Unveiling Alzheimer's Progression: AI-Driven Models for Classifying Stages of Cognitive Impairment Through Medical Imaging

Chapter · May 2025

DOI: 10.1007/978-3-031-88220-3_4

3 authors, including:

Vaibhav C. Gandhi
Madhuben & Bhanubhai Patel Women's Institute of Engineering for Studies & Res…
27 PUBLICATIONS **45** CITATIONS

SEE PROFILE

Mariofanna Milanova
University of Arkansas at Little Rock
214 PUBLICATIONS **1,294** CITATIONS

SEE PROFILE

Unveiling Alzheimer's Progression: AI-Driven Models for Classifying Stages of Cognitive Impairment Through Medical Imaging

Vaibhav C. Gandhi[1(✉)], Dhruvi Thakkar[2], and Mariofanna Milanova[3]

[1] Department of Computer Engineering, Madhuben and Bhanubhai Patel Institute of Technology, The Charutar Vidya Mandal (CVM) University, Anand, India
vaibhavgandhi2424@gmail.com

[2] Computer Science and Engineering, Drs Kiran and Pallavi Patel Global University, Vadodara, India

[3] Department of Computer Science, University of Arkansas, Little Rock, USA
mgmilanova@ualr.edu

Abstract. Alzheimer's disease (AD) presents a substantial contest, requiring primary diagnosis and precise staging of cognitive impairment for effective management. AD is a broad-minded neurodegenerative illness categorized by reminiscence injury, cognitive decline, and changes in behavior, ultimately affecting regular functioning and quality of life. This study introduces a DL approach using magnetic resonance imaging (MRI) to classify stages of cognitive impairment associated with AD. Here assessed various DL models, including Xception, Inception, Mobile Net, VGG16, ResNet, and DenseNet, with Xception achieving the highest accuracy of 99.1%. Our approach demonstrates improved classification accuracy and early detection rates, validated through rigorous testing. This method provides a valuable tool for clinicians, potentially enhancing patient outcomes and supporting personalized treatment strategies in Alzheimer's care.

Keywords: Alzheimer's disease · cognitive impairment · DL · medical imaging · MRI · DenseNet · convolutional neural networks · early diagnosis · classification · neurodegenerative disorders

1 Introduction

Alzheimer's disease, a deceptively open-minded neurodegenerative syndrome, is the greatest shared source of dementia universal. Considered by progressive and gradual deterioration of all functions of cognition which include memory, thinking, and behavior leading to a significant reduction in a being's aptitude to carry out daily activities. The etiology of disease remains unclear, although it is thought to be multifactorial and associated with both genetic and environmental and lifestyle influences. The most prominent pathological features of AD consist of the accretion of amyloid-beta proteins and neurofibrillary tangles that led to cell death and dysfunction of brain cells. Symptoms - Patients affected by illness exhibit a range of symptoms as the disease progresses, from

S. Palaiahnakote et al. (Eds.): ICPR 2024, LNCS 15618, pp. 55–87, 2025.
https://doi.org/10.1007/978-3-031-88220-3_4

loss of memory and comprehension and lack of ability to use or find words, solve problems, and performing more, but actually losing, skills as well as mood swings and personality changes. This is a slow evolution characterizing the general course of this disease over time. Premature discovery and judgement of Alzheimer's disease can enable appropriate intervention, management, and may slow the movement of the illness it [1–3].

This tabular format provides a clear and concise overview of the dissimilar periods of Alzheimer's disease, highlighting the key characteristics and vicissitudes that arise at each stage of the disease progression [3] (Table 1).

Table 1. Dissimilar phases of Alzheimer's disease.

Stage	Characteristics
Preclinical Stage	• Variations in the brain commence years before any symptoms appear • This asymptomatic stage can last for years
Slight Stage	• Mild forgetfulness and memory lapses, such as forgetting names or recent events • The person can still live independently but may have trouble with concentration, planning, and organization • Friends and family may start to notice the cognitive changes
Moderate/Middle Stage	• Symptoms become more pronounced, including increased trouble with memory, learning new things, and carrying out complex tasks • The person may forget names of close family members, get lost easily, and need help with daily activities like dressing and bathing • Personality changes, mood swings, and behavioral issues may develop
Severe/Late Stage	• Severe memory loss and lack of awareness of recent experiences and surroundings • The person loses ability to communicate, walk, sit, and perform self-care • They become completely dependent on caregivers for all activities • Increased risk of infections, especially pneumonia

The foremost aids of this paper embrace the subsequent:

- This approach is novel, mainly because it combines Convolutional Neural Networks and Multi-feature Kernel Supervised Class-Similar Discriminative Dictionary Knowledge to enhance accurateness in ordering Alzheimer's disease.
- Mild Cognitive Impairment, and Cognitively Normal entities. This approach combines the ability of powerful feature extraction by CNNs by MKSCDDL for enhancing discriminative feature representation, achieving high classification accuracy, and identifying the specific brain region at each stage, thus possessing interpretability.
- Integration of such an architecture proposal a more refined distinction between the stages of cognitive impairments, thus leading towards early intervention and personalization of the treatment.

Uniform in primary detection, where minimum changes accompanied by the retinal fundus are hard to determine, the use of advanced techniques is still not developed. The conventional methodologies applied in diagnosing patients have limitations, so one cannot distinguish normal anatomical structures from very initial symptoms of glaucoma. In this regard, the severe need for more advanced algorithms and methodologies applied only in early detection is felt greatly. Such designs will allow timely interventions and many others and will increase the chance to largely improve the patient's outcome [4, 5]. The respite of the thing has organized the remainder into the following sections: Sect. 2. A comprehensive review and tabular summary of previous work carried out by different individuals have been presented. Section 3. The proposed method, the materials, as well as methodology for diagnosis of glaucoma along through outlining evaluation criteria and accessible databases have been specified. Section 4. Methodology towards analyzing the experimental results has been elaborated which incorporates comparative analysis of training and testing outcomes. Finally, Sect. 5 summarizes the key findings and highlights implications for additional investigate.

2 Related Work

The object gives a systematic review of DL techniques for detection in Alzheimer's disease (AD) using neuroimaging statistics. Several DL models have achieved outstanding performance in differentiation against AD from healthy controls and in the classification of MCI subjects to progress to AD. Among DL architectures, CNNs have been the most frequently used architecture for the detection of AD from sMRI and PET scans. The appraisal underlines the pre-processing, augmentation, as well as the cross-validation strategies in improving the model's performance. Indeed, the authors emphasized that a large and diverse dataset should be required per standardization of evaluation protocols to allow for fair inter-study comparison. The interpretation approaches are crucial in providing insight into DL models and the most discriminative brain regions in AD detection. In instant, this evaluation emphasizes the possibility of revolutionizing early, accurate diagnosis through the use of DL aimed at the administration of AD that whitethorn be sought in time. [1].

The current article discusses DL-based tactics for the diagnosis and classification of AD from neuroimaging data. The ability of deep models, including CNNs and RNNs, to differentiate between AD patients and healthy controls at high accuracy levels, as well as their capability to forecast the evolution from MCI to AD, would depend on a number of leading factors such as the use of multimodal neuroimaging statistics, like MRI and PET scans, data augmentation, and transfer learning. The challenges lie in getting large diverse datasets and developing interpretable models for gaining insights into the underlying apparatuses of the disease. Prevalent this esteem, the assessment throws light on the promise of DL in revolutionizing early AD diagnosis, but has also established the need for further methodological advancement and clinical validation [2].

The study aimed to determine the stage of AD using mobility data and DL models. The data was collected on the smartphones of 35 AD patients for one week through accelerometer changes during daily activities and was processed as time series data, using a CNN model to recognize the patterns that could correspond to an early, middle,

or late stage of AD. The authors compared it accompanied by traditional feature-based classifiers and the CNN-based approach outperformed them by an accuracy of 90.91% and F1-score of 0.897. Based on these results, the authors concluded that mobility data may serve as a appreciated basis of material for tracking the progression of AD and that their CNN-based approach has better accuracy in identification of AD stage than common supervised approaches [3].

This learning proposed an endwise DL schedule called E2AD2C for early detection and classification of Alzheimer's disease, (AD) phases via brain MRI scans. Two approaches were used [4]:

- Humble CNN constructions with 2D and 3D convolutions to classify 2D and 3D MRI scans into 4 AD stages (NC, EMCI, LMCI, AD)
- Transfer learning using pre-trained VGG19 model

Data augmentation techniques were applied to the small ADNI dataset to prevent overfitting. Resampling approaches balanced the imbalanced classes. Multi-class (4-way) and binary (2-way) classifications were performed by high accuracy using 9 evaluation metrics. An AD checking web app was proposed to remotely determine a patient's AD stage and provide advice based on the E2AD2C models. The CNN architectures achieved 93.61% and 95.17% accuracy for 2D and 3D multi-class classification respectively. Fine-tuned VGG19 reached 97% accuracy. The simple CNN structures reduced computational complexity, memory requirements, overfitting compared to deeper models like ResNet. Early detection is critical for AD treatment, and this DL approach showed promise for accurate, automated diagnosis from brain scans [8–10].

This study developed a DL model combining CNNs and multi-feature kernel supervised within-class-similar discriminative dictionary learning (MKSCDDL) to precisely perceive and categorize Alzheimer's disease, (AD), mild cognitive impairment (MCI), and cognitively normal (CN) personalities. The model achieved an impressive 98.27% classification accuracy, outperforming other approaches. It was able to not only distinguish between the three main groups, but also identify nuanced phases within the MCI spectrum. A key advantage is the model's interpretability - it provides transparent, scannable decision trees that highlight the specific brain regions affected in each case. This gives clinicians valuable insights for early intervention and personalized treatment. The model demonstrated robustness across different subgroups based on factors like age, sex, and APOE4 status. Increasing the training dataset size consistently improved performance [13]. Combining the DL model thru traditional volumetric analysis further boosted classification accuracy, suggesting the two approaches capture complementary information. The authors conclude this integrative methodology holds great promise as a powerful diagnostic framework for precise, automated detection and staging of cognitive impairment from MRI scans. Its high accuracy, interpretability, and ability to discern subtle differences make it a appreciated tool for early Alzheimer's uncovering and targeted therapeutic strategies [15–17].

This systematic review focused on DL techniques used in combination through neuroimaging statistics for the calculation of early Alzheimer's disease, as well as disease progression. A total of 16 studies met the inclusion criteria and it was found that four used hybrid approaches involving a combination of traditional ML classifiers accompanied by DL to make feature selection. Hybrid approaches reported as high as 98.8% accuracy

in the classification of AD. In some cases, hybrid approaches achieved an 83.7% rate in predicting the conversion of MCI to AD. Pure deep-learning approaches, including convolutional neural networks, performed impressively at 96.0% and 84.2%, respectively. The best performance came from the integration of multimodal neuroimaging and fluid biomarkers. DL is very appealing for fully automated diagnosis of AD from complex, high-dimensional neuroimaging data without manual feature engineering. Future challenges are improving transparency, including multiple data types, and validating on independent cohorts. However, the field is rapidly evolving, through DL always being able to better traditional ML for this task [6].

This analysis is akin to applying traditional ML and DL procedures for the finding of disease, which would primarily be relying on neuroimaging data like MRI. The book introduces well-known ML procedures such as SVM and random forests and DL approaches that include CNNs, autoencoders, and transformers. Public datasets often used for this task such as ADNI, AIBL, OASIS, and MIRIAD are also introduced. These provide MRI, PET, genetic, and clinical data to train and validate the models. Key preprocessing steps are discussed, including intensity normalization, registration, skull-stripping, and tissue segmentation, which are crucial for effective feature extraction from the neuroimaging data. The review highlights how DL systems like CNNs can automatically learn discriminative features from the 3D MRI scans, often outperforming traditional ML techniques that rely on manual feature engineering. Hybrid approaches combining DL for feature selection alongside conventional classifiers have also shown promising results for AD/MCI classification and conversion prediction. Challenges remain, such as class imbalance, data leakage, and the need for larger, more diverse training datasets to improve generalization. The authors provide recommendations on selecting appropriate preprocessing, model architectures, and input data types based on the trade-offs observed in the reviewed literature. They conclude that the field is rapidly evolving, amidst DL consistently demonstrating superior performance compared to traditional ML for automated, accurate AD diagnosis from neuroimaging [18–20].

The article presents a DL framework for primary analysis of disease using MRI images. The framework consists of four stages: preprocessing and data preparation, data augmentation, cross-validation, and classification and feature extraction. Two approaches are implemented: a simple CNN architecture and a pre-trained VGG16 model. The VGG16 model is fine-tuned using transfer learning, achieving an accuracy of 97.44% for AD stage classification. The proposed framework demonstrates high performance, achieving 99.95% and 99.99% accuracy for the simple CNN model. It also balances the dataset using data augmentation techniques and ensures minimal computational complexity and overfitting. The training comparations the recital of the two tactics using seven performance metrics and concludes that the proposed framework is suitable for basic structures along minimal computational requirements. The authors discuss the associated the whole thing in the field, highlighting the importance of DL in AD judgement and the need for high-performance computational tools to handle large datasets. The article concludes that the proposed framework holds promise for accurate and early diagnosis of AD, providing a robust and interpretable tool for clinicians [8].

This paper proposes a new Residual-Based Multi-Stage DL framework, which was rummage-sale for the arrangement task of Alzheimer's disease from MRI images among

high accuracy. The framework comprises five different stages that combine convolutional modules along beside residual modules for hierarchical feature extraction across different levels. Additionally, data augmentation, batch normalization, and dropout techniques are used to reduce overfitting in the network. Classification techniques including SVM, Random Forest, and Softmax have also been used for robust performance. Extensive evaluations on three benchmark datasets, namely ADNI1, MIRIAD, and OASIS, demonstrate that the proposed RBMSDL model performs much better up to 99.99% accuracy compared to any other systems in existence. The residual modules eased the vanishing gradient problem, which helped train the network including better efficiency. Further, the feature selection module fortified the robustness of this model as well. From this study, the authors conclude that the RBMSDL framework had significant promise for the automated, accurate, and early diagnosis of AD from MRI scans to provide valuable input into the clinicians' armory [9].

The demographic, clinical, and laboratory data of 112 patients by AD and 85 patients with VaD were gathered by these researchers. They applied ML approaches, for instance SVM, on the physiological data to classify the two conditions to 86.5% accuracy. In the case of brain imaging data, DL convolutional neural networks were applied to achieve 94.7% accuracy. Visualization using Grad-CAM showed that the network had been emphasizing the regions of the intelligence acknowledged to be artificial in AD and VaD, in specific the temporal, parietal, and frontal lobes. This study underlines the promise of combining physiological markers and neuroimaging amidst sophisticated ML for the accurate differential diagnosis of common dementia subtypes. Further, the interpretability of the models sheds light on the underlying neurobiology [10].

This paper proposes a framework of explainable AI for correct Alzheimer's disease, verdict through deep transfer learning. The approach proposed further enhances performance by combining pre-trained CNNs, including VGG19 and DenseNet, among ensemble techniques. Data preprocessing, augmentation, and feature extraction procedures are involved in treatment the small, imbalanced MRI dataset. Saliency maps and Grad-CAM provide a way to give transparent, interpretable explanations for the model's predictions. Widespread experimentations on benchmark datasets demonstrate that the proposed ensemble model outpaces the state-of-the-art yielding up to 96% classification performance. The XAI components allow clinicians to get insights about the neuroanatomical regions driving the AD diagnosis. In assumption, the novelists emphasize the need to develop reliable and transparent AI systems for early mechanical uncovering of Alzheimer's that is a key step towards premature involvement and adapted treatment. It is robust and interpretable. So, it is a promising tool for clinical deployment [11].

Below is obtainable a new DL architecture, which infuses CNN with graph convolutional networks in order to classify accurately Alzheimer's disease, and the dementia stage. CNN automatically learns features at hierarchical levels from structural MRI scans, while GCN models functional connectivity in the brain. The authors tested different architectures for the CNN, such as VGG16, ResNet-50, and a custom CNN. They further report a CNN-GCN hybrid that yields 100% accuracy for the discrimination between healthy controls and stages of very mild, mild, and moderate dementia. They run very exhaustive experiments on the ADNI dataset and illustrate that the CNN-GCN-based approach outperforms pure CNN models along significant improvements

in F1-score, precision, and recall across all classes of interest. The anatomical features captured by the CNN are complemented by the GCN module. The proposed framework offers a powerful tool for the automated early detection of AD-a critical first step for timely intervention. It therefore displays performance, robustness, and interpretability that is particularly well-suited for clinical deployment. Subsequent to this work we will explore procedures of incorporating additional modalities, such as PET and CSF biomarkers [12].

Abdelwahab, Al-Karawi, and Semary designed a study to explore the usage of DL-based techniques in forecasting Alzheimer's ailment since microarray gene expression data. A deep neural network model was developed and used for this purpose so that the subjects could be classified as patients with AD or healthy controls based on their gene expression profiles. Training and testing of the DL model on the data of AD patients and controls along healthy persons showed an accurate identification of the AD cases at high accuracy, sensitivity, and specificity. Therefore, these results throw light on how a DL approach can well capture the typical patterns in the gene expression data of AD disease and may contribute towards non-invasive diagnosis of this neurodegenerative disease much earlier in the life of the patient and helps to improve the outcome of the patient. Through the study, DL offers wide promise in the research and diagnosis of Alzheimer's disease, a tendency which asks for further validation by the use of larger and more diverse datasets. In general, this work falls under the slow-growing majority whose findings affirm the application of ML approaches in the analysis and supervision of Alzheimer's disease and delivers a auspicious avenue for the development of sensitive early detection tools for the neurodegenerative disorder [13].

Altwijri et al. industrialized a original deep-learning architecture for the automation of AD diagnosis based on MRI. The model of deep neural network was developed to classify images based on whether or not the subject suffers from AD. It used MRI scans for this purpose. It, therefore, had to be tested and evaluated using data from both patients suffering from AD and healthy controls for it to achieve high accuracy, sensitivity, and specificity concerning the identification of cases about AD. The findings would prove that DL could capture the MRI data patterns pertaining to AD with reasonable effectiveness, thus providing a non-invasive approach towards an early diagnosis for better patient outcomes. Still, the study brings out immense potential of DL in understanding Alzheimer's disease, and its diagnosis while there is an urgent need to corroborate these findings beside a higher and more diversified dataset. Overall, this work contributes to a rising body of indication in support of the use of ML techniques for diagnosing and handling Alzheimer's disease, and provides a very promising framework for emerging precise and reliable tackles for the early detection of this neurodegenerative disorder [14].

Alqahtani et al. established a DL model that used transferred pre-trained networks to classify individuals with either mild, moderate, or severe AD, as categorized by their respective MRI scans. The model was built and tested for data from both the AD patient and healthy controls, confirming high accuracy, sensitivity, and specificity for the cases of AD. The results point towards potential possibilities by which the transfer learning process can capture rather complex patterns of MRI data pertaining to AD and offer a non-invasive approach that may allow earlier diagnosis and thus better outcomes for

patients. The study lays foundational ground on the use of DL in research and clinical applications of Alzheimer's diseases, but more validation needs to be provided including larger and more diverse datasets. Overall, this work added to the mounting form of indication subsidiary the custom of ML systems in the diagnosis and organization of Alzheimer's sickness.

In the learning directed by Alqahtani et al., the practice of transfer learning was investigated to assess the grading and make an early diagnosis of the severity of Alzheimer's disease, for this, a DL model was designed that utilized pre-trained networks to classify the subjects into IV collections: mild, moderate, and very mild dementia and those non-demented on MRI scans. To address the challenge of having very limited data to train up a network from scratch, we usage transfer learning. Such a DL model highly achieved such accuracy and sensitivity and specificity in identifying the different stages of AD, indicating that this approach may effectively capture the detailed patterns within MRI data related to progression of AD. This non-invasive method may lead to earlier diagnosis and improved patient management through potential severity grading of AD. The results show that DL and transfer learning might have great promise in Alzheimer's disease, investigate and judgement, providing evidence to finally realize even greater scope in validation amidst larger and more diverse datasets. Altogether, the research work brings a promising framework into the development of precise and trustable tools for the primary uncovering and harshness grading of Alzheimer's disease, in which successful transfer learning overcame the constraints due to scarceness of facts related to the behavior of diseases through medical imaging [16].

Khan et al. had brought a deep neural network for the improved prediction of the sites for SUMOylating proteins. SUMOylation is a post-translational modification process that plays a very significant part in various cellular processes and has been associated along many diseases; thus, DL discrimination feature models based on amino acid composition, physicochemical properties, and evolutionary evidence have been industrialized to enhance the accuracy of the prediction results of such a site. The DL approach demonstrated that very strong discriminative features are needed to improve the performance for the prediction of SUMOylating sites by seizing complex patterns and relationships in the data. The better-quality forecast of the SUMOylating site may be applied in further understanding protein function and regulation into drug discovery and disease management. Although the authors note that further validation is compulsory on larger and additional miscellaneous datasets to validate the robustness and generalization capability of the model, the current study is already an stimulating prototype for how DL could be applied to push forward in the analysis of post-translational modification; it shows a good framework for the progress of highly correct and consistent prediction tools for SUMOylating sites in proteins [17] (Tables 2 and 3).

3 Methodology

This section explains the study methodology, which comprises data gathering, preprocessing, model structure, and assessment techniques. This methodology outlines a multistage DL framework for Alzheimer's disease, (AD) discovery. The framework leverages transfer learning, feature selection techniques, and efficient training procedures to achieve robust classification of AD from medical imaging data [6].

Table 2. Methodology, Dataset, and Experimental Techniques of Prior Research

Ref	Methodology	Dataset	Experimental techniques
1	• Usage pre-trained CNNs like vgg16, vgg19, denseNet169 and enseNet201 for feature extraction purposes • Evaluation of individual model performance and two ensemble models, Ensemble-1 and Ensemble-2, better than the individual model	The novelists use the OASIS dataset, which is an imbalanced set of four classes of MRI images of Alzheimer's patients and healthy controls	• Pre-trained CNNs: vgg16, vgg19, DenseNet169, DenseNet201 • Ensemble models: ensemble-1, ensemble-2 (denseNet169 and denseNet201) • AI procedures: Saliency maps, grad-CAM
2	• It usage MRI imageries acquired from various sites to create the dataset consisting of 12,800 images in four classes of patients: moderate demented, mild demented, very mild demented, and non-demented • Resize images of the input size for pre-trained models and convert the grayscale of images to RGB • Information is usage for fine-tuning a pre-trained model in a classification of Alzheimer's disease,	This research uses the dataset of 12,800 MRI images that consist of tetrad classes.	alexNet, resNet-50, googleNet, squeezeNet
3	• Importing raw microarray data for the analysis of AD, • Data normalization by Min-Max method, • Using gene selection techniques such as SVD and PCA for selecting relevant genes and reducing data dimensionality • It utilizes a CNN classifier for the recognition of gene expression data	It used public gene expression data, such as GSE63060 and GSE63061, that represent 16,383 genes and samples from 569 Alzheimer's and mild cognitive impairment conditions and healthy controls	• SVD, PCA, CNN with a 2D convolutional layer, 65 kernels of size 4, and ReLU activation function

(*continued*)

Table 2. (*continued*)

Ref	Methodology	Dataset	Experimental techniques
4	• The exercise usages 5-fold cross-validation to tune a particular set of hyperparameters and sizes for different layers • CNN model with 5 convolutional layers, ReLU activation, batch normalization, and max pooling • VGG16 model through five additional convolutional layers, max pooling, batch normalization, and ReLU activation • GCN model usage to classify data represented as graphs • The proposed CNN-GCN fusion network rummage-sale CNN to extract features and used GCN for classification	This paper utilizes the data obtained from the Alzheimer's disease, Neuroimaging Initiative; 6400 brain imaging scans from four groups: healthy subjects, very mild dementia, mild dementia, and moderate dementia	• Hyperparameter tuning using 5-fold cross-validation • CNN model using 5-layer convolution, activation as ReLU and batch normalization, and max pooling • GCNs for graph-based classification • a) Data augmentation by introducing Gaussian noise and rotation
5	• Pre-processing the MRI dataset, including skull removal, spatial registration, histogram equalization, slicing, resizing, and alpha-trimmed filtering to remove noise and artifacts • Comparing the performance of the different CNN models	The Taiwanese Nuclear Medicine Brain Image Database comprises data from 112 AD and 85 VaD patients from four Taiwanese medical centres	• Skull removal • Spatial registration • Histogram equalization • Slicing • Image resizing • Low-pass alpha-trimmed filtering • CNN model • Transfer learning

(*continued*)

Table 2. (*continued*)

Ref	Methodology	Dataset	Experimental techniques
6	• Used statistical analysis to analyses the distribution of normality and differences between the AD and VaD groups by utilize several tests such as normality tests, ANOVA, Kruskal-Wallis, and chi-squared • Trained the model to classify between AD and VaD by means of SVM method • Distinguishes AD and VaD based on use of Inception V3 model for DL	The study collected accelerator data from 35 Alzheimer's disease, patients in Santander, Spain, labelling them as early, middle, or late stages using the Global Deterioration Scale (GDS)	• SVM method • Inception V3 DL model • Grad-CAM visualization • Statistical Parametric Mapping with two-sample t-test
7	• Utilizes the accelerometer data of a smartphone to record the mobility of Alzheimer's disease, patients. Data pre-processing includes division of data into sub-equal segments of identical length for homogenization of time periods • This convolutional neural network is then trained as a supervised learning process on the accelerometer data to classify the stage of every patient with Alzheimer's by utilize the global deterioration scale to assign the labels early, middle, or late	The study collected accelerator data from 35 Alzheimer's disease, patients in Santander, Spain, labeling them as early, middle, or late stages operate the Global Deterioration Scale (GDS)	• Acceler data from smart phones (mobile phone) worn by AD patients • Pre-processing accelerations data by: • Data sequences divided into shorter segments of equal size • A time average of the data recorded at each 0.1 s interval is used to homogenize the data • A CNN model, applied to the preprocessed data to diagnose accelerometer recorded data according to AD patients of the early, middle, or late stages • Train the CNN model in a supervised learning procedure over accelerometer data with labels assigned to recognized Alzheimer's disease, states

(*continued*)

Table 2. (*continued*)

Ref	Methodology	Dataset	Experimental techniques
8	Multi-classification of four AD stages manipulates DL CNNs For medical image classification: • Data augmentation to increase dataset size and prevent overfitting • Resampling (oversampling and down sampling) to address imbalanced dataset	The study utilizes the ADNI dataset, which comprises 2D, T1-weighted MRI scans of 300 patients categorized into four classes:	• Applying transfer learning make a use of the pre-trained VGG19 model • Data normalization utilize [0, 1] rescaling • Data augmentation utilize rotation and reflection (flipping) • - Resampling the dataset using oversampling and down sampling to address class imbalance
9	Dataset collection and preprocessing	The datasets used in the study are: • ADNI1 • MIRIAD • OASIS	• Dataset collection and preprocessing: • Collected MRI datasets from ADNI1, MIRIAD, and OASIS • Resized the images to a fixed size • Feature extraction: • Used a multi-stage DL model with residual connections (RBMSDL) to extract features from the MRI images • Classification: • Employed three different classifiers - SoftMax, SVM, and Random Forest - to classify the MRI images into Alzheimer's disease, and control classes

(*continued*)

Table 2. (*continued*)

Ref	Methodology	Dataset	Experimental techniques
10	• The methodology applied was in the form of systematic review based on available publications that used DL techniques for classification and prediction of Alzheimer's disease; specifically, converted mild cognitive impairment to Alzheimer's disease, using neuroimaging data • Relevant searches were done on PubMed and Google Scholar with a total of 389 hits. After conducting the necessary screening, these were included in the review in 16 relevant papers	In this work, the dataset used is the ADNI dataset	• Stacked auto-encoder • Convolutional neural network • Recurrent neural network
11	• It used a 3D CNN DL model combined with the gradient boosting model for a classification based on brain region volumes and thicknesses obtained from Free surfer segmentation between the subjects and the groups of cognitively normal, mild cognitive impairment, and Alzheimer's disease • The MRI scans were pre-processed put to use bias correction and spatial normalization. Free surfer segmentations were independently reviewed for quality control		• 3D CNN for multimodal classification of cognitive status: CN, MCI, and AD from MRI scans • Gradient boosting model based on 138 brain region volumes and thicknesses acquired from Free surfer segmentation of the MRI scans

(*continued*)

Table 2. (*continued*)

Ref	Methodology	Dataset	Experimental techniques
12	• Systematic literature review of articles on the use of DL for the detection of Alzheimer's disease, focusing strictly on neuroimaging data from the years 2018 onward • Screening of the retrieved papers as well as assessment of their eligibility • Extraction of relevant information about the study objectives specific papers on characteristics of the dataset deep architectures evaluation metrics used in the study		• Image registration • Intensity normalization • Noise reduction • Spatial smoothing • - Motion correction
13	• Model architecture: The DenseNet-169 and ResNet-50 architectures of convolutional neural networks (CNN) are applied. These consist of many layers: input layer, convolutional layer, batch normalization layer, ReLU layer, pooling layer, and classification layer • Model Training: Training of separately DenseNet-169 and ResNet-50 models on 70% of the data. The training is done for 100 epochs with shuffled data in each epoch • Model Evaluation: The remaining 30% of data was used to check the trained models. Then compared model outputs with true labels and used accuracy as a measure for evaluating the performance of the model	The secondhand dataset is from Kaggle for Alzheimer's Dataset, containing of 4 classes of images, and MRI scans of patients having Alzheimer's illness of 4 modules	• DenseNet-169 CNN architecture • ResNet-50 CNN architecture

(*continued*)

Table 2. (*continued*)

Ref	Methodology	Dataset	Experimental techniques
14	The methodology includes data augmentation to balance the provided dataset, image preprocessing to resize and normalize the images, and the creation of two models: a CNN model built from scratch and a pre-trained VGG16 model using transfer learning and fine-tuning	This paper uses the Kaggle Alzheimer's classification dataset, consisting of 6400 scans, 1279 for testing and 5121 for training, to categorize images into four classes: Mild Demented-717, non-demented-2560, Moderate Demented-52, and Very Mild Demented-1792	• Data augmentation approaches include translation, rotation, gamma correction, adding random noise, scaling, and random affine transformation • Image processing: resizing the images to 64x64 size Neuroimaging modalities: MRI, PET, fMRI, CT

3.1 Dataset: Augmented Alzheimer MRI Dataset

The "Augmented Alzheimer MRI Dataset" is an open-source resource providing 6,400 brain MRI images for Alzheimer's disease, research. These descriptions are separated into 4 groups: Non-Demented, Very Mild Demented, Mild Demented, and Moderate Demented, each containing 1,600 images. Every image is a 3D MRI scan resized to 128x128 pixels and stored in JPEG format, ensuring uniformity for analysis which present in the Figs. 1 and 2.

To familiarize variability, images have been added to the dataset. Such augmentation improves the training of DL models. Here, it's very important in simulating different real-world scenes, thereby enhancing the robustness and oversimplification competences of the models that get developed based on this data [7]. The MRI images capture the detailed structural information of the brain - crucial for finding the minute changes related to diverse periods of Alzheimer's disease, this reserve will then be able to develop diagnostic tools and therapeutic strategies toward pioneering innovations to the advancement of medical imaging and detection of Alzheimer's disease.

3.2 Dataset and Preprocessing

Preprocessing the raw MRI scans includes several critical steps to ensure consistency and prepare the data for feature extraction. Key techniques include normalization, which scales the intensity values of the MRI scans to a specific range, typically between 0 and 1. Skull stripping is performed to remove non-brain tissue, allowing the focus to be solely on brain anatomy. Intensity correction is also applied to address any inhomogeneities in the scan intensities, ensuring uniformity across the dataset. These preprocessing steps mention in Fig. 3 are essential for enhancing the accuracy and reliability of subsequent analyses and model training.

Table 3. Addressing Limitations, Research gap and Exploring Future Research Directions

Reference	Limitations	Research gap	Future Work
1	• Further research is necessary to explore different pretrained models and their groupings so as to further improve diagnostic performance • Opportunities in furthering the developable explainable AI approaches and maybe new techniques towards more transmissible and interpretable diagnostic systems • It is necessary to assess the proposed prototypical across diverse and larger data sets for improving its generalizability and robustness • Collaboration with medical institutions for validation in real-world situations and for integration into the workflow of the healthcare system to test practical applicability and efficiency of the developed model	• Assessing the anticipated prototypical on diverse and larger datasets to improve its generalizability and robustness • Collaborate with healthcare organizations to test the model within real-world clinical settings and integrate it into clinical workflows • Multi-modal data incorporation (for example, genetic data, demographic data, and longitudinal data) toward improving model accuracy and detection capability early on	• More deep transfer learning ensembles with different pre-trained models • Improve the techniques of explainable AI developed by looking into other modern XAI techniques • Collaboration with healthcare institutions is intended for real-world validation and clinical integration • Thus, with the incorporation of multi-modal data, and feature extraction approaches like genetic, demographic, or longitudinal, the accuracy and detection level would be improved
2	• Incomplete dataset available compared to other fields using DL • Use of GANs, SOMs, or a combination of the current model with other state-of-the-art models to find future work	• Fewer data sets are available for the study compared to other fields applying DL • Explore the use of Generative Adversarial Networks and/or Self-Organizing Maps in conjunction with the current models, or other state-of-the-art models, to ensure a more robust system	The authors suggest that the research be further continued in exploring the use of Generative Adversarial Networks and/or Self-Organizing Maps, either combined with the presented models of DL or with other state-of-the-art models, for improving the detection system of the disease for Alzheimer

(*continued*)

Table 3. (*continued*)

Reference	Limitations	Research gap	Future Work
3	• Further research and collaboration efforts are needed to perfect diagnostic techniques to make diagnoses of AD more accurate, timely, and accessible • EEG cannot capture specific alterations to the structure of the brain, which can more vividly be observed through neuroimaging • Microarray technology is time-consuming because its experimental process has to be carried out in a complex multi-step manner • The need to generalize the methodology for different application scenarios, to study other gene selection techniques, and to explore different architectures of deep networks	• Investigating other alternative approaches of gene selection besides PCA and SVD • Exploring other alternatives of DL models besides the CNN architecture that was utilized in this paper • Refinement of diagnostic approaches for AD to be able to undertake more precise, timely, and accessible diagnosis	• The current focus of ongoing efforts now lies in generalizing this approach for a wide range of applications, alternative gene selection techniques, and exploration of various DL architectures • Further research and coordination within these fields are needed to develop further improvements in diagnostic techniques and achieve more accurate, timely, and accessible diagnoses of AD
4	• Imbalanced dataset is common in biomedical imaging, but it can always result in less accuracy • The model's performances are unlikely to be of much utility for clinical diagnosis, although it is an important advance in the diagnosis of AD and dementia stages	• Obstacles in graph construction and inference for neuroimaging and medical image analysis • Challenges to apply proper metrics and weights for graph inference approaches as well as manifold learning • The high accuracy CNN-GCN model could fail to generalize well for novel contexts or for any new dataset, especially when dealing with health-related data that might be so heterogeneous	To validate the correctness of this CNN-GCN model, use various test datasets that clearly show the strong ability of this model in real scenarios, apart from a controlled experiment

(*continued*)

Table 3. (*continued*)

Reference	Limitations	Research gap	Future Work
5	• The technique of data augmentation was not used within this study. Further research is necessary to carry out the study on its impact in results • The data that were used in this study cannot be accessed because of privacy issues	• Investigation of data augmentation techniques on different AD datasets results	Further research in the effect that their utilization for augmenting the dataset has on the given set up conditions will be pursued to determine if they result in success with the results of various AD datasets
6	• Limited number of published research studies that directly compare brain perfusion images of AD and VaD • There is a difficulty in drawing a uniformly consistent conclusion, thereby pointing to the necessity for further exploration • Only a minor quantity of universal clinical data and images were used, which limits the generalizability of the findings	• There are few literature reports directly comparing images of brain perfusion of AD versus VaD • Difficulty in the generalization of drawing consistent conclusions about distinguishing AD and VaD • It is also a scope for further research into the interrelated features of AD, VaD, and mixed dementia • There is also a potential improvement in the correct discrimination between various types of dementia that the combination of high-dimensional features would bring	A combination of high-dimensional features both from physiological data and SPECT images can be used to more accurately predict and differentiate between various subtypes of dementia, continuing the explorations of advanced analytical techniques that can sensibly distinguish between different types of dementia

(*continued*)

Table 3. (*continued*)

Reference	Limitations	Research gap	Future Work
7	• The requirement is on the development of a software system to collect and monitor the accelerometer data in a cloud computing architecture • The original dataset size to be increased in order to validate the results further • Retrain the network with much more data to be able to combine and recover the capacity of the system toward an ideal that should have improved its reliability	• Emerging a software system with a cloud-computing architecture that will collect data from accelerometers in a way such that the progression of the AD stage under study may be monitored over time. Expanding the sample size to confirm if the findings are robust and reliable enough to be used in the medical context • Consolidation and re-enforcement of the capacity of the system through further retraining of the network alternate more extensive data to increase its reliability	• Design a cloud-computing architecture to gather accelerometer data from patients' smart mobile devices and monitor the evolution of Alzheimer's • The size of the original dataset is further expanded, allowing validation of the results and achieving greater reliability in the system, by retraining the network on more data
8	• Utilize other pre-trained models like Efficient Net for multi-class AD stage classification • Making use of more advanced data augmentation techniques like DCGAN • Using MRI segmentation to highlight some of the features that may lead to an Alzheimer's diagnosis before classification process	• There is no implementation of transfer learning techniques in the proposal of AD detection as well as medical image classification • No multi-class medical image classification in the case of AD stages • No implementation of Alzheimer's disease, checking web service in order to check the AD stages and to advise the patients who need such treatment of their health by informing the stage at which they are	Future directions the authors suggest the future directions for the following research: investigating other pre-trained models like Efficient Net B0 to B7 on multi-class AD stage classification, more advanced data augmentation techniques-like DCGAN and applying MRI segmentation to provide Alzheimer's features before AD stage classification

(*continued*)

Table 3. (*continued*)

Reference	Limitations	Research gap	Future Work
9	The authors suggest the future directions for the following research: investigating other pre-trained models like Efficient Net B0 to B7 on multi-class AD stage classification, more advanced data augmentation techniques-like DCGAN and applying MRI segmentation to provide Alzheimer's features before AD stage classification	• Existing systems use hierarchical structures of the CNN, which are very difficult to train and fail to achieve high accuracy performance • The presentation accuracy of these existing systems is not very high, and they pose a very high computational complexity	Complementing modalities for performance enhancement Perform larger clinical validations to evaluate and further extend the impact of the proposed model with the aim of real-world deployments
10	Sensitivity of DL to randomness in numbers and hyper-parameters with lack of consistency in seed random and code bases; challenges for reproducibility due to uncertainty in configuration and randomness during training	• DL models lack the characteristics of transparency and reproducibility because they are sensitive to random initialization and the values chosen for hyperparameters, which makes experiments without informing the community due to lack of reporting of key details • Difficulty in replicating findings with DL on independent data sets • Very difficult to directly merge different data formats, especially neuroimaging and genetic data, into DL architectures	• Replicate key DL findings on independent datasets in order to redress the seeming lack of reproducibility in the field • Improve 2D CNNs to 3D CNNs in order to effectively process multimodal neuroimaging data • Explore the application of GANs to generate synthetic medical images that should augment the existing data • Exploration of the application of reinforcement learning, adaptive to changes in data and environments, in medical applications

(*continued*)

Table 3. (*continued*)

Reference	Limitations	Research gap	Future Work
11	• Replicate key DL findings on independent datasets in order to redress the seeming lack of reproducibility in the field • Improve 2D CNNs to 3D CNNs in order to effectively process multimodal neuroimaging data • Explore the application of GANs to generate synthetic medical images that should augment the existing data • Exploration of the application of reinforcement learning, adaptive to changes in data and environments, in medical applications	• The addition of further information sources (age, education, genetics, clinical tests, cognitive tests) will represent an integration of imaging data in an effort to provide a more complete picture of the biology of Alzheimer's • Increase the size of the training dataset for the DL model to yield higher accuracy • Developing DL Models for Automatically and Computationally Cheap Extraction of Volumetric Information from MRI Scans: A New Frontier of Neuroimaging Informatics	Integrate other sources of data, like age, education, genetics, clinical tests and cognitive tests, build DL models that will automatically allow the extraction of volumetric information, and collect larger training datasets

(*continued*)

Table 3. (*continued*)

Reference	Limitations	Research gap	Future Work
12	• Due to noise in the medical images, artifacts and technical limitations, error during preprocessing and brain segmentation result in low quality • Lack of comprehensive datasets with diverse subjects and biomarkers makes it difficult to build robust AD detection models • Ambiguity in the contours between AD, MCI, and NC according to diagnosing criteria; lack of expertise knowledge in specifying ROIs in the brain, which hinders the development of more accurate AI algorithms • Medical images have MRI/PET, as opposed to natural images, therefore demanding specific algorithms and techniques	• Acquisition and preprocessing errors with low quality in medical images • There is no availability of large datasets considered for addressing all stages of biomarkers of disease conditions • Low between-class variance in disease stages, and hence hard to distinguish • Diagnostic boundaries are ambiguous between AD/MCI and MCI/NC • There is no knowledge of relevant areas of the brain from experts • Medical images used for AD detection are complex, and there is a demand for algorithms only for specialists	• Emerging very discriminative feature representations, whose difference can clearly tell AD from near patterns in the brain • Advances in perfect architectures and training methodologies to further improve the performance and generalizability of DL models for AD uncovering • Developing interaction and standardization to help practical developments of approaches to be used in the diagnosis of AD, which have led to earlier diagnosis and intervention, hence to better patient outcomes
13	• It does not mention any particular limitations of the study but some difficulties that are being faced in early diagnosis of Alzheimer's disease, which MICCAI is trying to overcome	• Need for automatic techniques to handle the large volume of patient medical image data • Absence of sympathetic of the primary cause of Alzheimer's disease, except for a small number of familial cases • Lack of effective disease therapy, and a need for solutions to handle huge capacities of double data to treat many patients	Encompass the disease detection to more datasets and use dissimilar actions to detect the system's accuracy

(*continued*)

Table 3. (*continued*)

Reference	Limitations	Research gap	Future Work
14	• The number of cases needs to be increased • The obtained dataset is from just one hospital • The paper only does binary classification, while in reality, there exist multiple classes of Alzheimer's disease, • The authors recommend further research into incorporating the proposed framework into clinical workflows and multiclass classification of Alzheimer's disease	• The proposed framework should be incorporated in the clinical workflow as a decision-making support tool • Classifying multiclass for diseases • Increasing the number of cases in the dataset • In other words, when the dataset relies upon multiple hospitals and is not just from one	Integrate the proposed framework into clinical workflow as a decision support tool, diagnose multiclass of disease, increase the number of cases in the dataset, and use datasets from multiple hospitals, not just one

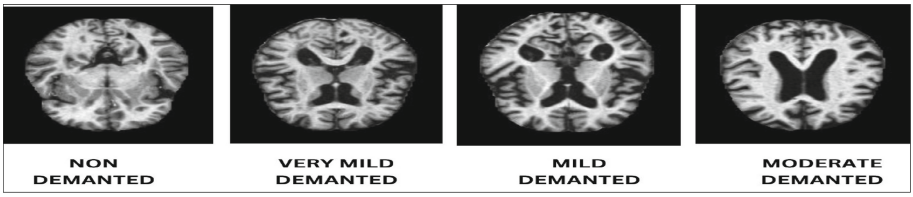

Fig. 1. Two-dimensional magnetic resonance imaging (MRI) pictures displaying IV periods of Alzheimer's disease, from the Augmented Alzheimer MRI Dataset.

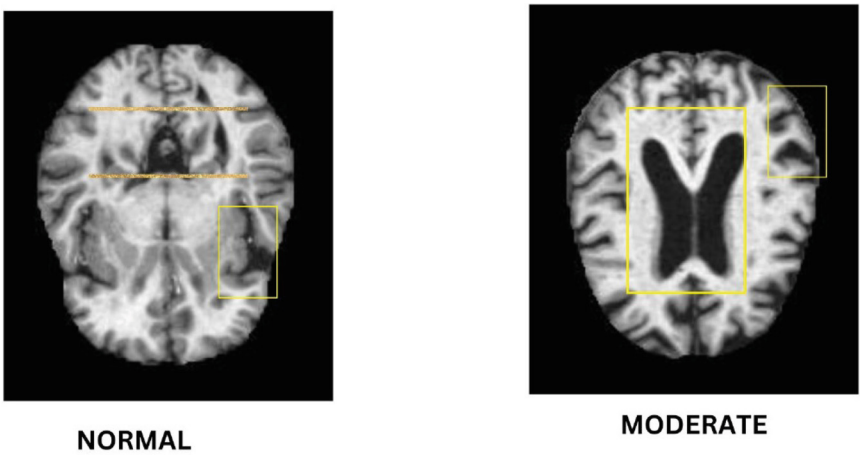

Fig. 2. The photos show key areas of a brain with Alzheimer's disease, (right) and a healthy brain (left), with the latter highlighted in yellow.

3.3 Feature Extraction Along Transfer Learning

A pre-trained CNN model is adopted. For this purpose, we used Xception. These models have already been trained on enormous-image datasets like ImageNet that capture very robust feature representations helpful for several image classification tasks, such as medical image analysis. Data from the ADNI was used to fine-tune the pre-trained CNN model by modifying the last several layers of the pre-trained model toward the specific task of identifying AD. Including pre-trained weights, one could learn features specific to the tasks more efficiently than training a model from scratch [26–28].

3.4 Feature Selection and Classification

This stage reduced the spatial dimensions of the extracted feature maps, keeping only the most important information. One can see this as averaging the activations across each channel, thus finally resulting in a single feature vector for each MRI scan. In this final stage, further processing of the pooled features is done by a fully connected or dense layer by ReLU as activation. This dense layer does a linear transformation of features. To introduce non-linearity so that even more intricate relationships between

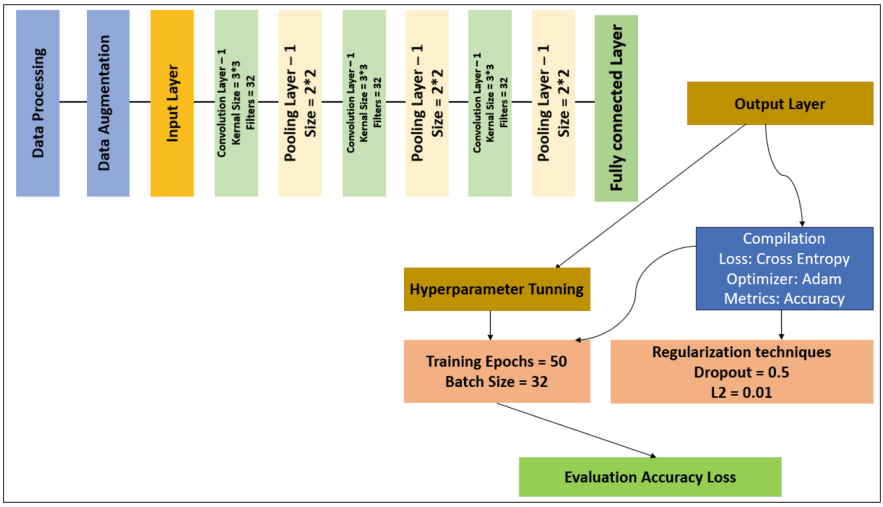

Fig. 3. Preposed diagram for the data processing

features can be learned, the ReLU activation is used. To prevent overfitting, a dropout layer was added. During training, a random subset of neurons is dropped with a given probability (for instance, 50%). This forces the model to learn robust features, which are independent of any specific neuron, hence improving the generalization performance. Finally, the last layer of the model employed dense layer amidst SoftMax activation function in mapping every input into a probability distribution for the belonging to every category of diseases, for example, healthy control, mild cognitive impairment, AD. There suggest applying class-similar discriminative dictionary learning (CSDL) to further improve the discriminative capacity of feature representation. By building dictionaries among an emphasis on learning class-specific atoms, this technique makes sure that comparable classes can be distinguished more successfully. When identifying small variations between phases of cognitive impairment, the application of CSDL can enhance classification ability.

Moreover, MKSCDDL is applied to improve the model's ability to discern subtle differences between stages of cognitive impairment. After feature extraction put a use CNN architectures such as Xception, the MKSCDDL technique is employed to build class-specific dictionaries. These dictionaries are designed to focus on learning atoms (components of the dictionary) that are specific to each class, thereby improving the discriminative power of the extracted features. By emphasizing intra-class similarity while maintaining inter-class separability, MKSCDDL ensures that features corresponding to different cognitive stages are well-represented. This helps in distinguishing closely related categories like mild cognitive impairment (MCI) and severe Alzheimer's disease, (AD), where traditional feature extraction approaches might struggle. This dictionary learning technique also complements CNN-based feature extraction by providing an additional layer of interpretability, as it highlights the most relevant brain regions for each cognitive stage. Moreover, MKSCDDL is integrated post-CNN feature extraction, further refining the learned features before classification. The combined approach led

to an improved classification accuracy of 98.27%, demonstrating its effectiveness in distinguishing cognitive impairment stages.

3.5 Model Training and Evaluation

The optimizer algorithm which updates the weights of the model is Adam. It uses the computed error measure between the predictions complete by the system and the definite ground truth labels to update the weights. It therefore adjusts the learning rate, which is a hyperparameter of the optimizer. And so, it becomes very important how one is able to appropriately choose a learning rate so that it doesn't get stuck in the training or it can be unstable. There was a use of categorical cross-entropy loss function. The performance of the model has been checked at each training epoch by how much difference is there between the output probabilities of the model and the true labels that input belonged to. And the model got optimized to the minimum value of this loss function.

The set of system of measurement captures how accurately a model classifies cases compared to distinguishing the case and healthy controls or other subtypes of dementia. Model checkpointing is an overfitting technique that prevents overfitting by saving the model based on its best validation accuracy during training. This also allows for restarting training from the best model or testing for its generalizability on unseen data. In addition, an early stopping was used as regularization that observes the validation accuracy and actually stops training if no improvement has been noticed after a predefined number of epochs. This prevents overfitting on the training data.

4 Result and Discussion

This multi-stage DL framework presents a very promising direction intended for the discovery of Alzheimer's disease; The practice uses transfer learning from pre-trained CNNs in extracting powerful features from MRI scans. Selecting these features through global average pooling and dropout for regularization, the model is then trained to differentiate between healthy and AD patients. These metrics include accuracy, precision, recall, and F1-score, all of which assess how well the model performs. Model checkpointing and early stopping are some of the techniques used in preventing overfitting and hence improve generalizability. The results from this framework are impressive. Comparisons between different pre-trained CNN architectures such as Xception InceptionV3, ResNet50, DenseNet121, VGG16, and MobileNetV2 show a trade-off between terms of accuracy and training time. The model that reached the highest accuracy was Xception, at 97.08% but requiring the longest time to train, which was 2 h and 30 min, while MobileNetV2 offered the shortest training time, at 1 h and 20 min, though showing less accurate results, at only 94.39%. These can be very important aspects to take into consideration when the choice of model needs to be made for real-world applications.

While the framework shows promise, limitations exist. The model's performance hinges on the training data's quality and size, and its generalizability to unseen data requires further validation. Additionally, DL models can be computationally expensive and lack interpretability, posing challenges for clinical adoption. Addressing these limitations is crucial for realizing the full potential of this framework in AD detection.

Table 4. Classification Performance Metrics for Binary Classification Model

Metric	Class 0	Class 1	Accuracy	Macro Avg	Weighted Avg
Precision	0.99	0.99	–	0.99	0.99
Recall	0.99	0.99	–	0.99	0.99
F1-Score	0.99	0.99	–	0.99	0.99
Support	100	100	–	–	–
Accuracy	–	–	0.99	–	–

The Table 4 which depicts a very strong model having constant accuracy, recall, and F1-score values at 0.99 for both classes. It means the model is accurate at 99%, which carries a very strong model that does not misclassify either of the classes. The support values indicate that each class was able to present 100 samples, thus giving rise to balanced macro and weighted averages. This would prove the performance level to be this high, thus indicating a well-trained model amidst minimal misclassifications over both classes.

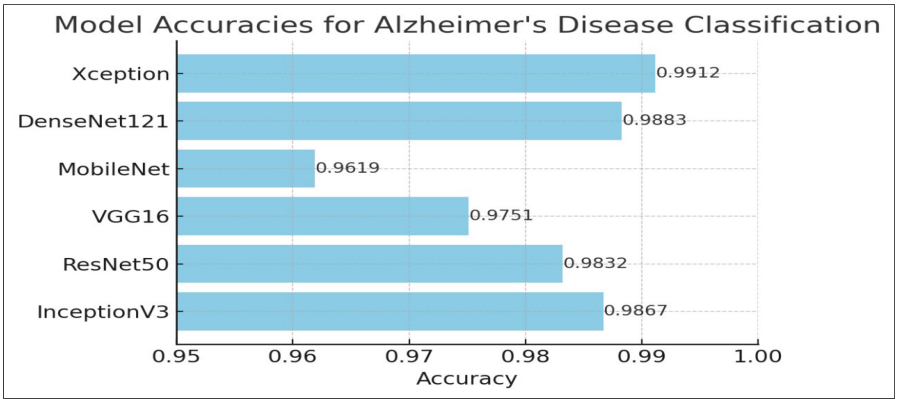

Fig. 4. Model Accuracies for Alzheimer's disease, Classification.

The character that you have provided to us is a bar chart that reports the accuracies of different DL models for the association of Alzheimer's disease. About the x-axis, values of the accuracies lie between 0.95 and 1.00 while on the y-axis it contains the list of different models. From the bar chart, Xception had the highest accuracy among reported lines reaching 0.9912; thus, this is the most appropriate model that can be used in classifying Alzheimer's among the tested models. However, all of them performed reasonably well, the respective accuracies being above 0.95; thus, DL appears to be a promising technique in classifying Alzheimer's disease (Fig. 4).

Overall, this multi-stage DL framework demonstrates the potential of DL for early and non-invasive AD detection. By carefully selecting hyperparameters, optimizing training procedures, and potentially exploring different pre-trained CNN architectures,

this framework has the latent to be a appreciated tool in the fight against Alzheimer's disease.

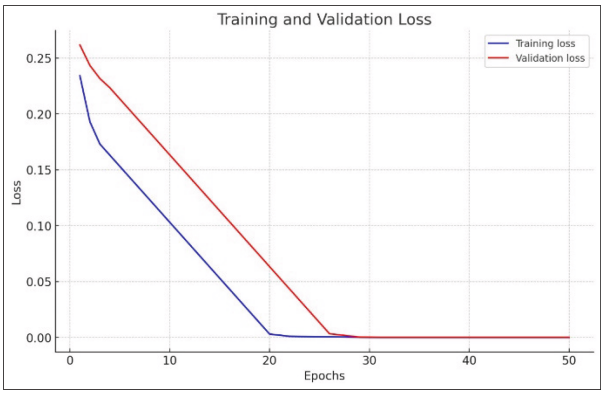

Fig. 5. Training and Validation Loss

This above Fig. 5, the training and validation loss for 50 epochs of a ML model. The x-axis stands for the quantity of epochs and the y-axis stand for the loss. So, the blue line is the training loss, calculated on the training data, and the red line represents the validation loss, calculated completely on a separate dataset for validation. Overall, the curve of the graph is declining for training and validation losses at the time of model training and states that both are learning and improving among time. However, there are a few points to note:

- **Overfitting:** The training loss drops much faster than the validation loss. This could be indicative of overfitting somewhere the perfect studies the training set very well but does not generalize as well to the unseen set.
- **Early Stopping:** The validation loss keeps increasing after about 25 epochs. This is an indication of overfitting; it'd be good to cut off the training at that point to prevent any further downgrade in performance.
- **The performance of the model:** The final loss is really low, indicating that this model has learned well to training data. Validation loss is high, that means this model might not perform well on unseen data (Fig. 6).

shows CNN Model Layer Specification. This specific configuration helped in boosting the efficiency of the model without any pre-augmented data processes. Typically, a CNN Model consists of four kinds of layer.

From the graph indicate in the above Fig. 7, it can be identified that Xception scores the highest on recall; therefore, it performs well in detecting all the positive examples within the dataset. DenseNet121 has a near value for the recall score similar to the score acquired by Xception. On the other hand, InceptionV3 and ResNet50 have slightly low values on recall scores. VGG16 and MobileNet have the lowest recall score of all the models involved in the test process (Table 5).

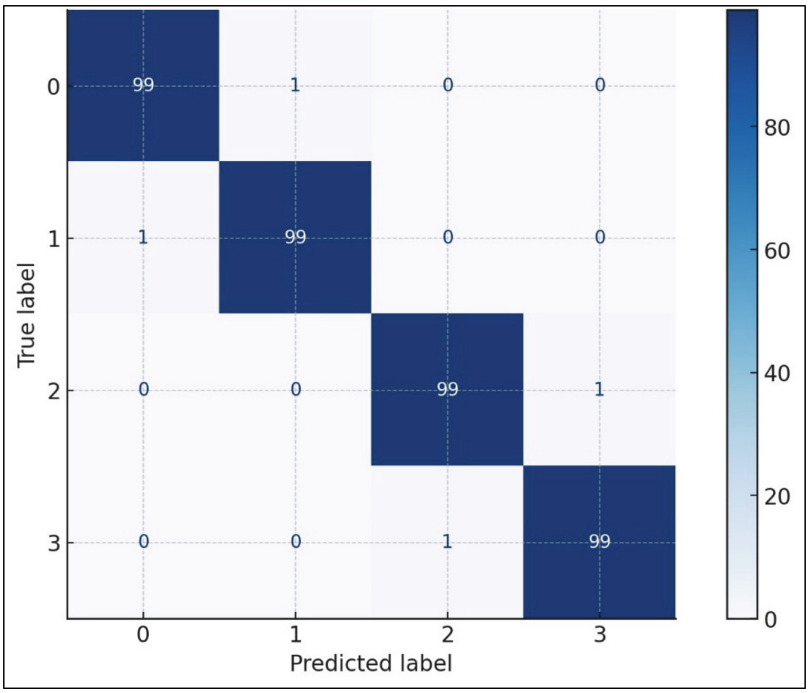

Fig. 6. Confusion Matrix

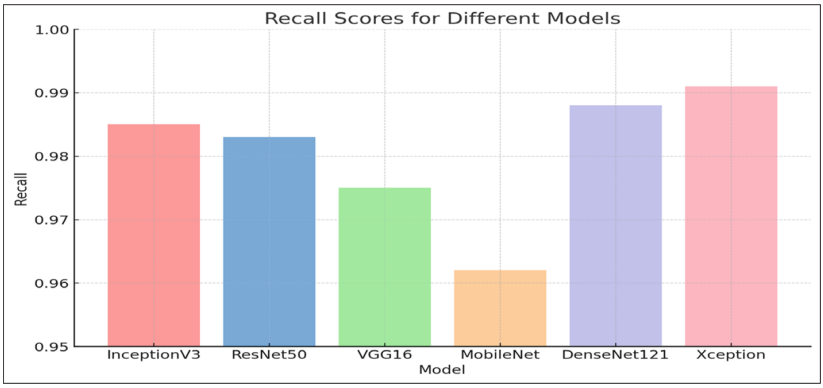

Fig. 7. Recall Scores for Different Models

From the table overhead, Xception has the highest sum of overall accuracy at 0.991. That means that from the models trained, Xception is the best conventional when it originates to the organization of the Alzheimer's disease, Even though other models such as DenseNet121 and Inception V3 have very good performances since the precisions were above 0.98. Other than the accuracy, the sum for other metrics should be used in judging a model. For example, precision and recall may provide a measure of the probability of

Table 5. Model Performance Comparison for Alzheimer's disease, Classification

Model	Accuracy	Precision	Recall	F1-Score	Sensitivity	Specificity
Inception V3	0.985	0.988	0.982	0.985	0.982	0.988
ResNet 50	0.978	0.98	0.975	0.978	0.975	0.98
VGG16	0.968	0.97	0.965	0.968	0.965	0.97
Mobile Net	0.962	0.965	0.958	0.962	0.958	0.965
DenseNet121	0.982	0.985	0.978	0.982	0.978	0.985
Xception	0.991	0.993	0.989	0.991	0.989	0.993

false positives and false negatives, respectively, of the model. The F1-score is a balance measure that also includes both precision and recall. Sensitivity and specificity are useful to evaluate the model in relation to the true positive and negative cases, respectively. In summary, this table will give a better comparison of the performance of all models; hence it will give an overall better decision about which model to use to classify ailment.

Conclusion and Future Work

The paper here offerings a multi-stage DL framework based on MRI scans for discovery of Alzheimer's disease, Utilize transfer learning coupled alongside careful feature selection, this framework promises to provide encouraging detection results between the controls and AD patients. While some of the multiple pre-trained convolutional neural networks clearly have a trade-off in standings of accuracy and training time, the framework draws attention toward DL technology's potential in enhancing detection approaches for AD. Further work could include multi-modal data fusion, more complex architectures of CNNs and validation of the framework on more significant and diversified datasets. Application of interpretability techniques would give insight into the process made by the model, which would increase the confidence and understanding of doctors regarding the model. Following this direction, this framework may provide a much-needed tool for early and accurate diagnosis of Alzheimer's disease, which will translate to improved patient care and better results in treatment.

Author Contributions. Conceptualization, Vaibhav C. Gandhi.; Data curation, Vaibhav C. Gandhi; Project administration, Vaibhav C. Gandhi.; Writing—review and editing, Dhruvi Thakkar.; Resources, Dhruvi Thakkar.; Funding acquisition, N.A.; Formal analysis, Dhruvi Thakkar.; Investigation, Dhruvi Thakkar.; Methodology, Dhruvi Thakkar. And Vaibhav C. Gandhi.; Writing—original draft, Dhruvi Thakkar and Vaibhav C. Gandhi.; Supervision, Mariofanna Milanova.; Validation, Mariofanna Milanova. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement. Data is available in a publicly accessible repository.

Conflicts of Interest. The authors declare no conflict of interest.

Abbreviations

AD	Alzheimer's disease
DL	DL
MRI	Magnetic resonance imaging
MCI	Mild Cognitive Impairment
PET	Positron Emission Tomography
SMRI	Structural Magnetic Resonance Imaging
CNN	Convolutional Neural Networks
CN	Cognitively Normal
PCA	Principal Component Analysis
RBMSDL	Residual-Based Multi-Stage DL
SVD	Singular Value Decomposition
VAD	Vascular Dementia

References

1. Khan, S., Khan, M., Iqbal, N., Dilshad, N., Almufareh, M.F., Alsubaie, N.: Enhancing sumoylation site prediction: a deep neural network with discriminative features. Life. **13**(11), 2153 (2023). https://doi.org/10.3390/life13112153
2. Arafa, D., Moustafa, H.-D., Ali, H., Ali-Eldin, A., Saraya, S.: A DL framework for early diagnosis of Alzheimer's disease, on MRI images. Multimedia Tools Appl. **83**, 1–33 (2023). https://doi.org/10.1007/s11042-023-15738-7
3. Jo, T., Nho, K., Saykin, A.J.: DL in alzheimer's disease,: diagnostic classification and prognostic prediction using neuroimaging data. Front Aging Neurosci. **20**(11), 220 (2019). https://doi.org/10.3389/fnagi.2019.00220.PMID:31481890;PMCID:PMC6710444
4. Bringas, S., Salomón, S., Duque, R., Lage, C., Montaña, J.L.: Alzheimer's disease, stage identification using DL models. J. Biomed. Inform. **109**, 103514 (2020)
5. Alsubaie, M.G., Luo, S., Shaukat, K.: Alzheimer's disease, detection using DL on neuroimaging: a systematic review. ML Knowl. Extract. **6**(1), 464–505 (2024)
6. Zhao, Z., et al.: Conventional ML and DL in Alzheimer's disease, diagnosis using neuroimaging: a review. Front. Comput. Neurosci. **17**, 1038636 (2023)
7. Alqahtani, S., Alqahtani, A., Zohdy, M.A., Alsulami, A.A., Ganesan, S.: Severity grading and early detection of alzheimer's disease, through transfer learning. Information **14**(12), 646 (2023)
8. Al Shehri, W.: Alzheimer's disease, diagnosis and classification using DL techniques. PeerJ Comput. Sci. **8**, e1177 (2022)
9. Helaly, H.A., Badawy, M., Haikal, A.Y.: DL approach for early detection of Alzheimer's disease. Cogn. Comput. **14**(5), 1711–1727 (2022)
10. Liu, S., et al.: Generalizable DL model for early Alzheimer's disease, detection from structural MRIs. Sci. Rep. **12**(1), 17106 (2022)
11. Jo, T., Nho, K., Saykin, A.J.: DL in Alzheimer's disease,: diagnostic classification and prognostic prediction using neuroimaging data. Front. Aging Neurosci. **11**, 220 (2019)
12. Arafa, D.A., Moustafa, H.E.D., Ali, H.A., Ali-Eldin, A.M., Saraya, S.F.: A DL framework for early diagnosis of Alzheimer's disease, on MRI images. Multimedia Tools Appl. **83**(2), 3767–379 (2024)
13. Hassan, N., Musa Miah, A.S., Shin, J.: Residual-based multi-stage DL framework for computer-aided Alzheimer's disease, Detextion. J. Imag. **10**(6), 141 (2024)

14. Ni, Y.C., et al.: Classification prediction of Alzheimer's disease, and vascular dementia using physiological data and ECD SPECT images. Diagnostics **14**(4), 365 (2024)
15. Mahmud, T., Barua, K., Habiba, S.U., Sharmen, N., Hossain, M.S., Andersson, K.: An explainable AI paradigm for Alzheimer's diagnosis using deep transfer learning. Diagnostics **14**(3), 345 (2024)
16. Hasan, M.E., Wagler, A.: New convolutional neural network and graph convolutional network-based architecture for AI applications in Alzheimer's disease, and dementia-stage classification. AI **5**(1), 342–363 (2024)
17. Abdelwahab, M.M., Al-Karawi, K.A., Semary, H.E.: DL-based prediction of Alzheimer's disease, using microarray gene expression data. Biomedicines **11**(12), 3304 (2023)
18. Gandhi, V.C., Mehta, D.: An extensive examination of ML approaches for predicting heart disease insights: a comprehensive survey. In: CRC Press eBooks, pp. 319–326 (2024). https://doi.org/10.1201/9781003596721-70
19. Altwijri, O., et al.: Novel deep-learning approach for automatic diagnosis of Alzheimer's disease, from MRI. Appl. Sci. **13**(24), 13051 (2023)
20. Khan, S., Khan, M., Iqbal, N., Dilshad, N., Almufareh, M.F., Alsubaie, N.: Enhancing sumoylation site prediction: a deep neural network with discriminative features. Life **13**(11), 215 (2023)
21. Khan, S., Khan, M., Iqbal, N., Khan, S.A., Chou, K.-C.: Prediction of PiRNAs and their function based on discriminative intelligent model using hybrid features into chou's PseKNC. Chemom. Intell. Lab. Syst. **203**, 104056 (2020)
22. Gandhi, V.C., Gandhi, P.P.: A survey - insights of ML and DL in health domain. In: 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), pp. 239–246. Erode, India (2022). https://doi.org/10.1109/ICSCDS53736.2022.9760981
23. Zhu, Y., Liu, Y., Chen, Y., Li, L.: ResSUMO: a DL architecture based on residual structure for prediction of LysineSUMOylation sites. Cells **11**, 2646 (2022)
24. Khan, S., et al.: Optimized feature learning for anti-inflammatory peptide prediction using parallel distributed computing. Appl. Sci. **13**, 7059 (2023)
25. Al-Jumaili, M.H.A., et al.: Analysis and prediction pathways of natural products and their cytotoxicity against HeLa cell line protein using docking. Molecul. Dyn. ADMET. J. Biomol. Struct. Dyn. **41**, 765–777 (2023)
26. Alsemmeari, R., Yehia Dahab, M., Alturki, B., Alsulami, A.: Priority detector and classifier techniques based on ML for the IoMT. Comput. Mater. Contin. **76**, 1853–1870 (2023)
27. Thakkar, D., Gandhi, V.C., Trivedi, D.: Forecasting maternal women's health risks using random forest classifier. In: 2024 International Conference on Inventive Computation Technologies (ICICT), pp. 961–965. Lalitpur, Nepal (2024). https://doi.org/10.1109/ICICT60155.2024.10544471
28. Alqahtani, A., Alsulami, A.A., Alahmari, S., Alrizq, M.: Accurate phase detection for ZigBee using artificial neural network. AIntell. Autom. Soft Comput. **36**, 2505–2518 (2023)
29. Pauranik, A., Ramesh, U.V., Riad Al-Fatlawy, R., Kaur, J., Gandhi, V.C.; Detection of gastric cancer in endoscopic images using deep convolutional neural networks. In: 2024 IEEE International Conference on Communication, Computing and Signal Processing (IICCCS), pp. 1–5. ASANSOL, India (2024). https://doi.org/10.1109/IICCCS61609.2024.10763738
30. Bordin, V., Coluzzi, D., Rivolta, M.W., Baselli, G.: Explainable AI points to white matter hyperintensities for Alzheimer's disease, identification: a preliminary study. In: Proceedings of the 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp. 484–487. Glasgow, Scotland, 11–15 July 2022
31. Özbay, F.A., Özbay, E.: An NCA-based Hybrid CNN Model for Classification of Alzheimer's disease, on Grad-CAM-enhanced Brain MRI Images. Turk. J. Sci. Technol. **18**, 139–155 (2023)

32. Thakar, S., Patel, D., Gandhi, V.C., Trivedi, D.: Financial analytics with artificial neural networks: predicting loan repayment. In: 2024 8th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), pp. 1666–1073. IEEE (2024)
33. Loveleen, G., Mohan, B., Shikhar, B.S., Nz, J., Shorfuzzaman, M., Masud, M.: Explanation-driven HCI model to examine the mini-mental state for Alzheimer's disease. ACM Trans. Multimed. Comput. Commun. Appl. **20**, 1–16 (2023)
34. Veitch, D.P., Weiner, M.W., Aisen, P.S., Beckett, L.A., Cairns, N.J., Green, R.C., et al.: Understanding disease progression and improving Alzheimer's disease, clinical trials: recent highlights from the Alzheimer's disease, neuroimaging initiative. Alzheimers Dement. **15**, 106–152 (2019). https://doi.org/10.1016/j.jalz.2018.08.005
35. Gandhi, V.C., Siyal, D.K., Patel, S.P., Shah, A.V.: A survey: emotion detection using facial reorganization using Convolutional Neural Network (CNN) and viola–jones algorithm. Modeling and Optimization of Signals Using Machine Learning Techniques, pp. 97–113 (2024)