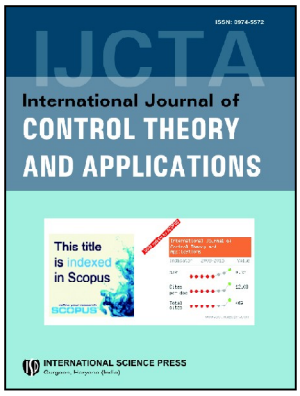

International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 10 • Number 26 • 2017

Enhancing Alzheimer Classification Using Genetic Selection

[*1]N. Shunmuga Karpagam and [2]S. Srinivasaraghavan

*[*1] Research Scholar, Department of Computer Science and Engineering, VEL TECH, Avadi, Chennai, Tamil Nadu, India*
E-mail: shunmugakarpagam.n@gmail.com
[2] Professor, VEL TECH, Avadi

Abstract: The usual kind of dementia is Alzheimer's disease (AD). This disease is wide spread among elderly people across the world. The prodromal stage and detecting the disease at an early stage is necessary to slow down the disease. Many new methods and development are cropping up for diagnosing the disease in its initial stages. Until recent diagnosis of AD and MCI, the multiple biomarkers were found responsive to the process i.e., structural MRI for brain atrophy measurement, functional imaging for hypo metabolism quantification, and cerebrospinal fluid (CSF) for quantification of specific proteins. Several tests were conducted to evaluate the features selection techniques such as Information Gain (IG), Minimum Redundancy Maximum Relevance (mRMR), Genetic Algorithm (GA) as well as Classification using Logit boost and Cascading Classifier.

Keywords: Alzheimer's disease (AD), Mild Cognitive Impairment (MCI), Information Gain (IG), Minimum Redundancy-Maximum Relevance (MRMR), Genetic Algorithm (GA).

1. INTRODUCTION

The condition AD occurs when the brain tissues undergo degeneration or in other words AD is a neurodegenerative disorder of brain tissue. When AD advances, it results in entire psychological activities loss. The recent epidemiological data shows that AD has affected around 26 million people around the world particularly about 100.000 in Algeria. The decrease in volume of gray matter (GM) with age and increase in cerebrospinal fluid (CSF) was detected through studies such as Anatomical magnetic resonance imaging (MRI). The necessary attention is given to patients affected by Alzheimer by quantifying the extent of atrophy in the cerebral cortex during the initial period of AD. The medial temporal lobe (MTA) is optically measured in clinical practice.

Conversely, visual analyses can be deceptive and might not provide true quantification of atrophy. Therefore, to automate the process several methods were developed in which the volume of hippocampal, volume of gray matter structures or thickness of cortical regions is measured. Whole-brain techniques to illustrate brain atrophy are considered to be suitable to differentiate AD from other neurodegenerative dementias [1]. Yet AD is a common neurodegenerative dementia and is rapidly growing among aged people. Definite diagnosis is possible only on post-mortem, and it necessitates histopathological confirmation of amyloid plaques as well as neurofibrillary tangles.

The task of diagnosing AD accurately in its early stages is very demanding as it can have serious effects during future treatments. The criteria for clinical diagnosis include clinical as well as neuropsychological measurement. Initially the presence of dementia is identified followed by the phenotype of Alzheimer. If the patients under clinical assessment are identified with AD at a prodromal stage are generally categorized as amnestic mild cognitive impairment (MCI). It is not a general rule that patients affected by amnestic MCI would acquire AD. The current research works are proposing many new criteria for early detection of AD at the prodromal phase of the syndrome. These criteria are derived from the clinical core of early episodic memory impairment as well as the existence of minimum one supplementary supportive aspect of abnormal MRI along with PET neuroimaging or abnormal cerebrospinal fluid amyloid along with tau biomarkers.

A constructive predictive assessment can be added to the diagnosis by utilizing neuro imaging techniques which also including utilizing structural MRI measurements to calculate medial temporal lobe atrophy as well as positron emission tomography employing fluoro deoxy glucose (FDG) or amyloid markers [2]. Early cholinesterase inhibitors are advantageous to AD patients and also these patients benefit with premature and precise diagnosis of AD. In last few years, there has been extensive investigation on the clinical diagnosis of AD and have resulted in amnestic MCI. In addition to neuropsychological examination, the diagnosis of AD is also supported by structural imaging. The progression of the AD disease is clear through pathological studies. According to it, AD initially occurs in the medial temporal lobe and in succession affects the entorhinal cortex, the hippocampus; the limbic system and it finally extend towards neocortical areas.

In the beginning a stage of AD, atrophy besides affecting the hippocampus or the entorhinal cortex it also affects other regions in AD as well as MCI patients. AD and MCI patients can be differentiated using methods such as whole-brain as it characterizes brain atrophy. This is useful in determining the possible evolution from healthy subjects towards developing AD. Further these methods must be able to provide individual predictive diagnosis to be used by the clinicians. According to recent classification methods, these methods are built in order to permit individual class prediction [3]. The developed countries are having increasing numbers of aged people with dementia thereby posing a major health problem.

AD is becoming a common occurrence in aged people and hence it is vital to control its progression by detecting the disease in its early stages. In this type of disease the brain cells are affected and subsequently lead to their degeneration. Hence, with a steady progress and improvement in medical imaging techniques, the variation in human brain structure along with its association with clinical diagnosis of AD can be affectively analysed. If the human brain is structurally abnormal then it can be detected using medical information from MRI. The MRI measurements are contemplated as an indicator of AD process and so the evolution of brain atrophy can be detected and tracked using structural MRI measurements. Patients affected with AD, undergo a change in their structure where the volume of the hippocampus is reduced.

Besides this, the commonly used methods for AD diagnosis are protracted. These methods are based on fine image segmentation and so require intervention by clinicians and also suffer from region of interest (ROI) segmentation's errors. Conversely, only with the help of hippocampus it is not possible to differentiate the subject with MCI and AD [4]. MRI, a neuro imaging method helps in visualizing brain structure with higher spatial resolution as well as the contrast among brain tissue types. Using structural MRI techniques, the volumetric transforms in brain regions related to AD as well MCI can be identified along with the illustration regarding the efficacy of such approaches in analysing such syndromes.

Especially, using the structural MRI it is possible to identify, AD as well as MCI related cross-sectional dissimilarities and longitudinal modifications in volume as well as size of definite brain regions, like the hippocampus, entorhinal cortex, with regional modifications in gray matter, white matter, even CSF on a voxel-by-voxel basis. According to recent developments, the MRI data is commonly used in machine learning experiments with the aim of categorizing patients as AD vs cognitively normal (CN) or MCI vs CN. The modern approach is to directly use network analysis or usage of machine learning on the voxels. Basically, it is difficult

to classify progressive neurodegenerative disease as the subjects span different stages of the disease from MCI to AD [5].

The specific volume of brain regions, particularly the hippocampus and entorhinal cortex display histopathogical alters in the initial period of AD. So such regions are considered for most of the MRI studies in MCI as well as AD. Computational neuro anatomy is employed in examining voxel-by-voxel brain variations in healthy aging, MCI as well as AD. According to studies, there are definite atrophy patterns that involve medial temporal lobe in MCI and AD. The significance of MRI is emphasized as an efficient surrogate marker of disease at the group-analysis level, i.e. the individuals with or without the pathology is clearly differentiated. Conversely, they have restricted diagnostic value, particularly in the initial brain pathology stage due to insufficient sensitivity and specificity for predicting the status of a given individual [6].

The remaining sections organized as: Section 2 reviews the related work in literature. Section 3 illustrates the proposed techniques. Section 4 discusses the experiment outputs and section 5 concludes the proposed work.

2. LITERATURE REVIEWS

Gao and Hui[7] investigated the possibility in applying the modern profound learning convolution neural network (CNN) to classify CT brain images. The main purpose is to subject these images into clinical applications. Finally, there categories are grouped together, which contain subjects' data with AD or lesion or normal ageing. This research uses both 2D and 3D CN owing to the features of CT brain images with larger thickness with depth (z) direction (~5mm).

Liu *et al.,* [8] suggested an inherent structure-based multi-view leaning technique employing various templates to classify AD/MCI. The subclasses are encoded using distinct codes where the original class as well as own distribution information is considered, trailed different by means of a feature selection model. The authors study the ensemble view of definite Support Vector Machine (SVM) classifiers on the features chosen in every view and integrate the outputs to obtain an absolute conclusion. The outputs from various tests on the Alzheimer's disease Neuroimaging Initiative database illustrate the promising performance of the method in classifying AD/MCI against other innovative multi-template-based methods.

Garali *et al.,* [9] introduced an innovative computer-aided diagnosis method for brain Positron Emission Tomography (PET) image classification in AD case. Initially, an atlas is used in segmenting brain images into ROI. By processing these regions with a few statistical parameters, the authors were able to categorize a Separation Power Factor (SPF) related to each region. SPF calculates the capacity of every region to isolate AD from Healthy Control (HC) brain images. To achieve better classification accuracy, the selected regions must be ranked according to their SPF and should be given as input to a SVM classifier rather than using the same number of ranked regions as inputs which are retrieved from four different conventional feature selection methods.

Garali *et al.,* [10] suggested a new approach to rank the effectiveness of brain regions and to isolate AD from healthy brains images. Initially, the brain images are mapped into 116 anatomical regions of interest. Then the Receiver Operating Characteristics curves are used in ranking the ability of regions to separate PET brain images. Twenty one selected regions are given as input to both SVM and Random Forest classifiers and then 142 brain PET images are evaluated. The results obtained from classification are better when compared to those obtained when using the whole 116 initial regions or when inputting the whole brain voxels. Further, the computational time was reduced.

Zhan *et al.,* [11] concentrated on anatomical brain networks processed using diffusion magnetic resonance images. In addition, suggested a modern feature extraction and classification frame derived from singular value decomposition of higher order along with thin logistic regression. From the experiments conducted on freely available dataset from the Alzheimer's disease Neuro imaging Initiative it has been proven that the suggested structure is capable of detecting distinct brain structure that helps to classify early versus late MCI.

Garali *et al.,* [12] built an innovative technique to rank the efficacy of brain region of interest for classifying healthy as well as AD brain. Based on anatomical atlas, the brain images are fragmented into 116 regions. A receiver Operating Characteristic curve is used in all regions, to sort the capability of the region to distinguish healthy from AD brain images. A set of chosen regions, based on their position was fed as input to a SVM. The classification outputs proved to be the same or better when compared to methods such as entire voxels and the 116 regions to the classifier as input features.

Rodrigues andSilveira[13] gave a comprehensive approach to the proposed study of overall brain pattern. In this method voxel-wise differences is used. Further it does not segment the images into regions of interest. In order to carry out the study, FDG-PET scans at the baseline,a 12-month followed-upon cognitively normal (CN), MCI along with AD subjects were extracted from the Alzheimer's disease neuro imaging initiative (ADNI) database. In order to identify AD and MCI, a blend of cross-sectional and longitudinal information can be used as it gives best classification results instead of utilizing cross-sectional data alone. Further, the longitudinal voxel-based analysis performed better than multi-region analysis.

Segovia and Phillips[14] introduced an innovative methodology for PET data analysis by which the automated differentiation among controls and AD patients can be improved. The images are categorized into small regions or parcels, described anatomically, geometrically or arbitrarily. The SVM classifier along with a cross-validation approach is used to approximate the precision of each single region. The purpose of the classifier is to increase the weight of the most discriminative regions during the final decision. In order to evaluate this method, a PET dataset containing images through healthy controls as well as AD patients was used. The classification output obtained using the suggested methodology was better than Principal Component and Independent Component Analysis that were based on computer systems.

Guan and Yang [15] suggested a technique derived from independent component analysis as well as support vector machine. It combines the images of sMRI and PET to classify the AD patients from HC. The simulation outputs demonstrated that the classification among AD and HC subjects from the Alzheimer's disease Neuroimaging Initiative database was acquired with an average precision of 96.53% for Multimodal images at baseline, when compared to 88.95% for only sMRI images as well as 89.44% for PET images alone. Using multimodal classification method the brain anomaly can be detected with greater accuracy and sensitivity during its initial stages.

Cabral and Silveira[16] suggested a distinct method of ensemble to focus on the problem of three class to classify AD, MCI and to Control Normal (CN) through PET brain images. The distinct image features are related to distinct brain voxels set. So the suggested preferred class classifiers will consider the information that the spatial brain degeneration patterns in AD alters as the disease advances with time. The authors used SVM and Random Forests (RF) as base classifiers to test this method on FDG-PET images from The Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The results from this SVM ensemble showed better performance than the equivalent single classifier with improved results.

3. METHODOLOGY

This section discusses the Feature selection using IG, MRMR, GA and Classification using logitboost and cascading classifier.

(A) Dataset

This article uses the data prepared by the Alzheimer's disease Neuro imaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). The ADNI was introduced in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies as well as non-profit organizations, as a $60 million, 5-year public-private

partnership. The main objective of ADNI is to examine serial MRI, PET, other biological indicators along with clinical and neuropsychological measurement can be pooled to assess the development of MCI as well as early AD. The intention of determining sensitive and specific markers for prime stage progression of AD is to help researchers and clinicians to innovate modern treatments together with regular monitoring. It is also aimed at reducing the time and cost involved.

The required eligibility criteria for ADNI are described at www.adni-info.org. In short, the age limit for the subjects ranges between 55-90 years, have a study partner capable of providing an independent assessment of functioning. The specific psychoactive medications are barred. The following are the general inclusion/exclusion criteria:

- **Healthy subjects:** Mini-Mental State Examination (MMSE) achieves amid 24-30, a Clinical Dementia Rating (CDR) of 0, non-depressed, non MCI, and non-demented;
- **MCI subjects:** MMSE scores between 24-30, a memory complaint, have objective memory loss measured through education adjusted scores on Wechsler Memory Scale Logical Memory II, a CDR of 0.5,significant level of impairment in some other cognitive domains is absent, the actions performed in day today life is conserved and absence of dementia; and
- **Mild AD:** MMSE results between 20-26, CDR of 0.5 or 1.0, meets the National Institute of Neurological and Communicative Disorders, Stroke and the Alzheimer's disease as well as Related Disorders Association (NINCDS/ADRDA) criteria for possible AD [17].

(B) Feature Extraction

The tissue is segmented and normalized utilising SPM12. (http://www.fil.ion.ucl.ac.uk/spm). To implement SPM12, image normalization, segmentation and bias correction are included. The creation of customized tissue probability maps (TPMs) reduces any prospective normalization as well as segmentation bias across the disease groups. The personalized TPMs are developed by normalizing all 380 images to the MNI template as well as by segmenting them into grey matter (GM), white matter (WM) as well as CSF employing the unified segmentation model in SPM5. Average probability maps of GM, WM and CSF can be developed using from all 380 GM, WM and CSF probability maps. Let be the vector of remaining GM voxels for the k-th subject (k=1,2,3....m). Likewise, the vectors for WM, , and CSF, , are developed. For each patient, the total quantity of retained tissue densities values was of the order [18].

(C) Feature Selection

In machine learning, feature selection is a global optimization issue. By employing feature selection, the quantity of features is reduced, irrelevant, noise filled and redundant features are discarded and above all it increases recognition accuracy. Despite being used for selecting relevant features, feature selection is also used to reduce computational overheads.

Information Gain (IG)

The main aim of IG criterion is to discover the quantity of unique data added by one feature to the entire features set. A feature's IGf can be calculated by $F(S \cup f) - F(S)$, wherein F (.) refers to the evaluator criteria while S represents the selected features subset. Features with more IG are preferred. Being a metric based technique, IG can be used to select the suitable split features in decision tree classifiers and it also denotes the extent to which entropy of data is decreased. Additionally it can also detect the values for every individual feature.

The entire feature basis receive IF values, depending on which the features are selected or discarded. Therefore, it is essential to initially set up threshold values for features selection as the features are selected only when IG values are greater than threshold ones. Let's assume a set A of comprising s samples and a set B of comprising k classes. If $P(B_i, A)$ is the fraction of samples in A which have class Bi, then, the expected data for class membership is given through:

$$Info(A) = -\sum_{i=1}^{k} P(B_i, A) * \log(P(B_i, A))$$

If the value of IG is greater, then there is a chance of getting pure classes in target class if splits are based on the parameter with the greatest gain.

Minimum Redundancy-Maximum Relevance (MRMR)

The mRMR denotes mutual information based technique and it selects features as per maximal statistical dependency criteria. Ther are challenges involved in directly implementing maximal dependency condition, so mRMR is approximated for maximization of dependency between joint distributions of chosen attributes and the classification parameter. Minimization of redundancy for discrete attributes and continuous attributes is given by:

For Discrete attributes: minWI, $$W_I = \frac{1}{|S|^2} \sum_{i,j \in S} I(i, j)$$

For Continuous attributes: minWc, $$W_C = \frac{1}{|S|^2} \sum_{i,j \in S} |C(i, j)|$$

Here $I(i, j)$ and $C(i, j)$ refer to mutual information as well as the correlation between f_i and f_j, correspondingly. Maximization of relevance for discrete as well as continuous attributes is given by:

For Discrete attributes: max VI, $$V_I = \frac{1}{|S|^2} \sum_{i \in S} I(h, i)$$

For Continuous attributes: max Vc, $$V_C = \frac{1}{|S|^2} \sum_{i} F(i, h)$$

Wherein h refers to the target class while $F(i, h)$ represents the F-statistic [19].

Genetic Algorithm (GA)

GA is a well-known optimization algorithm, which derives its basis from the Darwinian principle of natural selection. GA is stochastic global adaptive search method derived from natural selection mechanisms. This algorithm is compatible with other data mining techniques for optimization and performance amelioration. Initially in genetic algorithm, there is a random generation of initial population which includes several hundreds or thousands of potential solutions to the problem. The population is encoded and every individual in the population

is known as chromosome and every bit in the chromosome is called gene and has a value. The next step involved is genetic operators [20]. The generally used genetic operators are selection, crossover, and mutation. The fitness function is defined with the user to evaluate each chromosome in the population. GA unlike other optimization techniques begins with random initialization of initial population that consists of numerous chromosomes. Every chromosome denotes a solution to the problem and its performance is examined using a fitness function. GAs produce new individuals in cycles called generations; number of iterations are required to generate the population through reproduction.

$$fitness = W_A \times Accuracy + W_{nb} \times \frac{1}{N}$$

Here W_A and W_{nb} represents the weight of accuracy and the weight of N feature participated in classification where $N \neq 0$ respectively.

A population of competing feature transformation matrices is maintained by GA. The individual matrix in the population is evaluated by multiplying the input patterns by the matrix. The produced set of changed patterns is given as input to the classifier. The patterns are classified into training and testing set by the classifier. The former set trains the classifier whereas the latter evaluates the classification accuracy. The acquired accuracy is given back to GA as a degree of quality of the transformation matrix employed to acquire the transformed patterns sets. GA uses this data to search for a transformation to reduce the transformed patterns dimensions and increases the correctness of classification.

Typically, GA is made of three important phases: Selection, Crossover as well as Mutation. Every phase selects being from the existing population to represent them as parents and later uses them to generate offspring for the successive generation. The parents subjected to genetic operators yield children. If the generated younger ones are found better than their parents then they are included into the population. Chromosomes are candidate solutions that are typically denoted as strings of set length. A fitness or objective function reflects the goodness of every individual in population and measures the fitness of a chromosome. Chromosomes with low fitness values are discarded whereas the chromosomes with high fitness values are retained and taken to the next generation. The three fundamental operations are frequented for a number of generations until individuals representing the optimum solution to the problem are produced [21].

(D) Classification using Logitboost

Logistic regression is a statistical method employed in probabilistic classification. Simple logistic is a classifier that can build logistic regression linear models. LogitBoost with simple regression functions is used as learning base to adapt logistic models. The required number of iterations of LogitBoost can be calculated through cross-validation .This automatically selects the attributes. This technique is successfully used in predicting Alzheimer with the help of biomarkers identification [22].

Another well-known supervised learning algorithm is boosting which has become very successful in machine learning. The purpose boosting is to create'strong' classifiers by combining many "weak" classifiers. LogitBoost combines the aspect of Adaptive Boosting (AdaBoost) and Logistic regression. If AdaBoost is considered as a generalized additive model and if the cost functionalities of Logistic Regression are applied, then it arrives at the LogitBoost algorithm. An adaptive newton steps are used to fit an additive symmetric logistic model using utmost likelihood. Using this algorithm the expectation of the loss function to fit an additive logistic regression model to directly optimize a binomial log-likelihood is minimised and denoted by $-\log\left(1+e^{-2yF(x)}\right)$. This feature of Logit Boost alters linearly with the output error and is generally less responsive to noisy data [23].

(E) Cascading Classifier

There is a list of stages in cascade classifier and each stage consist a list of weak learners. The objects in the question are detected by moving a window over the image. The current location of the window defines specific regions. These regions are labelled by each stage of the classifier either as positive or negative – positive means the presence of an object and negative indicates that no specified object was found in the image. If an object is labelled as negative, it implies the classification of this particular region is finished as well as the location of the window is moved to the subsequent position. Conversely, for a positive result by labelling, the region is moved to the subsequent stage of classification. The final positive result is given by the classifier when all the stages (inclusive of the last stage) yield a result stating that the object is found in the image [24].

Many classification strategies were discussed particularly for the face detection field and it has shown improved performance. From the available techniques, the AdaBoost boosting algorithm gives a simple and at the same time an effective stage-wise learning method for the cascade design. A linear programming framework investigates the construction of a cascade of sparse linear classifiers $w'x+b$. Every stage in the cascade resolves a linear program which is expressed via hinge loss

$$\xi = \max\left\{0, 1 - y\left(w'x + b\right)\right\}$$

And the l_1 norm penalty or weighted -norm penalty

$$\|w\|_1^{\gamma} = \sum \gamma_i |w_i|$$

Here γ_i represents the weighing factor associated with the computational cost of i-th feature assumed the available information as cost and if not it becomes the regular l_1-norm by means setting all $\gamma = 1$. Though the linear programs at every phase may be resolved by using any general-purpose linear program solver, it is elaborated in the subsequent section that the column generation method for linear programs can also be applied to optimize each linear program in an incremental fashion similar to AdaBoost. Furthermore, the column generation boosting derivation is applicable to any linear program in spite of the trade-off choice factor between the detection rate as well as the false positive rate while AdaBoost needs modifications for an asymmetric re-weighing system [25].

4. RESULTS AND DISCUSSION

Experiment results conduct with IG, MRMR and GA using Logitboost and IG, MRMR and GA using Cascading Classifier.

Table 1
Classification Accuracy

Techniques	*Classification Accuracy*
IG- Logitboost	82.5
MRMR - LogitBoost	84.5
GA-Logitboost	92.5
IG-Cascading Classifier	86
MRMR-Cascading Classifier	88
GA-Cascading Classifier	95

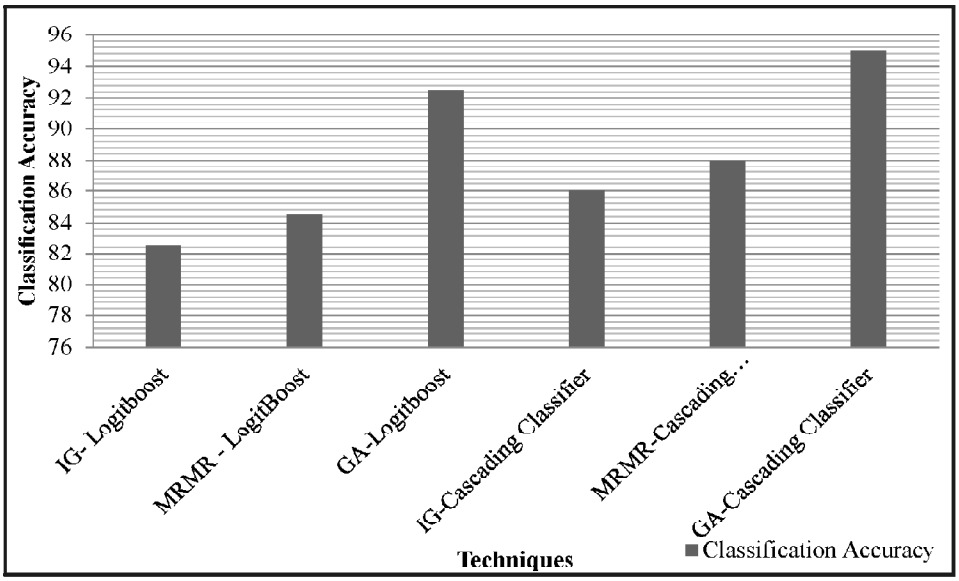

Figure 1: Classification Accuracy

From table 1 and figure 1, it can be observed that the classification accuracy has improved by 2.39% compared to IG- Logitboost and MRMR - LogitBoost.

Table 2
Average Precision

Techniques	*Average Precision*
IG- Logitboost	0.8
MRMR - LogitBoost	0.82
GA-Logitboost	0.9
IG-Cascading Classifier	0.81
MRMR-Cascading Classifier	0.85
GA-Cascading Classifier	0.93

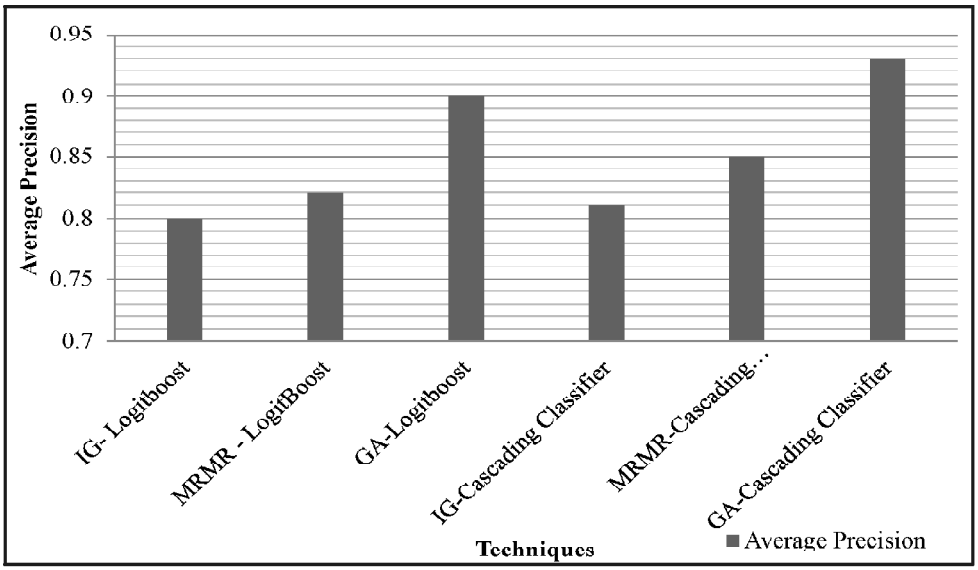

Figure 2: Average Precision

From table 2 and figure 2, it can be observed that the average precision has improved by 9.3% compared to MRMR - LogitBoost and GA-Logitboost.

Table 3
Average Recall

Techniques	*Average Recall*
IG- Logitboost	0.87145
MRMR - LogitBoost	0.89215
GA-Logitboost	0.9543
IG-Cascading Classifier	0.91625
MRMR-Cascading Classifier	0.92355
GA-Cascading Classifier	0.9715

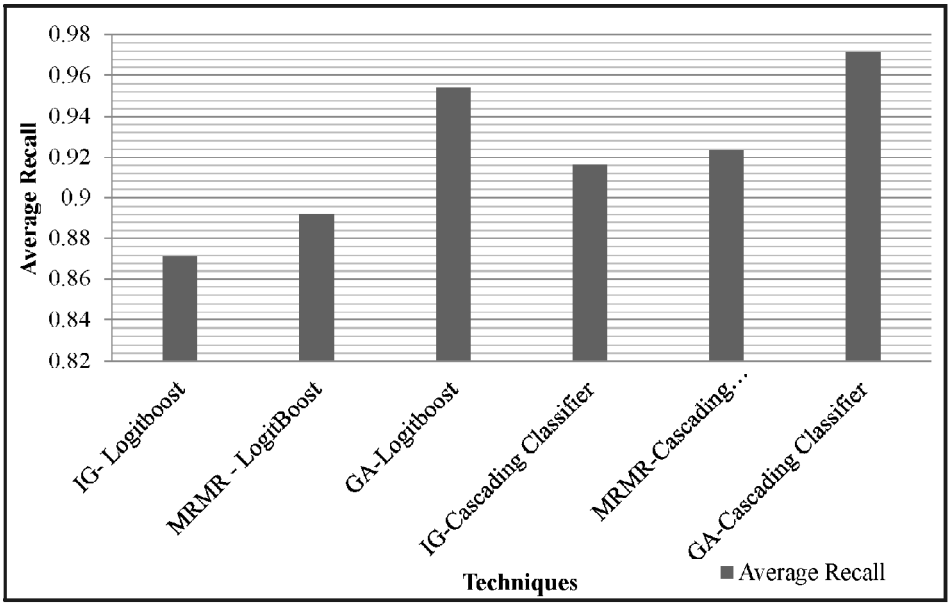

Figure 3: Average Recall

From table 3 and figure 3, it can be observed that the average recall has improved by 0.79% compared to IG-Cascading Classifier and MRMR-Cascading Classifier.

Table 4
Average Sensitivity

Techniques	*Average Sensitivity*
IG- Logitboost	0.9074
MRMR - LogitBoost	0.9143
GA-Logitboost	0.95835
IG-Cascading Classifier	0.92505
MRMR-Cascading Classifier	0.932
GA-Cascading Classifier	0.97415

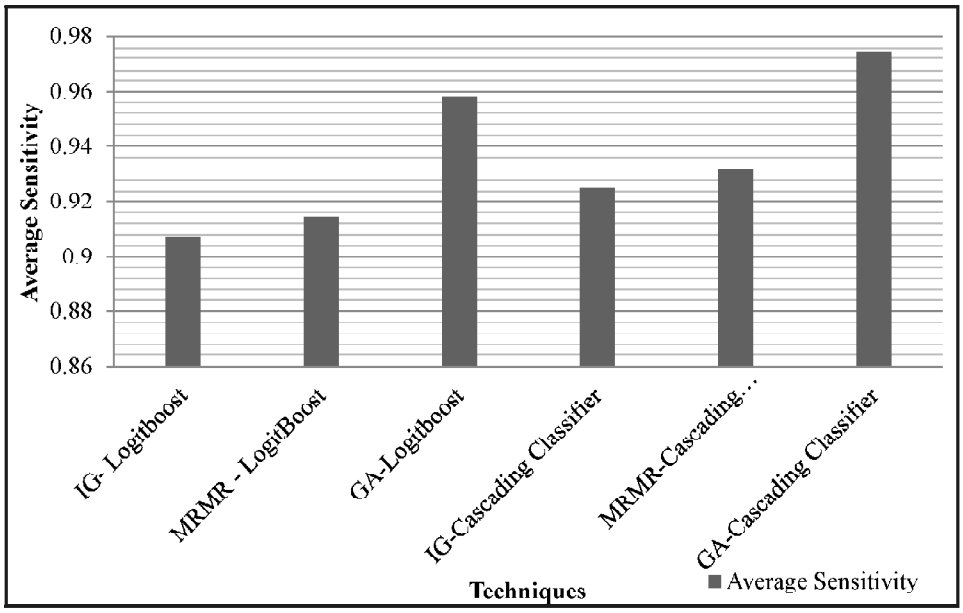

Figure 4: Average Sensitivity

From table 4 and figure 4, it can be observed that the average sensitivity has improved by 4.42% compared to MRMR-Cascading Classifier and GA-Cascading Classifier.

5. CONCLUSION

To delineate soft tissues MRIs can be used as they are the widely used medical imaging technique. This method can effectively identify Alzheimer's disease in its initial stages. The approach is examined in MCI patients and for other neuro degenerative diseases. The strength of the technique can be assessed in patients with images obtained from different MR scanners with several acquisition parameters. The data also indicates that it is reliable for both MCI and AD when compared to other existing cognitive screening scales. In this work feature selection is done using IG, MRMR, GA. The experiment is carried out for logit boost, cascading classifiers with IG, MRMR and GA. The result indicates that the classification accuracy has improved by 2.39% compared to IG-Logitboost and MRMR - LogitBoost.

REFERENCES

[1] Belmokhtar, N., &Benamrane, N. (2012). Classification of Alzheimer's disease from 3D structural MRI data. Age, 78(8), 69-96.

[2] Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehéricy, S., Habert, M. O., ... & Alzheimer's Disease Neuroimaging Initiative. (2011). Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. neuroimage, 56(2), 766-781.

[3] Magnin, B., Mesrob, L., Kinkingnéhun, S., Pélégrini-Issac, M., Colliot, O., Sarazin, M., ... &Benali, H. (2009). Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI.Neuroradiology, 51(2), 73-83.

[4] Ahmed, O. B., Benois-Pineau, J., Allard, M., Amar, C. B., Catheline, G., & Alzheimer's Disease Neuroimaging Initiative. (2015). Classification of Alzheimer's disease subjects from MRI using hippocampal visual features.Multimedia Tools and Applications, 74(4), 1249-1266.

[5] Natarajan, S., Joshi, S., Saha, B. N., Edwards, A., Khot, T., Moody, E., ... & Maldjian, J. A. (2012, December). A machine learning pipeline for three-way classification of alzheimer patients from structural magnetic resonance images of the brain. In Machine Learning and Applications (ICMLA), 2012 11th International Conference on (Vol. 1, pp. 203-208). IEEE.

[6] Davatzikos, C., Fan, Y., Wu, X., Shen, D., & Resnick, S. M. (2008). Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging. Neurobiology of aging, 29(4), 514-523.

[7] Gao, X. W., & Hui, R. (2016). A deep learning based approach to classification of CT brain images.

[8] Liu, M., Zhang, D., Adeli-Mosabbeb, E., & Shen, D. (2015). Inherent Structure Based Multi-view Learning with Multi-template Feature Representation for Alzheimer's Disease Diagnosis.

[9] Garali, I., Adel, M., Bourennane, S., &Guedj, E. (2015, August). A novel feature selection in the case of brain PET image classification. In Signal Processing Conference (EUSIPCO), 2015 23rd European (pp. 649-653). IEEE.

[10] Garali, I., Adel, M., Bourennane, S., &Guedj, E. (2015, September). Region-based brain selection and classification on pet images for Alzheimer's disease computer aided diagnosis. In Image Processing (ICIP), 2015 IEEE International Conference on (pp. 1473-1477). IEEE.

[11] Zhan, L., Liu, Y., Zhou, J., Ye, J., & Thompson, P. M. (2015, April). Boosting classification accuracy of diffusion MRI derived brain networks for the subtypes of mild cognitive impairment using higher order singular value decomposition. In 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI) (pp. 131-135). IEEE.

[12] Garali, I., Adel, M., Bourennane, S., Ceccaldi, M., &Guedj, E. (2016). Brain region of interest selection for 18FDG positrons emission tomography computer-aided image classification. IRBM, 37(1), 23-30.

[13] Rodrigues, F., &Silveira, M. (2014, August). Longitudinal FDG-PET features for the classification of Alzheimer's disease. In 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 1941-1944). IEEE.

[14] Segovia, F., & Phillips, C. (2014, June). PET imaging analysis using a parcelation approach and multiple kernel classification. In Pattern Recognition in Neuroimaging, 2014 International Workshop on (pp. 1-4). IEEE.

[15] Guan, J., & Yang, W. (2013, July). Independent component analysis-based multimodal classification of Alzheimer's disease versus healthy controls. In2013 Ninth International Conference on Natural Computation (ICNC) (pp. 75-79). IEEE.

[16] Cabral, C., &Silveira, M. (2013, July). Classification of Alzheimer's disease from FDG-PET images using favourite class ensembles. In 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 2477-2480). IEEE.

[17] Zhang, D., Wang, Y., Zhou, L., Yuan, H., Shen, D., & Alzheimer's Disease Neuroimaging Initiative. (2011). Multimodal classification of Alzheimer's disease and mild cognitive impairment. Neuroimage, 55(3), 856-867.

[18] Vemuri, P., Gunter, J. L., Senjem, M. L., Whitwell, J. L., Kantarci, K., Knopman, D. S., ... & Jack, C. R. (2008). Alzheimer's disease diagnosis in individual subjects using structural MR images: validation studies.Neuroimage, 39(3), 1186-1197.

[19] Chellammal. and Venkatachalam. (2016). Using Optimized Feature Selection For Classification of Brain MRI. IIOABJ, 7(9), 517-525.

[20] Yin, Z., Zhao, Y., Lu, X., &Duan, H. (2015). A Hybrid Intelligent Diagnosis Approach for Quick Screening of Alzheimer's Disease Based on Multiple Neuropsychological Rating Scales. Computational and mathematical methods in medicine, 2015.

[21] Kharrat, A., Messaoud, M. B., Abid, M., Gasmi, K., &Benamrane, N. (2010, July). Automated classification of magnetic resonance brain images using wavelet genetic algorithm and support vector machine. In Cognitive Informatics (ICCI), 2010 9th IEEE International Conference on (pp. 369-374). IEEE.

[22] Fernandez, I., Gonzalez, D., Gomez, A., Priore, P., Puente, J., &Parreno, J. (2011). Comparative analysis of artificial intelligence techniques for goods classification. In Proceedings of the 2011 International Conference on Artificial Intelligence, ICAI (pp. 1-7).

[23] Srinivasan, R. (2011). Learning Relationships Between Detected (Doctoral dissertation, Washington State University).

[24] Soo, S. (2014). Object detection using Haar-cascade Classifier.

[25] Bi, J., Periaswamy, S., Okada, K., Kubota, T., Fung, G., Salganicoff, M., & Rao, R. B. (2006, August). Computer aided detection via asymmetric cascade of sparse hyperplane classifiers. In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 837-844). ACM.