International Journal of Advanced Technology and Engineering Exploration, Vol 9(90)
ISSN (Print): 2394-5443 ISSN (Online): 2394-7454
http://dx.doi.org/10.19101/IJATEE.2021.875325

Research Article

Prediction of neurodegenerative disease using brain image analysis with multilinear principal component analysis and quadratic discriminant analysis

Arunraj Gopalsamy[1, 2*], B. Radha[3] and K. Haridas[4]
Research Scholar, Sree Saraswathi Thyagaraja College, Pollachi, TN, India - 642107[1]
Senior IT Architect, Kyndry, Bangalore, Karnataka, India – 560064[2]
Assistant Professor, Sri Krishna Arts and Science College, Coimbatore, TN, India-641008[3]
Associate professor, NGM College, Coimbatore, Pollachi - 642 001[4]

Received: 16-December-2021; Revised: 07-May-2022; Accepted: 10-May-2022

Abstract
Neuroimaging, a part of medical imaging is playing a crucial role in the field of neuroscience and psychology for analysing the structure as well as functions of the internal nervous system. It is the most evolving and prominent method in medical science as it not only focuses on the structure, but also the functional aspects of the brain. It helps in diagnosing metabolic and neurodegenerative diseases. More precisely, magnetic resonance imaging is one of the cutting-edge technologies in the medical verdict for the past four decades in diagnosing Neuro disorders. In recent years, machine learning and other high-end reckoning tools are used universally in assisting the automated identification of neurodegenerative diseases such as Parkinson's disease (PD), Alzheimer's disease (AD) and so on. However, similarities exist in the functional part of the nervous system, making it difficult for the learning algorithms to classify the disease. The main goal of this article is to predict neurodegenerative diseases in an effective way using functional analysis of brain images. The proposed model for predicting neurodegenerative diseases based on brain image analysis has two main phases. The first phase processes the image by applying pre-processing techniques and extracts the features from the images using wavelet transform and histogram gradients for further processing. The second phase applies machine learning algorithms such as multilinear principal component analysis and regularized discriminant analysis for selecting the features and identifying the prediction labels respectively. The experimental analysis has been performed with various brain image datasets such as DS-66, DS-75, DS-160 and DS-255. The obtained results show the 100% accuracy on DS-66, DS-75, and DS-160 datasets and 99.58% accuracy on DS-255 and take a minimum computational overhead of 0.025s. Thus, the proposed model offers improved results than many other existing models. Also, the performance analysis using various quality metrics endorses the significance of the proposed model in diagnosing neurodegenerative disease.

Keywords
Neurodegenerative disease prediction, Wavelet transforms, Histogram-oriented gradients, Multilinear principal component analysis, Regularized discriminant analysis, Brain MR images.

1.Introduction
Age-related disorders are becoming a real threat to humans and becoming increasingly prevalent for elders since the population of elderly people is increasing recently. This situation arises due to the development of technologies in healthcare. Neurodegenerative disease is quite a common term used to point out various health conditions of an individual that majorly affects the neurons in the human brain.

Neurons are the basic significant component in the nervous system that includes the brain. Neurons are also a crucial component due to their inability to reproduce themselves. Thus, once neurons are damaged, they cannot be replaced ostensibly. Neurodegenerative diseases slowly degenerate the cells in the nervous system end up with dead cells causing the disease incurable [1]. This normally affects the body movements, and mental functioning and even causes increased difficulty in moving, speaking as well as breathing [2].

Some of the most common neurodegenerative diseases are Alzheimer's disease (AD), Parkinson's

*Author for correspondence

disease (PD), prion disease, motor neuron diseases, amyotrophic lateral sclerosis, Huntington's disease, spinocerebellar ataxia, spinal muscular atrophy and even more [3]. Among these diseases, AD and PD are the top most common diseases. According to the statistical report, it is said that around 4,00,000 and 80,000 people were affected by mental disorders and PD respectively in Australia [4]. It is reported that 6.2 million Americans were affected by Alzheimer's disease by 2021 and around 1.2 million Americans could get affected by PD by the year 2030 [5]. In general, most neurodegenerative diseases are caused by inherited genetic changes, environmental factors such as pesticides, the chemicals used and air pollution, the combination of both genetics and environmental factors as well as the age factor.

The treatment can be effective only when the diagnosis is made accurately and promptly [6]. Several efforts have been taken to develop biomarkers to diagnose the disease and treat them through clinical trials. Thus, several methods have been used to identify neurodegenerative disorders. The clinical, as well as pathological diagnoses of the disease, involve voice analysis, gait movements and brain magnetic resonance (MR) images analysis. However, the increase in the population is really an obstacle in diagnosing and treating the disease. Thanks to the advancement of technology in the healthcare field, they strive hard to automatically diagnose diseases based on the history of data as early as possible [7]. Machine learning, a part of artificial intelligence is one of the powerful fields along with the use of other fields such as image processing, statistics and data mining allowing professionals to address the challenges by developing computer-aided diagnosis tools [8].

As an advancement in diagnosing neurodegenerative disease, neuroimaging is used widely comprising of various imaging biomarkers such as magnetic resonance imaging (MRI) and single photon emission computed tomography (SPECT). The traditional MRI images utilize a unique pulse sequence to capture T1-weighted (T1), T2-weighted (T2), proton-density weighted scans, SPECT is a functional nuclear imaging technique and so on [9]. Several research ideas have been proposed to use machine learning techniques to categorize pathological brain images from normal ones [10]. The overall idea is to process the images and extract the features using image processing and statistical methods for which machine learning models are applied to discriminate the diseased brain images. The most important factor that is considered to differentiate normal and diseased brain images is the axial symmetry that exists in normal brains. Thus, the asymmetry in the brain images indicates the abnormal or pathological brain image [11].

Different models have been proposed for diagnosing the disease based on the brain MR images [12]. These models vary by the methods utilized in feature extraction, feature selection and classification. The most significant phase is feature extraction and it contributes to the efficiency in attaining accurate diagnosing. Also, to classify the features as normal or abnormal, various techniques has been proposed in the literature. It can be grouped under supervised and unsupervised learning and they differ based on the requirements of training [13].

The main focus of this paper is to propose a model to predict the disease by categorizing the pathological images from normal ones and to investigate the performance using various experimental analyses. The proposed model is designed to predict the disease using MR images with two phases. The first phase helps to process the image by applying various pre-processing techniques and captures the features using discrete wavelet transform (DWT) and histogram oriented gradients (HOG). The second phase focuses on various machine learning algorithms such as multilinear principal component analysis (MPCA) for identifying significant features by reducing the dimensions and regularized discriminant analysis (RDA) for classifying the prediction labels. Various experimental analyses have been performed to analyse the performance of the proposed model in diagnosing neurodegenerative disease.

The paper is integrated into various sections that are as follows. Section 2 presents the various works related to the proposed study. Section 3 describes the proposed model with the architectural design and explains each step in the proposed model to classify the given input brain MR images. The experimental analysis for feature extraction and classification is presented in section 4. Section 5 reports the performance of the entire model by comparing the results with the existing model. Finally, the conclusion section is discussed in section 6.

2.Literature review

In the past two decades, the use of computer-aided software in healthcare has gained huge attention for the development of an automated system for

diagnosing the disease. It helps the physicians to treat the disease at the earliest and moderates the time. The earlier diagnosis of neurodegenerative disease through voice and speech data [14, 15] and gait parameters [16] has gained huge attention which are specific to particular diseases. However, the accurate diagnosis can be made for various neurodegenerative diseases through brain image analysis. In the literature, there exist several models for classifying pathological brain images from normal ones. A detailed survey was made that inspects and compares various deep learning mechanisms used in identifying neurological disorders. It was intended to identify AD, PD and schizophrenia by examining the brain MR images [17−19].

Each model in the literature differs from the other in extracting and selecting the features as well as classifying the samples. The most commonly used feature extraction techniques are intensity histogram (IH), gray level co-occurrence matrix (GLCM) [20], HOG [21], DWT, dense feature with speed up robust features (SURF), and scale invariant feature transform (SIFT), first order statistics (FOS), discrete cosine transform (DCT) and bag of words (BoW) [22]. With the analysis made on these feature extraction techniques, a model was proposed that makes use of DWT along with BoW for feature extraction. The various feature extraction methods were also analysed with different classifiers such as linear support vector machine (LSVM), radial support vector machine (RSVM), and polynomial linear support vector machine (PSVM), k-nearest neighbour (KNN), random forest (RF) and Adaboost. It was proved that their choice of feature selection mechanism offers better results with the RSVM classifier.

Similarly, a study was carried out that makes use of principal component analysis (PCA), MPCA, intensity summary statistics and laws' texture energy measure for feature extraction and were analysed using linear discriminant analysis (LDA), support vector machine (SVM), decision tree and RF classifiers for diagnosing PD [23]. For classifying pathological brain disease, various models such as two-dimensional discrete wavelet transform (2D-DWT), PCA and quadratic discriminant analysis (QDA) for feature extraction, feature selection and classification were anticipated [10]. The authors also used other approaches such as 2D-DWT, PCA and probabilistic principal component analysis (PPCA) using a random subspace ensemble (RSE) classifier [24] and AdaBoost with RF classifier (ADBRF) [25].

A model that utilizes DWT for feature extraction and SVM with polynomial (POLY) and radial basis function (RBF) kernel for the classification of brain MRI images was proposed [26]. Also, DWT was used with a naive Bayes classifier (NBC) to improve brain image categorization [27]. Most of the models employ DWT and PCA for feature extraction besides feature selection and are evaluated with different classifiers such as KNN, feed-forward back-propagation artificial neural network (FPANN) [28], forward neural network (FNN) with adaptive chaotic particle swarm optimization (ACPSO) [29] and scaled chaotic artificial bee colony (SCABC) [30], backpropagation neural network (BPNN) with scaled conjugate gradient (SCG) [31], kernel support vector machine (KSVM) with gaussian radial basis (GRB) kernel [32] and least squares support vector machine (LSSVM) [33]. A LSSVM was used with PCA in classifying the brain MR images that utilize Ripplet transform type for capturing the features [34].

The DWT based feature extraction approach has experimented with symmetric uncertainty ranking (SUR) and PPCA with various classifiers such as filtered classifiers (FC) and SVM [35]. The models have experimented with DS-75 and DS-160 datasets. A model that utilizes wavelet entropy (WE) along with spider web plots (SWP) for capturing the features and a probabilistic neural network (PNN) for label identification was suggested [36]. The approach WE has been experimented with the hybridization of biogeography-based optimization (BBO) and particle swarm optimization (PSO) using the FNN model (HBP-FNN) to classify the brain images [37]. Similarly, a discrete wavelet packet transform (DWPR) with Shannon entropy (SE) with generalized eigenvalue proximal support vector machine (GEPSVM) was suggested to classify MR images for effective diagnosis [38]. The same idea was explored with a variance of DWT, the dual-tree complex wavelet transform (DTCWT) with twin SVM was proposed [39].

A model was proposed to identify the tumour in the MR brain images in which feedback pulse-coupled neural network (FPCNN) was used for segmentation, DWT+PCA for feature extraction and artificial neural network for classification [40]. Prediction of PD at various stages was carried out using a convolutional neural network (CNN) classification algorithm with brain MRI images [41]. Another gait analysis based neurodegenerative disease prediction model that utilizes deep learning was proposed in which the CNN based approach offers better performance [42].

However, the model does not suitable for brain imaging since it was suitable for gait parameters. Stationary wavelet transform (SWT) was an alternate feature extraction technique for DWT, which was most widely used along with PCA with different classifiers such as an integrated model that uses artificial bee colony (ABC), PSO and ABC with standard PSO [43]. A fractional Fourier entropy (FRFE) for extracting features, dynamic pruning and Bayesian detection boundaries for selecting features and adaptive real-coded biogeography-based optimization (ARCBBO) for classification was anticipated [44]. Two different feature extraction techniques such as DWT and SWT were implemented with SUR feature selection and AdaBoost based support vector machine (ADBSVM) classifier [45].

From the extensive literature survey made for the classification of brain MR images, it is evident that most of the models utilize DWT or its deviances for extracting mass features from the brain MR images due to its efficiency. Also, the use of PCA is most common in dimensionality reduction. On the other hand, for classification, a wide range of algorithms was employed. However, most of the models suffer from computational overhead that may not be appropriate for producing effective results. Also, the reduction of features offered by many models may not be effective enough for increasing the performance of the underlying model. Moreover, the accuracy values are to be increased further. Thus, taking the limitations as the real motivation, the model has been proposed to predict neurodegenerative disease. It analyses the brain MR images using 2D-DWT, MPCA and QDA approach.

3.Methods

This study focuses on predicting neurodegenerative diseases by analysing the functional images of the brain. The overall framework of the proposed prediction model is depicted in *Figure 1*. The framework of the prediction model is segmented into two main levels recognized as the image processing phase and the learning phase. The image processing phase deals with the input image in which it applies pre-processing steps to enhance the image and filter the noises present in the image.

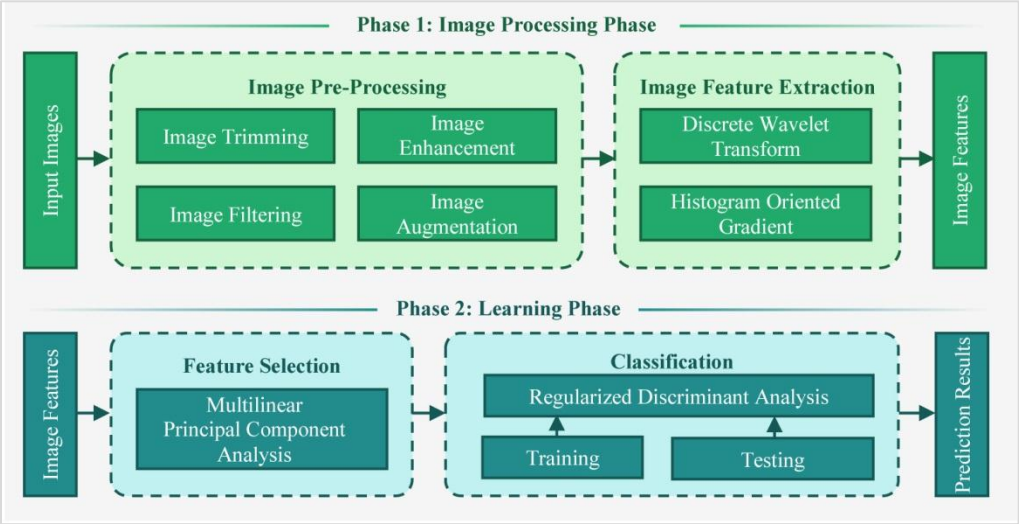

Figure 1 Overall design of the proposed prediction model

Then, the features are extracted from the image using the hybrid approach that makes use of DWT and HOG. The subsequent learning phase takes the extracted features as input with which the significant features are identified using MPCA. Finally, the model is trained and the class label for the test data is classified using RDA.

3.1Image processing phase

The input brain images cannot be directly processed by machine learning algorithms for automated prediction. The data may contain noise that must be eliminated. Additionally, the data in the form of the pixel must be converted to some form that is suitable for further analysis by applying statistics. This phase completely deals with the raw input image and

produces statistical data as an output that serves as a feature for the subsequent machine learning phases.

3.1.1 Image pre-processing

Any mining technique starts with the data pre-processing. It is evidently known that the quality of output obtained from any model clearly depends on the quality of the input data. Furthermore, in the case of image analysis and medical imaging, the image obtained from various scanners are susceptible to noise due to motion, spatial resolution, signal and contrast of the image and so on. These noises are to be removed to enhance the quality of the image thereby guaranteeing precise examination. Several image modalities are available for imaging biomarkers as listed in *Table 1* [9]. The pre-processing steps often depend on the type of input images given to the model. In general, the proposed model utilizes various steps to enhance image quality.

Table 1 Various image biomarkers for different disorders

Imaging modalities	Biomarker identification
Structural MRI	Volumetric and voxel-based morphometric examination
Diffusion Tensor MRI	Evaluation of white matter tract injury and microstructural integrity
Proton Magnetic Resonance Spectroscopy (PMRS)	Enumerating proton-holding brain metabolites
Single Photon Emission Computed Tomography (SPECT)	Assessment of nigrostriatal integrity
Positron Emission Tomography (PET)	Determining neuroinflammation, nigrostriatal functions, glucose metabolism, so on
Myocardial Scintigraphy	Dysautonomia
Transcranial Sonography	Gauging substantia nigra and lentiform nucleus echogenicity

Image conversion

Initially, the images are converted to grayscale having shades of black and white. This can be done using various methods. However, the luminosity method offers better results that are similar to the sensitivity of the human eye. Here the red, green and blue values at each pixel in the given input image can be converted to gray using weights based on the human perceiving capacity. The gray conversion using the luminosity method can be computed as in Equation 1.

$$g_{scale} = \text{Red} \times 0.299 + \text{Green} \times 0.587 + \text{Blue} \times 0.114 \quad (1)$$

Most of the input image contains a black background with brain images at the center. Processing the huge black background is of no use and thus the center of the image can be trimmed to reduce the size. Thus, the image with 128×128 pixels can be cropped to 50×50 pixels in which the brain image will be at the center with minimum black background.

Image enhancement

The cropped image must be enhanced by improving the contrast which can be done by adjusting the intensity values of the pixels. The proposed model applies histogram equalization (HE) to improve the contrast in the image. Though several techniques for improving the contrast have been proposed in the literature by enhancing the HE techniques, still HE is most widely used in medical image analysis due to its simplicity as well as effective performance [46]. HE aims at improving the image by distributing the dynamic range of pixel intensity using the aggregate density function and spreading the frequencies in the histogram [47]. The image can be improved using the function in Equation 2.

$$\text{HE}_k = \sum_{i=0}^{k} \frac{n_i}{N} \times (G_l - 1) \quad (2)$$

Here k varies from 0 to G_l-1, G_l represents the number of gray levels in the image, N represents the total number of pixels in the image and n_i specifies the number of occurrences of the i^{th} pixel.

Upon improving the contrast in the image, normalization can be performed to fit the given brain image into a particular template concerning size and shape [17]. The proposed model applies intensity normalization which decreases the disparity that occurred with the use of different scanners or the same scanner with different features as well as scanning the same object at different times. The intensity normalization Int_{norm} can be applied as in Equation 3 [48].

$$Int_{norm} = \frac{Int - Int_{min}}{Int_{\max} - Int_{min}} \quad (3)$$

Image filtering

Often the images obtained as the output from the aforementioned pre-processing step contain noises. Thus, it is necessary to perform denoising before further analysis. The proposed model employs Wiener filtering to remove the noise which produces the desired output by assuming additive noise and

noise spectra. The main advantage of Wiener filtering over other filtering techniques is that it reduces the mean squared error between estimated and desired output [24]. The Wiener filter can be applied to the image in the frequency domain as in Equation 4.

$$W(u,v) = \frac{H^*(u,v)P_s(u,v)}{|H(u,v)|^2 P_s(u,v)+P_n(u,v)} \quad (4)$$

Here, W and H are the DWT with H representing the blurring function and W represents the original image version after the Fourier transformation. The terms $P_s(u,v)$ and $P_n(u,v)$ specify the mean power of spectral density of the signal and noise respectively. In the case of inverse filtering, the blurring function must be determined with a lot of effort since it is not known exactly. Moreover, in some situations, the noise gets augmented and which may destroy the reconstruction. These issues are being solved by the Wiener filter.

Image augmentation

Prediction of neurodegenerative diseases involves a class imbalance problem as the number of samples with neurodegenerative diseases will be low in comparison with normal samples or vice versa. In order to balance the samples in the class and to improve the efficiency of the classifier, image augmentation is applied to the neurodegenerative samples obviously increasing the size of the training set [49]. This can be normally performed by applying transformations to the original image such as rotate, shift, zoom and so on to create duplicate images [23]. In the proposed model, the horizontal flip and brightness augmentation is applied randomly on the brain image corresponding to the PD patients to increase the minority class size and overcomes the imbalanced class problem.

The sample MR brain image of a female patient having a vague headache of age 20 years is shown in *Figure 2* [50]. The figure shows the output of the image after applying the various pre-processing steps on the (a) original image such as (b) cropping, (c) HE, (d) image normalization, (e) image filtering, and (f) horizontal flip augmentation.

3.1.2 Image feature extraction

Once the input image is pre-processed, then the next step to be performed is the extraction of features that are required for the analysis. The proposed model makes use of DWT and HOG for extracting the features of the image.

2D Discrete wavelet transform

DWT is considered to be the inevitable as well as an influential wavelet transform used in many fields of research. In general, wavelet, a mathematical function refers to the data in various forms based on frequency components based on the resolution [51]. It has become a common choice for extracting features from images. Moreover, it is also an effective means for classification since it contains the information of an image regarding frequency as well as time localization [22]. Simplicity in understanding and implementing the approach and minimum time consumption are the major advantages of DWT.

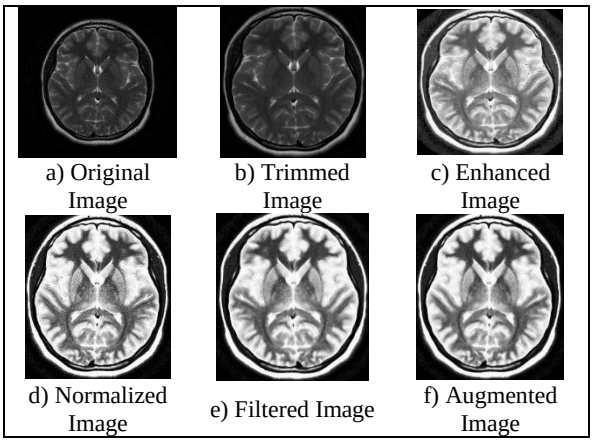

Figure 2 Image preprocessing steps for a sample brain image

The continuous wavelet transform of the signal x represented as the square integral function (t) concerning the wavelet (t) [10] can be given as in Equation 5.

$$W_\psi(a,b) = \int_{-\infty}^{\infty} x(t)\left(\frac{1}{\sqrt{2}}\psi * \left(\frac{t-a}{b}\right)\right)dt \quad (5)$$

Here $W_\psi(a,b)$ represents wavelet transform based on two main factors such as dilation and translation factors represented by the variables a and b. The operation * specifies the complex conjugate. It can be discretized by detaining the variables a and b to a separate lattice where $a=2^j$ and $b=2^jk$ to deliver the DWT as in Equation 6.

$$\left.\begin{aligned} A_{a,b}(n) &= DS(\textstyle\sum_n x(n) l_a * (n-2^a b)) \\ D_{a,b}(n) &= DS(\textstyle\sum_n x(n) h_a * (n-2^a b)) \end{aligned}\right\} \quad (6)$$

The variables A and D represent the coefficient of approximation and detailed components obtained from low pass filer (l) and the high pass filter (h) respectively with dilation and translation factors. DS represents the down sampling. Thus, low frequency components correspond to approximation and high frequency components correspond to detailed components [24].

In the proposed model, 2D DWT is utilized in which it computes the wavelet coefficient at four subbands with three levels of decomposition. These four subbands represents the approximation component of the image which is utilized for other levels in 2D DWT. Thus, the four subbands signifies approximation coefficients, horizontal detail coefficients, vertical detail coefficients and diagonal detail coefficients [52].

With an increase in the decomposition level, grainier details of approximation of an image in a compressed form can be obtained. Thus, wavelet transform provides clarity to an image in a layered approach. Though there exist various wavelet families in literature, Daubechies and Haar wavelet are the most popular among researchers [36]. The proposed model utilizes the Haar wavelet as it is simple to implement, has fast computation, and extracts coarser structural information from the given input. Besides, it is proved to be efficient in computing approximation coefficients with three-level decomposition [24]. The approximation component at level three decomposition represents the initial feature that is extracted using 2D-DWT.

Histogram oriented gradient

HOG is a popular feature engineering technique for extracting features from images. The overall idea is to describe the features based on the structure of the objects in the image. Here the image is partitioned into non-overlapping local portions called cells. For each cell, magnitude and direction are computed and are termed gradient and orientation. Finally, the differential IH for each region is figured out using gradients and the orientation of pixels by analysing the variations of the intensity [21]. The combination of the normalized HOG represents the feature vector for the given image [22]. Gradients for each pixel representing the small change in the directions of horizontal and vertical direction on the x-axis and y-axis are computed using the image filters [1,0,1] and $[-1,0,1]^T$ as in Equation 7.

$$\left.\begin{aligned} G_x &= \frac{\partial f(x,y)}{\partial x} = \frac{f(x+1,y)-(f(x-1,y)}{(x+1)-(x-1)} \\ G_y &= \frac{\partial f(x,y)}{\partial y} = \frac{f(x,y+1)-(f(x,y-1)}{(y+1)-(y-1)} \end{aligned}\right\} \quad (7)$$

Thus, for the given image, the gradient based magnitude and the orientation can be calculated as in Equations 8 and 9.

$$\Delta f(x,y) = \sqrt{{G_x}^2 + {G_y}^2} \quad (8)$$

$$\theta(x,y) = tan^{-1}\left(\frac{{G_y}^2}{{G_x}^2}\right) \quad (9)$$

The magnitude $\Delta f(x,y)$ represents the sum of directional directives of two dimensions concerning x and y; the orientation $\theta(x,y)$ specifies the angle concerning the positive horizontal axis. It is appropriately used in identifying the texture features in the image. The method has proven to be efficient in extracting the features of the MR brain image and is suitable for classification [21].

3.2Learning phase

Upon identifying the features of the input image in the first phase, the second phase applies various machine learning techniques for diagnosing the disease effectively. This phase constitutes two main processes such as feature selection and classification of test data. Selecting significant features is the most crucial step for improving the quality of the result. All the extracted features may not contribute to the prediction process. Eventually, involving those uninterested features reduces the model efficiency in terms of accuracy and execution time. Thus, identifying the interesting features improves the prediction accuracy as well as minimizes the time complexity. The proposed model applies MPCA for selecting significant attributes and employs RDA for classifying the results more effectively.

3.2.1Feature selection using MPCA

The feature selection technique can be employed in prediction models to solve the problem of the curse of dimensionality. PCA is the most popular and frequently employed in dimensionality reduction. A set of variables is explained using its variance and co-variance through linear combinations termed principle component [53]. This can be done by computing the difference between the values of the input feature vector with that of the mean value. With these values, the covariance matrix, eigenvalues and eigenvectors are computed. These eigenvectors are combined to form principle components [24]. Finally, the original variables are replaced with these estimated principle components. This method rapidly processes the data by creating variations in the dataset with the process of reducing the dimensions. The method creates three main effects such as orthogonalizing the input to make them uncorrelated, arranging the variation in descending order and eliminating the least dissimilarity value. The main drawback of PCA is that it processes the input image as a vector by creating the variance and covariance matrix. However, this process is complex and inefficient. The proposed model utilizes the variation of PCA called MPCA that concludes multilinear prediction onto a tensor subspace of lower dimensionality for identifying disparity present in the

original input [54]. The main advantage is that it reduces the dimension by preserving details [55]. Consider an image I with size m x m with n pixels varying from 1 to n. Then the MPCA is intended to optimize the function [23] as given in Equation 10.

$$minimize \sum_{j=1}^{n} \left\| (I_j - \bar{I}) - I_1 I_1^T (I_j - \bar{I}) I_2 I_2^T \right\|^2 \quad (10)$$

By finding I_1 $(n \times p)$ and I_2 $(n \times q)$ with p, q ≤ 0. Here, $\bar{I}$ specify the average value of the input and ||·|| specifies the input matrix in Frobenius form. Also, the original values of the image will be replaced by the low dimensional vector $I_1^T (I_j - \bar{I}) I_2$. Thus, the number of features in the image is reduced to *pq*. The singular value decomposition will be applied for 2(m x m) times which is very minimum than singular value decomposition in PCA which is applied for m^2xm^2 times.

3.2.2Regularized discriminant analysis classifier

Classification or class label prediction is the inevitable core process in prediction applications. In general discriminant analysis is a popular mathematical approach commonly employed in the classification approach. The features selected in the previous step are given as input for classification. The proposed approach utilizes the RDA, a generalized form of four different conditional probability Gaussian classifiers such as LDA, QDA, and diagonal linear discriminant analysis (DLDA) and diagonal quadratic discriminant analysis (DQDA) that are obtained by adjusting the parameters. The relation between various discriminant analyses is depicted in *Figure 3*.

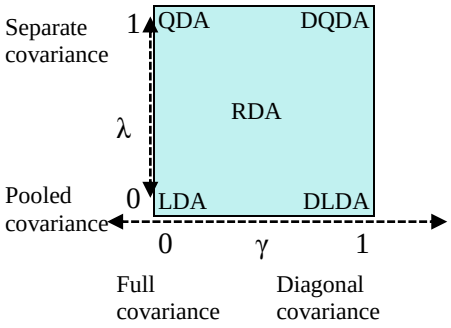

Figure 3 Relation between different discriminant analyses

The assumptions for the RDA method are as follows [56]. Consider a dataset with N samples and K classes, the number of samples in each class k and the data points can be represented as N_k and x_n. Then with the prior probability *(π_k=N_k/N)* and the mean of the class *(μ_k=$\sum x_n/N_k$)* the total covariance matrix can be computed as in Equation 11.

$$\hat{\Sigma} = \sum_k^K \sum_{n \epsilon k} \frac{(x_n - \mu_k)(x_n - \mu_k)^T}{N - K} \quad (11)$$

Here, in the case of LDA, it results in pooled or shared covariance in which the covariance matrix of each class is equal ($\sum_k = \sum$). On the other hand, In the case of QDA, the covariance matrix of each class has to be estimated separately. Thus, RDA is a generalized method of LDA and QDA and it makes use of both LDA and QDA concepts. In RDA, the covariance matrix of each class is separate and similar to QDA. Nonetheless, it tries to regularize the values towards shared covariance with the parameter λ like LDA. It is represented in Equation 12.

$$\hat{\Sigma}_k(\lambda) = \lambda \hat{\Sigma}_k + (1 - \lambda)\hat{\Sigma} \quad (12)$$

Here λ lies between 0 and 1 with which λ=0 specifies the LDA and λ=1 specifies the QDA [57]. Moreover, the covariance matrix can be further regularized with succeeding parameter γ towards diagonal results in Equation 13.

$$\left.\begin{aligned} \hat{\Sigma}(\gamma) &= (1 - \gamma)\hat{\Sigma} + \gamma \cdot dig(\hat{\Sigma}) \\ \hat{\Sigma}_k(\gamma) &= (1 - \gamma)\hat{\Sigma}_k + \gamma \cdot dig(\hat{\Sigma}_k) \end{aligned}\right\} \quad (13)$$

Thus, combining the orthogonal regularizations in (12) and (13) to find total covariance results in Equation 14.

$$\hat{\Sigma}_k(\lambda, \gamma) = \lambda(1 - \gamma)\hat{\Sigma}_k + (1 - \lambda)(1 - \gamma)\hat{\Sigma} + (1 - \lambda) dig(\hat{\Sigma}) + \lambda\gamma \cdot dig(\hat{\Sigma}_k) \quad (14)$$

The discriminant formula for classification is shown in Equation 15.

$$d_k(X) = (X - \bar{X})^T \hat{\Sigma}_k{}^{-1}(\lambda, \gamma)(X - \bar{X}_k) + log\left|\hat{\Sigma}_k(\lambda, \gamma)\right| - 2\log(\pi_k) \quad (15)$$

The RDA is used in the classification process of the proposed model as it is proved to be more effective in producing results than other discriminant analyses [58]. The overall flow of work of the proposed neurodegenerative disease prediction model using brain images is shown in *Figure 4*. The algorithm pseudocode for the proposed neurodegenerative disease prediction model using MR brain images is given in *Figure 5*.

4.Results

To analyse the performance of the proposed prediction model, simulation has been accomplished on a personal computer with the configuration as follows: Intel (R) Core (TM) with the i3-4005U CPU at the rate of 1.70 GHz having a memory of 4 GB RAM running under Windows 8 operating system installed with MATLAB 2014a.

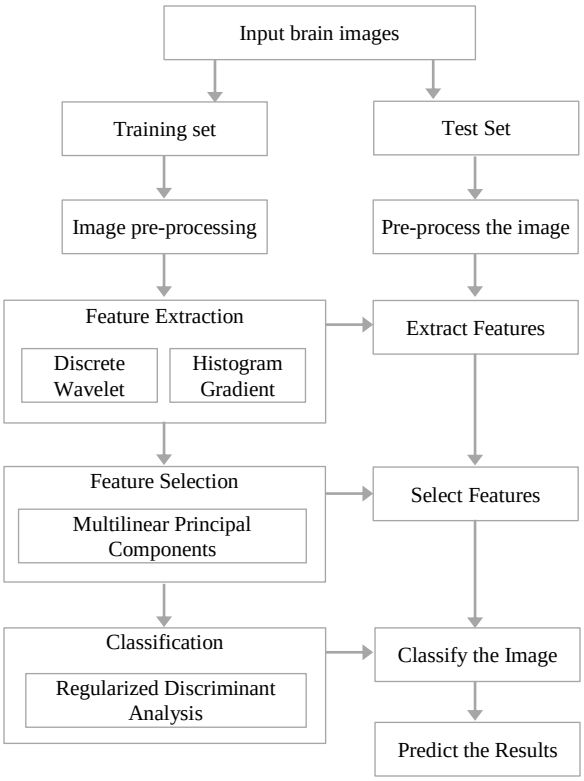

Figure 4 Workflow of the proposed model

Algorithm: Brain Image based Neurodegenerative Disease Prediction Model
Input: Brain Image Dataset D with N samples, Test Data t
Output: Predicted class label
Procedure BI_NDPM ()
// Image Preprocessing
For each image in D, *i* from 1 to N
Read the images and convert them to the grayscale image
Trim the image and apply Histogram Equalization
Apply Intensity Normalization and Weiner Filtering
If the class label is PD **then**
Apply horizontal flip image augmentation
End If
End For
//Feature Extraction Process
For each preprocessed image in the training set *i* from 1 to N
Apply 2D-DWT using 3l decomp. using Haar Wavelet
Extract and store wavelet coefficients as a feature vector
Apply HoG and normalize values
Normalize and store histogram values as a feature vector
End For
// Feature Selection Process
For each feature vector of the image in D, *i* from 1 to N
Apply MPCA on wavelet and histogram feature vectors
Store the reduced feature vector
End For
// Classification Process
For each feature vector of the image D, *i* from 1 to N
Apply the RDA classifier to train the model
End For
//Classify the test data
Preprocess the test data, extract and select the features
Classify the test data and predict the class labels
End Procedure

Figure 5 Algorithm pseudocode for the proposed prediction model

4.1Datasets used

Four standard brain MR image datasets have been used for the analysis of the proposed model as well as to validate the results with that of the existing models. The datasets DS-66, DS-75, DS-160, and DS-255 having 66, 75, 160 and 255 images each are utilized for the experiments. Each image in the datasets represents the T2-weighted MR brain images with an in-plane spatial resolution of 256×256 in the axial plane. The T2 model images are preferred over T1 and PET models due to high clarity and contrast. The datasets such as DS-66 and DS-160 are highly used for the classification of both normal and abnormal brain images. The abnormal brain images consist of seven types of brain diseases, including Glioma, Meningioma, Alzheimer's disease, Alzheimer's disease with visual agnosia, Pick's disease, sarcoma and Huntington's disease [59].

The DS-75 dataset is composed of 5 types of images such as normal brain, and bran images with neurodegenerative disease, cerebrovascular disease, neoplastic disease, and infectious disease. These images are equally distributed [35] and are available at the Harvard whole-brain atlas [59]. The dataset DS-255 is widely used recently that contains 255 brain images in which it has extra 4 types of diseased images apart from the 7 diseased images, including chronic subdural hematoma, cerebral toxoplasmosis, and herpes encephalitis, and multiple sclerosis [34].

Some of the diseased brain MR images from the whole brain atlas that falls under neurodegenerative disease and infectious disease are shown in *Figure 6*. Cross-validation is an assessment system used for assessing the prediction model's performance that helps to train the model with the underlying dataset. In general, it partitions the data and the process is carried out in terms of rounds. At each round, it arbitrarily selects the partitions for the training and test set.

In the proposed model, 6-fold stratified cross-validation is used for the dataset DS-66, and with the datasets DS-75, DS-160 and DS-255, 5-fold cross-validation has been implemented [18, 22, 25, 39]. The statistical analysis of the number of healthy (H) and pathological (P) brain images used in the experiments along with the test set and training set is shown in *Table 2*.

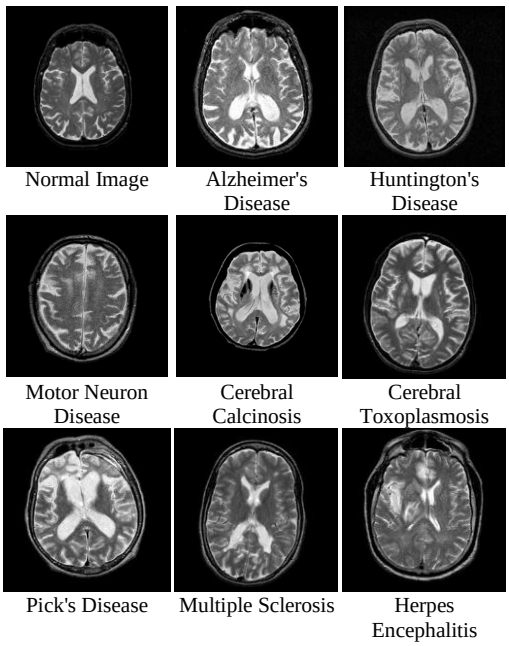

Figure 6 Sample brain MRI images from the whole brain atlas

Table 2 Statistical analysis of the datasets used

Dataset	Total		Training		Validation	
	H	P	H	P	H	P
DS-66	18	48	15	40	3	8
DS-75	15	60	12	48	3	12
DS-160	20	140	16	112	4	28
DS-255	35	220	28	176	7	44

The model is trained using the RDA classifier with a training set. For the dataset DS-66, out of 66 images, 55 images are used for training and 11 images for validation. Similarly, for DS-75, DS-160 and DS-255 datasets, 60, 128 and 204 images are used for training and 15, 32 and 51 images are used for validation.

4.2Parameters used

Eight different parameters have been used to analyse the performance of the proposed model with four different datasets.

4.2.1Parameters used for validation

A model can be proved to be effective only when its performance is assessed using some sort of evaluation metrics. There exist several performance metrics in the literature to significantly assess the proposed models [60]. In the proposed study, various performance metrics are used to evaluate its efficiency by analysing various phases in the proposed model and overall performance. The proposed model considers binary classes in which one is normal and another one is a diseased brain irrespective of the diseases. Some of the metrics used in the research article are accuracy, sensitivity, specificity, precision, F-measure, Matthew's correlation coefficient (MCC), Receiver operating characteristic (ROC) curve and error rate. Accuracy is the most frequently used metric which refers to the percentage of correctly classified samples concerning all the classes. Sensitivity refers to the fraction of correctly predicted positive class (disease). Specificity refers to the fraction of correctly predicted negative class (normal).

Precision is another metric that corresponds to the fraction of correctly classified positive samples among predicted positive samples. F-measure is particularly used for binary classification analysis that refers to the harmonic mean of the precision and sensitivity. MCC is a balanced measure that can be applied to assess the classifications of two classes that is suitable for unbalanced datasets that can be computed positive and negative predicted value, false and true positive rate, false positive and negative rate, false omission rate as well as false discovery rate. ROC curve plots the true positive rate against the false-positive rate at different thresholds which refers to the detection rate and false alarm rate. Error rate evaluates the inaccuracies in the prediction process. It indicates the number of samples that are misclassified in the underlying classification process.

4.3Analysis using conventional models

This section presents the various analysis carried out to prove the efficiency of each phase such as feature extraction and classification. The study includes the prediction of class labels as normal and diseased images. The proposed model undergoes various pre-processing steps such as cropping, contrast enhancement using HE, image normalization, image filtering and image augmentation. Then the feature extraction is performed using DWT and HOG approaches.

The dimensionality reduction on the extracted features of the images is carried out using MPCA, which is then trained and validated using an RDA classifier. The feature extracted using various approaches that are used habitually in a prediction model are evaluated using various classifiers. The results are compared with the proposed idea of using DWT + HOG with the DS-255 dataset. The various feature extraction used in the analysis are IH, GLCM, HOG, DWT, Dense feature with SIFT and SURF, DCT, FOS and BoW and these approaches are evaluated using various classifiers such an LSVM, RSVM, PSVM, KNN, RF and AdaBoost classifier

[22]. The accuracies calculated after the first run with different classifiers for various feature extraction approaches are presented in *Table 3*.

With LSVM classifier, DCT based feature extraction offers improved results. DWT with BoW offers good results with RSVM, PSVM and KNN classifiers. DWT with HOG and IH offers better results with RF and AdaBoost classifiers respectively. The feature extraction approach DWT + HOG used in the proposed model acquire better results of 93.46% with RF classifier whereas it acquires second top position for LSVM and AdaBoost classifiers with 98.68% and 97.75% respectively, and acquires third top position with RSVM, PSVM and KNN classifiers with 97.45%, 98.82% and 98.32% respectively. However, with the average values with different classifiers of the obtained results, the proposed method has a value of 97.41% and obtains the average rank of 2.33. This shows that the feature extraction used in the proposed model is better than many of the approaches such as DWT + BoW, SIFT + BoW and HOG. Thus, though the proposed model acquires the top position for a minimum number of classifiers when comparing the overall performance, it can be seen that the results are remarkable. The average values obtained for the accuracy and the rank for various classifiers are shown in *Figure 7*.

Table 3 Accuracy of different feature extraction approaches

Method	**LSVM**	**RSVM**	**PSVM**	**KNN**	**RF**	AdaBoost
IH	86.27	87.45	83.13	96.86	88.62	**98.45**
GLCM	86.27	87.84	81.17	89.90	86.27	87.95
HOG	98.43	86.27	98.43	58.82	92.54	96.07
DWT (Haar)	95.68	86.27	94.90	97.25	89.90	90.91
DWT (Daub)	98.43	86.27	97.25	97.65	91.76	93.72
Dense + SIFT	97.64	86.27	95.29	94.90	86.66	97.45
Dense + SURF	92.94	86.27	92.94	77.64	88.23	85.88
FOS	86.27	87.05	86.27	90.98	86.27	91.76
DCT	**98.82**	86.27	92.94	97.64	86.27	88.62
SIFT + BoW	86.27	99.21	**99.21**	98.43	87.05	93.72
DWT + BoW	91.37	**99.61**	**99.21**	**98.82**	90.58	94.90
DWT + HOG (Proposed)	98.68	97.45	98.82	98.32	**93.46**	97.75

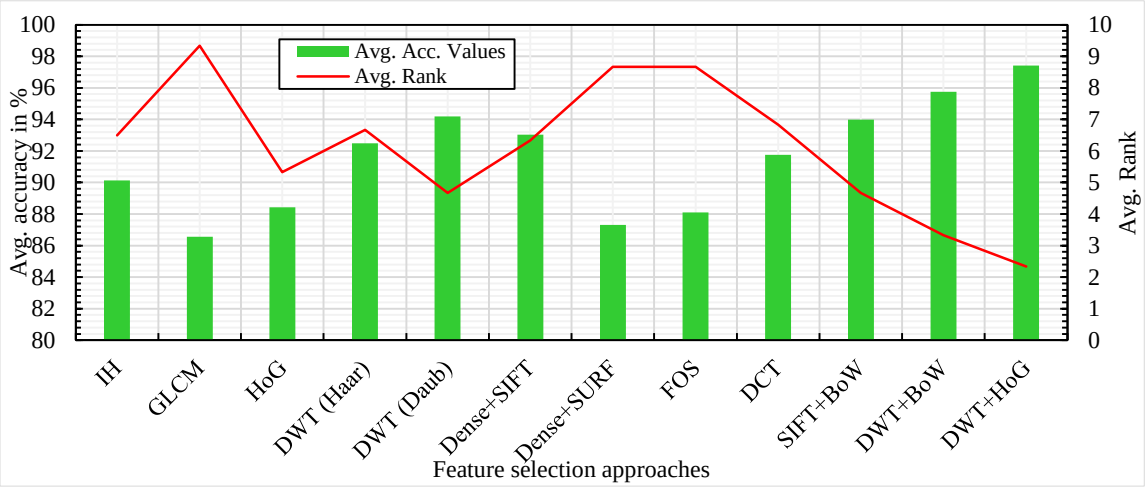

Figure 7 Performance analysis of different feature extraction approaches

Upon extracting features using the DWT + HOG approach, significant features are selected by applying a dimensionality reduction using MPCA. The selected features are trained using cross-validation by applying an RDA classifier. However, to analyse the performance of the RDA classifier, the results obtained for the selected features are trained using the RDA classifier and the obtained results are also compared with other classifiers such as LSVM, RSVM, PSVM, KNN, RF, AdaBoost, Logistic Regression (LR), QDA, and RDA classifiers. The various metrics such as accuracy (Acc.), sensitivity

(Se.), specificity (Sp.) and precision (Pre.) are employed to compare the performance of the models. The values obtained for DS-66, DS-160, DS-255 and DS-75 datasets for different classifiers after the first run are presented in *Table 4.*

From the study, the proposed RDA classifier offers better results for DS-66 and DS-75 datasets than for DS-160 and DS-255 datasets. Here, the LR and LSVM classifiers offer better results for DS-66 and DS-160 and RF classifier offers better results for DS-75. However, for DS-160 and DS-255, the RDA produces better results than many of the classier comparisons.

Table 4 Performance analysis of classification

Classifiers	DS-66				DS-160				DS-255				DS-75			
	Acc.	Se.	Sp.	Pre.	Acc.	Se.	Sp.	Pre.	Acc.	Se.	Sp.	Pre.	Acc.	Se.	Sp.	Pre.
LSVM	100	100	100	100	100	100	100	100	95.7	100	51.2	92.2	100	100	100	100
RSVM	98.3	100	97.4	96.3	100	100	100	100	99.7	100	98.5	99.6	98.3	98.7	69.3	96.3
PSVM	96.3	100	96.8	96.8	99.9	100	96.8	99.4	99.7	100	96.3	99.2	98.2	97.4	98.3	98.0
KNN	99.2	98.3	100	98.3	96.6	94.2	97.6	96.5	99.2	98.3	99.3	98.5	94.5	85.3	95.1	96.3
RF	85.2	83.3	84.2	90.1	97.5	94.3	80.3	95.8	91.9	98.3	68.7	92.7	100	100	100	100
AdaBoost	88.3	90.9	78.9	90.0	93.6	98.6	70.3	95.1	96.3	97.5	88.3	98.2	99.7	99.9	99.1	98.5
LR	100	100	100	100	100	100	100	100	98.3	100	99.1	98.8	98.5	99.9	85.7	98.3
QDA	100	100	100	100	99.0	100	97.7	98.3	97.3	93.3	91.3	94.3	98.9	99.4	98.4	97.3
RDA	100	100	100	100	99.3	100	98.6	99.5	99.5	98.6	96.6	99.4	100	100	100	100

When comparing the average accuracy, RDA, LR and RSVM take the top three positions with an average value greater than 99%. In the case of average sensitivity, LSVM, LR, RSVM and RDA takes the top 4 positions with values greater than 99%. For the average values of specificity and precision, RDA, KNN, PSVM and RDA, LR, and PSVM acquire the top three positions respectively. Thus, RDA and LR classifiers offer better performance than other classifiers. The average values obtained for the performance metrics used with different classifiers are shown in *Figure 8.*

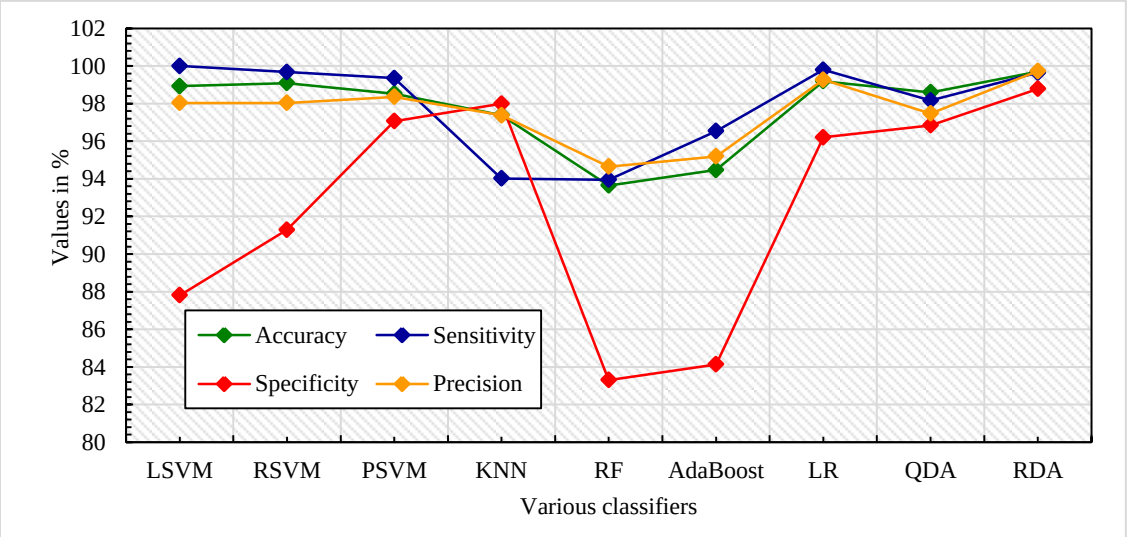

Figure 8 Average performance of different classifiers

4.4Analysis using existing models

This section discusses the performance analysis of the proposed model and the comparison with the other existing models in the field of the research study. *Table 5* shows the accuracy of various models for the datasets DS-66, DS-160 and DS-255. Most of the existing models iterate for 5 runs or 10 runs. The proposed model estimates the accuracy for 10 iterations. The results obtained also include the number of features selected for the classification. Each model is identified based on the methods used at each phase such as feature extraction (FE), feature selection (FS) and the classifiers (C) used in

classification. Thus, the methods are represented as FE + FS + C.

Additionally, an analysis has been made for evaluating the performance of the proposed model using various other metrics such as sensitivity, specificity, f-measure, MCC, ROC and error rate using the two datasets such as DS-75 and DS-160. The obtained results are also compared with the other existing models that were evaluated using these datasets in the literature. *Table 6* shows the obtained results under different performance metrics for the datasets DS-75 and DS-160. Most of the existing models iterate for 5 runs or 10 runs, though, the proposed model estimates the values using 10 iterations.

The proposed model is also evaluated for its efficiency based on the computation time. It is one of the important factors to be considered for analysing the model's efficiency. Here the computation time is identified for classifying the image as a normal or disease image and so the time is taken to train the model is not taken for analysis. Thus, the average time taken for the proposed model in processing images from the DS-255 dataset is presented in *Figure 9.*

The total computational time includes the time taken for feature extraction (0.014s), feature selection (0.006s) and classification (0.005s). Thus, the overall computation time taken by the proposed model is 0.025 seconds. This is minimum when compared with the other existing models specifically the model (DWT + BoW + SVM) has higher accuracy that even takes 0.027 seconds [22].

Table 5 Accuracy of prediction model comparison

Models (FE + FS + C)	**# of feat.**	**DS-66**	**DS-160**	**DS-255**
CNN + ReLu [19]	13	100	98.32	97.16
DWT+BoW+SVM [22]	-	100	100	99.61
DWT+PPCA+RSE [24]	13	100	100	99.2
DWT+PCA+RSE [24]	13	100	99.57	98.9
DWT+ PPCA+ADBRF [25]	13	100	100	99.53
DWT+PCA+ADBRF [25]	13	100	99.3	98.44
DWT+SVM+POLY [26]	4761	98	97.15	96.37
DWT+SVM+RBF [26]	4761	98	97.33	96.18
DWT+NBC [27]	7	-	87.5	-
DWT+PCA+KNN [28]	7	98	97.54	96.79
DWT+PCA+FPANN [28]	7	97	96.98	95.29
DWT+PCA+FNN+ACPSO [29]	19	100	98.75	97.38
DWT+PCA+FNN+SCABC [30]	19	100	98.93	97.81
DWT+PCA+BPNN+SCG [31]	19	100	98.29	97.14
DWT+PCA+KSVM+GRB [32]	19	100	99.38	98.82
DWT+KPCA+LSSVM [33]	7	-	87.5	-
RT+PCA+LSSVM [34]	9	100	100	99.39
DWT+SUR+SVM [35]	7	-	98.6	-
DWT+PPCA+FC [35]	13	-	100	-
DWT+WE+SWP+PNN [36]	3	100	99.88	98.9
WE+HBP-FNN [37]	6	100	100	99.49
DWPT+SE+GEPSVM [38]	16	99.85	99.62	98.78
DWPT+SE+GEPSVM [38]	16	100	100	99.33
FPCNN+DWT+PCA+FPANN [40]	7	100	98.88	98.43
SWT+PCA+IABAP-FNN [43]	7	100	99.44	99.18
SWT+PCA+ABC-SPSO-FNN [43]	7	100	99.75	99.02
FRFE+DP-MLP+ARCBBO [44]	12	100	99.19	98.24
FRFE+BDP-MLP+ARCBBO [44]	12	100	99.31	98.12
SWT+SUR+ADBSVM [45]	7	-	100	-
DWT+SUR+ADBSVM [45]	7	-	99.2	-
DWT+HOG+MPCA+RDA (Proposed)	13	100	100	99.58

Table 6 Accuracy of prediction model comparison

Models	Sensitivity		Specificity		F-Measure		MCC		ROC		Error Rate	
	DS-75	DS160	DS-75	DS160	DS-75	DS160	DS-75	DS160	DS-75	DS160	DS75	DS160
DWT + PPCA + ADBRF [25]	100	100	100	100	1	0.988	1	1	1	1	0	0
DWT + PCA + ADBRF [25]	100	95.2	100	100	1	0.994	1	0.972	1	1	0	0.7
DWT + NBC [27]	∞	∞	80	87.5	-	-	-	-	0.682	0.834	-	12.5
DWT + PCA + KNN [28]	100	90	100	98.5	1	0.975	1	0.886	1	0.934	2	2.46
DWT + PCA + FPANN [28]	93.7	89.5	100	97.8	0.99	0.969	0.96	0.855	1	0.935	3	3.02
DWT + KPCA + LSSVM [33]	100	∞	82.1	87.5	0.826	-	0.331	-	0.567	0.5	-	12.5
DWT + SUR + SVM [35]	100	90	100	96.4	1	0.975	1	0.815	1	0.932	0	1.4
DWT + PPCA + FC [35]	100	100	100	100	1	1	1	1	1	1	0	0
SWT + SUR + ADBSVM [45]	100	100	98.3	100	0.987	0.975	0.958	1	1	1	0	0
DWT + SUR + ADBSVM [45]	100	95.1	98.3	100	0.986	0.975	0.958	0.971	0.967	1	0.2	0.8
DWT + HOG + MPCA + RDA (Proposed)	100	100	100	100	1	1	1	1	1	1	0	0

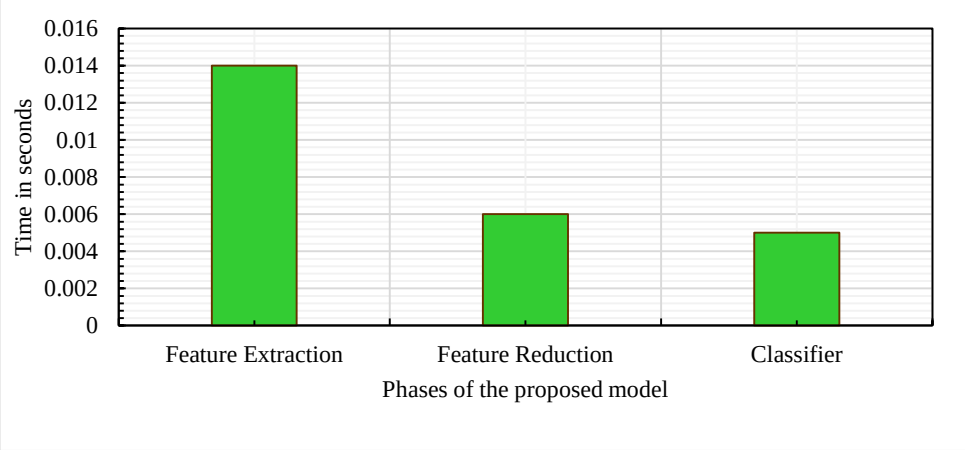

Figure 9 Average computation time of the proposed model

5.Discussion

This section discusses the insights and key findings from the experiments and results analysis, implications as well as the research limitations. The empirical analysis made for the proposed model in section 4 can be divided into two parts. The first part evaluates the approaches used in different phases with that of the other conventional approaches. The second part evaluates the overall proposed model with that of the other existing models identified from the literature review. These are discussed in sections 4.3 and 4.4 respectively.

In the first part, the accuracy of the feature extraction from brain images using various conventional methods with the DS-255 dataset has been evaluated. The feature extraction approaches are analysed with six different classifiers such as LSVM, RSVM, PSVM, KNN, RF and AdaBoost. It is found that the average accuracy, using different classifiers for the proposed DWT + HOG approach is 97.41% with an average rank of 2.33 which is better than other approaches.

The RDA classifier used in the proposed model has been analysed with other traditional classifiers suitable for the study such as LSVM, RSVM, PSVM, KNN, RF, AdaBoost, LR, and QDA. The evaluation has been made with four datasets such as DS-66, DS-160, DS-255, and DS-75 using four different metrics such as accuracy, specificity, sensitivity and precision. The proposed model offers 100% results for DS-66 and DS-75 and 99% for DS-160 datasets respectively. With DS-255, the model shows lower performance nonetheless it is better than the other classifiers used in the study.

With the overall evaluation, the proposed prediction model DWT + HOG + MPCA + RDA offers better results with 100% for DS-66 and DS-160 datasets. In the study, most of the existing models offer better results as 100% accuracy with DS-66 datasets. On the other hand, for the dataset DS-160 only 9 algorithms, including the proposed model offer 100% accuracy out of 30 models such as RT + PCA + LS-SVM, WE + HBP-FNN, DWPT + SE + GEPSVM, SWT + SUR + ADBSVM, DWT + BoW + SVM and DWT + PPCA + FC. With the DS-255 dataset, the model

DWT + BoW + SVM offers good results of 99.61% accuracy and the proposed model offers an accuracy of 99.58% which is still better than most of the existing algorithms. However, no prediction models offer 100% accuracy for the dataset DS-255 as it contains different types of diseased brain images.

Apart from accuracy, the sensitivity, specificity, F-Measure, MCC, the ROC and error rate have been analysed for the proposed model and compared with various existing models. From the performance analysis, it is clear that the proposed model can be considered one of the best models which has 100% results similar to the DWT + PPCA + FC model. The specificity and sensitivity for DS-75 and DS-160 is 100% and that of the values for F-measure, MCC and ROC is also 1 with a 0% error rate. It is evident that the proposed model is highly suitable for the prediction of the neurodegenerative disease dataset. Additionally, the model offers better results in minimum time duration. Thus, from the analysis, it is clear that the proposed model is superior to most of the existing models used for comparison.

While evaluating the various parameters used for evaluation, the model has better values for accuracy, precision, and sensitivity whereas the model has lower values for specificity. This shows that the proposed model has higher misclassification rates for the normal brain images. This arises due to the class imbalanced problem. The pathological images in DS-66, DS-75, Ds-160 and DS-255 are 72%, 80%, 87.5% and 86%, respectively which is very high when compared with the normal images in the datasets. Thus, it is identified that the performance of the model is highly correlated with the number of normal and pathological brain images in the training and test samples in the dataset. Balanced images in both classes improve the performance of the overall model.

5.1Limitations of research study

Though there exist several ways to predict neurodegenerative disorders such as analysing the speech signals and gait parameters, the proposed model is limited to brain images. Thus, it would be difficult to diagnose the disease at an early stage using the proposed model since the disease symptoms would not be seen by others. The model offers a limited accuracy rate for DS-255 datasets and still lacks 100% accuracy due to the inclusion of the number of diseased images. The proposed model has been evaluated using brain images having different diseases, but when it converges to the brain images having specific neurodegenerative diseases like PD, the model can provide good results. This shows that the proposed model offers more accurate results with a smaller number of diseased images. Thus, the datasets can be created with a greater number of normal brain images than pathological brain images for higher prediction accuracy. A complete list of abbreviations is shown in *Appendix I.*

6.Conclusion and future work

This paper proposes a prediction model for classifying neurodegenerative diseases. It operates in two phases in which the first phase is the image processing phase and the second is the machine learning phase. The brain MR images are pre-processed by applying enhancement and noise filtering techniques. The quality images are then fed to a feature extraction approach that fetches the features using DWT and HOG. The extracted features are then processed by machine learning algorithms that undergo dimensionality reduction using MPCA and classification using RDA. The model is analysed using various datasets such as DS-66, DS-75, DS-160 and DS-255 containing publicly available brain MR images. The experimental analysis has been performed which shows a better accuracy of 100% for DS-66, DS-75, DS-160 datasets and 99.58% for DS-255 with a minimum computation time of 0.025s. The comparison of the performance of the proposed model with various existing models with different quality metrics ensures the significance of the proposed model in diagnosing the neurodegenerative disease more than other models. The future works aim at providing 100% accuracy and evaluating the model with the real-time neurodegenerative especially PD dataset.

Acknowledgment

None.

Conflicts of interest

The authors have no conflicts of interest to declare.

Author's contributions statements

Arunraj Gopalsamy: Conceptualization and design of the work, data acquisition, implementation, interpretation and analysis, writing and editing original draft. **B. Radha:** Conceptualization and design of the work, Supervision, analysis and interpretation of results and review of results. **K. Haridas:** Conceptualization and design of the work, Investigation on the result analysis. All authors discussed the results, contributed to the final manuscript and provided critical feedback and helped shape the research, analysis and manuscript.

References

[1] https://www.physio-pedia.com/Neurodegenerative_Disease. 10 September 2021.

[2] Taylor JP, Brown RH, Cleveland DW. Decoding ALS: from genes to mechanism. Nature. 2016; 539(7628):197-206.

[3] Gitler AD, Dhillon P, Shorter J. Neurodegenerative disease: models, mechanisms, and a new hope. Disease Models & Mechanisms. 2017; 10(5):499-502.

[4] https://www.wehi.edu.au/research-diseases/development-and-ageing/neurodegenerative-disorders. 10 September 2021.

[5] https://www.niehs.nih.gov/research/supported/health/neurodegenerative/index.cfm. 10 September 2021.

[6] Balogh EP, Miller BT, Ball JR. The diagnostic process. Improving Diagnosis in Health Care. National Academies of Sciences, Engineering, and Medicine. 2015.

[7] Ahmed Z, Mohamed K, Zeeshan S, Dong X. Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. Database. 2020.

[8] Myszczynska MA, Ojamies PN, Lacoste A, Neil D, Saffari A, Mead R, et al. Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. Nature Reviews Neurology. 2020; 16(8):440-56.

[9] Saeed U, Compagnone J, Aviv RI, Strafella AP, Black SE Lang AE, et al. Imaging biomarkers in Parkinson's disease and Parkinsonian syndromes: current and emerging concepts. Translational Neurodegeneration. 2017; 6(1):1-25.

[10] Jha D, Kwon GR. Development of an efficient cascade pathological-brain detection system using a median filter and quadratic discriminant analysis. IEIE Transactions on Smart Processing and Computing. 2018; 7(2):140-7.

[11] Maji P, Kundu MK, Chanda B. Second order fuzzy measure and weighted co-occurrence matrix for segmentation of brain MR images. Fundamenta Informaticae. 2008; 88(1-2):161-76.

[12] Ain Q, Jaffar MA, Choi TS. Fuzzy anisotropic diffusion based segmentation and texture based ensemble classification of brain tumor. Applied Soft Computing. 2014; 21:330-40.

[13] Wang H, Fei B. A modified fuzzy C-means classification method using a multiscale diffusion filtering scheme. Medical Image Analysis. 2009; 13(2):193-202.

[14] Liu W, Liu J, Peng T, Wang G, Balas VE, Geman O, et al. Prediction of Parkinson's disease based on artificial neural networks using speech datasets. Journal of Ambient Intelligence and Humanized Computing. 2022:1-4.

[15] Nagasubramanian G, Sankayya M. Multi-variate vocal data analysis for detection of Parkinson disease using deep learning. Neural Computing and Applications. 2021; 33(10):4849-64.

[16] Shalin G, Pardoel S, Lemaire ED, Nantel J, Kofman J. Prediction and detection of freezing of gait in Parkinson's disease from plantar pressure data using long short-term memory neural-networks. Journal of Neuroengineering and Rehabilitation. 2021; 18(1):1-15.

[17] Noor MB, Zenia NZ, Kaiser MS, Mamun SA, Mahmud M. Application of deep learning in detecting neurological disorders from magnetic resonance images: a survey on the detection of Alzheimer's disease, Parkinson's disease and schizophrenia. Brain Informatics. 2020; 7(1):1-21.

[18] Wang S, Zhang Y, Zhan T, Phillips P, Zhang YD, et al. Pathological brain detection by artificial intelligence in magnetic resonance imaging scanning (invited review). Progress in Electromagnetics Research. 2016; 156:105-33.

[19] Jyotiyana M, Kesswani N. A study on deep learning in neurodegenerative diseases and other brain disorders. In rising threats in expert applications and solutions 2021 (pp. 791-9). Springer, Singapore.

[20] Varuna SN, Kumar TN. Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network. Brain Informatics. 2018; 5(1):23-30.

[21] Khalil M, Ayad H, Adib A. Performance evaluation of feature extraction techniques in MR-Brain image classification system. Procedia Computer Science. 2018; 127:218-25.

[22] Ayadi W, Elhamzi W, Charfi I, Atri M. A hybrid feature extraction approach for brain MRI classification based on Bag-of-words. Biomedical Signal Processing and Control. 2019; 48:144-52.

[23] Huang GH, Lin CH, Cai YR, Chen TB, Hsu SY, Lu NH, Chen HY, Wu YC. Multiclass machine learning classification of functional brain images for Parkinson's disease stage prediction. Statistical Analysis and Data Mining: The ASA Data Science Journal. 2020; 13(5):508-23.

[24] Jha D, Kim JI, Choi MR, Kwon GR. Pathological brain detection using weiner filtering, 2D-discrete wavelet transform, probabilistic PCA, and random subspace ensemble classifier. Computational Intelligence and Neuroscience. 2017.

[25] Nayak DR, Dash R, Majhi B. Brain MR image classification using two-dimensional discrete wavelet transform and AdaBoost with random forests. Neurocomputing. 2016; 177:188-97.

[26] Chaplot S, Patnaik LM, Jagannathan NR. Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. Biomedical Signal Processing and Control. 2006; 1(1):86-92.

[27] Zhou X, Wang S, Xu W, Ji G, Phillips P, Sun P, Zhang Y. Detection of pathological brain in MRI scanning based on wavelet-entropy and naive Bayes classifier. In international conference on bioinformatics and biomedical engineering 2015 (pp. 201-9). Springer, Cham.

[28] El-dahshan ES, Hosny T, Salem AB. Hybrid intelligent techniques for MRI brain images classification. Digital Signal Processing. 2010; 20(2):433-41.
[29] Zhang YD, Wang S, Wu L. A novel method for magnetic resonance brain image classification based on adaptive chaotic PSO. Progress in Electromagnetics Research. 2010; 109:325-43.
[30] Zhang YD, Wu L, Wang S. Magnetic resonance brain image classification by an improved artificial bee colony algorithm. Progress In Electromagnetics Research. 2011;116: 65-79.
[31] Zhang Y, Dong Z, Wu L, Wang S. A hybrid method for MRI brain image classification. Expert Systems with Applications. 2011; 38(8):10049-53.
[32] Zhang YD, Wu L. An MR brain images classifier via principal component analysis and kernel support vector machine. Progress in Electromagnetics Research. 2012; 130:369-88.
[33] Nayak DR, Dash R, Majhi B. Least squares SVM approach for abnormal brain detection in MRI using multiresolution analysis. In international conference on computing, communication and security 2015 (pp. 1-6). IEEE.
[34] Das S, Chowdhury M, Kundu MK. Brain MR image classification using multiscale geometric analysis of ripplet. Progress in Electromagnetics Research. 2013; 137:1-7.
[35] Devi S, Sahoo MN, Bakshi S. A novel privacy-supporting 2-class classification technique for brain MRI images. Biocybernetics and Biomedical Engineering. 2020; 40(3):1022-35.
[36] Saritha M, Joseph KP, Mathew AT. Classification of MRI brain images using combined wavelet entropy based spider web plots and probabilistic neural network. Pattern Recognition Letters. 2013; 34(16):2151-6.
[37] Zhang YD, 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. Progress In Electromagnetics Research. 2015;152: 41-58.
[38] Zhang Y, Dong Z, Wang S, Ji G, Yang J. Preclinical diagnosis of magnetic resonance (MR) brain images via discrete wavelet packet transform with Tsallis entropy and generalized eigenvalue proximal support vector machine. Entropy. 2015; 17(4):1795-813.
[39] Wang S, Lu S, Dong Z, Yang J, Yang M, Zhang Y. Dual-tree complex wavelet transform and twin support vector machine for pathological brain detection. Applied Sciences. 2016; 6(6).
[40] El-dahshan ES, Mohsen HM, Revett K, Salem AB. Computer-aided diagnosis of human brain tumor through MRI: a survey and a new algorithm. Expert Systems with Applications. 2014; 41(11):5526-45.
[41] Mozhdehfarahbakhsh A, Chitsazian S, Chakrabarti P, Chakrabarti T, Kateb B, Nami M. An MRI-based deep learning model to predict Parkinson disease stages. medRxiv. 2021.
[42] Berke Erdaş Ç, Sümer E, Kibaroğlu S. CNN-based severity prediction of neurodegenerative diseases using gait data. Digital Health. 2022; 1-14.
[43] Wang S, Zhang Y, Dong Z, Du S, Ji G, Yan J, et al. Feed-forward neural network optimized by hybridization of PSO and ABC for abnormal brain detection. International Journal of Imaging Systems and Technology. 2015; 25(2):153-64.
[44] 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. Journal of Medical Systems. 2016; 40(7):1-11.
[45] Ranjan Nayak D, Dash R, Majhi B. Stationary wavelet transform and AdaBoost with SVM based pathological brain detection in MRI scanning. CNS & Neurological Disorders-Drug Targets (Formerly Current Drug Targets-CNS & Neurological Disorders). 2017; 16(2):137-49.
[46] Vidyasaraswathi HN, Hanumantharaju MC. Review of various histogram based medical image enhancement techniques. In proceedings of the international conference on advanced research in computer science engineering & technology 2015 (pp. 1-6).
[47] Pandey A, Chandra R, Dutta MK, Burget R, Uher V, Minar J. Automatic detection of red lesions in diabetic retinopathy using shape based extraction technique in fundus image. In international conference on telecommunications and signal processing 2016 (pp. 538-42). IEEE.
[48] Bäckström K, Nazari M, Gu IY, Jakola AS. An efficient 3D deep convolutional network for Alzheimer's disease diagnosis using MR images. In 15th international symposium on biomedical imaging 2018 (pp. 149-53). IEEE.
[49] Esmaeilzadeh S, Yang Y, Adeli E. End-to-end parkinson disease diagnosis using brain MR-images by 3d-CNN. arXiv preprint arXiv:1806.05233. 2018.
[50] https://radiopaedia.org/cases/37605. Accessed on 7 October 2021.
[51] Herrera LJ, Rojas I, Pomares H, Guillén A, Valenzuela O, Baños O. Classification of MRI images for Alzheimer's disease detection. In international conference on social computing 2013 (pp. 846-51). IEEE.
[52] Wulandari P, Novitasari D, Asyhar A. Identification of Alzheimer's disease in MRI data using discrete wavelet transform and support vector machine. In proceedings of the international conference on mathematics and Islam 2018 (pp. 198-204).
[53] Johnson RA, Wichern DW. Applied multivariate statistical analysis. London, UK: Pearson; 2014.
[54] Lu H, Plataniotis KN, Venetsanopoulos AN. Multilinear principal component analysis of tensor objects for recognition. In international conference on pattern recognition 2006 (pp. 776-9). IEEE.
[55] Vasilescu MA, Terzopoulos D. Multilinear subspace analysis of image ensembles. In computer society conference on computer vision and pattern recognition, 2003 (pp. 1-7). IEEE.

[56] Fu R, Han M, Tian Y, Shi P. Improvement motor imagery EEG classification based on sparse common spatial pattern and regularized discriminant analysis. Journal of Neuroscience Methods. 2020.
[57] Lee CC, Huang SS, Shih CY. Facial affect recognition using regularized discriminant analysis-based algorithms. EURASIP Journal on Advances in Signal Processing. 2010:1-10.
[58] Krasoulis A, Nazarpour K, Vijayakumar S. Use of regularized discriminant analysis improves myoelectric hand movement classification. In 8th international IEEE/EMBS conference on neural engineering (NER) 2017 (pp. 395-8). IEEE.
[59] Johnson KA. The whole brain atlas. 2012.
[60] Bama SS, Ahmed MI, Saravanan A. A survey on performance evaluation measures for information retrieval system. International Research Journal of Engineering and Technology. 2015; 2(2):1015-20.

Arunraj Gopalsamy is currently pursuing Ph.D from Bharathiar University and his core interest includes Data Mining, Big Data Analytics, Cognitive application Development. He has written 2 books and also has done 3 Patent Publications. He is working as a Chief Architect in Kyndryl Solutions Private Limited, India.
Email: arunrajgs.phd@gmail.com

Dr. B. Radha is presently working as Assistant Professor in Sri Krishna college of Arts and Science, Coimbatore. She has completed her Ph.D from Anna University, Chennai. She has attended and published many papers in International/National conferences and also has written many books in different areas. Her publication is included in Scopus and web of science indexed journals. She has acted as a chairperson in many international conferences. Reviewer in IEEE as well few more. Her current research interests are Data Mining, Networking and Cloud Computing.
Email: radhakbr10@gmail.com

Dr. K. Haridas is presently working as Associate professor and Head, Dept of Computer Applications, NGM College, Coimbatore, India (affiliated to Bharathiar University, Coimbatore). He is a SBC and a member of Computer Society of India. He has published many papers in international/national journals and written books. His areas of interest include Data Mining and Networking. He has to his credit 21 years of teaching and research experience. His current research interests include Grid Computing, Cloud Computing, and Semantic Web.
Email: hari_das74@yahoo.com

Appendix I

S. No.	Abbreviation	Description
1	ABC	Artificial Bee Colony
2	ACPSO	Adaptive Chaotic Particle Swarm Optimization
3	AD	Alzheimer's Disease
4	ADBRF	Adaboost with Random Forest Classifier
5	ADBSVM	Adaboost based Support Vector Machine
6	ARCBBO	Adaptive Real-Coded Biogeography Based Optimization
7	BBO	Biogeography Based Optimization
8	BoW	Bag of Words
9	BPNN	Backpropagation Neural Networks
10	CNN	Convolutional Neural Network
11	DCT	Discrete Cosine Transform
12	DLDA	Diagonal Linear Discriminant Analysis
13	DQDA	Diagonal Quadratic Discriminant Analysis
14	DTCWT	Dual-Tree Complex Wavelet Transform
15	DWPR	Discrete Wavelet Packet Transform
16	DWT	Discrete Wavelet Transform
17	FC	Filtered Classifier
18	FNN	Forward Neural Network
19	FOS	First Order Statistics
20	FPANN	Feed-Forward Backpropagation Artificial Neural Network
21	FPCNN	feedback pulse-coupled neural network
22	FRFE	Fractional Fourier entropy
23	FS	Feature Selection
24	GEPSVM	Generalized Eigenvalue Proximal Support Vector Machine
25	GLCM	Gray Level Co-occurrence Matrix
26	GRB	Gaussian Radial Basis
27	HBP-FNN	Hybridization of BBO and PSO
28	HE	Histogram Equalization
29	HOG	Histogram Oriented Gradients
30	IH	Intensity Histogram
31	KNN	K-Nearest Neighbour
32	KSVM	Kernel Support Vector Machine
33	LDA	Linear Discriminant Analysis
34	LR	Logistic Regression
35	LSSVM	Least Squares Support Vector Machine
36	LSVM	Linear Support Vector Machine
37	MCC	Matthews Correlation Coefficient
38	MPCA	Multilinear Principal Component Analysis
39	MR	Magnetic Resonance
40	MRI	Magnetic Resonance Imaging
41	NBC	Naive Bayes Classifier
42	PCA	Principal Component Analysis
43	PD	Parkinson's Disease
44	PET	Positron Emission Tomography
45	PMRS	Proton Magnetic Resonance Spectroscopy

46	PNN	Probabilistic Neural Network
47	POLY	SVM with Polynomial Kernel
48	PPCA	Probabilistic Principal Component Analysis
49	PSO	Particle Swarm Optimization
50	PSVM	Polynomial Linear Support Vector Machine
51	QDA	Quadratic Discriminant Analysis
52	RBF	Radial Basis Function
53	RDA	Regularized Discriminant Analysis
54	RF	Random Forest
55	ROC	Receiver Operating Characteristic
56	RSE	Random Subspace Ensemble
57	RSVM	Radial Support Vector Machine
58	SCABC	Scaled Chaotic Artificial Bee Colony
59	SCG	Scaled Conjugate Gradient
60	SE	Shannon Entropy
61	SIFT	Scale Invariant Feature Transform
62	SPECT	Single Photon Emission Computed Tomography
63	SUR	Symmetric Uncertainty Ranking
64	SURF	Speed Up Robust Features
65	SVM	Support Vector Machine
66	SWP	Spider Web Plots
67	SWT	Stationary Wavelet Transform
68	WE	Wavelet Entropy