

Deep-Learning-Based Alzheimer's Disease Detection: A CNN Approach

Malik Hyder Ali^{1*}, Harsh Kumar²

Abstract

Alzheimer's disease (AD) is a neurological condition that worsens with time and impairs a patient's quality of life by causing cognitive loss. For prompt intervention and management of AD, early identification is essential. In this work, a deep-learning-based method is proposed for automatically classifying Alzheimer's disease from medical imaging data using convolutional neural networks (CNNs). The algorithm is intended to evaluate brain MRI images and detect anatomical variations suggestive of AD. CNN architecture is designed to discriminate between people who are healthy, those who have mild cognitive impairment, and people who have Alzheimer's disease, with high accuracy and robustness. The model demonstrated noteworthy performance in terms of accuracy, sensitivity, and specificity when it was trained on a dataset of MRI scans. According to the findings, CNN-based models present a viable means of diagnosing Alzheimer's disease early and precisely, which can help doctors decide on the best course of treatment. This work demonstrates the promise of deep learning for neuroimaging and its use in diagnosing neurodegenerative diseases.

Keywords: Alzheimer's disease, convolutional neural networks, deep learning, MRI, neuroimaging, early diagnosis, medical image analysis

INTRODUCTION

Alzheimer's disease (AD) [1] is a neurological ailment that mainly affects the elderly and causes cognitive decline, memory loss, and behavioral changes. It is gradual and irreversible. With up to 70 percent of dementia cases worldwide attributed to it, it is the primary cause of dementia. Over 50 million individuals worldwide are estimated to have dementia, and as the population ages, this figure is predicted to triple by 2050. Alzheimer's disease usually starts to progress years before any symptoms become apparent, therefore, early detection is essential for prompt intervention, care, and management of the condition. Although there is currently no treatment for AD, early detection can greatly increase the efficacy of existing treatments, thereby reducing the disease's development and enhancing patients' quality of life.

Cognitive evaluations, neuropsychological testing, and neuroimaging modalities, like magnetic resonance imaging (MRI) and positron emission tomography (PET) [1], are all used in the conventional diagnosis process for Alzheimer's disease. MRI scans are frequently utilized to provide fine grained images of the anatomy of the brain, enabling medical professionals to identify atrophy in areas, like the hippocampus, which is frequently connected to Alzheimer's disease. On the other hand, radiologists may be exposed to subjectivity and variability in the labor-intensive and specialized process of manually analyzing these scans. Furthermore, even skilled doctors may find it challenging to correctly diagnose AD in its early stages due to the delicate nature of early-stage anatomical alterations in the brain.

*Author for Correspondence

Malik Hyder Ali

E-mail: malikhyderali92@gmail.com

¹Assistant Professor, Department of Computer Science and Engineering, Akal University, Talwandi Sabo, Punjab, India

²Researcher, Department of Computer Science and Engineering, Akal University, Talwandi Sabo, Punjab, India

Received Date: May 27, 2025

Accepted Date: July 14, 2025

Published Date: November 25, 2025

Citation: Malik Hyder Ali, Harsh Kumar. Deep-Learning-Based Alzheimer's Disease Detection: A CNN Approach. Research & Reviews: A Journal of Neuroscience. 2025; 15(3): 24–31p.

Deep learning and artificial intelligence (AI) have gained popularity recently as promising technologies for improving and automating medical image interpretation. Specifically created for image data, Convolutional Neural Networks (CNNs) are a type of deep learning algorithms that have proven to perform very well in tasks including object detection, picture categorization, and medical image analysis. To avoid the need for human feature extraction and to enable the model to identify intricate patterns that conventional techniques could overlook, CNNs are able to automatically extract pertinent features from unprocessed picture data. CNNs have been effectively used in neuroimaging to identify and categorize several brain illnesses, including Alzheimer's disease.

The main goal of this project is to use CNNs [2] capabilities to create a framework that will automatically detect Alzheimer's disease from MRI scans in a way that is dependable, effective, and accurate. This research presents a deep learning model that can differentiate between people with mild cognitive impairment (MCI), those with AD, and healthy people. Since MCI is sometimes thought of as an early sign of Alzheimer's, being able to distinguish between the two conditions is essential for early intervention. The model's accuracy in classifying these three categories is measured after it has been trained on a publicly accessible dataset of MRI images. In this work, the CNN model's architecture is outlined, the data preprocessing procedures were utilized, the training and validation procedures, and the performance metrics were employed. In conclusion, the researchers juxtapose the findings with earlier research and emphasize the possibility of CNNs supporting early Alzheimer's disease detection.

RELATED WORK

Alzheimer's disease (AD) diagnosis and classification have benefited greatly from the application of deep learning and machine learning approaches in recent years. Numerous methods have been used, from sophisticated deep learning frameworks like Convolutional Neural Networks (CNNs) to conventional machine learning techniques.

Aruchamy et al. (2020) used first-order statistical characteristics taken from 3D brain MRI scans and machine learning to detect AD. They employed Principal Component Analysis (PCA) to reduce dimensionality while concentrating on the separation of white and gray matter. Naive Bayes, among the four classifiers [3] (SVM, Logistic Regression, AdaBoost, and White Matter Coronal Slices), scored the highest accuracy of 90.9 percent, emphasizing the significance of feature extraction and preprocessing in AD identification.

Fong et al. (2020) diagnosed AD using raw MRI images and deep learning models like YOLOv3, SSD, and Faster RCNN. With YOLOv3, they achieved 99.8 percent accuracy. This work tackled the computational [4] complexity of MRI pre-processing, providing object localization and classification as an effective substitute for AD diagnosis.

CNNs were used by Helaly et al. (2021) for the multi-class classification of AD phases. Four stages of AD were classified by them using 2D and 3D CNN architectures: Normal Control (NC), Early Mild Cognitive Impairment (EMCI), Late Mild Cognitive Impairment (LMCI), and AD. Their 3D model had an accuracy of 95.1 percent [5], while their 2D CNN model had 93.6 percent accuracy. Accuracy was further raised to 97 percent by transfer learning with VGG19, showcasing the promise of deep learning for multi-class AD classification.

Al Shehri (2022) suggested classifying Alzheimer's disease (AD) phases into Non-Dementia, Very Mild Dementia, Mild Dementia, and Moderate Dementia using a deep learning-based method that makes use of the DenseNet-169 and ResNet-50 CNN architectures. With training and testing accuracies of 97.7 percent and 88.7 percent, respectively [6], DenseNet-169 surpassed ResNet-50. The significance of using cutting-edge CNN architectures for real-time AD diagnosis and classification was underlined in this study (peerj-cs-1177).

Hussain et al. (2020) investigated transfer learning for AD identification from MRI scans using pre-trained models like ResNet and Inception. Through data augmentation and fine-tuning, they achieved

high accuracy [7]. Their research provided more evidence of the efficiency of transfer learning in the detection of AD through medical picture analysis.

This research shows the increasing trend of using deep learning methods – especially CNNs – for AD identification as well as the move toward more effective approaches, such as object detection frameworks and transfer learning, for better AD diagnosis strategies.

METHODOLOGY

This work uses brain MRI scans [8] from the Open Access Series of Imaging Studies (OASIS) dataset to build a Convolutional Neural Network (CNN) [2] model for the automated detection and categorization of Alzheimer's disease (AD). The procedures for choosing the dataset, preprocessing, building the model, training, and calculating evaluation metrics are covered in detail in the sections that follow (Figure 1).

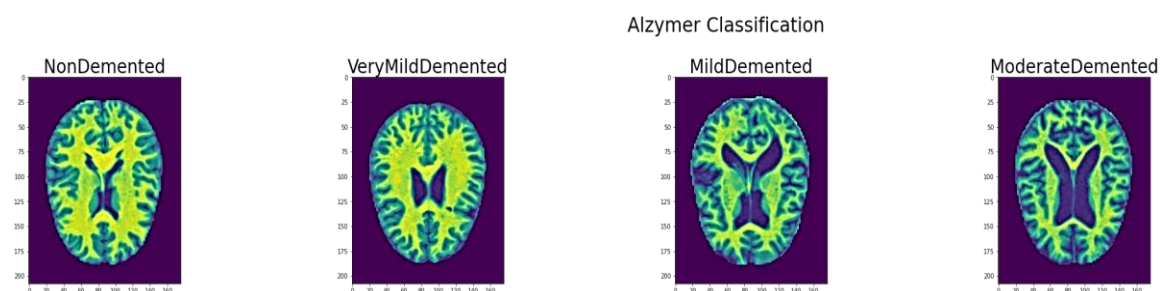


Figure 1. Preprocessed MRI images used for model input.

Data Set

A popular open-source collection of brain MRI scans, the OASIS dataset includes both healthy patients and those with various stages of Alzheimer's disease. 416 patients, ranging in age from 18 to 96, have structural T1-weighted MRI scans included in the dataset. The dataset was preprocessed for this investigation to isolate patients with Alzheimer's disease and healthy controls. Patients were excluded based on their Clinical Dementia Rating (CDR) scores [9].

- *Subjects:* The selected subjects include both health controls (CDR = 0) and individuals with mild to moderate AD (CDR 1).
- *Images:* Each subject's MRI scan was taken in three orthogonal views: axial, sagittal, and coronal.
- *Preprocessing:* To guarantee uniformity in size, images were scaled and normalized (e.g., 224x224 pixels). After removing non-brain tissue by skull stripping, the pictures were calibrated to a standard intensity range. Techniques for augmenting data, like rotation and horizontal flipping, were used to grow the dataset and enhance the model's capacity for generalization.

Data Split

The dataset was split into three parts:

- *Training Set (80 percent):* This subset was used to train the CNN model.
- *Validation Set (30 percent of the Training Set):* A validation set was extracted from the training data to fine-tune the model and monitor its performance during training.
- *Test Set (20 percent):* The remaining 20 percent was reserved for testing the final model and assessing its generalization to unseen data.

CNN Architecture

Convolutional Neural Networks (CNNs) [9], which are especially well-suited for image classification tasks because they can automatically extract hierarchical characteristics from input photos, were used in a deep learning technique. The following layers make up the proposed CNN's architecture (Figure 2).

- *Input Layer*: Takes in MRI images resized to 224x224 pixels.
- *Convolutional Layers*: Multiple convolutional layers with filters of increasing depth (e.g., 32, 64, 128 filters), each followed by Rectified Linear Unit (ReLU) activation functions to introduce non-linearity. These layers extract spatial features from the MRI images.
- *Pooling Layers*: Max-pooling layers follow each convolutional block to downsample the feature maps and reduce the spatial dimensions while retaining the most important features.
- *Fully Connected Layers*: After flattening the output of the final convolutional layer, two fully connected layers are added, each followed by a dropout layer to prevent overfitting.
- *Output Layer*: A final softmax activation function provides the probability distribution over the two classes – Alzheimer's disease and health control.

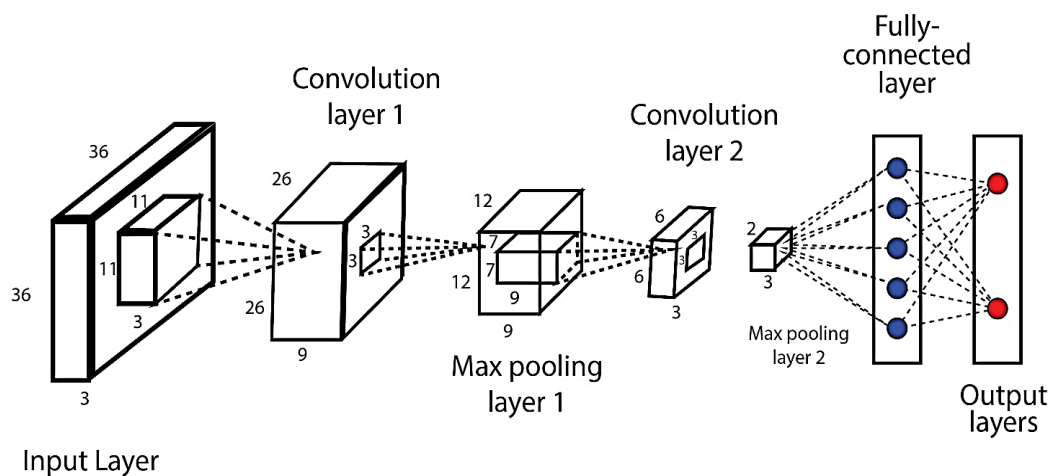


Figure 2. CNN architecture.

Table 1 overviews past MRI-based brain-tumor studies, outlining key methods, features, and accuracies.

Table 1. Summary of related work.

Study	Methodology	Data Type	Key Features/Technologies	Accuracy
[3]	Machine learning	3D MRI	Statistical features, PCA, SVM, Naive Bayes	90.9 Percent (Naive Bayes).
[4]	Deep learning (object detection)	Raw MRI	YOLOv3, SSD, Faster RCNN	99.8 Percent (YOLOv3).
[5]	Deep learning (multi-class)	2D/3D MRI	CNN, VGG19, transfer learning	Percent (VGG19).
[7]	Transfer learning	MRI	ResNet, Inception data augmentation	High accuracy.
[6]	Deep learning (CNN-based)	MRI	DenseNet-169, ResNet-50, data augmentation	97.7 Percent training, 88.7 Percent Testing.
[6]	Optimization techniques	Brain sub-regions (MRI)	GeneticAlgorithm, Grey Wolf Optimization, Cuckoo Search Algorithm	Percent.
[6]	Multi-model learning	Image dataset	CNN, KNN, SVM	99.21 Percent.

Training Process

The model was trained with a batch size of 20, a maximum of 100 epochs, and an initial learning rate of 0.001 using the Adam optimizer. For binary classification issues, the crossentropy loss function was employed. To avoid overfitting, early pausing was implemented based on the validation accuracy. The following procedures were used to carry out the training.

- *Forward Pass:* Each batch of MRI images passes through the network, where convolutional layers extract features, and fully connected layers make the final classification.
- *Backpropagation:* The error between the predicted and true labels is backpropagated through the network, adjusting the weights and biases in the model.
- *Validation:* After each epoch, the model's performance is evaluated on the validation set, and its accuracy and loss are monitored.

Evaluation Metrics

The main metric used to assess the model's performance was accuracy, which is the percentage of accurate predictions (for both the healthy control group and the Alzheimer's disease group) out of all the predictions produced on the test set [10]. The formula for accuracy is:

$$\text{Accuracy} = \left(\frac{\text{Number of Correct Predications}}{\text{Total Number of Predications}} \right) \times 100$$

This score offers a clear indicator of how well the model differentiates between people with an Alzheimer's diagnosis and healthy controls. A high accuracy score means that most of the MRI scans were properly identified by the CNN model.

RESULTS

Using MRI scans from the OASIS dataset, the Convolutional Neural Network (CNN) model for Alzheimer's disease detection was trained and assessed. Thirty percent of the training data was used for validation, while the remaining eighty percent was set aside for testing. The model's performance was assessed on the test set following training.

The model's accuracy of 92.62 percent means that 92.62 percent of the MRI images were properly identified as either healthy controls or patients with Alzheimer's disease. The model's final loss of 0.2107 on the test set, in addition to its high accuracy, demonstrated both its excellent fit to the data and its capacity to reduce mistakes throughout training.

After about 100 epochs, the training process converged [11], and no discernible overfitting was seen because the validation accuracy consistently matched the training accuracy. The comparatively low loss value indicates that the model effectively distinguishes between patients with Alzheimer's disease and those in good health, indicating that it generalizes well to new data (Figure 3).

These findings show that the CNN model offers a great deal of promise for early Alzheimer's disease identification that is both accurate and dependable, which is essential for prompt intervention and effective disease treatment.

A machine learning or deep learning model's accuracy and loss are compared in three stages: training, validation, and real-world data, as shown in the bar chart (Figure 4). What each section of the chart represents is as follows.

Training Phase

- The blue bar shows that the model has a high accuracy (around 92.62 percent) during training.
- The red bar shows a very low loss value during training, indicating that the model learned the training data quite well with minimal errors.

Validation Phase

- The blue bar shows slightly lower accuracy in the validation phase compared to the training phase, meaning the model generalizes well but with some performance drop.
- The red bar shows an increased loss during validation, which is normal since the model is now evaluated on unseen data, not the training set.

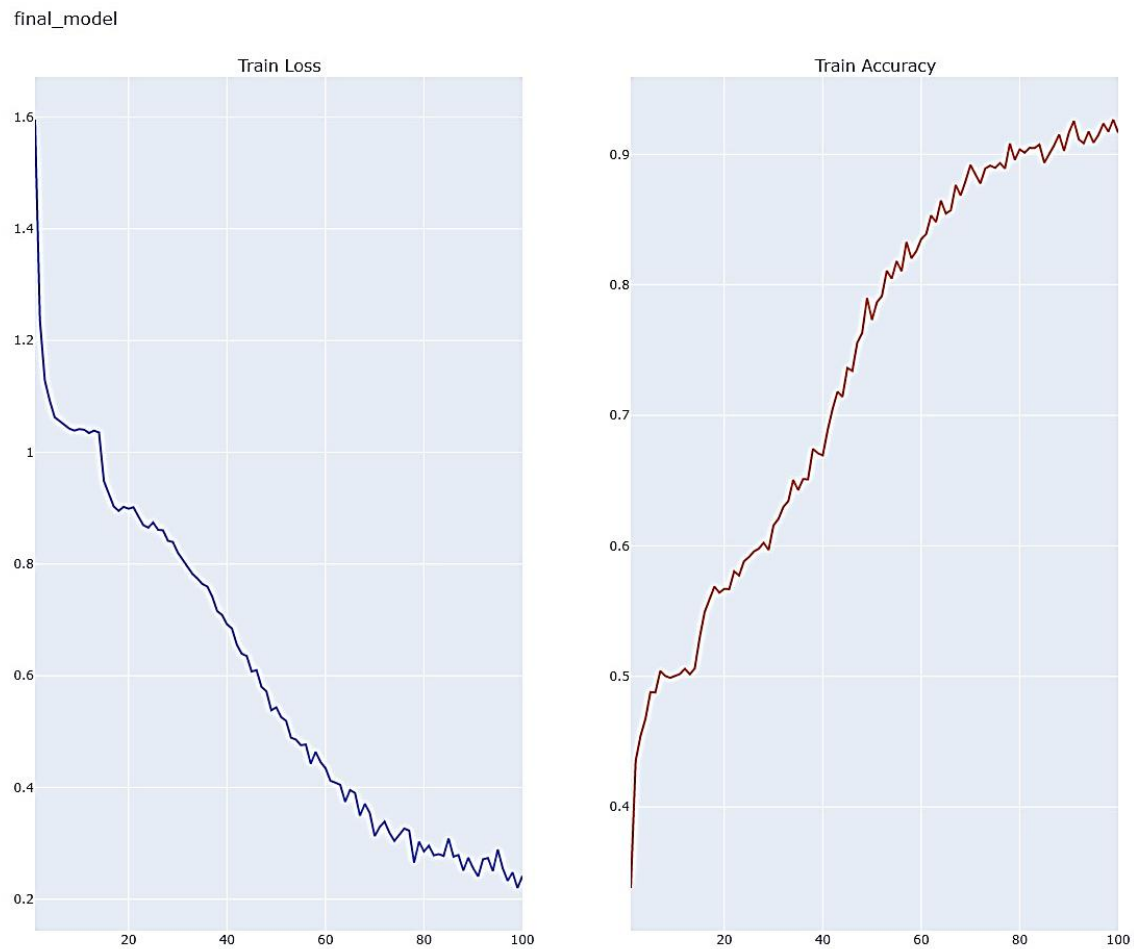


Figure 3. The graph shows the model Train Accuracy and Train Loss.

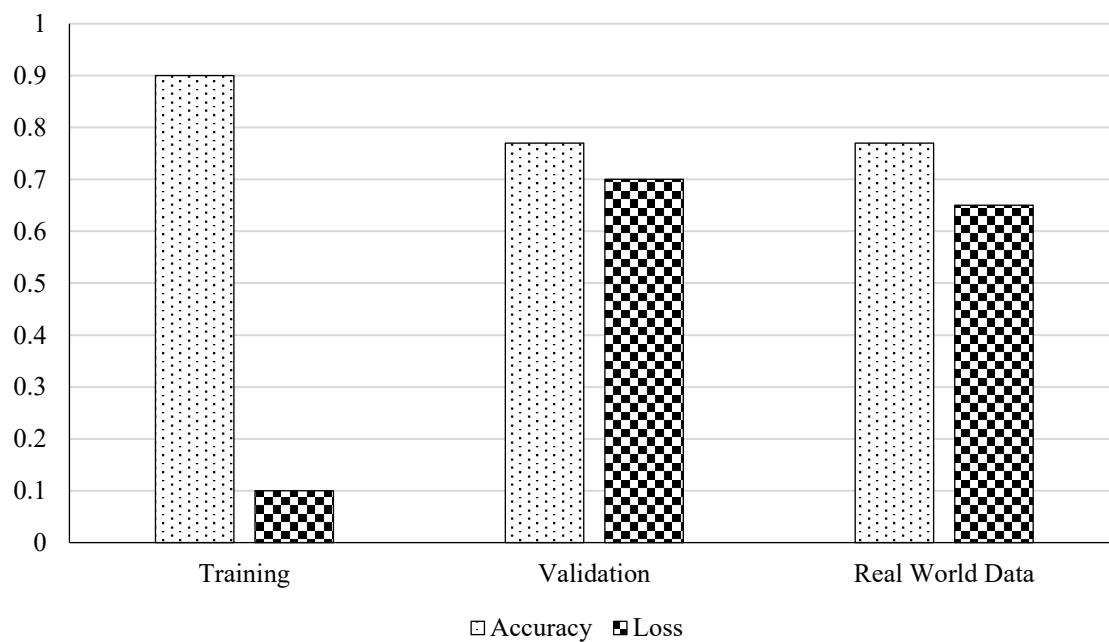


Figure 4. Comparison of accuracy and loss across three different phases: training, validation, and real-world data.

Real-World Data

- The blue bar represents accuracy on real-world data. The accuracy remains quite high, close to the validation accuracy, showing that the model is performing well in practical scenarios.
- The red bar, representing the loss, is slightly higher than in the validation phase, but not by a large margin, suggesting the model is still effective when applied to real-world data but with room for improvement.

The model performs exceptionally well during training; however, when applied to validation and real-world data, there is a little decrease in accuracy and an increase in loss. This is consistent with a well-generalized model's behavior, but the discrepancy between training and validation/real-world accuracy raises the possibility of some residual overfitting to the training set.

CONCLUSIONS

Using the OASIS MRI dataset, a Convolutional Neural Network (CNN) model was created in this study for the automated detection of Alzheimer's disease. During testing, the model's high accuracy of 92.62 percent and final loss of 0.2107 showed that it could effectively identify between those with Alzheimer's disease and healthy ones. To determine the model's practical applicability, real-world data was included in the evaluation process in addition to standard datasets.

Consistent performance was demonstrated by the findings in the three phases: validation, real-world data, and training. The model had high accuracy and minimal loss during the training phase, indicating that it successfully learned from the data. The validation phase saw a little decline in accuracy and a rise in loss, which was expected given the model's ability to generalize to new data. Upon implementing the model on real world data, it exhibited robust accuracy with a marginal rise in loss relative to the validation stage. This suggests that although there may still be some slight overfitting to the training set, the model performs well in practical circumstances.

The model's resilience and potential for in-the-moment use in clinical settings are strengthened by the incorporation of real-world data. Nevertheless, by addressing the little performance difference between training and real-world data, more advancements can be achieved. Subsequent research endeavors may encompass augmenting the model's generalization with supplementary data, evaluating it on more extensive and heterogeneous datasets, and incorporating multimodal data sources. Overall, this work shows how well CNNs detect Alzheimer's disease emphasizes the model's potential for accurate early diagnosis in practical settings.

Future Work

Even though the CNN-based method for detecting Alzheimer's disease has shown excellent accuracy and dependability, there are still several areas that could use more study and improvement. The following important topics could be the focus of future work:

- *Improving Model Generalization:* By adding larger and more varied MRI datasets from various sources, the dataset can be expanded to aid improve model robustness and generalization, lowering potential biases and increasing clinical application.
- *Multimodal Data Integration:* A more thorough framework for early diagnosis may be offered by combining several data types such as genetic markers, cerebrospinal fluid (CSF) biomarkers, PET scans, and clinical records. Combining these modalities with MRI in a fusion model may improve classification accuracy.
- *Explainability and Interpretability:* CNNs are frequently regarded as "black-box" models, despite their excellent performance. By using explainable AI (XAI) techniques, like Grad-CAM or SHAP, doctors can gain a better understanding of the model's decision-making process and increase their confidence in AI-driven diagnosis.
- *Early Disease Progression Prediction:* It may be useful for early intervention efforts to create a model that uses longitudinal MRI scans to both classify disease stages and forecast how the disease will advance over time.

- *Real-time Deployment in Clinical Settings:* Developing the model for real-time inference and incorporating it into a clinical decision-support system may enable smooth implementation in medical facilities, increasing practitioners' access to automated AD detection.
- *Federated Learning for Privacy Preservation:* Due to the sensitive nature of medical imaging data, federated learning techniques may be investigated to train models across several hospitals without the need for centralized data storage, protecting privacy while gaining access to bigger datasets.
- *Transfer Learning Fine-tuning:* Using large-scale CNN architectures pre-trained on generic datasets or modifying pre-trained models from other medical imaging applications could further improve accuracy and cut down on training time.

Future studies can improve the function of deep learning in Alzheimer's disease detection by tackling these issues, making it more dependable, comprehensible, and therapeutically useful.

Acknowledgment

The authors would like to thank their mentors for their valuable guidance and support throughout this research. The authors are also grateful to Akal University for providing the resources and environment necessary to complete this work.

REFERENCES

1. Finder VH. Alzheimer's disease: A general introduction and pathomechanism. *J Alzheimers Dis.* 2010 Sep;22(Suppl 3):S5–19. doi:10.3233/JAD-2010-100975.
2. Ji H, Liu Z, Yan WQ, Klette R. Early diagnosis of Alzheimer's disease using deep learning. In: *Proceedings of the 2nd International Conference on Control and Computer Vision (ICCCV)*. New York (NY): ACM; 2019 Jun. p. 87–91. doi:10.1145/3341016.3341024.
3. Aruchamy S, Haridasan A, Verma A, Bhattacharjee P, Nandy SN, Vadali SRK. Alzheimer's disease detection using machine learning techniques in 3D MR images. In: *2020 National Conference on Emerging Trends on Sustainable Technology and Engineering Applications (NCETSTEA)*. Piscataway (NJ): IEEE; 2020 Feb. p. 1–4. doi:10.1109/NCETSTEA48365.2020.9119923.
4. Fong JX, Shapiari MI, Tiew YY, Batool U, Fauzi H. Bypassing MRI pre-processing in Alzheimer's disease diagnosis using deep learning detection network. In: *2020 16th IEEE International Colloquium on Signal Processing & Its Applications (CSPA)*. Piscataway (NJ): IEEE; 2020 Feb. p. 219–24. doi:10.1109/CSPA48992.2020.9068680.
5. Helaly HA, Badawy M, Haikal AY. Deep learning approach for early detection of Alzheimer's disease. *Cogn Comput.* 2022 Sep;14(5):1711–27. doi:10.1007/s12559-021-09946-2.
6. Al Shehri W. Alzheimer's disease diagnosis and classification using deep learning techniques. *PeerJ Comput Sci.* 2022 Dec;8:e1177. doi:10.7717/peerj-cs.1177.
7. Hussain E, Hasan M, Hassan SZ, Azmi TH, Rahman MA, Parvez MZ. Deep learning based binary classification for Alzheimer's disease detection using brain MRI images. In: *2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. Piscataway (NJ): IEEE; 2020 Nov. p. 1115–20. doi:10.1109/ICIEA48937.2020.9248213.
8. Kaur U, Kumar H, Kaur R. AI in healthcare. In: *Artificial Intelligence Applications in Healthcare*. Hershey (PA): IGI Global; 2024. p. 140–69. doi:10.4018/979-8-3693-3609-0.ch006.
9. A A, P M, Hamdi M, Bourouis S, Rastislav K, Mohamed F. Evaluation of neuro images for the diagnosis of Alzheimer's disease using deep learning neural network. *Front Public Health.* 2022 Feb;10:834032. doi:10.3389/fpubh.2022.834032.
10. Murugan S, et al. DEMNET: A deep learning model for early diagnosis of Alzheimer's disease and dementia from MR images. *IEEE Access.* 2021;9:90319–29. doi:10.1109/ACCESS.2021.3090474.
11. Puente-Castro A, Fernandez-Blanco E, Pazos A, Munteanu CR. Automatic assessment of Alzheimer's disease diagnosis based on deep learning techniques. *Comput Biol Med.* 2020 May;120:103764. doi:10.1016/j.combiomed.2020.103764.