Multi-Features Fusion in Multi-plane MRI Images for Alzheimer's Disease Classification

**Cucun Very Angkoso[1,2]   Hapsari Peni Agustin Tjahyaningtijas[3]   Yudhi Adrianto[4]
Anggraini Dwi Sensusiati[5]   I Ketut Eddy Purnama[1,6,7]   Mauridhi Hery Purnomo[1,6,7]***

[1]Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Indonesia
[2]Informatics Department, University of Trunojoyo Madura, Bangkalan, Indonesia
[3]Department of Electrical Engineering, Universitas Negeri Surabaya, Indonesia
[4]Departement of Neurology, Universitas Airlangga, Indonesia
[5]Departement of Radiology, Universitas Airlangga, Indonesia
[6]Departement of Computer Engineering, Institut Teknologi Sepuluh Nopember, Indonesia
[7]University Center of Excellence on Artificial Intelligence for Healthcare and Society (UCE AIHeS), Indonesia
* Corresponding author's Email: hery@ee.its.ac.id

Abstract: Alzheimer's disease (AD) is characterized by severe memory loss, typical in dementia. This disease has serious public health consequences (high incidence, prevalence, and mortality rates), as well as significant health and social costs. Therefore, AD should be recognized as a disease rather than a natural occurrence that affects everyone, allowing for early detection and treatment to begin before the situation worsens. Previous studies have focused on shrinkage in certain brain locations (typically the hippocampus, cerebral cortex, and brain ventricles), ruling out the potential of atrophy in other areas. Therefore, our study proposed an algorithm for detecting and classifying AD which uses the whole brain area. While previous studies focused mostly on a single plane, we propose to exploit all planes (multi-plane) of the MRI, including the axial, coronal, and sagittal planes, to obtain more detailed whole-brain imaging characteristics. A multi-feature fusion of texture-based feature extraction, including First-order statistics (FOS), Gray Level Co-occurrence Matrix (GLCM), Local Binary Patterns (LPB), and Gray Level Run Length Matrix (GLRLM), were used as input of our classifier. We used T1 and T2-weighted structural MRI from the Alzheimer's Disease Neuroimaging Initiative (ADNI) MRI database and Brain-Atrophy (BA) MRI from the Airlangga University Hospital (AUH) to evaluate our proposed method. A significantly higher prediction accuracy confirms the effectiveness of the proposed method. As all of the extracted features were employed, the accuracy of the AD and CN classifiers with the multi-plane approach can be improved by up to 20.41% compared to the single-plane approach. In addition, our proposed method has the highest accuracy of 0.967 in binary classification tasks and 0.867 in multiclass classification tasks, outperforming previous works reported in the related references. Furthermore, the multi-plane analysis strategy has proven superior to the single-plane approach in all evaluations.

Keywords: Alzheimer's disease, Feature fusion, Medical image analysis, MRI, Multi-plane.

1. Introduction

Alzheimer's disease (AD) is one of the most common diseases in the elderly, and it becomes more common as people age. The neurological brain condition in AD differs from other types of dementia since it is progressive neurodegenerative, indicating that the brain's ability to function weakens with time [1]. Since not everyone with Mild Cognitive Impairment (MCI) develops into AD [2], it is critical to treat it as soon as possible to maintain their quality of life. However, according to 32 studies, 32% of people with MCI develop into AD after five years [3]. Thus, a comprehensive study on the early detection of AD classification is required to diagnose cognitive impairments at an early phase.

AD causes numerous brain areas to shrink (atrophy) as numerous neurons in the brain are

(a)

(b)

(c)
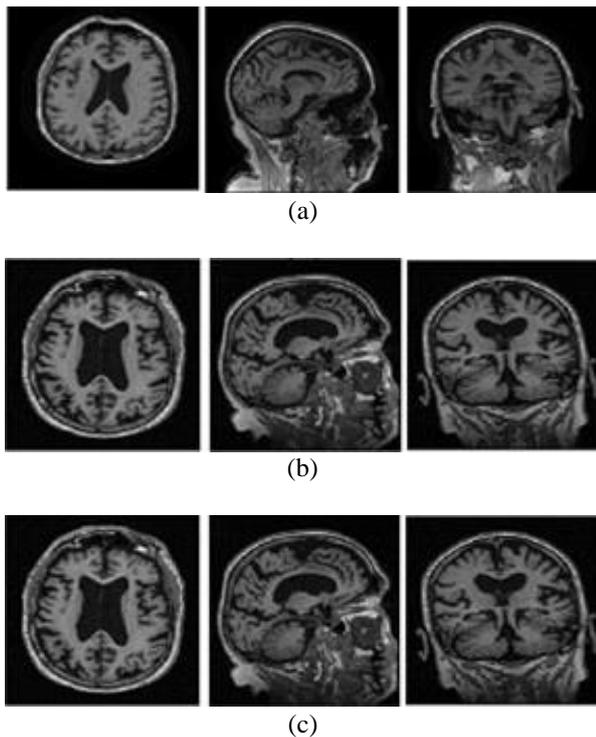
Figure. 1 Sample images from the database: (a) normal, (b) MCI, and (c) AD, from left to right: axial, coronal, sagittal

damaged and die, and connections between neuron networks are broken. The term "extensive brain atrophy" refers to a stage of brain atrophy in which the volume of the brain has significantly decreased.

As shown in Fig. 1, visible clinical conditions of the AD patients are the hippocampus and cerebral cortex atrophy, while the brain's ventricles swell. Furthermore, on the Region of Interest (ROI)-based observations, clinicians analyze Magnetic Resonance Imaging (MRI) of the brain to diagnose AD by evaluating clinical information in the patient, such as shrinking the hippocampus and cerebral cortex and enlarging the brain's ventricles [4-8]. As a result, MRI has become a supporting diagnostic tool for AD [3, 8-10]. MRI has a high spatial resolution that allows soft tissue structures to be seen with excellent clarity and detail compared to other imaging techniques. It also has the advantage of being a non-invasive diagnostic approach (without surgery). However, a comprehensive medical examination, including medical record, Mini-Mental State Examination (MMSE), clinical dementia rate (CDR), physical, and neurobiological investigations, is required for a complete diagnosis of AD [11].

The ROI-based approach to studying changes in the cerebrum and hippocampus has several advantages, but it also has some drawbacks, including:

a. The ROI-based method determines the ROI position based on the operator's knowledge and expertise, so the operator's experience significantly impacts the validity of detection.

b. There is no evidence that brain regions other than the hippocampus and entorhinal cortex certainly have no contribution to AD detection.

c. Based on actual data, the ROI auto-segmentation method produces data that is unsuitable in practice, causing examiners to segment the brain image manually.

Our study employs Machine Learning (ML) techniques to capture and develop new aspects that will support the classification of the entire brain. As a result, no specified ROI is demanded.

ML has been studied extensively in recent years for AD detection using various forms of data such as MRI, Positron Emission Tomography (PET), and functional MRI. Tong et al. used a Nonlinear Graph Fusion (NGF) approach and various modalities, including MRI, Fludeoxyglucose-PET (FDG-PET) CerebroSpinal Fluid (CSF), and Genes, to classify AD, MCI, and Cognitive Normal (CN) [12]. Zeng et al. then used MRI and performed a binary classification of AD vs. CN using the combination Switching Delayed Particle Swarm Optimization (SDPSO)-Support Vector Machine (SVM) Principal Component Analysis (PCA) method [13]. Vaithinathan et al. used MRI input, but the brain images were first segmented into Gray Matter (GM) and White Matter (WM). Then, the binary classification of AD versus CN was performed using the combined technique of SVM, Random Forest Classifier (RFC), and K-Nearest Neighbor (KNN) [14]. Peng et al. [15] offered another investigation using multimodal data, retaining the SVM as a classifier in categorizing AD and CN, with combined multimodal data including MRI, FGD-PET, and Single Nucleotide Polymorphisms (SNP). In addition, another SVM study in the multiclass classification task of AD, MCI, and CN had also been proposed [10, 11]. Furthermore, Jimenez-Mesa et al. (2020) categorized AD into four classes, AD, MCI, cMCI, and CN, using the SVM classifier for MRI data [16]. Our previous study observed deep learning approaches [17]; but, it requires substantially more data to achieve positive results than other ML techniques. Due to its complicated architectural model, it also has a high computational cost during the learning phase. Furthermore, deep learning will require expensive GPUs and other supporting hardware to speed up the training process.

Previous studies still have a lot of room for improvement, such as using single-plane MRI when three planes are available, improving the accuracy of

multiclass classification, and using multimodal data, which is impractical in the data collection process. Our proposed method provides a comprehensive investigation of all available MRI planes. This multi-plane analysis approach aims to obtain a detailed description of the investigated images based on the concept that MRI brain images are 3D images.

Unlike previous studies that used only a few statistical features, we used multiple order statistical feature extraction to obtain a detailed image description. In addition, we perform a feature selection process to keep computational efficiency and improve system accuracy by only using essential features and removing the unimportant ones. Finally, we evaluated the effectiveness of our approach using various ML algorithms for binary and multiclass classifications tasks.

We introduce new approaches to our proposed method in response to the drawbacks of earlier AD investigations. The following are a brief summary of our study's contribution:

a. Because atrophy is one of the clinical features of AD, we examined the entire brain image to avoid excluding areas that were also damaged by the atrophic condition.

b. We propose a simultaneous analysis of three MRI-plane based on the ML method for automatic AD detection.

c. We propose a new MRI slice selection method that cuts the need to process all slices.

d. Developed binary and multiclass AD classification based on multi-feature fusion FOS, second-order, and high-order statistical features (textural features) accompanied by the Boruta feature optimization and various classifiers.

The rest of this paper is structured as follows: Section 2 describe the dataset and the related methods we used. The method we used as our proposed solution for the topic is detailed in Section 3. Then, we presented the experimental results in several scenarios for different purposes. Next, we had results and discussion in Section 4, and in the last section, we present our findings as a conclusion.

2. Dataset

We collected MRI image data from Alzheimer's Disease Neuroimaging Initiative (ADNI) because it is readily accessible via the internet (http://www.loni.ucla.edu/ADNI) with organized and well-processed data. The ADNI data comprises T1 and T2 MRI images from 1.5T and 3T MRI scanners. In addition, for anonymization data, ADNI has transformed the image format to Neuroimaging
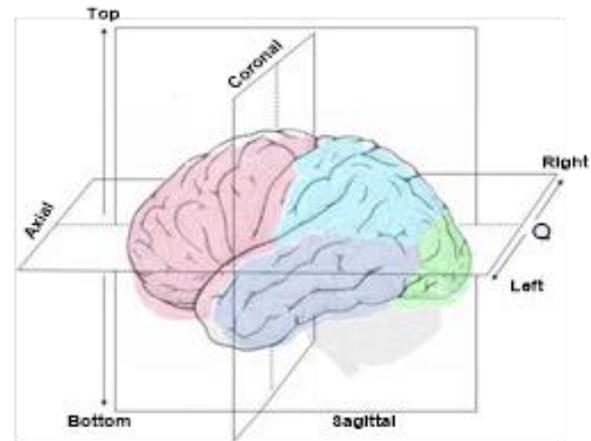
Figure. 2 Three planes in MRI brain imaging

Table 1. Demographic data of ADNI dataset

Data	AD	MCI	CN
Number of instances	100	100	100
Gender	55M/45F	50M/50F	50M/50F
MMSE +- SD	27.03+-2.60	21.88+-2.15	35.22+-1.24

Informatics Technology Initiative (Nifti). We used the baseline ADNI-1 dataset from a 1.5T Tesla scanner, preprocessed with Magnetization Prepared- RApid Gradient Echo (MP-RAGE) with a resolution of 256x256x170 voxels, to evaluate our proposed system.

We selected 300 ADNI images, with 100 images for each AD, MCI, and CN class. Table 1 shows our ADNI dataset's demographic distribution.

Furthermore, 20 MRI data of patients with Brain Atrophy (BA) from Airlangga University Hospital (AUH) in Surabaya, Indonesia, were gathered for comparison. The competent team of doctors at AUH Surabaya confirmed the atrophic conditions in all of the patients examined. We used extra data from 20 normal brain images to perform a binary classification task using predefined classifiers for this additional evaluation. All AUH data has been administratively reviewed and committed to subject privacy and data anonymity, ensuring that it is used in accordance with academic research standards. Furthermore, the ethical clearance was granted from the ethical committee of AUH with number 142/KEP/2020.

3. Methods of related works

The study of neuroimaging data has sparked a lot of interest since it has been shown to help with the early diagnosis of AD. The study is critical because, with proper treatment, people with MCI should not be converted to AD. An MRI scanner provides three

orientation planes: axial, coronal, and sagittal planes. Fig. 2 shows an illustration of each plane available in an MRI. However, most studies have concentrated on the axial plane only without involving the entire plane of MRI [11, 17, 18]. Therefore, we expect that using this entire plane will result in a more comprehensive description of AD characteristics and more accurate diagnostic performance. The following sections discuss relevant methodologies used in our study, especially AD classification using digital image analysis.

3.1 Features extraction

Feature extraction is crucial in medical image analysis, especially in the classification system. The extracted features are compact representations of the patterns containing only the most critical information for object recognition. Furthermore, feature extraction decreases the resources such as memory and computing power by presenting complex data in a more simple format [11]. The hippocampus, entorhinal cortex, and atrophy of the entire brain are some of the MRI markers that have been proved to have diagnostic significance with AD and have been used as features input for Machine Learning (ML) algorithms for predicting both MCI and AD [19]. Because there was no evidence of atrophy occurring selectively in specific locations, we decided to do feature extraction on the whole area of the MRI images.

In digital image studies, texture analysis is associated with the attempts to determine the characteristic of pixel patterns associated with rough, smooth, and wavy features due to spatial variations in pixel intensity and pattern [18]. This feature is one of the most important aspects of image data since the repeating pattern can be evaluated based on the intensity of the pixels or color variations. Texture analysis usually uses statistical calculation approaches based on the spatial distribution of gray level values around image pixels. First-order statistical (FOS) computations are related to the gray level distribution of the image, whereas second-, third-, and higher-order statistical calculations are approaches that consider the relationship between two, three, or more pixels. FOS cannot provide information about the relative positions of the image's various levels of gray. We need higher-order statistical calculations to examine the pattern relationships among pixels in the image.

This study combines FOS-based features with features from various statistical order calculations, i.e., Gray Level Co-occurrence Matrix (GLCM) [19],
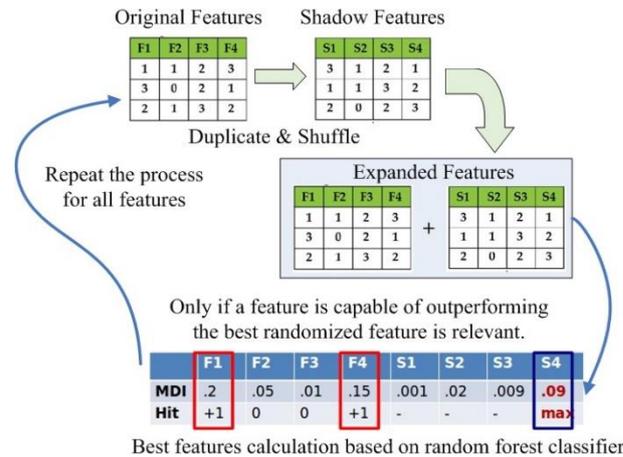
Figure. 3 Boruta feature selection process

Gray Level Run Length Matrix (GLRLM) [20], and Local Binary Patterns (LBP) [21]. GLCM is concerned with the relationship between two pixels and the quantity of all gray level value combinations that occur in a given direction and distance between them [18]. GLRLM, on the other hand, is more concerned with a collection of consecutive pixels in the same direction that has the same gray level [22]. Furthermore, GLRLM is usually calculated in four directions with a run-length histogram for each direction, whereas GLCM is usually calculated in eight directions [18]. We call our approach the multi-features fusion method.

In addition to FOS, GLCM, and GLRM, we also investigate the effectiveness of LBP-based feature extraction as a representative of high-order statistical texture. In contrast to the previous three textural features, LBP concentrates on intensity transition patterns within the region of interest's subregions by combining local structures structural and statistical examination [18]. LBP evaluates the intensity difference between a pixel and its eight neighbors. A direct comparison of each neighbor to the center pixel gives an 8-bit binary vector. The 8-bit binary-coded decimal value is then labeled at these pixels, and the texture descriptor is the histogram of these LBP labels

3.2 Features selection

The process of controlling the number of features in a classification model is recognized as feature selection (FS). The primary goal of FS is to eliminate redundant information and reduce the feature dimension. FS is required in order to determine which features are most relevant or which combinations of feature types are most important in distinguishing AD. There have been several approaches proposed for evaluating important features. Finding the optimal feature set using fisher criterion and t-test scores is

one of the FS techniques used in AD studies [23]. In terms of dimensionality reduction techniques such as Principal Component Analysis (PCA), which converts features to lower dimensions, FS selects and eliminates specific features without modifying them.

In general, FS techniques can only select non-redundant features, and they frequently overlook redundant features that are relevant and essential. The Boruta FS attempts to solve this issue [24, 25]. First, the Boruta technique creates duplicate input features called shadow features to expand the information. Then, the mean decrease impurity (MDI) is calculated from the expanded feature set (original features + shadow features) using a random forest classifier to quantify the importance of features. The higher MDI value indicates that the feature is more important than the others.

Next, Boruta FS analyzes whether the original feature is more important than its best shadow feature using Z-score evaluation at each iteration. Then, Boruta FS removes features that consider unimportant. Finally, the algorithm stops when all features have been validated or rejected or the random forest has reached a specified limit. Fig. 3 illustrates the Boruta feature selection method.

3.3 The classifier

A classifier is an algorithm used in classification tasks and a supervised learning approach that is used to identify categories in test data based on training data. ML has been used in several studies to diagnose AD using either binary or multiclass classification tasks. Despite the positive results of studies using supervised ML based on MRI to classify AD [10, 26, 27], there is still room for improvement.

The algorithm that can distinguish between people with brain diseases and healthy people would undoubtedly help physicians in their work. [28]. This algorithm is utilized after the feature extraction process. The extracted features are evaluated first to reduce the number of features processed without dropping system performance. Several classification methods were put to the test to classify three classes AD, MCI, and CN (healthy people). Several classifiers were chosen to examine the impact of the multi-plane approach compared to the single plane. We did not focus on studying the advantages of each classifier since our objective was to get the advantages of the multi-plane analysis approach from each. Each of the classifiers utilized in our study is highlighted in this section.

A decision tree is a supervised learning algorithm that is used when the independent and dependent variables have a nonlinear or complex nonlinear
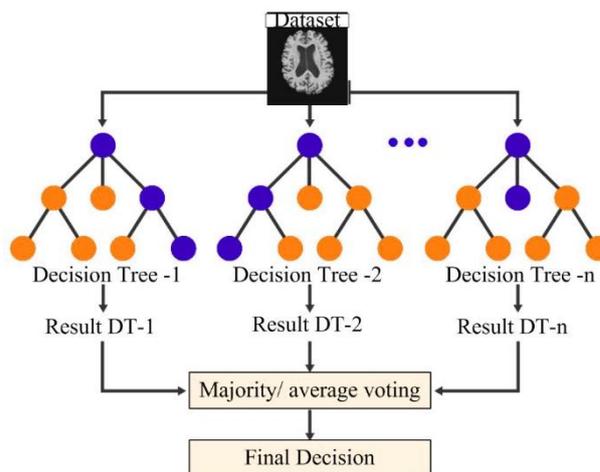
Figure. 4 Diagram of the random forest classifier

relationship. The first step is to separate the dataset into nodes/leaves. The node is then subdivided into different subnodes based on the number of target variables. A terminal node is a node that cannot be broken down further into subnodes [29].

AdaBoost is a highly efficient ensemble learning algorithm that could improve the classification accuracy of weak classifiers. AdaBoost combines a group of weak classifiers to form a robust classifier and selects the weak classifier with the lowest error [30].

Overfitting is a problem that frequently occurs in decision trees. As shown in Fig. 4, Random forest, also known as ensemble bagged tree, is a type of decision tree that reduces overfitting by combining multiple trees and selecting the best performer. The first step is to divide the dataset into N samples, after which general learning rules are applied to each sample. The predictions generated in this step are then combined through voting or by taking the average [31].

Gradient boosting is an efficient method for developing predictive models as part of supervised learning techniques. The model is suitable in machine learning algorithms for handling Regression and Distribution problems based on clusters of poor provisioning models found in most decision trees. This method extends the development of additive-boosting in which a faint model is built as a gradient descent algorithm on a function [32]. Modules are constructed using a correlative technique similar to other methods. Several single faint modules are used collectively to produce an accurate model.

Linear Discriminant Analysis (LDA) is a popular classification and dimension-reduction method in ML applications. The LDA method has been used to convert data into a lower-dimensional space to optimize the variance between classes while

MRI Scanner
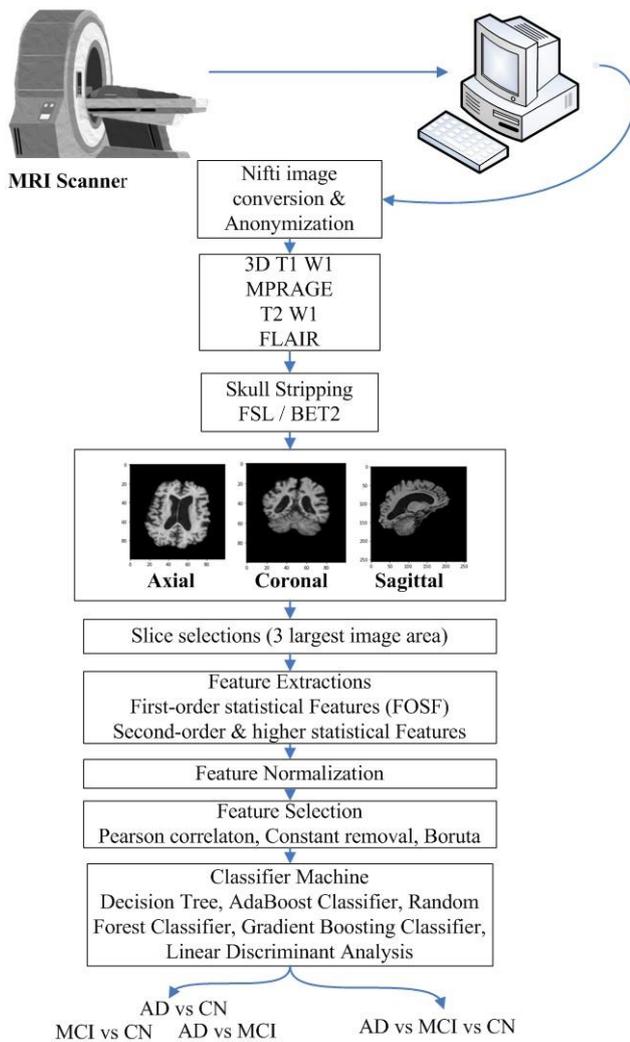
Figure. 5 Diagram of the proposed system

minimizing variation within classes, resulting in maximum class separation. The LDA technique has two types of approaches: class-dependent and class-independent approaches. In the class-dependent approach, each class is given its own lower-dimensional space into which its data is projected. The class-independent approach, on the other hand, considers each class separately. Each class is treated as a distinct entity from the rest of the classes in this approach. [33]. LDA provides predictions by evaluating the probability of new inputs to the probabilities of existing classes. The prediction class is the one with the highest likelihood.

4. Proposed method

4.1 Outline of the proposed method

The four main focuses of our proposed system are preprocessing the images, feature extraction, FS, and AD classification. The proposed method begins with MRI data collection, transformation into three planes (axial, coronal, and sagittal), and removing

non-brain areas using Brain Extraction Tool2 (BET2). Non-brain areas such as the skull and neck voxels must be excluded from MRI imaging since clinical evidence suggests the diseases appear in the brain area only. Our study used the Functional Magnetic Resonance Imaging (FMRIB) Software Library (FSL) (https://fsl.fmrib.ox.ac.uk/) to extract the brain area. The process continued with selecting three slices for each plane and multiple feature extraction. After normalization data, we used Boruta FS to minimize the features, and finally, all of the selected features were fed to the classifier. Fig. 5 shows a more detailed illustration of our proposed system design.

In addition, since the MRI scan image is in 3D format, a comprehensive analysis of all plane orientations is required to obtain a complete image description. Although 3D analysis of MRI images is possible, it requires significant computing power and resources. Multiple images of the image slice are used in 2D-MRI analysis. The proper slice selection method will ensure that the diagnostic system performs optimally. Clinical evidence of disease should be captured more correctly if MRI images with the largest brain area are used. Therefore, we propose taking the three largest slices in each plane using the largest proportion of brain imaging area. We sorted the S values from Eq.(1) to get the largest three slices of each plane.

$$S(i) = \frac{\sum p_i}{\sum (P_i - p_i)} \tag{1}$$

Where,

$S(i)$ = The brain image proportion on a certain slice
i = The slice number (i =1 … s)
p = The number of non-zero pixels
P = Total number of pixels

4.2 Textural feature

The first texture feature we used in this work is FOS. The statistical feature formula for this first feature group is relatively simple since it calculates individual pixel values without considering the neighboring relationship. The features comprise: Mean, Variance, Skewness, Kurtosis, Energy, and Entropy. The calculation of each feature is as follows: First, we need to calculate the image histogram (H_i) and let $f(x,y)$ be a grayscale image.

$$H_i = \frac{number\ of\ pixels\ with\ gray\ level\ i}{total\ number\ of\ pixels} \tag{2}$$

Next, our features in this first group can be derived using the formula below:

$$f_1 = \mu = \sum_i i H_i \qquad (3)$$

$$f_2 = \sigma = \sqrt{\sum_i (i-\mu)^2 H_i} \qquad (4)$$

$$f_3 = \sum_i \left(\frac{i-\mu}{\sigma}\right) H_i \qquad (5)$$

$$f_4 = \sum_i \left(\frac{i-\mu}{\sigma}\right)^4 H_i \qquad (6)$$

$$f_5 = \sum_i H_i^2 \qquad (7)$$

$$f_6 = -\sum_i H_i \ln[H_i] \qquad (8)$$

Where,
f_1 = Mean
f_2 = Variance
f_3 = Skewness
f_4 = Kurtosis
f_5 = Energy and,
f_6 = Entropy

Our second feature group is GLCM-based features, i.e., Angular Second Moment, Contrast, Correlation, Sum of Squares Variance, Inverse Difference Moment (Homogeneity), Sum Average, Sum Variance, Sum Entropy, Entropy, Difference Variance, Difference Entropy, Information Measures of Correlation1, Information Measures of Correlation2, and Maximal Correlation Coefficient.

The GLCM features are obtained from the co-occurrence matrix, which calculates the probability value (P) between two pixels (x_1, y_1) and (x_2, y_2) with the intensity of gray level i and gray level j within a given distance (d) and angle direction [θ]. Meanwhile, the GLRM features that we employ are as follows: Gray Level Non-Uniformity, Run Length Non-Uniformity, Run Percentage Low Gray Level Run Emphasis, High Gray Level Run Emphasis, Short Low Gray Level Emphasis, Short Run High Gray Level Emphasis, and Long-Run High Gray Level Emphasis. Detailed descriptions and formulations for each GLCM and GLRM feature may be found in [19] and [20].

Since the angle and distance parameters used to determine each feature varies, each GLCM and GLRM feature will have many features. We need to minimize the number of features we used, so we use Eqs. (9) and (10) to reduce the number of features on each feature to only two: the average and range values. Let f_i be the ith of a specific feature extracted from GLCM or GLRM, then we calculate:

Average-Feature
$$F_{Average} = \frac{\sum f_i}{total\ number\ of\ f} \qquad (9)$$

Range-Feature
$$F_{range} = f_{max} - f_{min} \qquad (10)$$

The last feature extraction is Energy and Entropy from LBP images, constructed at different scales (R=1,2,3 with the corresponding number of pixels=8,16,24). The LBP texture descriptor is calculated using eq.11. A detailed description of the LBP image descriptor can be found in [21].

$$LBP_{P,R}(x) = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c), if\ U(x) \le 2 \\ P+1,\ otherwise \end{cases} \qquad (11)$$

Where,

$$s(x) = \begin{cases} x, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

P = points are chosen on the circumference of the circle
R = the radius of the circle
g_c = the gray value of the center pixel
g_p = the gray value of the point, $p=0,…,P-1$
U = uniformity (number of spatial bitwise 0/1 transitions)

Following the feature extraction step, the feature data should be normalized to avoid significant variations in the data obtained. Our study used the Standard Scaler approach based on the Standard Normal Distribution (SND). The mean is set to 0, and the data is scaled to the unit variance. We did not use MinMaxScaler normalization because each feature obtained from statistical calculations of first, second, and high order varies significantly. This approach is unsuitable since our features are not derived directly from the RGB pixel intensity, which does not have a fixed range like 0 to 255.
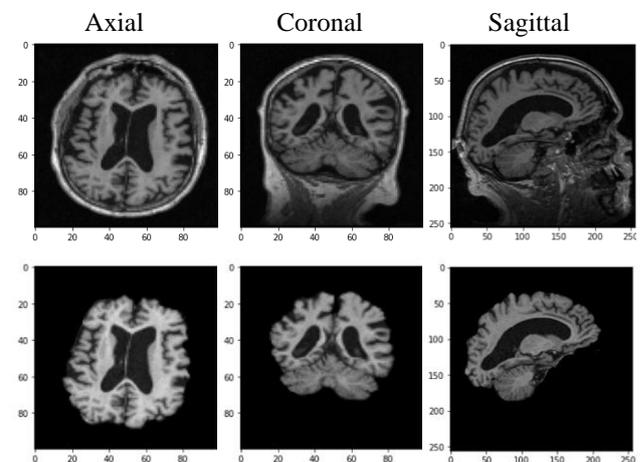
Figure. 6 The images: (top) original MRI images, (bottom) FSL-BET2 MRI images

Table 2. Evaluation scenarios on the proposed system

	Description	The objective
1	Binary classification using multi-feature fusion.	compare basic performance of multi-plane vs. single plane
2	A Boruta feature selection approach accompanies the binary classification.	to see how Boruta-FS affects system performance.
3	Multiclass classification AD vs. MCI vs. CN.	evaluate the system performance when handling the multiclass classification tasks.

5. Results and discussion

The analysis needs to focus only on the brain image area, so anything else should be discarded. After cleaning the non-brain portion of each MRI image, three images were taken using eq.1 for each plane, providing nine slices of image for each MRI image. Then, we adjusted all selected images resolution to 100×100 pixels. Several result images from the preprocessing steps can be seen in Fig. 6. Finally, as indicated in Table 2, we evaluated the proposed method using three different experimental scenarios for different purposes. BET [36] was used to remove non-brain tissue from the whole head image using BET option "B" with f=0.1 [37] to remove non-brain tissue from the complete head image.

5.1 Performance evaluation

The classifier output (predicted results) is compared to the original label to evaluate the classifier's performance (actual results). The accuracy, precision, specificity, sensitivity, and f1-score performance measures were used to examine the performance of our proposed system. Five classifiers were chosen to evaluate the effectiveness of our approach, to demonstrate the superiority of the MP approach. All classifiers were tested using 5-fold cross-validation to assure classification performance, with 20% of the data utilized as test data and 80% as training data for each fold.

$$Accuracy = \frac{True\ Negative + True\ Positive}{Total\ Data} \quad (12)$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (13)$$

$$Sensitivity = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (14)$$

$$Specificity = \frac{True\ Negative}{True\ Negative + False\ Positive} \quad (15)$$

$$F1 = \frac{2x\ Precision\ x\ recall}{Precision + recall} \quad (16)$$

The "True" variable indicates the number of correctly identified subjects to their class, and the "False" means the number of incorrectly identified subjects.

Our GLCM features consist of 28 features comprising 14 mean-features and 14 range-features, with parameter, distance d=1, and $\theta=0^0, 45^0, 90^0, 135^0$. The angle parameters used for GLRM-based features are the same as for GLCM-based features ($0^0, 45^0, 90^0$, and 135^0). Same as GLCM-based features, each feature obtained from GLRM comprises two feature values: the feature-mean and feature-range. Our last set of features is LBP-based features. The detailed LBP features consist of energy and entropy features with various parameters [R=1, P=8], [R=2, P=16], and [R=3, P=24]. Thus, 14 features were extracted from GLRM and LBP, respectively, eight features and six features. As a result, the total number of features obtained became 48 when six features from FOS were added. The 48 features of all observed images are normalized using the SND technique before the data is fed into the classifier. Our first evaluation scenario uses all normalized features as inputs of the five classifiers using two sub-scenarios: the single-plane and the multiple-plane approaches. In this scenario, the feature selection step has been bypassed because the goal of the present study is to evaluate the system's performance using all of the features in our proposed method.

| (a) | (b) | (c) |
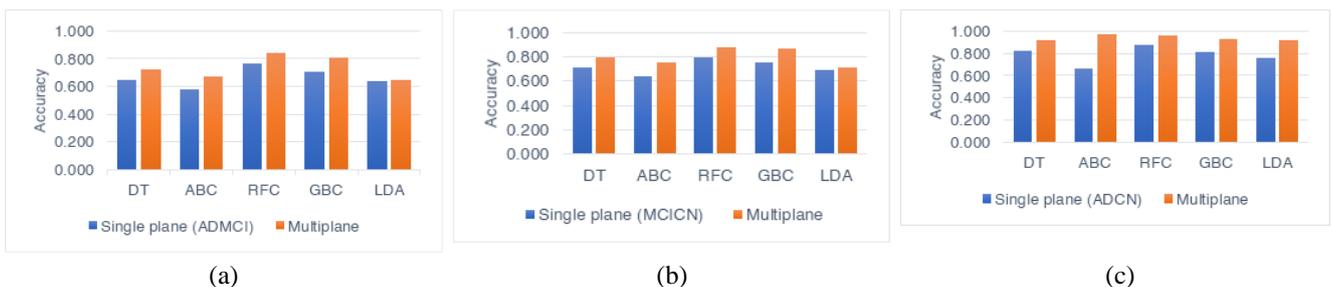
Figure. 7 The binary classification results: (a) AD vs. MCI, (b) MCI vs. CN, and (c) AD vs. CN using 48 fusion features

Table 3. Complete classification results on AD vs. MCI class using five classifiers

Classifier (AD MCI)	Accuracy		Sensitivity		Specificity		Precision		f1-score	
	SP	MP	SP	MP	SP	MP	SP	MP	SP	MP
DT	0.650	0.725	0.600	0.617	0.700	0.833	0.667	0.787	0.632	0.692
ABC	0.575	0.675	0.600	0.617	0.533	0.750	0.569	0.706	0.592	0.649
RFC	0.767	**0.842**	0.800	**0.833**	0.733	**0.850**	0.750	**0.847**	0.774	**0.840**
GBC	0.708	0.808	0.683	0.783	0.733	0.833	0.719	0.825	0.701	0.803
LDA	0.642	0.650	0.633	0.650	0.650	0.750	0.644	0.688	0.638	0.668

SP= Single plane, MP= Multi-plane, DT = Decision Tree, ABC = AdaBoost Classifier, RFC = Random Forest Classifier, GBC = Gradient Boosting Classifier

Table 4. Complete classification results on MCI vs. CN class using five classifiers

Classifier (MCI CN)	Accuracy		Sensitivity		Specificity		Precision		f1-score	
	SP	MP	SP	MP	SP	MP	SP	MP	SP	MP
DT	0.717	0.792	0.650	0.733	0.783	0.850	0.750	0.830	0.696	0.778
ABC	0.642	0.750	0.683	0.783	0.600	0.717	0.631	0.734	0.656	0.758
RFC	0.792	**0.875**	0.800	**0.900**	0.783	**0.850**	0.787	**0.857**	0.793	0.878
GBC	0.750	0.867	0.783	0.883	0.717	0.850	0.734	0.855	0.758	0.869
LDA	0.692	0.708	0.717	0.733	0.650	0.700	0.677	0.705	0.704	0.711

Table 5. Complete classification results on AD vs. CN class using five classifiers

Classifier (AD CN)	Accuracy		Sensitivity		Specificity		Precision		f1-score	
	SP	MP	SP	MP	SP	MP	SP	MP	SP	MP
DT	0.825	0.917	0.800	0.933	0.850	0.900	0.842	0.903	0.820	0.918
ABC	0.667	0.967	0.667	0.967	0.667	0.967	0.667	0.967	0.667	0.967
RFC	0.875	**0.958**	0.833	**0.917**	0.917	**1.000**	0.909	**1.000**	0.869	**0.957**
GBC	0.808	0.933	0.800	0.917	0.817	0.950	0.814	0.948	0.807	0.932
LDA	0.758	0.917	0.767	0.917	0.750	0.917	0.754	0.917	0.760	0.917

Fig. 7 shows the accuracy evaluation of single plane approach and multi-plane approach using 48 features using several classifiers. The result indicates that system accuracy obtained proves that a classifier based on multi-plane feature fusion can outperform single plane performance. In addition, complete classification results of AD vs. MCI, AD vs. CN, and MC vs. CN classes employing a total of 48 features in five classifiers are shown in Table 3, Table 4, and Table 5.
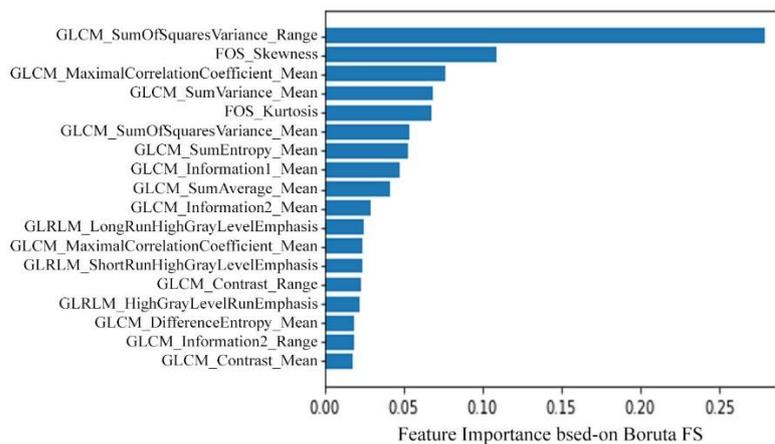
Figure. 8 The importance level of the 18 selected features

Table 6. Improving performance accuracy of five classifiers when employing multi-plane approach

Classifier	Improved Accuracy		
	AD vs MCI	AD vs CN	MCI vs CN
DT	11.54%	11.15%	10.46%
ABC	17.39%	44.98%	16.82%
RFC	9.78%	9.49%	10.48%
GBC	14.12%	15.47%	15.60%
LDA	1.25%	20.98%	2.31%
Average	10.82%	20.41%	11.13%

The classification results using five classifiers indicate that our proposed multi-plane analysis can improve the classifier's performance of the single plane approach. The RFC with a multi-plane approach has the best overall performance with an accuracy of 0.958 and a sensitivity of 0.917 when performing binary classification AD vs. CN.

When single-plane was used as a baseline comparison, the AD vs. CN classification had the most considerable average improvement in accuracy of 20.41 %, as shown in Table 6. The classifier's overall performance has increased as a contribution of the multi-plane analysis approach. The next step is to add a feature selection process to exclude the less essential features while keeping a high system performance even with fewer features. Since system performance is our top priority; the FS should not be employed if it degrades.

5.2 Feature selection evaluation

Several objectives of FS implementation are to eliminate redundant and irrelevant data, improve learning accuracy, reduce computational costs, and improve understanding of observed data.

Our study used Boruta FS to select features extracted from the whole MRI plane. Therefore, an evaluation of the use of FS is required to verify whether Boruta FS is essential in our study**.** Our study reveals that it is unnecessary to use all textural features (FOS, second-order, and higher-order statistical features) to achieve optimal classifier performance. The implementation of Boruta FS resulted in only 18 features being selected from a total of 48. The selected features include two features from FOS, 13-features based on GLCM, and three features from GLRM, but all features from LBP have been discarded. Fig. 8 shows a visualization of the importance level of the 18 features selected by Boruta FS.

Fig. 9 shows that the distribution of feature GLCM_SumOfSquaresVariance_Range, which is the most important feature, has a lower intersection
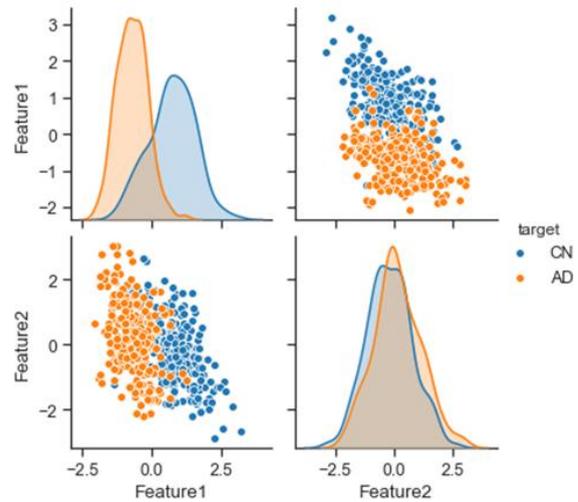
Figure. 9 The distribution of selected features (highest importance vs. the lowest importance)

between classes than the intersection between classes in the distribution of GLCM_Contrast_Mean feature. Thus, Boruta FS successfully demonstrated its ability to select important features while eliminating less important features.

Boruta FS uses RFC to select relevant features by analyzing the degree of system accuracy reduction when a certain feature is removed. Our studies show that the proposed method improves the classifier's performance, especially when the RFC categorizes AD vs. CN. Figure 10 presents a comparison between a single plane and multi-plane accuracy based on 18 features.

In addition, the results show that classifiers based on multiplane feature fusion of selected Boruta FS features can outperform single planes in overall performance. Furthermore, as shown in Table 5, 6, and 7, using a multiplane approach has been proved to outperform the usage of a single plane approach in all classifier performance. The classifier performance on AD vs. CN classification task is the most effective compared to AD vs. MCI and MCI vs. CN is another finding from this studyTable 9 demonstrates that RFC with a multiplane approach has the best classifier among all performances. The RFC with single plane approach accuracy performance of 0.875 improves to 0.917 when the Boruta-FS process is used, while the RFC with multiplane approach accuracy performance of 0.958 improves to 0.967 when the Boruta-FS process is used. The improved performance of the multiplane approach is claimed, with the added benefit of reducing the number of features processed in the classifier for increased resource efficiency. Apart from improving initial performance without FS, the multiplane analysis technique has been shown to be superior to the single plane approach once more by using Boruta FS.

(a)   (b)   (c)
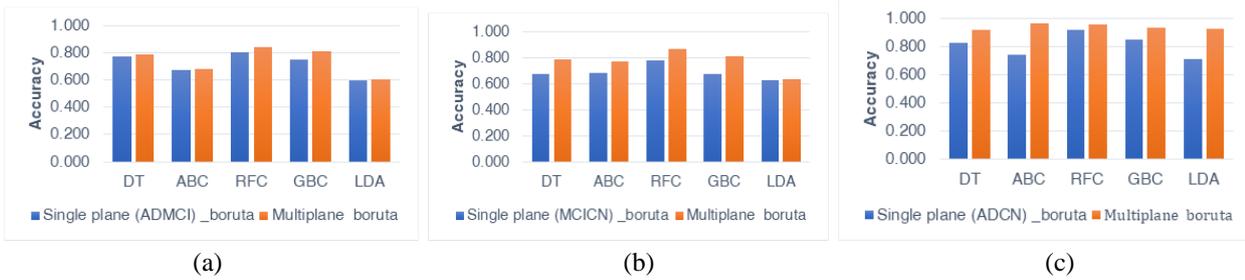
Figure. 10 The binary classification results: (a) AD vs. MCI, (b) MCI vs. CN, and (c) AD vs. CN using 18 fusion features

Table 7. Complete classification results on AD vs. MCI class using Boruta FS

Classifier (AD MCI)	Accuracy		Sensitivity		Specificity		Precision		f1-score	
	SP	MP	SP	MP	SP	MP	SP	MP	SP	MP
DT	0.775	0.792	0.750	0.733	0.800	0.850	0.789	0.830	0.769	0.778
ABC	0.675	0.683	0.600	0.650	0.750	0.717	0.706	0.696	0.649	0.672
RFC	0.800	**0.842**	0.817	**0.833**	0.783	**0.850**	0.790	**0.847**	0.803	**0.840**
GBC	0.750	0.808	0.800	0.783	0.700	0.833	0.727	0.825	0.762	0.803
LDA	0.600	0.608	0.533	0.550	0.667	0.667	0.615	0.623	0.571	0.584

Table 8. Complete classification results on MCI vs. CN class using Boruta FS

Classifier (MCI CN)	Accuracy		Sensitivity		Specificity		Precision		f1-score	
	SP	MP	SP	MP	SP	MP	SP	MP	SP	MP
DT	0.675	0.783	0.750	0.833	0.600	0.733	0.652	0.758	0.698	0.794
ABC	0.683	0.767	0.650	0.845	0.717	0.683	0.696	0.729	0.672	0.785
RFC	0.775	**0.867**	0.783	**0.850**	0.767	**0.883**	0.770	**0.879**	0.776	**0.864**
GBC	0.675	0.808	0.717	0.833	0.633	0.783	0.662	0.794	0.688	0.813
LDA	0.625	0.633	0.700	0.667	0.550	0.600	0.609	0.625	0.651	0.645

Table 9. Complete classification results on AD vs. CN class using Boruta FS

Classifier (AD CN)	Accuracy		Sensitivity		Specificity		Precision		f1-score	
	SP	MP	SP	MP	SP	MP	SP	MP	SP	MP
DT	0.825	0.917	0.817	0.900	0.833	0.933	0.831	0.931	0.824	0.915
ABC	0.742	0.958	0.683	0.917	0.800	0.967	0.774	0.967	0.726	0.941
RFC	0.917	**0.967**	0.883	**0.967**	0.950	**1.000**	0.946	**1.000**	0.913	**0.983**
GBC	0.850	0.933	0.783	0.917	0.917	0.950	0.904	0.948	0.839	0.932
LDA	0.708	0.925	0.650	0.917	0.767	0.933	0.736	0.932	0.690	0.924

5.3 Performance evaluation on multiclass classification

The task of multiclass categorization is a challenging problem, especially when compared to the binary classifier's performance. Multiclass classification tasks are generally more computationally expensive than binary classification tasks, but we need to improve both. classification of AD vs. MCI vs. CN. Our experiments showed that when using the single plane technique, the best accuracy is 0.767, and when using the multi-plane analysis approach, the highest accuracy is 0.867. The results show that all of the classifiers investigated in this multiclass classification task have improved their accuracy performance. ABC achieved the highest accuracy improved performance throughout this multiclass classification task, about 82.02%, jumping from 0.406 to 0.739.
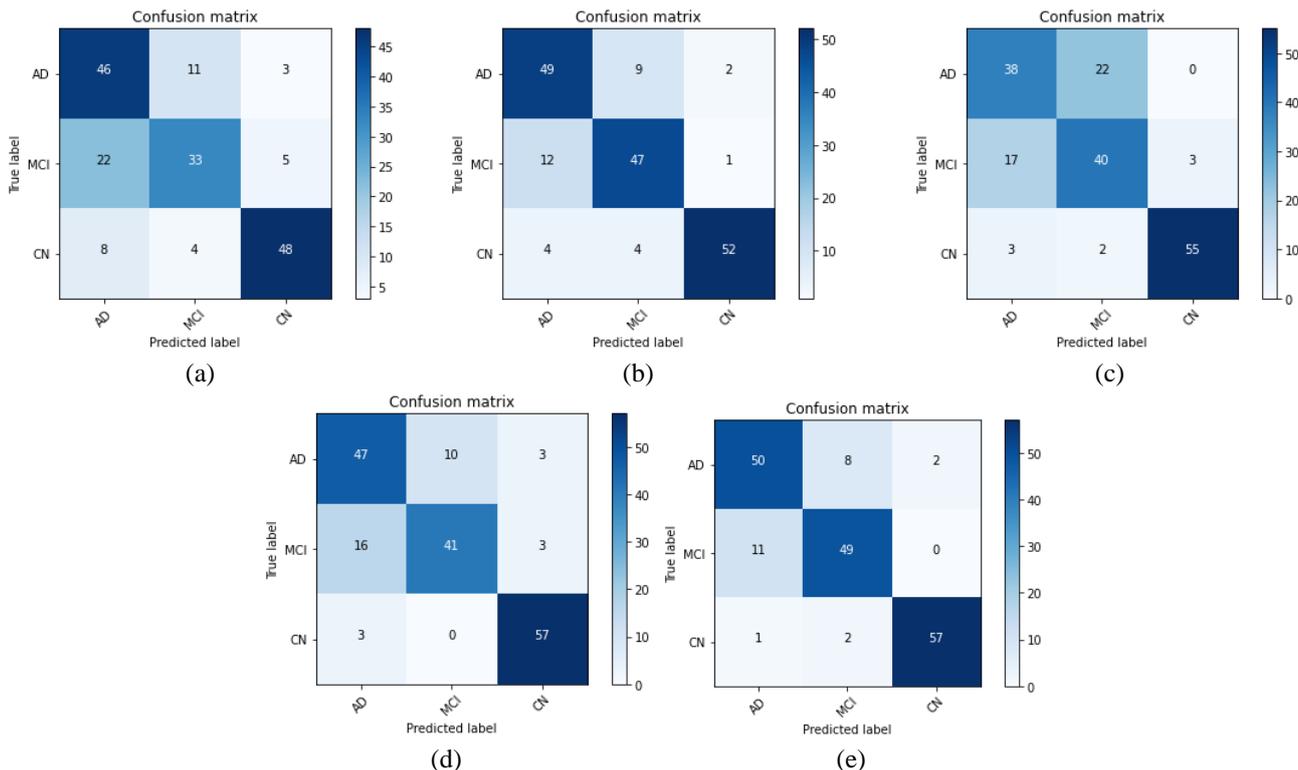
Figure. 11 The confusion matrix AD vs. MCI vs. CN of classification results using Boruta FS on five classifiers; (a) Decision Tree, (b) AdaBoost Classifier, (c) Random Forest Classifier, (d) Gradient Boosting Classifier, and (e) Linear Discriminant Analysis

Table 10. Classification results on multiclass AD vs. MCI vs. CN using Boruta FS

Model		Acc	Precision			Sensitivity			F1-score		
			AD	MCI	CN	AD	MCI	CN	AD	MCI	CN
SP	DT	0.633	0.705	0.630	0.569	0.717	0.567	0.617	0.711	0.596	0.592
	ABC	0.406	0.536	0.310	0.396	0.500	0.367	0.350	0.517	0.336	0.372
	RFC	0.767	0.797	0.792	0.714	0.817	0.633	0.750	0.853	0.704	0.732
	GBC	0.606	0.636	0.576	0.600	0.700	0.567	0.550	0.667	0.571	0.574
	LDA	0.550	0.559	0.492	0.604	0.633	0.483	0.533	0.594	0.487	0.566
MP	DT	0.706	0.605	0.688	0.857	0.767	0.550	0.800	0.676	0.611	0.828
	ABC	0.739	0.655	0.625	0.948	0.633	0.667	0.917	0.644	0.645	0.932
	RFC	**0.867**	**0.806**	**0.831**	**0.966**	**0.833**	**0.817**	**0.950**	**0.820**	**0.824**	**0.958**
	GBC	0.822	0.754	0.783	0.945	0.817	0.783	0.867	0.784	0.783	0.904
	LDA	0.806	0.712	0.804	0.905	0.783	0.683	0.950	0.746	0.739	0.927

The confusion matrix was applied to quantify method performance in detail, which effectively deals with multiclass classification tasks. It gives a clear overview of how well the proposed method identified the classes based on the supplied data. This tool will also make it easier to observe the data on classes that have been misclassified. Fig. 11 shows the confusion matrix obtained from the classification results. CN class is most accurately predicted according to the five confusion matrices, followed by AD class, and finally, MCI class. The results can be explained by the fact that

AD and CN classes have vastly different physical and anatomical characteristics in clinical conditions, unlike MCI, a transitional class between CN and AD. MCI patients have clinically suffered brain damage, and if they do not receive proper treatment, they may develop AD.

5.4 Comparison with previous work

This section explores previous studies on AD classifications and discusses our proposed method,

Table 11. The proposed method and prior comparable studies

Study	Dataset	Modality	Classification Type	Method	Classifier	Accuracy
Tong et al. (2017) [12]	ADNI	MRI, FDG-PET CSF, Gene	AD vs. MCI vs. CN	SP	NGF	0.602
Zhang et al. (2017) [10]	ADNI	ROI MRI	AD vs. CN	SP	SVM	0.883
Altaf et al. (2018) [11]	ADNI	MRI + clinical	AD vs. MCI vs. CN	SP	SVM	0.798
Zeng et al. (2018) [13]	ADNI	MRI	AD vs. CN	SP	SDPSO-SVM-PCA	81.25
Peng et al. (2019) [15]	ADNI	MRI + FDG-PET + SNP	AD vs. CN	SP	SVM	0.961
Vaithinathan et al. (2019) [14]	ADNI	MRI (GM,WM)	AD vs. CN	SP	SVM + RFC +KNN	87.39
Jimenez-Mesa et al. (2020) [16]	ADNI	MRI	AD vs. MCI vs. cMCI vs. CN	SP	SVM	0.670
Our work (2022)	ADNI	MRI	AD vs. CN	MP	RFC	0.967
		MRI	AD vs. MCI	MP	RFC	0.842
		MRI	MCI vs. CN	MP	RFC	0.867
		MRI	AD vs. MCI vs. CN	MP	RFC	0.867
	AUH	MRI	BA vs. Normal	MP	RFC	0.942

which is more effective for distinguishing AD, MCI, and CN. Table 11 compares the proposed method to several past studies, showing that it is more precise since it has higher accuracy than the others. In addition, our proposed method is also effective in identifying BA patients in the dataset from AUH.

There has not been much study on the multiclass classification of AD vs. MCI vs. CN, compared to the binary classification of AD vs. CN, where the classifier is shown to be better at handling this task. Therefore, the comparative study table does not display all of their results, but only the best of them, i.e., AD vs. CN binary classification, if they do not perform the multiclass AD vs. MCI vs. CN classification

The binary classification of AD vs. CN [10, 13-15] was shown to have the best accuracy of 0.961 by Peng et al. [15]. However, these results were acquired using multi-modality, i.e., MRI, FDG-PET, and SNP, so technically, it would require more effort to obtain these varied data from each diagnosed patient. Meanwhile, the multiclass classification of AD vs. MCI vs. CN in [11, 12, 16] revealed that Altaf [11], with an accuracy of 0.798, had the highest accuracy.

We compared single-plane and multi-plane approaches, demonstrating that our proposed method outperforms all investigated single-plane approaches. In addition to the binary classification task, the multiclass classification results showed that our

proposed method is superior. The additional comparison data, particularly the BA vs. Normal classification task on the AUH dataset, reveals that our proposed method performance is comparable to the AD vs. CN classification, 0.942.

6. Conclusions

Our study presents a novel method with a multi-plane approach based on MRI images using multiple texture-based features derived from multiple statistical order calculations, referred to as a multi-feature fusion technique for diagnosing AD. The number of features can be reduced by using Boruta FS, resulting in only 18 of the 48 possible features being processed. In addition to the binary classification of AD vs. CN, MCI vs. Our CN, and AD vs. MCI, we also evaluated our proposed method for the multiclass classification of AD vs. MCI vs. CN.

The following are some of the most important findings from our study by involving all three planes: axial, coronal, and sagittal. The single plane technique, which focuses on the axial plane, can be improved by utilizing the entire available plane. We also presented automatic slice selection on three orthogonal planes, eliminating the requirement to process the full slice in the classification. The results show that using all of the extracted features (multi-feature fusion), our multi-plane classification system

could enhance the performance of a single-plane system.

The average performance improvement for all classifiers is 10.82% for AD vs. MCI, 20.41% for AD vs. CN, and 11.13% for MCI vs. CN. Second, the results of FS based on Boruta FS can improve the classifier's performance, allowing for a reduction in the number of features processed and, consequently, a reduction in computational cost. According to the evaluation results, our new approach outperformed all previous methods in the related work compared in this study. In the previous study results, the highest accuracy scores on the AD vs. CN task were 0.961, while the highest scores on the multiclass classification task were 0.798; however, in our work, we received higher results of 0.967 on AD vs. CN and 0.867 on multiclass classification tasks. Thus, our new approach has proven to be effective in both binary and multiclass classification tasks for detecting Alzheimer's disease.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization, I Ketut Eddy Purnama, Mauridhi H. Purnomo; methodology, Cucun Very Angkoso; software, Cucun Very Angkoso; validation, I Ketut Eddy Purnama; formal analysis, Cucun Very Angkoso, I Ketut Eddy Purnama and Mauridhi H. Purnomo; investigation, Cucun Very Angkoso; resources, Cucun Very Angkoso, Yudhi Adrianto and, Anggraini Dwi Sensusiati; data curation, Hapsari Peni Agustin Tjahyaningtijas, Yudhi Adrianto and, Anggraini Dwi Sensusiati; writing original draft preparation, Cucun Very Angkoso, I Ketut Eddy Purnama, and Hapsari Peni Agustin Tjahyaningtijas; writing—review and editing, Cucun Very Angkoso and Hapsari Peni Agustin Tjahyaningtijas; visualization, Cucun Very Angkoso; supervision, I Ketut Eddy Purnama, Mauridhi H. Purnomo; project administration, I Ketut Eddy Purnama and Mauridhi H. Purnomo; funding acquisition, I Ketut Eddy Purnama and Mauridhi H. Purnomo.

Acknowledgments

References

[1] J. Liu, M. Li, W. Lan, F. X. Wu, Y. Pan, and J. Wang, "Classification of Alzheimer's Disease Using Whole Brain Hierarchical Network", *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, Vol. 15, No. 2, pp. 624-632, 2018.

[2] T. Glozman, J. Solomon, F. Pestilli, and L. Guibas, "Shape-attributes of brain structures as biomarkers for Alzheimer's disease", *Journal of Alzheimer's Disease*, Vol. 56, No. 1, pp. 287-295, 2017.

[3] A. Ward, S. Tardiff, C. Dye, and H. M. Arrighi, "Rate of Conversion from Prodromal Alzheimer's Disease to Alzheimer's Dementia: A Systematic Review of the Literature", *Dementia and Geriatric Cognitive Disorders Extra*, Vol. 3, No. 1, pp. 320-332, 2013.

[4] L. Sørensen, C. Igel, A. Pai, I. Balas, and C. Anker, "Differential diagnosis of mild cognitive impairment and Alzheimer' s disease using structural MRI cortical thickness , hippocampal shape, hippocampal texture, and Volumetry", *NeuroImage: Clinical*, Vol. 13, pp. 470-482, 2017.

[5] M. Ruthirakuhan, K. L. Lanctôt, M. D. Scipio, M. Ahmed, and N. Herrmann, "Biomarkers of agitation and aggression in Alzheimer's disease: A systematic review", *Alzheimer's and Dementia*, Vol. 14, No. 10, pp. 1344-1376, 2018.

[6] G. M. Juan, G. S. Guell, and G. Piella, "A survey on machine and statistical learning for longitudinal analysis of neuroimaging data in Alzheimer's disease", *Computer Methods and Programs in Biomedicine*, Vol. 189, 2020.

[7] J. P. Lerch, J. Pruessner, A. P. Zijdenbos, D. L. Collins, S. J. Teipel, H. Hampel, and A. C. Evans, "Automated cortical thickness measurements from MRI can accurately separate Alzheimer's patients from normal elderly controls", *Neurobiology of Aging*, Vol. 29, No. 1, pp. 23-30, 2008.

[8] A. H. Syaifullah, A. Shiino, H. Kitahara, R. Ito, M. Ishida, and K. Tanigaki, "Machine Learning for Diagnosis of AD and Prediction of MCI Progression From Brain MRI Using Brain Anatomical Analysis Using Diffeomorphic Deformation", *Frontiers in Neurology*, Vol. 11, No. February, pp. 1-13, 2021.

[9] J. Islam and Y. Zhang, "Brain MRI analysis for Alzheimer's disease diagnosis using an

ensemble system of deep conVolutional neural networks", *Brain Informatics*, Vol. 5, No. 2, 2018.

[10] J. Zhang, M. Liu, L. An, Y. Gao, and D. Shen, "Alzheimer's disease diagnosis using landmark-based features from longitudinal structural MR images", *IEEE Journal of Biomedical and Health Informatics*, Vol. 21, No. 6, pp. 1607-1616, 2017.

[11] T. Altaf, S. Muhammad, N. Gul, M. Nadeem, and M. Majid, "Multi-class Alzheimer' s disease classification using image and clinical features", *Biomedical Signal Processing and Control*, Vol. 43, pp. 64-74, 2018.

[12] T. Tong, K. Gray, Q. Gao, L. Chen, and D. Rueckert, "Multi-modal classification of Alzheimer's disease using nonlinear graph fusion", *Pattern Recognition*, Vol. 63, pp. 171-181, 2017.

[13] N. Zeng, H. Qiu, Z. Wang, W. Liu, H. Zhang, and Y. Li, "A new switching-delayed-PSO-based optimized SVM algorithm for diagnosis of Alzheimer's disease", *Neurocomputing*, Vol. 320, pp. 195-202, 2018.

[14] K. Vaithinathan and L. Parthiban, "A Novel Texture Extraction Technique with T1 Weighted MRI for the Classification of Alzheimer's Disease", *Journal of Neuroscience Methods*, Vol. 318, pp. 84-99, 2019.

[15] J. Peng, X. Zhu, Y. Wang, L. An, and D. Shen, "Structured sparsity regularized multiple kernel learning for Alzheimer's disease diagnosis", *Pattern Recognition*, Vol. 88, pp. 370-382, 2019.

[16] C. J. Mesa, I. A. Illan, A. M. Martin, D. C. Barnes, F. J. M. Murcia, J. Ramirez, and J. M. Gorriz, "Optimized one vs one approach in multiclass classification for early alzheimer's disease and mild cognitive impairment diagnosis", *IEEE Access*, Vol. 8, pp. 96981-96993, 2020.

[17] C. V. Angkoso, H. P. A. Tjahyaningtijas, M. H. Purnomo, and I. K. E. Purnama, "Multiplane ConVolutional Neural Network (Mp-CNN) for Alzheimer's Disease Classification", *International Journal of Intelligent Engineering and Systems*, Vol. 15, No. 1, pp. 329-340, 2022, doi: 10.22266/ijies2022.0228.30.

[18] M. K. Ghalati, A. Nunes, H. Ferreira, P. Serranho, and R. Bernardes, "Texture Analysis and its Applications in Biomedical Imaging: A Survey", *IEEE Reviews in Biomedical Engineering*, Vol. 15, pp. 222–246, 2022.

[19] R. M. Haralick, I. Dinstein, and K. Shanmugam, "Textural Features for Image Classification", *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-3, No. 6, pp. 610-621, 1973.

[20] M. M. Galloway, "Texture analysis using gray level run lengths", *Computer Graphics and Image Processing*, Vol. 4, No. 2, pp. 172-179, 1975.

[21] D. C. He and L. Wang, "Texture Unit, Texture Spectrum, and Texture Analysis", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 28, No. 4, pp. 509-512, 1990.

[22] R. W. Conners and C. A. Harlow, "A theoretical comparison of texture algorithms", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-2, No. 3, pp. 204-222, 1980.

[23] I. Beheshti and H. Demirel, "Feature-ranking-based Alzheimer's disease classification from structural MRI", *Magnetic Resonance Imaging*, Vol. 34, No. 3, pp. 252-263, 2016.

[24] M. B. Kursa and W. R. Rudnicki, "Feature selection with the boruta package", *Journal of Statistical Software*, Vol. 36, No. 11, pp. 1-13, 2010.

[25] M. J. Hasan, J. Kim, C. H. Kim, and J. M. Kim, "Health state classification of a spherical tank using a hybrid bag of features and K-Nearest neighbor", *Applied Sciences (Switzerland)*, Vol. 10, No. 7, 2020.

[26] S. Minhas, A. Khanum, F. Riaz, A. Alvi, and S. A. Khan, "A Nonparametric Approach for Mild Cognitive Impairment to AD Conversion Prediction: Results on Longitudinal Data", *IEEE Journal of Biomedical and Health Informatics*, Vol. 21, No. 5, pp. 1403-1410, 2017.

[27] K. P. M. Niyas and P. Thiyagarajan, "Feature selection using efficient fusion of Fisher Score and greedy searching for Alzheimer's classification", *Journal of King Saud University - Computer and Information Sciences*, 2021.

[28] E. E. Tripoliti, D. I. Fotiadis, and M. Argyropoulou, "A supervised method to assist the diagnosis and monitor progression of Alzheimer's disease using data from an fMRI experiment", *Artificial Intelligence in Medicine*, Vol. 53, No. 1, pp. 35-45, 2011.

[29] R. B. Mofrad, N. S. M. Schoonenboom, B. M. Tijms, P. Scheltens, P. J. Visser, W. M. V. D. Flier, and C. E. Teunissen, "Decision tree supports the interpretation of CSF biomarkers in Alzheimer's disease", *Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring*, Vol. 11, pp. 1-9, 2019.

[30] S. Saravanakumar and P. Thangaraj, "A Computer Aided Diagnosis System for Identifying Alzheimer's from MRI Scan using

Improved Adaboost", *Journal of Medical Systems*, Vol. 43, No. 3, 2019.

[31] S. I. Dimitriadis, D. Liparas, and M. N. Tsolaki, "Random forest feature selection, fusion and ensemble strategy: Combining multiple morphological MRI measures to discriminate among healhy elderly, MCI, cMCI and alzheimer's disease patients: From the alzheimer's disease neuroimaging initiative (ADNI) data", *Journal of Neuroscience Methods*, Vol. 302, pp. 14-23, 2018.

[32] P. Lodha, A. Talele, and K. Degaonkar, "Diagnosis of Alzheimer's Disease Using Machine Learning", 2018.

[33] A. Tharwat, T. Gaber, A. Ibrahim, and A. E. Hassanien, "Linear discriminant analysis: A detailed tutorial", *AI Communications*, Vol. 30, No. 2, pp. 169-190, 2017.