic Gabor Face Detection by Neural Net work, 2010/Automatic Brain Tissue segmentation In MR Images by Neural Net work, 2009/Automatic Defect Detection in Steel Sheet in MOBARAKE STEEL FACTORY, Isfahan,2008/Automatic Label Reading in Cars, 2008/Automatic Growth Perediction of Tumors, By Inteligent Methods , 2008/Automatic Chromosome Obvolute in Microscopic Image By fuzzy Logic, 2009/Automatic Separation Chromosome Obvolute in Microscopic Image By fuzzy Logic, 2009/Parameter Optimization in Factory, 2010/Chaos and Fractal Signal Prediction In Biological System, 2010 and etc.