EL-SAYED A. EL-DAHSHAN, ABDEL-BADEEH M. SALEM, AND TAMER H. YOUNIS

ABSTRACT. This paper presents two hybrid techniques for the classification of the magnetic resonance human brain images. The proposed hybrid technique consists of three stages, namely, feature extraction, dimensionality reduction, and classification. In the first stage, we have obtained the features related with MRI images using discrete wavelet transformation (DWT). In the second stage, the features of magnetic resonance images (MRI) have been reduced using principles component analysis (PCA) to the more essential features. In the classification stage, two classifiers based on supervised machine learning have been developed. The first classifier based on feed forward back-propagation artificial neural network (FP-ANN) and the second classifier based on k-nearest neighbor (k-NN). The classifiers have been used to classify subjects as normal or abnormal MRI human images. A classification with a success of 95.6% and 98.6% has been obtained by the two proposed classifiers FP-ANN and k-NN respectively. This result shows that the proposed hybrid techniques are robust and effective compared with other recently work.

1. INTRODUCTION

Magnetic resonance imaging (MRI) is often the medical imaging method of choice when soft tissue delineation is necessary. This is especially true for any attempt to classify brain tissues [1]. The most important advantage of MR imaging is that it is non-invasive technique [2]. The use of computer technology in medical decision support is now widespread and pervasive across a wide range of medical area, such as cancer research, gastroenterology, hart diseases, brain tumors etc. [3, 4]. Fully automatic normal and diseased human brain

Received by the editors: March 1, 2009.

²⁰⁰⁰ Mathematics Subject Classification. 68T10, 62H30.

¹⁹⁹⁸ CR Categories and Descriptors. I.5.2 [Pattern Recognition]: Design Methodology – Pattern analysis; I.4.9 [Image Processing and Computer Vision]: Applications – General.

Key words and phrases. MRI human brain Images; Wavelet Transformation (WT); Principle Components Analysis (PCA); Feedforawd-Backpropagation Neural Network (FP-ANN); k-Nearest Neighbors, Classification.

56 EL-SAYED A. EL-DAHSHAN, ABDEL-BADEEH M. SALEM, AND TAMER H. YOUNIS

classification from magnetic resonance images (MRI) is of great importance for research and clinical studies.

Recent work [2, 5] has shown that classification of human brain in magnetic resonance (MR) images is possible via supervised techniques such as artificial neural networks and support vector machine (SVM) [2], and unsupervised classification techniques unsupervised such as self organization map (SOM) [2] and fuzzy c-means combined with feature extraction techniques [5]. Other supervised classification techniques, such as k-nearest neighbors (k-NN) also group pixels based on their similarities in each feature image [1, 6, 7, 8] can be used to classify the normal/pathological T2-wieghted MRI images. We used supervised machine learning algorithms (ANN and k-NN) to obtain the classification of images under two categories, either normal or abnormal.

Wavelet transform is an effective tool for feature extraction, because they allow analysis of images at various levels of resolution. This technique requires large storage and is computationally more expensive [4]. Hence an alternative method for dimension reduction scheme is used. In order to reduce the feature vector dimension and increase the discriminative power, the principal component analysis (PCA) has been used. Principal component analysis is appealing since it effectively reduces the dimensionality of the data and therefore reduces the computational cost of analyzing new data. To perform the classification on the input data the k-NN and artificial network classifier have been used.

The contribution of this paper is the integration of an efficient feature extraction tool and a robust classifier to perform a more robust and accurate automated MRI normal/abnormal brain images classification. Also, this paper focuses on a comparison of our results with a similar study (by others) using supervised and unsupervised methods [2, 5].

This paper is organized as follows: A short description on the input dataset of images is presented in Sections 2 and methods for feature extraction and reduction as well for classification are presented in Sections 3. Section 4 contains results and discussion while conclusions and future work are presented in Section 5.

2. The proposed hybrid techniques

The proposed hybrid techniques based on the following techniques, discrete wavelet transforms DWT, the principle components analysis PCA, FP-ANN, and k-NN. It consists of three stages: (1) feature extraction stage, (2) feature reduction stage, and (3) classification stage. The proposed hybrid technique for MRI image classification is illustrated in Fig. 1. In the following sections, a review of basic fundamental of k-NN, principal component analysis, and wavelet decomposition are introduced.



FIGURE 1. The methodology of the proposed technique

2.1. Feature extraction scheme using DWT. The proposed system uses the Discrete Wavelet Transform (DWT) coefficients as feature vector. The wavelet is a powerful mathematical tool for feature extraction, and has been used to extract the wavelet coefficient from MR images. Wavelets are localized basis functions, which are scaled and shifted versions of some fixed mother wavelets. The main advantage of wavelets is that they provide localized frequency information about a function of a signal, which is particularly beneficial for classification [9]. A review of basic fundamental of Wavelet Decomposition is introduced as follows:

The continuous wavelet transform of a signal $\mathbf{x}(t)$, square-integrable function, relative to a real-valued wavelet, $\Psi(t)$ is defined as [10]:

(1)
$$W_{\Psi}(a,b) = \int_{-\infty}^{\infty} f(x) * \Psi_{a,b}(t) dx$$

where

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}((t-a)/b)$$

and the wavelet $\Psi_{a,b}$ is computed from the mother wavelet Ψ by translation and dilation, wavelet, a the dilation factor and b the translation parameter (both being real positive numbers). Under some mild assumptions, the mother wavelet Ψ satisfies the constraint of having zero mean [11, 12]. The eq. (1) can be discretized by restraining a and b to a discrete lattice $(a = 2^b, a \in R_+, b \in R)$ to give the discrete wavelet transform (DWT).

The discrete wavelet transform (DWT) is a linear transformation that operates on a data vector whose length is an integer power of two, transforming it into a numerically different vector of the same length. It is a tool that separates data into different frequency components, and then studies each component with resolution matched to its scale. DWT can be expressed as [13].

(2)
$$DWT_{x(n)} = \begin{cases} d_{j,k} = \sum (x(n)h * j(n-2jk)) \\ d_{j,k} = \sum (x(n)g * j(n-2jk)) \end{cases}$$

The coefficients $d_{j,k}$ refer to the detail components in signal x(n) and correspond to the wavelet function, whereas $a_{j,k}$ refer to the approximation components in the signal. The functions h(n) and g(n) in the equation represent the coefficients of the high-pass and low-pass filters, respectively, whilst parameters j and k refer to wavelet scale and translation factors. The main feature of DWT is multiscale representation of function. By using the wavelets, given function can be analyzed at various levels of resolution [14]. Fig. 2 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 subband (LL, LH, HH, HL) images at each scale. (Fig.2). Subband image LL is used only for DWT calculation at the next scale. To compute the wavelet features in the first stage, the wavelet coefficients are calculated for the LL subband using Harr wavelet function.

2.2. Feature reduction scheme using PCA. One of the most common forms of dimensionality reduction is principal components analysis. Given a set of data, PCA finds the linear lower-dimensional representation of the data such that the variance of the reconstructed data is preserved [12, 15]. Using a system of feature reduction based on a combined principle component analysis on the feature vectors that calculated from the wavelets limiting the feature vectors to the component selected by the PCA should lead to an efficient classification algorithm utilizing supervised approach. So, the main idea behind using PCA in our approach is to reduce the dimensionality of the wavelet coefficients. This leads to more efficient and accurate classifier.

The following algorithm is used to find out the principal components of the input matrix to the neural network. Now the input matrix consists of



FIGURE 2. DWT schematically

only these principal components. The size of the input matrix is reduced from (1024) to (7). Algorithm 1 shows the involved steps for extracting principal components of the input vector to the two classifiers.

Therefore, the feature extraction process was carried out through two steps: firstly the wavelet coefficients were extracted by the DWT and then the essential coefficients have been selected by the PCA (see Fig.3.)



FIGURE 3. Schematic diagram for the used feature extraction and reduction scheme

3. Developing the supervised learning Classification

3.1. k-Nearest Neighbors based Classifier. One of the simplest classification techniques is the k- Nearest Neighbour classifier. Classification of an

60 EL-SAYED A. EL-DAHSHAN, ABDEL-BADEEH M. SALEM, AND TAMER H. YOUNIS

Algorithm 1 PCA algorithm

Let X be an input data set (X: matrix of dimensions M X N). Perform the following steps:

- 1: Calculate the empirical mean $u[m] = (1/N) \sum_{i=1}^{N} X[m, n]$.
- 2: Calculate the deviations from the mean and store the data in the matrix $B[M \ N]$:, B=X-u.h, where h is a 1 x N row vector of all 1's: h[n] =1 for n=1...N.
- 3: Find the covariance matrix C: $C = (1/N)B.B^*$.
- 4: Find the eigenvectors and eigenvalues of the covariance matrix $V^{-1}CV = D$, where V: the eigenvectors matrix, D: the diagonal matrix of eigenvalues of C, $D[p,q] = \lambda_m$ for p=q=m is the mth eigenvalues of the covariance matrix C.
- 5: Rearrange the eigenvectors and eigenvalues $\lambda_1 \ge \lambda_2 \ge \lambda_3 \ge \dots \ge \lambda_N$.
- 6: Choosing components and forming a feature vector. Save the first L columns of V as the M x L matrix W: W[p,q]=V[p,q] for n=1...M, q=1...L where $1 \le L \le M$.
- 7: Deriving the new data set The eigenvectors with the highest eigenvalues are projected into space, this projection results in a vector represented by fewer dimension (L < M) containing the essential coefficients.

input feature vector X is done by determining the k closest training vectors according to a suitable distance metric. The vector X is then assigned to that class to which the majority of those k nearest neighbours belong to [12, 16]. The k-NN algorithm is based on a distance function and a voting function in k nearest neighbors, the metric employed is the Euclidean distance. The knearest neighbor classifier is a conventional nonparametric supervised classifier that is said to yield good performance for optimal values of k [15]. Like most guided learning algorithms, k-NN algorithm consists of a training phase and a testing phase. In the training phase, data points are given in a n-dimensional space. These training data points have labels associated with them that designate their class. In the testing phase, unlabeled data are given and the algorithm generates the list of the k nearest (already classified) data points to the unlabeled point. The algorithm then returns the class of the majority of that list [15, 17]. Algorithm 2 describes the k-NN steps.

3.2. Artificial Neural Network based Classifier. An ANN is a mathematical model consisting of a number of highly interconnected processing elements organized into layers, geometry and functionality of which have been

Algorithm 2 k-NN algorithm

- 1: Determine a suitable distance metric.
- 2: In the training phase: Stores all the training data set P in pairs(according to the selected features) P = (yi, ci), i=1. ...n, where yi is a training pattern in the training data set, ci is its corresponding class and n is the amount of training patterns.
- 3: During the test phase: Computes the distances between the new feature vector and all the stored features (training data).
- 4: The k nearest neighbors are chosen and asked to vote for the class of the new example. The correct classification given in the test phase is used to assess the correctness of the algorithm. If this is not satisfactory, the k value can be tuned until a reasonable level of correctness is achieved.

resembled to that of the human brain. The ANN may be regarded as possessing learning capabilities in as much as it has a natural propensity for storing experimental knowledge and making it available for later use [18].

The neural network which was employed as the classifier required in this study had three layers, as shown in Fig (4). The first layer consisted of 7 input elements in accordance with the 7 feature vectors that selected from the wavelet coefficients by the PCA. The number of neurons in the hidden layer was four. The single neuron in the output layer was used to represent normal and abnormal human brain (see Fig. 4).



FIGURE 4. The architecture of used ANN

62 EL-SAYED A. EL-DAHSHAN, ABDEL-BADEEH M. SALEM, AND TAMER H. YOUNIS

The most frequently used training algorithm in classification problems is the back-propagation (BP) algorithm, which is used in this work also. The details of back-propagation (BP) algorithm are well documented in the literature [18].

The neural network has been trained to adjust the connection weights and biases in order to produce the desired mapping. At the training stage, the feature vectors are applied as input to the network and the network adjusts its variable parameters, the weights and biases, to capture the relationship between the input patterns and outputs [18].

4. Case study

In this section, the proposed hybrid techniques have been implemented on a real human brain MRI dataset. All the input dataset (total images is 70: 60 images are abnormal and 10 normal) used for classification consists of axial, T2-weighted, 256 -256 pixel MR brain images, 60. These images were collected from the Harvard Medical School website (http:// med.harvard.edu/AANLIB/) [19]. Fig.5 shows some samples from the used data for normal and pathological brain: a- normal, b- Glioma, c- Metastatic bronchogenic carcinoma, d-Alzheimer's disease, visual agnosia.



FIGURE 5. Samples from the used data

The algorithm described in this paper is developed locally and successfully trained in MATLAB version 7.1 using a combination of the Image Processing Toolbox, Wavelet toolbox (The MathWorks) for MATLAB. We performed all the computations of DWT+PCA+FP-ANN and DWT+PCA+k-NN classification on a personal computer with 1.5 MHz Pentium IV processor and 384 MB of memory, running under Windows-2000TM operating system. The programs can be run/tested on many different computer platforms where MATLAB is available. Algorithm 3 depicts the steps of the proposed two classifiers.

Algorithm 3 Pseudocode of the used hybrid techniques

Input: 256x256 brain images Parameters: N is number of images

Stage (1): Features Extraction using DWT FOR Loop on i=1 to N

Read the input images, Resize the images, and apply the DWT for the 3rd level using "Haar" family to extract the wavelet coefficients, Put the wavelet coefficients in a matrix X [MxN]

End Loop of i

ENDFOR

Repeat the above loop for the test image to extract its wavelet coefficients. Concatenate the feature coefficients of the training images and the test image in one matrix.

Stage (2): PCA Features reduction Loop on J =1 to N Apply PCA transformation (according to algorithm 1) on the obtained wavelet coefficients. Put the new dataset in a matrix Y End Loop on j

Stage (3): Classification Using two supervised techniques: Classifier 1: based on ANN

Create the design of neural network with feed forward back-propagation algorithm. Create target vector. Train the net with the selected dataset and the desired target. Input the Features of test image on. Trained the neural network.Classify it.

Output: Normal or abnormal brain.

Classifier 2: based on k-NN Loop for g =1 to 5, where 5= k-nn level. For j=1 to N Apply the k-NN algorithm PCA. END Loop j END Loop on g Classify test image. Output: Normal or abnormal brain

5. Results and discussions

In this section, we present the performance evaluation methods used to evaluate the proposed approaches. Finally, we will show the experimental results and examine the performance of the proposed classifiers for the MRI dataset mentioned above. We evaluate the performance of the proposed method in terms of confusion matrix, sensitivity, specificity and accuracy. The three terms are defined as follows [20]:

Sensitivity (true positive fraction) is the probability that a diagnostic test is positive, given that the person has the disease.

(3)
$$Sensitivity = \frac{TP}{TP + FN}$$

Specificity (true negative fraction) is the probability that a diagnostic test is negative, given that the person does not have the disease.

(4)
$$Specificity = \frac{TN}{TN + FP}$$

Accuracy is the probability that a diagnostic test is correctly performed.

(5)
$$Accurrcy = \frac{TP + TN}{TP + TN + FP + FN}$$

Where:

TP (True Positives) - Correctly classified positive cases,

TN (True Negative)- Correctly classified negative cases,

FP (False Positives) -Incorrectly classified negative cases, and

FN (False Negative) - Incorrectly classified positive cases.

Table 1 shows the classification rates for performing the proposed hybrid approach. In this experiment two classifiers based on supervised machine learning are presented for MRI normal/abnormal human brain classification. In the proposed methods using DWT, first three levels coefficients of decomposition of MR images with Harr as mother wavelet are computed. These coefficients are used for feature extraction. PCA is used for feature selection and NN and k-NN classifiers for MRI normal/abnormal human brain classification are used in methods 1 and 2, respectively.

For reducing the complexity of the system, PCA was used for feature reduction which was described in Section 3. The dimension of the feature vector was reduced from 1024 to 7 with the PCA algorithm. Limiting the feature vectors to the component selected by the PCA leads to an increase in accuracy rates and decreases the time complexity. In this experimental, MRI dataset that have healthy and diseased brain are classified by the proposed classifiers. The experimental results of the proposed classifiers are compared in the Table 1, which shows the percentage classification for the two different image classes. The analysis of the experimental results shows that classification accuracy 97% is achieved with the FP-ANN classifier and classification accuracy 98

To evaluate the effectiveness of our methods we compare our results with recently results [2, 5] for the same MRI datasets. Table 2 gives the classification accuracies of our method and the recently results. This comparison shows that our system has high classification accuracy and less computation due to the feature reduction based on the PCA.

6. Conclusions and Future works

In this study, we have developed a medical decision support system with normal and abnormal classes. The medical decision making system designed by the wavelet transform, principal component analysis (PCA), and supervised learning methods (FP-ANN and k-NN) that we have built gave very promising results in classifying the healthy and patient brain. The benefit of the system is to assist the physician to make the final decision without hesitation.

According to the experimental results, the proposed method is efficient for classification of the human brain, normal or abnormal. Our proposal produced 95.9% sensitivity rate and 96% specificity rate for FP-ANN classifier and 96% sensitivity rate and 97% specificity rate for k-NN classifier. SOM and SVM [2, 5] produced the similar results. ANN method gained the worst sensitivity and specificity rate.

Our results have compared to the results reported very recently based on the same T2-wieghted MRI database. Our method can be employed for all types of MR images T1-wieghted, T2-wieghted, and proton density (T1- T2-PD). This research developed two hybrid techniques, DWT+PCA+FP-ANN and DWT+PCA+k-NN to classify the human brain MR images. The stated results show that the proposed method can make an accurate and robust classifier. The classification performances of this study show the advantages of this technique: it is easy to operate, noninvasive and inexpensive. The limitation of this work is that it requires fresh training each time whenever there is an increase in image database. The extension of the developed techniques for processing the pathological brain tissues (e.g. lesions, tumors) is the topic of future research.

Hybrid technique	TP	TN	FP	FN	Sensi-	Speci-	Accu-
					tivity	ficity	rracy
					(%)	(%)	(%)
DWT+PCA+ANN	58	9	2	1	98.3	81.8	95.7
DWT+PCA+k-NN	60	9	1	0	98.4	100	98.6

TABLE 1. Classification rates for the used classifiers

66 EL-SAYED A. EL-DAHSHAN, ABDEL-BADEEH M. SALEM, AND TAMER H. YOUNIS

TABLE 2. Classification performance (P) comparisons for the proposed technique and the recently works for the same MR images datasets.

Technique	P(%)
Our hybrid technique DWT+PCA+ANN	95.7
Our hybrid technique $DWT + PCA + k-NN$	98.6
DWT+SOM [2]	94
DWT+SVM with linear kernel [2]	96.15
DWT+SVM with radial basis function based kernel [2]	98

References

- L. M. Fletcher-Heath, L. O. Hall, D. B. Goldgof, F. R. Murtagh; Automatic segmentation of non-enhancing brain tumors in magnetic resonance images; Artificial Intelligence in Medicine 21 (2001), pp. 43-63.
- [2] Sandeep Chaplot, L.M. Patnaik, N.R. Jagannathan; Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network; Biomedical Signal Processing and Control 1 (2006), pp. 86-92.
- [3] F. Gorunescu; Data Mining Techniques in Computer-Aided Diagnosis: Non-Invasive Cancer Detection; PWASET Volume 25 November 2007 ISSN 1307-6884, PP. 427-430.
- [4] S. Kara, F. Dirgenali; A system to diagnose atherosclerosis via wavelet transforms, principal component analysis and artificial neural networks; Expert Systems with Applications 32 (2007), pp. 632-640.
- [5] M. Maitra, A. Chatterjee; Hybrid multiresolution Slantlet transform and fuzzy cmeans clustering approach for normal-pathological brain MR image segregation, Med Eng Phys (2007), doi:10.1016/j.medengphy.2007.06.009.
- [6] P. Abdolmaleki, Futoshi Mihara, Kouji Masuda, Lawrence Danso Buadu; Neural networks analysis of astrocytic gliomas from MRI appearances' Cancer Letters 118 (1997), pp. 69-78.
- [7] T. Rosenbaum, Volkher Engelbrecht, Wilfried Kro?lls, Ferdinand A. van Dorstenc, Mathias Hoehn-Berlagec, Hans-Gerd Lenard; MRI abnormalities in neurofibromatosis type 1 (NF1): a study of men and mice; Brain & Development 21 (1999), pp. 268-273.
- [8] C. Cocosco, Alex P. Zijdenbos, Alan C. Evans; A fully automatic and robust brain MRI tissue classification method; Medical Image Analysis 7 (2003), pp. 513-527.
- [9] K. Karibasappa, S. Patnaik, "Face Recognition by ANN using Wavelet Transform Coefficients", IE(I) Journal-CP, 85,(2004), pp. 17-23, .
- [10] P.S.Hiremath, S. Shivashankar, and Jagadeesh Pujari; Wavelet Based Features for Color Texture Classification with Application to CBIR; IJCSNS International Journal of Computer Science and Network Security, 6 (2006), pp. 124-133.
- [11] I. Daubechies, "Ten Lectures on Wavelets," Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 1992.
- [12] A. Sengur, An expert system based on principal component analysis, artificial immune system and fuzzy k-NN for diagnosis of valvular heart diseases Comp. Biol. Med. (2007), doi: 10.1016/j.compbiomed.2007.11.004.

- [13] D. Bouchaffra, J. Tan, Structural hidden Markov models for biometrics: Fusion of face and Fingerprint; Pattern Recognition 41 (2008), pp. 852 - 867.
- [14] M. Kocionek, A. Materka, M. Strzelecki P. Szczypinski Discrete wavelet transform derived features for digital image texture analysis, Proc. of Interational Conference on Signals and Electronic Systems, 18-21 September 2001, Lodz, Poland, pp. 163-168
- [15] R. O. Duda, P. E. Hart, and D. G. Stork, Pattern Classi cation, New York: John Wiley and Sons, Inc., 2001.
- [16] F. Latifoglu, K. Polat, S. Kara, S. Gunes; Medical diagnosis of atherosclerosis from Carotid Artery Doppler Signals using principal component analysis (PCA), k-NN based weighting pre-processing and Artificial Immune Recognition System (AIRS); Journal of Biomedical Informatics 41 (2008), pp. 15-23.
- [17] O'Farrell, M., 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" Sensors and Actuators B: Chemical, Vol 111-112C (2005), pp. 354-362.
- [18] S. Haykin, Neural Networks: A comprehensive Foundation, Prentice Hall, 1999.
- [19] Harvard Medical School, Web: data available at http://med.harvard.edu/ AANLIB/.
- [20] Kemal Polat, Bayram Akdemir, Salih Gnes; Computer aided diagnosis of ECG data on the least square support vector machine; Digital Signal Processing 18 (2008), pp. 25-32.

Physics Department, Faculty of Science, Ain Shams University, Abbassia, Cairo 11566, Egypt

E-mail address: e_eldahshan@yahoo.com

FACULTY OF COMPUTER AND INFORMATION SCIENCE, AIN SHAMS UNIVERSITY, AB-BASSIA, CAIRO, EGYPT

E-mail address: absalem@asunet.shams.edu.eg and abmsalem@yahoo.com

FACULTY OF ENGINEERING, MISR UNIVERSITY FOR SCIENCE AND TECHNOLOGY, 6TH OCTOBER CITY, CAIRO, EGYPT

E-mail address: tyounis@must.edu.eg