Multi class disorder detection of magnetic resonance brain images using composite features and neural network

Vandana V. Kale¹ · Satish T. Hamde² · Raghunath S. Holambe²

Received: 27 November 2018 / Revised: 16 January 2019 / Accepted: 7 February 2019 / Published online: 4 March 2019
© Korean Society of Medical and Biological Engineering 2019

Abstract

Brain disorder recognition has becoming a promising area of study. In reality, some disorders share similar features and signs, making the task of diagnosis and treatment challenging. This paper presents a rigorous and robust computer aided diagnosis system for the detection of multiple brain abnormalities which can assist physicians in the diagnosis and treatment of brain diseases. In this system, we used energy of wavelet sub bands, textural features of gray level co-occurrence matrix and intensity feature of MR brain images. These features are ranked using Wilcoxon test. The composite features are classified using back propagation neural network. Bayesian regulation is adopted to find the optimal weights of neural network. The experimentation is carried out on datasets DS-90 and DS-310 of Harvard Medical School. To enhance the generalization capability of the network, fivefold stratified cross validation technique is used. The proposed system yields multi class disease classification accuracy of 100% in differentiating 90 MR brain images into 18 classes and 97.81% in differentiating 310 MR brain images into 6 classes. The experimental results reveal that the composite features along with BPNN classifier create a competent and reliable system for the identification of multiple brain disorders which can be used in clinical applications. The Wilcoxon test outcome demonstrates that standard deviation feature along with energies of approximate and vertical sub bands of level 7 contribute the most in achieving enhanced multi class classification performance results.

Keywords Magnetic resonance imaging · Discrete wavelet transform · Gray level co-occurrence matrix · Back propagation neural network · Multi class classification

1 Introduction

Magnetic resonance imaging (MRI) technique uses magnetic field and radio frequency (RF) signals to produce the image of anatomical structure which shows the presence of disease. It is non-invasive technique and does not use harmful radiations. It provides superior contrast for different brain tissues

compared to other imaging modalities [1]. These properties of MRI are more useful for brain pathology diagnosis and treatment [2]. However, the amount of MR imaging data is voluminous for manual interpretation, hence computer aided detection (CAD) systems are emerging to reduce the manual workload thus improving the diagnostic accuracy of radiologists [3, 4].

To classify MR brain images, researchers have proposed various useful techniques. El-Dahshan et al. [5] used discrete wavelet transform (DWT) as feature extraction and principal component analysis (PCA) as feature reduction tool for MRI images. They attained classification accuracy of 98% using k-nearest neighbour (k-NN) as classifier. Das et al. [6] utilised combination of ripplelet transform (RT) and multi scale geometric analysis (MGA) to extract features from images. Using least square-SVM (LS-SVM) classifier, they achieved average classification accuracy greater than 99%. Saritha et al. [7] proposed a new methodology using spider web plot areas as features and probabilistic neural network as classifier. They achieved 100% binary classification

✉ Vandana V. Kale
vandanav_kale@yahoo.co.in

Satish T. Hamde
sthamde@sggs.ac.in

Raghunath S. Holambe
rsholambe@sggs.ac.in

¹ Department of Instrumentation Engineering,
AISSMS's Institute of Information Technology, Pune,
Maharashtra 411001, India

² Department of Instrumentation Engineering, SGGS
Institute of Engineering and Technology, Nanded,
Maharashtra 431606, India

accuracy. El-Dahshan et al. [8] suggested a method using feedback pulse coupled neural network (FPCNN) for MR image segmentation and the discrete wavelet transform (DWT) for features extraction. Yang et al. [9] introduced biogeography based optimization (BBO) method to optimize the weights of the classifier. Using wavelet energy as feature, they achieved binary classification accuracy of 97.78%. Zhang et al. [10] employed entropy of approximation wavelet coefficients as feature. They trained neural network classifier with hybridized BBO and PSO method and obtained significant accuracy. Zhang et al. [11] calculated Shannon entropy (SE) and Tsallis entropy (TE) of extracted wavelet packet coefficients of MR images. They used different combinations of SVM classifier and entropies to test four diagnosis methods. Nayak et al. [12] used probabilistic principal component analysis (PPCA) to reduce approximation wavelet coefficients from MR brain images. Adaboost with random forests (ADBRF) was used for binary classification. Nayak et al. [13] suggested a system using fast discrete curvelet transform (FDCT), principal component analysis (PCA) and least squares SVM (LS-SVM) with three different kernels to perform binary classification. Zhang et al. [14] suggested a novel system employing fourier entropy (FRFE) as feature and multi-layer perceptron (MLP) as classifier. They obtained average classification accuracy of 99.53% using neural network classifier. Gudigar et al. [15] utilised multi-resolution analysis techniques along with SVM classifier to classify MR brain images and they obtained 97.38% classification accuracy using shearlet coefficients. Authors [16], extended the concept of binary classification to multi class disease classification. They employed DWT and PCA as feature extraction and reduction tool respectively. Using different classification models, they perform multi disease classification and achieved 95.70% accuracy. In [17], authors utilised deep neural network to classify multiple brain diseases using DWT coefficients as feature.

The literature review reveals that the most of the research has been undertaken on binary classification. The widely used feature extraction methods include discrete wavelet transform and its variants. The classification features involved are mostly wavelet coefficients, multi resolution analysis techniques' coefficients, energy and entropy of these coefficients. For classification, most of the systems employed SVM, neural network and their variants. Using these hybrid systems, the researchers achieved promising accuracies in binary classification. However, the research should be explore in the field of multi disease classification, as it is challenging task because of sharing of similar features by many brain disorders. Along with accuracy other performance measures namely sensitivity, specificity of CAD systems should be enhanced as they assist physicians in diagnosis and treatment of MR brain diseases [4]. To address these problems, in this work, instead of using

classification features individually, composite features are used, as different features extract different information from image and it results in more competent feature set. The classification of this feature set is carried out using BPNN in which optimization, training functions, stopping criteria are chosen not to over fit the network and to model the training data with 100% accuracy. Also, Wilcoxon test is used to rank the features as per their discriminative capacity in disease classification. The contributions of this work can be outlined as follow.

- The proposed system has utilised the potential of composite features namely energy of wavelet coefficients, GLCM features, intensity feature to classify multiple brain abnormalities.
- The Wilcoxon test is used to find out ranks of features as per their discriminative power in differentiating multiple diseases.
- Rigorous experimentations are carried out on BPNN classifier to prove the competency and efficiency of the proposed method.

This paper is organized as follows: Materials are presented in Sect. 2. Section 3 leads to methodology of proposed system. The experimental results with discussion are delineated in Sect. 4. The work is concluded in Sect. 5.

2 Materials

We worked on datasets DS-90 and DS-310 containing 90 and 310 T2-weighted MR brain images respectively obtained from the website of Harvard Medical School [18]. The images are of size 256×256 in-plane resolution in axial plane. The dataset DS-90 consists of 17 classes of diseased and 1 class of normal brain images. DS-310 consists of 5 classes of diseased and 1 class of normal brain images. Sample images of diseased and normal MR brain are shown in Fig. 1.

3 Methodology

The proposed system consists of major components, feature extraction and classification. Figure 2 depicts the overall block diagram of the proposed system. MR images are decomposed to level-7 using DWT to form 22 sub bands. Wavelet Energy of each sub band, four GLCM features and one intensity feature form feature matrix of MR images. Features are classified using BPNN and their rank is decided by Wilcoxon test. All these components are discussed below.

Fig. 1 Sample of brain MRIs: **a** Alzheimer, **b** Alzheimer with visual agnosia, **c** cerebral calcinosi, **d** Huntington, **e** motor neuron, **f** Picks disease, **g** AIDS dementia, **h** cerebral toxoplasmosis, **i** Creutzfeldt–

Jakob, **j** herpes encephalitis, **k** lyme encephalopathy, **l** multiple sclerosis, **m** glioma, **n** meningioma, **o** metastatic adenocarcinoma, **p** metastatic bronchogenic carcinoma, **q** sarcoma, **r** normal brain

Fig. 2 Block diagram of proposed CAD methodology

3.1 Feature extraction

The feature set consists of energy of wavelet coefficients, GLCM textural features and intensity feature of MR images. The extracted GLCM textural features are contrast, correlation, energy and homogeneity. Standard Deviation is used as intensity feature.

3.1.1 Energy of wavelet sub bands

Wavelets have emerged as powerful mathematical tools. Wavelet transform provides localized frequency information of data. It can analyses the data at multiple scales [19]. These features of wavelet transform are used by many researchers for feature extraction [20]. Wavelets decomposes image into sub bands. The energy of sub band is computed using following expression.

$$E_k = \sum_{i=1}^M \sum_{j=1}^N (x_k(i, j))^2 \quad (1)$$

where E_k is the energy of the k th sub band and k varies from 1 to 7. $x_k(i, j)$ is the corresponding sub band with dimension $M \times N$.

The energy reveals strength of image information in respective sub band. In this work, MR image is decomposed to level 7 using db4 wavelet resulting into 22 sub bands. Energies of 22 sub bands are considered as features.

3.1.2 Gray level co occurrence matrix (GLCM)

Haralick et al. [21] have introduced GLCM approach in 1973. Researchers have widely used the GLCM approach for texture classification. In this work, four textural features of GLCM namely contrast, correlation, energy and homogeneity are calculated. Let $P(i, j)$ is the GLCM of an image I . Each element (i, j) in GLCM specifies the number of times, the pixel of image with gray level value i occurred horizontally adjacent to a pixel with value j .

In this work, the direction of neighbouring pixel with respect to pixel of reference is set as an angle 0° and the distance is set as 1. Using these parameters, the adjacent pixels get select and local details of abnormalities in image can be detected effectively.

Contrast: This property of image is measure of intensity contrast between pixel and its neighbour over whole image. It is calculated as,

$$\text{Contrast} = \sum_{i,j} |i - j|^2 p(i, j) \quad (2)$$

Correlation: Correlation indicates how pixel is correlated with its neighbour over the whole image. It is given by equation,

$$\text{Correlation} = \sum_{i,j} \frac{(i - \mu_i)(j - \mu_j)p(i, j)}{\sigma_i \sigma_j} \quad (3)$$

where, μ indicates mean and σ indicates standard deviation.

Energy: Energy of image is the sum of squared elements of GLCM.

$$\text{Energy} = \sum_{i,j} (p(i, j))^2 \quad (4)$$

Homogeneity: Homogeneity measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. It shows uniformity in intensity values.

$$\text{Homogeneity} = \sum_{i,j} \frac{p(i, j)}{1 + |i - j|} \quad (5)$$

3.1.3 Intensity based feature

Standard Deviation: Standard deviation is measure of spread of gray level around the mean. It is given by following equation.

$$\text{StandardDeviation} = \sqrt{\frac{\sum_i \sum_j (I(i, j) - \mu)^2}{N}} \quad (6)$$

where, $I(i, j)$ is image matrix, μ indicates mean and N is total number of pixels in image.

The feature matrix consists of 27 features of all images which is given as input to BPNN classifier.

3.2 Wilcoxon test

The Wilcoxon test is a non parametric test for independent samples [22, 23]. Let, X and Y are independent samples with different sample sizes n_x and n_y respectively. Then the rank sum of sample is calculated using Wilcoxon test. The Wilcoxon test is equivalent to the mann whitney U-test. U is the

number of times y precedes x in an ordered arrangement of the elements in samples X and Y. In the following way, mann whitney U-test is related to the Wilcoxon rank sum statistic,

$$U_x = n_x n_y + \frac{n_x(n_x + 1)}{2} - T_x \quad (7)$$

where T_x , is sum of ranks for sample X with size n_x ; U_x , is U value based on sample size n_x .

The U statistic of Wilcoxon test is used to rank classification features. Out of 27 classification features used in this work, extracted feature with highest discriminative power has highest U value and highest feature rank. Feature ranking decides the discriminative capacity of feature.

3.3 Classification based on back propagation neural network (BPNN)

Neural networks (NNs) are biologically inspired [24]. NNs can learn from previous experience and generalize using past examples to new ones. They have achieved considerable success when applied to solve difficult, multivariate, non linear, diverse problems and noisy domains [24, 25]. Ojha et al. [26] has given excellent review on feed forward neural network.

BPNN with single hidden layer and sufficient number of hidden neurons can approximate virtually any function to any degree of accuracy [27, 28]. The selection of optimal number of neurons in hidden layer is an important issue in the design of BPNN. A network with few neurons may not capture the input and output patterns accurately. If network is having excess of hidden neurons then it can lead to poor generalization and over fit the training data. To overcome over fitting issue of NN, bayesian regularization algorithm is used in this work. This training function provides a measure of how many weights and biases out of total network parameters are being effectively used by the network, thus avoiding the over fitting issue of network [27]. In this work, hyperbolic tangent sigmoid transfer function is used in hidden layer and output layer as activation function. It ranges from -1 to $+1$ so, it can prevent training difficulties caused by saturation of the sigmoid functions [27]. It is used in NN where speed is important [22]. The BPNN architecture details are tabulated in Table 1. The structure of two layer network with X_D input neurons, N_H hidden neurons and Y_K output neurons is depicted in Fig. 3. For simplicity biases are not shown in the figure. The overall network function is given by following equation,

$$Y_k = h_2 \left(\sum_{j=0}^M W_{kj}^{(2)} \left(h_1 \left(\sum_{i=0}^D W_{ji}^{(1)} x_i \right) \right) \right) \quad (8)$$

where x_i is input variable and $i = 1, 2, \dots, D$; Y_K , is output variable and K is total number of outputs; $W^{(1)}$ and $W^{(2)}$ represents weight parameters of first and second layer of network

Table 1 BPNN parameters

Parameters	BPNN
Output layer activation function	Hyperbolic tangent sigmoid
Hidden layer activation function	Hyperbolic tangent sigmoid
Performance function	Mean squared normalized error
Optimization method	Levenberg–Marquardt algorithm
Training function	Bayesian regulation backpropagation
Minimum performance gradient	1×10^{-10}
Stopping criteria	Minimum gradient

Fig. 3 Network diagram for two layer Neural Network

respectively; M , is total number of parameters of network; $h_1(\cdot)$ and $h_2(\cdot)$, are activation functions of first and second layer of network respectively.

The training phase of the network progresses by adaptively adjusting the values of weights and biases of the network based on the performance function, given by following formula,

$$E = \beta \frac{1}{2} \sum_{n=1}^K (Y_n - t_n)^2 + \alpha \frac{1}{N} \sum_{j=1}^N W_j^2 \quad (9)$$

where Y_n , is the output produced by network; t_n , is the target output; K , is total number of input-target pairs; α and β , are performance function parameters; N , is total number of network weights and biases; W_j , is network weights and biases.

The performance function causes the network to have smaller weights and biases and this forces the network

response to be smoother and less likely to over fit, improving generalization ability of network. The key point is optimal selection of parameter α and β . This is achieved using bayesian regularization. When the value of E gets minimized, the training phase of network stops and the network represents optimal values of weights and biases which are further used for testing phase. In this work, BPNN with single hidden layer and 26 hidden neurons are used in case of dataset DS-90. Dataset DS-310 is having more images resulting in large number of training equations so, two hidden layers with 25 and 10 hidden neurons in first and second layer respectively are used to train BPNN model. The number of hidden neurons are decided by performing experimentation.

4 Experimental results with discussion

To validate the proposed scheme, experimentation has been carried out using combination of wavelet and neural network toolboxes of MATLAB 2013a, running under windows 7 operating system with 1.60 GHz, dual core processor and 2 GB RAM.

4.1 Cross validation setting

Stratified cross validation (SCV) is commonly used for rearranging the dataset into different folds so that each fold fully represents the whole dataset. It helps classifier to be more generalize [27]. Running SCV more times gives more accurate results.

In this work, we have used 5-fold stratified cross validation technique to arrange MR brain images into 5 folds. One of the folds used for testing and remaining 4 folds used for training. The process is repeated until each fold is used for testing. The setting of training and test images is shown in Table 2. The 5-fold SCV has repeated for 75 runs with the aim of improving generalization ability of the network [12, 13, 15] which results in reliable and robust system.

4.2 Feature ranking using Wilcoxon test

The absolute value of U-statistic of Wilcoxon test is used to rank the features. The values are plotted as shown in

Table 2 Training and test image setting

Dataset	Total images	Total diseases	Total classes	k-fold SCV	Total training images	Total testing images
DS-90	90	17	18	5	72	18
DS-310	310	5	6	5	248	62

Fig. 4. Feature 1 is the energy of approximation wavelet coefficients of level 7. Features 2 to 22 represent energies of horizontal, vertical and diagonal sub bands of levels 1 to 7 respectively. Features 23 to 26 are GLCM extracted four features namely correlation, homogeneity, contrast and energy. Feature 27 represents intensity feature namely standard deviation. The graph demonstrates that the standard deviation is the most significant feature for classification. Energies of approximate and vertical sub bands of level-7 are the 2nd and 3rd ranked discriminative features. The textural features contrast, correlation and energy have major role in classification. In summary, the combination of wavelet, GLCM texture and intensity feature of MR images lead to enhance the diagnostic accuracy of brain diseases.

4.3 Comparison with existing methods

To evaluate the performance of proposed multi class disease classifier, different classification methods are compared on Harvard Medical School brain dataset. Table 3

presents the comparative results of multi class disease classification. Authors [16], used various methods to differentiate 310 images into 6 brain image classes. Initially, (DWT + PCA + kNN), (RT + PCA + LS-SVM (RBF)), (DWPT + GEPSVM), these methods were proposed by

Fig. 5 Performance measures of existing binary classification methods on Harvard Medical School brain dataset DS-90

Fig. 4 The absolute value of U-statistic of Wilcoxon test of features

Table 3 Multi class disease classification accuracy of existing methods on Harvard Medical School brain dataset

Paper	Method	Number of images	Number of brain diseases	Number of classes to distinguish	Accuracy (%)
[16]	DWT + PCA + kNN	310	5	6	83.87
[16]	RT + PCA + LS-SVM (RBF)	310	5	6	86.02
[16]	DWPT + GEPSVM	310	5	6	88.92
[16]	DWT + PCA + RF	310	5	6	95.70
[17]	DWT + PCA + Deep NN	66	3	4	96.97
Proposed method	Composite features + BPNN	90	17	18	100
		310	5	6	97.81

Table 4 Training and test results of 10 runs out of 75 runs of dataset DS-90 (training, test sets and weight-biases are different in each fold)

Run	Fold	Training performance error	Testing performance error	Network performance error	Epochs	Effective parameters	Actual neurons	Approximate neurons	Classification accuracy (%)
1	1	1.8470×10^{-09}	1.1322×10^{-03}	2.2644×10^{-04}	28	258	5.22	6	100
	2	5.2184×10^{-09}	5.1575×10^{-06}	1.0357×10^{-06}	25	275	5.59	6	100
	3	2.6690×10^{-09}	1.5899×10^{-03}	3.1798×10^{-04}	30	275	5.59	6	100
	4	2.5154×10^{-09}	6.8949×10^{-06}	1.3810×10^{-06}	35	256	5.17	6	100
	5	3.0968×10^{-09}	1.8870×10^{-03}	3.7741×10^{-04}	29	260	5.26	6	100
2	1	7.7201×10^{-09}	3.5996×10^{-04}	7.1998×10^{-05}	25	275	5.59	6	100
	2	2.6332×10^{-09}	1.8072×10^{-03}	3.6144×10^{-04}	25	273	5.54	6	100
	3	2.4845×10^{-09}	2.4095×10^{-04}	4.8192×10^{-05}	29	273	5.54	6	100
	4	6.3178×10^{-10}	1.3842×10^{-04}	2.7684×10^{-05}	30	266	5.39	6	100
	5	3.1226×10^{-09}	8.1992×10^{-05}	1.6401×10^{-05}	32	255	5.15	6	100
3	1	2.4792×10^{-09}	6.9302×10^{-04}	1.3861×10^{-04}	28	268	5.43	6	100
	2	5.8932×10^{-09}	5.3566×10^{-05}	1.0718×10^{-05}	30	268	5.43	6	100
	3	7.8613×10^{-09}	2.7043×10^{-04}	5.4093×10^{-05}	32	254	5.13	6	100
	4	2.8418×10^{-09}	5.8998×10^{-03}	1.1800×10^{-03}	29	270	5.48	6	100
	5	2.2405×10^{-09}	1.0694×10^{-04}	2.1391×10^{-05}	30	265	5.37	6	100
4	1	2.4940×10^{-09}	2.5727×10^{-03}	5.1455×10^{-04}	29	255	5.15	6	100
	2	2.6913×10^{-09}	1.7599×10^{-03}	3.5199×10^{-04}	27	248	5.00	5	100
	3	3.7703×10^{-09}	1.9469×10^{-03}	3.8938×10^{-04}	28	269	5.46	6	100
	4	1.5557×10^{-09}	7.4870×10^{-03}	1.4974×10^{-03}	39	261	5.28	6	100
	5	3.7431×10^{-09}	4.1156×10^{-03}	8.2313×10^{-04}	25	286	5.83	6	100
5	1	3.7775×10^{-09}	9.3240×10^{-03}	1.8648×10^{-03}	25	263	5.33	6	100
	2	2.2907×10^{-09}	7.9196×10^{-04}	1.5839×10^{-04}	38	259	5.24	6	100
	3	3.9774×10^{-09}	2.7866×10^{-04}	5.5736×10^{-05}	29	257	5.20	6	100
	4	1.9201×10^{-09}	4.0172×10^{-05}	8.0359×10^{-06}	32	271	5.50	6	100
	5	2.2804×10^{-09}	3.0386×10^{-06}	6.0954×10^{-07}	26	271	5.50	6	100
6	1	6.3866×10^{-09}	1.6673×10^{-04}	3.3347×10^{-04}	26	283	5.76	6	100
	2	2.7670×10^{-09}	5.2337×10^{-04}	1.0468×10^{-04}	26	272	5.52	6	100
	3	1.9947×10^{-09}	2.0087×10^{-04}	4.0176×10^{-05}	30	270	5.48	6	100
	4	2.4039×10^{-09}	3.8848×10^{-03}	7.7697×10^{-04}	29	247	4.98	5	100
	5	4.2223×10^{-09}	4.0325×10^{-03}	8.0650×10^{-04}	25	264	5.35	6	100
7	1	2.3232×10^{-09}	4.3732×10^{-04}	8.7467×10^{-05}	36	259	5.24	6	100
	2	2.7109×10^{-09}	2.8828×10^{-05}	5.7678×10^{-06}	31	262	5.30	6	100
	3	7.1933×10^{-09}	7.2253×10^{-04}	1.4451×10^{-04}	23	280	5.70	6	100
	4	3.0473×10^{-09}	4.3541×10^{-05}	8.7106×10^{-06}	26	284	5.78	6	100
	5	4.0524×10^{-09}	5.5358×10^{-03}	1.1072×10^{-03}	26	271	5.50	6	100
8	1	2.5156×10^{-09}	4.7339×10^{-05}	9.4699×10^{-06}	29	270	5.48	6	100
	2	1.9023×10^{-09}	2.0428×10^{-03}	4.0856×10^{-04}	29	256	5.17	6	100
	3	4.2356×10^{-09}	1.2223×10^{-03}	2.4446×10^{-04}	31	248	5.00	5	100
	4	2.4042×10^{-09}	8.0379×10^{-03}	1.6076×10^{-03}	29	254	5.13	6	100
	5	1.4850×10^{-09}	1.0754×10^{-04}	2.1510×10^{-05}	26	262	5.30	6	100
9	1	3.0244×10^{-09}	6.0535×10^{-05}	1.2109×10^{-05}	29	265	5.37	6	100
	2	1.3259×10^{-09}	2.4919×10^{-03}	4.9838×10^{-04}	38	267	5.41	6	100
	3	1.5360×10^{-09}	5.0027×10^{-03}	1.0005×10^{-03}	31	274	5.57	6	100
	4	3.5911×10^{-09}	4.3539×10^{-03}	8.7077×10^{-04}	25	273	5.54	6	100
	5	3.8356×10^{-09}	1.1563×10^{-03}	2.3127×10^{-04}	29	266	5.39	6	100

Table 4 (continued)

Run	Fold	Training performance error	Testing performance error	Network performance error	Epochs	Effective parameters	Actual neurons	Approximate neurons	Classification accuracy (%)
10	1	5.7878×10^{-09}	2.1846×10^{-03}	4.3692×10^{-04}	27	287	5.85	6	100
	2	1.5139×10^{-09}	1.2345×10^{-03}	2.4691×10^{-04}	28	264	5.35	6	100
	3	6.0237×10^{-09}	1.6717×10^{-03}	3.3435×10^{-04}	22	270	5.48	6	100
	4	1.3166×10^{-09}	2.3029×10^{-03}	4.6059×10^{-04}	36	266	5.39	6	100
	5	4.4638×10^{-09}	1.5829×10^{-03}	3.1658×10^{-04}	31	252	5.09	6	100

Table 5 Network performance results of 14 cases of dataset DS-90 (training, test sets are same and weight-biases are different in each case)

Case	Training performance error	Testing performance error	Network performance error	Epochs	Effective parameters	Actual neurons	Approximate neurons	Classification accuracy (%)
1	1.80×10^{-08}	1.36×10^{-03}	2.71×10^{-04}	24	307	6.28	7	100
2	4.00×10^{-09}	1.48×10^{-04}	2.95×10^{-05}	27	262	5.30	6	100
3	3.94×10^{-09}	1.91×10^{-04}	3.82×10^{-05}	27	273	5.54	6	100
4	3.80×10^{-09}	1.07×10^{-04}	2.14×10^{-05}	27	267	5.41	6	100
5	5.42×10^{-09}	2.34×10^{-04}	4.69×10^{-05}	25	270	5.48	6	100
6	3.47×10^{-09}	7.85×10^{-04}	1.57×10^{-04}	26	275	5.59	6	100
7	3.78×10^{-09}	2.88×10^{-04}	5.76×10^{-05}	26	273	5.54	6	100
8	2.75×10^{-09}	2.23×10^{-04}	4.46×10^{-05}	47	266	5.39	6	100
9	1.77×10^{-09}	1.10×10^{-03}	2.20×10^{-04}	28	260	5.26	6	100
10	2.70×10^{-09}	3.71×10^{-04}	7.42×10^{-05}	29	271	5.50	6	100
11	1.09×10^{-08}	1.93×10^{-04}	3.86×10^{-05}	26	279	5.67	6	100
12	5.18×10^{-09}	3.20×10^{-04}	6.39×10^{-05}	27	296	6.04	7	100
13	5.24×10^{-09}	3.02×10^{-03}	6.04×10^{-04}	31	272	5.52	6	100
14	2.06×10^{-09}	1.78×10^{-04}	3.57×10^{-05}	28	276	5.61	6	100

authors [5, 6, 11] respectively for binary classification. Authors [16] used these methods along with their proposed method (DWT + PCA + RF) for multi class classification and attained highest average accuracy of 95.70% with their proposed approach. Authors [17], attained 96.97% classification accuracy over 4 brain image classes of 66 images. On dataset DS-310, our proposed method achieved higher accuracy of 97.81% than existing approaches. Table 3, reveals that classification accuracy of 100% is yielded by our proposed method in differentiating 90 images into 18 brain image classes. Also, in binary classification, our proposed method achieved higher classification performances metrics of 100% than existing method [9] as shown in Fig. 5.

4.4 Back propagation neural network (BPNN) classifier

4.4.1 Experimentation on dataset DS-90

Following two types of experiments have carried out on dataset DS-90:

- (1) The BPNN is retrained for 375 folds (75 runs $\times$ 5 fold = 375 folds), with different initial weights-biases, and training, test sets. However, these different conditions can lead to different solutions for the same problem [22]. But, under these conditions, we achieved 100% accuracy over training and test sets in each fold. It shows the robustness of system. Randomly 10 runs are selected whose performance results are tabulated in Table 4. It provides information about training performance error, testing performance error, network performance error, number of epochs (iterations), number of effective parameters (weights-biases) and approximate

Table 6 Training and test results of 10 runs out of 75 runs of dataset DS-310 (training, test sets and weight-biases are different in each fold)

Run	Fold	Training performance error	Testing performance error	Network performance error	Training classification accuracy (%)	Testing classification accuracy (%)
1	1	6.33×10^{-10}	1.01×10^{-2}	2.03×10^{-3}	100	96.77
	2	4.78×10^{-10}	5.18×10^{-3}	1.04×10^{-3}	100	98.39
	3	2.66×10^{-10}	1.15×10^{-2}	2.30×10^{-3}	100	96.77
	4	7.15×10^{-10}	5.77×10^{-3}	1.15×10^{-3}	100	98.39
	5	2.56×10^{-10}	1.49×10^{-5}	2.97×10^{-6}	100	100.00
2	1	7.20×10^{-10}	5.80×10^{-6}	1.16×10^{-6}	100	100.00
	2	3.44×10^{-10}	3.59×10^{-3}	7.18×10^{-4}	100	98.39
	3	4.85×10^{-10}	2.33×10^{-3}	4.66×10^{-4}	100	100.00
	4	3.16×10^{-10}	1.03×10^{-2}	2.06×10^{-3}	100	96.77
	5	4.79×10^{-10}	1.06×10^{-2}	2.13×10^{-3}	100	96.77
3	1	4.76×10^{-10}	5.36×10^{-3}	1.07×10^{-3}	100	98.39
	2	6.71×10^{-10}	9.53×10^{-3}	1.91×10^{-3}	100	96.77
	3	6.36×10^{-10}	5.38×10^{-3}	1.08×10^{-3}	100	98.39
	4	3.57×10^{-10}	1.08×10^{-2}	2.16×10^{-3}	100	96.77
	5	3.80×10^{-10}	6.14×10^{-3}	1.23×10^{-3}	100	98.39
4	1	1.63×10^{-10}	1.08×10^{-2}	2.16×10^{-3}	100	96.77
	2	2.05×10^{-10}	4.59×10^{-5}	9.19×10^{-6}	100	100.00
	3	3.37×10^{-10}	1.07×10^{-2}	2.15×10^{-3}	100	96.77
	4	5.52×10^{-10}	5.21×10^{-4}	1.04×10^{-4}	100	100.00
	5	4.86×10^{-10}	7.79×10^{-3}	1.56×10^{-3}	100	96.77
5	1	6.32×10^{-10}	2.78×10^{-4}	5.55×10^{-5}	100	100.00
	2	8.25×10^{-10}	6.13×10^{-3}	1.23×10^{-3}	100	98.39
	3	2.75×10^{-10}	5.07×10^{-3}	1.01×10^{-3}	100	98.39
	4	4.46×10^{-10}	1.07×10^{-2}	2.15×10^{-3}	100	96.77
	5	5.51×10^{-10}	1.03×10^{-2}	2.06×10^{-3}	100	96.77
6	1	2.58×10^{-10}	1.08×10^{-2}	2.16×10^{-3}	100	96.77
	2	9.77×10^{-10}	5.73×10^{-3}	1.15×10^{-3}	100	98.39
	3	6.08×10^{-10}	1.20×10^{-2}	2.41×10^{-3}	100	96.77
	4	6.42×10^{-10}	1.56×10^{-2}	3.12×10^{-3}	100	95.16
	5	2.92×10^{-10}	9.75×10^{-6}	1.95×10^{-6}	100	100.00
7	1	4.62×10^{-10}	1.14×10^{-2}	2.29×10^{-3}	100	96.77
	2	5.78×10^{-10}	5.47×10^{-3}	1.09×10^{-3}	100	98.39
	3	4.71×10^{-10}	2.57×10^{-3}	5.13×10^{-4}	100	98.39
	4	9.27×10^{-10}	9.21×10^{-3}	1.84×10^{-3}	100	96.77
	5	6.84×10^{-10}	2.85×10^{-3}	5.71×10^{-4}	100	98.39
8	1	5.93×10^{-10}	2.80×10^{-3}	5.61×10^{-4}	100	98.39
	2	8.23×10^{-10}	8.41×10^{-4}	1.68×10^{-4}	100	100.00
	3	3.40×10^{-10}	1.14×10^{-2}	2.29×10^{-3}	100	95.16
	4	5.93×10^{-10}	7.99×10^{-3}	1.60×10^{-3}	100	96.77
	5	2.47×10^{-10}	6.42×10^{-4}	1.28×10^{-4}	100	100.00
9	1	4.55×10^{-10}	2.80×10^{-3}	5.59×10^{-4}	100	98.39
	2	3.28×10^{-10}	9.77×10^{-3}	1.95×10^{-3}	100	96.77
	3	1.12×10^{-10}	1.92×10^{-3}	3.84×10^{-4}	100	98.39
	4	4.61×10^{-10}	1.12×10^{-2}	2.25×10^{-3}	100	96.77
	5	5.33×10^{-10}	1.20×10^{-2}	2.39×10^{-3}	100	95.16

Table 6 (continued)

Run	Fold	Training performance error	Testing performance error	Network performance error	Training classification accuracy (%)	Testing classification accuracy (%)
10	1	4.70×10^{-10}	5.14×10^{-3}	1.03×10^{-3}	100	98.39
	2	3.39×10^{-10}	7.51×10^{-3}	1.50×10^{-3}	100	96.77
	3	2.27×10^{-10}	3.67×10^{-3}	7.33×10^{-4}	100	98.39
	4	5.54×10^{-10}	1.07×10^{-2}	2.13×10^{-3}	100	96.77
	5	3.52×10^{-10}	1.07×10^{-2}	2.15×10^{-3}	100	96.77

neurons used by network model and accuracy of network. Thus under diverse initial conditions of network, the performance errors are minimal and we can observe 100% classification accuracy. Also, in each fold the approximate number of neurons i.e. number of neurons used by network is almost same. It indicates the system is reliable and accurate.

- (2) The BPNN is retrained with different initial weight-biases, and training—test sets are kept same. With neural network such tests are generally carried out if satisfactory results are not achieved. By changing initial weight -biases, the network might produce different results [22]. It means, for same diagnosis problem, system can produce different results. It imply, with same training, test sets and varying initial weight-biases, if a network can produce same results continuously, the network is reliable and robust. We have performed this experiment 14 times and attained 100% accuracy in multiple class disease detection. The performance results of these 14 cases are tabulated in Table 5. The performance errors are minimal and in each case the number of neurons used by network is approximately same. It indicates the system is robust and reliable.

5 and 6 reveal competency of proposed method in classifying different brain disorders.

5 Conclusion

In this paper, we have proposed a method utilising wavelet, GLCM extracted textural features and intensity feature of MR images and BPNN as classifier for the diagnosis of different MR brain diseases. The experimental results show that the composite features could achieve 100% multi class disease classification accuracy on dataset DS-90 and 97.81% on dataset DS-310 which are higher than the existing approaches. The standard deviation, energies of approximate and vertical sub bands of level 7 are proved to be first three dominant features for classification as per Wilcoxon test. Also, GLCM textural features contrast, correlation and energy are significant classification features. The rigorous experimentation using BPNN classifier reveals the robustness and effectiveness of proposed method in diagnosis of multi class brain disorders and it can be applied for practical use.

4.4.2 Experimentation on dataset DS-310

The BPNN is retrained for 75 runs, with different initial weights-biases and training, test sets. The results of 10 runs out of 75 runs are tabulated in Table 6. Table 6 provides information about training performance error, testing performance error, network performance error, training classification accuracy and testing classification accuracy. The performance errors are minimal. The average classification accuracy in training phase is 100% and in testing phase is 97.81%.

The average computation time for each image is 0.9428 s. Table 3 and Fig. 5 explain the efficiency of our proposed method over state-of-the-art methods. Tables 4,

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Sprawls P. Magnetic resonance imaging: principles, methods and techniques. Montreat: Sprawls Education Foundation; 2000.
2. Chaplot S, Patnaik LM, Jagannathan NR. Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. Biomed Signal Process Control. 2006;1:86–92. <https://doi.org/10.1016/j.bspc.2006.05.002>.
3. Fujita H, Uchiyama Y, Nakagawa T, Fukuoka D, Hatanaka Y, Hara T, Lee GN, Hayashi Y, Ikeda Y, Gao X, Zhou X. Computer-aided diagnosis: The emerging of three CAD systems induced by Japanese health care needs. Comput Methods Programs Biomed. 2008;92(3):238–48. <https://doi.org/10.1016/j>

- cmpb.2008.04.003 (Medical Imaging and Medical Informatics (MIMI), 0169–2607).**
4. Doi K. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Comput Med Imaging Gr.* 2007;31(4):198–211. <https://doi.org/10.1016/j.compmedimag.2007.02.002> **(Computer-aided Diagnosis (CAD) and Image-guided Decision Support, 0895–6111).**
 5. El-Dahshan E-SA, Hosny T, Salem A-BM. Hybrid intelligent techniques for MRI brain images classification. *Digit Signal Process.* 2010;20(2):433–41. <https://doi.org/10.1016/j.dsp.2009.07.002>.
 6. Das S, Chowdhury M, Kundu MK. Brain MR image classification using multiscale geometric analysis of ripplelet. *Prog Electromagn Res.* 2013;137:1–17. <https://doi.org/10.2528/PIER13010105>.
 7. Saritha M, Joseph KP, Mathew AT. Classification of MRI brain images using combined wavelet entropy based spider web plots and probabilistic neural network. *Pattern Recognit Lett.* 2013;34:2151–6. <https://doi.org/10.1016/j.patrec.2013.08.017>.
 8. El-Dahshan E-SA, Mohsen HM, Revett K, Salem A-BM. Computer-aided diagnosis of human brain tumor through MRI: a survey and a new algorithm. *Expert Syst Appl.* 2014;41(11):5526–45. <https://doi.org/10.1016/j.eswa.2014.01.021>.
 9. Yang G, Zhang Y, Yang J, Ji G, Dong Z, Wang S, Feng C, Wang Q. Automated classification of brain images using wavelet-energy and biogeography-based optimization. *Multimed Tools Appl.* 2015;75:1560115617. <https://doi.org/10.1007/s11042-015-2649-7>.
 10. Zhang Y, Wang S, Dong Z, Phillip P, Ji G, Yang J. Pathological brain detection in magnetic resonance imaging scanning by wavelet entropy and hybridization of biogeography-based optimization and particle swarm optimization. *Prog Electromagn Res.* 2015;152:4158.
 11. Zhang Yudong, Dong Zhengchao, Wang Shuihua, Ji Genlin, Yang Jiquan. Preclinical diagnosis of magnetic resonance (MR) brain images via discrete wavelet packet transform with Tsallis entropy and generalized eigenvalue proximal support vector machine (GEPSVM), *Entropy*, 17, 4, 2015;1795–1813, Multidisciplinary Digital Publishing Institute
 12. Nayak DR, Dash R, Majhi B. Brain MR image classification using two-dimensional discrete wavelet transform and AdaBoost with random forests. *Neurocomputing.* 2016;177(Supplement C):188–97. <https://doi.org/10.1016/j.neucom.2015.11.034>.
 13. Nayak DR, Dash R, Majhi B. Pathological brain detection using curvelet features and least squares SVM. *Multimed Tools Appl.* 2016. <https://doi.org/10.1007/s11042-016-4171-y>.
 14. Zhang Y, Sun Y, Phillips P, Liu G, Zhou X, Wang S. A multilayer perceptron based smart pathological brain detection system by fractional Fourier entropy. *J Med Syst.* 2016;40:1–11. <https://doi.org/10.1007/s10916-016-0525-2>.
 15. Gudigar A, Raghavendra U, San TR, Ciaccio EJ, Acharya UR. Application of multiresolution analysis for automated detection of brain abnormality using MR images: a comparative study. *Future Gen Comput Syst.* 2019;90:359–67.
 16. Siddiqui MF, Mujtaba G, Reza AW, Shuib L. Multi-class disease classification in brain mris using a computer-aided diagnostic system. *Symmetry.* 2017;9(3):37.
 17. Mohsen H, El-Dahshan E-SA, El-Horbaty E-SM, Salem A-BM. Classification using deep learning neural networks for brain tumors. *Future Comput Inform J.* 2018;3(1):68–71.
 18. Harvard Medical School: Database. <http://www.med.harvard.edu/AANLIB>. Accessed 2016.
 19. Vetterli M, Herley C. Wavelets and filter banks: theory and design. *IEEE Trans Signal Process.* 1992;40:2207–32. <https://doi.org/10.1109/78.157221>.
 20. Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans Pattern Anal Mach Intell.* 1989;11:674–93. <https://doi.org/10.1109/34.192463>.
 21. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern.* 1973;SMC-3:610–21. <https://doi.org/10.1109/TSMC.1973.4309314>.
 22. Mark HB, Hagan MT, Demuth, HB. Neural network toolbox™ user's guide. The Mathworks Inc; 1992.
 23. Lowry R. Concepts and applications of inferential statistics, VassarStats website
 24. Haykin S. Neural networks: a comprehensive foundation. Upper Saddle River: Prentice Hall PTR; 1994.
 25. Bishop CM. Neural networks for pattern recognition. Oxford: Oxford University Press; 1995.
 26. Ojha VK, Abraham A, Snášel V. Metaheuristic design of feedforward neural networks: a review of two decades of research. *Eng Appl Artif Intell.* 2017;60:97–116.
 27. Hagan M, Demuth HB, Beale MH, De Jess O. Neural network design. 2nd ed. Boston: PWS Publishing; 1996.
 28. Okut H. Bayesian regularized neural networks for small n big p data. In: Artificial neural networks-models and applications. London: InTech; 2016.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.