A Novel Method for Diagnosing Alzheimer's Disease from MRI Scans using the ResNet50 Feature Extractor and the SVM Classifier

Farhana Islam[1], Md. Habibur Rahman[2], Nurjahan[3], Md. Selim Hossain[4], Samsuddin Ahmed[5]
Department of Educational Technology, Bangabandhu Sheikh Mujibur Rahman Digital University[1]
Department of Internet of Things and Robotics Engineering, Bangabandhu Sheikh Mujibur Rahman Digital University[2,3,5]
Department of Electronics and Communication Engineering, Hajee Mohammad Danesh Science and Technology University[4]
Bangladesh

***Abstract*—Alzheimer's disease (AD), a chronic neurodegenerative brain disorder, caused by the accumulation of abnormal proteins called amyloid, is one of the prominent causes of mortality worldwide. Since there is a scarcity of experienced neurologists, manual diagnosis of AD is very time-consuming and error-prone. Hence, automatic diagnosis of AD draws significant attention nowadays. Machine learning (ML) algorithms such as deep learning are widely used to support early diagnosis of AD from magnetic resonance imaging (MRI). However, they provide better accuracy in binary classification, which is not the case with multi-class classification. On the other hand, AD consists of a number of early stages, and accurate detection of them is necessary. Hence, this research focuses on how to support the multi-stage classification of AD particularly in its early stage. After the MRI scans have been preprocessed (through median filtering and watershed segmentation), benchmark pre-trained convolutional neural network (CNN) models (AlexNet, VGG16, VGG19, ResNet18, ResNet50) carry out automatic feature extraction. Then, principal component analysis is used to optimize features. Conventional machine learning classifiers (Decision Tree, K-Nearest Neighbors, Support Vector Machine, Linear Programming Boost, and Total Boost) are deployed using the optimized features for staging AD. We have exploited the Alzheimer's disease Neuroimaging Initiative(ADNI) data set consisting of AD, MCIs (MCI), and cognitive normal (CN) classes of images. In our experiment, the SVM classifier performed better with the extracted ResNet50 features, achieving multi-class classification accuracy of 99.78% during training, 99.52% during validation, and 98.71% during testing. Our approach is distinctive because it combines the advantages of deep feature extractors, conventional classifiers, and feature optimization.**

Keywords*—*Alzheimer's disease; brain images; machine learning; deep learning; brain disorder; ADNI dataset

I. Introduction

The neurological illness known as Alzheimer's disease(AD) affects the central nervous system and gradually worsens memory and cognitive function over time [1], [2]. Eventually causing the affected person to lose the ability to learn new information and to retain previously learned information [3] which severely impedes people's daily lives such as failing to recognize the family members and performing essential daily activities leaving the patients with anxiety, aggressiveness, or childish behavior [4]–[6]. Studies [7]–[11] shown that the neurological deterioration of this disease includes the accumulation of abnormal beta-amyloid proteins and phosphorylated tau resulting in depreciation of the hippocampus and cerebral cortex while expanding the ventricles that leads affecting brain regions involved in remembering, thinking, planning, and decision-making.

Usually, AD symptoms appear after the age of 60 with rare exceptions that emerge relatively early at the age of 30 to 50 years in individuals with gene mutation [12]. However, the transition from a healthy state to AD takes several years [13] while going through three different stages, namely, normal controlled (NC), mild cognitive impairment (MCI), and AD. Among the three stages of Alzheimer's, MCI is the symptomatic stage, progressing to its most severe form over time. Since it leads a patient to experience a set of symptoms [14] it incurs huge costs for their proper care and treatment [15]. Therefore, early detection of the disease is essential for initiating treatments, minimizing brain cell damage, and enhancing the quality of life of affected individuals and their families

In the conventional diagnostic system, Alzheimer's patients can be diagnosed the late stages of the disease's progression. In the early stages, the symptoms are similar to those of normal aging. Also, in the conventional system, it is difficult to determine the stages of the disease which may prevent the patient from starting treatment earlier. Besides this, the conventional diagnostic system is limited by the availability of expert physicians and medical tools.

There are studies for automating the diagnosis of this disease. Conventional machine learning and deep learning-based approaches are proposed [16] to classify AD and their stages from different modalities of data. These Machine learning techniques specifically, deep learning techniques are gaining success in the early diagnosis of AD from magnetic resonance imaging (MRI) modality having better accuracy in binary classification while suffering in multiclass classification [2], [17]–[22]. Conventional machine learning leverages handcrafted features while deep learning methods automatically extract features in regression and classification tasks. Studies have shown that the use of conventional machine learning and deep learning techniques combines the strengths of each to create a more accurate and reliable diagnostic tool [3].

Deep Learning models combined with MRI data can give a high degree of diagnostic accuracy of age-related cognitive decline (ARCD) in dementia patients [4], [21]. It has been argued that deep learning approaches produce the sufficient information necessary to correlate AD sample data [13]. Deep learning enables the characterization of AD in MRI images by generating computational models with multiple processing layers. It automatically retrieves its necessary information from input images, without the intervention of the expert who labels the information, as in a standard Machine Learning model [23]. Besides the conventional machine learning models demonstrated state-of-the-art performance in classification and regression tasks if the feature is provided. Considering the classification performance of conventional machine learning models and the automatic feature extraction capacity of deep learning models, specifically CNN, we utilized the strength of both approaches in our study to get better performance in multi-class classification. In this work, we have selected structured MRI (sMRI) data rather than multimodal or other single modal data considering the benefits mentioned in [21]. The data were collected from Alzheimer's Diseases Neuroimaging Initiatives (ADNI) database (adni.loni.usc.edu). Here, a robust and efficient machine learning model has been proposed for analyzing brain MRI images. There are five main phases in this work: (a) MRI Preprocessing (b) Region clustering (c) feature extraction (d) feature optimization and (e) classification of AD into one of its three stages. At first, preprocessing was performed. Preprocessing was necessary to alleviate the problem of low contrast and enhance image quality. The preprocessing tasks include skull removal, intensity normalization so that the mean is zero and variance is one, and image enhancement with histogram equalizations, and mean and median filtering techniques. For region clustering, we have experimented with otsu, edge-based clustering, k-means, region growing, morphology-based clustering, and fuzzy C-means algorithms and found the watershed algorithm suitable. From the clustered images we have selected 64 three-view patches of size 128 by 128 for further analysis.

To alleviate the problem of low contrast and enhance image quality watershed algorithm has been applied to the MRI image. For clustering, a region-based clustering technique that performs better than other state of art techniques has been chosen. The clustered image is further processed to extract features through the use of multiple deep-learning techniques. The principal component analysis was performed to find fine-tuned optimized features. Finally, these features are then input into a machine learning algorithm to classify the disease into its three major AD phases. The main contribution of our work are: 1) Combining the strength of both conventional and deep machine learning techniques for achieving better accuracy in multi-class classification of AD stages. 2) Improved performance with single modality structural MRI (sMRI) analysis without computing the whole brain. 3) Addressing dataset inconsistency and enhancing contrast quality and visibility through the use of contrast amplification techniques. 4) Selection of region clustering technique to find uniform samples for feature extraction that exhibits improved performance compared to conventional techniques.

The paper is organized as follows: Section II introduces the materials and methods including chosen dataset. Section III includes result analysis. Section IV incorporates the related works and discussion. Finally, the conclusion is drawn in Section V.

II. Materials and Methods

The workflow for the proposed framework of Alzheimer's detection mechanism has been divided into several steps such as data collection, data preprocessing, region clustering, feature extraction, feature optimization, classification, and evaluation presented in Fig. 1.

First, the brain MR images have been collected from ADNI. The collected images are then preprocessed through several preprocessing techniques such as intensity normalization, image resizing, contrast enhancement techniques, etc. After completing the pre-processing step, the region clustering algorithms such as C-means, threshold-based otsu clustering, K-means, morphology-based, edge-based, watershed, region-growing, and k-means cluster-based methods have been applied to find out the distinct region for analysis.

Several deep learning techniques such as VGG16, VGG19, Alexnet, Resnet18, and Resnet50 have been applied to extract features from the three view samples selected from clustered images. Then features are optimized by using principal component analysis. Finally, the extracted images are then fed into five different ML techniques such as ensemble-based LP-Boost and TotalBoost, tree-based decision tree (DT), distance-based k-nearest neighbor (KNN), and Support Vector Machine (SVM) methods for the classification into three different stages of Alzheimer's.

A. Dataset

In this study, a subset of the ADNI database has been considered for the experiment. The database was established in 2004 as a result of a public-private partnership with the collaboration of Dr. Michael W. Weiner. The objective of the ADNI dataset was to find the MRI, PET, clinical and neuropsychological assessments, and another biological marker behind the development of MCI and AD. The dataset comprises of 2042 brain MR images representing three different stages of AD such as AD, CN, and MCI. The details of the data are provided in the Table I.

TABLE I. Demographic Information of the ADNI1:Complete 2Yr 1.5T Dataset

Class Label	Nunmber of Scan	Male Subject	Female Subject	Age (Avg. +-std.)
CN	567	271	296	75.12+-8.10
MCI	1206	797	409	76.81+-5.51
AD	269	137	132	75.73+-7.17

The data imbalance problems were avoided by duplicating the MRIs. As we have sampled three view patches from segmented regions to ensure the representation of each significant region the repeated MRIs do not bias the model performance. We have considered total of 1546 MRIs for the experiment (CN-470, MCI-477, AD-599).

B. Data Preprocessing

In our work, at first, we removed the skull from the MRI images. Then we performed intensity normalization so that the mean intensity is zero while keeping the intensity

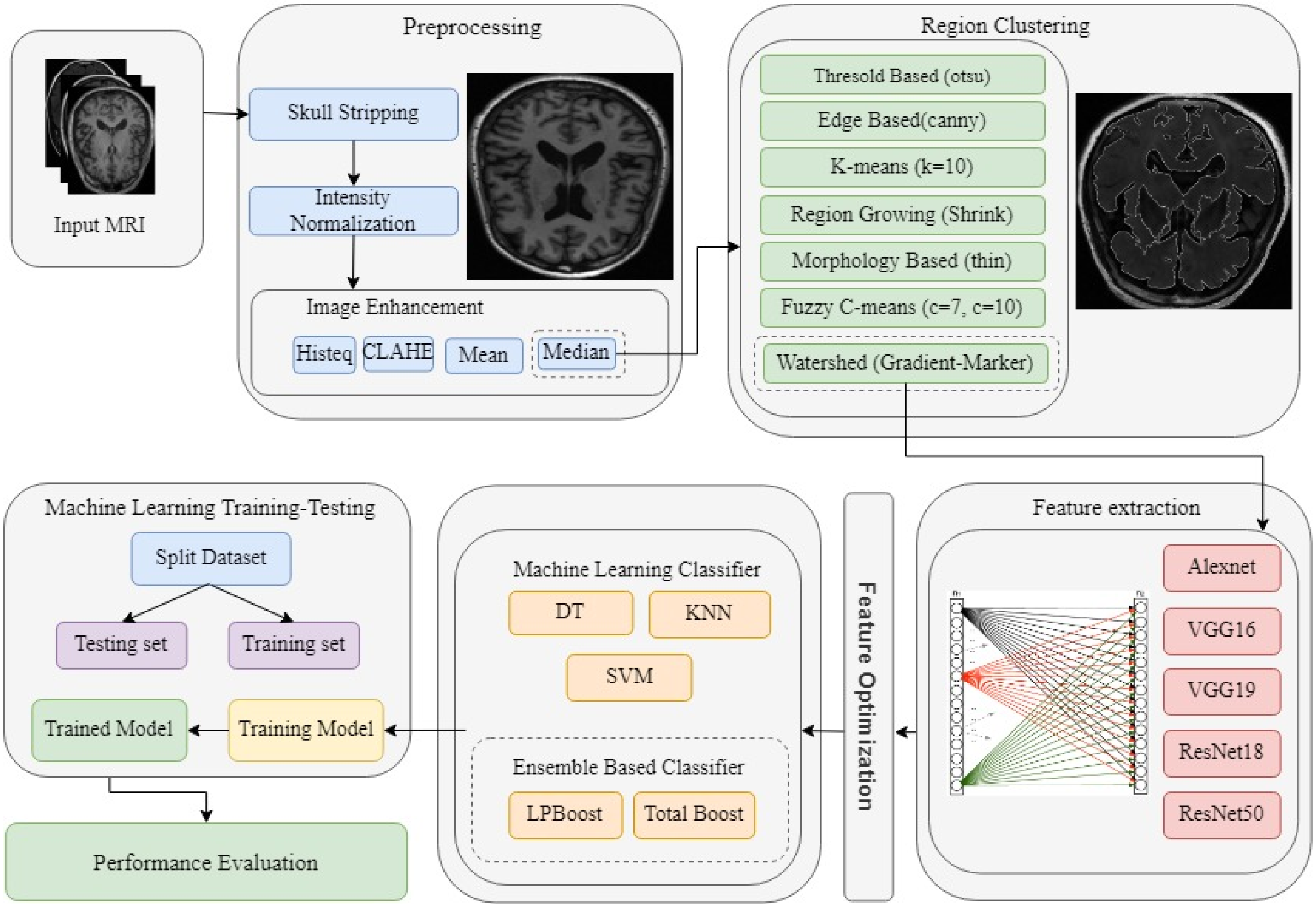

Fig. 1. Conceptual flow of the proposed model.

variance one. We used several pre-processing techniques for contrast enhancement like histogram equalization, contrast limiting adaptive histogram equalization (CLAHE), mean, and median filtering techniques. These techniques are widely used preprocessing methods for medical imaging [24], [25]. The effects of these techniques have been depicted in Fig. 2. Table II represents the comparison of the performance of the preprocessing techniques in terms of mean structural similarity (MSSIM), peak signal-to-noise Ratio (PSNR), and root mean square error (RMSE). It is found that the Median filter outperforms other techniques.

TABLE II. Comparison of Various Preprocessing Techniques

Preprocessing Technique	MSSIM	PSNR	RMSE
Intensity Transformation	0.9940	12.6433	0.2333
Histogram Equalization	0.9386	3.1293	0.6975
Contrast Limited Adaptive Histogram Equalization (CLAHE)	0.9856	9.1310	0.3495
Mean Filter (3 by 3)	1	32.7397	0.0231
Median Filter (3 by 3)	1	37.5815	0.0132

C. Region Clustering

In this work, we have applied several region clustering algorithms such as Threshold Based OTSU methods, Edge Based CANNY filter, region-based region-grow method, Morphological Based THIN filter, K-means Clustering (k=4), Fuzzy Based C-means Clustering (c=4), Watershed with sobel filter considering their wide acceptance in medical imaging [26], [27]. To choose the appropriate method for our system we have calculated the evaluation metrics PSNR, SSIM, and RMSE of these clustering algorithms. In Fig. 3 different output images after using various clustering techniques have been represented. It has been proclaimed here that the Watershed-based clustering technique provides a better image than other techniques. The performance of image enhancement techniques is measured based on evaluation metrics PSNR, MSSIM, and RMSE scores. Table III represents the comparison of different pre-processing techniques. Based on the experimental result it has been found that the watershed algorithm outperforms other algorithms. Here the highest value of MSSIM and PSNR as well as the lowest value of RMSE has been considered to select the method for the system.

D. Sample three View Patch and Feature Extraction

From the segmented images, we have sampled three view patches as inspired from [2], [21], [22] for further analysis. From each segment of an MRI, we have generated 16 uniformly random three-view patches of size 128 by 128 by 3.

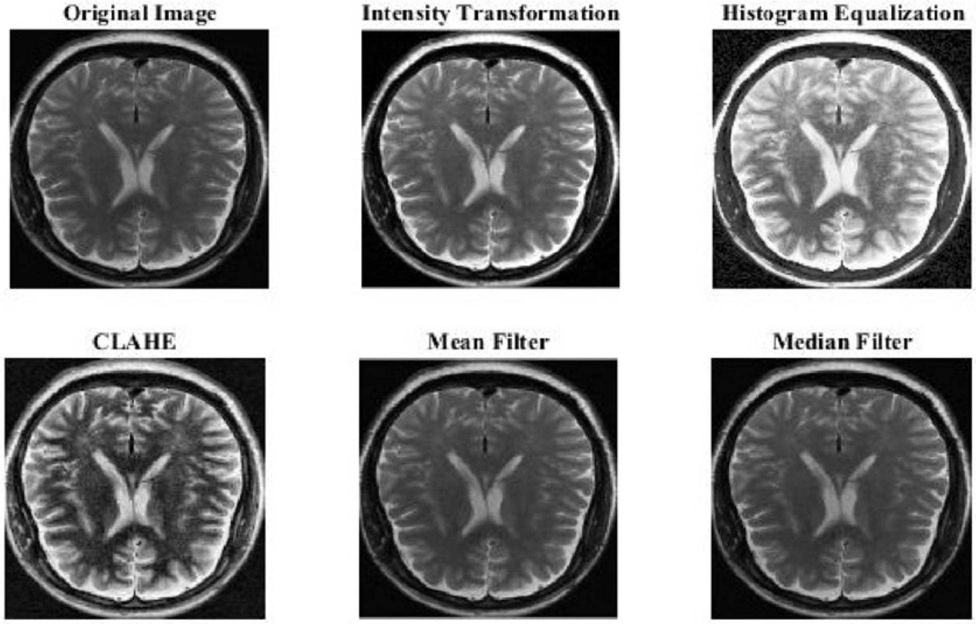

Fig. 2. Effects of various enhancement techniques.

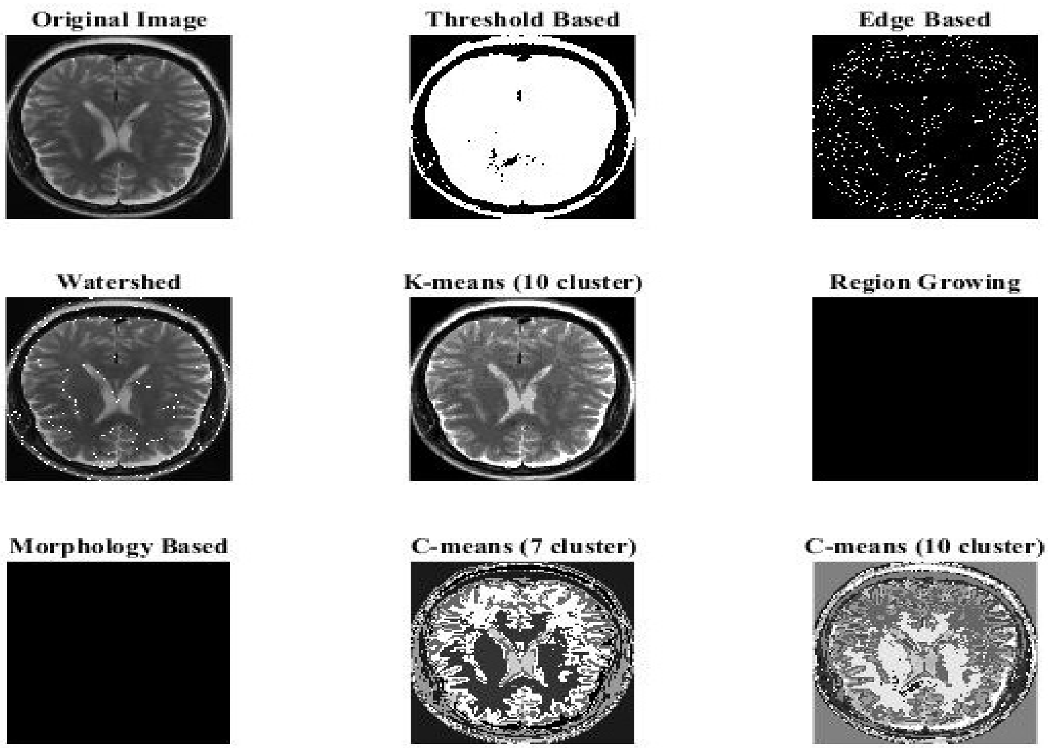

Fig. 3. Images with different clustering techniques.

Then, these are fed to benchmark CNN. In this paper, we have deployed five different benchmark CNN models such as AlexNet, VGG16, VGG19, ResNet18, and ResNet50. We have got the best results using ResNet50. The CNN generated a feature vector of size 262,144. We have applied principal component analysis for selecting optimal 8192 features from 262144 extracted by CNN.

E. Classification

In this paper, we have used five different classifiers (considering their classification performance as reputed in [28], [29]) such as DT [30], SVM [31], KNN [32], Linear programming boosting (LPBoost), and TotalBoost [33] after the feature extraction through different techniques.

TABLE III. PERFORMANCE ANALYSIS TABLE FOR IMAGE SEGMENTATION TECHNIQUES

Clustering Technique	MSSIM	PSNR	RMSE
Threshold-based (otsu)	0.9878	9.9152	0.3193
Edge-based (canny)	0.9957	12.1325	0.2474
Watershed (Gradient and Marker)	0.9989	16.3936	0.1515
K-means clustering (4 cluster)	0.9935	12.3958	0.2400
Region growing (shrink)	0.9963	14.5663	0.1869
Morphology based (thin)	0.9959	12.7946	0.2292
Fuzzy C-means clustering (4 clusters)	0.9289	2.5403	0.7464

III. RESULTS

In this experiment, we have investigated the overall classification accuracy including the individual precision, recall, f1-score, accuracy, and misclassification rate. At first, for each model deep CNN based algorithm such as AlexNet, VGG16, VGG19, ResNet18, ResNet50 were used to extract the enhanced discriminative features. Then ensemble-based TotalBoost, tree-based DT, KNN, and SVM methods were applied for classification. To identify the classification errors of the algorithm, we have calculated the confusion matrix for each method.

A. Performance Evaluation

To evaluate the performance of the models we have considered several metrics such as precision, negative predictive value (NPV), sensitivity, efficiency, f1 score, and accuracy. The number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) from the confusion matrix are used to define the performance metrics using the following equations from (1) to (6).

$$Accuracy(x,y) = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

$$PPV(x,y) = \frac{TP}{TP+FP} \tag{2}$$

$$NPV(x,y) = \frac{TN}{TN+FN} \tag{3}$$

$$Recall\ or\ Sensitivity\ or\ TPR(x,y) = \frac{TP}{TP+FN} \tag{4}$$

$$Efficiency\ or Specificity\ or\ TNR(x,y) = \frac{TN}{TN+FP} \tag{5}$$

$$F_1\ Score(x,y) = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{6}$$

B. Deep Feature Extraction using AlexNet

Experiments show that the AlexNet+ML classifier can successfully classify the different phases of AD. The overall classification accuracy for DT classifier achieved 81.5%. SVM, LPBoost, and TotalBoost attained 95.1%, 80.3%, and 81.6% accuracy individually. On the other hand, KNN reached the highest accuracy with 95.8% among the others. This result assures that the classification is performed correctly. Table IV illustrates the performance metrics of ML classifier with AlexNet. The detailed measurements of CN, MCI and AD classes are presented sequentially. Among all the classifier, KNN gained the highest average accuracy with 97.20%. Similarly, it reduced the minimum average error rate with 2.80% compared to the DT, SVM, LPBoost, and TotalBoost.

C. Deep Feature Extraction using VGG16

With features extracted by VGG16; DT, KNN, SVM, LPBoost and TotalBoost achieved the overall classification accuracy by 81.2%, 90.9%, 93.5%, 79.0% and 75.4% respectively. On the other hand, SVM reached the highest accuracy with 93.5% among the others. Table V illustrates the performance metrics of the ML classifier with VGG16. The detailed Among all the classifiers, SVM gained the highest average accuracy with 95.69%. Similarly, it reduced the minimum average error rate by 4.31% compared to the DT, KNN, LPBoost, and TotalBoost.

D. Deep Feature Extraction using VGG19

It has been found that SVM achieved the highest classification accuracy with 96.1%. DT, KNN, LPBoost, and Total Boost gained 79.9%, 92.6%, 84.8%, and 82.2% classification accuracy respectively. Table VI shows the Illustration of the performance metrics of ML classifier with ResNet50. Among all the classifiers, SVM gained the highest average accuracy with 97.41% minimum average error rate of 2.59%.

E. Deep Feature Extraction using ResNet18

The ResNet18+ML classifier model shows the classification of 3 different AD phases. SVM achieved the highest classification accuracy with 91.3%. DT, KNN, LPBoost, and Total Boost gained 75.1%, 90.0%, 79.0%, and 74.4% classification accuracy respectively. An illustration of the performance metrics of the ML classifier with ResNet18 is given in Table VII. Among all the classifiers, SVM gained the highest average accuracy with 94.17% minimum average error rate of 5.83%.

F. Deep Feature Extraction using ResNet50

It has been observed that SVM achieved the highest classification accuracy with 98.1%. Other classifiers such as DT, KNN, LPBoost, and Total Boost achieved 81.6%, 91.5%, 85.8%, and 81.6% classification accuracy. From the Table VIII we can observe that SVM gained 98.71% average accuracy. So, the average error rate is 1.29.

In Fig. 4 comparison of different CNN models has been shown. This figure has represented the performance of different CNN models based on accuracy and error rate. Here, ResNet50 with SVM has been provided with a high accuracy rate which is 98.71% and an error rate is 1.29% for the dataset. Based

TABLE IV. Performance Metrics for Three Class ML Classifier using AlexNet

Model	Class	Accuracy	Precision	NPV	Recall	Efficiency	F1 Score
AlexNet+ DT	Alzheimer's	0.9838	0.9833	0.9841	0.9752	0.9894	0.9793
	Cognitive Normal	0.8285	0.6737	0.8972	0.7442	0.861	0.7072
	MCI	0.8188	0.7447	0.8512	0.6863	0.8841	0.7143
	Average	**0.877**	**0.8005**	**0.9108**	**0.8019**	**0.9115**	**0.8002**
AlexNet+ KNN	Alzheimer's	1	1	1	1	1	1
	Cognitive Normal	0.9579	0.8947	0.986	0.9659	0.9548	0.929
	MCI	0.9579	0.9681	0.9535	0.901	0.9856	0.9333
	Average	**0.9719**	**0.9543**	**0.9798**	**0.9556**	**0.9801**	**0.9541**
AlexNet+ SVM	Alzheimer's	1	1	1	1	1	1
	Cognitive Normal	0.9515	0.9263	0.9626	0.9167	0.9671	0.9215
	MCI	0.955	0.9149	0.9674	0.9247	0.963	0.9198
	Average	**0.9688**	**0.947**	**0.9766**	**0.9471**	**0.9767**	**0.9471**
AlexNet+ LPBoost	Alzheimer's	0.9838	0.9667	0.9947	0.9915	0.9792	0.9789
	Cognitive Normal	0.8026	0.7263	0.8364	0.6635	0.8732	0.6935
	MCI	0.8188	0.6702	0.8837	0.7159	0.8597	0.6923
	Average	**0.8684**	**0.7877**	**0.9049**	**0.7903**	**0.904**	**0.7882**
AlexNet+ TotalBoost	Alzheimer's	0.945	0.8667	0.9947	0.9905	0.9216	0.9244
	Cognitive Normal	0.8155	0.8737	0.7897	0.6484	0.9337	0.7444
	MCI	0.8706	0.6915	0.9488	0.8553	0.8755	0.7647
	Average	**0.877**	**0.8106**	**0.9111**	**0.8314**	**0.9103**	**0.8112**

TABLE V. Performance Metrics for Three Class ML Classifier using VGG16

Model	Class	Accuracy	Precision	NPV	Recall	Efficiency	F1 Score
VGG16 + DT	Alzheimer's	0.9450	0.9333	0.9524	0.9256	0.9574	0.9295
	Cognitive Normal	0.8479	0.7474	0.8925	0.7553	0.8884	0.7513
	MCI	0.8317	0.7234	0.8791	0.7234	0.8791	0.7234
	Average	**0.8748**	**0.8013**	**0.9080**	**0.8014**	**0.9083**	**0.8014**
VGG16+ KNN	Alzheimer's	0.9935	1	0.9894	0.9836	1	0.9917
	Cognitive Normal	0.9159	0.7368	0.9953	0.9859	0.895	0.8434
	MCI	0.9094	0.9681	0.8837	0.7845	0.9845	0.8667
	Average	**0.9396**	**0.9016**	**0.9561**	**0.9180**	**0.9598**	**0.9006**
VGG16+ SVM	Alzheimer's	1	1	1	1	1	1
	Cognitive Normal	0.9353	0.8632	0.9673	0.9213	0.9409	0.8913
	MCI	0.9353	0.9255	0.9395	0.87	0.9665	0.8969
	Average	**0.9569**	**0.9296**	**0.9689**	**0.9304**	**0.9691**	**0.9294**
VGG16+ LPBoost	Alzheimer's	0.9709	0.9333	0.9947	0.9912	0.9592	0.9614
	Cognitive Normal	0.8155	0.7263	0.8551	0.6900	0.8756	0.7077
	MCI	0.7929	0.6702	0.8465	0.6563	0.8545	0.6632
	Average	**0.8598**	**0.7766**	**0.8988**	**0.7792**	**0.8964**	**0.7774**
VGG16+ TotalBoost	Alzheimer's	0.9256	0.8167	0.9947	0.9899	0.8952	0.895
	Cognitive Normal	0.7767	0.8421	0.7477	0.597	0.9143	0.6987
	MCI	0.8058	0.5851	0.9023	0.7237	0.8326	0.6471
	Average	**0.8360**	**0.7480**	**0.8816**	**0.7702**	**0.8807**	**0.7470**

on this result it can be notified that SVM and KNN perform better than other classifiers. The performance of the ensemble classifier is not that much efficient for AD classification.

IV. Discussion

The main objective of this work is to diagnose of AD in the early stages accurately. The comparative study of some of the recent state-of-the-art works in this field with our proposed

TABLE VI. PERFORMANCE METRICS FOR THREE CLASS ML CLASSIFIER USING VGG19

Model	Class	Accuracy	Precision	NPV	Recall	Efficiency	F1 Score
VGG19 + DT	Alzheimer's	0.9515	0.9500	0.9524	0.9268	0.9677	0.9383
	Cognitive Normal	0.8123	0.7263	0.8505	0.6832	0.8750	0.7041
	MCI	0.8350	0.6809	0.9023	0.7529	0.8661	0.7151
	Average	**0.8662**	**0.7857**	**0.9017**	**0.7876**	**0.9029**	**0.7858**
VGG19+ KNN	Alzheimer's	0.9871	1	0.9788	0.9677	1	0.9836
	Cognitive Normal	0.9320	0.8211	0.9813	0.9512	0.9251	0.8814
	MCI	0.9320	0.9362	0.9302	0.8544	0.9709	0.8934
	Average	**0.9503**	**0.9191**	**0.9634**	**0.9244**	**0.9653**	**0.9194**
VGG19+ SVM	Alzheimer's	1	1	1	1	1	1
	Cognitive Normal	0.9612	0.9368	0.972	0.9368	0.9720	0.9368
	MCI	0.9612	0.9362	0.9721	0.9362	0.9721	0.9362
	Average	**0.9741**	**0.9577**	**0.9813**	**0.9577**	**0.9814**	**0.9577**
VGG19+ LPBoost	Alzheimer's	0.9644	0.9167	0.9947	0.9910	0.9495	0.9524
	Cognitive Normal	0.8544	0.8421	0.8598	0.7273	0.9246	0.7805
	MCI	0.8770	0.766	0.9256	0.8182	0.9005	0.7912
	Average	**0.8986**	**0.8416**	**0.9267**	**0.8455**	**0.9248**	**0.8413**
VGG19+ TotalBoost	Alzheimer's	0.9450	0.8583	1	1	0.9175	0.9238
	Cognitive Normal	0.8479	0.7368	0.8972	0.7609	0.8848	0.7487
	MCI	0.8511	0.8617	0.8465	0.7105	0.9333	0.7788
	Average	**0.8813**	**0.8189**	**0.91457**	**0.8238**	**0.9119**	**0.8171**

TABLE VII. PERFORMANCE METRICS FOR THREE CLASS ML CLASSIFIER USING RESNET18

Model	Class	Accuracy	Precision	NPV	Recall	Efficiency	F1 Score
ResNet18 + DT	Alzheimer's	0.9159	0.8833	0.9365	0.8983	0.9267	0.8908
	Cognitive Normal	0.7767	0.6737	0.8224	0.6275	0.8502	0.6497
	MCI	0.8091	0.6596	0.8744	0.6966	0.8545	0.6776
	Average	**0.8339**	**0.7389**	**0.8778**	**0.7408**	**0.8771**	**0.7393**
ResNet18+ KNN	Alzheimer's	0.9968	0.9917	1	1	0.9947	0.9958
	Cognitive Normal	0.8997	0.8000	0.9439	0.8636	0.914	0.8306
	MCI	0.9029	0.883	0.9116	0.8137	0.9469	0.8469
	Average	**0.9331**	**0.8916**	**0.9518**	**0.8924**	**0.9519**	**0.8911**
ResNet18+ SVM	Alzheimer's	0.9871	0.9833	0.9894	0.9833	0.9894	0.9833
	Cognitive Normal	0.9126	0.8632	0.9346	0.8542	0.939	0.8586
	MCI	0.9256	0.8723	0.9488	0.8817	0.9444	0.877
	Average	**0.9418**	**0.9062**	**0.9576**	**0.9064**	**0.9576**	**0.9063**
ResNet18 + LPBoost	Alzheimer's	0.9547	0.9000	0.9894	0.9818	0.9397	0.9391
	Cognitive Normal	0.7929	0.8632	0.7617	0.6165	0.9261	0.7193
	MCI	0.8317	0.5745	0.9442	0.8182	0.8354	0.6750
	Average	**0.8598**	**0.7792**	**0.8984**	**0.8055**	**0.9004**	**0.7778**
ResNet18 + TotalBoost	Alzheimer's	0.9320	0.8333	0.9947	0.9901	0.9038	0.9050
	Cognitive Normal	0.7476	0.8316	0.7103	0.5603	0.9048	0.6695
	MCI	0.8091	0.5426	0.9256	0.7612	0.8223	0.6335
	Average	**0.8296**	**0.7358**	**0.8769**	**0.7705**	**0.8770**	**0.7360**

model has been shown in Table IX.

Jain et al. [34] utilized VGG19 features for classification using DT and demonstrated 86.62% overall accuracy with a sensitivity of 78.76% and a specificity of 90.29. The authors computed the whole brain in their work. Our method demonstrated higher performance with the VGG16+PCA+DT pipeline in reduced sampled brain region (accuracy 87.48%, sensitivity 80.13%, and specificity 90.83%). Pueto-Castro et

TABLE VIII. PERFORMANCE METRICS FOR THREE CLASS ML CLASSIFIER USING RESNET50

Model	Class	Accuracy	Precision	NPV	Recall	Efficiency	F1 Score
ResNet50 + DT	Alzheimer's	0.9482	0.9250	0.9630	0.9407	0.9529	0.9328
	Cognitive Normal	0.8317	0.6842	0.8972	0.7471	0.8649	0.7143
	MCI	0.8511	0.8085	0.8698	0.7308	0.9122	0.7677
	Average	**0.8770**	**0.8059**	**0.9100**	**0.8062**	**0.9100**	**0.8049**
ResNet 50+ KNN	Alzheimer's	0.9968	0.9917	1	1	0.9947	0.9958
	Cognitive Normal	0.9450	0.9053	0.9626	0.9149	0.9581	0.9101
	MCI	0.9482	0.9255	0.9581	0.9063	0.9671	0.9158
	Average	**0.9633**	**0.9408**	**0.9736**	**0.9404**	**0.9733**	**0.9406**
ResNet50+ SVM	Alzheimer's	0.9968	0.9917	1	1	0.9947	0.9958
	Cognitive Normal	0.9806	0.9684	0.9860	0.9684	0.9860	0.9684
	MCI	0.9838	0.9787	0.9860	0.9684	0.9907	0.9735
	Average	**0.9871**	**0.9796**	**0.9907**	**0.9789**	**0.9904**	**0.9792**
ResNet50 + LPBoost	Alzheimer's	0.9741	0.9333	1	1	0.9594	0.9655
	Cognitive Normal	0.8576	0.9158	0.8318	0.7073	0.9570	0.7982
	MCI	0.8835	0.7021	0.9628	0.8919	0.8809	0.7857
	Average	**0.9050**	**0.8504**	**0.9315**	**0.8664**	**0.9324**	**0.8498**
ResNet50 + TotalBoost	Alzheimer's	0.9547	0.8833	1	1	0.9310	0.9381
	Cognitive Normal	0.8155	0.8842	0.7850	0.6462	0.9385	0.7467
	MCI	0.8608	0.6596	0.9488	0.8493	0.8644	0.7425
	Average	**0.8770**	**0.8090**	**0.911267**	**0.831833**	**0.9113**	**0.8091**

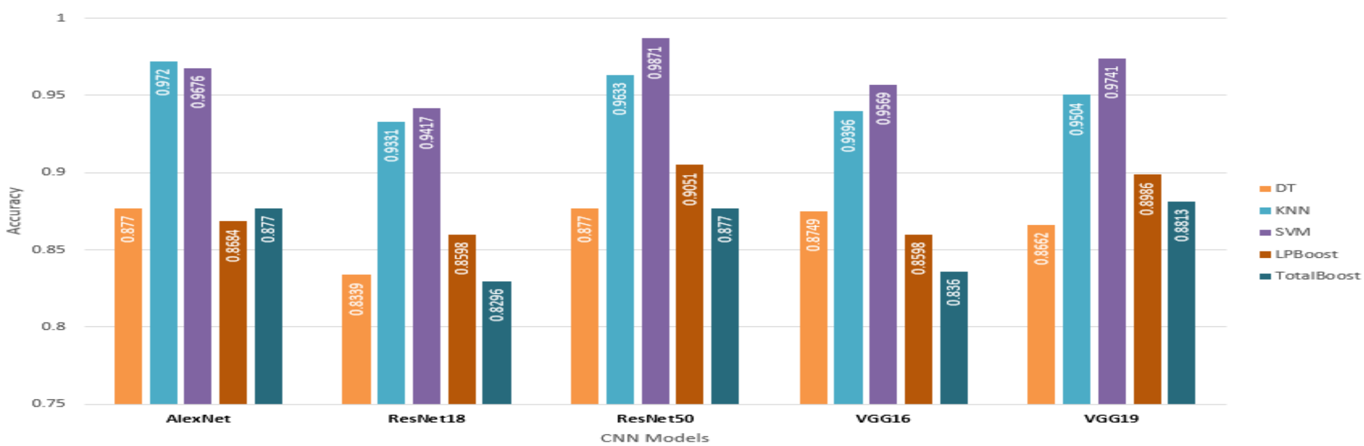

Fig. 4. Comparison of the performance of applied techniques consists of CNN feature extractor with ML classifier.

al. [35] exploited OASIS dataset and deployed RESNET18 with SVM classifiers on the whole brain. The method demonstrated a sensitivity of 58.66% and specificity of 80.21% while combining the RESNET 18 features with DenseNet121 features Odusami et al. [36] showed more than 98% in all performance measures. Feng et al. [6] utilized 3DCNN with SVM and showed 92% accuracy with standard deviation of 2. Raju et al. [37] have shown higher performance with the same method and same dataset(97% above in terms of accuracy, precision and recall). Abdulazeem et al. [38] designed a CNN classifier and demonstrated 97.50% accuracy while Hazarika et al. [39] demonstrated 88.66% accuracy with CNN based hybrid model. They have computed the whole brain. In our work, the CNN model ResNet50 along with SVM classifier has achieved comparable performance with 98.71% accuracy, 97.96% precision, 99.07% NPV, 97.89% recall, 99.04% specificity, and 97.92% f1 score. It is evident from the Table IX that our proposed model outperforms other works such as [6], [34]–[39]. Moreover, in comparison to the whole brain computation of the studies we have computed features from 128 by 128 by 3 slices of MRIs.

V. CONCLUSIONS

In this paper, we have presented a pipeline for classifying an MRI into one of its three stages(AD, MCI, CN). We have leveraged the benefits of the capacity of deep learning

TABLE IX. COMPARISON WITH STATE-OF-THE ART WORKS

Study	Dataset with stages	Modality	Feature Extraction with Classifier	Performance metrics
Jain et al. [34]	ADNI-150 subjects (AD-50,CN-50, MCI-50)	sMRI	VGG16	Accuracy: 95.73% Precision:96.33% Recall: 96% F1 score: 95.66%
Pueto-Castro et al. [35]	OASIS-416 (AD-2, CN-316, MCI-98); ADNI-1743 (AD-287, CN-525, MCI-921)	MRI	Resnet 18 and SVM	Accuracy: 78.72% Precision: 68.96 % Recall: 58.66% Specificity: 80.21% F1 score: 60.79%
Odusami et al. [36]	ADNI (AD,CN, MCI)	MRI	Resnet18 and DenseNet121 with Randomized weight	Accuracy: 98.21% Precision: 98.14 % Recall: 98.14%
Feng et al. [6]	ADNI-469 subjects (AD-153, MCI-157, CN-159)	MRI	3D-CNN with SVM	Accuracy: 92.11%± 2.31
Raju et al. [37]	ADNI-465 subjects (AD-132, MCI-181, CN-152)	MRI	3D-CNN with SVM	Accuracy: 97.77% Precision: 97.93% Recall: 97.76% F1 score: 97.80
Abdulazeem et al. [38]	ADNI-211,655 (After augmentation)	MRI	CNN	Accuracy: 97.50%
Hazarika et al. [39]	ADNI- 150 subjects (CN:50, MCI: 50, AD: 50)	MRI	Custom CNN based Hybrid Model	Accuracy: 84.66% Precision: 88.33% Recall: 87.66% F1 score: 88.33%
Proposed	**ADNI-1546 (CN-470, MCI-477, AD-599)**	**MRI**	**Resnet50 +SVM**	**Accuracy: 98.71%** **Precision: 97.96 %** **NPV: 99.07%** **Sensitivity/Recall: 97.89%** **Specificity: 99.04%** **F1 Score: 97.92%**

methods in feature extraction and the classification strength of conventional ML methods. In our method, we have optimized benchmark CNN-extracted features from three view patches by PCA that are generated from segmented regions of MRI enabling us to avoid whole-brain computation. We have demonstrated state-of-the-art performance exploited on the ADNI dataset. Our work showed that the RESNET50-PCA-SVM pipeline suits well for this multi-class classification task.

REFERENCES

[1] S. Rayaprolu, L. Higginbotham, P. Bagchi, C. M. Watson, T. Zhang, A. I. Levey, S. Rangaraju, and N. T. Seyfried, "Systems-based proteomics to resolve the biology of Alzheimer's disease beyond amyloid and tau," *Neuropsychopharmacology*, vol. 46, no. 1, pp. 98–115, Jan. 2021. [Online]. Available: https://www.nature.com/articles/s41386-020-00840-3

[2] S. Ahmed, B. C. Kim, K. H. Lee, H. Y. Jung, and for the Alzheimer's Disease Neuroimaging Initiative, "Ensemble of roi-based convolutional neural network classifiers for staging the alzheimer disease spectrum from magnetic resonance imaging," *PLOS ONE*, vol. 15, no. 12, pp. 1–23, 12 2020. [Online]. Available: https://doi.org/10.1371/journal.pone.0242712

[3] A. Shukla, R. Tiwari, and S. Tiwari, "Review on alzheimer disease detection methods: Automatic pipelines and machine learning techniques," *Sci*, vol. 5, no. 1, 2023. [Online]. Available: https://www.mdpi.com/2413-4155/5/1/13

[4] J. Islam and Y. Zhang, "A novel deep learning based multi-class classification method for alzheimer's disease detection using brain mri data," in *Brain Informatics*, Y. Zeng, Y. He, J. H. Kotaleski, M. Martone, B. Xu, H. Peng, and Q. Luo, Eds. Cham: Springer International Publishing, 2017, vol. 10654, pp. 213–222. [Online]. Available: http://link.springer.com/10.1007/978-3-319-70772-3_20

[5] S. Lahmiri and A. Shmuel, "Performance of machine learning methods applied to structural MRI and ADAS cognitive scores in diagnosing Alzheimer's disease," *Biomedical Signal Processing and Control*, vol. 52, pp. 414–419, Jul. 2019. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S174680941830199X

[6] W. Feng, N. V. Halm-Lutterodt, H. Tang, A. Mecum, M. K. Mesregah, Y. Ma, H. Li, F. Zhang, Z. Wu, E. Yao, and X. Guo, "Automated MRI-Based Deep Learning Model for Detection of Alzheimer's Disease Process," *International Journal of Neural Systems*, vol. 30, no. 06, p. 2050032, Jun. 2020. [Online]. Available: https://www.worldscientific.com/doi/abs/10.1142/S012906572050032X

[7] "2022 Alzheimer's disease facts and figures," *Alzheimer's & Dementia*, vol. 18, no. 4, pp. 700–789, Apr. 2022. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/alz.12638

[8] T. Altaf, S. M. Anwar, N. Gul, M. N. Majeed, and M. Majid, "Multi-class Alzheimer's disease classification using image and clinical features," *Biomedical Signal Processing and Control*, vol. 43, pp. 64–74, May 2018. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1746809418300508

[9] T. J. Saleem, S. R. Zahra, F. Wu, A. Alwakeel, M. Alwakeel, F. Jeribi, and M. Hijji, "Deep Learning-Based Diagnosis of Alzheimer's Disease," *Journal of Personalized Medicine*, vol. 12, no. 5, p. 815, May 2022. [Online]. Available: https://www.mdpi.com/2075-4426/12/5/815

[10] E. M. Ali, A. F. Seddik, and M. H. Haggag, "Automatic Detection and Classification of Alzheimer's Disease from MRI using TANNN," *International Journal of Computer Applications*, vol. 148, no. 9, pp. 30–34, Aug. 2016. [Online]. Available: https://www.ijcaonline.org/archives/volume148/number9/25787-2016911320

[11] S. Sarraf, D. D. DeSouza, J. Anderson, G. Tofighi, and f. t. A. D. N. Initiativ, "DeepAD: Alzheimer's Disease Classification via Deep Convolutional Neural Networks using MRI and fMRI," Jan. 2017. [Online]. Available: https://www.biorxiv.org/content/10.1101/070441v4

[12] M. Tanveer, B. Richhariya, R. U. Khan, A. H. Rashid, P. Khanna, M. Prasad, and C. T. Lin, "Machine Learning Techniques for the Diagnosis of Alzheimer's Disease: A Review," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 16, no. 1s, pp. 1–35, Jan. 2020. [Online]. Available: https://dl.acm.org/doi/10.1145/3344998

[13] G. Battineni, N. Chintalapudi, F. Amenta, and E. Traini, "A Comprehensive Machine-Learning Model Applied to Magnetic Resonance Imaging (MRI) to Predict Alzheimer's Disease (AD) in Older Subjects," *Journal of Clinical Medicine*, vol. 9, no. 7, p. 2146, Jul. 2020. [Online]. Available: https://www.mdpi.com/2077-0383/9/7/2146

[14] R. Karim, A. Shahrior, and M. M. Rahman, "Machine learningbased tristage classification of Alzheimer's progressive neurodegenerative disease using PCA and mRMR administered textural, orientational, and spatial features," *International Journal of Imaging Systems and Technology*, vol. 31, no. 4, pp. 2060–2074, jun 30 2021.

[15] "Alzheimer's and Dementia — alz.org," https://www.alz.org/alzheimer_s_dementia, [Accessed 04-May-2023].

[16] S. Pallawi and D. K. Singh, "Study of alzheimer's disease brain impairment and methods for its early diagnosis: a comprehensive survey," *Int. J. Multim. Inf. Retr.*, vol. 12, no. 1, p. 7, 2023. [Online]. Available: https://doi.org/10.1007/s13735-023-00271-y

[17] A. Alberdi, A. Aztiria, and A. Basarab, "On the early diagnosis of alzheimer's disease from multimodal signals: A survey," *Artif. Intell. Medicine*, vol. 71, pp. 1–29, 2016. [Online]. Available: https://doi.org/10.1016/j.artmed.2016.06.003

[18] G. Martí-Juan, G. Sanroma-Guell, and G. Piella, "A survey on machine and statistical learning for longitudinal analysis of neuroimaging data in alzheimer's disease," *Comput. Methods Programs Biomed.*, vol. 189, p. 105348, 2020. [Online]. Available: https://doi.org/10.1016/j.cmpb.2020.105348

[19] R. A. Hazarika, A. K. Maji, S. N. Sur, B. S. Paul, and D. Kandar, "A survey on classification algorithms of brain images in alzheimer's disease based on feature extraction techniques," *IEEE Access*, vol. 9, pp. 58 503–58 536, 2021. [Online]. Available: https://doi.org/10.1109/ACCESS.2021.3072559

[20] D. A. Arafa, H. E. Moustafa, A. M. T. Ali-Eldin, and H. A. Ali, "Early detection of alzheimer's disease based on the state-of-the-art deep learning approach: a comprehensive survey," *Multim. Tools Appl.*, vol. 81, no. 17, pp. 23 735–23 776, 2022. [Online]. Available: https://doi.org/10.1007/s11042-022-11925-0

[21] S. Ahmed, K. Y. Choi, J. J. Lee, B. C. Kim, G.-R. Kwon, K. H. Lee, and H. Y. Jung, "Ensembles of patch-based classifiers for diagnosis of alzheimer diseases," *IEEE Access*, vol. 7, pp. 73 373–73 383, 2019.

[22] S. Ahmed, K. H. Lee, and H. Y. Jung, "Robust hippocampus localization from structured magnetic resonance imaging using similarity metric learning," *IEEE Access*, vol. 10, pp. 7141–7152, 2022.

[23] J. J. Prado and I. Rojas, "Machine Learning for Diagnosis of Alzheimer's Disease and Early Stages," *BioMedInformatics*, vol. 1, no. 3, pp. 182–200, Dec. 2021. [Online]. Available: https://www.mdpi.com/2673-7426/1/3/12

[24] B. M. Rashed and N. Popescu, "Critical analysis of the current medical image-based processing techniques for automatic disease evaluation: Systematic literature review," *Sensors*, vol. 22, no. 18, p. 7065, 2022.

[25] K. P. Shankar and S. P. Shyry, "A survey of image pre-processing techniques for medical images," in *Journal of Physics: Conference Series*, vol. 1911, no. 1. IOP Publishing, 2021, p. 012003.

[26] H. Mittal, A. C. Pandey, M. Saraswat, S. Kumar, R. Pal, and G. Modwel, "A comprehensive survey of image segmentation: Clustering methods, performance parameters, and benchmark datasets," *Multimedia Tools and Applications*, vol. 81, no. 24, p. 35001–35026, 2021.

[27] M. Fayez, S. Safwat, and E. Hassanein, "Comparative study of clustering medical images," in *2016 SAI Computing Conference (SAI)*, 2016, pp. 312–318.

[28] E. Miranda, M. Aryuni, and E. Irwansyah, "A survey of medical image classification techniques," in *2016 International Conference on Information Management and Technology (ICIMTech)*, 2016, pp. 56–61.

[29] Z. Solatidehkordi and I. Zualkernan, "Survey on recent trends in medical image classification using semi-supervised learning," *Applied Sciences*, vol. 12, no. 23, 2022. [Online]. Available: https://www.mdpi.com/2076-3417/12/23/12094

[30] L. Rokach and O. Maimon, *Decision Trees*. Boston, MA: Springer US, 2005, pp. 165–192. [Online]. Available: https://doi.org/10.1007/0-387-25465-X_9

[31] N. Cristianini and E. Ricci, *Support Vector Machines*. Boston, MA: Springer US, 2008, pp. 928–932. [Online]. Available: https://doi.org/10.1007/978-0-387-30162-4_415

[32] A. Mucherino, P. J. Papajorgji, and P. M. Pardalos, *k-Nearest Neighbor Classification*. New York, NY: Springer New York, 2009, pp. 83–106. [Online]. Available: https://doi.org/10.1007/978-0-387-88615-2_4

[33] M. K. Warmuth, J. Liao, and G. Rätsch, "Totally corrective boosting algorithms that maximize the margin," in *Proceedings of the 23rd international conference on Machine learning - ICML '06*. Pittsburgh, Pennsylvania: ACM Press, 2006, pp. 1001–1008. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1143844.1143970

[34] R. Jain, N. Jain, A. Aggarwal, and D. J. Hemanth, "Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images," *Cognitive Systems Research*, vol. 57, pp. 147–159, Oct. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389041718309562

[35] A. Puente-Castro, E. Fernandez-Blanco, A. Pazos, and C. R. Munteanu, "Automatic assessment of Alzheimer's disease diagnosis based on deep learning techniques," *Computers in Biology and Medicine*, vol. 120, p. 103764, May 2020. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0010482520301384

[36] M. Odusami, R. Maskeliūnas, and R. Damaševičius, "An Intelligent System for Early Recognition of Alzheimer's Disease Using Neuroimaging," *Sensors*, vol. 22, no. 3, p. 740, Jan. 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/3/740

[37] M. Raju, V. P. Gopi, V. S. Anitha, and K. A. Wahid, "Multi-class diagnosis of Alzheimer's disease using cascaded three dimensional-convolutional neural network," *Physical and Engineering Sciences in Medicine*, vol. 43, no. 4, pp. 1219–1228, Dec. 2020. [Online]. Available: https://link.springer.com/10.1007/s13246-020-00924-w

[38] Y. AbdulAzeem, W. M. Bahgat, and M. Badawy, "A CNN based framework for classification of Alzheimer's disease," *Neural Computing and Applications*, vol. 33, no. 16, pp. 10 415–10 428, Aug. 2021. [Online]. Available: https://link.springer.com/10.1007/s00521-021-05799-w

[39] R. A. Hazarika, A. K. Maji, D. Kandar, E. Jasinska, P. Krejci, Z. Leonowicz, and M. Jasinski, "An Approach for Classification of Alzheimer's Disease Using Deep Neural Network and Brain Magnetic Resonance Imaging (MRI)," *Electronics*, vol. 12, no. 3, p. 676, Jan. 2023. [Online]. Available: https://www.mdpi.com/2079-9292/12/3/676