Journal of Medical Systems (2019) 43: 76
https://doi.org/10.1007/s10916-018-1147-7

IMAGE & SIGNAL PROCESSING

A Computer Aided Diagnosis System for Identifying Alzheimer's from MRI Scan using Improved Adaboost

S. Saravanakumar[1] · P. Thangaraj[2]

Received: 2 November 2018 / Accepted: 19 December 2018 / Published online: 13 February 2019

Abstract

The recent studies in Morphometric Magnetic Resonance Imaging (MRI) have investigated the abnormalities in the brain volume that have been associated diagnosing of the Alzheimer's Disease (AD) by making use of the Voxel-Based Morphometry (VBM). The system permits the evaluation of the volumes of grey matter in subjects such as the AD or the conditions related to it and are compared in an automated manner with the healthy controls in the entire brain. The article also reviews the findings of the VBM that are related to various stages of the AD and also its prodrome known as the Mild Cognitive Impairment (MCI). For this work, the Ada Boost classifier has been proposed to be a good selector of feature that brings down the classification error's upper bound. A Principal Component Analysis (PCA) had been employed for the dimensionality reduction and for improving efficiency. The PCA is a powerful, as well as a reliable, tool in data analysis. Calculating fitness scores will be an independent process. For this reason, the Genetic Algorithm (GA) along with a greedy search may be computed easily along with some high-performance systems of computing. The primary goal of this work was to identify better collections or permutations of the classifiers that are weak to build stronger ones. The results of the experiment prove that the GAs is one more alternative technique used for boosting the permutation of weak classifiers identified in Ada Boost which can produce some better solutions compared to the classical Ada Boost.

Keywords Magnetic Resonance Imaging (MRI) · Alzheimer's Disease (AD) · Voxel-Based Morphometry (VBM) · Principal Component Analysis (PCA) · Genetic Algorithms (GA) and Greedy Search

Introduction

The Alzheimer's Disease (AD) has not always been related with aging. This is a dementia type that causes issues related to thinking, behaviour, and memory. The signs of the AD normally develop very slowly and keep worsening over time. They can also get very severe and can interfere with day to day life and may, at times even result in death. There has not been any cure identified for the disease. In the year 2006 alone, there have been 26.6 million people all the world over suffering from this. It has been predicted that AD can affect at least 1 in about 85 people by the year 2050 and a minimum of 43% of the existing cases will possibly need a high level of care. With the constantly evolving world, the burdens, as well as the impacts that are caused by the AD on society and also on families has been significantly increasing. In the United States, the healthcare on people suffering from this costs roughly around $100 billion every year and has further been predicted to cost around $1 trillion each year by the end of the year 2050 [1].

An early, as well as accurate identification of the AD, can be extremely beneficial for managing this disease. Currently, there have been several neurologists and researchers giving a considerable amount of time to this goal of early detection and this has been showing some very promising results. The Magnetic Resonance Imaging (MRI) which is a technique in imaging produces some very greater image quality of the anatomical human body structure. This refers mainly to the brain where there is some rich information provided for clinically diagnosing this and also for biomedical research. The MRI and its diagnostic values have been enhanced to a great extent by the accurate classification that is automated using the Magnetic Resonance (MR) images. It has been playing a

This article is part of the Topical Collection on *Image & Signal Processing*

✉ S. Saravanakumar
sar112113118@gmail.com; saravanakumarme85@gmail.com

[1] Research Scholar, Anna University, Chennai, Tamilnadu, India

[2] Department of Computer Science and Engineering, Bannari Amman Institute of Technology, Sathyamangalam, Erode, India

critical role in the detection of the subjects of the AD from the Normal Elder Controls (NC).

A Structural MRI (sMRI) has an important part to play in the distinction of the subjects of the AD from that of the NC and in the distinction of the subjects of Mild Cognitive Impairment (MCI) that afterwards change to the AD and those who may not. In before times, there was a lot of work that was done by hand or even semi-manually measuring the priori Region of Interest (ROI) which was based on the fact where the patients of the AD had suffered cerebral atrophy on being evaluated to the NCs. Majority of these ROI-based analyses tend to focus on the hippocampus regions or the entorhinal cortex. Once the comparison between the ROIs of various subjects is done, the examiners tend to discover some valuable information and also provided valuable information for purposes of diagnosis. But the methods based on the ROI may suffer from certain limitations. Firstly, the focus tends to be on the ROIs that are pre-determined on the basis of their prior knowledge as opposed to being data-driven or being exploratory. Second, the accuracy of an early detection may depend to a great extent on the examiner and his experience. Thirdly, its efficiency is quite low and fourthly, any mutual information exchanged among the voxels may be challenging to operate [2].

Contrastingly, all methods that are multivariate tend to consider the relationship existing among voxels across the entire brain and so does not need any specific ROIs. These methods are generally time-consuming and are expertise-demanding, thus making a manual analysis impossible to achieve. Thus, the methods of automated classification are found desirable. The VBM is a method that was established for an automated and unbiased study of the MRI scans that are structural. A variety of settings of the diseases that include the analysis of a VBM of the loss of brain tissues in the AD are employed. There is an argument that any imperfect MRI scan registration into a common template may result in a false estimation of the atrophy. Further, the errors of tissue classification at the time of a segmentation of the brain tissues can produce a thin rim of the periventricular grey matter that is artificial [3].

The automated method of measuring brain atrophy based on VBM independently maps all the tissue loss in grey matter in a voxel-by-voxel pattern. This method also uses the standardization techniques to for the tissues to obtain pattern similar to that used in well-designed neuroimaging. The main benefit of that of the VBM over investigates is that it was based on an ROI and the VBM produces some unbiased results through exploring the entire brain. This particular approach has shown a very high level of accuracy in terms of AD discrimination and controls contrasted to the ROI-based approach. This VBM work had been undertaken for an evaluation of the discrimination ability of several premature AD patients from age-matched controls with a software program which was newly developed [4].

The main idea behind the automated classifier has described as: The Principle Component Analysis (PCA) has been appealing as it has effectively reduced the data dimensionality and further the cost of computation of the analysis of a new data. These Principal Components (PCs) are extorted from the MRI data and are part of the natural groupings of the circuit of the brain that is involved at the time the AD is progressing. This is a task of supervised learning to automatically classify the AD, the NC, and the MCI. The Ada Boost has been one effective method that improves the performance of the base classifiers empirically and also theoretically.

This work has proposed the PCA, the Ada Boost, the GA and the Greedy Search Algorithm that is based on the VBM approach to distinguish the AD from the MRI. The rest of the work has been structured thus: all related work in the literature has been discussed in Section 2. The techniques employed are discussed in Section 3. Section 4 discusses the empirical outcomes and the work is concluded in Section 5.

Related Works

Gorji and Haddadnia [5] had presented a new method that was on the basis of Pseudo Zernike moments (PZMs) to diagnose the individuals with MCI from the AD and the (HC) groups that make use of a structural MRI. This method made use of the PZMs for extracting some discriminative information from the MR images in the AD, the MCI, and the HC groups. There were two different categories of such Artificial Neural Networks (ANNs), based on recognition of pattern and the Learning Vector Quantization (LVQ) based networks, for classifying information which was extracted from MRIs. Mahanand et al., [6] further presented a new approach for the detection of the AD from the MRI by using a Meta-cognitive Radial Basis Function Network (McRBFN) classifier. The work also proposed the Recursive Feature Elimination (RFE) based approach using an efficient method of classification called Projection based Learning-McRBFN (referred to as the PBL-McRBFN-RFE) for identifying all meaningful biomarkers along with a power of prediction for the detection of the AD in male members.

Chyzhyk et al., [7] had further proposed another evolutionary method of a wrapper feature selection by means of employing an Extreme Learning Machines (ELM) as its base classifier that consisted of the GA discovering the feature combination space. The GA fitness function is a mean accuracy of evaluation made with cross-validation for an individual feature selection. There is also a marginal accuracy distribution of the classification which corresponds to features employed to compute feature saliency. All the unprocessed features extorted as voxel selection from an anatomical brain MRI were used.

Beheshti et al., [8] had also developed yet another novel Computer-Aided Diagnosis (CAD) system that employed feature ranking along with the GA to analyse the MRI data. By

using this system, the conversion of the MCI-to-AD that was among one to three years prior to any clinical analysis could be predicted.

Zhang et al., [9] had also developed one more novel system in machine learning which can automatically diagnose using the brain MRI. Firstly, the brain imaging will get processed and this includes the skull stripping along with the spatial normalization. Secondly, an axial slice is chosen from among a volumetric image, and a Stationary Wavelet Entropy (SWE) is completed for extracting all the texture features. Thirdly, there is a neural network with a single-hidden-layer that is employed in place of a classifier. At last, the predator-prey Particle Swarm Optimization (PSO) had been presented for training all the biases and also the weights of this classifier.

Methodology

For this work, an Alzheimer's disease Neuroimaging Initiative (ADNI) database [10] had been utilized and the data that was used for the article preparation was attained from the ADNI database (which was adni.loni.usc.edu). This ADNI had been initiated in the year 2003 as the public-private partnership, which is directed by Michael W. Weiner, MD a public investigator. The main objective of that of the ADNI is to investigate the approaches that can combine various markers like serial MRI or PET or any other biological markers with the clinical or the neuropsychological assessment for assessing and predicting the condition of MCI or the early AD. The PCA also has some more methods of feature reduction. The PCA's primary purpose [11] was the reduction of a huge size of its data space (the observed variables) to that of a small feature space dimensionality (the independent variables). As a consequence, the PCA method was seen as a mathematical procedure to extract some relevant information from a dataset that was large.

A feature extraction by making use of the curvelet transform [12] is the multi-scale and the multi-directional transform that was presented by Candes and Donoho for overcoming limitations of a conventional wavelet transform. The curvelet transforms further provide an optimal representation of the objects having curve singularities. This only means that the curvelet transform will require small numbers of coefficients that represent either a line or a curve in a particular medical image. The Ada Boost classifier and the optimization that employs the GA and the methods of greedy search have been discussed here.

AdaBoost Algorithm

The Ada Boost had been proposed by Freund and Schapire in the year 1995 as a very efficient algorithm in the field of ensemble learning. It was used for the purpose of boosting the performance of a classification of that of the weak learners. This was made by means of combining a set of functions of weak classifiers which was in order to form a strong classifier. The Ada Boost normally combines all weak classifiers taking into consideration the distribution of weight of training samples to ensure more of the weight gets featured to the samples misclassified under the earlier iterations. The ultimate strong classifier obtains the perception form which is nothing but a weighted combination of the weak classifiers trailed by a threshold [13].

The pseudo-code for the Ada Boost algorithm further gets an input set (x_1, y_1), ..., (x_m, y_m) wherein every x_i is part of a domain or an instance X and every label y_i a label set Y. As soon as the T rounds of the Ada Boost training is complete, the T numbers of all of the weak classifiers h_t and their ensemble weights α_t are capitulated by learning for constituting their final and strong classifiers.

Given : *a training set* $(x_1, y_1), ..., (x_m, y_m)$
where $x_i \in X, y_i \in Y = \{-1, +1\}$
Initialize weights
$w_{1,i} = 1/m$
For $t = 1, ..., T$:
1. *Normalize the weights* :
$w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^{n} w_{t,j}}$
so that w_t *is a probability distribution*.
2. *For each feature* j, *train a classifier* h_j
which is restricted to using *a* single *feature*.
The error is evaluated with respect to w_t :
$\varepsilon_j = \sum_i w_i |h_j(x_i) - y_i|$
3. *Choose the weak classifier* h_t,
with the lowest error ε_t.
4. *Update the weights* :
$w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$
where $e_i = 0$ *if example* x_i *is classified*
correctly, $e_i = 1$ *otherwise*, *and* $\beta_t = \varepsilon_t / (1 - \varepsilon_t)$.
Output the final strong classifier :
$$H(x) = sign\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right)$$
where $\alpha_t = \log(1/\beta_t)$.

The Adaboost does not perform well in case the data set consists of outliers and noisy data hence it is not suitable to application where its data is corrupted in noisy environment. Similarly overfitting problem occurs when the number of rounds T (boosting steps) is high. Which means as the number of iterations increases the correctly classified samples gets lost in the mist of the misclassified samples hence performance is dropped.

Genetic Algorithm (GA) Based AdaBoost

Here in this work, the GAs get employed as techniques of search for exploring all possible weak classifier

permutations which can outperform the Ada Boost permutation. The GA is the computation model in machine learning that obtains its manners from a allegory of the evolution. For producing a solution that is found to be better than the Ada Boost, the candidate solutions are symbolized as individuals and their act duly calculated based on their training set samples [14].

Since these candidate solutions generally overcome through several generations, only triumphant solutions tend to stay alive. The fittest individual that is found the final population leads to a better solution and the GAs are normally the general techniques of search compared to a greedy search restricting a search space and being compared to an exhaustive search. Both the mutation, as well as the crossover, will make the GA progress arbitrarily within the search space. But the fitness function tends to concentrate on solutions that are found to be better in the current population. The ability can further restrict the genetic algorithms waste time of some useless solutions. For the same reason, an evolutionary search has been deemed to be the middle layer existing between that of the greedy search and an thorough search. Both the representation and the fitness functions are two basic conditions that form a GA.

The Representation

Each single theory chosen by the Ada Boost in every round t_i denotes the element of hypotheses sets which is H. This means the $h_i = h_1, h_2, ..., h_M$, wherein the M denotes the number of hypotheses and the $t_i = t_1, t_2, ..., t_T$ wherein the T denotes the actual number of these rounds. These hypotheses can also be a complex learning similar to the neural networks, the Support Vector Machines (SVM) or certain simple learners like the thresholds on the 1D histograms, the nearest neighbourhood classifier or the k-means classifier. In this entire work, the weak hypotheses get to be the adaptive thresholds on the 1D histograms produced by a feature that is predefined.

The Fixed Length Encoding: this includes a string of the T integer values wherein each value or the gene of the individual can represent the hypothesis h_i. T denotes the actual number of the rounds of boosting performed. Every element of the individual iFs is a weak learner from the set H selected by the boosting process. The operations of both mutation, as well as the crossover, will change both the order and also the value of the gene permutations which can produce various individuals. This type of a representation is duly closed under the genetic operators such as mutation or crossover and the mutation can change a single gene value which is t_i to t_j wherein the i, j = 1, 2, · · ·, T (as depicted in Fig. 1).

A Variable Length Encoding: the actual length of an encoding based on its initial value obtained from the integer string. Firstly, there is an individual gene that can be any number which characterizes individual genes and the other individuals remaining are inside a similar structure like the individuals in a coding of a fixed length. The individual element (first element in the string decides the parameter T) initially decides on the number of boosting steps and based on this, the GA identifies optimum classifier numbers that are boosted (as depicted in Fig. 2).

Fitness Function

Let, Ii denote the individual from a population at generation Gj wherein j = 1, 2, ...,G denotes the actual number of the generations that are evolved by this population. The fitness function's general structure will be as below:

Require : *training images* $TR = (x_1, y_1), (x_2, y_2), ..., (x_N, y_N)$ *where* $y_i = 0, 1$ *for negative and positive examples respectively. An individual* I_i *from population at certain generation where* $I_{i,j}$ *is the element (gene) of the individual.*
Ensure : a value $f(I_i)$ *that represents fitness(strength) of individual* I_i.
Initialize weights $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$ *for* $y_i = 0, 1$ *respectively, where m and l are the number of negatives and positives respectively.*
Initialize if representation is variable length *encoding,* $T = I_{i,0}$, *otherwise* $T = LENGTH(I_i)$
Repeat[1−3] *for* $t = 1, 2, ..., T$:
1 : *Normalize weights,*
$w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^{n} w_{t,i}}$
so that w_t *is a probability distribution.*
2 : *Train classifier* h_j *is restricted to* using single *feature* $I_{i,t}$*.The error is evaluated with respect to* $w_t, \varepsilon_j = \sum_i w_i |h_j(x_i) - y_i|$.
3 : *Update weights* :
$w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$
where $e_i = 0$ *if example* x_i *is classified correctly,* $e_i = 1$ *otherwise, and* $\beta_t = \frac{\varepsilon_t}{1-\varepsilon_t}$.
Form final strong classifier

$$h(x) = \begin{cases} 1 \text{ if } \sum_{t=1}^{T} \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha_t \\ 0 \qquad \text{otherwise} \end{cases}$$

Evaluate a function with respect to final strong classifier.
$f(I_i) = FFUNCTION(I_i, h(x), TR) = DR - FA$
where

$$DR = \frac{number\ of\ false\ positives}{number\ of\ negative\ samples}.$$

$$FA = \frac{number\ of\ true\ positives}{number\ of\ positive\ samples},$$

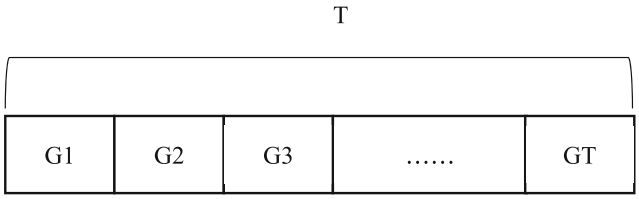

Fig. 1 Length of the encoding is T which is the number of rounds of boosting steps

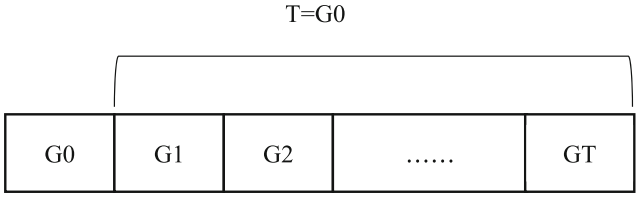

Fig. 2 Length of this encoding is T + 1 that is the number of rounds of boosting steps

Initialization

The Random Initialization: there have been several GAs that begin with a arbitrary population. The random initialization is that system used extensively in various experiments.

A Partial Initialization along with the Ada Boost: One more approach that had been proposed was the initialization of all the individuals that were found in the initial population with a certain proportion of that of a final classifier produced by the Ada Boost once definite iterations were complete. This concept commences the GA by employing an initial population which was more qualified enhancing an Ada Boost solution.

Initialization of the Population by an Unselected Hypotheses: for the actual purpose of identifying certain precise solutions, the initial population has to be large adequate to be able to envelop all hypotheses within the genetic pool. For the Ada Boost, in every stage, all hypothesis need to be sorted and further recorded based on their error rate. Any hypothesis having the lowest error rate is generally chosen and injected into a genetic pool to be a step in initialization. For this phase, the first individual is initiated with a hypothesis with the lowest number of errors and the next individual is initialized using a hypothesis that has the second lowest errors. This particular scheme of initialization generally improves the performance of the Ada Boost.

Greedy Search Based AdaBoost

A greedy algorithm will be the one which constructs the object X step by step. In certain cases, this can construct the best globally by means of repeatedly selecting the option that is locally the best.

The Greedy algorithms generally have many advantages over that of the other approaches which are: Simplicity: the greedy algorithms are normally easier to be described and to code up. Efficiency: they may be implemented in a more efficient manner compared to the other algorithms.

The Greedy algorithms also have many drawbacks: Difficult to design: as soon as the right greedy approach is identified, the designing of greedy algorithms become easy. But, identifying the right approach can be very challenging. Difficult to verify: the showing this greedy algorithm to be correct, will generally need a nuanced argument. This is an effective algorithm that can create a classifier that is strong from a weak one. The boosting work by means of frequently running a weak classifier on several teaching examples is modelled from new training pools. The Ada Boost has been proved theoretically and is shown empirically as an effective method which improves the accuracy of classification. By using a certain procedure of weight, the Ada Boost can focus on the data points that are misclassified in the earlier iterations and thus minimizes any training errors [15].

As the Ada Boost has been a greedy algorithm which and focuses intentionally on the error minimization there are several other investigations on issues of overfitting for the Ada Boost. The conclusion from such early investigation is that in practice, the Ada Boost hardly ever over fits training data. The investigations made recently have implied that the phenomenon may be connected to the fact that this process of greedy search that is used in the Ada Boost can implicitly increase the margin of classification. But the other investigations have

Table 1 Summary of Results

	PCA Adaboost	PCA-GA Adaboost	PCA- Greedy AdaBoost
Classification accuracy	0.8553	0.9234	0.8723
True Positive Rate - Normal	0.9143	0.96	0.92
True Positive Rate - MCI	0.7375	0.85	0.8
True Positive Rate - AD	0.575	0.75	0.6
True Negative Rate - Normal	0.8119	0.875	0.8302
True Negative Rate - MCI	0.9397	0.9708	0.9377
True Negative Rate - AD	0.9335	0.9735	0.9531

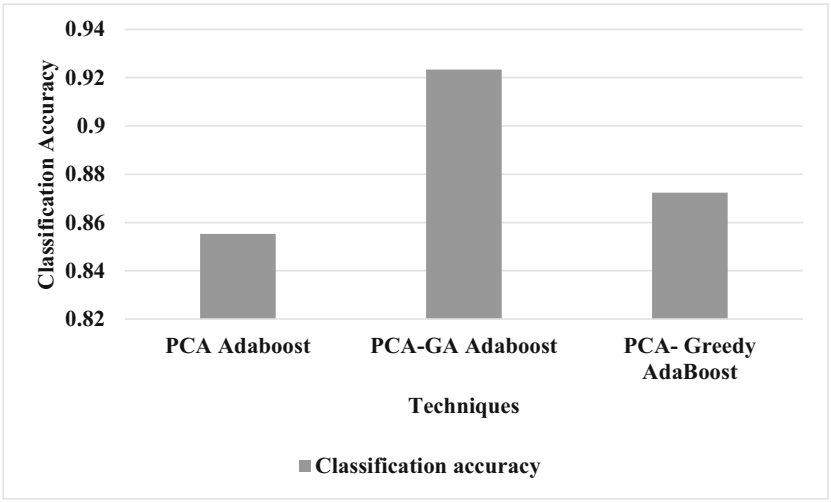

Fig. 3 Classification Accuracy for identifying Alzheimer'susing PCA-GA Adaboost

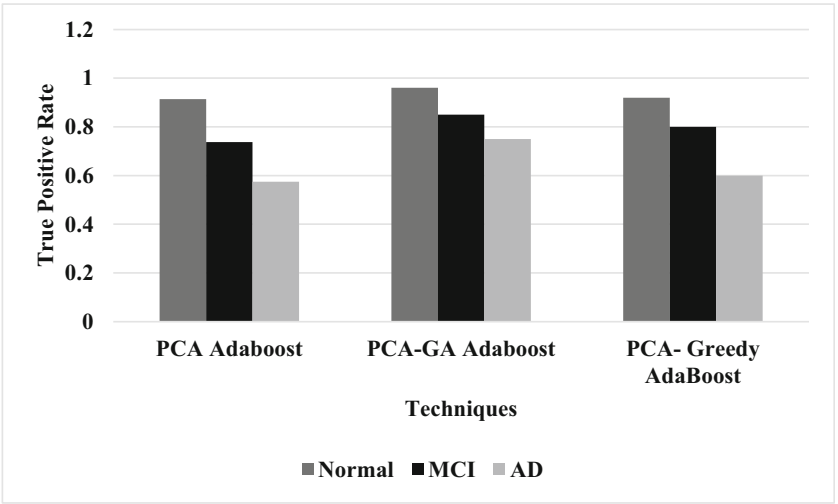

Fig. 4 True Positive Rate for identifying Alzheimer's PCA-GA Adaboost

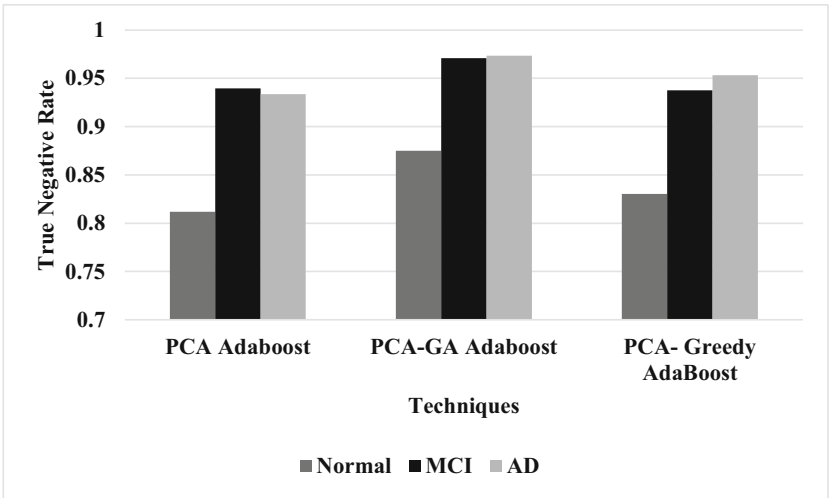

Fig. 5 True Negative Rate for identifying Alzheimer's PCA-GA Adaboost

proved that the Ada Boost may also have problems in overfitting especially when the data get noisy.

For the selection of a learning algorithm that is weak, in each step of an algorithm, the Ada Boost will select the classifier having the lowest error will divide the search space in relation to the learner. The Ada Boost further proposed another greedy search in the space of the hypotheses by means of choosing the one that has as its heuristic the lowest error. For this point, the is a question in which the selection is the ONLY best classifier where every step can lead to the algorithm to identify an optimal solution.

Results and Discussion

In this section, the PCA-Adaboost, PCA-GA-Adaboost and PCA-Greedy-Adaboost methods are used. The GA and Greedy based Adaboost classifier optimization (The number of classification trees (range 25–300), depth of tree (1 to 6) are optimized. The summary of results as shown in Table 1. The classification accuracy, true positive rate for normal, MCI and AD and true negative rate for normal, MCI and AD as shown in Figs. 3, 4, 5.

From the Fig. 3, it can be observed that the PCA-GA Adaboost has higher classification accuracy for identifying Alzheimer's by 7.65% for PCA Adaboost and by 5.69% for PCA- Greedy AdaBoost.

From the Fig. 4, it can be observed that the PCA-GA Adaboost has higher average true positive ratefor identifying Alzheimer's by 13.92% for PCA Adaboost and by 9.83% for PCA- Greedy AdaBoost.

From the Fig. 5, it can be observed that the PCA-GA Adaboost has higher average true negative rate for identifying Alzheimer's by 4.87% for PCA Adaboost and by 3.54% for PCA- Greedy AdaBoost.

Conclusion

The VBM is another very useful approach that investigates the neuro-structural changes in the brain. For this work, there is a better method of search to boost the usage of an evolutionary search that was proposed. This system and the GAs, will be able to consecutively enhance the hypotheses coded as the genes on various chromosomes (every chromosome will represent the final classifier that is boosted). The function of performance evaluation for calculating the fitness values in every boosting step was introduced for deciding the individuals and their strength in the space of the hypotheses. For this work, a boosting technique which targets the solving of partial problems in the Ada Boost has been proposed. This finds a very elegant way to boost one bunch of classifiers in order to form one such "better classifier" compared to the ensemble classifiers. This algorithm has employed a greedy search over the space of the hypotheses for identifying a "good" and suboptimal solution. The results have proved that this PCA-GA Ada boost had a higher accuracy of classification by about 7.65% for the PCA Ada boost and further by about 5.69% for the PCA- Greedy Ada Boost.

References

1. Zhang, Y., Dong, Z., Phillips, P., Wang, S., Ji, G., Yang, J., and Yuan, T. F., Detection of subjects and brain regions related to Alzheimer's disease using 3D MRI scans based on eigenbrain and machine learning. Frontiers in Computational Neuroscience 9:66, 2015.
2. Zhang, Y., Wang, S., and Dong, Z., Classification of Alzheimer disease based on structural magnetic resonance imaging by kernel support vector machine decision tree. Progress In Electromagnetics Research 144:171–184, 2014.
3. Karas, G. B., Burton, E. J., Rombouts, S. A. R. B., Van Schijndel, R. A., O'Brien, J., Scheltens, P. H. et al., A comprehensive study of gray matter loss in patients with Alzheimer's disease using optimized voxel-based morphometry. Neuroimage 18(4):895–907, 2003.
4. Hirata, Y., Matsuda, H., Nemoto, K., Ohnishi, T., Hirao, K., Yamashita, F. et al., Voxel-based morphometry to discriminate early Alzheimer's disease from controls. Neuroscience letters 382(3): 269–274, 2005.
5. Gorji, H. T., and Haddadnia, J., A novel method for early diagnosis of Alzheimer's disease based on pseudo Zernike moment from structural MRI. Neuroscience 305:361–371, 2015.
6. Mahanand, B. S., Babu, G. S., Suresh, S., & Sundararajan, N. (2015). Identification of imaging biomarkers responsible for Alzheimer's disease using a McRBFN classifier. In Cognitive Computing and Information Processing (CCIP), 2015 International Conference on (pp. 1–4). IEEE.
7. Chyzhyk, D., Savio, A., and Graña, M., Evolutionary ELM wrapper feature selection for Alzheimer's disease CAD on anatomical brain MRI. Neurocomputing 128:73–80, 2014.
8. Beheshti, I., Demirel, H., Matsuda, H., and Initiative, A.'s. D. N., Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm. Computers in biology and medicine 83:109–119, 2017.
9. 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.
10. Popuri, K., Balachandar, R., Alpert, K., Lu, D., Bhalla, M., Mackenzie, I. R. et al., Development and validation of a novel dementia of Alzheimer's type (DAT) score based on metabolism FDG-PET imaging. NeuroImage: Clinical 18:802–813, 2018.
11. Nandi, D., Ashour, A. S., Samanta, S., Chakraborty, S., Salem, M. A., and Dey, N., Principal component analysis in medical image processing: a study. International Journal of Image Mining 1(1):65–86, 2015.

12. Mikki, M. (2013). Brain Tumor Detection Based On Curvelet and Artificial Neural Network (Doctoral dissertation, Islamic University–Gaza).
13. Dezhen, Z., & Kai, Y. (2008). Genetic algorithm based optimization for adaboost. In 2008 International Conference on Computer Science and Software Engineering (pp. 1044–1047). IEEE.
14. Yalabik, I., & Fatos, T. V. (2007). A pattern classification approach for boosting with genetic algorithms. In Computer and information sciences, 2007. iscis 2007. 22nd international symposium on (pp. 1–6). IEEE.
15. Jin, R., Liu, Y., Si, L., Carbonell, J. G., & Hauptmann, A. (2003). A new boosting algorithm using input-dependent regularizer.