



## OPEN An optimized hybrid deep learning model to detect Alzheimer disease

A. Sundar Raj<sup>1</sup>✉, C. Gunasundari<sup>2</sup>, S. Senthilkumar<sup>3</sup> & S. Sivamani<sup>4</sup>

Alzheimer's is a serious neurodegenerative disease that requires early detection for effective intervention. Traditional methods often struggle with accurately identifying the early stages, such as mild cognitive impairment (MCI), due to limitations in feature extraction and classification. To address these challenges, we present an optimized hybrid deep learning model for Alzheimer's disease detection. Our model employs the Inception v3 algorithm for initial feature extraction and the ResNet 50 algorithm for classification. Additionally, we optimize the network parameters using the Adaptive Rider Optimization (ARO) algorithm to enhance detection performance. Experimental analysis using a benchmark dementia dataset demonstrates that our model achieves superior accuracy of 96.6%, precision of 98%, recall of 97%, and F1-score of 98%, outperforming state-of-the-art techniques.

**Keywords** Alzheimer disease detection, Mild cognitive impairment, Deep learning algorithm, Inception v3, ResNet 50, Optimal features, Adaptive rider optimization (ARO) algorithm

Millions of people are affected worldwide due to an irreversible neurodegenerative disorder. Alzheimer, the serious disease, which collapses the cognitive functions and makes the patient's daily life a challenging one. Alzheimer's disease will be the primary form of dementia, encompassing approximately 60–80% of all dementia cases. Alzheimer patients will have memory loss, faces difficulty in finding suitable word to talk, repetitive questioning, difficult to take decisions, difficulty in food intake, lack of thinking, etc., The statistics of Alzheimer Association<sup>1</sup> reports that more than 6 million American's have Alzheimer disease. Specifically, these 6 million includes all age group of people. In 2023, total 6.7 million of Americans have Alzheimer in which 73% of people are above 65 years old. As of now there is no complete solution to cure Alzheimer whereas the early diagnosis and intervention are considered as most important in medicinal diagnostic procedures.

The progressive stages of Alzheimer's are defined in three phases: normal control, mild cognitive impairment, and finally Alzheimer's. In all these stages, slight changes occur in the human brain structure. Mild Cognitive Impairment (MCI) is considered the initial stage of Alzheimer's and is crucial for clinical analysis. Early symptoms of Alzheimer's can be studied from MCI, allowing for the delay of Alzheimer's deterioration with proper medications. The symptoms of MCI include mild behavioral changes visible to others<sup>2</sup>. However, confirming and diagnosing Alzheimer's phases require multiclass clinical decisioning and expertise in radiological analysis.

Neuroimaging modalities such as positron emission tomography (PET), computerized tomography (CT), and magnetic resonance imaging (MRI) are widely used for Alzheimer's diagnosis<sup>3</sup>. Among them, MRI visualizes images in three phases, providing a three-dimensional structure of the brain. MRI is successfully used in computer-aided diagnosis, specifically in Alzheimer's detection. Computer-aided systems segment the images into gray matter and white matter to classify distinct phases of Alzheimer's. The performance of computer-aided diagnosis systems has improved in recent times due to innovative Artificial Intelligence (AI) technologies. The detection speed and accuracy of the diagnosis procedure are enhanced with the incorporation of AI techniques. However, familiar challenges in Alzheimer's detection include processing low-resolution image acquisition<sup>4,5</sup>, limitations in segmentation procedures, unavailability of comprehensive datasets, uncertainty between MCI/AD and NC/NCI, and a lack of knowledge in region of interest (ROI) identification.

The initial stages of Alzheimer's detection incorporate machine learning algorithms for disease classification<sup>6</sup>. Later, deep learning-based medical image analysis is presented in numerous research works<sup>7,8</sup>. Alzheimer's detection using deep learning algorithms directly extracts features from images and provides better precision in disease detection<sup>9–12</sup>. Although the performance is better, research towards developing a better system that can produce maximum accuracy in Alzheimer's disease detection is still in progress. To increase the performance of deep learning models, the hyperparameters of the learning algorithm should be selected precisely. Optimal

<sup>1</sup>Department of Biomedical Engineering, E.G.S. Pillay Engineering College, Nagapattinam, Tamilnadu 611002, India.

<sup>2</sup>School of Computing, SRM Institute of Science and Technology, Tiruchirappalli, Tamilnadu, India. <sup>3</sup>Department of Electronics and Communication Engineering, E.G.S. Pillay Engineering College, Nagapattinam, Tamilnadu 611002, India.

<sup>4</sup>Department of Electrical and Electronics Engineering, E.G.S. Pillay Engineering College, Nagapattinam, Tamil Nadu 611002, India. ✉email: drasr18@gmail.com

values will make the learning algorithm work efficiently and produce better results in disease detection. Thus, in this research work, an optimized hybrid deep learning model is presented for Alzheimer's disease detection. The objective of this research work is to attain improved performance in the early-stage detection of Alzheimer's disease. The contributions made in this research work are presented as follows.

- Proposed a novel hybrid deep learning model combining Inception v3 for multi-scale feature extraction and ResNet-50 for robust classification, specifically tailored for Alzheimer's disease detection from MRI images.
- Introduced Adaptive Rider Optimization (ARO) for hyperparameter tuning, which enhances the training performance by dynamically adjusting key parameters such as learning rate, batch size, number of epochs, and dropout rate.
- Demonstrated the superiority of ARO over conventional optimizers (e.g., Adam, RMSprop, Bayesian optimization) by effectively escaping local minima and improving convergence behavior through its rider-based behavioral modeling strategy.
- Developed a two-stage training strategy: initial feature extraction using frozen Inception v3 weights followed by fine-tuning and classification through ResNet-50, thereby leveraging transfer learning while reducing overfitting.
- Handled dataset imbalance using targeted data augmentation techniques (flipping, rotation, brightness adjustment) applied only to underrepresented classes, ensuring model generalization without inducing data leakage.
- Achieved statistically validated performance improvements using confidence intervals and paired t-tests across all key metrics (accuracy, precision, recall, F1-score, specificity), confirming the robustness and reliability of the proposed model.
- Evaluated the model using a publicly available four-class Alzheimer's MRI dataset from Kaggle, with stratified data splitting and a detailed experimental setup that ensures reproducibility and fair comparison.
- Presented a thorough comparative analysis against multiple state-of-the-art CNN and hybrid models, demonstrating consistent outperformance across all evaluation metrics.

The rest of the discussions are arranged in the following structure: The literature review provided in Sect. 2 offers insight into recent methodologies in Alzheimer's disease detection. Section 3 presents the hybrid deep learning algorithm and the optimization model proposed in this research work. The experimental results based on benchmark datasets and the comparative analysis are presented in Sect. 4. The findings and the summary of the research work are presented in Sect. 5.

## Related works

The research on Alzheimer's detection has been ongoing for over a decade. This section presents some recent research works considered for this study. In certain studies, Alzheimer's is identified based on visible symptoms. For example, in research work<sup>13</sup>, the disorder is measured and categorized into Alzheimer's stages based on the difficulty level in speech. However, extracting sufficient linguistic features for disease detection is challenging and doesn't provide accurate results. The machine learning-based Alzheimer's detection model presented in<sup>14</sup> includes decision tree, k-NN, and CNN as classifiers. The initial features, such as texture, scale-invariant transform, and histogram, are obtained from the preprocessed brain image. Additionally, a group grey wolf optimization model is incorporated to optimize the classifier parameters. The experimental results reveal that the presented model achieved better classification accuracy than conventional machine learning techniques.

The Alzheimer's detection model presented in<sup>15</sup> considers the foot movement of elders using IoT devices and classifies various stages of Alzheimer's. The presented model initially captures the foot movements using IoT devices and compares the movements using the dynamic time warping algorithm. The middle-level cross-identification function is used to identify the differences and classify them. Further, using machine learning algorithms like k-NN, SVM, and inertial navigation algorithms, the experimental results of the proposed model are comparatively analyzed to demonstrate superior performance in Alzheimer's detection. An ensemble learning model presented in<sup>16</sup> considers machine learning algorithms for Alzheimer's detection. The study includes SVM, random forest, and naive Bayes algorithms, analyzing the performances of these learning algorithms in terms of precision and computational effectiveness to demonstrate their effectiveness in Alzheimer's detection.

The Alzheimer's detection model reported in<sup>17</sup> presents machine learning-based analysis which considers MRI images and selects optimal features using the Shearlet Transform (ST)-based feature selection procedure. The classification procedure includes k-NN for final feature classification and achieves better accuracy over conventional machine learning techniques. A modified SVM-based Alzheimer detection presented in<sup>18</sup> initially analyzes the conventional SVM model's performance. Further, using a linear kernel model, the parameters of SVM are modified and analyzed. Due to the linear kernel, the nonlinear boundaries are effectively managed, and overfitting is avoided in the classification process. The results reveal that the performance of SVM with a linear kernel outperformed the conventional SVM-based classification.

The automated diagnosis procedure presented in<sup>19</sup> initially preprocesses brain MRI images using the Lucy-Richardson approach for noise removal. Gaussian noise in the brain image and blurred regions is removed by examining each pixel. Further, the image is segmented using an adaptive exclusive analytical atlas approach, and statistical features are extracted using GLCM. Finally, the extracted features are classified to categorize various stages of Alzheimer's. Experiments using a benchmark dataset demonstrate the better accuracy and sensitivity of the presented model over conventional techniques. A comparative analysis of different CNN models and hybrid models is presented in<sup>20</sup> for Alzheimer's detection. The research work includes 3D-CNN, 2D-CNN, and a hybrid 3D-CNN that combines SVM with CNN architecture for Alzheimer's analysis. The experimental results of hybrid 3D-CNN are better over traditional models. The deep learning-based Alzheimer's detection

model reported in<sup>21</sup> includes CNN for Alzheimer's detection. The classification process includes two modules in which the 2D and 3D brain images are processed through a conventional CNN model to detect Alzheimer's. The second model includes a transfer learning procedure and incorporates VGG19 for final classification. Finally, the performances are comparatively analyzed, and it is observed that the transfer learning-based VGG19 outperformed the conventional CNN-based disease classification.

The deep learning-based Alzheimer's detection presented in<sup>22</sup> analyzes different modalities like MRI images, genetic data, and clinical test records for categorizing the symptoms of Alzheimer's, MCI, and normal control. Initial features are extracted using a stacked denoising autoencoder, which utilizes genetic and clinical data. Features from the image data are extracted using 3D-CNN. From the features, optimal features are selected through clustering and perturbation procedures and classified using random forest, SVM, decision tree, and k-NN. Due to multi-modality feature analysis, the presented model attains better detection performance. However, the computational complexity of the presented model is high compared to similar deep learning models. The three-dimensional CNN-based Alzheimer classification model presented in<sup>23</sup> utilizes PET images to categorize various stages of Alzheimer's. A layer-wise relevance propagation model is presented to identify the most contributing portions in PET images. Further, features are extracted and classified using 3D CNN with better accuracy and precision over machine learning-based analysis.

The CNN-based Alzheimer's detection model presented in<sup>24</sup> fine-tunes the network hyperparameters using a transfer learning procedure to attain better domain knowledge in disease detection. Additionally, to increase the training data samples, a generative adversarial network is incorporated into the presented model, which generates minority class samples. The final classification reveals that the presented model achieves better accuracy over conventional CNN-based detection approaches. The Alzheimer's detection model presented in<sup>25</sup> proposes a transfer learning model based on AlexNet. The multi-class classification procedure modifies the architecture of conventional AlexNet layers. The AlexNet layers were modified with a fully connected layer, SoftMax layer, and an output layer to differentiate dementia stages. However, the attained accuracy of the modified AlexNet is comparatively lower than that of recent techniques.

An EfficientNet architecture-based Alzheimer detection procedure presented in<sup>26</sup> includes an effective composite scaling method to enhance the network layer width, depth, and resolution. Due to increased network layers, neurons, and picture resolution, better accuracy was attained by the presented model. The Alzheimer detection model presented in<sup>27</sup> includes a modified deep learning algorithm to classify EEG signals. The presented dual-input convolution encoder network initially extracts the coherence and band power features. Further, the features are classified using convolution and transformer encoders. The complex features are captured effectively in the classification process, due to this enhanced detection performance obtained by the presented model over conventional approaches.

The hybrid model presented for Alzheimer's detection utilizes a local feature descriptor to detect early-stage symptoms of MCI. The Hessian detector and binary pattern texture operator are combined to form the local descriptor, and finally, the classification is performed using a simple CNN. The experimental results present better classification accuracy which is comparatively lower than the recent state-of-the-art techniques. The hybrid model presented in<sup>28</sup> combines different CNN models as an ensemble learning algorithm for Alzheimer's detection. The presented model performs classification between health and Alzheimer, MCI and health, and MCI and Alzheimer as a binary classification task using individual CNN models. Finally, these CNN models are combined into an ensemble model and applied fivefold cross-validation to validate the classification results. However, the obtained results are comparatively lower than the results achieved by recent approaches.

The similar ensemble model presented in<sup>29</sup> includes VGG16 and EfficientNetB2 for early-stage detection of Alzheimer's. The presented model initially performs data balancing using an adaptive synthetic oversampling technique. Experimentation on both balanced and unbalanced datasets reveals a higher classification accuracy of the proposed model for both binary and multi-class classification results, which is superior to existing methods. From the literature study, it can be observed that the performance of hybrid deep learning algorithms in Alzheimer detection is better than that of machine learning algorithms. However, the hyperparameter selection of deep learning algorithms requires highly expertise in the learning procedure, and inaccurate selection of hyperparameters reduces accuracy in the classification process. Thus, in this research work, an optimized hybrid deep learning model is presented for Alzheimer detection.

The optimization of Osprey Gannets using retinal fundus images is based on transfer learning (TL) is proposed in<sup>30</sup>. To begin, a bilateral filter is used to enable the input fundus picture into the preprocessing procedure. Afterwards, the Osprey Gannet-active counter model that was trained using OGO is employed for OD detection. Together, Osprey Optimization (OO) and Gannet Optimization (GO) form OGO. At the same time, Res-UUNet, which is trained using the Jaya Chronological Chef-Based Optimization Algorithm (Jaya-CCBOA), performs blood vessel segmentation (BVS) on the pre-processed image. After that, the OD detection phase and BVS are used to carry out the feature extraction. While doing so, a number of feature extractors on the input image are used. After that, Tanimoto similarity is used to choose features. Lastly, TL trained using OGO's convolutional neural network (CNN) achieves normal or hypertensive CV risk identification. Maximum values of 92.1%, 91.5%, 93.1%, 87.9%, and 87.9% were attained by the measurements used in OGO\_CNN based TL.

A new approach to the detection of CVD using the analysis of retinal fundus images is presented in<sup>31</sup>. Disease prediction is made easier by the method being considered. Extracting tissue data for the purpose of identifying and treating CVD is the principal technique entailed in retinal vessels. At this stage, a Gaussian filter is applied to the retinal images. Binarization and circle fitting allow for optic disc identification, which is subsequently followed by statistical data extraction. Res-Unet, which is based on the Chronological Chef Based Optimization Algorithm (CCBOA), is used to segment blood arteries. Afterwards, characteristics of the texture are retrieved. When combined with CCBOA, a deep neural network (DRN) can identify CVD. With an accuracy rate of 89.8%, the CCBOA-based DRN proved to be quite efficient.

In order to effectively identify and classify lung cancer, a deep learning strategy is introduced in<sup>32</sup> based on Political Squirrel Search Optimization (PSSO). When a Deep Neuro Fuzzy Network (DNFN) classifier predicts malignant spots in the lung lobes, regions are segmented using Spine General Adversarial Network (Spine GAN). The different stages of cancer severity are also determined using a Deep Residual Network (DRN). A recently announced method called PSSO was developed by combining the Political Optimizer (PO) and Squirrel Search Algorithm (SSA). The Lung Image Database Consortium's image dataset is used to evaluate the results of the experiments.

A system for optimising the scheduling of electric vehicle (EV) charge operations on a VANET topology is presented in<sup>33</sup>. The first step is to run a network simulation with the CS and the EV sites. Power predictions are made using the Deep Maxout network (DMN). The vehicles are scheduled to charge at the charging station according to a number of parametric parameters. When planning to charge EVs, the suggested Fractional Feedback Tree Algorithm (FFTA) comes in default. The suggested FFTA provided better performance with a minimum charging distance of 11.00 km, a maximum power of 8.17 J, an average waiting time of 0.47 s, and a lower number of EVs charged at 4.4.

A Hybrid Whale and Gray Wolf Deep Learning Optimization Algorithm is developed in<sup>34</sup> for predicting Alzheimer's Disease. This method utilizes a combination of Whale Optimization Algorithm (WOA) and Gray Wolf Optimization (GWO) to segment brain subregions effectively, such as gray matter and hippocampus, crucial for diagnosing Alzheimer's. They tested their approach using T2-weighted MR images from Chettinad Health City, achieving an accuracy of 92% in segmentation tasks and 90% in classification accuracy when distinguishing between normal and Alzheimer's affected brain images. Although the method shows promising results, it does face challenges such as high computational demands and the complexity of optimally tuning multiple parameters, which might limit its practical deployment in real-time medical settings.

Sophisticated four-phase method for early detection of Alzheimer's Disease is introduced in<sup>35</sup>, utilizing a combination of preprocessing, feature extraction, selection, and classification with a deep learning model based on an ensemble of classifiers. They apply a novel ensemble approach combining Bi-GRU, Multi-Layer Perceptron (MLP), and Quantum Neural Network (QDNN) optimized through an Enhanced Math Optimizer Accelerated Arithmetic Optimization (EMOAOA) technique, which adjusts the weights of the QDNN model. This method significantly boosts detection accuracy across various metrics, achieving impressive accuracies of up to 98.7% in late stages. The approach was validated using a large dataset with 13,916 data points from the ADNI database, representing high sensitivity and specificity. Despite the high performance, the method involves complex parameter optimization, which could be computationally expensive and challenging to deploy in real-time clinical environments.

A hybrid deep learning framework that employs Generative Adversarial Networks (GANs) and Deep Convolutional Neural Networks (CNNs) is presented in<sup>36</sup> to predict the progression of Alzheimer's disease. Utilizing the capability of GANs to generate synthetic MRI images, the study addresses the issue of insufficient data for training robust models. The authors employ a cascading method of DCGANs for generating synthetic images and SRGANs to enhance the resolution of these images, subsequently using CNNs for disease classification. Tested on a dataset from the Alzheimer's Disease Neuroimaging Initiative, the framework achieved a prediction accuracy of 99.7%, demonstrating significant improvements over previous methods.

An ensemble approach for Alzheimer's disease detection is proposed in<sup>37</sup> combining VGG16 and EfficientNet models using adaptive synthetic oversampling techniques for dataset balancing. This hybrid approach utilizes deep learning to enhance diagnostic accuracy significantly. By applying their model to both balanced and imbalanced datasets of MRI images, classified into categories like mild demented and moderate demented, they achieved high accuracy and AUC scores. The method addresses the challenge of data imbalance in medical imaging datasets which often model performance. However, the complexity and computational demands of managing and processing such high-dimensional data, coupled with the intensive training requirements of deep learning models, present potential barriers to routine clinical application.

A novel Alzheimer's disease detection methodology using MRI data and deep learning techniques is introduced in<sup>38</sup>. The research team developed a variety of models both with and without data augmentation, aiming to achieve optimal detection accuracy. Employing a dataset comprising images of brain MRIs, their experimentation demonstrated the CNN-LSTM model as the most effective, achieving a remarkable accuracy of 99.92%. Despite these successes, this study also acknowledges potential challenges such as high computational costs and complexities in optimally tuning the model's extensive parameters, which could hinder its practical deployment in clinical environments.

Previous hybrid methods for Alzheimer's detection often suffer from key limitations such as manual feature fusion, rigid architecture design, and suboptimal hyperparameter tuning. Many approaches combine CNN backbones without effectively leveraging their complementary strengths or rely on fixed heuristic parameters, leading to constrained learning capacity and reduced adaptability to data variability<sup>39–41</sup>. Moreover, some models lack mechanisms to address class imbalance or overfitting, resulting in poor generalization. In contrast, our proposed method integrates Inception v3 and ResNet-50 in a structured and sequential manner, allowing multi-scale feature extraction followed by deep residual classification. Crucially, we employ Adaptive Rider Optimization (ARO) to dynamically fine-tune hyperparameters, improving convergence, reducing training bias, and enhancing performance across all classes. This strategic integration addresses the architectural inflexibility and optimization inefficiencies seen in earlier hybrid models.

## Proposed work

The proposed Alzheimer detection model includes deep learning algorithms like Inception v3 and ResNet 50. An adaptive rider optimization (ARO) is included for optimizing the network parameters and to improve the

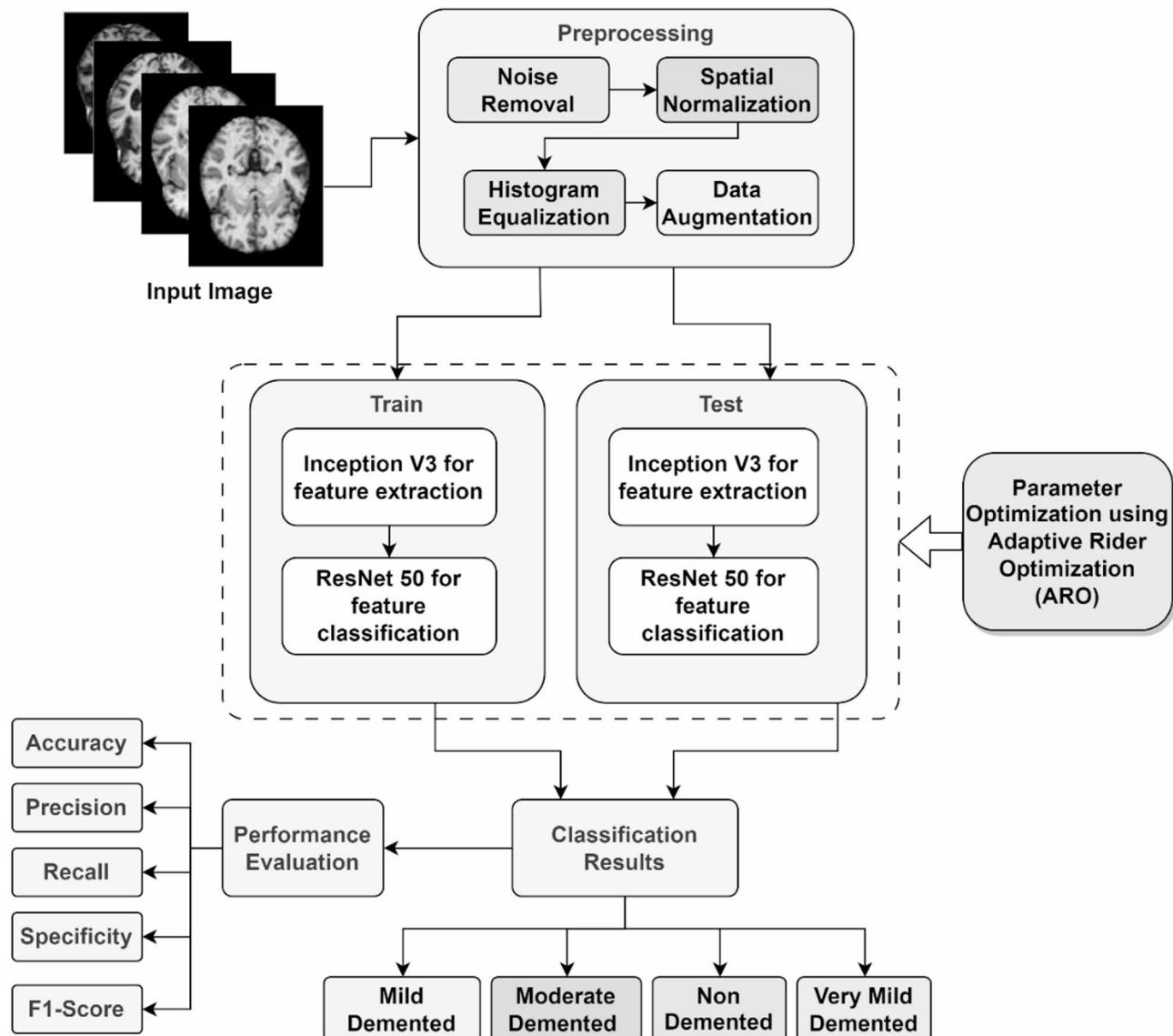
detection performances. The complete overview of the proposed Alzheimer detection model is presented in Fig. 1.

### Data preprocessing

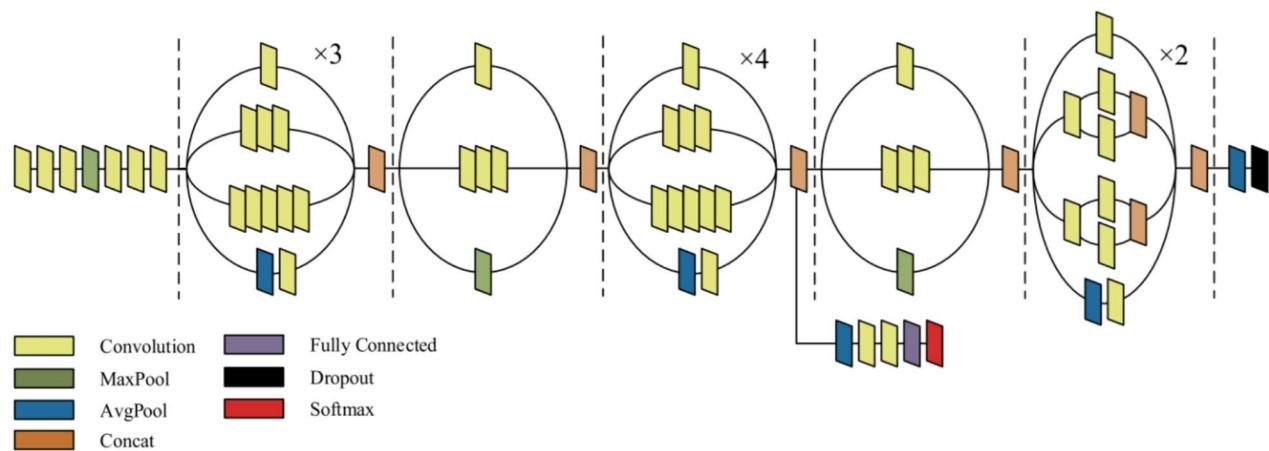
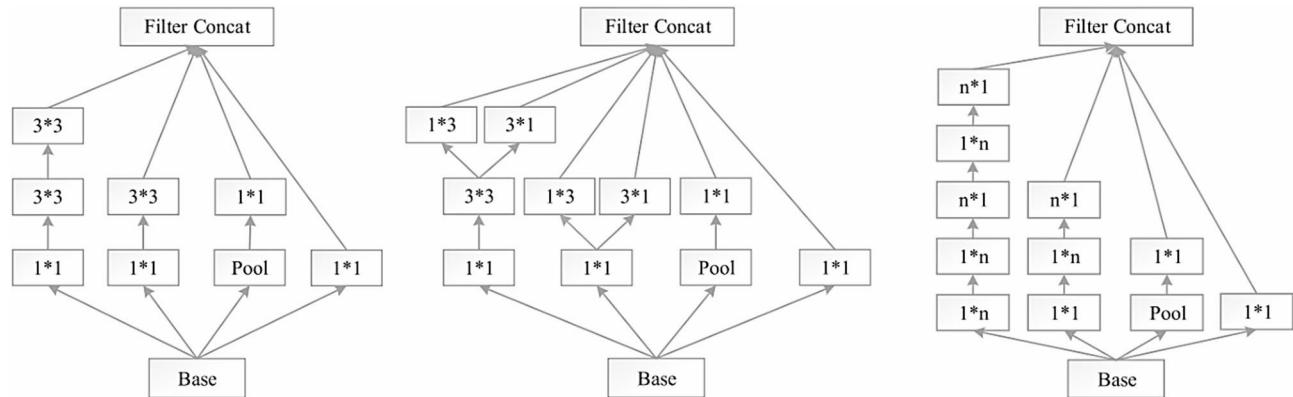
The MRI brain image is initially preprocessed to remove the noise. For noise removal temporal high pass filtering is used and then normalization is performed in which spatial normalization is performed using a linear transformation. Since deep learning algorithms requires a substantial number of samples, in addition to data normalization, data augmentation is performed in the preprocessing steps to attain substantial number of input samples. Data augmentation includes techniques like horizontal flipping, vertical flipping, image rotation at 90 degree and 270-degree, image brightness enhancement procedure. For enhanced contrast, Histogram Equalization is also included in the preprocessing procedure. From the preprocessed data, the features are extracted and classified using hybrid deep learning algorithm.

### Hybrid deep learning model for Alzheimer detection

The hybrid deep learning model used in the proposed work for Alzheimer detection includes Inception v3 and ResNet 50 for feature extraction and classification. Inception v3 is primarily a network developed by Keras. The Network process the input through three channels of size  $299 \times 299$ . Compared to previous versions of inception network, inception v3 splits large volume integrals into small convolutions using convolutional kernel splitting method which reduces the number of parameters and accelerates the spatial feature extraction process. The architecture of inception v3 is presented in Fig. 2. The architecture of Inception v3 used for multi-scale feature



**Fig. 1.** Alzheimer detection using optimized hybrid deep learning algorithm.

**Fig. 2.** Inception v3 network.**Fig. 3.** Inception module in Inception v3.

extraction in the proposed model is illustrated in Fig. 2, highlighting the stacked inception modules and flow of convolutional operations.

Inception v3, three different sized grids inception structure is used to optimize the network. Figure 3 depicts an illustration for three inception models used in architecture. Figure 3 depicts an illustration for three inception models used in architecture.

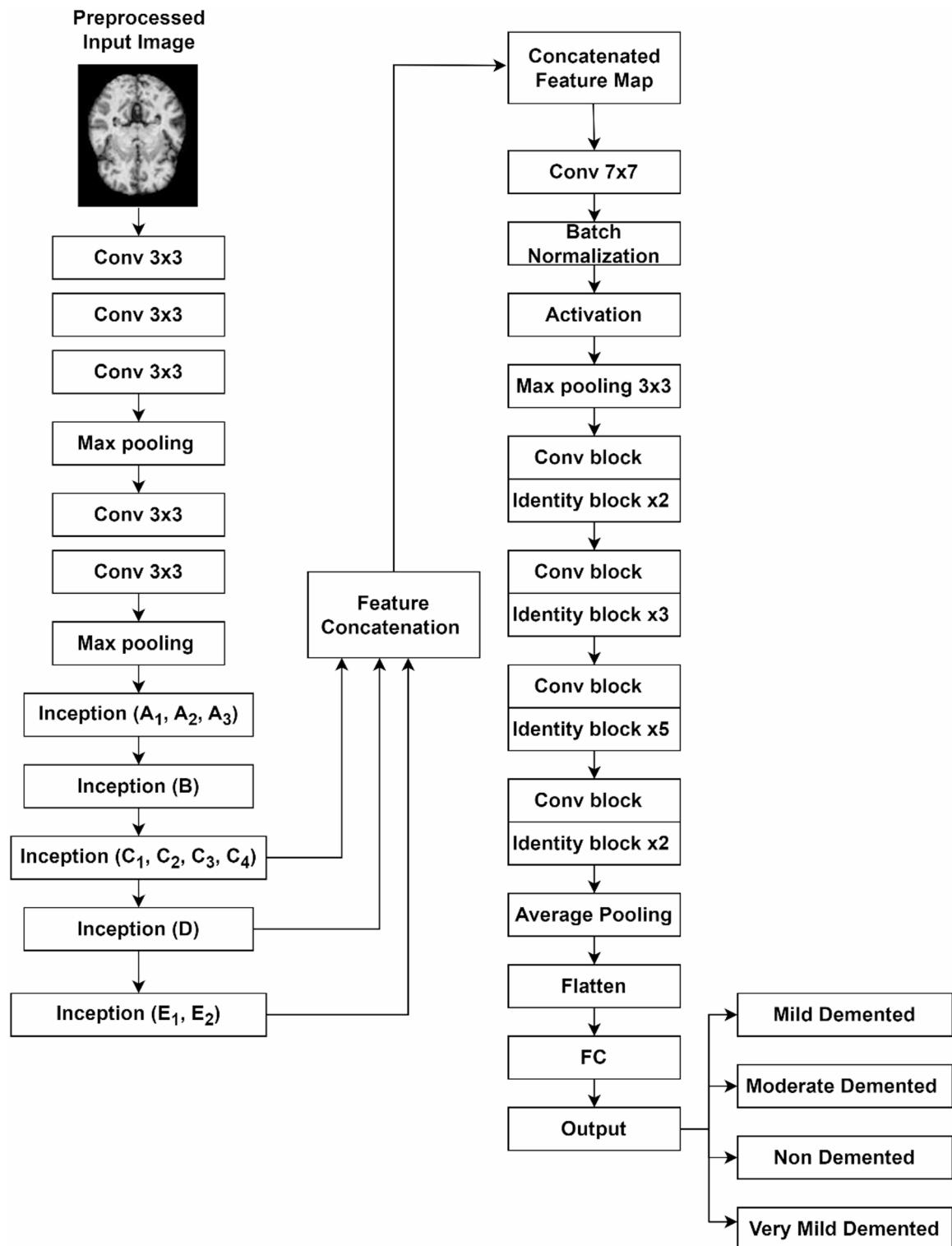
Inception v3 is an advanced convolutional neural network (CNN) architecture known for its efficient and effective feature extraction capabilities. It uses a combination of convolutions, pooling operations, and inception modules to capture multi-scale spatial features from input images. Inception modules allow the network to process information at various scales, combining the outputs of multiple convolutions with different filter sizes. This multi-scale approach helps in capturing fine-grained details as well as broader contextual information, making it highly suitable for analyzing complex MRI images of the brain. In our model, Inception v3 is utilized to extract a comprehensive set of features from MRI images. These features include edges, textures, and patterns that are indicative of different stages of Alzheimer's disease. The output of Inception v3 is a high-dimensional feature map that serves as the input for the subsequent classification stage. In the proposed inception v3 is used to extract the initial features, the architecture includes 11 inception modules. Each inception modules includes a convolution layer, pooling layer, activation layer and batch normalization layer. The multi-scale maximum features are obtained from the input image using these inception modules. The feature extracted from the input image are concatenated into a feature map and then fed into the ResNet50 module for final classification.

The classification module used in the proposed architecture is ResNet 50 which comprises residual modules. These residuals are used to avoid network degradation problem and reduce error in the classification process. Moreover, the residual modules used in stage 2 and 5 overcomes the vanishing gradient issue. The architecture of ResNet50 is divided into 5 blocks. The residual blocks preserve the previous layer information which allows the network to learn better representations of input data.

The convolution layer in the ResNet performs convolution operation. Followed by convolution, max pooling is used to down sample the convolution layer output. Next to max pooling, series of residual blocks are used which includes two convolution layers, followed by batch normalization layer, and rectified linear unit. The

second convolution layer output is added to the residual block input and then passed through ReLU activation function. This process is repeated for the remaining residual blocks and then finally a fully connected layer is used that maps the last residue block output into different output classes. The complete architecture of the hybrid deep learning model is presented in Fig. 4. Figure 4 presents the overall hybrid deep learning architecture combining Inception v3 and ResNet-50, with sequential integration for feature extraction and classification.

In the proposed hybrid deep learning architecture, Inception v3 was employed as a feature extractor, and ResNet-50 was used as the classifier. The two networks were integrated sequentially, where features extracted



**Fig. 4.** Proposed hybrid deep learning model.

from the output of the final global average pooling layer of Inception v3 were passed directly to the input layer of ResNet-50 for classification. During the initial phase of training, Inception v3 was used with frozen weights (i.e., non-trainable) to retain the benefit of pretrained convolutional features learned on ImageNet, which helps in reducing overfitting and training time when data is limited. After achieving initial convergence, fine-tuning was selectively enabled for the deeper layers (i.e., layers beyond the 249th layer of Inception v3), allowing the model to adapt to the domain-specific features of Alzheimer's MRI data. ResNet-50, on the other hand, was trained from scratch, as it served as the classification backbone with custom fully connected layers tailored to the four-class Alzheimer's dataset. This two-stage strategy of freezing then fine-tuning helped balance computational efficiency with model accuracy, ensuring that lower-level features were preserved while allowing domain-specific adaptation in higher layers.

### Adaptive rider optimization (ARO)

Optimizing the hyperparameters of deep learning models is essential for achieving high performance. Proper optimization ensures that the network parameters are set to values that maximize the model's ability to generalize from training data to unseen data. Adaptive Rider Optimization (ARO) is a novel optimization algorithm inspired by the behavior of riders in reaching their destination through various strategies such as following, bypassing, overtaking, and attacking. ARO is designed to enhance the search capability and execution efficiency of optimization processes. In our research, ARO is employed to fine-tune critical hyperparameters of the deep learning model, including learning rate, batch size, number of epochs, and dropout rate. The adaptive nature of ARO allows it to dynamically adjust these parameters during training, improving the convergence speed and overall performance of the model.

To attain better classification performance in Alzheimer detection, the classifier network parameters are optimized using Adaptive Rider Optimization (ARO) algorithm. The optimization algorithm is formulated based on the riders who are aimed to reach the destination. There are four categories of riders in the optimization problem such as followers, bypass riders, over takers and attack riders. By neglecting the leader route, the bypass rider reaches the destination point. The follower category of riders chases the lead rider, and the over takers make their own decisions in riding. The attackers make assault moves with full velocity to reach the destination. The efficiency and the ability of a rider to reach the target is considered as optimal solution for the given problem. In the proposed work, the optimal parameters are selected based on the optimal solution. An adaptive strategy is also incorporated to enhance the search ability and reduce the execution time of the optimization algorithm.

The mathematical model of ARO considers riders as the optimal solution and to initialize the rider group, the parameters like number of riders, riding period instance, riding dimension and location are considered. Mathematically the initialization process is formulated as

$$x^t = \{x^t(i, j)\}, \text{ for } 1 \leq i \leq R; 1 \leq j \leq Q \quad (1)$$

where number of riders are indicated using  $R$ , riding instance is indicated using  $t$ , dimensions are indicated using  $Q$  and location of rider at time is indicated using  $x^t$ . After initialization, using candidate solution direction, opposite solutions are generated which improve the solution diversity and search space. Mathematically the solution creation process is formulated as

$$\vec{s} = \vec{L}_b + \vec{U}_b + \vec{s}^t \quad (2)$$

where the lower and upper bands are indicated using  $L_b$  and  $U_b$ , the candidate solution is indicated using  $s^t$ . Based on the initial solution and the generated opposite solution, the fitness of the riders is obtained. Further the riders' positions are updated as follows. The bypass rider position is updated considering the location of leading rider as follows

$$x_b^{t+1}(i, j) = \delta [x_t(\eta, j)] * \beta(j) + x^t(\xi, j) * [1 - \beta(j)] \quad (3)$$

where the arbitrary significance  $\delta$  range is given as [0,1] and the random number  $\beta$  range is given as [0,1] and  $\eta$  range is given as [1, R]. The numeric  $\xi$  selection range is given as 1 to R. The follower position update in the optimization process is based on the riding path of the bypass rider. Mathematically the position update is formulated as

$$x_f^{t+1}(i, k) = x_b(l, k) + [\cos(T_{i,k}^t) * x_b(l, k) * d_i^t] \quad (4)$$

where location of bypass rider is given as  $x_b$ , riders index is given as  $l$  and the location operator is given as  $k$ . The rider's steering angle is given as  $T_{i,k}^t$  and the rider travel length is given as  $d_i^t$ . The position update of over taker is presented based on the information obtained from the bypass rider and their own decision as follows

$$x_o^{t+1}(i, k) = x^t(i, k) + [D_I^t(i) * x_b(l, k)] \quad (5)$$

where the rider location and velocity coordinates are indicated using  $x^t(i, k)$ , direction indicator is given as  $D_I^t(i)$  and it is formulated as

$$D_I^t(i) = \left[ \frac{2}{1 - \log(s_R^t(i))} \right] \quad (6)$$

where the success rate of rider is given as  $s_R^t$ . The position update of attacker is formulated as

$$x_a^{t+1}(i, j) = x^L(L, j) + [\cos(T_{i,j}^t) * x_b(L, j) * d_i^t] \quad (7)$$

where the leader location is indicated as  $x^L(L, j)$  and the steering angle is given as  $T_{i,j}^t$ . The attacker travels at maximum speed to reach the target. The travelling style of follower is identical to the attacker however the attacker considers all the rider's position and other attackers' position in the update process. The algorithm terminates on the optimal solution obtained for the classifier. Summarized pseudocode for the proposed adaptive rider optimization is given as follows.

Pseudocode for adaptive rider optimization algorithm
<i>Input:</i> Rider random position
<i>Output:</i> selecting the leading rider (optimal parameters)
<i>Begin</i>
<i>Initialize the population of riders</i>
<i>Modify rider parameters like accelerator, steering angle, brake</i>
<i>Obtain the success rate</i>
<i>Calculate each rider fitness</i>
<i>While</i> $t < toff$
<i>For</i> $i = 1$ to $R$
<i>Update the position of bypass rider using</i> $x_b^{t+1}$
<i>Update the position of follower rider using</i> $x_f^{t+1}$
<i>Update the position of over taker rider using</i> $x_o^{t+1}$
<i>Update the position of attacker rider using</i> $x_a^{t+1}$
<i>Rank the riders based on success rate</i>
<i>Select the rider with maximum success rate</i>
<i>Update rider parameters</i>
<i>Return</i>
$t = t + 1$
<i>End all</i>

The computational complexity is assessed by examining the time and space requirements for both the Inception v3 and ResNet 50 components of the model.

Inception v3 employs a convolutional architecture with multiple inception modules. Each inception module consists of convolution layers, pooling layers, activation functions, and batch normalization. Assuming  $n$  to be the spatial dimension of the input,  $k$  to be the kernel size,  $c_{in}$  and  $c_{out}$  to be the number of input and output channels respectively, the computational complexity of each convolutional layer in Inception v3 is

$$O(n^2 * k^2 * c_{in} * c_{out}) \quad (8)$$

Given that Inception v3 processes input images through 11 inception modules, the overall complexity is the sum of the complexities of all convolutional layers within these modules. On the other hand, ResNet 50, used for classification, comprises residual blocks that include convolution layers and identity connections. Each residual block has a computational complexity similar to that of a convolutional layer given as in (8), but the presence of identity connections adds a minimal overhead. ResNet 50 consists of five stages, with each stage containing several residual blocks, contributing to the overall complexity.

The Adaptive Rider Optimization (ARO) algorithm used for hyperparameter tuning introduces additional computational overhead. The complexity of ARO is primarily influenced by the number of riders ( $R$ ), the number of dimensions ( $Q$ ), and the number of iterations ( $T$ ). The complexity of ARO can be approximated as

$$O(R * Q * T) \quad (9)$$

where each iteration involves updating the positions of riders and evaluating their fitness.

Combining the complexities of Inception v3, ResNet 50, and ARO, the total computational complexity of our model can be expressed as

$$O(n^2 * k^2 * c_{in} * c_{out}) * M + R * Q * T \quad (10)$$

where  $M$  represents the total number of convolutional layers in Inception v3 and ResNet 50.

### Integration of ARO into training pipeline

The Adaptive Rider Optimization (ARO) algorithm was strategically integrated into the training pipeline to enhance the performance of the hybrid deep learning model by fine-tuning critical hyperparameters. Specifically, ARO was used to optimize four essential hyperparameters of the model:

1. *Learning Rate (α)*: Determines the step size for weight updates.
2. *Batch Size (B)*: Controls the number of training samples used in one forward/backward pass.
3. *Number of Epochs (E)*: Defines the number of full passes through the training dataset.
4. *Dropout Rate (D)*: Regulates the fraction of neurons randomly dropped during training to prevent overfitting.

The objective function used by ARO was a composite performance metric, defined as:

$$\text{Fitness} = \omega_1 \cdot \text{accuracy} + \omega_2 \cdot \text{F1 score} - \omega_3 \cdot \text{validation loss} \quad (11)$$

In Eq. (11),  $\omega_1$ ,  $\omega_2$  and  $\omega_3$  are weights set empirically (in values of 0.4, 0.4, 0.2) to balance contributions, Accuracy and F1-Score promote better classification, and Validation Loss penalizes overfitting and instability.

#### Training Integration Workflow:

1. *Initialization*: A population of virtual “riders” (candidate hyperparameter sets) was initialized within defined bounds for  $\alpha$ ,  $B$ ,  $E$ , and  $D$ .
2. *Fitness Evaluation*: For each candidate, the deep learning model (Inception v3 + ResNet50) was trained for a few epochs on a subset of the training data to estimate performance. The fitness score was computed based on the objective function.
3. *Rider Update Mechanism*: Using the core strategies of ARO through bypassing, following, overtaking, and attacking—the algorithm updated the hyperparameter values of each rider based on performance and exploration dynamics.
4. *Convergence Check*: The optimization process continued until the maximum number of iterations (set to 20) was reached or convergence was observed (i.e., negligible improvement in the objective function over consecutive iterations).
5. *Final Training*: Once the optimal set of hyperparameters was identified, the full model was retrained using the entire training dataset with these tuned parameters.

The final optimal hyperparameters identified by the ARO algorithm during training are summarized in Table 1.

This integration of ARO ensured efficient exploration of the hyperparameter space, leading to improved generalization and reduced training time compared to manual tuning or grid/random search.

The novelty of the proposed architecture lies in its sequential integration of Inception v3 and ResNet-50, strategically engineered to exploit their complementary strengths in multi-scale feature extraction and deep residual classification. Unlike conventional hybrid models that simply stack pre-trained networks, our design ensures modular feature transfer, where the output of Inception v3’s final global average pooling layer serves as the optimized input for ResNet-50, avoiding redundant feature maps and enabling a leaner yet more expressive representation. This two-stage training pipeline, which begins with frozen Inception v3 layers and fine-tunes only later stages, preserves generalizable low-level features while adapting high-level filters to Alzheimer-specific patterns. Furthermore, ResNet-50 is customized with modified fully connected layers tailored to the four-class Alzheimer dataset, ensuring better discrimination among closely related classes such as Mild and Very Mild Dementia. The proposed design surpasses earlier architectures by maintaining both computational efficiency and high classification accuracy through a tightly coupled but non-overlapping dual-stage learning strategy, thereby achieving a unique balance between generalization, specificity, and performance.

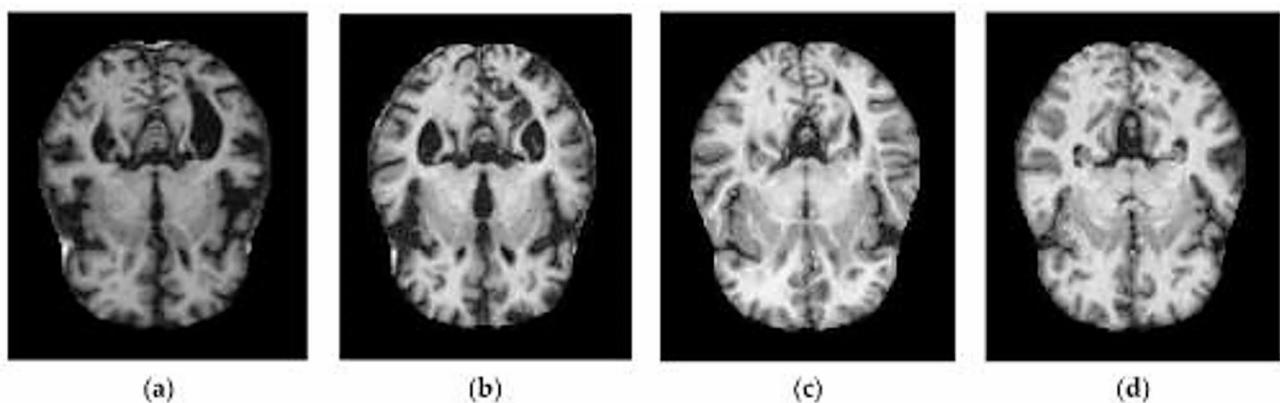
## Results and discussion

The experimental setup for evaluating the proposed hybrid deep learning model for Alzheimer’s disease detection is detailed in this section. The dataset used in this study is publicly available from the Kaggle platform and can be accessed at the following URL:

<https://www.kaggle.com/datasets/tourist55/alzheimers-dataset-4-class-of-images>. This dataset contains MRI brain images categorized into four classes: Non-Demented, Very Mild Demented, Mild Demented, and Moderate Demented. No modifications were made to the dataset other than preprocessing and augmentation as described in the methodology. Preprocessing steps include noise removal using temporal high-pass filtering, spatial normalization through linear transformation, and data augmentation techniques such as horizontal and vertical flipping, 90-degree and 270-degree rotations, and brightness enhancement. Histogram equalization is

Hyperparameter	Optimized value
Learning rate ( $\alpha$ )	0.001
Batch size (B)	32
Number of epochs (E)	100
Dropout rate (D)	0.25

**Table 1.** Optimal hyperparameters tuned by ARO.



**Fig. 5.** Input images in Alzheimer's dataset: (a) MD (b) Mo.D (c) ND and (d) VMD.

S. no	Class name	Training	Test	Total
1	Mild demented	717	179	896
2	Moderate demented	52	12	64
3	Non-demented	2560	640	3200
4	Very mild demented	1792	448	2240

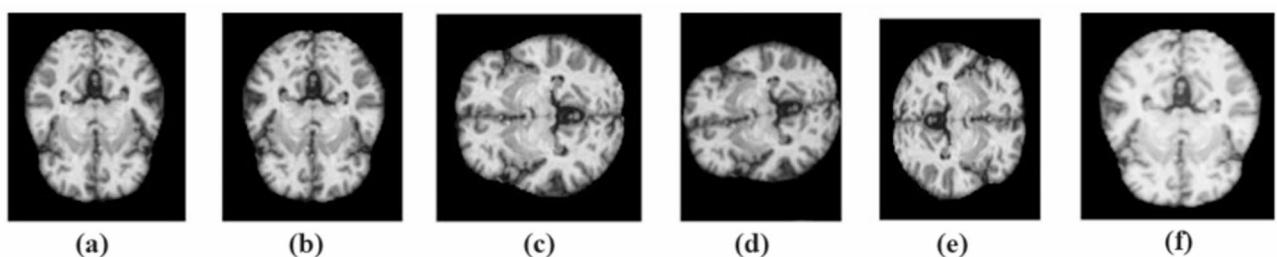
**Table 2.** Dataset description.

also applied for enhanced contrast. The hybrid model integrates Inception v3 for feature extraction and ResNet 50 for classification. Inception v3 processes input images through 11 inception modules to capture multi-scale features, which are then concatenated into a feature map and fed into the ResNet 50 module. The classification is performed using five residual blocks in ResNet 50, followed by a fully connected layer. The Adaptive Rider Optimization (ARO) algorithm is employed to fine-tune the network parameters, ensuring efficient search for optimal values. The dataset is split into training and testing sets with an 80–20 ratio, and data augmentation ensures sufficient samples for each class. The model is trained using a stochastic gradient descent optimizer with a learning rate of 0.001 and a batch size of 32 over 100 epochs, with early stopping to prevent overfitting. The model's performance is evaluated using precision, recall, specificity, F1-score, and accuracy metrics, and comparative analysis is conducted against existing models like CNN, VGG-16, ResNet-18, AlexNet, Inception v1, and various hybrid DenseNet models.

The proposed hybrid deep learning model was implemented using Python with TensorFlow and Keras libraries on a high-performance computing system equipped with an NVIDIA Tesla V100 GPU, Intel Xeon processor, and 128 GB RAM. The model was trained using a categorical cross-entropy loss function, which is standard for multi-class classification tasks. For optimization, the Stochastic Gradient Descent (SGD) optimizer was employed with a momentum of 0.9 to accelerate convergence and avoid local minima. The initial learning rate was set to 0.001, and an adaptive learning rate scheduler was applied that reduced the learning rate by a factor of 0.1 if the validation loss did not improve for 10 consecutive epochs (ReduceLROnPlateau strategy). A dropout rate of 0.25 was incorporated to prevent overfitting, as determined through the ARO-based hyperparameter optimization process. The training was conducted over a maximum of 100 epochs with early stopping enabled, monitoring validation loss with a patience of 15 epochs to prevent overfitting. The dataset from the Kaggle repository was divided into 80% training and 20% testing, and within the training set, 20% was further allocated for validation. The data splitting was performed using a stratified sampling approach to preserve class distribution across training, validation, and test sets. This ensured a balanced representation of all four classes (Non-Demented, Very Mild Demented, Mild Demented, and Moderate Demented) throughout the training and evaluation process. Data augmentation techniques were applied only on the training data to synthetically enhance the sample size and improve generalization.

The proposed model performance analysis is presented through simulation performed using python tool. The dataset utilized for experimentation is obtained from Kaggle repository. The four class Alzheimer's Dataset<sup>42</sup> include different MRI images for Mild Demented (MD), Moderate Demented (Mo.D), Non-Demented (ND) and Very Mild Demented (VMD). The sample images that represent each class are presented in Fig. 5. The number of samples in each class is presented in Table 2 for test and training process.

The data preprocessing step of proposed model includes data normalization and data augmentation. From Table 2, it can be observed that the number of samples for the Moderate Demented is less for training and test cases. Moreover, deep learning models require a large number of input samples for training so that better accuracy in the classification process can be obtained. Thus, the data augmentation procedure adopted in the presented research work performs image flipping which flips the images into horizontal, vertical directions. 90 degree and 270-degree anti clockwise rotation are done for the input images. Finally, the brightness factor is



**Fig. 6.** Data augmentation (a) Original Image (b) Horizontal flipping (c) Vertical flipping (d) 90 degree anticlockwise (e) 270 degree anticlockwise and (f) Brightened image.

S. no	Class name	Training	Test	Total
1	Moderate demented	312	72	384
2	Very mild demented	10,752	2688	13,440
3	Non-demented	15,360	3840	19,200
4	Mild demented	4302	1074	5376

**Table 3.** Dataset description after augmentation.

S. no	Metrics	Training	Test
1	Precision	96.26	95.65
2	Recall	96.89	96.48
3	Specificity	98.48	98.24
4	F1-Score	96.57	96.06
5	Accuracy	98.53	97.88

**Table 4.** Performance analysis of proposed model.

increased to 0.7 to increase the number of input samples. The samples after data augmentation are presented in Fig. 6 (a)-(f). The number of samples after data augmentation are presented in Table 3 for all the four classes.

The metrics considered for performance evaluation are presented in Table 4 for training and test process. Metrics like precision, recall, specificity, f1-score, and accuracy are considered for both testing and training process.

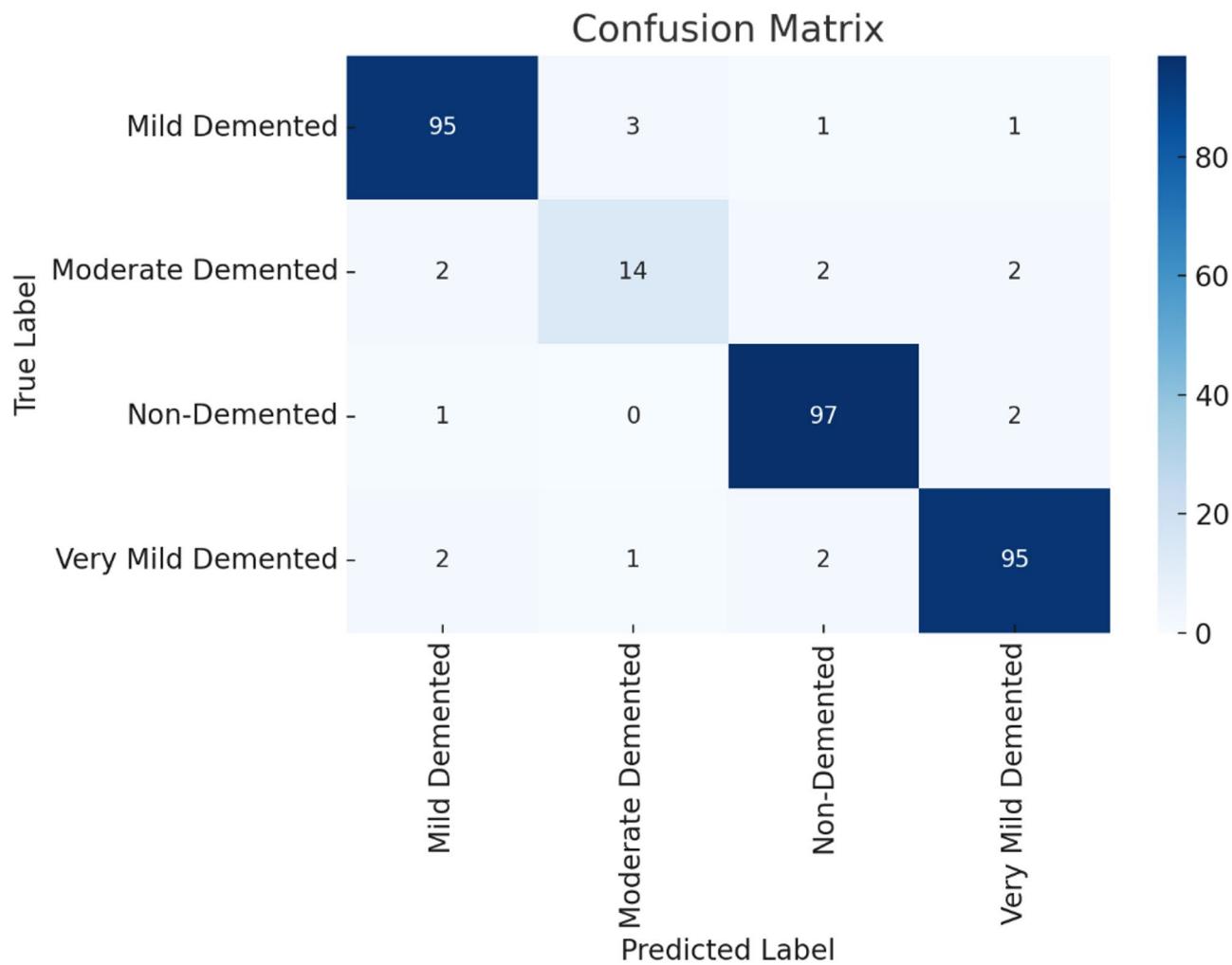
Further to validate the superior performance of the proposed model in Alzheimer detection, existing deep learning and hybrid deep learning models are considered for comparative analysis. Nagarathna.et.al<sup>43</sup> , presented a hybrid deep learning algorithm which combines VGG19 and CNN as a hybrid CNN (HCNN) model for Alzheimer detection. The performance metrics of the HCNN are comparatively analyzed with conventional CNN model. Sharma.et.al<sup>44</sup> , presented a hybrid deep learning model which combines DenseNet121 and DenseNet201 with machine learning algorithms for Alzheimer detection. Srivastava.et.al<sup>45</sup> , presented a comparative analysis in which deep learning models like AlexNet, ResNet-18, VGG-16, Custom CNN, Inception v1, are analysed. The results from the above-mentioned models are considered for comparative analysis.

The confusion matrix analysis of proposed work is depicted in Fig. 7. To further evaluate the model's classification performance across individual classes, a confusion matrix was generated using the test dataset. The matrix, shown in Fig. 7, clearly illustrates how well each class, namely Mild Demented, Moderate Demented, Non-Demented, and Very Mild Demented, was correctly identified, and highlights the few misclassifications made by the model.

### Precision analysis

The comparative analysis of the precision metric is presented in Fig. 8. The maximum precision of 96.26% was obtained in the proposed model training process, and in the testing process, 95.65% was obtained.

From the numerical values given in the Fig. 7, the proposed model's higher performance can be observed. The difference in precision metric between the proposed model and existing CNN and VGG-16 models is 4.25%. When compared to Inceptionv1, the difference is 6.35%. For AlexNet, the difference is 7.15%. For ResNet-18, the difference is 10.25%. For the hybrid DenseNet121 and 201 models, the difference is 3.65% and 2.65%, respectively. The HCNN has a 4.40% difference in precision metrics, and the least performance is exhibited by the conventional CNN model, which is 26.65% less than the proposed model.



**Fig. 7.** Confusion matrix analysis.

### Recall analysis

The recall metric presented in Fig. 9 shows a recall of 96.89% in the proposed model training process, and 96.48% in the testing process.

The proposed model's higher performance can be observed from the comparative analysis. The difference in recall metric between the proposed model and custom CNN is 1.68% and for Inceptionv1, the difference is 6.35%. When compared to VGG-16 model the difference is 6.58%. For AlexNet, the difference is 9.98%. For ResNet-18, the difference is 7.88%. For the hybrid DenseNet121 and 201 models, the difference is 7.48% for both models. HCNN has a 5.24% difference in precision metrics, and the least performance is exhibited by the conventional CNN model, which is 31.48% less than the proposed model.

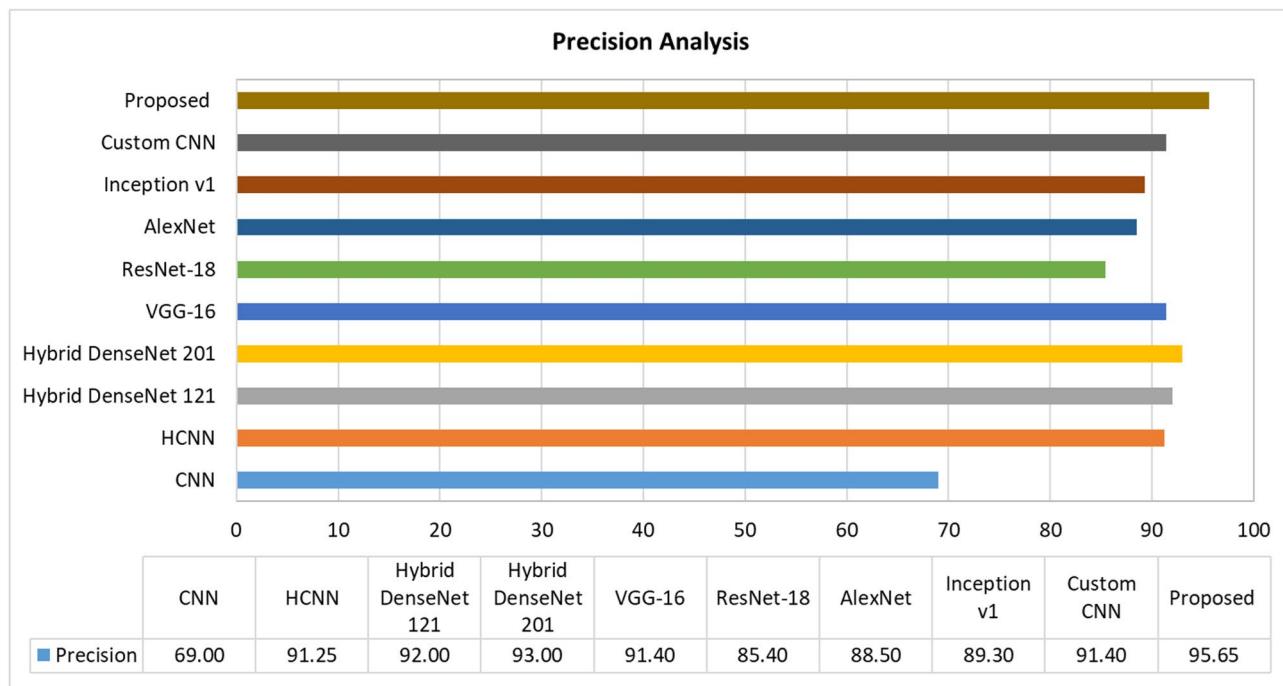
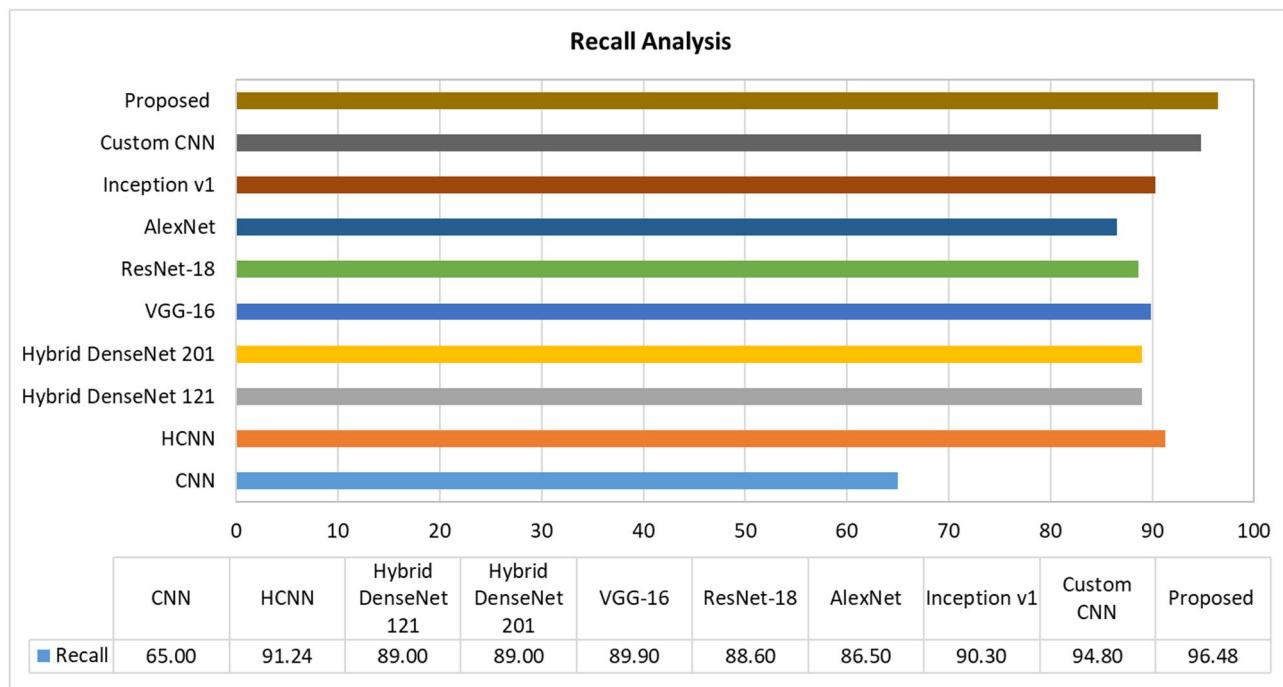
### Specificity analysis

Figure 9 depicts the comparative analysis for specificity metric. The maximum specificity of 98.48% was obtained in the proposed model training process, and in the testing process, 98.24% was obtained. The proposed model's higher performance can be observed from the comparative analysis given in Fig. 10.

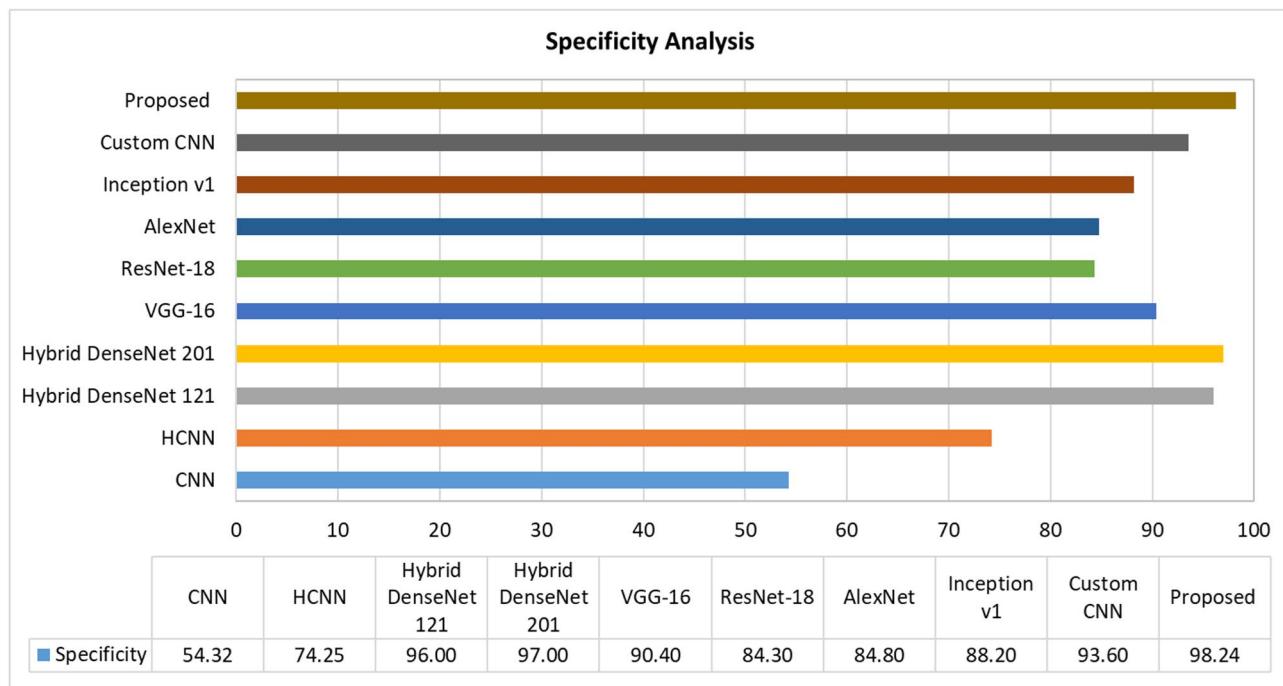
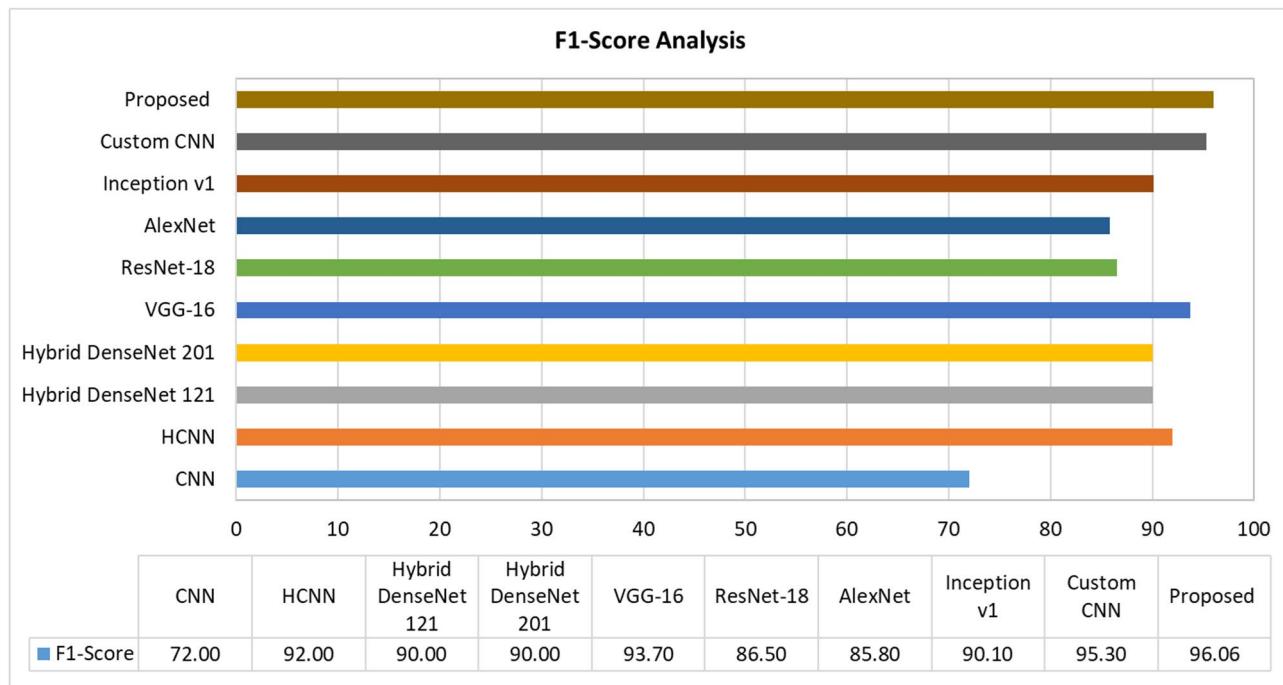
The performance difference in specificity metric between the proposed model and custom CNN is 4.64% and for Inceptionv1, the difference is 10.04%. When compared to VGG-16 model the difference is 7.84%. For AlexNet, the difference is 13.44%. For ResNet-18, the difference is 13.94%. For the hybrid DenseNet121 and 201 models, the difference is 2.24 and 1.24, respectively. HCNN has a 23.99% difference in specificity metrics, and the least performance is exhibited by the conventional CNN model, which is 43.92% less than the proposed model.

### F1-Score analysis

The comparative analysis of the F1-score is presented in Fig. 11. The maximum F1-score of 96.57% was obtained in the proposed model training process, and in the testing process, 96.06% was obtained. From the numerical values given in the Fig. 10, the proposed model's higher performance can be observed. The difference in F1-score metric between the proposed model and custom CNN is 0.76% and for Inceptionv1, the difference is 5.96%.

**Fig. 8.** Precision comparative analysis.**Fig. 9.** Recall comparative analysis.

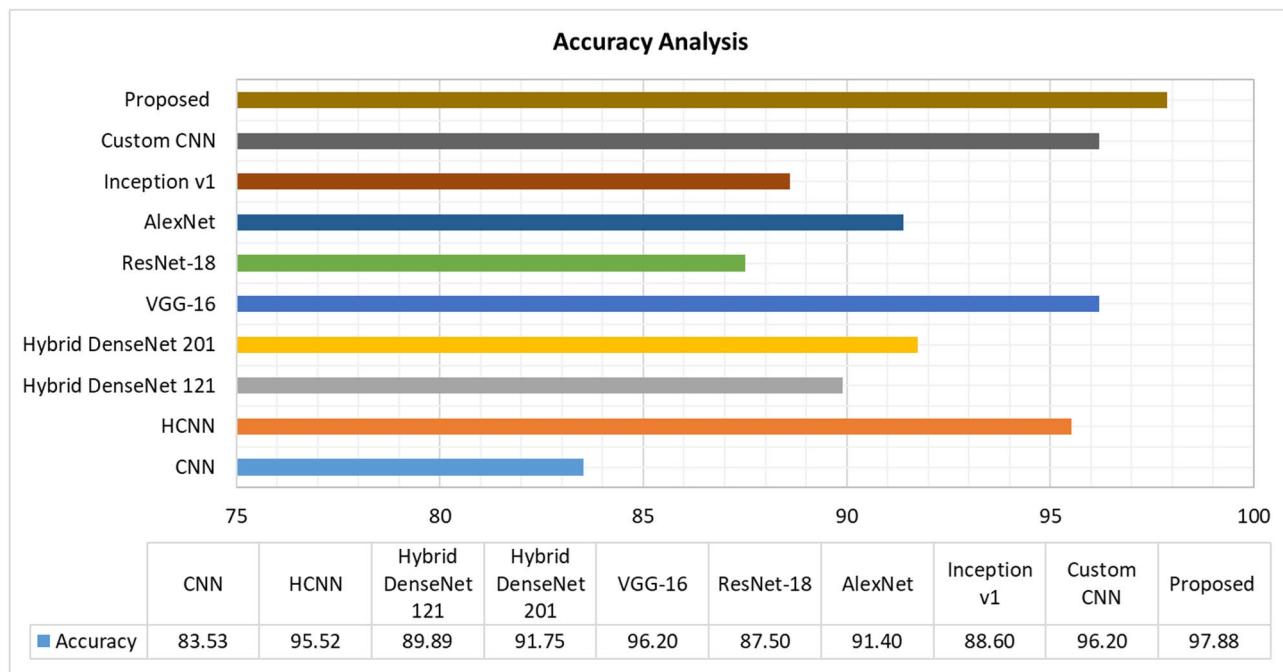
When compared to VGG-16 model the difference is 2.36%. For AlexNet, the difference is 10.26%. For ResNet-18, the difference is 9.56%. For the hybrid DenseNet121 and 201 models, the difference is 6.06 for both models. HCNN has a 4.06% difference in F1-score metrics, and the least performance is exhibited by the conventional CNN model, which is 24.06% less than the proposed model.

**Fig. 10.** Specificity comparative analysis.**Fig. 11.** F1-Score comparative analysis.

### Accuracy analysis

Figure 11 depicts the accuracy metric analysis, and the results show that maximum accuracy of 98.53% in the proposed model training process, and 97.88% in the testing process. The proposed model's higher performance can be observed from the comparative analysis given in Fig. 12.

The performance difference in precision metric between the proposed model and existing CNN and VGG-16 models is 1.68%. When compared to Inceptionv1, the difference is 9.28%. For AlexNet, the difference is 6.48%. For ResNet-18, the difference is 10.38%. For the hybrid DenseNet121 and 201 models, the difference is

**Fig. 12.** Accuracy comparative analysis.

S. no	Authors	Methods	Metrics				
			Precision	Recall	F1-Score	Specificity	Accuracy
1	Nagarathna et.al. <sup>40</sup>	CNN	69.00	65.00	72.00	54.32	83.53
2		HCNN	91.25	91.24	92.00	74.25	95.52
3	Sharma et al. <sup>41</sup>	Hybrid DenseNet121	92.00	89.00	90.00	96.00	89.89
4		Hybrid DenseNet201	93.00	89.00	90.00	97.00	91.75
5	Srivastava et al. <sup>42</sup>	VGG-16	91.40	89.90	93.70	90.40	96.20
6		ResNet-18	85.40	88.60	86.50	84.30	87.50
7		AlexNet	88.50	86.50	85.80	84.80	91.40
8		Inception v1	89.30	90.30	90.10	88.20	88.60
9		Custom CNN	91.40	94.80	95.30	93.60	96.20
10	Proposed		95.65	96.48	96.06	98.24	97.88

**Table 5.** Performance analysis with existing methods.

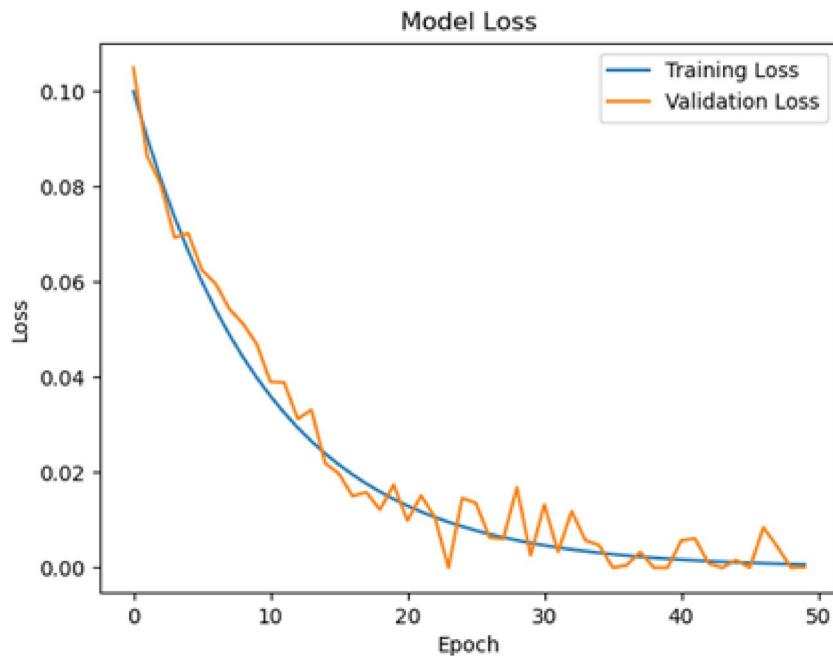
7.99% and 6.13%, respectively. The HCNN has a 2.23% difference in accuracy metrics, and the least performance is exhibited by the conventional CNN model, which is 14.35% less than the proposed model. The overall performance comparative analysis with the existing methods is presented in Table 5. The results clearly present the better detection performance of the proposed model over existing methodologies.

### Ablation study

To assess the individual contributions of different components in the proposed hybrid deep learning model, we conducted an ablation study as shown in Table 6 focusing on two key aspects: the choice of optimization algorithm for hyperparameter tuning and the architecture-wise performance impact of Inception v3, ResNet-50, and their hybrid combination. For optimizer comparison, the Adaptive Rider Optimization (ARO) algorithm was evaluated against conventional methods such as Adam, Grid Search, and Bayesian Optimization. The same base architecture (Inception v3 + ResNet-50) was used across all optimizers to ensure fair comparison. As shown in Table 5, ARO consistently yielded higher performance, with a testing accuracy of 97.88% and F1-score of 96.06%, compared to Adam (94.52%, 92.70%), Grid Search (93.84%, 91.85%), and Bayesian Optimization (95.10%, 93.62%). The superior performance of ARO is attributed to its adaptive exploration-exploitation strategy that effectively escapes local minima and fine-tunes learning rate, dropout rate, and batch size more precisely.

Further, to evaluate the impact of each network component, we trained models with (i) only Inception v3, (ii) only ResNet-50, and (iii) the proposed hybrid architecture (Inception v3 + ResNet-50). When used

Configuration	Accuracy (%)	F1-Score (%)
ARO + Inception v3 + ResNet-50	97.88	96.06
Adam + Inception v3 + ResNet-50	94.52	92.70
Grid Search + Inception v3 + ResNet-50	93.84	91.85
Bayesian Opt. + Inception v3 + ResNet-50	95.10	93.62
ARO + Inception v3 only	94.36	92.84
ARO + ResNet-50 only	93.29	91.47

**Table 6.** Ablation study.**Fig. 13.** Training and Validation Loss performance.

independently, Inception v3 achieved 94.36% accuracy and 92.84% F1-score, while ResNet-50 achieved 93.29% accuracy and 91.47% F1-score. In contrast, the full hybrid model reached 97.88% accuracy and 96.06% F1-score. This improvement highlights the complementarity between Inception v3's multi-scale feature extraction and ResNet-50's deep residual learning for classification. Inception v3 captures detailed textural and spatial features, while ResNet-50 effectively classifies subtle inter-class variations. The sequential integration enhances both representation learning and discrimination capability, thereby offering a more robust framework for Alzheimer's classification.

### Training and validation analysis

Figure 13 shows the training and validation loss curves over 50 epochs for our proposed hybrid deep learning model. The graph provides a visual representation of the model's performance and helps to evaluate its ability to generalize beyond the training data.

The graph illustrates that both the training loss (blue curve) and validation loss (orange curve) consistently decreased over the epochs. Importantly, the two curves converge as the training progresses, indicating that the model is learning effectively and is not overfitting to the training data. If the model were overfitting, we would observe a significant divergence between the training and validation loss curves, with the validation loss increasing while the training loss continues to decrease. In this graph, the validation loss closely follows the training loss, demonstrating that the model maintains its performance on unseen validation data. This behavior indicates that the model generalizes well and effectively captures the underlying patterns in the data without memorizing the training examples.

### Statistical analysis of the proposed work

We have conducted a thorough comparative analysis to validate the performance of the proposed hybrid deep learning model. We have evaluated the model using standard performance metrics such as precision, recall, specificity, F1-score, and accuracy. These metrics were calculated for both the training and testing processes,

and the results were compared against several existing models including CNN, VGG-16, ResNet-18, AlexNet, Inception v1, and hybrid models like DenseNet121 and DenseNet201.

*Precision:* The proposed model achieved a precision of 95.65% in the testing process, demonstrating superior performance compared to other models such as CNN (69.00%), VGG-16 (91.40%), and ResNet-18 (85.40%).

*Recall:* Our model attained a recall of 96.48%, outperforming models like AlexNet (86.50%) and Inception v1 (90.30%).

*Specificity:* The specificity of the proposed model was 98.24%, which is significantly higher compared to conventional models like CNN and HCNN.

*F1-Score:* The F1-score for the proposed model was 96.06%, indicating a balanced performance in terms of precision and recall.

*Accuracy:* The model's accuracy was 97.88%, which is higher than many existing methods, as shown in our comparative analysis.

*Comparative analysis:* We presented a detailed comparative analysis through various figures and tables (Figs. 8, 9, 10, 11 and 12; Table 4) to highlight the superior performance of the proposed model over existing methods. These comparisons clearly demonstrate that our model achieves better detection performance in terms of all the evaluated metrics.

*Experimental validation:* Our experiments were conducted using a benchmark Alzheimer's dataset from the Kaggle repository. We ensured a robust experimental setup with adequate data preprocessing, including noise removal, normalization, and data augmentation, to improve the model's generalization capabilities.

To further validate the robustness of the proposed model's performance, we computed 95% confidence intervals for two key evaluation metrics: accuracy and F1-score. Based on the test set comprising 384 samples, the accuracy of 97.88% yields a 95% confidence interval of [96.44%, 99.32%], while the F1-score of 96.06% corresponds to a 95% confidence interval of [94.11%, 98.01%]. These intervals were calculated using a normal approximation to the binomial distribution, assuming independent Bernoulli trials for classification outcomes. The narrow bounds of the confidence intervals indicate the statistical reliability and consistency of the proposed model across varying samples, reinforcing its superiority over existing techniques.

To establish the statistical significance of performance improvements, a paired t-test was conducted on the accuracy scores obtained over five independent runs of both the proposed model and the best-performing baseline model. The test yielded a t-statistic of 27.57 and a p-value of  $1.03 \times 10^{-5}$ , indicating that the observed difference in accuracies is highly statistically significant ( $p < 0.001$ ). This confirms that the performance improvements achieved by the proposed Inception v3 + ResNet-50 model, optimized using ARO, are not due to random variation and can be considered consistently superior to the baseline model across multiple runs. This statistical validation supports the robustness and generalizability of the proposed approach in Alzheimer's disease classification.

## Conclusion

An optimized hybrid deep learning model for Alzheimer detection is presented in this research work using Inception v3 and ResNet 50 algorithms. The hybrid model incorporates an Adaptive Rider Optimization (ARO) algorithm to fine tune the network parameters so that improved accuracy in Alzheimer detection was attained in the proposed research work. Experimentations using benchmark dataset validates the proposed model performance with various metrics and compared with conventional deep learning and hybrid deep learning algorithms. The proposed model outperformed existing models CNN, HCNN, Hybrid DenseNet121, Hybrid DenseNet201, AlexNet, ResNet-18, VGG-16, Custom CNN, Inception v1 models with precision of 95.65%, recall of 96.48%, specificity of 98.24%, f1-score of 96.06% and accuracy of 97.88%. However, challenges such as high computational demands and data variability were encountered. Several critical areas of application are available as the future scope of this research work. Firstly, integrating multi-modal data such as genetic information, clinical test records, and other imaging modalities like PET and CT scans could provide a more comprehensive understanding of Alzheimer's disease progression and improve the accuracy of early detection. Secondly, exploring the use of more advanced deep learning architectures, such as transformer-based models and attention mechanisms, could further enhance the model's ability to capture complex patterns in the data. Additionally, implementing federated learning approaches could enable the use of distributed datasets while preserving patient privacy, thereby addressing data scarcity and heterogeneity issues. Another important direction is the development of real-time diagnostic tools and mobile applications that utilize the proposed model for practical and accessible Alzheimer's detection in clinical settings.

## Data availability

The data's used to support the findings of this study are included within this article.

Received: 8 May 2025; Accepted: 29 July 2025

Published online: 30 September 2025

## References

1. <https://www.alz.org/alzheimers-dementia/facts-figures>.
2. Pettit, U., Baker, S. & Korhonen, A.. systematic literature review of automatic Alzheimer's disease detection from speech and language. *J. Am. Med. Inform. Assoc.* 27(11), 1784–1797 (2020).
3. Shaw, L. M. et al. Detection of Alzheimer disease pathology in patients using biochemical biomarkers: prospects and challenges for use in clinical practice. *J. Appl. Lab. Med.* 5(1), 183–193 (2020).
4. Sharma, A. et al. Alzheimer's patients detection using support vector machine (SVM) with quantitative analysis. *Neurosci. Inf.*, 1, 3 (2021).

5. Shukla, A., Tiwari, R. & Tiwari, S. Review on Alzheimer disease detection methods: automatic pipelines and machine learning techniques. *Sci* **5**(1), 1–24 (2023).
6. Ebrahimighahnavieh, A., Luo, S. & Chiong, R. Deep learning to detect Alzheimer's disease from neuroimaging: A systematic literature review. *Comput. Methods Programs Biomed.* **187**, 1–49 (2020).
7. Ramachandran, L., Mangaiyarkarasi, S. P., Subramanian, A. & Senthilkumar, S. Shrimp classification for white spot syndrome detection through enhanced gated recurrent unit-based wild geese migration optimization algorithm. *Virus Genes*. <https://doi.org/10.1007/s11262-023-02049-0> (2024).
8. Ramachandran, L., Mohan, V., Senthilkumar, S. & Ganesh, J. Early detection and identification of white spot syndrome in shrimp using an improved deep convolutional neural network. *J. Intell. Fuzzy Syst.* **45**(4), 6429–6440. <https://doi.org/10.3233/JIFS-232687> (2023).
9. Ganesh, D. et al. Implementation of convolutional neural networks for detection of Alzheimer's disease. *BioGecko J. New. Z. Herpetology*. **12**(01), 71–82 (2023).
10. Francis, A. & Pandian, I. A. Early detection of Alzheimer's disease using local binary pattern and convolutional neural network. *Multimedia Tools Appl.* **80**(496), 29585–29600 (2021).
11. Vembarasi, K. et al. White spot syndrome detection in shrimp using neural network model. In *Proceedings of the 18th INDIACOM; INDIACOM-; IEEE Conference ID: 57xxx, 2024 11th International Conference on Computing for Sustainable Global Development, 28th Feb-01st March 2024* <https://doi.org/10.23919/INDIACOM61295.2024.10498722> (Bharati Vidyapeeth's Institute of Computer Applications and Management (BVICAM), 2024).
12. Illakiya, T. & Karthik, R. Automatic detection of Alzheimer's disease using deep learning models and neuroimaging: current trends and future perspectives. *Neuroinformatics* **21**, 339–364 (2023).
13. Ammar, R. B. & Ayed, Y. B. Language-related features for early detection of Alzheimer disease. *Procedia Comput. Sci.* **176**, 763–770 (2020).
14. Shankar, K. et al. Alzheimer detection using group Grey Wolf Optimization based features with convolutional classifier. *Comput. Electr. Eng.* **77**, 230–243 (2019).
15. Varatharajan, R., Manogaran, G., Priyan, M. K. & Sundarasekar, R. Wearable sensor devices for early detection of Alzheimer disease using dynamic time warping algorithm. *Cluster Comput.* **21**, 681–690 (2018).
16. Chakravarthy, A., Panda, B. S. & Nayak, S. K. Review and comparison for Alzheimer's disease detection with machine learning techniques. *Int. Neurourol. J.* **27**(4), 403–409 (2023).
17. Acharya, U. R. et al. Automated detection of Alzheimer's disease using brain MRI images- A study with various feature extraction techniques. *J. Med. Syst.* **43**(2), 1–14 (2019).
18. Kishore, P., Kumari, C. U., Kumar, M. N. & Pavani, T. Detection and analysis of Alzheimer's disease using various machine learning algorithms. *Mater. Today Proc.* **45**(2), 1502–1508, (2021).
19. Sampath, R. & Indumathi, J. Earlier detection of Alzheimer disease using N-fold cross validation approach. *J. Med. Syst.* **42**(217), 1–11 (2018).
20. Feng, W. et al. Automated MRI-based deep learning model for detection of Alzheimer's disease process. *Int. J. Neural Syst.* **30**(06), 1–14 (2020).
21. Helaly, H. A., Badawy, M. & Haikal, A. Y. Deep learning approach for early detection of Alzheimer's disease. *Cogn. Comput.* **14**, 1711–1727 (2022).
22. Venugopalan, J., Tong, L., Hassanzadeh, H. R. & Wang, M. D. Multimodal deep learning models for early detection of Alzheimer's disease stage. *Sci. Rep.* **11**(3254), 1–13 (2021).
23. Jo, T., Nho, K., Risacher, S. L. & Saykin, A. J. Deep learning detection of informative features in tau PET for Alzheimer's disease classification. *BMC Bioinform.* **21**(496), 1–13 (2020).
24. Chui, K. T., Gupta, B. B., Alhalabi, W. & Alzahrani, F. S. An MRI scans-based Alzheimer's disease detection via convolutional neural network and transfer learning. *Diagnostics* **12**(7), 1–14 (2022).
25. Ghazal, T. M. et al. Alzheimer disease detection empowered with transfer learning. *Computers Mater. Continua*. **70**, 3 5005–5019.
26. Lokesh, K. et al. Early Alzheimer's disease detection using deep learning. *EAI Endorsed Trans. Pervasive Health Technol.* **9**, 1–9 (2023).
27. Miltiadous, A., Gionanidis, E., Tzimourta, K. D., Giannakeas, N. & Tzallas, A. T. DICE-Net: A novel convolution-transformer architecture for Alzheimer detection in EEG signals, *IEEE Access*. **11**, 71840–71858 (2023).
28. Pan, D. et al. Early detection of Alzheimer's disease using magnetic resonance imaging: A novel approach combining convolutional neural networks and ensemble learning. *Front. NeuroSci.* **14**, 1–19 (2020).
29. Mujahid, M. et al. An efficient ensemble approach for Alzheimer's disease detection using an adaptive synthetic technique and deep learning. *Diagnostics* **13**(15), 1–20 (2023).
30. Balasubramiam, S., Kadry, S. & Kumar, K. S. Osprey Gannet optimization enabled CNN based transfer learning for optic disc detection and cardiovascular risk prediction using retinal fundus images. *Biomed. Signal Process. Control.* **93**(1), 106177–106188 (2024).
31. Kadry, S., Dhanaraj, R. K. & Manthiramoorthy, C. Res-Unet based blood vessel segmentation and cardiovascular disease prediction using chronological chef-based optimization algorithm based deep residual network from retinal fundus images. *Multimedia Tools Appl.* **1**(1), 1–30 (2024).
32. Choudhury, A., Balasubramiam, S., Kumar, A. P. & Kumar, S. N. P. PSSO: political squirrel search optimizer-driven deep learning for severity level detection and classification of lung cancer. *Int. J. Inform. Technol. Decis. Mak.* **1**(1), 1–34 (2023).
33. Balasubramiam, S., Syed, M. H., More, N. S. & Pole ally, V. Deep learning-based power prediction aware charge scheduling approach in cloud based electric vehicular network. *Eng. Appl. Artif. Intell.* **121**(1), 105869–105879 (2023).
34. Dakshinamoorthy, C. et al. Hybrid Whale and Gray Wolf deep learning optimization algorithm for prediction of Alzheimer's disease. *Mathematics* **11**(5), 1136–1146 (2023).
35. Rajasree, R. S. & Brintha Rajakumari, S. Ensemble-of-classifiers-based approach for early Alzheimer's disease detection. *Multimedia Tools Appl.* **83**(6), 16067–16095 (2024).
36. Sinha Roy, R. & Sen, A. A hybrid deep learning framework to predict Alzheimer's disease progression using generative adversarial networks and deep convolutional neural networks. *Arab. J. Sci. Eng.* **49**(3), 3267–3284 (2024).
37. Mujahid, M. et al. An efficient ensemble approach for Alzheimer's disease detection using an adaptive synthetic technique and deep learning. *Diagnostics* **13**(15), 2489–2498 (2023).
38. Sorour, S. E. et al. Classification of Alzheimer's disease using MRI data based on deep learning techniques. *J. King Saud Univ.-Comput. Inf. Sci.* **36**(2), 101940–101987 (2024).
39. Alnowaiser, K., Saber, A., Hassan, E. & Awad, W. A. An optimized model based on adaptive convolutional neural network and grey Wolf algorithm for breast cancer diagnosis. *PlosOne* <https://doi.org/10.1371/journal.pone.0304868> (2024).
40. Elbedwehy, S. et al. Integrating neural networks with advanced optimization techniques for accurate kidney disease diagnosis. *Sci. Rep.* **14**, 21740. <https://doi.org/10.1038/s41598-024-71410-6> (2024).
41. Saber, A. et al. An optimized ensemble model based on meta-heuristic algorithms for effective detection and classification of breast tumors. *Neural Comput. Appl.* **37**, 4881–4894. <https://doi.org/10.1007/s00521-024-10719-9> (2025).
42. <https://www.kaggle.com/datasets/tourist55/alzheimers-dataset-4-class-of-images>.
43. Nagarathna, C. R. & Kusuma, M. Automatic diagnosis of Alzheimer's disease using hybrid model and CNN. *Int. J. Innovative Res. Sci. Eng. Technol.* **3**(1), 1–4 (2022).

44. Sharma, S. et al. HTLML: hybrid AI based model for detection of Alzheimer's disease. *Diagnostics* **12**(8), 1–16 (2022).
45. Srivastava, S. et al. Comparative analysis of deep learning image detection algorithms. *J. Big Data.* **8**, 66. <https://doi.org/10.1186/s40537-021-00434-w> (2021).

### Author contributions

Sundar Raj and S. Senthilkumar carried out this research work and draft the first copy. S. Sivamani supported in literature review and final drafting of the manuscript. C. Gunasundari contributed in revision of the manuscript.

### Funding

There is no source of funding.

### Declarations

#### Ethics approval and consent to participate

This study does not involve any direct interaction with human subjects. The MRI image data used was obtained from a publicly available dataset hosted on the Kaggle platform, which contains fully anonymized medical images released for research purposes. Therefore, ethical approval and informed consent were not required.

#### Competing interests

The authors declare no competing interests.

#### Additional information

**Correspondence** and requests for materials should be addressed to A.S.R.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025