Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 4, April 2021: 2069-2083

A Voxel Based Morphometry Approach for Identifying Alzheimer From MRI Images Using an Optimized PSO Algorithm

N.Balaraman[1], Dr. B.Rajalingam[2], Dr. R.Santhoshkumar[3], Dr. P. Santosh Kumar Patra[4]

ABSTRACT

Alzheimer's Disease (AD) is a commonly occurring brain disorder that affects elderly people. It is a progressive, neurodegenerative brain disorder that attacks neurotransmitters, and causes dementia. For the evaluation of normal ageing and AD, Voxel Based Morphometry (VBM) using structural brain Magnetic Resonance Imaging (MRI) has been widely used. This VBM of MRI has data that has been segmented as Gray Matter (GM), White Matter (WM), and Cerebro-Spinal Fluid (CSF) partitions. Anatomical standardization of all the images to the same stereotactic space is done. It makes use of linear affine transformation as well as non-linear warping, smoothing and at last performs statistical analysis. The work suggests the following- Particle Swarm Optimization (PSO) based AdaBoost, using Principal Component Analysis (PCA) for feature reduction and feature extraction using curvelet transform classifier optimization. It is not completely possible by the curvelet transform to characterize the high dimensional signals that contain hyper plane singularities, lines or curves. For decreasing the data set dimensions that contain several interrelated variables, PCA is an effective tool and it can also retain most of the differences. The work also presents an improvised AdaBoost algorithm that is based on optimizing the sample space search. In order to find a threshold in AdaBoost algorithm, more time is needed for comparing samples while working with data on a large scale while making use of the decision stump as a weak classifier. This work makes use of the PSO algorithm in order to change and also choose the most optimal feature in sample space for weak classifiers to reduce computation time. It has been shown via empirical outcomes that the suggested technique performs better compared to the other techniques.

Keywords: Alzheimer's Disease (AD), Voxel Based Morphometry (VBM), Magnetic Resonance Imaging (MRI), Curvelet Transform, Principal Component Analysis (PCA), Particle Swarm Optimization (PSO) and Adaboost Classifier.

1. INTRODUCTION

Usually, AD affects people who are over 65 years of age; yet, early symptoms of this fatal neurodegenerative disorder can be detected before 65 years of age. Neuron cells in the brain die when two abnormal protein fragments known as plagues and tangles in the brain are deposited. The first region to be affected by AD is the hippocampus. This is where the memories are initially formed. The initial symptoms of the disease include issues with memory like problems in finding words and also in the process of thinking. Patients who suffer from AD have an issue

---

[1]Assistant Professor, [2,3]Associate Professor, [4]Principal & Professor
[1,2,3,4] Department of Computer Science and Engineering, St. Martin's Engineering College, Secunderabad, Telangana, India
[1]padmabala.raj@gmail.com, [2]rajalingam35@gmail.com,[3]santhoshkumar.aucse@gmail.com

with behavioural changes, personality variations, lack of initiatives, even affecting the routine functions at home or in the workplace, slowly leading to their death. As AD progresses, the volume of the brain progressively decreases and most of the brain functions are eventually affected [1]. AD will have severe socioeconomic implications as the number of elderly people in developing countries is on the rise. A recent report suggests that the number of people affected by AD will double in the coming two decades; also, every other person aged above 85 will suffer from AD by 2050. Thus, early and precise detection of ADHD is extremely important. Traditionally, a neuropsychological examination in support of structural imaging is used for the diagnosis of AD [2].

One of the most popular brain imaging techniques for obtaining accurate information about the shape and the volume of the brain is structural MRI which is non-invasive. MRI can not only detect any irregularity in the brain but also provide better soft tissue differentiation, high spatial resolution and also better contrast, compared to Computed Tomography (CT) and Positron Emission Tomography (PET) scans. Besides providing rich information regarding the brain condition of the patient, it is used actively for AD identification. For formulating a computer aided brain disease diagnosis system with neuro images like the MRI, PET as well as functional MRI (fMRI) and Diffusion Tensor Imaging (DTI), presently, machine learning and pattern classification techniques have been widely used. It has been demonstrated that in clinical practice, structural MRI is a highly standardized imaging technique, and this has also proven to be useful for tracking the varying clinical phases of the AD. Hence, based on structural Magnetic Resonance (MR) images, this technique is evaluated. Many features like intensities or GM densities, texture measures, morphometry, group comparison of cortical thickness are extracted from the structural MRI of the entire brain. Compared to the techniques that just make use of single features, this combination of different types of features can improvise the precision of AD diagnosis [3].

Voxel based method that has been used for comparing the volumes of GM between different groups of subjects is VBM. These are studied with MRI. In common anatomical space spatial modifications of the MRI scans of all subjects that have been researched has been done with this method. Images are restricted by linear as well as non-linear transformations in such a spatial normalization process to a customized template that has specially been formulated for the work or by a standardized template. Following this, every subject's images have been automatically characterized as GM, WM and cerebral spinal fluid partitions and a Gaussian filter is used to smooth this [4]. For each and every voxel, GM segments are then compared statistically between the groups; to show the location of the voxel clusters where many changes in mean grey matter volumes exist, statistical maps are produced. These maps are at a predefined statistical inference level. The advantages are presented by the voxel based approach compared to the Region-of-interest techniques. The reason is as it is completely automated, independent and has the ability to investigate the presence of AD related morphometric GM anomalies across the entire brain in huge AD sample subjects, comparative with healthy controls having a high capacity to reproduce highly.

For eliminating the mis-segmented non-GM voxel areas, optimized VBM technique has been formulated. Additional steps for pre-processing are used for removing non-brain voxels before the anatomical standardization. This is followed by segmentation. To enable accurate segmentation, the optimized standardization parameters are reapplied to the original, whole brain structural images in native space. The standardized whole brain structural images that are optimal are then divided into gray and WM then the CSF partitions and subject to a second extraction of standardized segmented gray/WM images. As a few of the non-brain voxels from the scalp, skull or the venous sinuses in the normalized brain have the possibility to still be exterior to the brain margins on partitioned gray/ WM images, the brain extraction step is repeated [5]. One of the most commonly encountered decision-making task of human activity is classification. An issue in classification takes place when an object should be allocated to a predefined group or a class on the basis of the attributes that are observed and related to that object. Several of the industrial problems have been identified as classification problems like character recognition, speech recognition, medical diagnosis, predicting bankruptcy, weather forecasting and stock market prediction. Both mathematically and in non-linear fashion, these classification problems can be solved. The precision and the distribution of data properties and capabilities of the models determine the difficulty in solving such of the problems mathematically [6].

Another name that is given for classification is supervised learning- this is because the instances are mentioned with familiar labels, as opposed to unsupervised learning wherein labels are unknown. A set of features or attributes that may be either continuous or categorical represent every data set instance that is used by supervised and unsupervised learning techniques. The process wherein a model is constructed from the training set which comprises database instances and associated labels is referred to as classification. The class label of the testing instances is predicted by the resulting model wherein the predictor features value is known. This supervised learning technique is a common task executed by intelligent techniques, there are several techniques that have been developed [7]. For identifying the AD from MRI, this work suggests an optimized AdaBoost classifier or the PSO AdaBoost in the VBM approach. The remainder of the study has been classified into the following sections: The techniques of this work are explained in the third section; the related works in literature are discussed in the second section. The empirical outcomes are discussed in the fourth section and the work is concluded in the fifth section.

2. RELATED WORKS

A new and non-linear metric learning technique for improvising biomarker identification for AD and Mild Cognitive Impairment (MCI) has been developed by Shi et al., [8]. The suggested method that has been developed using limited optimization framework makes use of a smooth non-linear feature space transformation; this makes the mapped data more linearly separable for Support Vector Machines (SVMs). As the Thin-Plate Apline (TPS) has a considerable versatility and also power for representation to spawn complex yet smooth deformations, it is selected as the geometric model. Additionally, the cross sectional and the longitudinal estimated from brain MRIs are integrated using a deep network based feature fusion strategy through stacked Denoising Sparse Auto-Encoder (DSAE). An ensemble of the SVMs that integrates bagging at the same time not replacing and feature selection has been proposed by Sorensen et al., [9]. In multivariate classification of dementia, SVM is almost all pervading and hence it is of great value to evaluate the potential benefits of this classifier ensemble. The ensemble SVM uses either a linear or a Radial Basis Function (RBF) kernel. It has attained multi-class classification precision of 55.6%in the former and 55.0% inthe latter case, in the challenge test set (60 Normal Control (NC), 60 MCI, 60 converting MCI (cMCI), 60 AD). This has resulted in 3rd rank in the challenge. Similar feature sizes are attained for the other two kernels. The MRI features that are selected most often are the volumes of two hippocampus sub areas left subiculum and right subiculum.

For predicting the diagnosis of subjects who suffered from MCI by using their baseline structural MRIs, a new Convolutional Neural Network (CNN) based on VBM analysis was suggested by Çitak-ER et al., [10]. The baseline structural MRIs of patients have been made use for the prediction model's training and evaluation. Significant Volume of Interests (VOIs)that associated with the GM damage can be generated from the baseline structural MRIs. Hence, for extracting prognostic characteristics from MRIs, a CNN was trained using a set of Convolutional feature detectors that are obtained by training a patch-based auto encoder. An accuracy of about 78.7% was achieved using this work which was about 4% more than a reference article so that the risk that a patient could develop AD with MCI could be predicted. Jha et al., [11] proposed a Dual-Tree Complex Wavelet Transform (DTCWT) for extracting features from an image. The PCA was used for reducing the dimensionality of feature vector. This decreased feature vector was forwarded to the Feed Forward Neural Network (FNN) so that AD and Healthy control (HC) could be differentiated from the input MRIs.High and reproducible accuracy rates of $90.06 \pm 0.01\%$ with a sensitivity of $92.00 \pm 0.04\%$, a specificity of $87.78 \pm 0.04\%$, and a precision of $89.6 \pm 0.03\%$ with 10-fold cross-validation resultedwith these proposed and implemented pipelines, which demonstrate improvements in classification output when compared to that of recent investigations.

In order to separate pathological brain from normal brains in MRI scanning, Ranjan Nayak et al., [12] has suggested a system of categorization that is automated. For increasing the identification of the affected areas in brain MRIs, the system uses contrast limited adaptive histogram equalization scheme.For extraction of features from pre-processed images, two-dimensional stationary wavelet transform is harnessed. Using the energy as well as entropy values, the feature vector is constructed which is calculated from the level-2 Stationary Wavelet Transform (SWT) coefficients. After this, the symmetric uncertainty ranking filter is used for selecting the relevant and the uncorrelated features. After this, to the suggested AdaBoost with SVM classifier, the selected features are given as inputs. Here, the SVM is used as the base classifier of AdaBoost algorithm A new machine learning system which can automatically diagnose from the MRIs was suggested by Zhang et al., [13]. The first step was the processing of brain images which also comprised skull stripping as well as spatial normalization. In the second step from the volumetric image, one axial slice was chosen. The texture features could be extracted using the Stationary Wavelet Entropy (SWE). Next, as the classifier, one hidden layer NN was utilized. Finally, for training the weights and the biases of a classifier, a predator-prey PSO has been suggested. A total precision of$92.73\pm1.03\%$, a sensitivity of $92.69\pm1.29\%$, and a specificity of $92.78\pm1.51\%$ has been yielded by the classification.

An automatic dementia MRI classification that used the machine learning techniques was presented by Valarmathy & Vanitha [14]. MRI images from Open Access Series of Imaging Studies (OASIS) dataset are used for evaluating. Discrete Wavelet Transform (DWT) is used for MRI images that are segmented and features are extracted from the segmented image. Artificial Immune System (AIS) is used for proposing feature selection; The correlation based feature selection is searched in the solution space. The selected features are then classified as dementia or non-dementia by Naïve Bayes, Classification and Regression Tree (CART), C4.5 and K

Nearest Neighbour (KNN). For classifying the brain tumor of the given MRI brain image as either normal or not, a hybrid optimization classification technique has been suggested by Ahmed et al., [15]. For achieving the MRI classification accuracy by means of selecting the optimal ANN parameters, the suggested system makes use of a Gray Wolf Optimizer (GWO) combined with a supervised Artificial Neural Network (ANN) classifier. By making use of the receiver operating character analysis, GWO–ANN classification system performance that has been introduced is compared to the traditional Neural Network (NN) classifier [26-29].

With the objective of differentiating normal and abnormal brains in MRI scanning, a new PSO and ABC based classification system has been suggested by Wang et al., [16]. For extracting the features from the brain MRIs, SWT has been used which is invariant to translation and even when there was a slight translation in the image, it has performed well. The next step was reducing the SWT coefficients using the PCA. It has been shown via the ten runs of the K-fold cross validation outcome that the suggested PSO and ABC Feed Forward Neural Network (HPA-FNN) hybridization lead to superior performance compared to the two suggested classifiers and also the state-of-the-art techniques that existed, with respect to the accuracy of classification. The Binary PSO (BPSO) approach has been extensively studied by Wang et al., [17]. Three of the new variants have been suggested - BPSO with Mutation and Time-varying acceleration coefficients (BPSO-MT), BPSO with Mutation (BPSO-M), and BPSO with Time-varying acceleration coefficients (BPSO-T). First the Wavelet Entropy (WE) features have been extracted from both the approximation as well as the detailed sub bands of 8-level decomposition. Later, for selecting the features, this work suggests the use of BPSO-M, BPSO-T, and BPSO-MT. These features that have been selected were finally fed to a Probabilistic Neural Network (PNN) [30-35].

3. METHODOLOGY

This work makes use of the Alzheimer's disease Neuro imaging Initiative (ADNI) database. The various schemes that have been discussed are- classifier using AdaBoost and proposed Adaboost optimized PSO methods, PCA feature reduction methods, feature extraction using curvelet transform and optimization using PSO.

3.1 Alzheimer's disease Neuroimaging Initiative (ADNI) Database

The ADNI database (adni.loni.usc.edu) has been utilized for getting the data in formulating this article. This has been spearheaded by Michael W. Weiner, MD, the principal investigator, ADNI was launched in 2003 as a partnership of public and private enterprise. The objective of the ADNI is testing using PET, biological markers, and clinical and neuropsychological assessment can be used together for following up with the progression of the MCI as well as early onset of AD [18].

3.2 Feature Extraction Using Curvelet Transform

The ridgelet transform is extended by the curvelet transform to multi scale analysis. Hence, it begins with the definition of a ridgelet transform [19]. The continuous ridgelet coefficients for a given image are expressed as (1):

$$f(a,b,\theta) = \int\int \psi_{a,b,\theta}(x,y) f(x,y) dxdy \tag{1}$$

Here, a is the scale parameter where a > 0, b ∈ R is the translation parameter and θ ∈ [0, 2π). A ridgelet can be defined as (2):

$$\psi_{a,b,\theta}(x,y) = a^{-\frac{1}{2}\psi}\left(\frac{x\cos\theta + y\sin\theta - b}{a}\right) \tag{2}$$

Ridgelets do not vary along the lines and the ridgelet orientation is represented by θ. When the curvelet and the ridgelet transforms are both compared, in terms of capturing the edge information, it is noticed that curvelets at all the scaled can capture the edge information better and more tightly than wavelets. Ridgelet based curvelet transform is a hybrid of the wavelet transform and the Radon transform. The curvelet technique involves the decomposition of the input image into a set of sub bands. Each of this is then divided into many blocks for the analysis of ridgelet. The Radon transform, and single dimensional wavelet transform is used for the implementation of the ridgelet transform. One of the processes during the ridgelet transform is spatial partitioning that makes use of overlapping windows so that the blocking effects can be avoided. This leads to a lot of redundancy. Also, this process consumes a lot of time and this makes it infeasible for analysing the data in huge databases.

3.3 Feature Reduction Using PCA

A tool that can change the present input features into a novel lower dimension feature space is the PCA, which although makes use of the data that is extracted from the image will not improvise upon the original image. The largest eigenvectors of the correlation matrix are used in PCA for transforming the input feature space into a lower dimensional feature space. The most widely used subspace projection method is the PCA. PCA finds the linear lower-dimensional representation of the data such that the variance of the data is preserved when a data set has been given [20]. The feature vectors are limited to the component that is chosen by the PCA by making use of a system of feature reduction based PCA. This in turn results in an inefficient classification algorithm. Hence, decreasing the wavelet coefficients dimensionality is the chief idea of using the PCA and this leads to a more effective as well as a precise classifier. As much information as possible as is measured by the variance, is squeezed by the PCA into the initial main components. In certain other cases, the main parts can store the major variance is extremely miniscule. This can be looked upon as a great achievement in data manipulation. Any number of useful applications can result as the transformation can be performed very quickly on contemporary hardware and it is invertible.

3.4 Particle Swarm Optimization (PSO)

Modelled for simulating the social behaviour of bird flocks, the PSO is a population based stochastic optimization algorithm. Akin to EA as both are population based with a fitness function associated with every individual. Additionally, the arithmetic crossover operator used in the EAs is similar to the adjustments of the individuals in the PSO. Nonetheless, rather than the survival of the fittest, PSO is influenced by the simulation of the social behaviour. Yet another difference exists which is that every candidate gains from its past while in EAs this is absent. Being easy to implement, PSO has been used widely for solving several optimization problems like the discrete and the continuous optimization problems [21]. A swarm of individuals referred to as particles roam around the search space in PSO. A candidate solution to the optimization problem is represented by every particle. The position of a particle is influenced by best position visited by itself and the experience of the neighbourhood particles which refers to the position of the best particle in the neighbourhood. The best position of the neighbourhood is called the global best particle in the entire swarm. Gbest PSO refers to the resulting algorithm. The algorithm is usually known as lbest PSO for small sized neighbourhoods. A fitness function is used that changes depending on the optimization problems, the measure of how far a particle is from the global optimum is measured. Each particle in the swarm is represented by the following characteristics:

The existing location of the particle: x_i

Current speed of the particle: v_i

Best position of the particle: y_i

Neighbourhood best position of the particle: $\hat{y}_i$

The best position or the position that results from the best fitness value of a particle i visited so far is its personal best position. Let the objective function be denoted by f [22]. Then at time step t, the personal best of a particle is updated as (3):

$$y_i(t+1) = \begin{cases} y_i(t) \ \ if \ f(x_i(t+1)) \geq f(y_i(t)) \\ x_i(t+1) \ \ if \ f(x_i(t+1)) < f(y_i(t)) \end{cases} \tag{3}$$

The entire swarm determined the best particle for the gbest model by selecting the best personal best position. If the position of the global best particle is denoted by the vector $\hat{y}$, then (4):

$$\hat{y}(t) \in \{y_0, y_1, ..., y_s\} = \min\{f(y_0(t)), ..., (y_s(t))\} \tag{4}$$

Where the swarm size is shown as s. for each dimension j ∈ 1,…,N_d, The velocity update step is specified. Hence, $v_{i,j}$ represents the jth element of the velocity vector of the ith particle. The following equation (5) is used for calculating the velocity of the particle i:

$$v_{i,j}(t+1) = wv_{i,j}(t) + c_1 r_{1,j}(t)(y_{i,j}(t) - x_{i,j}(t)) + c_2 r_{2,j}(t)(\hat{y}_j(t) - x_{i,j}(t)) \tag{5}$$

Inertia weight =*w*, acceleration constants are c1 and c2 and $r_{1,j}(t)$, $r_{2,j}$ ~U (0, 1).

3.5 AdaBoost Classifier

There are T rounds in which AdaBoost works and a weak learner is trained during every round. The algorithm increases sample weights after training and this is predicted as misclassification and correspondingly the sample weight decreases which is correctly classified. This is how the correctly classified samples have lesser chances of being used in the next iteration whereas the chances for the incorrectly classified samples increases [23]. AdaBoost takes a training samples of $S = (x_1, y_1), \ldots\ldots, (x_n, y_n)$ with size N as input, where each sample x_i is a vector of values for domain of space X, and y_i is label of each sample x_i that belongs to label space of Y. When the algorithm begins iteration initially, the weights are initialized uniformly across the training set; The weights of every example that are not classified correctly are increased in every iteration. This is how a weak learner would pay attention to the hard samples on the data set. There is a weight vector which is associated with AdaBoost on the training samples; Here, it updates the weights of every sample using a weight function in every iteration (6).

$$W_{t+1,i} = W_{t,i}\beta_t^{1-bi} \tag{6}$$

It calculates and uses an error rate $\in_j$ of this classifier according to the weight of each sample in (7), to adjust the probability distribution for training samples:

$$\in_j = \sum_{i=1} W_{t,i} bi \tag{7}$$

Every classifier is weighed as per its accuracy during the training phase helps construct a final strong classifier H (x). Hence, the focus of the algorithm is on pattern samples that are hard to classify. This work focuses on binary classification problems in which Y = {1, +1}. The pseudo code of the algorithm is as follow:

$Given\ N\ examples\ (x_1,\ y_1),..(x_i,\ y_i)..,(x_N,\ y_N)$
$where\ y_i \in \{-1,\ 1\}$
$Initialize\ w_{1,i} = 1/2m,\ 1/2l\ for\ y_i = 0.1$
$respectively,\ where\ m\ and\ l\ are\ the\ number\ of$
$negatives\ and\ positives\ respectively.$
$for\ t = 1,...,T\ do$
$(1)\ for\ each\ feature\ j,\ train\ a\ classifier\ h_j()$
$(2) Evaluate\ the\ error\ of\ the\ classifier\ \in_t = \sum_{i=0} w_{t,i}.b_i$
$(3)\ Choose\ a\ classifier\ h_t()\ with\ lowest\ error\ \in_t$
$Update\ weights:\ w_{t+1,i} = w_{t,i}\beta_t^{1-b_i}$
$where\ b_i = 0\ if\ h_t(x_i) = y_i, b_i = 1$
$with\ \beta_t = \in_t /(1- \in_t)$
$end\ for$
$Output\ strong\ classifier:$

$$H(x) = \begin{cases} 1\ \ if\ \ sign(\sum_{i=0} \alpha_t h_t)\ is\ positive \\ -1 \qquad\qquad\qquad otherwise \end{cases}$$

$with\ \alpha_t = \log(1/\beta_t)$

3.6 Proposed PSO-Adaboost Algorithm

The most popular boosting technique used is the AdaBoost. Here, a strong classifier is built by combining weak classifiers. A weak classifier is a simple classifier that cannot precisely classify the training set despite having the best classification function. An optimal weak classifier and the classification error is calculated at every iteration of the AdaBoost algorithm. All the chosen weak classifiers are comprised in the strong final classifier and are weighted. EAs and heuristics are applied inside the AdaBoost framework for decreasing the training time of the AdaBoost. This helps to supplant the immense search [24]. In order to select the feature that can best minimize the classification error, exhaustive search is performed in every round of AdaBoost which is a type of optimization problem. Additionally, though there is a wide use of the average means in literature, the AdaBoost does not state the method in which the decision threshold is calculated. These two issues are tackled in this work by allowing the PSO to design the weak classifier –which means that finding its feature as well as the decision threshold is an optimization process [25].

This example involves the weak classifier containing features (type, x, y, w, h)as well as two centroids C− and C+. These centroids can approximately represent the means of negative and positive paradigms in 1D feature. The first five parameters (type, x, y, w, h) are integer. Their values are restricted by the dimension of the detection sub-window, while the centroid parameters can be real numbers. This is how the entire problem is converted into a constrained optimization problem. For instance, for the second type (two vertical rectangles, the constraint is x + 2w ≤ W and for the first type feature (with two horizontal rectangles, the constraint is y + 2h

≤ H. The label of the examples that are displayed in Fig 1 are decided by the two centroids C+ and C-

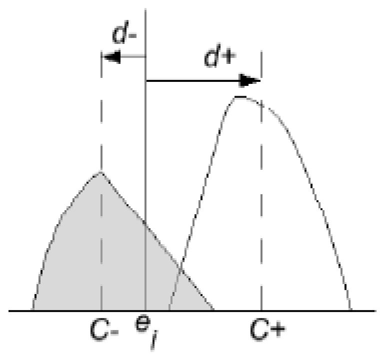

Figure 1C− and C+ are found using PSO, example feature ei is labelled with class C− because d− < d+.

For a given example feature e_i, the distance |C− e_i| to the centroids are calculated.The class having the least distance to its centroid is labelled with the example. It is to be noted that for accurate classification, the comparative proximity between the centroids is more important than the exact positions. The two points which can minimize the distance of the inner class and also maximize the distance across class on the respective feature axis are considered as optimum centroids. Or, it is trying to find two centroids (C− and C+) so that the class separation criteria in (8) can be maximized:

$$J(C_-,C_+)=\frac{|C_- - C_+|}{\frac{1}{m}\sum_{i=1}^{m}|C_- - e_i^-| + \frac{1}{l}\sum_{i=1}^{l}|C_+ - e_i^+|} \qquad (8)$$

Here, the number of negatives is represented by m and the number of positives by l. The criteria is maximized using PSO indirectly such that the error is decreased instead of maximizing the criteria which is distance based. This is akin to the notion of minimizing the class separation which is denoted by Fisher discriminant formula and at the same time minimizing the classification error.Equation (8) however, is subject to a single feature. It is hoped that this process of finding the centroids of examples is more precise than finding the centroids merely by averaging the examples responses [26]. Figure 2 demonstrated particle encoding which has 2 centroids and feature parameters; these can label the instances as positive class or negative class. Its aim is optimization of the 7 parameters as per fitness function. Minimizing the weighted error rate which is akin to the original AdaBoost is the fitness function (9)

$$\begin{aligned} &\in_j = \sum_{i=1} w_{i,t}.b_i \\ &\text{where } b_i = 0 \text{ if } h_t(x_i) = y_i, \\ &\qquad b_i = 1 \text{ otherwise} \end{aligned} \qquad (9)$$

type	x	y	w	h	c-	C+

Figure 2 Particle Encoding

The weak classifier's parameters optimize integers while the basic PSO is used for non-discrete optimization. Additionally, the feature parameters in this work are subject to constraints that demonstrate the sub window directions and those particles that go against the constraint are penalized.

4. RESULTS AND DISCUSSION

This section makes use of the Adaboost, PCA-Adaboost, PSO-Adaboost and PCA-PSO-Adaboost. In the PSO based Adaboost classifier optimization, the number of classification trees (range 25-300), Depth of tree (1 to 6) are optimized. The results have been summarized in table 1. In figures from 3 to 5, the classification accuracy, true positive rate for normal, MCI and AD and true negative rate for normal, MCI and AD are shown

Table 1 Summary of Results

	Adaboost	PCA-Adaboost	PSO-Adaboost	PCA - PSO-Adaboost
Classification Accuracy	0.8298	0.8553	0.883	0.9149
True Positive Rate – Normal	0.8943	0.9143	0.9257	0.9543
True Positive Rate – MCI	0.675	0.7375	0.8	0.825
True Positive Rate – AD	0.575	0.575	0.675	0.75
True Negative Rate – Normal	0.7857	0.8119	0.8426	0.8807
True Negative Rate – MCI	0.9282	0.9397	0.9538	0.9655
True Negative Rate – AD	0.9175	0.9335	0.9487	0.9662

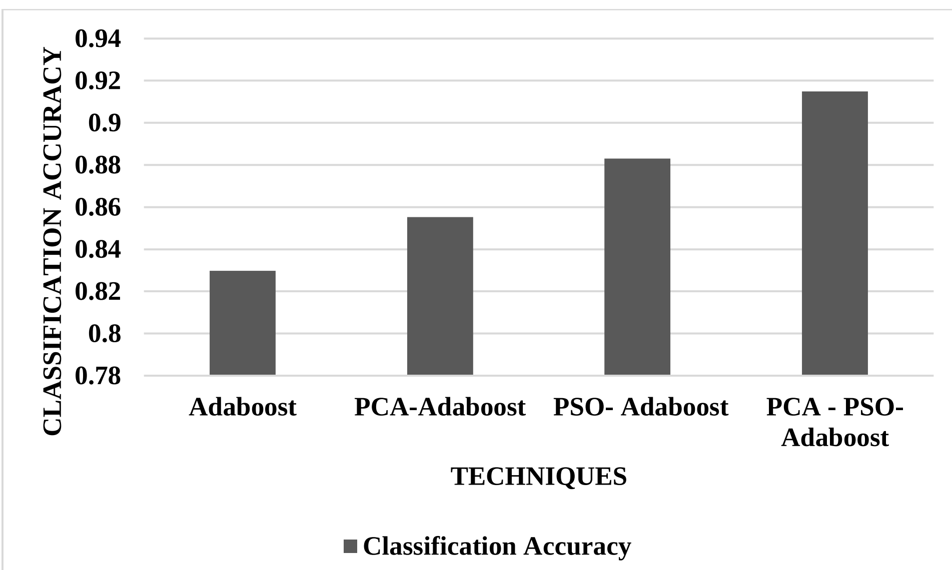

Figure 3 Classification accuracy for PCA-PSO-Adaboost

From the figure 3, it can be observed that the PCA-PSO-Adaboost has improved classification accuracy by 16.47% than Adaboost, by 12.72% than PCA-Adaboost and by 5.21% than PSO-Adaboost.

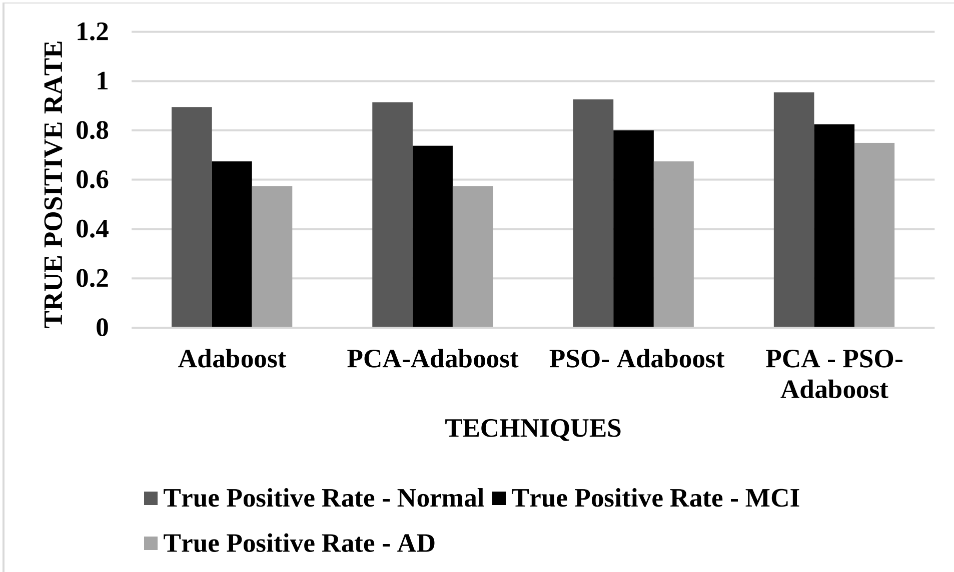

Figure 4 True Positive Rate for PCA-PSO-Adaboost

From the figure 4, it can be observed that the PCA-PSO-Adaboost has higher average true positive rate by 9.75% for Adaboost, by 6.73% for PCA-Adaboost and by 3.54% for PSO-Adaboost.

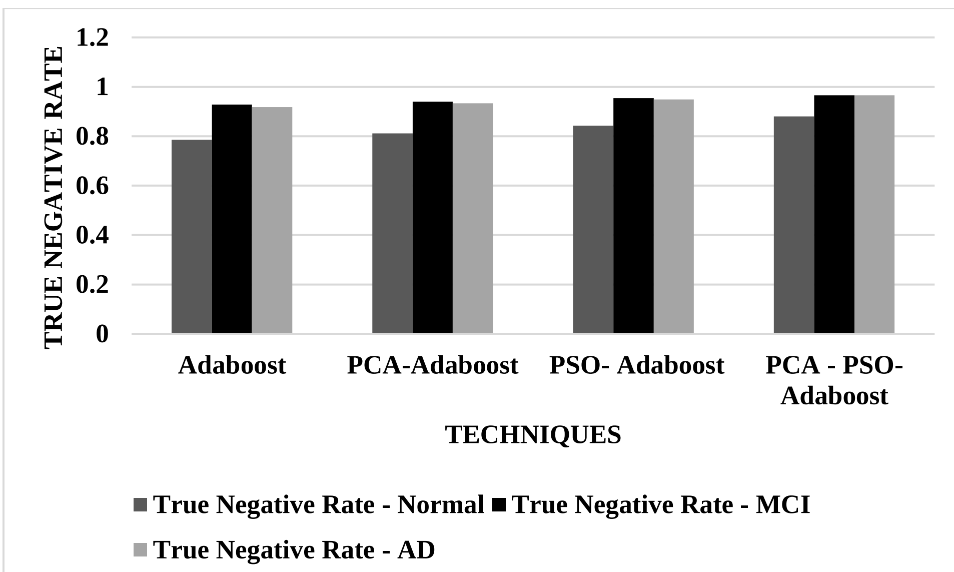

Figure 5 True Negative Rate for PCA-PSO-Adaboost

From the figure 5, it can be observed that the PCA-PSO-Adaboost has higher average true negative rate by 6.64% for Adaboost, by 4.63% for PCA-Adaboost and by 2.42% for PSO-Adaboost.

5. CONCLUSION

By means of the VBM data processing method for MRI, the localized brain volume anomalies can be detected in AD sets when compared to the control sets, in an automatic manner across the complete brain. The misinterpretation of the significant difference relative to the standard VBM is reduced using an optimized VBM technique. This work makes use of an AdaBoost classifier optimized PSO technique so that AD can be detected using the MRIs. An improvised AdaBoost algorithm has been shown in this work wherein by using the PSO algorithm, the time to construct a weak classifier is decreased. The weak classifier is nothing but a decision stump wherein the in-depth search is supplanted with PSO based optimized search. It has been shown by empirical outcome that when PSO is applied to the decision stump, the time that is consumed by the AdaBoost is better than the basic AdaBoost. Thus, in case of a large scale problem, these evolutionary algorithms can decrease the time to find the best solution and also improve the algorithm performance Results show that the PCA-PSO-Adaboost has higher classification accuracy by 16.47% for Adaboost, by 12.72% for PCA-Adaboost and by 5.21% for PSO-Adaboost.

REFERENCES

1. Lama, R. K., Gwak, J., Park, J. S., & Lee, S. W. (2017). Diagnosis of Alzheimer's Disease Based on Structural MRI Images Using a Regularized Extreme Learning Machine and PCA Features. Journal of healthcare engineering, 2017.
2. Jha, D., & Kwon, G. R. (2016). Alzheimer disease detection in MRI using curvelet transform with KNN. The Journal of Korean Institute of Information Technology, 14(8).
3. Xiao, Z., Ding, Y., Lan, T., Zhang, C., Luo, C., & Qin, Z. (2017). Brain MR Image Classification for Alzheimer's Disease Diagnosis Based on Multifeature Fusion. Computational and mathematical methods in medicine, 2017.
4. Busatto, G. F., Diniz, B. S., & Zanetti, M. V. (2008). Voxel-based morphometry in Alzheimer's disease. Expert review of neurotherapeutics, 8(11), 1691-1702.
5. Matsuda, H. (2013). Voxel-based morphometry of brain MRI in normal aging and Alzheimer's disease. Aging and disease, 4(1), 29.
6. Sathya, R., & Abraham, A. (2013). Comparison of supervised and unsupervised learning algorithms for pattern classification. Int J Adv Res Artificial Intell, 2(2), 34-38.
7. Bhavsar, H., & Ganatra, A. (2012). A comparative study of training algorithms for supervised machine learning. International Journal of Soft Computing and Engineering (IJSCE), 2(4), 2231-2307.
8. Shi, B., Chen, Y., Zhang, P., Smith, C. D., Liu, J., & Alzheimer's Disease Neuroimaging Initiative. (2017). Nonlinear feature transformation and deep fusion for Alzheimer's Disease staging analysis. Pattern recognition, 63, 487-498.
9. Sørensen, L., Nielsen, M., & Alzheimer's Disease Neuroimaging Initiative. (2018). Ensemble support vector machine classification of dementia using structural MRI and mini-mental state examination. Journal of neuroscience methods.
10. Çitak-ER, F., Goularas, D., Ormeci, B., & Alzheimer's Disease Neuroimaging Initiative. (2017). A novel Convolutional Neural Network Model Based on Voxel-based Morphometry of Imaging Data in Predicting the Prognosis of Patients with Mild Cognitive Impairment. Journal of Neurological Sciences (Turkish), 34(1).
11. Jha, D., Kim, J. I., & Kwon, G. R. (2017). Diagnosis of Alzheimer's disease using dual-tree complex wavelet transform, PCA, and feed-forward neural network. Journal of healthcare engineering, 2017.
12. Ranjan Nayak, D., Dash, R., & Majhi, B. (2017). 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), 16(2), 137-149.

13. Zhang, Y., Wang, S., Sui, Y., Yang, M., Liu, B., Cheng, H.,... & Gorriz, J. M. (2017). Multivariate Approach for Alzheimer’s Disease Detection Using Stationary Wavelet Entropy and Predator-Prey Particle Swarm Optimization. Journal of Alzheimer's Disease, (Preprint), 1-15.
14. Valarmathy, S., & Vanitha, N. S. (2017). Feature Selection Optimization Using Artificial Immune System Algorithm for Identifying Dementia in MRI Images. Journal of Medical Imaging and Health Informatics, 7(1), 73-78.
15. Ahmed, H. M., Youssef, B. A., Elkorany, A. S., Saleeb, A. A., & El-Samie, F. A. (2018). Hybrid gray wolf optimizer–artificial neural network classification approach for magnetic resonance brain images. Applied Optics, 57(7), B25-B31.
16. Wang, S., Zhang, Y., Dong, Z., Du, S., Ji, G., Yan, J.,... & Phillips, P. (2015). Feed-forward neural network optimized by hybridization of PSO and ABC for abnormal brain detection. International Journal of Imaging Systems and Technology, 25(2), 153-164.
17. Wang, S., Phillips, P., Yang, J., Sun, P., & Zhang, Y. (2016). Magnetic resonance brain classification by a novel binary particle swarm optimization with mutation and time-varying acceleration coefficients. Biomedical Engineering/Biomedizinische Technik, 61(4), 431-441.
18. Alam, S., & Kwon, G. R. (2017). Alzheimer disease classification using KPCA, LDA, and multi-kernel learning SVM. International Journal of Imaging Systems and Technology, 27(2), 133-143.
19. Bains, A. K., & Sandhu, P. (2015). Image denoising using curvelet transform. Int. J. Current Eng. Technol, 5(1), 490-493.
20. Abdullah, N., Chuen, L. W., Ngah, U. K., & Ahmad, K. A. (2011, November). Improvement of MRI brain classification using principal component analysis. In Control System, Computing and Engineering (ICCSCE), 2011 IEEE International Conference on (pp. 557-561). IEEE.
21. Omran, M. G. (2005). Particle swarm optimization methods for pattern recognition and image processing (Doctoral dissertation, University of Pretoria).
22. Mohammadpour, M., Ghorbanian, M., & Mozaffari, S., AdaBoost performance improvement using PSO algorithm. In Information and Knowledge Technology (IKT), 2016 Eighth International Conference on (pp. 273-275). IEEE.
23. Hamdi, N., Auhmani, K., Hassani, M. M. R., & Elkharki, O. (2015). An efficient gentle adaboost-based approach for mammograms classification. Journal of Theoretical and Applied Information Technology, 81(1), 138.
24. Mohemmed, A., Johnston, M., & Zhang, M. (2011). Particle swarm optimisation based AdaBoost for object detection. Soft Computing, 15(9), 1793-1805.
25. Mohemmed, A. W., Zhang, M., & Johnston, M. (2009, May). Particle swarm optimization based adaboost for face detection. In Evolutionary Computation, 2009. CEC'09. IEEE Congress on (pp. 2494-2501). IEEE.
26. R. Santhoshkumar, M. KalaiselviGeetha, J. Arunnehru, ‘SVM-KNN based Emotion Recognition of Human in Video using HOG feature and KLT Tracking Algorithm, International Journal of Pure and Applied Mathematics, vol. 117, No. 15, 2017, pp.621-624, ISSN: 1314-3395.

27. P. Deepan and L.R. Sudha, Effective utilization of YOLOv3 model for aircraft detection in Remotely Sensed Images, Materials Today: Proceedings, Elsevier Publisher, 20121, ISSN 2214-7853, https://doi.org/10.1016/j.matpr.2021.02.831
28. B.Rajalingam, R. Santhoshkumar (2020) "Intelligent Multimodal Medical Image Fusion with Deep Guided Filtering", Multimedia Systems, Springer-Verlag GmbH Germany, part of Springer Nature 2020
29. R. Santhoshkumar, M. KalaiselviGeetha, 'Deep Learning Approach: Emotion Recognition from Human Body Movements', Journal of Mechanics of Continua and Mathematical Sciences (JMCMS), Vol.14, No.3, June 2019, pp.182-195, ISSN: 2454-7190.
30. Dr.B.Rajalingam, Dr.R.Santhoshkumar, Dr. G. Govinda Rajulu, Dr. R. Vasanthselvakumar, Dr. G. JawaherlalNehru, Dr. P. Santosh Kumar Patra, "Survey On Automatic Water Controlling System For Garden Using Internet Of Things (Iot)" The George Washington International Law Review, Vol.- 07 Issue -01 April-June 2021.
31. R. Santhoshkumar, M. KalaiselviGeetha, 'Vision based Human Emotion Recognition using HOG-KLT feature' Advances in Intelligent System and Computing, Lecture Notes in Networks and Systems, Vol.121, pp.261-272, ISSN: 2194-5357, Springer.
32. P.Deepan and L.R. Sudha, "Object Classification of Remote Sensing Image Using Deep Convolutional Neural Network", The Cognitive Approach in Cloud Computing and Internet of Things Technologies for Surveillance Tracking Systems, pp.107-120, 2020. https://doi.org/10.1016/B978-0-12-816385-6.00008-8.
33. R. Santhoshkumar, M. KalaiselviGeetha, 'Human Emotion Prediction Using Body Expressive Feature', Microservices in Big Data Analytics, IETE Springer Series, ISSN 2524-5740, 2019, (Springer), https://doi.org/10.1007/978-981-15-0128-9_13
34. P.Deepan and L.R. Sudha, "Remote Sensing Image Scene Classification using Dilated Convolutional Neural Networks", International Journal of Emerging Trends in Engineering Research, Vol. 8, No.7, pp.3622-3630, 2020, ISSN: 2347-3983.
35. R. Santhoshkumar, M. KalaiselviGeetha, 'Emotion Recognition System for Autism Children Using Non-verbal Communication', International Journal of Innovative Technology and Exploring Engineering (IJITEE), Vol.8, No.8, June 2019, pp.159-165, ISSN: 2278-3075.