ORIGINAL ARTICLE

Efficient Algorithm for Distinction Mild Cognitive Impairment from Alzheimer's Disease Based on Specific View FCM White Matter Segmentation and Ensemble Learning

Soheil Ahmadzadeh Irandoost * , **Faeze Sadat Mirafzali Saryazdi**

Department of Medical Physics and Biomedical Engineering, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran

*Corresponding Author: Soheil Ahmadzadeh Irandoost
Email: soheil.ahmadzade@yahoo.com

Received: 10 August 2022 / Accepted: 20 August 2022

Abstract

Purpose: Alzheimer's Disease (AD) is in the dementia group and is one of the most prevalent neurodegenerative disorders. Between existing characteristics, White Matter (WM) is a known marker for AD tracking, and WM segmentation in MRI based on clustering can be used to decrease the volume of data. Many algorithms have been developed to predict AD, but most concentrate on the distinction of AD from Cognitive Normal (CN). In this study, we provided a new, simple, and efficient methodology for classifying patients into AD and MCI patients and evaluated the effect of the view dimension of Fuzzy C Means (FCM) in prediction with ensemble classifiers.

Materials and Methods: We proposed our methodology in three steps; first, segmentation of WM from T1 MRI with FCM according to two specific viewpoints (3D and 2D). In the second, two groups of features are extracted: approximate coefficients of Discrete Wavelet Transform (DWT) and statistical (mean, variance, skewness) features. In the final step, an ensemble classifier that is constructed with three classifiers, K-Nearest Neighbor (KNN), Decision Tree (DT), and Linear Discriminant Analysis (LDA), was used.

Results: The proposed method has been evaluated by using 1280 slices (samples) from 64 patients with MCI (32) and AD (32) of the ADNI dataset. The best performance is for the 3D viewpoint, and the accuracy, precision, and f1-score achieved from the methodology are 94.22%, 94.45%, and 94.21%, respectively, by using a ten-fold Cross-Validation (CV) strategy.

Conclusion: The experimental evaluation shows that WM segmentation increases the performance of the ensemble classifier, and moreover the 3D view FCM is better than the 2D view. According to the results, the proposed methodology has comparable performance for the detection of MCI from AD. The low computational cost algorithm and the three classifiers for generalization can be used in practical application by physicians in pre-clinical.

Keywords: Alzheimer's Disease; Fuzzy C Means; Ensemble.

DOI: https://doi.org/10.18502/fbt.v10i3.13158

Frontiers in BIOMEDICAL TECHNOLOGIES

1. Introduction

Alzheimer's Disease (AD) is one of the neurodegenerative disorders that is characterized by progressive cognitive deterioration [1]. AD is a type of dementia that leads to problems with memory, thinking, and behavior [2]. AD disorder is the most common cause of neurodegenerative dementia and affects those over 65 years old. It is expected that the progression of AD will double within 20 years and that one out of every 85 people will be afflicted with AD by 2050 and reach 131 million people [3-6]. AD disease was named after Dr. Alois Alzheimer in 1906 [7]. According to recent studies, four stages have been introduced for AD, such as pre-dementia, early, moderate, and advanced [8]. There are several methods for the diagnosis of AD that include mental status, physical exams, and neurological exams like MRI, fMRI, PET, and CT [2, 9, 10]. The brain MRI shows the structures, shrinkage, and any other structural variations that might cause cognitive dysfunction [8]. White Matter (WM) was identified as a critical feature of Alzheimer's Disease (AD) by neuroimaging evidence [11]. WM atrophy in mild AD has been observed before [12], and another study shows that the pattern of WM volume decrement aids in finding the underlying pathologic mechanism in AD [13]; other researchers reported that AD patients had a greater annual decrease in temporal WM volume than controls [14, 15]. Recently, machine learning has played an important role in improving medical disorder diagnosis [16]. One of the critical applications of that is the classification of neurodegenerative diseases with biomedical data such as MRI and EEG [16]. Early detection of these diseases is always important and effective [16]. Several machine learning algorithms, including the K-nearest neighbor (KNN) algorithm, the Artificial Neural Network (ANN), the Support Vector Machine (SVM), the Nave Bayesian (NB), the Ensemble, and regression models have been used to classify AD, MCI, and Cognitive Normal (CN) [16-18]. Sometimes, the classification by one individual classifier does not provide the best result, so there were attempts to use an ensemble of classifiers to achieve better classification results [19]. The strategy in an ensemble classifier is to create a set of classifiers and merge their decisions [20]. Although using various machine learning techniques is beneficial, it is noteworthy that user interaction with them can be extremely beneficial in solving difficult problems [21, 22].

Contributions: We describe three key contributions to our work. (i) The first contribution compares two distinct FCM clustering perspectives (2D and 3D) to determine the optimum view based on performance. (ii) The second contribution is introducing the ensemble classifier with three efficient machines (KNN, DT, and LDA) to increase the capability used in the practical application of tracking progressive AD with minimum input data cost and low computational cost. (iii) The third contribution is to evaluate the effect of WM segmentation in MCI detection with an ensemble classifier.

Novelty: In this study, three aspects make our methodology distinct. The first is (i) The FCM is an algorithm used to segment WM from MRI as before, but the effect of viewpoint, such as slice-based (2D) and volume-based (3D) in MCI detection with ensemble machine learning, is reported in this study that is different from others; (ii) previous research employed a variety of machine learning classifiers to achieve the best results; the majority of them used SVM machines, however, the purposed ensemble classifier differs from them, despite having comparable performance; (iii) according to the literature, most of best approaches for detecting MCI from AD were examined with an unequal number of AD and MCI samples. In this study, we tested our methodology using a better way (equal AD and MCI samples) to reduce bias in performance evaluation.

In this study, we proposed our methodology in three steps; first, segmentation of WM from T1 MRI with FCM according to two specific viewpoints (3D and 2D). In the second step, two groups of features are extracted, approximate coefficients of Discrete Wavelet Transform (DWT) with three-level decomposition, and statistical (mean, variance, skewness) features. In the final step, an ensemble classifier that is constructed with three simple classifiers, K-Nearest Neighbor (KNN), Decision Tree (DT), and Linear Discriminant Analysis (LDA) was used to distinguish MCI from AD. The proposed method has been tested by using T1-weighted MR imaging of AD and MCI, which have been chosen from the ADNI dataset. The work of this paper is summarized as follows: in Section II, the overall methods of the introduced methodology are explained. In the next section, the evolution of the proposed algorithm's performance is presented with a variety of assessment measurements such as sensitivity, specificity, accuracy, precision, and f1-score. Section IV contains a discussion based on the

results, and Section V contains the conclusion and future work of the current study.

2. Materials and Methods

The block diagram of the purpose methodology framework is shown in Figure 1. We used a multi-view unsupervised fuzzy clustering method to segment WM from T1 MRI. Then, we applied DWT to FCM output to extract some features and used PCA to reduce the dimension of each subject data set from 32×32×20 to 7×20. In the next step, three statistical features are extracted, so in total, the ensemble classifier gets 10 inputs to detect MCI from AD patients. Finally, we evaluate our method with a well-known ADNI dataset, according to popular measurements in machine learning.

2.1. Dataset

The proposed method has been evaluated by using 1,280 slices (samples) from 64 patients with MCI (32) and AD (32) that were randomly selected from the dataset to prevent bias in our study. All the brain MRIs are from the T1-weighted 1.5T sagittal of the ADNI dataset [23]. A sample dataset that has been used in this study is shown in Figure 2.

2.2. WM Segmentation

To segment WM from the extracted T1 images, FCM segmentation is used. The fuzzy clustering algorithms establish the expression of the uncertainty and can explain the ambiguity in the brain MRI [24]. FCM has produced satisfactory performance in segmenting WM clusters compared to others, and it has a simple form and low computational cost that it uses more easily than other methods in practical and real-time implementation [24]. Assume the image pixel is X = {x1, x2,.., xN} where xi shows the intensity of the image pixel. FCM transforms the image segmentation process into optimization to separate similar pixel data from each other [24]. The clustering problem is dividing N pixels into C classes. The expression FCM is [24] (Equations 1, 2):

$$J_{FCM} = \sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}^{m} d^2(x_j, z_i) \quad (1)$$

$$d^2(x_j, z_i) = ||x_j - z_i||^2 \quad (2)$$

The m parameter determines the ambiguity. Equations 3 and 4 give the formula of cluster center z_i and degree of membership u_{ij} [24].

$$z_i = \frac{\sum_{j=1}^{N} u_{ij}^{m} x_j}{\sum_{j=1}^{N} u_{ij}^{m}} \quad (3)$$

$$u_{ij} = \frac{1}{\sum_{r=1}^{c}[(\frac{d(x_j, z_i)}{d(x_j, z_r)})^{\frac{2}{m-1}}]} \quad (4)$$

The WM segmentation procedure is as follows : (1) input the twenty T1 slices in each subject; (2) set the FCM hyper-parameters (fuzzy factor m = 2, number of clusters C = 5) and randomly initialize the cluster centers and membership degrees; (3) Using Equations 3 and 4, update the clustering centers z and membership degrees u; (4) make the objective function converge until error is satisfied; (5) extract the WM cluster from other groups [24]. With those steps, we segment WM based on two different viewpoints, two dimensions (2D), and three dimensions (3D). In the 2D viewpoint, FCM segments each slice of MRI into five different clusters, but in the 3D viewpoint, FCM segments each volume of MRI in a specific subject. Table 1 shows the hyper-parameters and variables that use WM segmentation with FCM.

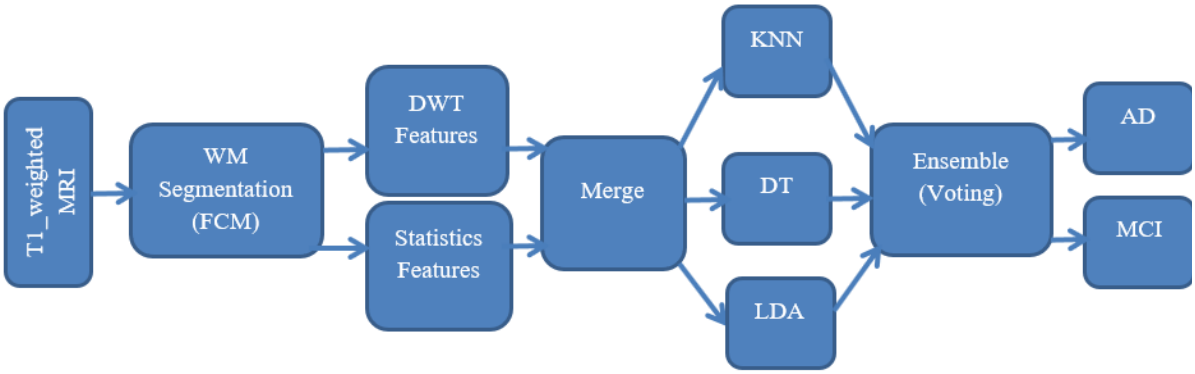
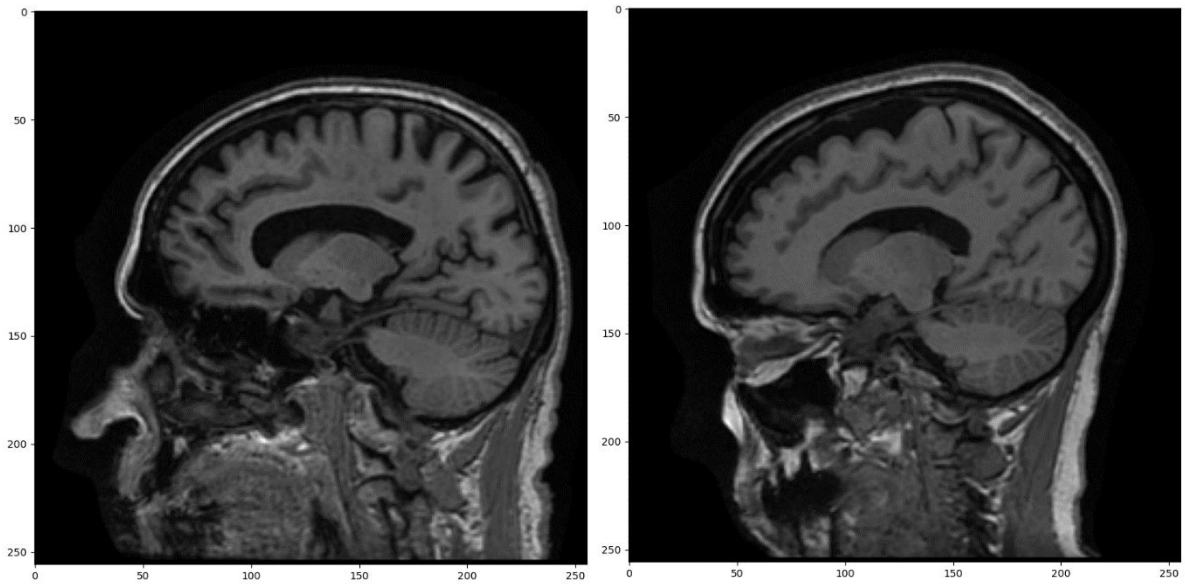


Figure 1. Block diagram of purpose methodology



a b

Figure 2. A sample of the dataset that was used in our study; a and b respectively represent the one T1 MRI (sagittal plane) slice with all of the brain tissues (skull, scalp, CSF, WM, and GM) of AD and MCI patients

Table 1. Explain parameters that used in FCM

Parameter	Explain	Parameter	Explain
X	Data	x_j	jth pixel
C	Total Cluster numbers	N	Number of pixels
u_{ij}	The degree of pixel x_i belonging to the jth cluster	z_i	ith cluster center
d_{ij}	Euclidean distance from sample point x_j to cluster center z_i	m	Fuzzy facto

2.3. Feature Extraction

The DWT is an image processing technique that provides a space-frequency representation of a given image and is often used in feature extraction methods for many classification purposes for medical applications [2, 25, 26]. The DWT gains its popularity for being efficient for classification since it extracts the structural information from an input dataset [2]. At each level of the decomposition, four sub-bands are obtained (LL, LH, HH, and HL). The LL is considered an approximation of the image, while the LH, HL, and HH are considered the vertical, horizontal, and diagonal details of the image [27, 28]. We have various types of mother wavelet functions, for example, Haar, Daubechies, Symlets, and Coiflets [28, 29]. In this work, we computed three levels of decomposition for the approximation coefficient based on the Haar basis function wavelet; Figure 3 shows the three-level decomposition of DWT of one sample.

The number of extracted features from the 3rd level of approximation is 1024 (32×32) features which is a high dimension of features to be used resulting in more classification complexity [2]. PCA is the most popular method between projections to subspace techniques [30]. This method provides the suboptimal solution with a low computational cost and lower computational complexity [31]. Therefore, the main idea behind using PCA in our approach is to reduce the dimensionality of the wavelet coefficients which results in a more efficient and accurate classifier [2]. The following algorithm is used to find out the principal components of the input matrix to the ensemble classifier, now the input matrix consists of only these principal components. The size of features

is reduced from 1024 to 7 [31]. The PCA method can be described below [25]:

a) Calculating the mean of the data

b) Constructing the covariance matrix

c) Calculating the eigenvalue and the eigenvector

d) Projecting the data to a new space based on eigenvectors and eigenvalues

Another group of features is statistical features; we use three moments: mean, standard deviation, and skewness for each of the T1_weighted MRI data that these parameters represent the average, standard deviation, and asymmetry of the distribution of grayscale images, respectively [32]. The equation of these statistical features is shown below [32] (Equations 5-7):

$$\mathrm{E} = \frac{1}{N}\sum_{j=1}^{N} I_j \tag{5}$$

$$\mathrm{R} = \sqrt[2]{\frac{1}{N}\sum_{j=1}^{N}(I_j - E)^2} \tag{6}$$

$$\mathrm{S} = \sqrt[3]{\frac{1}{N}\sum_{j=1}^{N}(I_j - E)^3} \tag{7}$$

The E parameter represents the mean, N is the total number of pixels, and I represent the intensity of each pixel's grayscale image value. In Equations 6 and 7, R represents standard deviation, and S represents the skewness of the image.

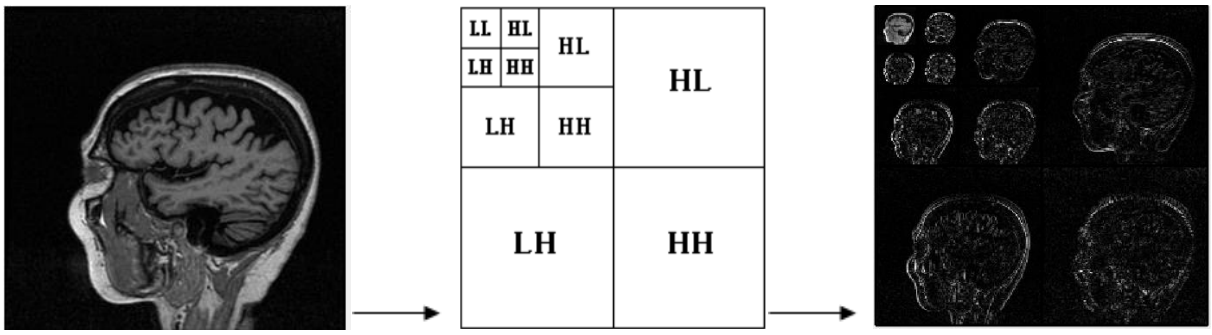


Figure 3. The LL is an approximation to the three-level decomposition of DWT of T1-weighted MRI images based on the Haar basis function

2.4. Ensemble Classifier

From the literature, it is noted that the consideration ability of the Computer-Aided-Brain-Diagnosis (CABD) system mainly depends on the type and quality of the classifier system, and sometimes a multi-machine performs better than a single machine [8]. Therefore, in this work, we use an ensemble classifier to classify T1-weighted MRI images for the detection of MCI from AD. For this purpose, we take three classifiers: KNN, DT, and LDA.

2.5. K-Nearest Neighbor

A K-Nearest Neighbor is a simple data mining algorithm with a variety of applications for image processing [33]. The KNN saves all samples and classifies new inputs according to similarity measures with distance calculation [34]. Here, K is the key hyper-parameter for the algorithm [8]. Let there exist two feature vectors of D dimensions, M = (M1, M2,..., Mn)T and M = (N1, N2,..., Nn)T, then the Euclidean distance can be shown to be [8] (Equation 8):

$$\mathrm{d}(M,N) = \sqrt{(M_1 - N_1)^2 + (M_2 - N_2)^2 + \cdots (M_n - N_n)^2} \tag{8}$$

The nearest k points are determined. Testing data points are classified according to specified nearest neighbors [19].

2.6. Decision Tree

A decision tree classification is a flowchart-like tree, in which the DT procedure starts with a root

node; leaf nodes are for the label of classes, and the intermediate nodes are referred to as the group of non-leaf nodes [35, 36]. The classification splitting process was continued based on the data values of the respective node [36]. In machine learning, the DT learns in the training procedure and the performance of the classifier is assessed during the testing phase [36]. Depth and cost function are the two most important hyper-parameters for the DT classifier, and the best value of these parameters is chosen using a grid search in a specific range.

2.7. Linear Discriminant Analysis

Bayesian learning is a statistical method to learn the structure of data for different purposes. Bayesian methods provide several structural learning algorithms [19]. The LDA classifier is a simple probabilistic classifier based on Bayesian learning. The LDA classifier uses a linear hyperplane to discriminate between classes according to the covariance matrix. This algorithm projects features to new space with linear transformation and finds the best coordinate that increases between class separability while decreases within class separability and takes high accuracy in many applications [25, 37, 38].

2.8. Voting

There are two main strategies for combining classifiers: fusion and selection [39, 40]. We assume that each classifier is located in a distinct region of the possible space. Therefore, when an instance is submitted for classification, the ensemble classifier coincides with the decision given by the classifier responsible for the region of the space to which the instance belongs [41]. In classifier fusion, all of the decisions are combined in some manner to make the ensemble classification [20]. Classifier fusion algorithms include combining methods, such as the average, majority vote, weighted majority vote, and the Borda Count [20]. In this work, we use a majority vote to fusion the results of all classifiers to classify T1_weighted MRI images.

3. Results

The dataset of this study was the Alzheimer's Disease Neuroimaging Initiative (ADNI), a multicenter study designed to develop clinical, imaging, and other biomarkers for the early detection of AD. In this work, we used T1 MRI data (in sagittal view) of 64 subjects; half of them are AD and the others are MCI (Table 2). Different patients have various slice numbers (166–180), so for the experiment and to get comparable results, 20 slices are extracted from each subject. Therefore, overall we have a dataset of 1,280 samples in 2D viewpoint clustering and a dataset of 64 samples in 3D viewpoint clustering. As explained before in our methodology, the WM has been segmented with FCM into two different views (Figure 4). The next two groups of features (DWT and Statistics) were extracted from WM. The dimension of DWT features was decreased to 7 by PCA and then added to 3 statistical features.

Table 2. The properties of the subject and the dementia state

Parameters	AD	MCI
Number of subjects	32	32
Male/Female	19/13	21/11
Age	75.93	74.97

The parameters boundary of classifiers is shown in Table 3. The parameter's optimal value is founded on the grid search method in the setting bound. For evaluating the performance, 10-fold cross-validation was applied to the ensemble classifier, then averaged in all folds to get the final result. The accuracy, precision, sensitivity (recall), specificity, and f1-score of the 3D view and 2D view ensemble machines are respectively represented in Tables 4 and 5. To evaluate the role of WM segmentation with FCM in the performance of classification, we implement all procedures with and without FCM and show the output in Figure 5. All the experiments were based on Python software, which was simulated on a core i7 processor with a 3.6 GHz speed, 8 GB of RAM, and the Windows 10 operating system.

4. Discussion

From the above results in Tables 4 and 5, the best performance was for the 3D viewpoint and achieved accuracy, sensitivity, and specificity of about 94.14%, 96.78%, and 91.54%, respectively. From the experimental information in Table 6 and Figure 5, it can be determined that our methodology has better

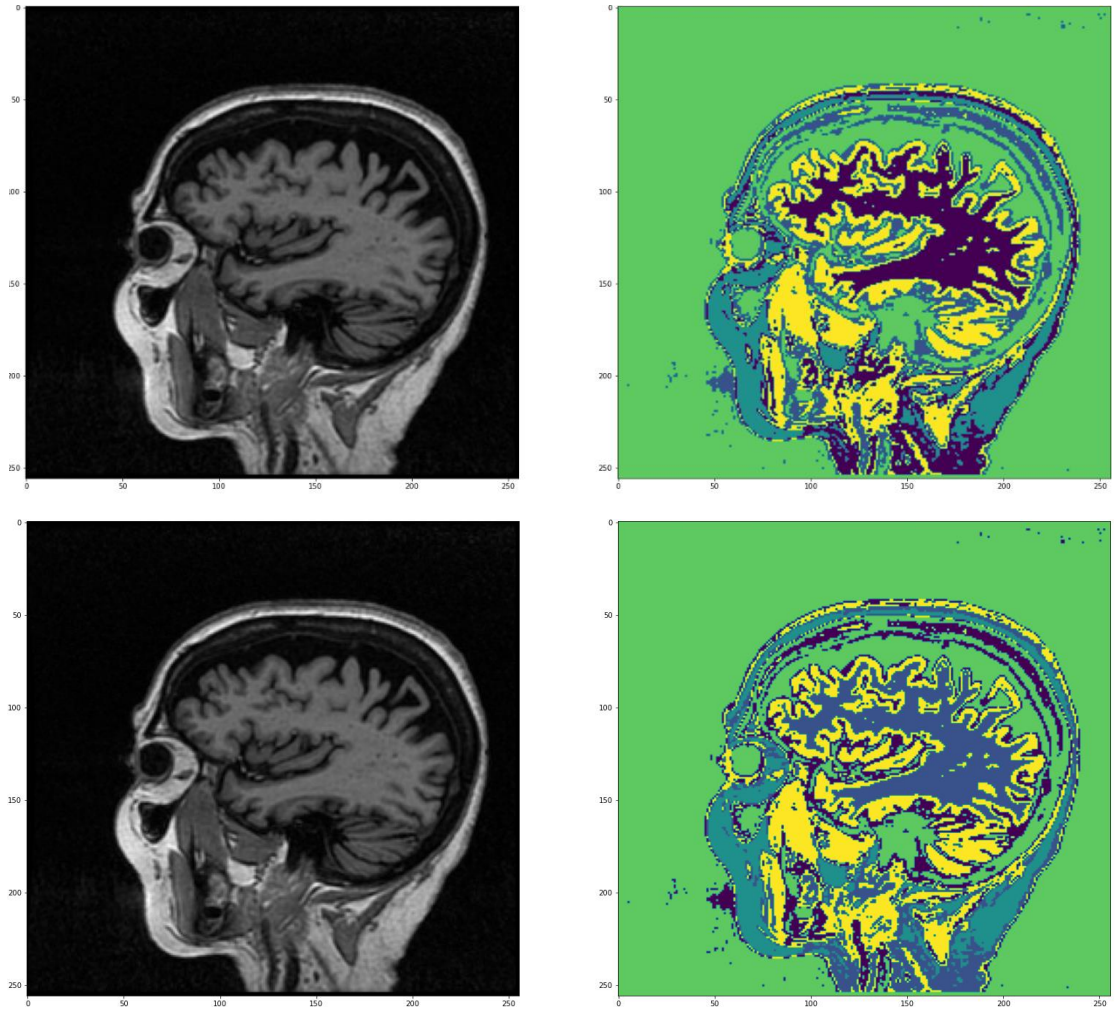


Figure 4. Segmentation of WM with multi-view FCM, first row 2D view clustering, second row 3D view clustering

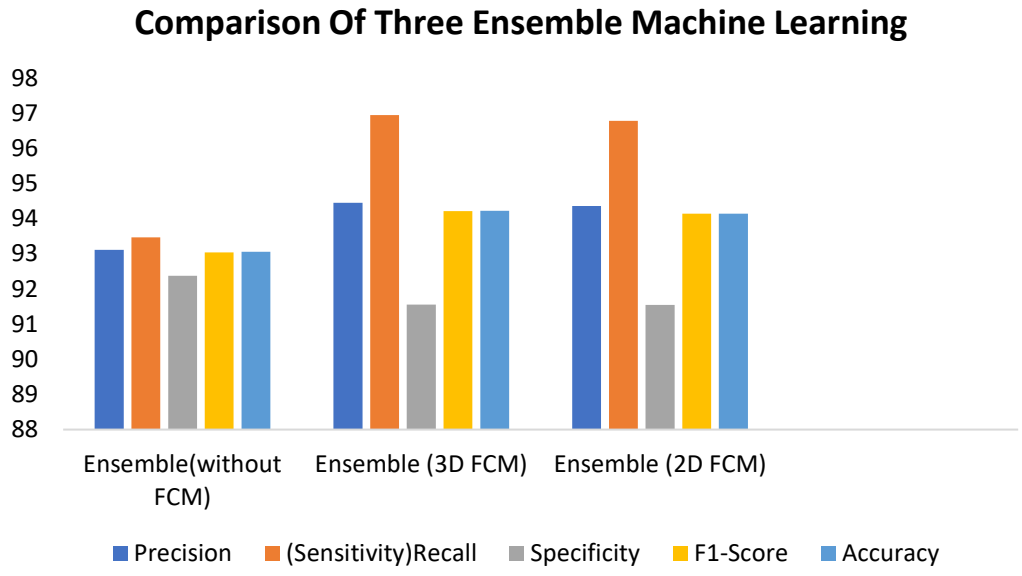


Figure 5. Compare Ensemble performance in three modes: Ensemble without WM segmentation (Ensemble (without FCM)), Ensemble with 2D FCM WM segmentation (Ensemble (2D FCM), and Ensemble with 3D FCM WM segmentation (Ensemble (3D FCM)

Table 3. The boundary of hyper-parameters used in our ensemble classifier and optimum value

Classifier	Parameter setting bound	Optimum
DT	Criterion = {'gini', 'entropy'}	entropy
	Maximum depth = {2,3,4,5,6,7,8,9,10,11,12,13,14,15}	8
KNN	Number of Nearest Neighbor K = {3,5,7,9,11,13,15}	9

Table 4. The performance of the 3D view ensemble classifier for distinct MCI from AD

Classifier	Precision	Sensitivity (Recall)	Specificity	F1-Score	Accuracy
KNN	65.57	65.16	64.88	64.75	64.77
DT	97.61	97.37	97.86	97.58	97.58
LDA	87.86	94.14	80.02	85.89	85.78
Ensemble	**94.45**	**96.94**	**91.55**	**94.21**	**94.22**

Table 5. The performance of the 2D view ensemble classifier for distinct MCI from AD

Classifier	Precision	(Sensitivity)Recall	Specificity	F1-Score	Accuracy
KNN	65.57	65.16	64.88	64.75	64.77
DT	97.67	97.51	97.86	97.66	97.66
LDA	87.86	94.14	80.02	85.89	85.78
Ensemble	**94.36**	**96.78**	**91.54**	**94.14**	**94.14**

performance when compared to other methods. Moreover, in our methodology, we need to choose the correct cluster label for WM segmentation and, as it has been shown, 3D clustering is less time-consuming than 2D because, in the 2D viewpoint, experts select the region of interest at all slices, while in the 3D viewpoint, they select once. The other feature of the purpose method is the possibility of physician interaction along the WM segmentation by reforming FCM output, which counts as a great property. It can increase the accuracy of the classification and improve the confidence of clinicians to use it in practical application [15]. Compared with other MCI distinctions from AD classifiers, the computational cost of this ensemble classification is low while the training procedure is shorter than 4 sec. According to the number of classifiers, the generalization is increased, which is an important point, especially in practice where we may work with various data sets and one classifier may fail in performance as before. This shows our method can be used in real-time clinical practice or pre-clinical implementation. Furthermore, despite some other methods, the number of cases in both groups of AD and MCI is equal in this work. That is of help to have a worthy classification result with less bias. Since in the real situation, the number of samples may be too small to learn a model, novel techniques exist that work with small datasets, and in our technique, the training model works efficiently with a relatively small sample size and high speed [15].

5. Conclusion

In this paper, to further improve the classification of AD disease, the purpose methodology is applied in this study. Based on the literature, few studies work on the detection of MCI from AD, and one of the important weaknesses of most past studies was high computational cost and low interaction with experts that could be used, such as a neurologist [15]. This study introduces a new methodology that, in addition to having a comparable performance for distinction MCI form AD, benefits from WM segmentation for interaction with experts, takes primary 3D FCM with few features in the feature extraction step, and fast

Table 6. The comparison of the purpose method and other best similar methods to classify MCI from AD with machine learning algorithms

Year	Modality	Machine Learning	Dataset	Validation	performance		
					Accuracy	Sensitivity	Specificity
2015	MRI	Linear SVM [42]	401 MCI 188 AD	2-fold	85.41	85.59	85.11
2015	FDG-PET + MRI	Linear SVM [43]	111MCI 70 AD	10-fold	84	87	81
2015	MRI	ANN [44]	172 MCI 180 AD	10-fold	94.88	94.18	95.55
2017	MRI	Multi kernel SVM [44]	102 MCI 89 AD	10-fold	75.12	73.92	77.24
2017	MRI	RBF SVM [45]	401 MCI 188 AD	10-fold	85	85	86
2017	MRI	Multi kernel SVM [46]	136 MCI 200 AD	10-fold	89.63	91.55	86.25
2018	MRI	Group lasso SVM [47]	210 MCI 137 AD	5-fold	65.7	63.2	67.3
2018	MRI	Multi kernel SVM [48]	280 MCI 200 AD	10-fold	90.41	92.83	88.82
2018	MRI	rMLTFL [49]	192 MCI 102 AD	10-fold	76.7	61.4	81.8
2018	MRI	MFN [50]	221 MCI 142 AD	10-fold	73.83	64.08	80.09
2019	MRI	SVM+RF+KNN [51]	396 MCI 189 AD	K=1:50-fold	63.41	57.29	65.35
2019	MRI	Ensemble linear discriminant [52]	607 MCI 249 AD	10-fold	92.8	95.8	88.3
2019	MRI+PET +SNP	Multi kernel SVM [53]	93 MCI 49 AD	10-fold	76.9	65.9	82.7
2022	**MRI**	**KNN+DT+LDA**	**32 MCI 32 AD**	**10-fold**	**94.14**	**96.78**	**91.54**

Abbreviation: MFN: Multi-Feature-based Network; rMLTFL: robust Multi-Label Transfer Learning; RBF: Radial Basis Function; RF: Random Forest; FDG: Fluorodeoxyglucose; SNP: Single Nucleotide Polymorphisms

ensemble classifier to increase the efficiency of our algorithm. These advantages show our method can be used in practical real-time clinical and preclinical applications. For feature study, more levels of AD progression such as MCI converters (MCIc), non-converters (MCInc), and amnestic MCI (aMCI) vs. non-amnestic MCI (naMCI) could be considered in classification and also using other recent FCM algorithms for WM segmentation. We are seeing some of the other tissue, especially the skull, in the clustering procedure, which could decrease it with skull removal or brain segmentation algorithms to improve the performance of classification.

References

1- Dennis J Selkoe and Peter J Lansbury, "Alzheimer's disease is the most common neurodegenerative disorder." *Basic Neurochemistry: molecular, cellular and medical aspects*, Vol. 6pp. 101-02, (1999).

2- Heba Mohsen, El-Sayed A El-Dahshan, El-Sayed M El-Horbaty, and Abdel-Badeeh M Salem, "Classification of brain MRI for Alzheimer's disease based on linear discriminate analysis." *Egyptian Computer Science Journal*, Vol. 41 (No. 3), pp. 44-52, (2017).

3- Fabio Previtali, Paola Bertolazzi, Giovanni Felici, and Emanuel Weitschek, "A novel method and software for automatically classifying Alzheimer's disease patients by magnetic resonance imaging analysis." *Computer methods and programs in biomedicine*, Vol. 143pp. 89-95, (2017).

4- KR Kruthika, Akshay Pai, HD Maheshappa, and Alzheimer's disease Neuroimaging Initiative, "Classification of alzheimer and MCI phenotypes on MRI data using SVM." in *International Symposium on Signal Processing and Intelligent Recognition Systems*, (2017): *Springer*, pp. 263-75.

5- Debesh Jha, Ji-In Kim, and Goo-Rak Kwon, "Diagnosis of Alzheimer's disease using dual-tree complex wavelet transform, PCA, and feed-forward neural network." *Journal of healthcare engineering*, Vol. 2017, (2017).

6- Barbara J Grabher, "Effects of Alzheimer disease on patients and their family." *Journal of nuclear medicine technology*, Vol. 46 (No. 4), pp. 335-40, (2018).

7- Hanns Hippius and Gabriele Neundörfer, "The discovery of Alzheimer's disease." *Dialogues in clinical neuroscience*) , (2022).

8- U Rajendra Acharya *et al.*, "Automated detection of Alzheimer's disease using brain MRI images–a study with various feature extraction techniques." *Journal of Medical Systems*, Vol. 43 (No. 9), pp. 1-14, (2019).

9- Sitara Afzal *et al.*, "Alzheimer disease detection techniques and methods: a review".(2021).

10- Ajax E George *et al.*, "CT diagnostic features of Alzheimer disease: importance of the choroidal/hippocampal fissure complex." *American journal of Neuroradiology*, Vol. 11 (No. 1), pp. 101-07, (1990).

11- Sara E Nasrabady ,Batool Rizvi, James E Goldman, and Adam M Brickman, "White matter changes in Alzheimer's disease: a focus on myelin and oligodendrocytes." *Acta neuropathologica communications*, Vol. 6 (No. 1), pp. 1-10 , (2018).

12- MLF Balthazar, CL Yasuda, FR Pereira ,T Pedro, BP Damasceno, and F Cendes, "Differences in grey and white matter atrophy in amnestic mild cognitive impairment and mild Alzheimer's disease." *European journal of neurology*, Vol. 16 (No. 4), pp. 468-74, (2009).

13- Xiaojuan Guo *et al.*, "Voxel-based assessment of gray and white matter volumes in Alzheimer's disease." *Neuroscience letters*, Vol. 468 (No. 2), pp. 146-50, (2010).

14- Lars Frings *et al.*, "Longitudinal grey and white matter changes in frontotemporal dementia and Alzheimer's disease." *PloS one*, Vol. 9 (No. 3), p. e90814, (2014).

15- Muhammad Tanveer *et al.*, "Machine learning techniques for the diagnosis of Alzheimer's disease: A review." *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, Vol. 16 (No. 1s) ,pp. 1-35, (2020).

16- Soheil Ahmadzadeh Irandoost and Fatemeh Asadi, "Classification of brain MRI for Alzheimer's disease detection based on ensemble machine learning." *Iranian Journal of Radiology*, Vol. 16 (No. Special Issue), (2019).

17- Stanisław Adaszewski, Juergen Dukart, Ferath Kherif, Richard Frackowiak, Bogdan Draganski, and Alzheimer's Disease Neuroimaging Initiative, "How early can we predict Alzheimer's disease using computational anatomy?" *Neurobiology of aging*, Vol. 34 (No. 12), pp. 2815-26, (2013).

18- Stefan Klöppel *et al.*, "Automatic classification of MR scans in Alzheimer's disease." *Brain*, Vol. 131 (No. 3), pp. 681-89, (2008).

19- KR Kruthika, HD Maheshappa, and Alzheimer's Disease Neuroimaging Initiative, "Multistage classifier-based approach for Alzheimer's disease prediction and retrieval." *Informatics in Medicine Unlocked*, Vol. 14pp. 34-42, (2019).

20- M Paz Sesmero, Agapito I Ledezma, and Araceli Sanchis, "Generating ensembles of heterogeneous classifiers using stacked generalization." *Wiley interdisciplinary reviews: data mining and knowledge discovery*, Vol. 5 (No. 1), pp. 21-34, (2015).

21- Andreas Holzinger, "Interactive machine learning for health informatics: when do we need the human-in-the-loop?" *Brain Informatics*, Vol. 3 (No.2, pp. 119-31, (2016).

22- Andreas Holzinger, Markus Plass, Katharina Holzinger, Gloria Cerasela Crişan, Camelia-M Pintea, and Vasile Palade, "Towards interactive Machine Learning (iML): applying ant colony algorithms to solve the traveling salesman problem with the human-in-the-loop approach." in *International Conference on Availability, Reliability, and Security*, (2016): *Springer*, pp. 81-95.

23- [Online]. Available: http://adni.loni.usc.edu./

24- Lei Hua, Yi Gu, Xiaoqing Gu, Jing Xue, and Tongguang Ni, "A novel brain MRI image segmentation method using an improved multi-view fuzzy c-means clustering algorithm." *Frontiers in Neuroscience*, Vol. 15p. 662674 , (2021).

25- Saruar Alam, Goo-Rak Kwon, Ji-In Kim, and Chun-Su Park, "Twin SVM-based classification of Alzheimer's disease using complex dual-tree wavelet principal coefficients and LDA." *Journal of healthcare engineering*, Vol. 2017, (2017).

26- AB Al-Khafaji, "Classification of MRI Brain Images using Discrete Wavelet Transform And K-NN." *International Journal Of Engineering Sciences & Research Technology*, Vol. 4 (No. 11), (2015).

27- Hashem Kalbkhani, Mahrokh G Shayesteh, and Behrooz Zali-Vargahan, "Robust algorithm for brain magnetic resonance image (MRI) classification based on GARCH variances series." *Biomedical Signal Processing and Control*, Vol. 8 (No. 6), pp. 909-19, (2013).

28- Debesh Jha, Ji-In Kim, Moo-Rak Choi, and Goo-Rak Kwon, "Pathological brain detection using weiner filtering, 2D-discrete wavelet transform, probabilistic PCA, and random subspace ensemble classifier." *Computational intelligence and neuroscience*, Vol. 2017, (2017).

29- Alexia Tzalavra *et al.*, "Comparison of multi-resolution analysis patterns for texture classification of breast tumors based on DCE-MRI." in *International Workshop on Machine Learning in Medical Imaging*, (2016): *Springer*, pp. 296-304.

30- Bekir Karlık, Yücel Koçyiğit, and Mehmet Korürek, "Differentiating types of muscle movements using a wavelet based fuzzy clustering neural network." *Expert Systems*, Vol. 26) No. 1), pp. 49-59, (2009).

31- N Hema Rajini and R Bhavani, "Classification of MRI brain images using k-nearest neighbor and artificial neural network." in *2011 International conference on recent trends in information technology (ICRTIT)*, (2011): *IEEE*, pp.563-68.

32- Muhammad Nazir, Fazli Wahid, and Sajid Ali Khan, "A simple and intelligent approach for brain MRI classification." *Journal of Intelligent & Fuzzy Systems*, Vol. 28 (No. 3), pp. 1127-35, (2015).

33- Xindong Wu *et al.*, "Top 10 algorithms in data mining." *Knowledge and information systems*, Vol. 14 (No. 1), pp. 1-37, (2008).

34- El Mehdi Benyoussef, Abdeltif Elbyed, and Hind El Hadiri, "3D MRI classification using KNN and deep neural network for Alzheimer's disease diagnosis." in *International Conference on Advanced Intelligent Systems for Sustainable Development*, (2018): *Springer*, pp. 154-58.

35- M Evanchalin Sweety and G Wiselin Jiji, "Detection of Alzheimer disease in brain images using PSO and Decision Tree Approach." in *2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies*, (2014): *IEEE*, pp. 1305-09.

36- S Naganandhini and P Shanmugavadivu, "Effective diagnosis of Alzheimer's disease using modified decision tree classifier." *Procedia Computer Science*, Vol. 165pp. 548-55, (2019).

37- Usha BL and Supreeth HSG, "Brain Tumor detection and identification in brain MRI using supervised learning: A LDA based classification method." *Image*, Vol. 1p. 2, (2107).

38- Yong Zhang, Xiaobo Zhou, Rochelle M Witt, Bernardo L Sabatini, Donald Adjeroh, and Stephen TC Wong, "Dendritic spine detection using curvilinear structure detector and LDA classifier." *NeuroImage*, Vol. 36 (No. 2), pp. 346-60, (2007).

39- Lorenza Saitta, "Integrated architectures for machine learning." in *Advanced Course on Artificial Intelligence*, (1999): *Springer*, pp. 218-29.

40- Ludmila I Kuncheva, "Switching between selection and fusion in combining classifiers: An experiment." *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, Vol. 32 (No. 2), pp. 146-56, (2002).

41- Xingquan Zhu, Xindong Wu, and Ying Yang, "Dynamic classifier selection for effective mining from noisy data streams." in *Fourth IEEE International Conference on Data Mining (ICDM'04)*, (2004): *IEEE*, pp. 305-12.

42- Laila Khedher, Javier Ramírez, Juan Manuel Górriz, Abdelbasset Brahim, and IA Illán, "Independent component analysis-based classification of Alzheimer's disease from segmented MRI data." in *International Work-Conference on the Interplay between Natural and Artificial Computation*, (2015): *Springer*, pp. 78-87.

43- Andrés Ortiz, Jorge Munilla, Ignacio Álvarez-Illán, Juan M Górriz, Javier Ramírez, and Alzheimer's Disease Neuroimaging Initiative, "Exploratory graphical models of functional and structural connectivity patterns for Alzheimer's Disease diagnosis." *Frontiers in computational neuroscience*, Vol. 9p. 132, (2015).

44- Saruar Alam, Goo-Rak Kwon, and Alzheimer's Disease Neuroimaging Initiative, "Alzheimer disease classification using KPCA, LDA, and multi-kernel learning SVM." *International Journal of Imaging Systems and Technology*, Vol. 27 (No. 2), pp. 133-43, (2017).

45- Laila Khedher, Ignacio A Illán, Juan M Górriz, Javier Ramírez, Abdelbasset Brahim, and Anke Meyer-Baese, "Independent component analysis-support vector machine-based computer-aided diagnosis system for Alzheimer's with visual support." *International journal of neural systems*, Vol. 27 (No. 03), p. 1650050, (2017).

46- Jin Liu, Jianxin Wang, Bin Hu, Fang-Xiang Wu, and Yi Pan, "Alzheimer's disease classification based on individual hierarchical networks constructed with 3-D texture features." *IEEE transactions on nanobioscience*, Vol. 16 (No. 6), pp. 428-37, (2017).

47- Zhuo Sun, Yuchuan Qiao, Boudewijn PF Lelieveldt, Marius Staring ,and Alzheimer's Disease NeuroImaging Initiative, "Integrating spatial-anatomical regularization and structure sparsity into SVM: Improving interpretation of Alzheimer's disease classification." *NeuroImage*, Vol. 178pp. 445-60, (2018).

48- Jin Liu, Min Li ,Wei Lan, Fang-Xiang Wu, Yi Pan, and Jianxin Wang, "Classification of Alzheimer's disease using whole brain hierarchical network." *IEEE/ACM transactions on computational biology and bioinformatics*, Vol. 15 (No. 2), pp. 624-32, (2016).

49- Bo Cheng, Mingxia Liu, Daoqiang Zhang, and Dinggang Shen, "Robust multi-label transfer feature learning for early diagnosis of Alzheimer's disease." *Brain imaging and behavior,* Vol. 13 (No. 1), pp. 138-53, (2019).

50- Weihao Zheng, Zhijun Yao, Yuanwei Xie, Jin Fan, and Bin Hu, "Identification of Alzheimer's disease and mild cognitive impairment using networks constructed based on multiple morphological brain features." *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging,* Vol. 3 (No. 10), pp. 887-97, (2018).

51- Krishnakumar Vaithinathan, Latha Parthiban, and Alzheimer's Disease Neuroimaging Initiative, "A novel texture extraction technique with T1 weighted MRI for the classification of Alzheimer's disease." *Journal of neuroscience methods,* Vol. 318pp. 84-99, (2019).

52- Ali Ezzati *et al.*, "Optimizing machine learning methods to improve predictive models of Alzheimer's disease." *Journal of Alzheimer's Disease,* Vol. 71 (No. 3), pp. 1027-36, (2019).

53- Jialin Peng, Xiaofeng Zhu, Ye Wang, Le An, and Dinggang Shen" ,Structured sparsity regularized multiple kernel learning for Alzheimer's disease diagnosis." *Pattern recognition,* Vol. 88pp. 370-82, (2019).