scientific reports

OPEN

An integrated deep learning model for early and multi-class diagnosis of Alzheimer's disease from MRI scans

Emanuel Raju Vinukonda & B. N. Jagadesh✉

Alzheimer's disease (AD) is a progressive neurodegenerative disorder that severely affects memory, behavior, and cognitive function. Early and accurate diagnosis is crucial for effective intervention, yet detecting subtle changes in the early stages remains a challenge. In this study, we propose a hybrid deep learning-based multi-class classification system for AD using magnetic resonance imaging (MRI). The proposed approach integrates an improved DeepLabV3+ (IDeepLabV3+) model for lesion segmentation, followed by feature extraction using the LeNet-5 model. A novel feature selection method based on average correlation and error probability is employed to enhance classification efficiency. Finally, an Enhanced ResNext (EResNext) model is used to classify AD into four stages: non-dementia (ND), very mild dementia (VMD), mild dementia (MD), and moderate dementia (MOD). The proposed model achieves an accuracy of 98.12%, demonstrating its superior performance over existing methods. The area under the ROC curve (AUC) further validates its effectiveness, with the highest score of 0.97 for moderate dementia. This study highlights the potential of hybrid deep learning models in improving early AD detection and staging, contributing to more accurate clinical diagnosis and better patient care.

A loss of cognitive function is a common feature of many cognitive issues that are referred to as dementia. Dementia can manifest in three different levels: mild dementia, severe dementia, and health control (no dementia)[1]. For therapeutic and drug development applications, models with the ability to evaluate the state of the disease and automate the classification process are essential[2]. Many of a person's everyday tasks may still be able to be completed alone, and only mild cognitive functioning issues may be present in the early stages of the disorder[3]. However, as the illness becomes more severe, these challenges worsen, and people may need more help doing things such as dressing, taking a shower, and feeding themselves[4]. When later stages of dementia develop, the person's health and well-being are maintained by requiring 24-hour care and may become dependent on others for necessities[5]. Clinical doctors find it difficult to determine how a disease is progressing based only on behavior and other external symptoms. Therefore, a disease classification tool can help with drug research and advance early diagnosis[6].

The examination of brain cell images is one method for creating models to aid in the classification of dementia. This is a sensible strategy because dementia develops as a result of damage to brain cells[7]. This interruption may, therefore, have a detrimental effect on someone's conduct, feelings, and mental processes[8]. Nonetheless, since several individuals survive well into their nineties without suffering from dementia symptoms, it is not regarded as a normal part of ageing[9]. Although there are many other kinds of dementia, the most prevalent type is Alzheimer's disease (AD). Amyloid plaque accumulation and the formation of neurofibrillary tangles in the brain are two of the primary symptoms of AD that are still recognized today[10]. Furthermore, the illness also results in a reduction in the connections between neurons, which are in charge of transmitting information from the brain to other areas of the body and from one area of the brain to another[11]. Furthermore, it's thought that several other complex alterations in the brain also have a role in the onset and course of AD[12]. AD is connected to damage to specific types of brain cells in particular brain regions. Certain proteins that are abnormally abundant both within and outside of brain cells impair brain cell health and interfere with intercellular communication[13].

School of Computer Science and Engineering, VIT-AP University, Vijayawada 522237, India. ✉email: nagajagadesh@gmail.com

Since the hippocampus, the part of the brain involved in learning and memory is often the first to suffer damage, memory loss is among the first signs of AD[14].

Convolutional filters are utilized to discern the features of an intricate AD image. These features of an image are its corners, edges, or textures. After locating a feature, each filter determines its values during training[15]. CNNs use several convolutional filters to identify a multitude of features simultaneously. By concentrating on distinct elements, each filter aids in the network's learning of hierarchical characteristics, ranging from simple edges to intricate patterns[16,17]. After convolutional layers, pooling layers are added to minimize the spatial dimensions of activation maps while preserving crucial information[18]. For the dementia stage's classification, the convolutional filters may be able to identify image features after multiple layers[19,20].

Everyone experiences AD symptoms differently in the beginning. Researchers are looking into biomarkers, which are biological markers of disease detected in blood, cerebrospinal fluid, and brain scans, to detect early brain alterations in patients with MCI who may be more susceptible to AD. As a diagnostic and therapeutic tool, this work advances the understanding of the alterations in the brain associated with dementia. We propose to use hybrid deep learning models to identify different stages of dementia and AD to accomplish this goal. Given that the signs and progression patterns of dementia vary depending on the stage, understanding the stages of dementia aids medical practitioners in providing the appropriate care and medications. For this reason, this work presents a deep learning technology called Improved DeepLabV3+-Modified ResNext to predict different phases of dementia. The proposed hybrid approach (IDeepLabV3 + and EResNext) is not a mere combination of models but a carefully designed architecture that leverages the strengths of each model. IDeepLabV3 + precisely segments the lesion regions, ensuring that only the most relevant areas are used for classification. The EResNext model then efficiently captures both local and global spatial relationships, allowing for a more refined classification of different AD stages. The synergy between these two models enhances the overall accuracy and robustness compared to individual models. We introduce a feature selection mechanism based on average correlation plus error probability, which systematically eliminates redundant and irrelevant features. Instead of traditional pre-processing, we incorporate a unique combination of image enhancement and augmentation techniques, including a specialized contrast-limited adaptive histogram equalization (CLAHE) with an overlapping average filter. This enhancement significantly improves image quality, allowing the model to learn more discriminative features from MRI images. Unlike previous studies that focus on binary AD classification, our approach provides a multi-class classification of AD (moderate dementia (MOD), mild dementia (MD), very mild dementia (VMD), and non-dementia (ND)). Additionally, we conducted an extensive ablation study and statistical significance testing (ANOVA) to demonstrate the contribution of each module to the final performance.

Research contributions

- A hybrid deep learning-based approach is introduced in this research that combines the strengths of the Improved DeepLabV3+ (IDeepLabV3+) model for precise segmentation of AD-related regions and the Enhanced ResNeXt (EResNext) model for classification.
- We employ image enhancement methods, such as the application of average filters alongside Contrast-Limited Adaptive Histogram Equalization (CLAHE). These techniques improve the visualization of brain structures, aiding in better segmentation and classification.
- We address class imbalance issues by implementing data augmentation strategies, including shifting, flipping, and rotating images at various angles, ensuring a more balanced and representative dataset.
- From the segmented images, we extract both local and global features using the LeNet-5 model, capturing essential patterns pertinent to AD.
- To refine the feature set, we apply a feature selection method based on average correlation plus error probability, effectively eliminating less significant features and retaining those most relevant for accurate classification.
- We conduct a comprehensive evaluation of our hybrid deep learning model against state-of-the-art techniques using recognized performance metrics. This comparison demonstrates the efficacy and superiority of our approach in accurately detecting and classifying different stages of Alzheimer's Disease.

Paper organization

The remaining research is arranged as follows: A thorough review of the most recent research on the classification of AD is provided in Sect. "Literature survey". Section "Proposed methodology" provides a detailed description of the materials and procedures used in this proposed research. The results analysis is shown in Sect. "Result and discussion". Lastly, Sect. "Conclusion" concludes the work.

Literature survey

Two hybrid models, CNN-Conv1D-LSTM and hybrid recurrent ensemble network (HReENet), were presented by Ayus et al.[21] for the categorization of AD. CNN is used as a feature extractor, and Conv1D-LSTM is used as a classifier in the hybrid CNN-Conv1D-LSTM model. Meanwhile, CNN, LSTM, and CNN-Conv1D-LSTM predicted values are used in the HReENet ensemble model. Using 5-fold cross-validation, the analysis of proposed models is performed. HReENet is used to enable precise AD diagnosis.

The moderately demented and non-demented MRI images were segmented and object-recognized using a CNN technique called Mask-RCNN by Unde et al.[22], using data from the train and test datasets. The first uses of Mask-RCN were object instance segmentation and object recognition in MRI images. A framework for detecting specific characteristics of AD from MRI images was created by Singh et al.[23] utilizing the CNN. The synthetic minority oversampling strategy is used to avoid the problem of class imbalance. To forecast the various phases of dementia based on MRI, an STCNN model is suggested.

By using hybrid deep learning (DL) approaches, Matlani et al.[24] created a novel method for diagnosing AD. Improved adaptive wiener filtering (IAWF) is used for image pre-processing to enhance the obtained images. Next, the features are extracted using a powerful hybrid technique called Principal Component Analysis, which extracts the essential features from images without the need for image segmentation. PCA-NGIST stands for Normalized Global Image Descriptor. Subsequently, the Improved Wild Horse Optimization algorithm (IWHO) is employed to identify the optimal features. Ultimately, a hybrid BiLSTM-ANN is used to determine the disease.

Roy et al.[25] developed a Forward Attention-based deep network (FA-VGG16) for classifying breast histopathology images. Solving class imbalance boosts performance to 97.7% accuracy. The attention modules in FA-VGG16 enhance feature extraction, which is evident from its progressive ablation study performance drop. The promise of FA-VGG16 to improve breast cancer detection through attention processes is numerically validated by its 97.7% accuracy for binary classification and 92.4% accuracy for quaternary classification.

Difficulties such as constrained generalizability, limited interpretability, high computational demands, class imbalance, and limited accuracy are faced by most of the existing deep learning models. Nonetheless, our proposed hybrid deep learning model successfully overcomes these drawbacks and offers advancements in several areas. Although prior deep learning models have demonstrated some degree of accuracy in AD classification, the proposed model (IDeepLabV3+-EResNext) outperforms them in terms of accuracy and capability. One of the challenges with AD classification has been class imbalance within datasets. To address this problem, prior techniques included feature extraction techniques and pre-trained models. Similarly, our approach efficiently addresses class imbalance and improves classification accuracy by utilizing conventional data augmentation techniques like shifting, flipping, and rotating with multiple angles. In clinical settings, limited computational resources may be necessary for effective AD diagnosis. Our proposed model is to offer thorough and effective AD diagnosis while considering the clinical environment and, if needed, optimizing computing resources. The interpretability of deep learning models has been a constraint, impeding comprehension of the decision-making process underlying forecasts. In the proposed hybrid deep learning model, our model integrates multiple modules, including Dilated Residual Network, squeeze excitation (SE) module, Atrous Spatial Pyramid Pooling (ASPP), MultiStepLR strategy, and label smoothing strategy, into the features and regions that impact its predictions, improve interpretability, and establish reliability. The generalizability of the approach was limited by the frequent focus of earlier research on particular datasets or cohorts. Local and global features from AD datasets can be extracted using a LeNet 5 model.

Proposed methodology

As seen in Fig. 1, Resources for MRI image-based AD diagnosis are provided in this section to identify the disease early in its phases of development. In this research, we have proposed a novel multi-class classification system for AD disease diagnosis based on a hybrid deep learning system. The publicly accessible ADNI dataset is where the MRI images of AD are sourced from. For multi-class classification, the five essential stages are used in the proposed model, including (1) Pre-processing, (2) Segmentation, (3) Feature extraction, (4) Feature selection, and (5) Classification. First, pre-processing is done to enhance the quality of the dataset. By overlapping the average filter and the CLAHE approach, MRI images of AD are improved, and some data augmentation techniques such as shifting, flipping, and rotating with many angles, are applied to solve the class imbalance problem and balance the dataset. After pre-processing, the interested lesion region of the AD is segmented using the IDeepLabV3 + model. From the segmented images, the local and global features are extracted using the LeNet-5 model. To reduce the dimensionality of the features, we have used the feature selection method, which is based on average correlation plus error probability for selecting the essential features. For classifying the AD diseases, the selected features are sent to the EResNext model, which classifies the MRI images into MD, MOD, VMD, and ND. Utilizing efficient models such as Depthwise Separable Convolutions in the IDeepLabV3 + model and Grouped Convolutions in EResNeXt significantly decreases parameter size while preserving performance in our proposed research. Incorporating SE modules ensures that only the most relevant feature maps contribute to classification, reducing unnecessary feature propagation and the number of trainable parameters.

The IDeepLabV3 + model is employed to precisely segment the AD-affected regions from MRI scans, thereby isolating key lesion areas and reducing background noise. This focused segmentation allows the subsequent EResNext model to concentrate on the most informative features for AD classification. EResNext, enhanced with mechanisms such as Squeeze-and-Excitation, efficiently captures channel-wise dependencies and subtle variations in the segmented features, which are crucial for distinguishing between different AD stages. The synergistic combination of these models improves the overall sensitivity and specificity by ensuring that only the most relevant, high-quality features contribute to the final classification, ultimately leading to better detection and staging of AD.

Dataset collection

The Alzheimer's Dataset applied in this research was derived from[26] a 2D picture and slice obtained from the ADNI database[27]. Most research has either used it exclusively or in combination with other data sets. Increasing the amount of relevant data and testing the usefulness of MRI and other biomarkers in tracking the progression of AD are the primary goals of the ADNI task. The ADNI dataset includes imaging and clinical metadata from a diverse population. The subjects in our research were selected based on the availability of MRI scans, with the following demographic characteristics: Age Range: 55–90 years (covering early-onset and late-onset AD cases); Gender Distribution: Both male and female subjects were included, ensuring balanced representation; Cognitive Scores: Subjects were assessed using neuropsychological tests such as the Mini-Mental State Examination (MMSE) and Clinical Dementia Rating (CDR). Medical History The dataset includes additional biomarkers such as APOE genotype and cerebrospinal fluid (CSF) analysis for a subset of participants.
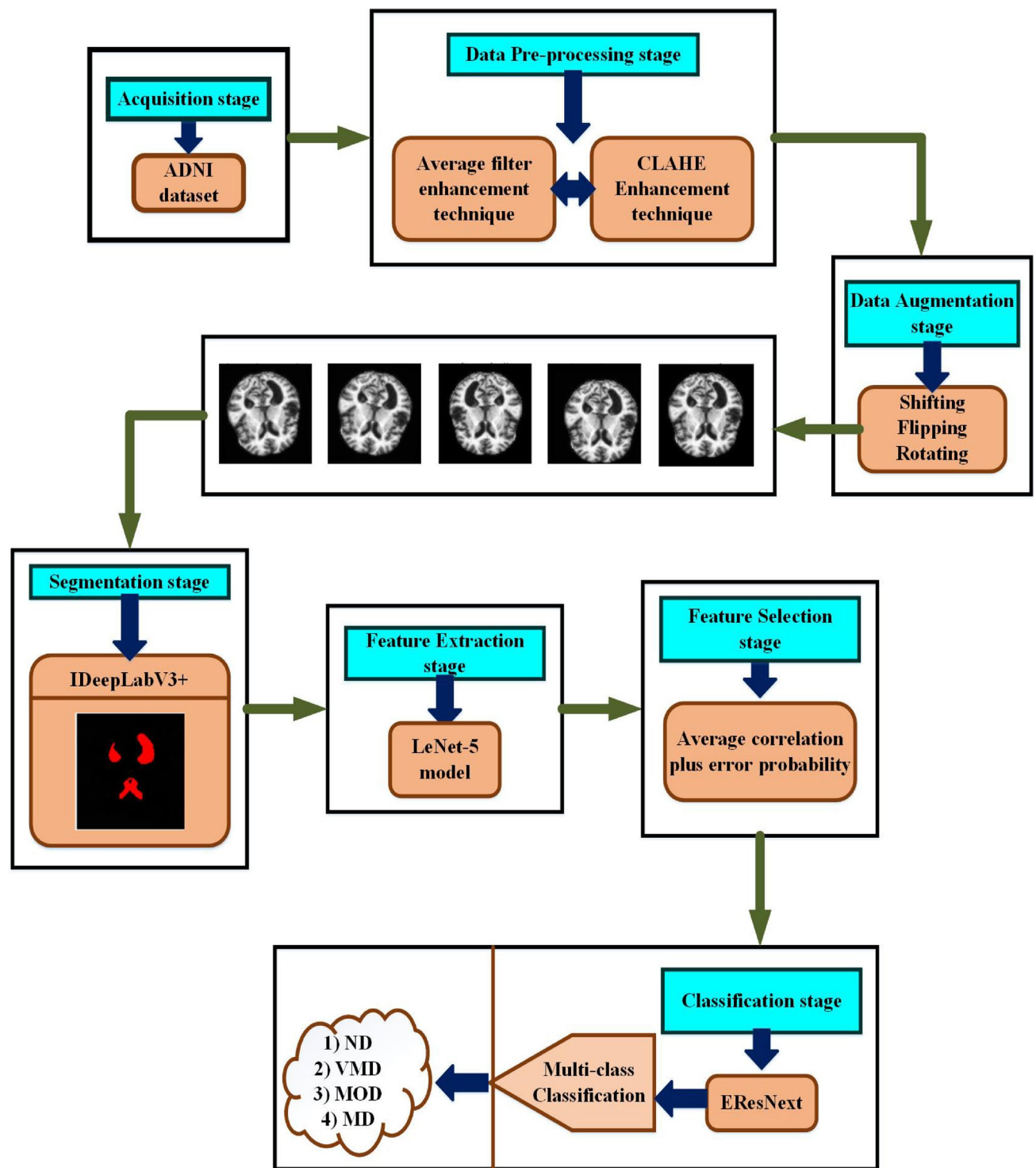
Fig. 1. A methodology framework for diagnosing MRI images to identify early stages of AD progression.

6400 MRI images make up the data collection, and every image is used in the research. There are four classes of data in the used dataset: MD, MOD, VMD, and ND. The data set was initially pre-processed and organized for the experiment. The amount of data and sample images used in the research are shown in Table 1. Figure 2 shows sample MRI images of different stages of AD.

Image pre-processing

The performance of the proposed systems is impacted by a variety of factors that might cause noise in MRI images, including patient movement during the scan, brightness, reflections, low contrast, and the acquisition

Class	No of samples
ModerateDemented (MOD)	64
MildDemented (MD)	896
VeryMildDemented (VMD)	2240
NonDemented (ND)	3200

Table 1. Number of samples in the ADNI dataset.
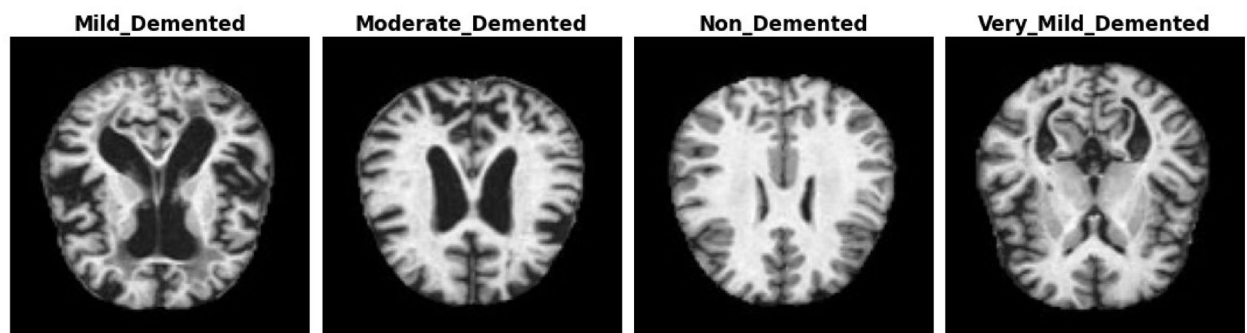
Fig. 2. ADNI dataset's sample MRI images.

of MRI data from several devices. Thus, improving MRI images yields better images, which help achieve high performance. The average filter and CLAHE are two methods that this research used to enhance MRI images.

First, we applied the average filter to improve the MRI images. The filter method states that the filter always functions by altering a pixel that has, on average, fifteen neighboring pixels. The central pixel gets replaced for the average of the 15 neighboring pixels. Equation (1) illustrates the candidate's labour, which doesn't stop working until the image is finished:

$$F(z) = \frac{1}{N} \sum_{i=0}^{N-1} y(z-1) \tag{1}$$

Where the improved image is represented as $F(z)$, the past input is depicted as $y(z-1)$, and the image's pixels are depicted as N. Second, the CLAHE[28] approach has improved low-region contrast in MRI images. The technique enhances the contrast of the borders of the affected regions by uniformly distributing the light pixels of the MRI images. Every time, the method uses a different pixel based on the transformation function to replace the original pixel. Each central pixel's neighboring pixels are captured based on the contrast of the image. The contrast of the image will increase when the goal pixel value is less than the neighbor's pixels and drop when the goal pixel value is larger than the neighbors' pixels. MRI images were thereby enhanced. An enhanced MRI image sample from the Alzheimer's dataset across all classes is displayed in Fig. 3. Finally, an even better image was obtained by combining the image improved for the CLAHE approach with the image improved for the average filter.

To enhance the quality of the MRI images, we employed image enhancement techniques by combining an average filter with CLAHE. This approach improves image contrast and preserves important structural details, making it easier for the IDeepLabV3 + model to segment lesion regions accurately. Additionally, to address the class imbalance problem, we applied data augmentation techniques such as shifting, flipping, and rotation at multiple angles, ensuring more significant variability in training data and improving the model's ability to generalize across different cases. These pre-processing steps enhance segmentation accuracy and contribute to better feature extraction for the LeNet-5 model, leading to improved classification performance in the EResNext model. This ultimately results in a higher classification accuracy and improved reliability of AD staging.

Data augmentation and balancing dataset

A small number of images in the dataset is one of the CNN models' limitations. CNN requires a large image dataset to train the systems. Furthermore, the unbalanced dataset is regarded as a restriction of CNN because it influences its performance and shifts the accuracy towards the class with the majority of images. As a result, by artificially raising the dataset's image size from the same images in the dataset, the data augmentation technique addresses the issues raised. Compared to the increase in the majority class, the data augmentation method balances the unbalanced dataset by producing a higher percentage of images. Thus, by balancing the dataset, this approach increases the number of images during training. Multiple methods are employed to enhance the images in the dataset collection, including flipping, shifting, and multi-angle rotation. The MD class saw an increase of six artificial images from each image in this data augmentation process; the MD class saw an increase of 97 artificial images from each image; the ND class saw an increase of two artificial images from each image;
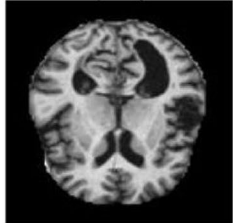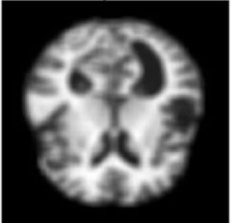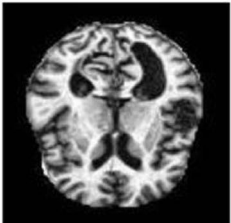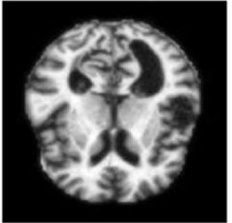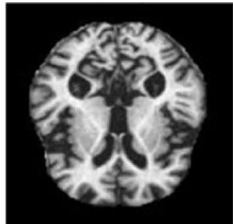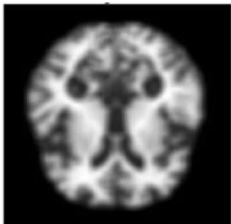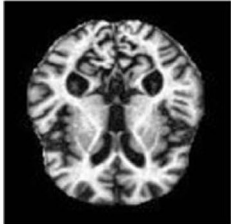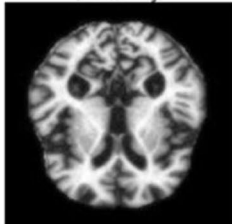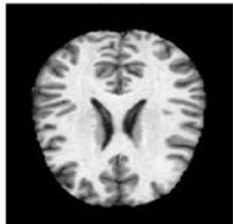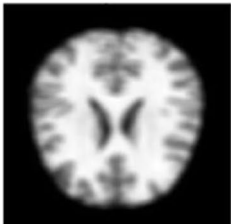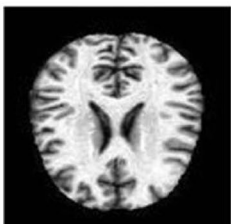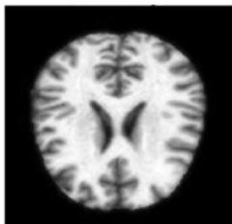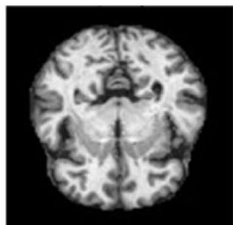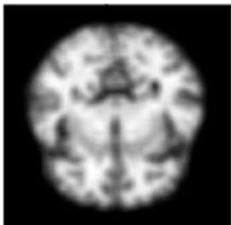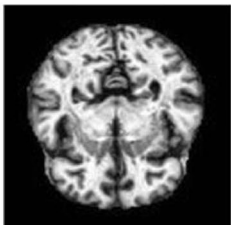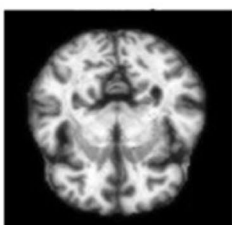
Fig. 3. Improved MRI image samples for the AI dataset.

and the VMD class saw an increase of two artificial images from each image. The data augmentation of the brain MRI images is shown in Fig. 4. The data is enhanced by the use of flipping, rotating, and shifting.

Data splitting

The dataset was initially divided into training (80%), validation (10%), and testing (10%) subsets using stratified random sampling. This ensured that the class distribution was preserved in each subset. Following this, data augmentation techniques including flipping, shifting, and multi-angle rotation were applied exclusively to the training set to address class imbalance. The validation and testing sets remained unaltered for unbiased evaluation. Table 2 presents the distribution of samples for each set, comparing the number of samples.

Segmentation

In this research, we segmented the relevant AD region using the improved DeepLabV3+ (IDeepLabV3+) model. To accomplish different segmentation tasks, the mask boundaries provided by the dataset source were kept unchanged when the ground-truth maps were rearranged. While the basic DeepLabV3+[29] model performs well in segmentation, it is not as effective in segmenting smaller lesion regions and fails to capture complicated patterns in the impacted lesion regions. To improve the structure of the basic DeepLabV3 + model, we must incorporate atrous spatial pyramid pooling (ASPP) and dilated residual networks.

This section provides an overview of the encoder and decoder structure that is widely employed in the majority of cutting-edge models for semantic segmentation. The proposed DeepLabV3 + model's primary parts are described, along with the ASPP, decoder module, and dilated residual network, which serves as the
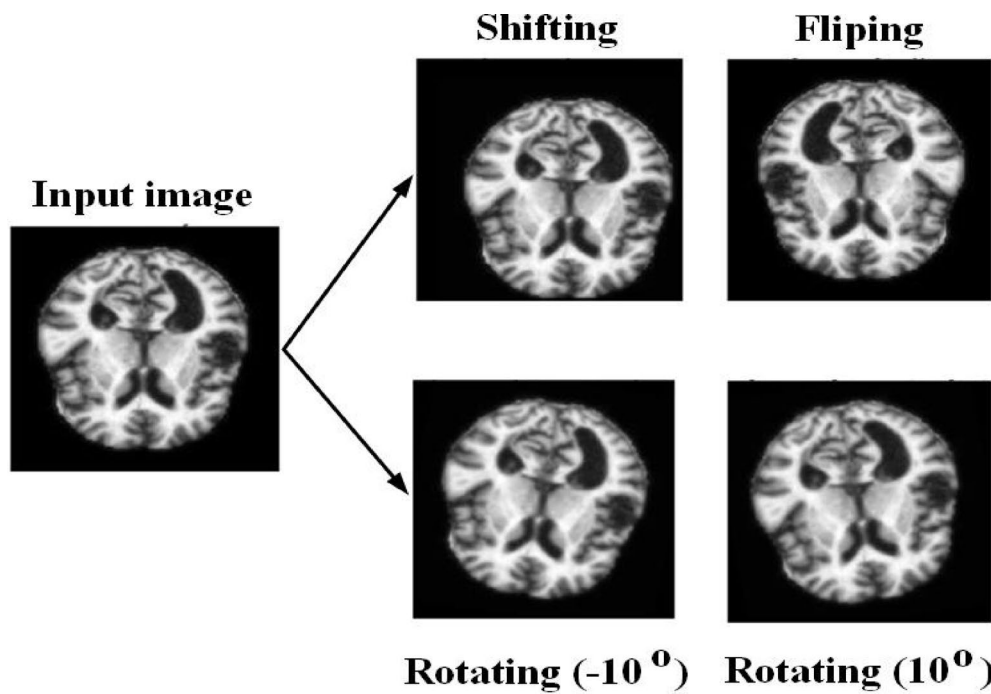
Fig. 4. Data augmentation.

Class	Training (80%)	Testing (10%)	Validation (10%)
ModerateDemented (MOD)	4268	10	10
MildDemented (MD)	4416	80	80
VeryMildDemented (VMD)	5092	327	327
NonDemented (ND)	5382	223	223

Table 2. The dataset distribution for each class in training, testing, and validation.

model's backbone. The structure of the IDeepLabV3 + segmentation network utilized in this research is depicted in Fig. 5. We outline the advantages of our model compared to traditional UNet and its variants (UNet++, Attention UNet, Residual UNet, etc.). UNET relies on simple skip connections, which may not capture fine-grained lesion details effectively. IDeepLabV3 + enhances segmentation using Atrous Spatial Pyramid Pooling (ASPP) and Dilated Residual Networks, enabling multi-scale feature extraction and better lesion segmentation. UNet + + introduces dense skip pathways to improve feature fusion. However, it still struggles with small lesion regions. The proposed IDeepLabV3 + model incorporates dilated convolutions to capture fine details at multiple resolutions without significantly increasing computational overhead. Attention UNet improves segmentation by applying attention gates to suppress irrelevant regions. Our proposed model takes a different approach by enhancing contextual feature learning through ASPP and dilated residual networks, which provide a more robust solution for segmenting complex AD lesion regions. ResUNet improves UNET's learning capability by using residual connections. However, its feature extraction remains limited by kernel size and lacks effective multi-scale context aggregation. IDeepLabV3 + overcomes this limitation by integrating ASPP, which extracts multi-scale contextual information and improves segmentation accuracy.

Encoder and decoder architecture
Blocks with encoder and decoder functions make up the majority of proposed techniques for semantic segmentation. The encoder portion gradually reduces the image's sample size before presenting it to the input. The high-level information is captured by progressively lowering the feature map resolution. An effective decoder module is used to create a segmentation map from the encoder's small feature maps that match the original image's dimensions. By utilizing the decoder to gather spatial information, an overall architecture that is more robust and faster is produced. A deep learning-based efficient encoder and decoder structure is used for semantic segmentation. The ResNet served as the encoder block's backbone in this research.

Dilated residual network for encoder
In particular, deep learning performs better than traditional techniques for classifying large-scale images. Popular CNN architectures with different processing units and topological variations have been proposed recently. It
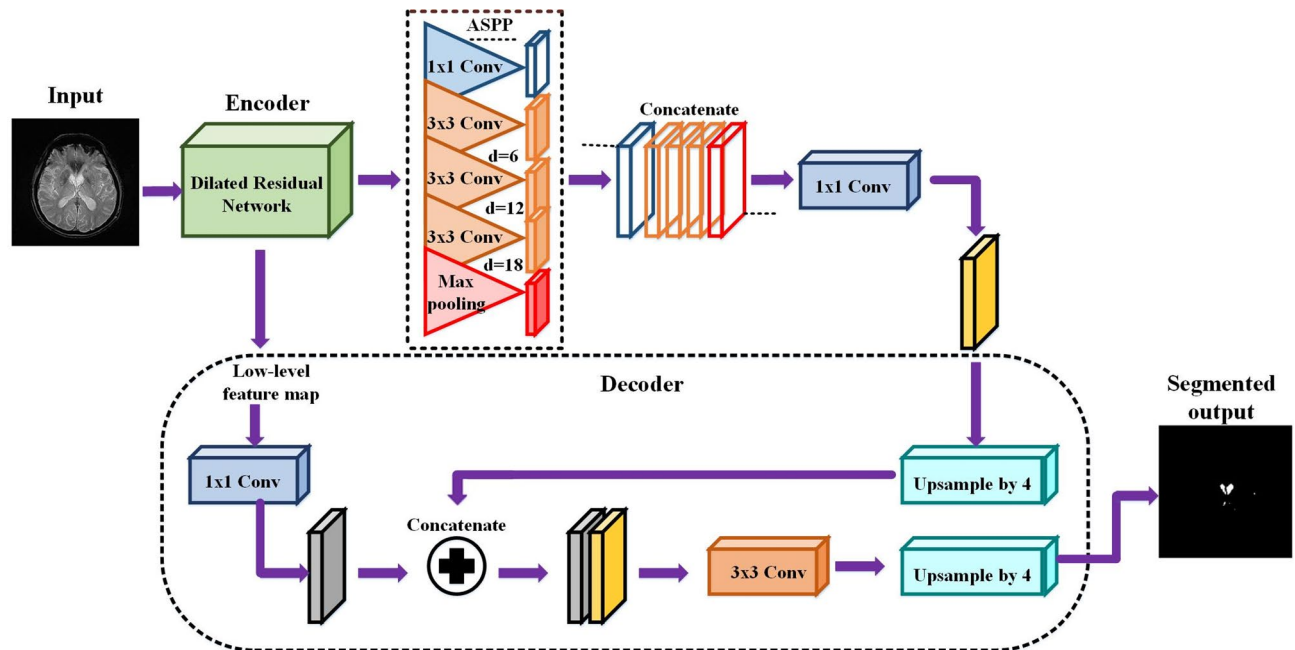
Fig. 5. Improved DeepLabV3+ (IDeepLabV3+) architecture for AD segmentation.

was noted that performance does not rise in direct proportion to training depth despite the training procedure becoming more challenging. The gradient problem with vanishing and exploding is the primary cause of this. The remaining components were included to make it easier for networks to be trained and to make possible deeper architectures. Despite their more profound architecture, Semantic segmentation problems frequently employ ResNet topologies as feature extractors because of their lower computing cost. The networks have a basic structure even though they have different numbers of layers, such as ResNet-50, ResNet-34, and ResNet-18. ResNet uses five convolution blocks with spatially resolved outputs. Consequently, it has been determined that using the most basic of these architectures is sufficient.

The segmentation performance of the ResNet-18 structure is negatively impacted by the feature maps with a small scale produced following the encoder, where each convolutional block is built on the residual block. For obtaining highly spatially-reflective feature maps, the stride rate in the convolutional layer of the last two blocks of the design is set to 1 instead of 2. However, the convolution process attempted to compensate for the decrease of the receptive field by using the proper dilation factor when the stride factor was reduced. Standard convolution operations are represented by dilation rate (d = 1), whereas up-sampled convolution filters are defined by d > 1.

Dilation convolutions, in which between each filter weight are zeros, are used to keep the receptive field and the feature map in balance based on the established d rates. Dilated convolution can increase the final 16×16 feature maps' dimensions produced by the original ResNet architecture for an image of 512×512 pixels. In CNN, the standard down-sampling factor, also known as the output stride (OS), is 32. The OS stands for the ratio of the input to output image's spatial resolution. In this research, 16 was chosen as the OS ratio using the dilation approach. To attain a reduced OS rate while achieving a more extensive feature extraction, more substantial segmentation maps were attempted to be obtained.

Atrous spatial pyramid pooling (ASPP)
The success of spatial pyramid pooling (SPP) and atrous convolutional processes served as the impetus for the development of the ASPP approach. The SPP was created to get around deep learning's restriction that input photos have a set size. Using the ASPP approach, feature maps produced by the encoder at various atrous rates are resampled. The outputs are then concatenated after the feature maps are subjected to a parallel convolution filter at various atrous rates. The region of interest's boundaries may disappear, even though the final feature map encodes an extensive amount of semantic information. The ASPP method is an accurate and efficient way to capture multi-scale information. With dilation rates of d = 6, 12, and 18, the ASPP module comprising a 1 $\times 1$ convolution, 3×3 convolutions, and a max-pooling layer in parallel was employed in this research. AD abnormalities of varied densities and sizes have been attempted to be segmented with high sensitivity using depth-wise convolution instead of ordinary convolution.

Decoder module
From the encoder's low-scale feature maps, original image size segmentation maps are produced by the decoder module. OS = 16 down-samples the encoder feature map compared to the original image. The semantic segmentation map is created by first up-sampling the encoder feature maps by factor 4. From the encoder backbone, the intermediate low-level features are concatenated with these feature maps in the same spatial dimension. Concatenating high-level data from ASPP with low-level features from the backbone network, which

are rich in spatial information, improves segmentation performance. To generate the final segmentation output, this feature map is up-sampled by 4 and then undergoes 3×3 convolution. Based on the discussion, Fig. 6 shows that the proposed segmentation method correctly separates the relevant regions of AD.

IDeepLabV3 + serves as the core segmentation module in our proposed framework for AD detection. Its primary function is to accurately segment affected brain regions from MRI scans, which is crucial for precise feature extraction and subsequent classification. The model improves upon the basic DeepLabV3 + by integrating Dilated Residual Networks, which enhances the feature propagation and capturing fine-grained details of small lesion regions. ASPP is used to extract multi-scale contextual information to improve segmentation accuracy. Preservation of mask boundaries, which ensures that smaller lesions are effectively segmented without loss of structural details.

These enhancements allow IDeepLabV3 + to overcome the limitations of standard segmentation models, particularly in detecting subtle and complex lesion patterns associated with early-stage AD. As a result, the segmented outputs provide more refined and informative features for classification, improving the overall accuracy and reliability of the proposed system.

By predicting the number of foreground and background pixels, the auxiliary branch helps the network better understand structural and spatial relationships in MRI images. This additional supervision encourages the model to focus on relevant regions of interest while reducing sensitivity to background noise and improving feature representation. The auxiliary task introduces an additional learning constraint, reducing over-fitting by forcing the network to learn complementary information about the image structure.

Unlike traditional methods that require extra labeled data for improved performance, this auxiliary branch leverages existing data to generate additional training signals. It acts as a form of self-supervision, enabling the model to extract richer features without the need for external datasets or manual annotations. The auxiliary objective provides additional gradient feedback, leading to faster convergence during training. This results in a more stable optimization process, reducing the likelihood of the model getting stuck in suboptimal solutions.
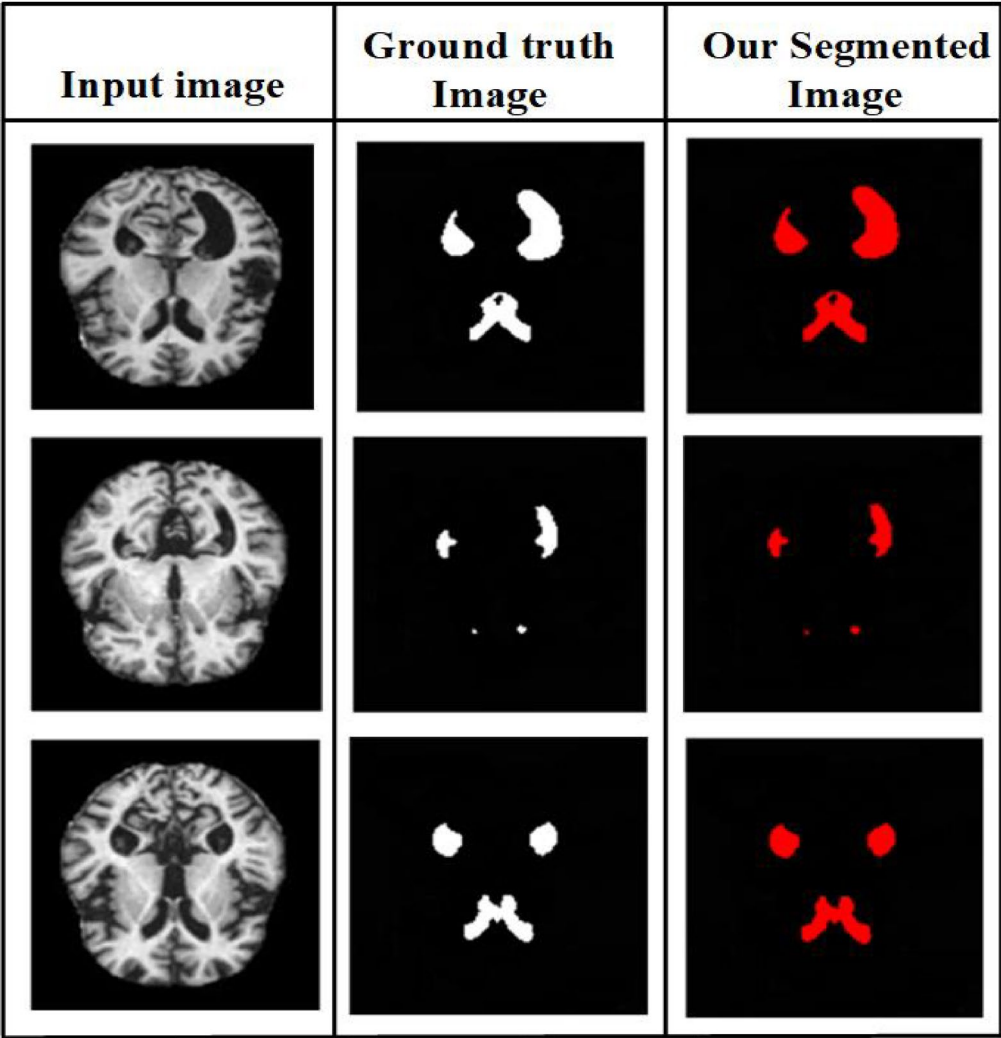

Fig. 6. Segmentation results of the proposed approach.

By integrating this auxiliary branch, the model gains more substantial feature learning capabilities, improved generalization, and enhanced robustness while maintaining computational efficiency.

Feature extraction

To extract local and global features, the segmented image is processed using the LeNet-5 model[30]. The local features such as Intensity variations (gray matter, white matter), texture patterns (smoothness, roughness, and homogeneity), shape descriptors (curvature, elongation, or roughness), and local histograms (the statistical properties of tissue or pathology within specific regions of the image), and Global Features such as Intensity statistics (mean, variance, skewness, and kurtosis of pixel intensities), Histogram features (histogram shape, entropy, and energy), Spatial relationships, Shape descriptors, and Texture descriptors are extracted.

Depending on the needs of the application, several CNN types can be employed as feature extractors. For our purposes, a pre-trained LeNet-5 network on the ImageNet database is utilized as the feature extraction network. Seven layers make up LeNet-5: an input layer, two convolution (ConV) layers, two fully connected (FC) layers, two pooling (Pool) layers, and an output layer. The LeNet-5 has multiple weighted layers that are predicated on the concept of utilizing shortcut connections to bypass the convolution layer blocks. The fundamental "bottleneck" blocks often adhere to two design principles: If the output feature size is the same, use the same number of filters; if it is half, use twice as many filters. Additionally, the convolution layers with a stride of two do the down-sampling, and the rectified-linear-unit (ReLU) activation occurs before batch normalization, which happens after every convolution.

If the dimensions of the input and output are the same, an identity shortcut is employed; if the dimensions grow, a projection shortcut is used to match the dimensions using 1×1 convolutions. The feature-extraction network utilizes the 49 convolutional layers from LeNet-5 solely that lack average pooling and fully connected layers, as indicated in Table 3, because the feature selection and classification process receives the feature map produced by the last convolutional layer.

The LeNet-5 model was chosen for feature extraction due to its ability to efficiently capture both local and global structural patterns from segmented MRI images, which are essential for accurate AD classification. It extracts local features such as intensity variations, texture patterns, and shape descriptors, as well as global features like intensity statistics, histogram features, and spatial relationships. By leveraging these comprehensive feature representations, LeNet-5 enhances the classification capability of the EResNext model, improving its ability to distinguish subtle pathological changes across different AD stages. This integration contributes to the model's high classification accuracy, demonstrating its effectiveness in early AD detection and staging.

Feature selection

The most significant step in the deep learning process is feature selection. This process's primary objective is to identify a dataset's most valuable properties and eliminate its worthless ones. It was found in this study that not every feature that was recovered was equally valuable for categorizing AD.

Our approach refines the feature set by eliminating redundant and less informative features, ensuring that only the most discriminative features contribute to classification. The average correlation metric helps remove highly correlated features, which often provide redundant information, thereby reducing model complexity and over-fitting. Meanwhile, the error probability criterion ensures that the retained features maximize class separability, leading to a more compelling feature set for classification.

The collected dataset was highly challenging to work with because of its enormous feature vector space of 5,486,400 ($254 \times 21,600$). The resolution of this issue involved optimizing the functionality's size to provide a consistent and faithful representation of all data, leading to a suitable categorization with a low error rate. Since PCA made it easier to convert the input data linearly, it produced good results on data that had been divided along linear dimensions.

Furthermore, it was employed in the feature selection process. The most significant feature set, which was less functional than the original feature vector space, was found using PCA. Unfortunately, because the PCA was unable to retain a large number of discrete data points, this optimized feature set did not give an accurate image of the complete data set. Furthermore, an unsupervised methodology was the PCA method; nevertheless, the dataset of AD diseases was labelled, and the PCA outcomes on the labelled data were not as encouraging. Effective feature selection techniques[31], precisely average correlation (AC) and probability of error (POE), are applied to choose the best features from this extensive, multi-dimensional dataset acquired from AD disease

Layer Name	Input size	Output size	Step size	Convolution kernel size	Pooled area
Input layer	32*32	28*28	1	5*5	–
ConV	6@28*28	6@14*14	2	–	2*2
Pool	6@14*14	16@10*10	1	5*5	–
ConV	16@10*10	16@5*5	2	–	2*2
Pool	16@5*5	120@1*1	1	5*5	–
FC	1*120	1*84	–	–	–
FC	1*84	1*7	–	–	–
Output	1*7	–	–	–	–

Table 3. LeNet-5 feature-extraction network.

to solve this challenge. In comparison to PCA, this method performed better and produced a sub-dataset with adequate features for the large dataset. Using Python 3.9, the proposed approach, out of 254 features, 10 optimized features were chosen. Mathematically, (POE + AC) is defined as

$$PA(f_w) = \frac{Missclassified Samples}{Total Samples} \tag{2}$$

$$f_2 = f_w : Min_w[PA(f_w) + Correlate(f, f_w)] \tag{3}$$

$$f_k = f_w : Min_w\left[PA(f_w) + \frac{1}{N-1}|Correlate(f_1, f_w)|\right] \tag{4}$$

Upon deploying (POE + AC) on the obtained dataset, 10 optimized features were identified for additional processing, which was then used to deploy the optimized fused hybrid-feature dataset to the classification network model for multi-class classification. This process was done using a (POE + AC)-based methodology.

Classification

With the aid of a deep learning network, the classification of AD is the last stage of the research. The newly developed technique improves the proposed approach's overall performance while maximizing the cancer prediction rate. The most used classification techniques for AD detection and classification are SVM, Random Forest, and Deep Learning Neural Networks. These traditional classification algorithms have a low classification rate and are not appropriate for use in the process of subsequent severity diagnosis. The MRI images in this research are classified as MD, MOD, ND, or VMD using the Enhanced ResNext (EResNext) model. ResNeXt is a convolutional neural network architecture that extends the ResNet framework by incorporating a new dimension called 'cardinality,' referring to the number of parallel paths within each residual block.

The fundamental ResNext model[32] yields good classification results; it can also identify channel-wise dependencies and prevent degradation caused by the network becoming deeper or larger. Channel dependencies might not be able to classify objects accurately, which would weaken the model. Over-fitting is more likely when there is an excessive reliance on particular channels. It is frequently advantageous to create models that can efficiently utilize data from several channels while retaining flexibility and adaptation to various data settings to lessen these drawbacks. Consequently, we use the squeeze excitation (SE) module, MultiStepLR technique, and label smoothing technique to improve the structure of the ResNext model.

The Enhanced ResNeXt (EResNext) based classification network model is provided to simplify classifying MD, MOD, ND, and VMD, and the complexity of the multi-class AD classification process is addressed by this approach. Furthermore, the accuracy is increased by utilizing three optimization strategies. First, to enhance the network's capacity for learning, the MultiStepLR technique is applied to modify the learning rate dynamically. Second, the label smoothing technique optimizes the one-hot label, which can lessen the network's reliance on the actual labels' probability distribution and enhance the network's capacity for prediction. Ultimately, the MRI dataset transfer learning technique simplifies the transfer learning process and enhances the network's capacity for generalization.

There are similarities between the attention mechanism and the SE (Squeeze and Excitation) module. Through the squeeze and excitation operations, it uses the scale operation to reweight features after generating and capturing channel-wise dependencies. This module's selective emphasis on tumor features and suppression of useless ones are its main advantages. The classification procedure uses grouped convolutions. Cardinality is used to limit the number of groups while maintaining the same topology across all groups. Figure 7 displays the IResNext network that is utilized in the proposed framework.

Conv_bn is the batch-normalization layer that was added to the convolution layer; the reduction ratio is $r = 16$, the cardinality is 32, and the Excitation operation's fully connected layer is called fc. In addition to utilizing
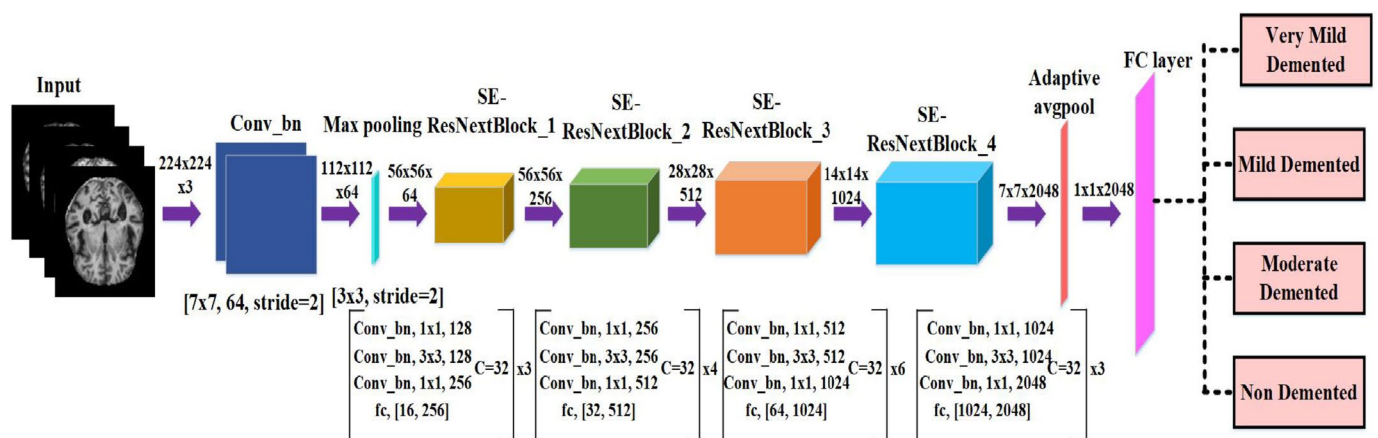

Fig. 7. The enhanced ResNext (EResNext) model structure.

the benefits of the reweighted feature, the IResNeXt network expands its width and depth. As a result, the network can accurately and consistently classify AD classes.

The multisteplr strategy

It is often acknowledged that the most significant hyper-parameter is the learning rate to adjust when training CNN, as it has a direct impact on the network's capacity to learn and how it trains. The training algorithm may diverge if the learning rate is too high. Conversely, a learning rate that is too low will limit the network's capacity to learn and cause the training algorithm to converge slowly. A more extensive range of the optimal value will be fluctuated by the training algorithm even if it reaches the convergence state during training with a single learning rate. Nevertheless, the training algorithm's fluctuation range will be lower as the learning rate decreases with increasing epochs. Thus, to achieve the learning rate's dynamic adjustment, the MultiStepLR strategy is employed.

After the chosen epoch is reached, the rate of learning will decrease linearly. The formula is as follows:

$$\alpha' = \alpha \times \gamma^{[milestones]} \tag{5}$$

Where the specified epoch is [milestones], γ is 0.1.

The label smoothing strategy

One technique for regularization is label smoothing. Usually, it's employed to keep the network from growing overconfident. The ground-truth label distribution $q(k) = \delta_{k,y}$ is primarily changed throughout its implementation. Utilizing label smoothing, the sample's ground-truth label distribution is replaced by

$$q'=(k) = (1-\in)\delta_{k,y} + \in u(k) \tag{6}$$

Where $u(k)$ is commonly interpreted as the uniform distribution $u(k) = \frac{1}{k}$ since it is independent of the training sample x, and the smoothing parameter is described as . The cross-entropy loss is as follows:

$$L(q',p) = -\sum_{k=1}^{k} log(p(k)q'(k)) \tag{7}$$

$$= (1-\in)L(q,p) + \in L(u,p) \tag{8}$$

Consequently, it is the same as increasing the cross entropy loss by one penalty item. The degree of variation between the predicted distributions, $p(k)$ and $u(k)$, is represented by $u(k) = \frac{1}{k}$ and $L(u,p)$.

The ground-truth label's dependence on the network is lessened by smoothing the distribution of the ground-truth label. Network learning could be successfully directed by using the label smoothing technique when the network's learned features are not sufficient to distinguish between positive and negative data. This research assigns 0.1 to the smoothing parameter .

The transfer learning strategy

To address the over-fitting problem resulting from insufficient training data, a transfer learning technique was presented. The majority of pre-trained networks used in medical image classification systems nowadays were trained using the ImageNet dataset. However, ImageNet is a part of the natural image collection and is not significantly related to medical imaging. Low classification accuracy on medical images could result from the pre-trained network's poor performance for high-level features and its high capacity for generalizing low-level features, which was trained on the ImageNet dataset. Using medical images, a transfer learning technique is suggested as a solution to this issue. Before applying high capacity for generalizing low-level features, pre-train the network on various types of medical image datasets. To increase classification performance and efficiency, the network is retrained.

Result and discussion

This section goes over the many performance metrics that were utilized to assess the proposed approach and contrast it with the most advanced existing techniques.

The multi-class classification of AD is classified by using the proposed hybrid deep learning system (IDeepLabV3+-EResNext). For experiment analysis, the MRI brain imaging dataset (ADNI) is utilized, with 80% of the data being used for training and 20% for testing. With the help of an i5 processor and 8 GB of RAM, the outcomes were obtained using the Python 3.9 platform.

Parameter settings

The hyper-parameter settings for this research project are shown in Table 4. The proposed hybrid deep learning models' parameters are optimized using the ADAM optimizer.

Performance metrics

Comprehending the system's functionality is crucial. Various assessment measures are applied to this. While TNs successfully detected negative cases and TPs accurately classified positive instances, FPs and FNs are incorrectly predicted as negative and positive, respectively.

The percentage of accurately classified predictions made by a model over the total number of classified occurrences is known as accuracy

Hyperparameters	Values
Maximum epoch size	100
Activation function	Softmax
Drop out factor	0.4
Number of iterations	100
Initial learning rate	0.0001
Batch size	10
Loss function	Cross entropy

Table 4. Hyper-parameter settings of the proposed model.

Accuracy: 0.9812				
Class name	Precision	Recall	F1-score	Support
MD	1.00	0.97	0.99	80
MOD	1.00	1.00	1.00	10
ND	0.98	0.99	0.98	327
VMD	0.99	0.97	0.98	223
Accuracy				640
Macro avg	0.99	0.99	0.99	640
Weighted avg	0.98	0.98	0.98	640

Table 5. Performance analysis of the proposed multi-class classification.

$$Accuracy = \frac{TP + TN}{FN + TN + FP + TP} \qquad (9)$$

$$Precision = \frac{TP}{TP + FP} \qquad (10)$$

$$Recall = \frac{TP}{TP + FN} \qquad (11)$$

$$F1 - Score = 2 * \frac{precision * recall}{precision + recall} \qquad (12)$$

Experimental results

Table 5 provides information regarding the effectiveness of the proposed multi-class classification of AD and shows the proposed classification system's classification reports. For the ADNI dataset, the proposed hybrid IDeepLabV3+-IResNext model achieved an overall classification accuracy of 98.12%, precision of 99.25%, recall of 98.25%, and F1-score of 98.75%.

The accuracy and loss curves of the proposed classification system for training and validation are shown in Fig. 8. After 40 iterations, the training and validation learning rates stabilize as the number of epochs increases. The output of the curves, which run through smoothly during the entire experiment, demonstrates the curves' effectiveness. The loss rate of the model is similarly represented by the loss curve. They level off and stay largely stable with no obvious trend after 45 epochs. Based on the experimental results, a maximum accuracy of 98.29% is obtained for training, and a maximum accuracy of 98.12% is acquired for testing. Testing and training both reach their maximum level of efficiency. In training and testing, the model achieves a loss of 0.021 and 0.049, respectively.

The proposed multi-class AD classification approach's test data confusion matrix is displayed in Fig. 9. It shows that it accurately recognizes 313 patients with Very mildly affected and 337 healthy persons. It correctly identified 286 out of 301 patients with a moderate case and 255 out of 274 patients with a mild case. Thus, the efficiency of the proposed approaches is confirmed. The proposed classification approach accurately and highly accurately classifies the MRI brain images based on the confusion matrix.

Some test results for the proposed hybrid model are displayed in Fig. 10. All test samples were correctly identified since the proposal can achieve an accuracy of 99.18%.

The ROC-AUC graph is typically used to evaluate the diagnostic test. The relationship between the false positive rate and the true positive rate is displayed on the ROC-AUC graph. The multi-class classification system's Receiver Operating Characteristic (ROC) curve, which distinguishes between the ND, VMD, MD, and MOD categories, is shown in Fig. 11. More specifically, the AUC for the ND class was 0.95578, the AUC for VMD was 0.92319, and the AUC for MD was 0.9351. By comparison, the MD class had the most excellent AUC value, 0.97964. These AUC values highlight how well the model performs in differentiating between the various stages of dementia.
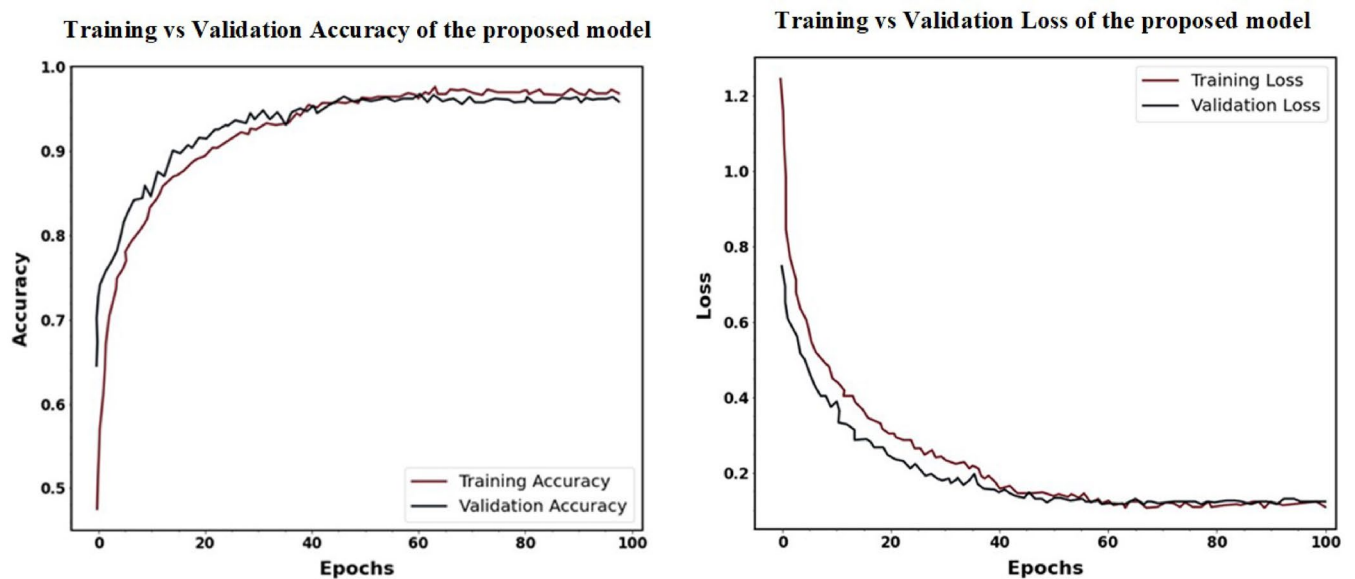
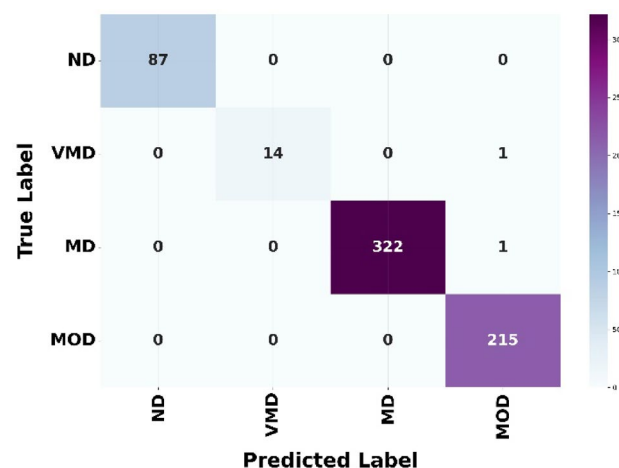**Fig. 8**. Accuracy and loss curve for a proposed classification system.

Fig. 9. Confusion matrix for multi-class classification.

Performance comparison

Comparing the metrics employed in the research with values from the metrics of the classification findings utilized in this work, this part conducts a comparative analysis with other comparable works. The works that make use of the same dataset represent the most recent research in the field. Using the same dataset as this work, Table 6 summarizes current research on multi-class classification of AD.

It is evident from the analysis that the proposed model performs better in a variety of metric analyses. To assess the proposed model's superiority over the existing models, Mask R-CNN, SVM, ResNet-50, VGG16, and InceptionV3-CNN are compared with it. It is evident from the analysis that, in comparison to the existing models, the proposed model performs better. Better results have been obtained by the proposed algorithm as a result of efficient data training and feature learning. The increased accuracy rate indicates that multi-class classification AD may be discriminated more effectively, which will enhance the model's performance.

With 98.12% accuracy, 99.25% precision, 98.25% recall, and 98.75% F1-score, the proposed classification system performs well. Given its early level of AD classification, the IDeepLabV3+-EResNext is a robust model because of its greater precision and recall, which indicate that it is effective in minimizing false positives and false negatives. With a high F1 score, the IDeepLabV3+-EResNext exhibits a strong balance between precision and recall, demonstrating its efficacy in identifying actual positive cases while preventing false negatives and positives. Based on these metrics, the proposed system seems to be a very accurate and reliable model for AD classification overall.

Due to its efficient learning of data attributes, the proposed model has outperformed the existing models in comparison. Through the effective extraction of features that accurately reflect the inter-scale variability of the
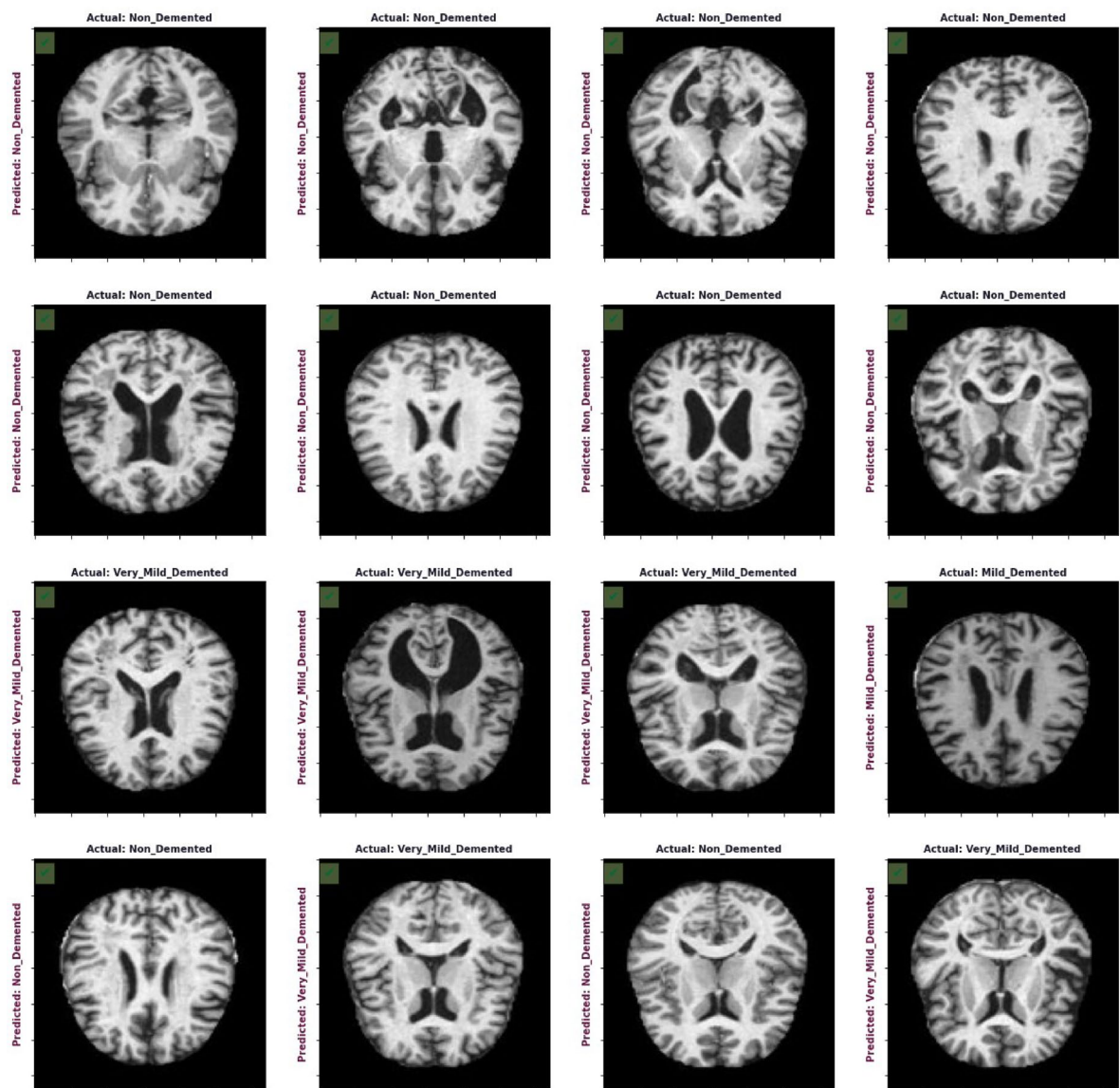
Fig. 10. Prediction result of the proposed hybrid model.

diseases, the proposed model enhances classification performance. The proposed model's high classification accuracy allows for automatic diagnosis and pre-screening of AD based on multi-class classification.

The IDeepLabV3 + model enhances the segmentation of brain regions affected in early AD, allowing for a more detailed structural analysis. This segmentation helps capture subtle cortical atrophy and white matter changes that might not be easily detectable by conventional imaging methods. The use of the LeNet-5 model for feature extraction allows the model to capture fine-grained spatial patterns associated with early-stage AD. The integration of the Squeeze-and-Excitation (SE) module in EResNext improves sensitivity to minor variations in brain structures by recalibrating the importance of features. The MultiStep Learning Rate (MultiStepLR) and label smoothing techniques enhance the model's ability to generalize across subtle early-stage abnormalities.

The enhanced performance of the proposed hybrid (IDeepLabV3+-EResNext) model in predicting AD is due to its well-chosen features and network design, which provides doctors with a trustworthy instrument. Because of the model's high precision and recall, diagnosis errors are minimized in both positive and negative cases by ensuring precise identification. Its precision and memory are highlighted by its capacity to maintain a balanced F1 score, which helps practitioners identify actual positive situations while averting false positives and negatives. The effectiveness of EResNext model increases its practical effectiveness by giving medical practitioners a reliable and accurate detecting tool for rapid intervention and individualized patient treatment in real-world situations.

Increasing the number of convolutional layers enhances the model's ability to capture complex and hierarchical features. Shallow networks may struggle to extract deep spatial and texture-based features, while excessively deep networks may introduce redundant information. More convolutional layers increase computational complexity and training time. However, beyond a certain depth, the performance gain becomes marginal, and diminishing returns or even over-fitting may occur, especially in smaller datasets. Very deep networks can suffer
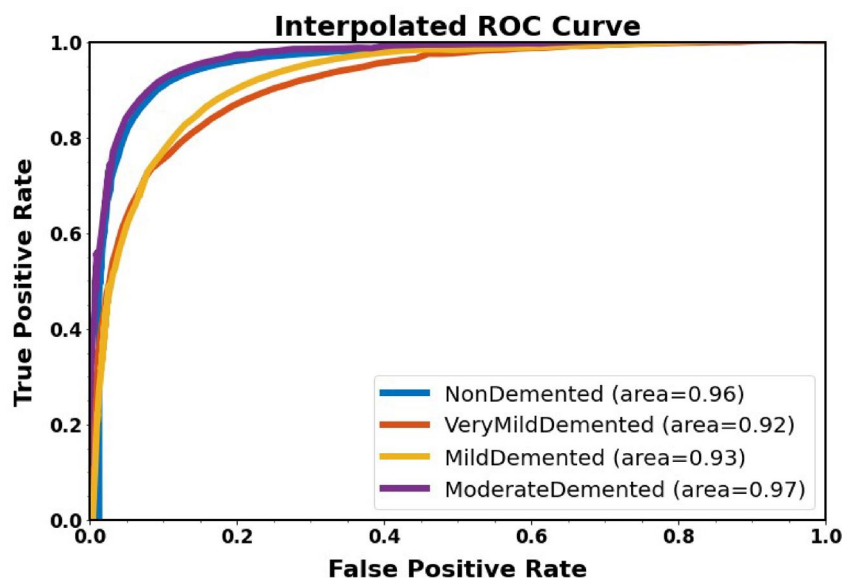
Fig. 11. Multi-class ROC for Alzheimer's classification.

Reference	Year	Approach	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Unde et al.[22]	2024	Mask R-CNN	97.46	–	–	–
Shobha et al.[33]	2024	SVM	94.49	77.42	94.49	81.08
Alqahtani et al.[34]	2023	ResNet- 50	94.36	94.34	94.36	94.35
Sharma et al.[35]	2022	Neural networks with VGG16	90.40	90.00	90.40	90.40
Rana et al.[36]	2023	InceptionV3-CNN	97.31	97.00	97.00	97.00
Proposed research		IDeepLabV3+-EResNext	98.12	99.25	98.25	98.75

Table 6. Comparative results of AD MRI images with different models.

from gradient vanishing or exploding issues, which may lead to unstable training or difficulty in convergence. Proper architectural design, such as using residual connections or batch normalization, helps mitigate these issues. Through experimental analysis, we evaluated models with varying convolutional depths and observed that an optimal number of convolutional layers balances feature extraction and computational efficiency. Future research could explore adaptive architectures where the model dynamically adjusts the number of convolutional layers based on input complexity. Combining different convolutional depths within a multi-scale or ensemble-based approach may improve classification performance. Automated model selection techniques such as NAS can optimize the depth and configuration of convolutional layers to maximize performance while minimizing computational overhead. Exploring advanced regularization methods, dropout strategies, and optimization techniques to mitigate over-fitting and enhance generalization for deeper models.

Clinical integration of the proposed deep learning model for early Alzheimer's diagnosis

The model can be embedded within existing hospital Picture Archiving and Communication Systems (PACS) or radiology workflow software to process MRI scans automatically. Standard MRI data formats (such as DICOM) can be directly fed into the system, allowing for smooth interoperability with clinical imaging protocols. Our improved DeepLabV3+ (IDeepLabV3+) model ensures precise lesion segmentation, which can assist radiologists in identifying early structural changes associated with AD. This can be integrated into clinical decision-support systems (CDSS), providing radiologists with additional insights rather than replacing expert judgment. The LeNet-5 feature extraction and the novel feature selection method (based on average correlation and error probability) enable efficient processing of MRI features relevant to AD staging. The model's feature selection enhances interpretability, which can be valuable in clinical discussions and second-opinion assessments. The Enhanced ResNext (EResNext) model classifies AD into four stages, aiding clinicians in precise staging and disease monitoring. The classification results can be presented in a structured report format, making it easier for neurologists and radiologists to incorporate them into their assessments. A confidence score can be provided to improve trust and explainability for clinicians. The model can be deployed via on-premise AI workstations within hospitals, enabling real-time inference. Validation studies in collaboration with neurologists and radiologists will ensure clinical acceptance, reliability, and robustness before widespread adoption.

By incorporating these strategies, our proposed deep learning model can enhance clinical workflows by providing early and accurate AD diagnosis, improving patient outcomes through timely intervention, and supporting healthcare professionals with AI-driven decision-making tools.

Practical implications

Achieving high accuracy, precision, and recall in AD classification has significant practical implications in clinical settings. These performance metrics directly impact patient outcomes, clinical decision-making, and healthcare efficiency.

High accuracy ensures reliable classification of patients into appropriate AD stages, minimizing misdiagnosis and facilitating early therapeutic interventions. Early detection is crucial for slowing disease progression and improving patient quality of life. High precision minimizes false positives, reducing unnecessary psychological distress and medical procedures for patients who do not have AD. High recall ensures that fewer cases of AD are missed, preventing delays in diagnosis and treatment. This is particularly critical for early-stage AD, where timely intervention can significantly alter disease progression. A highly accurate system reduces the need for repeated scans and additional diagnostic tests, optimizing hospital resources. Reliable classification enables better tracking of disease progression, helping clinicians adjust treatment strategies based on disease evolution.

Challenges and solutions for deploying AI-based AD diagnosis in healthcare

AI models trained on specific datasets may not generalize well to diverse patient populations due to variations in MRI protocols, scanner types, and demographic differences. Expanding our training dataset to include multi-center, multi-population data and applying domain adaptation techniques can enhance model robustness. Regular updates and fine-tuning should be performed to maintain accuracy across different clinical settings.

Deploying deep learning models in hospitals requires computational resources, which may not be readily available in all healthcare institutions. Cloud-based deployment options can provide scalability, allowing hospitals with limited computational capacity to access the model remotely. Edge computing and optimized lightweight models can also reduce hardware requirements for on-premise deployment. AI tools must seamlessly integrate into hospital PACS and radiology software without disrupting existing workflows. Developing API-based integration with PACS and EHR systems ensures smooth data exchange. Additionally, user-friendly interfaces with automated reporting can help clinicians efficiently interpret model outputs.

Clinical reliability of the proposed model

In this research, we assess the reliability of the proposed model for clinical deployment by evaluating its performance, generalizability, and integration into real-world workflows. The model achieves high accuracy (98.12%) and an AUC of 0.97, minimizing false diagnoses and ensuring robust classification. Validation of MRI datasets enhances their adaptability to diverse imaging conditions. Explainability techniques, such as visualized and confidence scores, improve transparency, aiding clinician trust. The model outperforms existing methods and is designed for seamless integration into PACS and radiology workflows. These aspects collectively demonstrate its potential for reliable AD diagnosis in clinical settings.

Statistical validation of the proposed model

To verify the effectiveness of the proposed hybrid deep learning-based multi-class classification system, a one-way Analysis of Variance (ANOVA) test was conducted. This test compared the classification performance of the proposed method against baseline models, including traditional deep learning approaches.

We have obtained an F-statistic (F) = 12 and a p-value (p) = 0.02 from the ANOVA test, indicating a statistically significant difference in classification performance among the models. This result confirms the effectiveness of the proposed approach, demonstrating its superiority over baseline methods in distinguishing between the four stages of AD. The F-statistic of 12 indicates a significant variance in classification performance across different models, confirming that the proposed approach offers superior accuracy. The p-value of 0.02 is less than the standard threshold of 0.05, proving that the observed improvements in classification accuracy are statistically significant and not due to random variation.

These results demonstrate that the proposed IDeepLabV3 + for lesion segmentation, LeNet-5 for feature extraction, and Enhanced ResNext (EResNext) for classification significantly outperform conventional models in distinguishing between the four AD stages. This statistical validation reinforces the reliability and robustness of the proposed methodology in improving early AD detection and staging, contributing to more accurate clinical diagnosis and patient care.

Ablation study

The ablation study systematically evaluates the impact of each component within the proposed framework, demonstrating how their integration enhances AD classification performance. The ablation study of the proposed research is given in Table 7. Initially, the pre-processing stage, incorporating Average Filtering and CLAHE, improves image quality, but without segmentation, feature extraction, or classification, its contribution remains limited. When segmentation is introduced using the IDeepLabV3 + model, the isolation of lesion regions provides a more structured input for analysis. However, without feature extraction and classification, the accuracy does not significantly improve. Similarly, incorporating only feature extraction with the LeNet-5 model enhances the ability to capture local and global characteristics but remains suboptimal without segmentation, as irrelevant regions may still impact classification.

A notable improvement is observed when segmentation and feature extraction are combined, as the segmented lesion regions provide meaningful feature representations that contribute to classification. However, the highest accuracy is achieved when all components—pre-processing, segmentation, feature extraction, and

Pre-processing	Segmentation	Feature extraction	Classification	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
	✓	✓	✓	96.54	98.25	97.62	96.10
✓		✓	✓	95.29	96.76	96.42	96.46
✓			✓	95.45	95.30	95.45	95.56
✓	✓		✓	97.43	97.35	97.46	96.42
✓	✓	✓	✓	98.12	99.25	98.25	98.75

Table 7. The ablation study of the proposed approach.
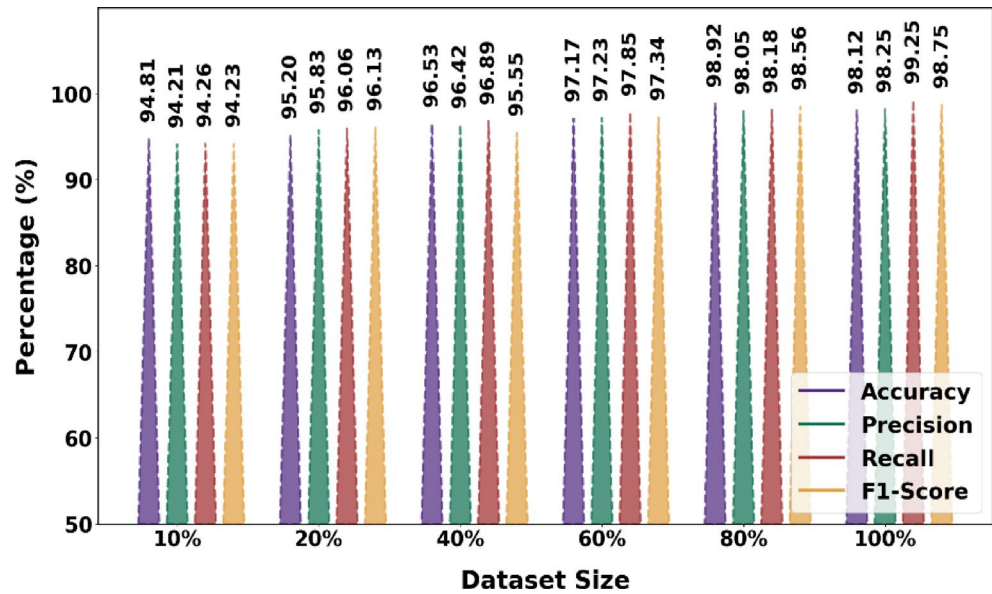
Fig. 12. Performance variation with increasing dataset size.

classification—are integrated. The final classification using the EResNext model, along with feature selection based on average correlation plus error probability, refines the extracted features, ensuring that only the most significant ones contribute to decision-making. This comprehensive approach maximizes F1-score, recall, precision, and accuracy, validating the effectiveness of each stage. The results confirm that segmentation helps focus on relevant regions, feature extraction enhances representational power, and the classification module ensures robust decision-making, making the entire pipeline the most effective configuration for AD classification.

Data growth study
Figure 12 shows the performance variation with increasing dataset size. The data growth study evaluates the impact of dataset size on the proposed model's performance. The figure illustrates the variations in F1-score, recall, precision, and accuracy as the dataset size increases from 10 to 100%. Initially, at 10% of the dataset, the model's accuracy rate is 94.81%, with corresponding precision, recall, and F1-score values of 94.21%, 94.26%, and 94.23%, respectively. As the dataset size increases, performance metrics show a consistent upward trend, indicating that a more extensive dataset contributes to improved model generalization and robustness. At 20%, accuracy improves to 95.20%, and similar trends are observed in precision, recall, and F1-score. With 40% and 60% of the dataset, the model continues to benefit from additional training data, reaching an accuracy of 96.53% and 97.17%, respectively. The highest performance is observed at 100% dataset utilization, where the model's results are 98.75% F1-score, 98.25% recall, 99.25% precision, and 98.12% accuracy. This confirms the effectiveness of using a sufficiently large dataset for training deep learning models, reducing overfitting risks and ensuring better classification results. The results further validate the importance of data augmentation in enhancing model performance, especially in cases where collecting large datasets is challenging.

Limitations
Despite augmentation techniques, the dataset size remains a constraint, which may affect generalization; future work will explore more extensive and more diverse datasets. While deep learning models offer high accuracy, ensuring explainability for clinical adoption remains a challenge; future research will integrate Explainable AI (XAI) techniques. The proposed hybrid model requires substantial computational resources; optimizing the architecture for efficiency without compromising accuracy is an area for further exploration.

Future research plans

We plan to adapt our methodology to other conditions, such as Parkinson's disease and multiple sclerosis, by refining segmentation and classification strategies. Integration with Multi-Modal Imaging: Combining MRI with other modalities, such as PET or fMRI, could enhance diagnostic accuracy and provide deeper insights into disease progression. Developing a real-time AI-driven diagnostic system that integrates with hospital workflows and assists radiologists in decision-making.

The adaptability of our proposed model to other types of tumors for accurate diagnosis and treatment planning faces several challenges. Different tumors exhibit diverse morphological, textural, and intensity-based features, which may impact segmentation and classification accuracy. To address this, domain adaptation techniques and transfer learning can be explored in the future to fine-tune the model for various tumor types. Clinicians require interpretable AI models for decision-making; incorporating explainable AI (XAI) techniques in the future research works, such as saliency maps and Grad-CAM, can enhance model transparency and trustworthiness. Different tumor types may require analysis from multiple imaging modalities (e.g., MRI, CT, PET). A potential solution is developing multi-modal fusion techniques to integrate information from various sources.

Conclusion

MRI images were used to analyze the automatic prediction of AD difficulties with multi-class classifications, and a hybrid deep learning approach was created. First, pre-processing used an image enhancement technique called average filter and CLAHE. IDeepLabV3 + model is used to segment the interested region of AD. The features of the segmented images are extracted using the LeNet-5 model. Next, feature selection was performed based on average correlation plus error probability. Finally, the EResNext model was used for classification, i.e. for diagnosing diseases. With accuracy, AUC value, loss, F1-score, precision, and recall measures of 99.12%, 94.50%, 0.022, 98.75%, 99.25%, and 98.25%, respectively, the proposed model performs exceptionally well. In comparison to state-of-the-art models, our proposed model (IDeepLabV3+-EResNext) showed superior performance, efficiency, and accuracy. Our method's increased accuracy when employing proper network architecture selection indicates its potential use in forecasting various AD stages across a range of age groups. In future studies, we will combine several datasets utilizing advanced data mining techniques to improve the efficacy and performance of AD prediction at early stages by using various datasets and stages.

Data availability

The datasets used and/or analyzed during the current study available from the corresponding author on reasonable request.

References

1. Bansal, D. & Chhikara, R. *Classification of Dementia Using Statistical First-Order and Second-Order Features* 235–255 (Blockchain and Deep Learning for Smart Healthcare, 2023).
2. Balakrishnan, N. B., Pillai, A. S., Panackal, J. J. & Sreeja, P. S. Alzheimer's Disease detection and classification using optimized neural network. *Comput. Biol. Med.* **187**, 109810 (2025).
3. Mahjoubi, M. A. et al. Optimizing ResNet50 performance using stochastic gradient descent on MRI images for Alzheimer's disease classification. *Intelligence-Based Medicine*, **11**, 100219 (2025).
4. Katkam, S., Tulasi, V. P., Dhanalaxmi, B. & Harikiran, J. *Multi-class Diagnosis of Neurodegenerative Diseases Using Effective Deep Learning Models with Modified DenseNet-169 and Enhanced DeepLabV3+.* (IEEE Access, 2025).
5. Gopinadhan, A. Advancing Alzheimer's disease detection: integrating machine learning and image analysis for accurate diagnosis. *Int. J. Intell. Syst. Appl. Eng.* **12** (7s), 350–363 (2024).
6. Vyas, A., Aisopos, F., Vidal, M. E., Garrard, P. & Paliouras, G. Identifying the presence and severity of dementia by applying interpretable machine learning techniques on structured clinical records. *BMC Med. Inf. Decis. Mak.* **22** (1), 271 (2022).
7. Olle Olle, D. G., Zoobo Bisse, J., Alo'o, A. & G Application and comparison of K-means and PCA based segmentation models for alzheimer disease detection using MRI. *Discover Artif. Intell.* **4** (1), 1–14 (2024).
8. Kina, E. *TLEABLCNN: Brain and Alzheimer's Disease Detection Using Attention based Explainable Deep Learning and SMOTE Using Imbalanced Brain MRI* (IEEE Access, 2025).
9. Joy, M. A. M., Nasrin, S., Siddiqua, A. & Farid, D. M. ViTAD: Leveraging modified vision transformer for Alzheimer's disease multi-stage classification from brain MRI scans. *Brain Res.* **1847**, 149302 (2025).
10. Shastry, K. A., Shastry, S., Shastry, A. & Bhat, S. M. A multi-stage efficientnet based framework for Alzheimer's and Parkinson's diseases prediction on magnetic resonance imaging. *Multimedia Tools and Applications* (2025).
11. Arumugam, J., Prasanna Venkatesan, V. & Beigh, T. MRI-Based biomarker in the diagnosis of Alzheimer's disease using Attention-UNet. *SN Comput. Sci.* **6** (3), 1–15 (2025).
12. Xue, C. et al. AI-based differential diagnosis of dementia etiologies on multi-modal data. *Nat. Med.* **30** (10), 2977–2989 (2024).
13. Sriharsha AV. Using Deep Learning to Classify and Diagnose Alzheimer's Disease. *Artificial Intelligence and Cybersecurity in Healthcare.*, 413–428 (2025).
14. Chen, Y. et al. Automated Alzheimer's disease classification using deep learning models with Soft-NMS and improved ResNet50 integration. *J. Radiat. Res. Appl. Sci.* **17**(1), 100782 (2024).
15. Poyraz, M. et al. BrainNeXt: novel lightweight CNN model for the automated detection of brain disorders using MRI images. *Cogn. Neurodyn.* **19** (1), 1–17 (2025).
16. Pradhan, N., Sagar, S. & Jagadesh, T. Advance Convolutional Network Architecture for MRI Data Investigation for Alzheimer's Disease Early Diagnosis. *SN Comput. Sci.* **5**(1), 167 (2024).
17. Dharwada, S., Tembhurne, J. & Diwan, T. An Optimal Weighted Ensemble of 3D CNNs for Early Diagnosis of Alzheimer's Disease. *SN Comput. Sci.* **5**(2), 252 (2024).
18. Lu, S. Y., Zhang, Y. D. & Yao, Y. D. An efficient vision transformer for Alzheimer's disease classification using magnetic resonance images. *Biomed. Signal Process. Control* **101**, 107263 (2025).
19. Li, X., Gong, B., Chen, X., Li, H. & Yuan, G. Alzheimer's disease image classification based on enhanced residual attention network. *PloS One.* **20** (1), e0317376 (2025).

20. Manochandar, T., Diderot, & P. Kumaraguru. Deep Learning-Based Magnetic Resonance Image Segmentation and Classification for Alzheimer's Disease Diagnosis. *International Journal of Image and Graphics*. **25** (03), 2550026 (2025).
21. Ayus, I. & Gupta, D. A novel hybrid ensemble based Alzheimer's identification system using deep learning technique. *Biomed. Signal Process. Control*. **92**, 106079 (2024).
22. Unde, M. & Rathore, A. S. Brain MRI image analysis for Alzheimer's disease diagnosis using mask R-CNN. *Int. J. Intell. Syst. Appl. Eng.* **12** (13s), 137–149 (2024).
23. Anjali, Singh D., Pandey O.J., Dai H.N. STCNN: Combining SMOTE-TOMEK with CNN for Imbalanced Classification of Alzheimer's Disease. *IEEE Sens. Lett.* **8**, 1–4 (2024).
24. Matlani, P. BiLSTM-ANN: early diagnosis of Alzheimer's disease using hybrid deep learning algorithms. *Multimedia tools and applications*, **83** (21), 60761–60788 (2024).
25. Roy, S., Jain, P. K., Tadepalli, K. & Reddy, B. P. Forward attention-based deep network for classification of breast histopathology image. *Multim. Tools Appl.* **83**, 88039–88068 (2024).
26. Kumar, S. Alzheimer MRI dataset. https://www.kaggle.com/datasets/marcopinamonti/alzheimer-mri-4-classes-dataset?utm_source. Accessed 18 May 2022.
27. Alzheimer's Disease Neuroimaging Initiative (ADNI). ADNI data. https://adni.loni.usc.edu/data-samples/access-data/. Accessed 18 May 2021.
28. Kalirajan, K. An intelligent magnetic resonance imagining-based multistage Alzheimer's disease classification using swish-convolutional neural networks. *Med. Biol. Eng. Comput.* **63** (3), 885–899 (2025).
29. Peng, H. et al. Semantic segmentation of litchi branches using DeepLabV3 + model. *IEEE Access.* **8**, 164546–164555 (2020).
30. Zhang, J., Yu, X., Lei, X. & Wu, C. A novel deep LeNet-5 convolutional neural network model for image recognition. *Comput. Sci. Inform. Syst.* **19** (3), 1463–1480 (2022).
31. Naeem, S. et al. Machine-learning based hybrid-feature analysis for liver cancer classification using fused (MR and CT) images. *Appl. Sci.* **10**(9), 3134 (2020).
32. Pant, G., Yadav, D. P. & Gaur, A. ResNeXt convolution neural network topology-based deep learning model for identification and classification of Pediastrum. *Algal Res.* **48**, 101932 (2020).
33. Shobha, S. & Karthikeyan, B. R. Classification of Alzheimer's disease using transfer learning and support vector machine. *Int. J. Intell. Syst. Appl. Eng.* **12** (15s), 498–508 (2024).
34. 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).
35. Sharma, S., Guleria, K., Tiwari, S. & Kumar, S. A deep learning based convolutional neural network model with VGG16 feature extractor for the detection of Alzheimer Disease using MRI scans. *Meas. Sens.* **24**, 100506 (2022).
36. Rana, M. M. et al. A robust and clinically applicable deep learning model for early detection of Alzheimer's. *IET Image Proc.* **17** (14), 3959–3975 (2023).

Acknowledgements

Author contributions

All authors contributed equally.

Funding

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this article. The ADNI study was approved by the institutional review boards of all participating institutions, and written informed consent was obtained from all participants or their authorized representatives. This article does not contain any studies with human participants or animals performed by any of the authors.

Additional information

Correspondence and requests for materials should be addressed to B.N.J.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.