IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Mathematical Modelling of Engineering Problems
Vol. 12, No. 9, September, 2025, pp. 3216-3226
Journal homepage: http://iieta.org/journals/mmep

Enhancing AD Classification with Deep Learning: A Study of Transfer Learning and Snake Optimization on MRI Data

Sarah Riyadh Adnan[1], Aqeel Majeed Breesam[2*], Alaa Hussein Abdulaal[3]

[1] High Institute for Infertility Diagnosis Assisted Reproductive Technologies, Al-Nahrain University, Baghdad 10072, Iraq
[2] Institute of Medical Technology/Baghdad, Middle Technical University, Baghdad 10001, Iraq
[3] Department of Electrical Engineering, College of Engineering, Al-Iraqia University, Baghdad 10054, Iraq

Corresponding Author Email: aqeelmajeed@mtu.edu.iq

https://doi.org/10.18280/mmep.120925

Received: 2 February 2025
Revised: 30 April 2025
Accepted: 8 May 2025
Available online: 30 September 2025

Keywords:
MRI-based AD diagnosis, deep learning, metaheuristic optimization, Snake Optimization Algorithm (SOA), transfer learning, MobileNet

ABSTRACT

Alzheimer's disease functions as the leading dementia disorder and creates a major health problem for millions of patients worldwide. Effective preventive intervention requires AD detection during the Mild Cognitive Impairment stage. The study employs VGG16 together with MobileNet architectures to classify Alzheimer's disease through MRI image analysis. The adoption of transfer learning for pre-trained models allowed us to modify MobileNet using the Snake Optimization Algorithm (SOA) for superior performance outcomes. Accurate AD classification through deep learning technology depends on transfer learning combined with hyperparameter optimization mechanisms, which process image datasets as input. When the MobileNet model operated with the SOA optimizer, the system reached a 97.71% accuracy, outperforming the results obtained from the VGG16 model. Our optimized model achieved superior performance across all other metrics with both high precision and recall rates in addition to reaching a 97.71% accuracy in AD stage diagnosis. The MobileNet+SOA algorithm exhibits higher precision and accuracy rates than its counterparts for MRI image diagnosis, as shown by comparative performance evaluation. The combination of deep learning methods, transfer learning and hyperparameter optimization produces an efficient solution for MRI image-based Alzheimer's disease classification. The MobileNet+SOA model is potentially a successful system of AI-based AD diagnosis, which will be applicable in detecting AD earlier to control this harmful neurological disorder.

1. INTRODUCTION

The most prevalent type of dementia is the Alzheimer disease (AD) affecting millions of people across the world. No cure has been discovered despite a lot of research done to stop or reverse its course [1]. Neuroimaging-based classification of early-stage Alzheimer diseases is rather problematic because of the existence of subtle changes in the brain, a large data dimension, and variable manifestations of the disease. The traditional machine learning approaches, which use manually derived parameters, tend to have low discrimination between early-stage Alzheimer disease and moderate cognitive impairment (MCI) and normal aging.

The number of Americans with Alzheimer disease is estimated in 2020 at about 6 million, and the number is expected to increase to 14 million by the middle of the century [2]. Diagnosis of AD at an early age is important because it is possible to treat the illness prior to the development of clinical symptoms. MCI is a transitional disorder between typical aging and Alzheimer disease and is a disease that is progressive affecting 20 percent of the elderly over the age of 65 and 35 percent of them escalate to develop Alzheimer disease three to five years later [2, 3]. The best method of diagnosis of the Alzheimer is the autopsy [4].

AD has enormous financial effects with the healthcare expenses expected to reach USD 1.1 trillion by 2050 [2] and USD 305 billion in 2020. The following alarming statistics show the necessity of better diagnostic tools. Therefore, the objective of this research is to prevent the development of the Alzheimer disease through developing advanced computer tools in detecting the disease at an early stage.

Decision making algorithms, which have the capacity to differentiate between AD, MCI, and normal cognitive functioning are needed to detect the disease at its early phases. Conventional classification methods have acute disadvantages (poor overfitting of small MRI data, as well as extensive human feature engineering) [5]. Additionally, there is a challenge in differentiating between fine-grained transition boundaries of AD and MCI because of standard diagnostic procedures.

The diagnosis of Alzheimer disease in the case of the combined use of contemporary neuroimaging and machine learning (ML) is promising. The effectiveness of machine learning methods has been proven by such contests as the AD large data challenge [6] and MCI prediction challenges [7]. Nevertheless, the strength and reliability of such procedures can still be enhanced [8, 9].

Deep learning (DL) is a form of machine learning (ML), which has increased in popularity in medical imaging workplaces as it can extract complex data without human intervention [10]. DL methods are excellent at combining data across various locations in the brain, representations of learning that cannot be seen on an MRI screen, and finding concealed patterns in MRI scans. Compared to other imaging methods, convolutional neural networks (CNNs) have been shown to be more useful in tasks related to the classification of Alzheimer disease tasks than conventional methods.

Hyperparameter tuning has a strong impact on the performance of DL models. Snake Optimization Algorithm (SOA) as an algorithm that emulates the snake motions in nature offers a feasible solution to optimizing hyperparameters automatically to enhance the accuracy of a classification task without necessarily involving excessive amounts of human labor.

Transfer learning (TL), particularly for small datasets, has emerged as a powerful strategy for improving deep learning models [11, 12]. TL improves generalization and accelerates convergence by leveraging knowledge from related fields. TL-based approaches are particularly effective at distinguishing between progressive MCI (pMCI) and stable MCI (sMCI) [13, 14].

This research uses deep learning architectures VGG16 and MobileNet and transfer learning and optimizes hyperparameters by applying the SOA to specifically address AD classification problems. Our method will eliminate current technique restrictions by leveraging deep learning model feature extraction abilities to optimize their performance for AD classification from MRI images. We integrate advanced computational techniques to build a better diagnostic method for detecting AD in early stages and its subsequent management.

2. RELATED WORK

Deep learning and machine learning-based approaches for Alzheimer's disease diagnosis have lately received substantial attention in computer vision and medical imaging research. To do this, machine learning algorithms that use image or voxel intensity, tissue density, and form as feature input test classifiers can discriminate between AD patients with MCI and cognitively normal (CN) people.

A new deep learning technique for identifying AD compared to a healthy control was presented by Sarraf et al. [15]. The study’s sample consisted of 15 healthy people serving as a control group and 28 AD patients who were gathered from the ongoing multicenter AD Neuroimaging Initiative. Skull stripping, motion correction, registration, denoising, and spatial smoothing with a full-width-at-half maximum value of 5 mm were all included in the preprocessing. After preprocessing, the data was inputted into the Le-Net model with the results of 96.85 percent accuracy. Another study by Mathew et al. [16] employed 158 MRI images (71 NC and 87 patients) of our Alzheimer disease dementia (ADD) to present the early diagnosis of the AD. Normalization, resizing, deforming, and flipping were incorporated in the preprocessing stage to enhance better learning. The feature extraction methods (Principal Component Analysis (PCA) and Discrete Wavelet Transform) and the classification method (Support Vector Machine (SVM)) were employed when analyzing the data, reaching the accuracy of 84% when comparing AD to CNs and 91 when comparing MCI to CFs.

At the same time, Hosseini-Asl et al. [17] suggested a deep three-dimensional convolutional neural network (3D-CNN) to diagnose the Alzheimer disease (AD). They tested their model on MRI data of 70 AD, MCI, and NC cases accessed through the ADNI data. The 3D-CNN model identified the local features on the 3D input images, thus, allowing the identification of features to be learnt effectively in the classification process.

Convolutional autoencoder (CAE), a CAD-Dementia dataset of T1-weighted MRI scans of AD, CN, and NC individuals was used to train the model. Skull peeling and spatial normalization comprised preprocessing. Features from the CAD-Dementia dataset were used as biomarkers in the fine-tuning to identify AD in the ADNI dataset. A ten-fold cross-validation produced a classification accuracy of 97.6% when comparing AD with NC.

Ju et al. [18] created a deep neural network for an AD diagnostic task using MRI and textual data (age, gender, and genetic). Using MRI scans of 91 patients with mild cognitive impairment (MCI) and 79 normal controls, together with the matching genetic data from the ADNI-2 dataset, they assessed our proposed technique. They also looked into the relationships between ApoE genotype, age, sex, and MCI. Data Processing and Analysis for Brain Imaging (DPABI) was used for data preparation [19, 20].

In order to do this, they fed correlation coefficient data and Rf-MRI time-series data into LDR, LR, and SVM models (authors their findings indicated that incorporating correlation coefficient data increased test accuracy. The accuracy, sensitivity, and specificity of the LDR model were 65%, 66%, and 67.72%, respectively. Accuracy in the LR model was 71.38%, with a sensitivity of about 77% and specificity of about 62%. Its specificity is 64%, sensitivity of the model is 79%, and accuracy is 78.91%. With an accuracy of 86.47%, sensitivity of 92%, and specificity of 81%, as determined by correlation coefficient data, the autoencoder network demonstrated superior performance.

Deep learning models were used by Farooq et al. [21] for the multi-class categorization of AD. They divided the data into four classes using the ADNI dataset. These classes included 33, 22, 449, and 45 MRI images, respectively. While ResNet-18 and ResNet-152 attained accuracies of 98.01% and 98.14%, respectively, GoogLeNet yielded an accuracy of 98.88%.

A straightforward and effective method for identifying AD using brain MRIs and a three-dimensional convolutional neural network architecture (3D ConvNet) was reported by Bäckström et al. [22]. They extracted automatic features after completing preprocessing operations such as cortex reconstruction, edge clipping, picture resizing, and intensity normalization. The study made use of 1190 MRI images and 340 people from the ADNI dataset, which included 196 AD patients (of whom 103 were male and 96 were female) and 141 normal controls (of whom 75 were male and 66 were female). The model obtained an accuracy of 98.78%.

Gautam et al. [23] introduced a one-class classification (OCC) method that needs training data to come from only one class. By adding the lowest variance data to the OCC design, they improved the classifier’s capacity to generalize and decreased intra-class variation. Tests was done on eighteen reference datasets showed that the suggested technique beat previous methods by more than 5% in F1 score. The primary benefit of the one-class classifier is its efficacy in scenarios when there are either extremely few or no data samples available from other classes.

In the study conducted by Liu et al. [24], a framework consisting of two deep learning models was introduced. The first model is a multi-task deep CNN intended for AD classification and hippocampal segmentation. A binary segmentation mask of the hippocampal region is produced by this model. Nevertheless, it was discovered that the characteristics this multi-task model learnt were insufficient for precise AD classification. In order to make up for these shortcomings, 3D patch hippocampal characteristics were derived using the centroid as a guide. In order to train features for AD classification, the second model, a 3D-DenseNet, was used to differentiate between three classes for AD/NC classification, the suggested strategy outperformed the voxel-wise (86.1%) and area of interest (ROI) (84.7%) characteristics, achieving a classification accuracy of 88.9%.

Functional MRI (fMRI) data from the ADNI dataset was used by Kazemi and Houghten [25] to categorize the various phases of AD. They gathered information from 197 participants—107 women and 90 men—during five courses. With an average accuracy of 97.63%.

Tajbakhsh et al. [26] investigated which approach—training a CNN from scratch or using a fine-tuned CNN approach—is more successful for medical image analysis. They experimented with both approaches and found that, in terms of medical picture classification, detection, and seg- mentation, the optimized method on the ImageNet dataset performed better than training from scratch. Large labeled training datasets, which are sometimes hard to come by in the medical in- dustry, along with a great deal of experience, memory use, and processing power are all necessary for training a CNN from scratch. On the other hand, a CNN that had been trained beforehand using the ImageNet dataset yielded encouraging outcomes for a range of uses, such as the interpretation of medical images.

Ebrahimi-Ghahnavieh et al. [27] used transfer learning to identify AD from MRIs in the ADNI dataset. They performed MRI scan trials with 132 participants per group (AD; NC). They combined recurrent neural networks (RNN) with CNNs. Moreover, identifying improved sequence associations of input photos was the primary goal. After feeding the characteristics into one of our CNNs, we trained an RNN on top of it to increase accuracy.

Using MRI data, Wang et al. [28] presented a 3D CNN-based model using DenseNet. With better information and gradient propagation, these dense connections in the 3D-CNN minimized overfitting and made training easier by bridging the gap between feature extractions caused by the intrinsic lack of data. The authors combined base classifiers using a fusion approach to create an ensemble-based model with a 97.19% accuracy. Table 1 shows a comparison of related works to AD diagnosis.

Table 1. Comparison of related work on AD diagnosis

Ref.	Method	Dataset	Accuracy (%)	Advantages and Disadvantages
[15]	Le-Net	ADNI	96.85	High accuracy but limited to small sample sizes.
[16]	SVM and PCA	ADD	84	Effective for early diagnosis; Lower accuracy compared to deep learning models.
[17]	3D-CNN	ADNI	Not specified	Extracts local features effectively; accuracy not specified.
[18]	Deep Neural Network	ADNI-2	78.91	Incorporates genetic data, but moderate accuracy.
[19, 20]	Autoencoder	CAD-dementia and ADNI	97.6	High accuracy with cross-dataset validation; computationally intensive.
[21]	GoogLeNet, ResNet	ADNI	Up to 98.88	Very high accuracy; requires substantial computational resources.
[22]	3D ConvNet	ADNI	98.78	High accuracy; preprocessing may introduce data loss.
[23]	One-class Classifier	Multiple datasets	Not specified	Good for limited data scenarios; may not generalize well across diverse datasets.
[24]	Multi-task CNN and 3D-DenseNet	ADNI	88.9	Good for AD/NC classification; initial features may be insufficient without further tuning.
[25]	fMRI analysis	ADNI	97.63	High accuracy; fMRI data may not be widely available.
[26]	CNN (Fine-tuning)	ImageNet and medical images	Varies	Lower resource requirement than training from scratch; dependent on pre-trained model relevance.
[27]	RNN and CNN	ADNI	Not specified	Aims to improve sequence learning; complexity may hinder practical application.
[28]	3D CNN with DenseNet	ADNI	97.19	Reduces overfitting with dense connections; complex model structure.

3. PROPOSED METHODOLOGY

Specific researchers have created a thorough deep learning system, which detects multiple stages of AD through magnetic resonance imaging neuroimaging data. Data acquisition and exploratory data analysis form the first steps of this methodology since they provide vital perspectives about both the input data's distribution and its quality. 3,714 T1-weighted MRI scans exist in the dataset, which are categorized as NonDemented, MildDemented and VeryDemented. All images underwent processing that included scaling them to 224 pixels by 224 pixels as well as RGB conversion for activation with pre-trained convolutional neural networks. During the EDA process experts examined the data sets while running statistical tests to find imbalanced classes alongside unique features. Normalization of pixel values followed by the application of horizontal flips and rotations occurred after data preprocessing. The categorical labels underwent one-hot encoding during this process. An 80/20 ratio was used for stratified partitioning, which let the model evaluate its performance in a standardized way.

The central mechanism adopts transfer learning with

VGG16 and MobileNet, which were pre-trained on ImageNet images. Both models functioned as embedded feature extractions that received their initial classification layers replaced by newly created dense layers for processing AD stage categories. The trained networks had global average pooling layers which were followed by four serial fully connected layers consisting of nodes with decreasing numbers (1024, 512, 256, 128) with ReLU activation. A Softmax output layer with three neurons was added as the last component of the model structure for multi-class prediction. During training the researchers kept the first network layers frozen while focusing on developing the additional layers with AD dataset information. Models benefited from the integration of generalizable features acquired from broad-scale data, which they applied to AD-specific MRI scan characteristics. The training lasted for 10 epochs using 32 batch instances for prediction while the Adam optimizer and categorical cross-entropy defined the loss parameter.

The model performance was strengthened through using the SOA to optimize learning rate and dropout rate, together with dense layer size. SOA functions as a nature-inspired algorithm that builds its operation off snake movement patterns, which adapt and use sinusoidal behavior in multidimensional search spaces. The evaluation of candidate hyperparameter sets through validation accuracy takes place in the SOA. The algorithm repeatedly adjusts the velocity and positioning of every snake following the global best solution influence through a sinusoidal exploration mechanism. The MobileNet model achieved better performance after retraining it with the most effective hyperparameters discovered through the PSO algorithm. Figure 1 illustrates the proposed work.

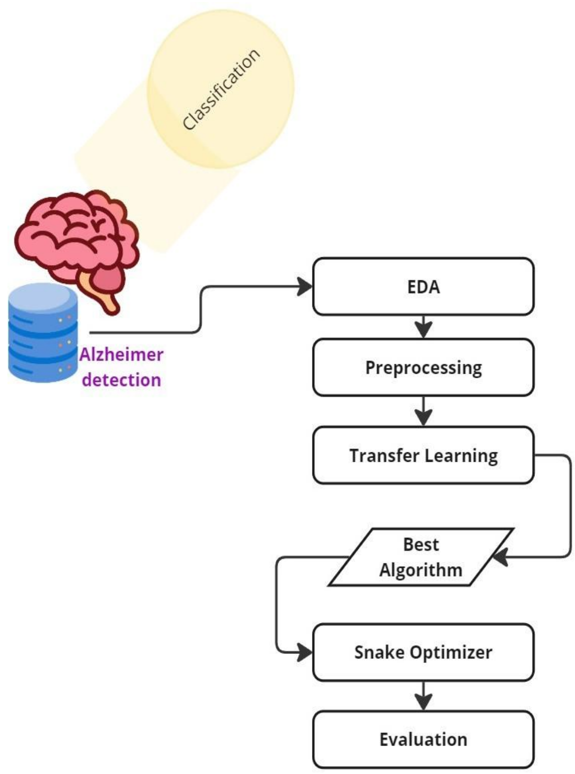

Figure 1. Proposed scheme

The evaluation metrics involved accuracy, precision, recall and F1-score computation where the results were macro-averaged across all three classes to achieve balanced measurement of performance. The enhanced deep learning framework proves its ability to identify and classify early-stage AD right after optimization.

A Python-based approach adopted TensorFlow together with Keras frameworks for developing the model. The dataset was divided into an 80-20 split of training and testing sections and it contained an extra validation group obtained from the training data. The research experiments operated from a GPU-based system platform. The assessment utilized accuracy together with precision and recall and F1-score metrics that performed an average calculation across all classes to maintain balance during evaluation of data sets with unbalanced classes.

3.1 Dataset overview

Our research used an open-access MRI neuroimaging database, which was developed for classifying AD. The database features 3714 T1-weighted magnetic resonance imaging (MRI) scans that received classification labels based on three clinical diagnostic categories, which describe cognitive decline development stages from NonDemented, MildDemented and VeryDemented. The classified dataset utilizes definitive AD diagnostic stages, so it produces a meaningful multi-class classification system similar to the medical diagnostic procedures faced by clinicians.

The imaging data originated from established repositories for medical images before the images underwent a preparatory step, which included both skull-stripping operations and intensity normalization tasks. All images received pre-processing treatment by being resized to 224 × 224 pixels and being converted to RGB color mode even though they originated as grayscale scans. Before inputting the images to VGG16 and MobileNet networks we performed this conversion because both networks need three-channel images as their source data.

The structural parameters of this dataset show mild bias since it contains 1,216 NonDemented scans while MildDemented scans reach 1,792 images and VeryDemented scans total 706 images. The research data shows good clinical accuracy because healthcare professionals routinely examine more patients with MildDemented conditions. This visual representation in Figure 2 shows MRI cutting planes from each class to represent their structural and intensity differences. The anatomical differences between samples in cognitive processing centers become noticeable in these examples which supports accurate model functioning during training and inference. The varied content of this database enables deep learning models to become effective while they demonstrate multispectral capabilities for AD detection at an early stage.

3.2 EDA and preprocessing

A proper Exploratory Data Analysis (EDA) was performed in advance to uncover the structural features alongside visual elements and distribution imbalances throughout the dataset. The underlying database consists of 3,714 T1-weighted brain MRI images containing clinical labels of NonDemented, MildDemented and VeryDemented cognitive stage classifications. The three stages of AD organize into separate categories which the labels represent. The main goal of Exploratory Data Analysis included two objectives: first displaying representative images from each class category as shown in Figure 2 and second applying analysis techniques to examine statistical properties which would guide further preprocessing steps and model development.

EDA began with investigating the class label distribution which showed a moderate imbalance with 1,216 NonDemented images and 1,792 MildDemented images and 706 VeryDemented images. The model would have mostly detected the images as the Mild Demented which is the largest segment of the data, and this would have complicated the early and advanced dementia detection without equalizing the data distribution. To reduce the imbalance in classes, the data augmentation was used. The outcomes of EDA indicated that there was brightness and contrast difference across classes depending on the pixel intensity histograms, which highlighted the necessity of intensity normalization, to ensure the same contrast and enhance neural network training performance.

The qualitative anatomical analysis and the histogram evaluation provided in EDA have demonstrated apparent structural dissimilarities in medial temporal lobe and the ventricular areas associated with AD progression. Preprocessing was done to guarantee spatial and textual consistency by resizing images to standardized 224x224 RGB sizes rather than highly compressed ones. MRI images were rearranged into three RGB in order to match the input of an already trained convolutional network like VGG16 or MobileNet.

The results of the research informed a thorough method on the construction of the preprocessing system of deep learning-based classification. The core steps included:

- Resizing all images to 224 × 224 × 3 dimensions.
- The data range normalization operates on pixels between values 0 to 1 for better numerical accuracy.
- Using one-hot encoding labels become suitable for multi-class classification through the application of Softmax activation. To convert these categorical labels into numerical form, we apply the binary vector method known as one-hot encoding. Given a set of C distinct classes, we use a coding system that transforms each label y into a numerical vector as follows:

$$y_i = \begin{cases} 1, \text{if class } i \text{ corresponds to the given label} \\ 0, \quad \text{otherwise} \end{cases}$$

- For instance, considering three classes are NonDemented (ND), MildDemented (MD), and VeryDemented (VD) a scan labeled as MildDemented (MD) would be encoded as: [0,1,0].

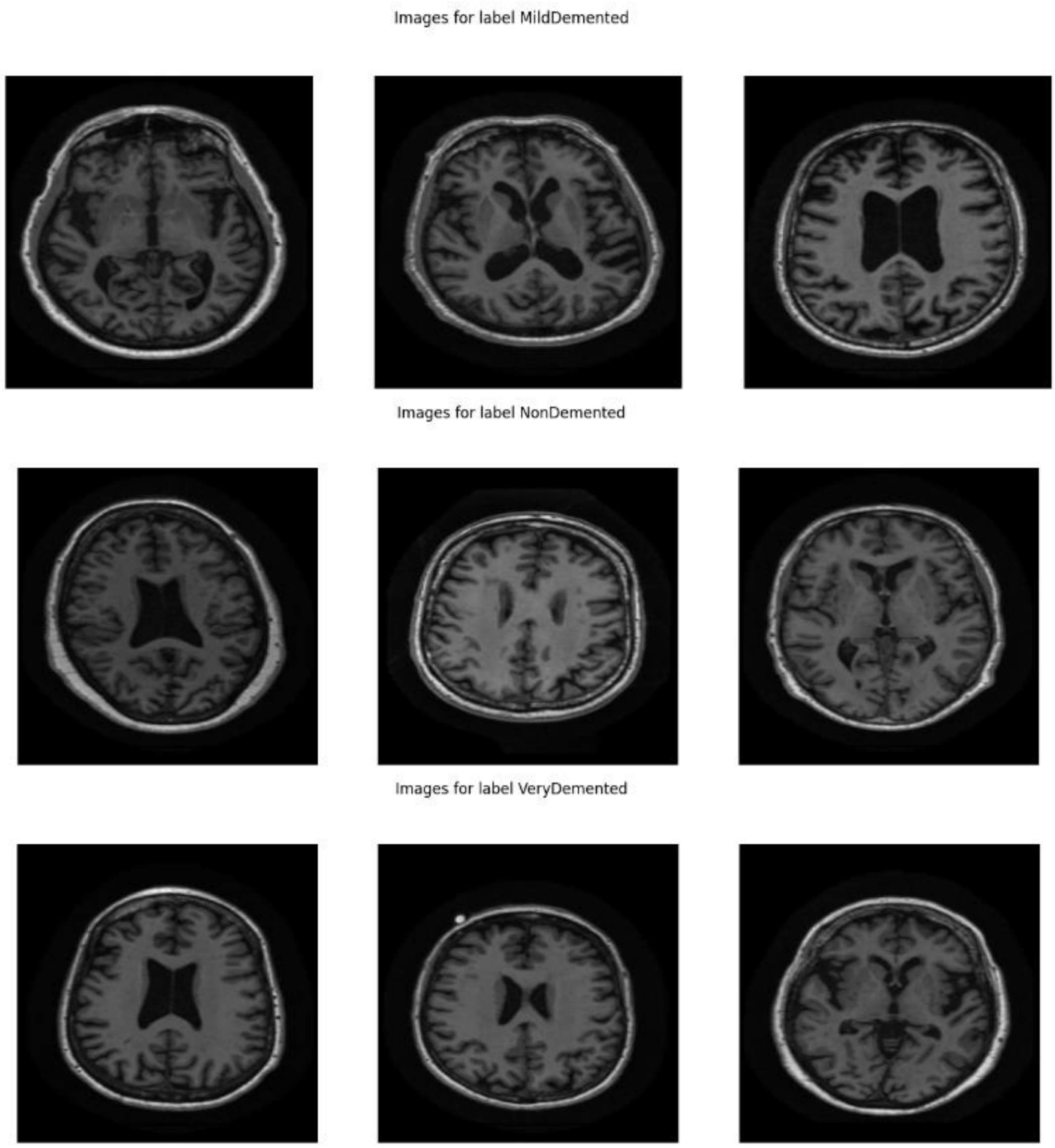

Figure 2. Sample MRI images from different classes

- The dataset was divided through stratified splitting into 80% training data and 20% test data subsets for maintaining consistent class distribution throughout different folds.

The model received augmented data through random horizontal flips and small-angle rotations extending from ±10° plus zoom adjustments and brightness transformations that strengthened training capacity while addressing class unbalance issues. The preprocessing techniques obtained their direction from EDA outcomes and established critical components for creating unbiased and sturdy models.

Through the EDA process researchers gained essential data knowledge while developing essential preprocessing methods for their operation. The standardized integrated strategy allowed our input data to prepare effectively for training deep learning models dedicated to AD classification.

3.3 Transfer learning

In this section, we discuss the major area of AD classification using transfer learning. Pre-training a model simply means that we will train our own custom dataset with some pre-opened models to perform really well on tasks of another new type, from all the past collections trained already. This substantially reduces the effort of training models in similar tasks and improves overall accuracy using previously learned features from a related domain.

We used two different pre-trained models: VGG16 and MobileNet on the classification of AD. They were pre-trained on large image datasets and are popular for different tasks of generalized image recognition.

VGG16 is one of those deep Convolutional Neural Networks that has been used a lot in image classification challenge. It contains several convolutional layers, which act as filters to learn important features from the input images. In our implementation, given the pre-trained VGG16 without top layers for ImageNet classification (implemented by Keras). We did not replace the feature extractor with another CNN architecture but rather used custom fully connected layers to re-purpose it for AD classification. These layers allow the model to learn features particular to this problem making it more capable of distinguishing between different levels of dementia.

In the experiments, another deep learning model called MobileNet was also used because of its lightweight architecture which makes it more suitable to run in resource-constrained environments. MobileNets uses depthwise separable convolutions to reduce the number of parameters the net is modified in such a way that reduces a huge number of parameter and retains same accuracy. Similar to the VGG16, we replaced the top layers with our custom intermediate layers on MobileNet.

In both configurations, we retained the initial layers frozen during training rest of the model using Generative Adversarial Networks to force and constrain PCA transformation from base inputs. These custom layers were trained on the AD dataset to recognize certain patterns differentiating between stages of dementia.

In the post training phase, performance evaluation on accuracy, precision or recall and similar metrics for both the models are performed. We compared the outcomes of these models to establish an optimal method for AD classification by 6.

Transfer learning has been applied in this context and is shown to be effective using pre-trained models, which can achieve state of the art results with minimal training data. This renders it a significant medical image analysis tool. We will provide the findings of each model and discuss them in reference to the interpretation of their meaning when it comes to detection of AD.

VGG16 and MobileNet are selected as they complement each other. The de-facto implementation of VGG16 and its deep architecture, along with its high efficiency in medical imaging, provide an excellent feature extractor, whereas the lightweight design of MobileNet makes it the best architecture to be used on real-time applications and limited performance. Additionally, to validate the effectiveness of our optimizer choice, we conducted comparative experiments where the SOA was evaluated against traditional optimizers such as Adam and SGD. Table 2 presents the model's parameters and values.

Table 2. Model parameters and values

Parameter	VGG16	MobileNet
Pre-trained Weights	ImageNet	ImageNet
Input Size	224 × 224 × 3	224 × 224 × 3
Global Average Pooling	Yes	Yes
Dense Layer 1 Units	1024	1024
Dense Layer 2 Units	512	512
Dense Layer 3 Units	128	256
Dense Layer 4 Units	64	128
Activation Function	ReLU	ReLU
Output Layer Activation	Softmax	Softmax
Optimizer	Adam	Adam
Loss Function	Categorical Crossentropy	Categorical Crossentropy
Batch Size	32	32
Epochs	10	10
Trainable Layers	Custom Dense Layers	Custom Dense Layers

3.4 Snake optimizer

In this study we applied SOA as a hyperparameter optimization strategy for improving performance in the MobileNet architecture for multi-class AD classification. SOA serves as a new biological metaheuristic that uses snake network behavior to discover solutions within complex search spaces which adapt their bodies while remaining aware of environmental conditions. This design suits deep learning model optimization because it helps experts find perfect generalization performances through several dependent hyperparameter adjustments.

All snakes in the population serve as potential solutions because each contains one distinct hyperparameter configuration for the MobileNet model. The algorithm launches its operation by randomly placing snakes across the hyperparameter space where every position represents individual sets of hyperparameters values. Each snake element in the population receives its unique initial velocity direction, which allows it to shift through the search territory. The MobileNet model receives its present set of hyperparameters from each snake in order to conduct training operations during each sequence. The model uses validation accuracy to evaluate the solutions, which have been assessed for fitness.

Snakes who reach the best validation accuracy when tested become the global best solution after which all other snakes adjust their movements based on this position. The software implements position and velocity updating procedures that follow these rules:

$$\text{velocity}_i = \text{velocity}_i + (\text{best\_position} - \text{position}_i) \times \text{learning\_factor} \quad (1)$$

$$\text{position}_i = \text{position}_i + \text{velocity}_i + \sin(\text{iteration}) \times \text{sinusoidal\_factor} \quad (2)$$

Here, velocity_i represents the velocity of snake i, position_i denotes the current position of snake i in the hyperparameter space, and best_position is the position of the best-performing snake.

The learning factor is the weighting of how much a snake's speed gets adapted by moving towards to top snakes position, and the sinusoidal factor puts some periodic behavior into how we update our position.

Finally, an optimal set of hyperparameters is obtained as the snake with the highest fitness after a predefined number optimization cycles. These hyperparameters are then put to use for fine tuning of the deep learning model which boosts its performance. The best model is then subjected to further physical and mathematical testing on an independent test set of data to provide functionality in terms of accuracy, precision also recalls.

The MobileNet model has the following hyperparameters that were optimized using the SOA in our implementation:

- **Learning Rate:** This is important in regulating how fast the backpropagation converges.
- **Batch Size:** It affects the estimation of gradient and stability of training.
- **Dropout Rate:** This was introduced to regularize the network to reduce overfitting by randomly disabling a portion of neurons at any time.
- **The Dense Layers:** This changes the depth of the network and affects the network in its knowledge of abstract representations.
- **Number of Units in Dense Layers:** Determines the learning capacity of each layer by controlling the number of neurons.
- **L2 Regularization Parameter:** Helps reduce model complexity and overfitting by penalizing large weights.

The results obtained with the optimized model show better classification result which proved its prowess regarding to performance SOA.

To sum up, the optimal algorithm for hyperparameter optimization of deep learning is reliable to get higher accuracy with better generalization. This use case of the algorithm for Alzheimer classification highlights its ability to improve complex models having many hyperparameters.

Algorithm 1. SOA

1: **Input:** Population size *n*, number of iterations *T*, learning factor α, sinusoidal factor β
2: **Output:** Best hyperparameters best_position
3: Initialize population of *n* snakes, each with random positions and velocities
4: Evaluate fitness of each snake based on validation accuracy of the model
5: Identify the best snake best_snake with highest fitness
6: **for** iteration = 1 to *T* **do**
7: **for** each snake *i* in the population **do**
8: Update velocity: $\text{velocity}_i = \text{velocity}_i + \alpha \times (\text{best\_position} - \text{position}_i)$
9: Update position: $\text{position}_i = \text{position}_i + \text{velocity}_i + \beta \times \sin(\text{iteration})$
10: Evaluate new fitness of snake *i*
11: **if** new fitness of snake *i* is better than best fitness **then**
12: Update best_snake and best_position
13: **end if**
14: **end for**
15: **end for**
16: **Return:** best_position as the optimal hyperparameters

3.5 Evaluation metrics

One of key ingredients in evaluation (I also talked a bit on this), is metrics, as importance comes attached with its critical role especially for medical imaging domain where everything revolves around life and death. Metrics: Common metrics used in the diagnosis of AD through machine learning approaches are Accuracy, Precision Recall and F1-Score. These metrics are used to generate understanding of the model performance at different angles such as overall correctness, ability to detect positive cases and trade-off between precision recall [29].

$$Accuracy = \frac{TN+TP}{TP+TN+FP+FN} \quad (3)$$

$$Precision = \frac{TP}{FP+TP} \quad (4)$$

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

$$F1 - Score = 2 \times \frac{PRE \times REC}{PRE+REC} \quad (6)$$

- **Accuracy** score represents how many of the test cases were classified correctly on all test cases taken together.
- **Precision** is important in medical diagnostics as it determines the model's ability to accurately anticipate positive labels, preventing false positives.
- **Recall** refers to a model's ability to retrieve all true positives, ensuring that illness cases are also included.
- **F1-Score** is a weighted harmonic mean of precision and recall, offering a composite measure that prioritizes imbalanced classes with big differences across datasets.

4. EXPERIMENTS RESULTS

The study assessed deep learning models VGG16 and MobileNet when used for multiple AD class identification through MRI image analysis. Moreover, the results distinguish between the performance strength of VGG16 and MobileNet models with and without utilization of SOA. Medical imaging results require evaluation through accuracy, precision, recall and F1-score measurements because wrong positives and wrong negatives produce critical outcomes in this field.

The VGG16 model delivered 91.39% accuracy in its operations as shown in Table 3. According to the detailed classification report the precision score for Mild Demented cases reached 0.95 while the recall metric reached 0.91 and F1-score existed at 0.93. The NonDemented category obtained values measuring 0.95 for precision and 0.91 for F1-score and 0.88 for recall. The VeryDemented class exhibited lower model efficiency reflected through 0.82 precision and 0.96

recall and an F1-score of 0.88. VGG16 demonstrates excellent performance detecting VeryDemented cases yet generates numerous wrong positive diagnoses shown by its poor precision value. The model demonstrates weak performance stability across different groups of subjects.

MobileNet demonstrated superior performance than VGG16 according to all measurement criteria where it achieved a total accuracy score of 96.50% as shown in Table 3. The MildDemented category achieved an F1-score of 0.97 together with a precision level of 0.99 and recall measurement of 0.95. The detection metrics for NonDemented equaled 0.93, 0.99, and 0.96 and VeryDemented metrics showed 0.95, 0.96, and 0.95. The confusion matrix in Table 4 demonstrates the model's excellent reliability by properly identifying 355 cases of MildDemented along with 176 instances of NonDemented and 186 VeryDemented cases among the total 743 instances. MobileNet achieved superior class distribution together with enhanced generalization capabilities by reducing the number of wrong negative outcomes and incorrect positive predictions. The model achieves good performance due to its lightweight structure and efficient depthwise separable convolutions that eliminate parameter redundancies and enhance prediction generalization capabilities.

Table 3. Summary of experiment results

Model	Accuracy (%)	Precision	Recall	F1-Score
VGG16	91.39	0.91	0.92	0.91
MobileNet	96.50	0.96	0.97	0.96
MobileNet + Snake optimizer	97.71	0.97	0.98	0.98

Table 4. Comparison of accuracy between related work and proposed work

Ref.	Method	Dataset	Accuracy (%)
[16]	SVM	ADD	84
[17]	3D-CNN	ADNI	Not specified
[18]	Deep neural network	ADNI-2	78.91
Proposed Work	MobileNet + Snake Optimizer	ADNI	97.71

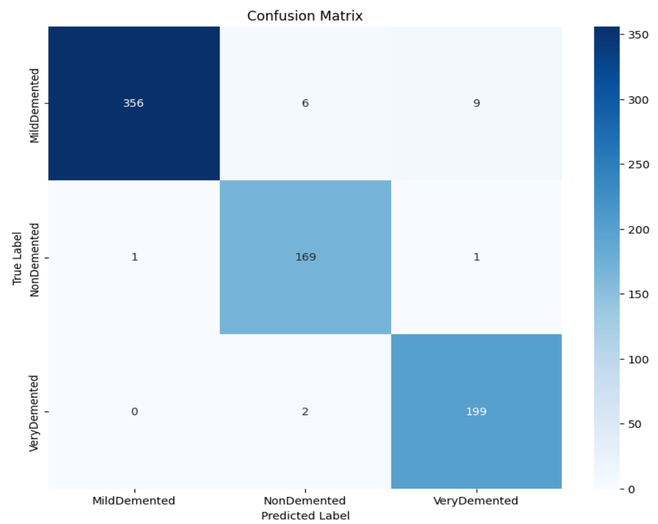

Figure 3. Confusion matrix of the VGG16 model for Alzheimer's disease stage classification

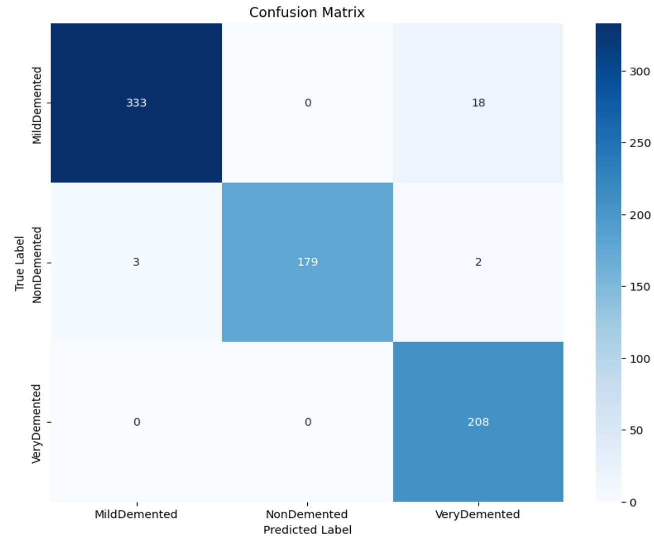

Figure 4. Confusion matrix of the MobileNet model for Alzheimer's disease stage classification

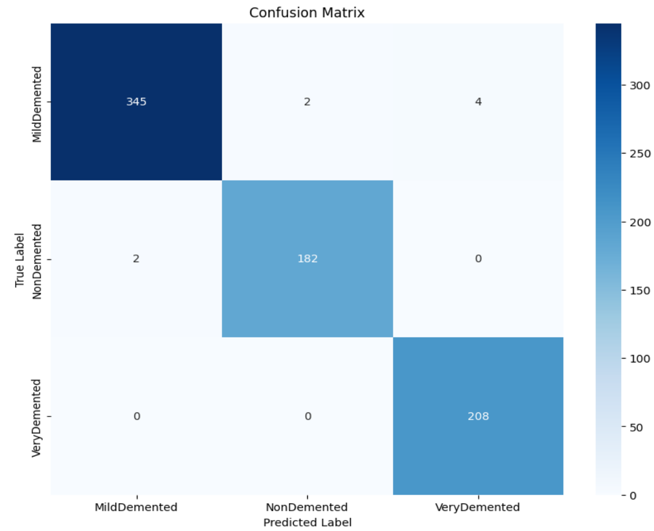

Figure 5. Confusion matrix of the hybrid MobileNet model optimized with Snake algorithms for AD stage classification

MobileNet achieved a 97.71% accuracy level after the SOA implementation as shown in Table 3. The assessment metrics for MildDemented category showed precision at 0.99 and recall at 0.97 along with an F1-score of 0.98. The accuracy scores for NonDemented amounted to 0.92, 0.99, 0.95 while VeryDemented achieved a perfect or near-perfect accuracy of 1.00, 0.99, 1.00.

The evaluation of the proposed models was conducted through the analysis of confusion matrices, which provide a comprehensive overview of classification performance across different AD stages. Figure 3 presents the confusion matrix for the VGG16 model, where a moderate number of misclassifications is observed, particularly between MildDemented and VeryDemented categories. Figure 4 represents the confusion matrix of the MobileNet model, which is better in terms of classification results and less misclassifications as well as class separation in comparison to VGG16. The hybrid MobileNet-Snake model with the best balanced and accurate results is presented in Figure 5 and it almost perfectly separates the Mild Demented, Non Demented, and Very Demented classes. The significant decrease in the rate of misclassifications confirms the efficiency of the suggested hybridization optimization strategy in enhancing the accuracy of the prediction of the AD stage.

The SOA enhanced key parameters such as the learning rate as well as dropout through which the system reached greater performance in terms of convergence and robust working.

The vast result of these studies makes MobileNet a potent solution that is enhanced with the optimization of SOA. Tough constraints of VGG16 have been experienced because of its heavy architecture when carrying out tasks with small dataset. The optimized structure of MobileNet was a contributory factor with SOA parameter optimization to attain excellent balanced performance outcomes across all categories of AD.

Table 4 analyzes the similarities and differences of the proposed method and some of the existing methods of AD classification. It provides the summarization of major studies, outlining their models, datasets, and stated accuracy. With the help of ADD dataset, one study with a Support Vector Machine (SVM) achieved 84 percent accuracy, and Ju et al. [18] used deep neural networks on the ADNI-2 dataset and achieved 78.91 percent accuracy. Hosseini-Asl et al. [17] used a 3D-CNN architecture using the ADNI dataset but failed to provide precise results in the form of accuracy.

With MobileNet algorithm optimised with the SOA on the ADNI data, we obtained significantly a better accuracy of 97.71. This improvement highlights the theoretical and practical importance of our strategy. The framework proposed consists of a combination of the characteristics of transfer learning and the optimization of hyperparameters on the basis of SOA-based metaheuristic optimization to enhance the generalization and classification results of machine and deep learning models. The findings validated the effectiveness and the strength of the approach as SOA recognizes the best configurations that the manual tuning or grid search have overlooked. In mobile applications, MobileNet has an efficient architecture that is both lightweight and provides a high-level of accuracy and Computer speed, thus suitably applicable in clinical settings, which require both accuracy and scalability. The results of Table 4 show the numerical benefits and the approach of improving the rapid and accurate diagnosis of AD.

5. CONCLUSION

The research paper has constructed and tested a novel deep learning algorithm to identify multi-class AD using T1-weighted MRI neuroimaging data. The basic worth of this undertaking lies in the integration of MobileNet framework with SOA to make autonomous adjustment of the hyperparameters. The proposed system had a classification accuracy of 97.71% with ImageNet transfer learning, best learning rates and dropout rates, and alteration of dense layer parameters. MobileNet performance with physiologically motivated optimization gave better performance results compared to VGG16 due to the synergistic effect of the two.

The model of the MobileNet + SOA was found to be very precise with better recall and F1-scores in the NonDemented, MildDemented and VeryDemented stages of depression, and thus proves to be able to identify simple degrees of cognitive decline as well as the advanced levels. The confusion matrices confirmed the consistent conclusions method of the classification because it showed low misclassification errors and equivalent sensitivity values among the classes. Early identification of mild cognitive impairment with the help of this model is important to clinical practice due to the fact that it allows medical workers to identify the necessary intervention to prevent the development of the disease and improve the quality of the life of a patient.

This framework shows that deep learning with the help of metaheuristic optimization can lead to effective and successful diagnostic tools that can scale intelligently and quickly. The implementation of MobileNet enables real-time utilization in limited medical settings since such a low-cost computing model is cost effective. The solidware optimization as a service system gives researchers a consistent way to optimize deep learning models using automatic configuration adjustments, which leads to increased repeatability on test datasets.

The findings of the given study can be utilized to initiate further research. The framework could use the benefit of adding the PET imaging data such as cerebral spinal fluid biomarkers and genetic markers because the addition of this data would provide an added information that enhances diagnostic accuracy. The application of this methodology to Parkinson's Disease or Huntington's Disease in further researches should identify the scope of its application to various neurodegenerative diseases and clinical consequences. Explainable methods (Grad-CAM or SHAP) that will be used with the system will enhance its clinical adoption potential as they will show the clinicians interpretable model decisions.

The study reveals the effectiveness of MobileNet + SOA-based framework in enhancing the diagnosis of Alzheimer through a high level of accuracy. The model offers a bright perspective of implementation in the smart healthcare systems since it demonstrates high performance rates and functioning efficiency and adaptability in the neurodegenerative disorder analysis and therapy planning.

REFERENCES

[1] Hardy, J. (1997). Amyloid, the presenilins and Alzheimer's disease. Trends in Neurosciences, 20(4): 154-159. https://doi.org/10.1016/S0166-2236(96)01030-2

[2] Alzheimer's Association. (2021). Facts and Figures. Alzheimer's Disease and Dementia. https://www.alz.org/alzheimers-dementia/facts-figures, accessed on 16 January 2021.

[3] Jicha, G.A., Parisi, J.E., Dickson, D.W., Johnson, K., et al. (2006). Neuropathologic outcome of mild cognitive impairment following progression to clinical dementia. Archives of Neurology, 63(5): 674-681. https://doi.org/10.1001/archneur.63.5.674

[4] Klöppel, S., Stonnington, C.M., Chu, C., Draganski, B., Scahill, R.I., Rohrer, J.D., Fox, N.C., Jack, Jr, C.R., Ashburner, J., Frackowiak, R.S. (2008). Automatic classification of MR scans in Alzheimer's disease. Brain, 131(3): 681-689. https://doi.org/10.1093/brain/awm319

[5] Rathore, S., Habes, M., Iftikhar, M.A., Shacklett, A., Davatzikos, C. (2017). A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. NeuroImage, 155: 530-548. https://doi.org/10.1016/j.neuroimage.2017.03.057

[6] Allen, G.I., Amoroso, N., Anghel, C., Balagurusamy, V., et al. (2016). Crowdsourced estimation of cognitive decline and resilience in Alzheimer's disease.

Alzheimer's & Dementia, 12(6): 645-653. https://doi.org/10.1016/j.jalz.2016.02.006
[7] Sarica, A., Cerasa, A., Quattrone, A., Calhoun, V. (2016). A machine learning neuroimaging challenge for automated diagnosis of mild cognitive impairment. Journal of Neuroscience Methods. https://www.kaggle.com/c/mci-prediction, accessed on 18 Jan. 2021.
[8] LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. Nature, 521(7553): 436-444. https://doi.org/10.1038/nature14539
[9] Samper-González, J., Burgos, N., Bottani, S., Fontanella, S., et al. (2018). Reproducible evaluation of classification methods in Alzheimer's disease: Framework and application to MRI and PET data. NeuroImage, 183: 504-521. https://doi.org/10.1016/j.neuroimage.2018.08.042
[10] Ali, S.M., Abbosh, Y.M., Breesam, A.M., Ali, D.M., Alhummada, I.A. (2024). Heart diseases classification through deep learning techniques: A review. AIP Conference Proceedings, 3232: 020022. https://doi.org/10.1063/5.0236126
[11] Mesnil, G., Dauphin, Y., Glorot, X., Rifai, S., Bengio, Y., Goodfellow, I., Lavoie, E., Muller, X., Desjardins, G., Warde-Farley, D., Vincent, P., Courville, A., Bergstra, J. (2012). Unsupervised and transfer learning challenge: A deep learning approach. In Proceedings of the 2011 International Conference on Unsupervised and Transfer Learning workshop, Washington, USA, pp. 97-111.
[12] Raina, R., Ng, A.Y., Koller, D. (2006). Constructing informative priors using transfer learning. In Proceedings of the 23rd International Conference on Machine Learning, New York, United States, pp. 713-720. https://doi.org/10.1145/1143844.1143934
[13] Zhou, L., Wang, Y., Li, Y., Yap, P. T., Shen, D., Alzheimer's Disease Neuroimaging Initiative (ADNI). (2011). Hierarchical anatomical brain networks for MCI prediction: Revisiting volumetric measures. PloS One, 6(7): e21935. https://doi.org/10.1371/journal.pone.0021935
[14] Ker, J., Wang, L., Rao, J., Lim, T. (2017). Deep learning applications in medical image analysis. IEEE Access, 6: 9375-9389. https://doi.org/10.1109/ACCESS.2017.2788044
[15] Sarraf, S., DeSouza, D. D., Anderson, J., Tofighi, G. (2016). DeepAD: Alzheimer’s disease classification via deep convolutional neural networks using MRI and fMRI. BioRxiv, 070441. https://doi.org/10.1101/070441
[16] Mathew, J., Mekkayil, L., Ramasangu, H., Karthikeyan, B.R., Manjunath, A.G. (2016). Robust algorithm for early detection of Alzheimer's disease using multiple feature extractions. In 2016 IEEE Annual India Conference (INDICON), Bangalore, India, pp. 1-6. https://doi.org/10.1109/INDICON.2016.7839026
[17] Hosseini-Asl, E., Keynton, R., El-Baz, A. (2016). Alzheimer's disease diagnostics by adaptation of 3D convolutional network. In 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, pp. 126-130. https://doi.org/10.1109/ICIP.2016.7532332
[18] Ju, R., Hu, C., Li, Q. (2017). Early diagnosis of Alzheimer's disease based on resting-state brain networks and deep learning. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 16(1): 244-257. https://doi.org/10.1109/TCBB.2017.2776910
[19] Razzak, M.I., Imran, M., Xu, G. (2018). Efficient brain tumor segmentation with multiscale two-pathway-group conventional neural networks. IEEE Journal of Biomedical and Health Informatics, 23(5): 1911-1919. https://doi.org/10.1109/JBHI.2018.2874033
[20] Razzak, M.I., Naz, S., Zaib, A. (2018). Deep learning for medical image processing: Overview, challenges and the future. In Classification in BioApps. Lecture Notes in Computational Vision and Biomechanics. https://doi.org/10.1007/978-3-319-65981-7_12
[21] Farooq, A., Anwar, S., Awais, M., Rehman, S. (2017). A deep CNN based multi-class classification of Alzheimer's disease using MRI. In 2017 IEEE International Conference on Imaging systems and techniques (IST), Beijing, China, pp. 1-6. https://doi.org/10.1109/IST.2017.8261460
[22] Bäckström, K., Nazari, M., Gu, I.Y.H., Jakola, A.S. (2018). An efficient 3D deep convolutional network for Alzheimer's disease diagnosis using MR images. In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, pp. 149-153. https://doi.org/10.1109/ISBI.2018.8363543
[23] Gautam, C., Mishra, P.K., Tiwari, A., Richhariya, B., Pandey, H.M., Wang, S., Tanveer, M., Alzheimer’s Disease Neuroimaging Initiative. (2020). Minimum variance-embedded deep kernel regularized least squares method for one-class classification and its applications to biomedical data. Neural Networks, 123: 191-216. https://doi.org/10.1016/j.neunet.2019.12.001
[24] Liu, M., Li, F., Yan, H., Wang, K., Ma, Y., Shen, L., Xu, M., Alzheimer’s Disease Neuroimaging Initiative. (2020). A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer’s disease. Neuroimage, 208: 116459. https://doi.org/10.1016/j.neuroimage.2019.116459
[25] Kazemi, Y., Houghten, S. (2018). A deep learning pipeline to classify different stages of Alzheimer's disease from fMRI data. In 2018 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), St. Louis, MO, USA, pp. 1-8. https://doi.org/10.1109/CIBCB.2018.8404980
[26] Tajbakhsh, N., Shin, J.Y., Gurudu, S.R., Hurst, R.T., Kendall, C.B., Gotway, M.B., Liang, J. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE Transactions on Medical Imaging, 35(5): 1299-1312. https://doi.org/10.1109/TMI.2016.2535302
[27] Ebrahimi-Ghahnavieh, A., Luo, S., Chiong, R. (2019). Transfer learning for Alzheimer's disease detection on MRI images. In 2019 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT), Bali, Indonesia, pp. 133-138. https://doi.org/10.1109/ICIAICT.2019.8784845
[28] Wang, S., Wang, H., Shen, Y., Wang, X. (2018). Automatic recognition of mild cognitive impairment and alzheimers disease using ensemble based 3d densely connected convolutional networks. In 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, pp. 517-523. https://doi.org/10.1109/ICMLA.2018.00083

[29] Najjar, E., Alkhaykanee, N.A., Breesam, A.M. (2025). Classification of Arabic documents with five classifier models using machine learning. International Journal on Informatics Visualization, 9(1): 365-370. https://doi.org/10.62527/joiv.9.1.2539