Deep Learning with Predictive Uncertainty for Alzheimer's Disease Detection from MRI using EfficientNet

Ranga Sai Varun Talluri
School of Computing
Amrita Vishwa Vidyapeetham
Amaravati, India
av.sc.u4aie23141@av.students.amrita.edu

P V G P Tejasv
School of Computing
Amrita Vishwa Vidyapeetham
Amaravati, India
av.sc.u4aie23147@av.students.amrita.edu

Venkata Raghavendra Pranay Tangirala
School of Computing
Amrita Vishwa Vidyapeetham
Amaravati, India
av.sc.u4aie23142@av.students.amrita.edu

G Lavanya
School of Computing
Amrita Vishwa Vidyapeetham
Amaravati, India
g_lavanya@av.amrita.edu

V. M. K. Prasad Goura
School of Engineering
Amrita Vishwa Vidyapeetham
Amaravati, India
g_vmkprasad@av.amrita.edu

***Abstract*—Alzheimer's disease (AD) is a progressive neurodegenerative disease that causes memory loss and cognitive decline, and thus early detection is crucial to provide good care. We propose a deep-learning approach using EfficientNetB3 through transfer learning to classify brain scan MRI images into four stages: non-demented, very mild, mild, and moderate dementia. The dataset was completely processed, augmented, and balanced for training. Our EfficientNetB3 model attained a test accuracy of 96.9% (99.5% CI: 95.73%–97.88%) with high precision and F1-scores in all classes, outperforming several recent state-of-the-art methods. We additionally evaluated predictive uncertainty using predictive entropy and compared deterministic softmax entropy estimates with exploratory Monte Carlo (MC) Dropout. Per-sample predictions and uncertainty metrics were logged to CSV for reproducibility and further analysis. These results show that EfficientNet-based methods have potential for providing accurate, automated classification of AD stages to enable more rapid and accurate clinical decision-making.**

***Index Terms*—Alzheimer's Disease, Deep Learning, EfficientNet, Transfer Learning, MRI Classification, Early Diagnosis**

I. Introduction

The brain is the most complex organ in the human body. It controls the central nervous system and is responsible for mental processes, physical actions, learning, memory and emotional reactions. Anything that disrupts normal function can impact the functioning of the whole body. The hippocampus is found deep in the temporal lobe and is a critical part of the brain for memory formation, learning, and emotion processing. Its damage causes neurological and psychiatric disorders, such as epilepsy, Alzheimer's, and depression [1]. AD is the most frequently occurring form of dementia, and it brings about a gradual loss of memory and a decline in thinking and understanding in the individuals who suffer from it. The disorder is experienced as a total cognitive decline that has an impact on both mental skills and behavior. This disrupts normal daily activities [2]. In 2021, the global population of people with dementia reached 57 million people, including 60% who live in low- and middle-income countries. In every year, the world records about 10 million new cases. The World Health Organization (WHO) states that AD represents the primary dementia type. This makes up between 60 and 70 percent of all cases [3]. AD ranks as the world's second most severe brain disorder, which starts by damaging the hippocampus before it spreads to different brain areas, leading to neuron death and major brain tissue shrinkage. The disease shows progression through a permanent neurodegenerative process. This causes dementia that has no available treatment. Early detection and treatment enable patients to manage disease progression and its effects [4]. The main symptoms of AD appear as memory problems, communication difficulties, and decreased mental abilities for decision-making and thinking. The symptoms increase gradually and eventually obstruct the person from doing their daily activities. If the disease is detected at an early stage, it will be easier for the doctors to optimize the treatment methods [5]. In its early stages, AD first appears as moderate cognitive impairment (MCI), which develops into stable MCI, progressive MCI, and AD [6]. Medical imaging and computer-based methods have been the most reliable tools for detecting AD in its early stage [5]. The diagnosis of AD has been drastically enhanced by the adoption of advanced technology such as Machine Learning (ML), Deep Learning (DL) algorithms, and magnetic resonance imaging (MRI). Taking advantage of the web and digital technology has become an option to address these problems. One of the emerging areas of scientific research is the application of artificial intelligence (AI). AI is capable of recognizing objects and providing meaningful and efficient insights, such as classifying images to detect different disease types. This type of object recognition can

be achieved using DL techniques, which can automatically extract important features [7]. Recently, various ML and DL algorithms have been applied on MRI images to classify them, detect early AD, and for applications in computer vision and healthcare [4].

The main aim of this work is to develop a DL technique for an early-stage AD classification on MRI scan images. This technique uses a Convolutional Neural Network (CNN) to extract discriminative brain features and achieves strong classification accuracy with little manual effort. After preprocessing and analyzing the evaluation metrics, the proposed algorithm demonstrated strong performance in distinguishing between different AD stages, supporting more accurate diagnosis and early treatment.

II. Literature Review

A lot of research has been carried out and is still being carried out on the diagnosis and detection of AD, a condition that slowly damages brain cells over time. The detection of AD during its initial stages proves challenging because early symptoms show up as very mild signs. Early diagnosis is crucial, as timely treatment can help slow its progression, although there is currently no cure. ML and DL algorithms have shown importance in many fields such as sentiment analysis, speech enhancement, cybersecurity, image classification, and healthcare. So, the authors suggest that these techniques can also help address the challenge of early AD [8]. While several ML/DL approaches have demonstrated effectiveness for early Alzheimer's detection, Attur et al. [9] presented a Fuzzy Inference System (FIS) that better handles uncertainty and can classify disease stages (Normal, MCI, Mild, Moderate, Severe) with competitive accuracy on a modest Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. Sujatha et al. [10] presented an ML-based framework for AD detection. Early works using MRI-based analysis demonstrated that longitudinal MRI data from the Open Access Series of Imaging Studies (OASIS) are well-suited for evaluating traditional ML classifiers and identifying dementia-related patterns. Moving toward richer clinical modeling, Khan et al. [11] described a three-tiered hybrid framework that uses correlation-based feature grouping, sequential selection, and a two-level stacking strategy, enabling more accurate discrimination among cognitively normal (CN), MCI, and AD subjects. The growing importance of ensemble learning is evident in studies by Moodely et al. [12], Shah et al. [13], and Uddin et al. [14]. Their findings indicate that combining classifiers through hybrid or soft-voting approaches improves robustness and predictive accuracy when applied to important biomarkers including the Mini-Mental State Examination (MMSE), Clinical Dementia Rating (CDR), volumetric brain measures, and other clinical indicators. Similarly, Gani Lakshmi et al. [15] compared classical ML models with boosting techniques using clinical features extracted from the OASIS MRI dataset. Their results suggest that soft-voting ensembles perform better than individual models, mainly because they handle noise and small sample sizes more effectively. They also noted that weighted model averaging can serve as a lightweight and dependable approach for early-stage AD screening. Syed et al. [16] strengthened the ensemble approach by identifying a compact and discriminative Cerebrospinal Fluid (CSF) biomarker subset through Recursive Feature Elimination (RFE) and L1 regularization and employing a calibrated weighted ensemble of Support Vector Machine (SVM) and logistic regression that produced high sensitivity, precision, and AUC scores for early-stage screening. Broader modeling of disease progression was explored by El-Sappagh et al. [17], who integrated cognitive, MRI-derived, and neuropsychological features within a heterogeneous ensemble optimized through accuracy–diversity balancing, resulting in improved multi-stage AD prediction performance across a large ADNI population. At the same time, the solutions based on DL also become important because of their ability to automatically extract features from neuroimaging data. Jeyalakshmi et al. [18] applied EfficientNet models to capture structural features for the early AD detection. In a comprehensive study, Srividhya et al. [19] examined 26 pretrained Keras models over Structural Magnetic Resonance Imaging (sMRI) dataset and found out that ResNet-50v2 performed best for five-level AD staging, validating the advantage of transfer learning coupled with data-balancing techniques like Synthetic Minority Oversampling Technique (SMOTE). Sequential DL modeling was enhanced by Ebrahimi-Ghahnavieh et al. [20], with a CNN–Long Short-Term Memory (LSTM) hybrid model that extracted spatial and slice-level temporal features of MRI scans to surpass traditional CNN baselines. This verified that the combination of spatial and sequential features improve the performance of AD detection. Transfer learning has been a reliable approach in various studies, such as that of Khagi et al. [21], who evaluated AlexNet, GoogLeNet, ResNet50, and a custom CNN on OASIS MRI data and demonstrated that small, well-tuned models can obtain competitive accuracy. Salehi et al. [22] and Gurpreet Singh et al. [23] have similarly shown deep CNNs are able to identify early structural brain abnormalities for strong prospects of reliable automated diagnosis. Complementing these architectures, Baskar et al. [24] compared DenseNet-169 and ResNet-50 for multi-stage dementia classification and reaffirmed the advantage of denser connectivity patterns, with DenseNet-169 achieving the highest testing accuracy among the evaluated models.

III. Methodology

This section describes the methodology for classifying brain MRI scans from the OASIS dataset into 4 classes. The entire process contains data preparation, preprocessing, augmentation, and model architectures that use EfficientNet via transfer learning. Training strategies, evaluation metrics, and uncertainty estimation techniques are all explained to ensure reproducibility and reliable performance assessment is in place.

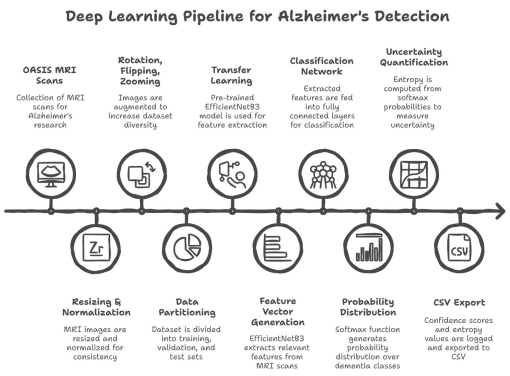
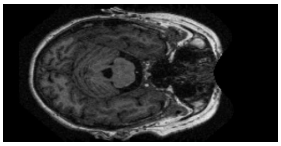


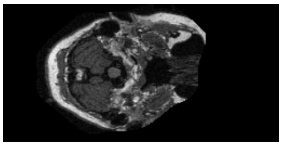
Fig. 1: Deep learning pipeline for dementia classification

A. Dataset Description

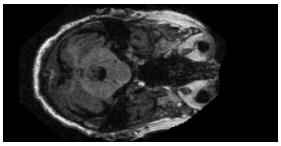
The dataset used in this study is the OASIS MRI dataset, available publicly through the Kaggle platform. The dataset contains enough data variety and quality standards to build and test models for AD classification.



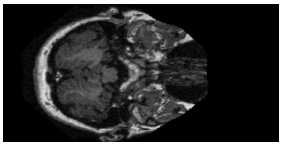
(a) Non Demented



(b) Very Mild Demented



(c) Mild Demented



(d) Moderate Demented

Fig. 2: Sample MRI Brain Scans from the Alzheimer's Disease Dataset

B. Data Preprocessing and partitioning

Two different dataset splitting methods were implemented for the two models. The first model used stratified random sampling to select samples from each class independently, which resulted in balanced train, validation, and test sets according to the 80–10–10 split. To ensure reproducibility, we fixed the random seed to 42 for Python's random module, NumPy, and TensorFlow before sampling and splitting. The second model used a fixed subset of 15,000 images, which were divided into training, validation, and test sets based on the 80–10–10 ratio to produce 12,000 training images, 1,500 validation images, and 1,500 test images.

All the MRI images were resized to meet the input size requirement of each EfficientNet model by adjusting them to 300×300 pixels for EfficientNet-B3 and 224×224 pixels for EfficientNet-B0. The identical preprocessing method has been used for both models, and the pixel values have been normalized to the [0, 1] interval by dividing by 255.

C. Data Augmentation Strategy

o improve generalization, data augmentation was applied only to the training sets. The first model used geometric and photometric transformations including random rotations (±25°), translations (±20%), shear, zoom, horizontal flipping, and brightness adjustment. The second model employed a similar but slightly stronger augmentation strategy with increased rotation and shear ranges to enhance robustness.

D. Model Architecture

Both models used transfer learning with pretrained EfficientNet architectures. EfficientNetB3 was chosen for the first experiment due to its balanced trade-off between model complexity and computational efficiency, using ImageNet pretrained weights for effective feature extraction on medical images. Experiment 2 was conducted with EfficientNetB0, a much lighter model with less parameters (5.3M vs. 12M in B3) that offers faster training and inference while still providing strong classification power. Both models were initialized with ImageNet weights to fine-tune pretrained features for the Alzheimer's MRI classification.

The proposed architectures have the same structure, consisting of a frozen EfficientNet backbone, a Global Average Pooling 2D layer, a dense layer containing 512 ReLU units, a dropout layer with a rate of 0.4, and a softmax output layer for the four-class dementia classification. This replaces the default ImageNet head with a task-specific classifier for MRI-based medical imaging. Method 1 employs EfficientNetB3 with the size of 300×300×3 as input. During the initial training phase, all convolutional layers remain frozen to preserve pretrained ImageNet features, while the custom classification head learns domain-specific patterns. Dropout regularization helps control overfitting and improves generalization. Method 2 employs EfficientNetB0 with input dimensions of 224×224×3 while retaining the same classification head design. Its lighter architecture provides faster training and inference while maintaining strong representational capacity for the dementia classification task. calibrated, and a complete Bayesian evaluation is left for future work.

E. Class Imbalance Handling

To address class imbalance in the dataset, class weights were calculated using the balanced class weight method, where weights are inversely proportional to class frequencies in the training data.

F. Training Strategy

1) Two-Stage Training Protocol: Both models were trained in two phases to adapt pretrained EfficientNet features for medical image classification. EfficientNetB3 was trained for 10 epochs with the backbone frozen (learning rate 5×10^{-4}), followed by 45 fine-tuning epochs using a lower learning rate (1×10^{-5}). EfficientNetB0 underwent 5 epochs of frozen-backbone training and 10 epochs of fine-tuning with the same learning rates. Mixed-precision training, early stopping, learning-rate scheduling, and model checkpointing were used to improve efficiency and prevent overfitting.

2) Training Optimization Techniques: Both models employed training callbacks to improve efficiency and prevent overfitting. Early stopping monitored validation accuracy and restored the best weights after a patience of seven epochs for Model 1 and five epochs for Model 2. ReduceLROnPlateau reduced the learning rate by a factor of 0.3 when validation loss plateaued for three epochs. Model checkpointing was used to save the weights, achieving the highest validation accuracy. For Model 1, a higher learning rate (5×10^{-4}) was used during frozen-backbone training to quickly adapt the new classification layers. During fine-tuning in Model 2, the backbone was unfrozen and a lower learning rate (1×10^{-5}) was applied to update weights more conservatively, ensuring stable convergence and improved generalization.

3) Mixed Precision Training: Mixed precision training was used to accelerate computation while maintaining numerical stability. The `mixed_float16` policy performs most operations in 16-bit floating-point precision and reserves 32-bit precision for numerically sensitive calculations. This reduces the amount of memory used and increases the throughput of the GPU, making training faster and more efficient.

4) Hyperparameters: The hyperparameters were chosen to maintain stable optimization and fast training. EfficientNetB3 was trained with a batch size of 16, and EfficientNetB0 was assigned a batch size of 32 to fit within the available GPU memory. Both models were trained with the Adam optimizer using learning rates of 5×10^{-4} for frozen-backbone training and 1×10^{-5} for fine-tuning. Categorical cross-entropy was adopted as the loss function, and a dropout rate of 0.4 was applied in the classification head to improve generalization.

G. Model Evaluation

1) Performance Metrics: To assess the model's performance, we used an independent test set and tracked both accuracy and categorical cross-entropy loss. The checkpoint that achieved the highest validation accuracy was used for final evaluation. Accuracy measures the proportion of correctly classified samples out of the total number of samples evaluated.

$$\text{Accuracy}_{\text{multi-class}} = \frac{\sum_{i=1}^{C} TP_i}{\sum_{i=1}^{C} \big(TP_i + FP_i + FN_i\big)}, \quad (1)$$

The categorical cross-entropy loss L quantifies the divergence between true class distribution and the predicted probability distribution of the model.

$$L = -\sum_{i=1}^{C} y_i \log(p_i), \quad (2)$$

In this equation:

- C represents the total number of classes,
- y_i is the ground-truth label for class i, encoded as a one-hot vector (i.e., $y_i = 1$ only for the correct class and 0 otherwise),
- p_i denotes the predicted probability assigned by the model to class i.

Along with accuracy and categorical cross-entropy loss, we report precision, recall, and F1-score for each class to better understand how well the model distinguishes between the various dementia classes. For a given class $i \in \{1, \ldots, C\}$, these metrics are defined in [17].

For multi-class evaluation, macro-averaged and weighted-averaged scores are also reported. The macro-averaged F1-score is given by:

$$\text{F1}_{\text{macro}} = \frac{1}{C} \sum_{i=1}^{C} \text{F1}_i, \quad (3)$$

and the weighted F1-score is computed as:

$$\text{F1}_{\text{weighted}} = \frac{1}{N} \sum_{i=1}^{C} n_i \, \text{F1}_i, \quad (4)$$

where n_i is the support (number of true samples) of class i and $N = \sum_{i=1}^{C} n_i$ is the total number of test samples.

Here, TP_i (true positives) denotes the number of samples belonging to class i that are correctly predicted as class i, FP_i (false positives) denotes the number of samples from other classes that are incorrectly predicted as class i, and FN_i (false negatives) denotes the number of samples from class i that are incorrectly predicted as another class. C is the total number of classes in the dementia staging task.

2) Statistical Confidence Analysis: Prediction uncertainty was quantified using two-sided 99.5% confidence intervals computed via the binomial proportion test. Overall accuracy was treated as the observed success rate across independent test predictions, and confidence intervals were estimated using SciPy's Clopper-Pearson exact method, providing a conservative yet statistically reliable measure of performance uncertainty.

H. Uncertainty Quantification and Result Logging

Along with point estimates of standard metrics (accuracy, precision, recall, and F1-score), we estimated predictive uncertainty for the EfficientNetB3 model using *predictive entropy* derived from the model's softmax output. For a single image with predicted class probabilities $p = [p_1, p_2, \ldots, p_C]$, the predictive entropy $H(p)$ is defined as:

$$H(p) = -\sum_{i=1}^{C} p_i \log(p_i), \quad (5)$$

where the natural logarithm is used and the entropy values obtained are in nats. Predictive entropy, at a deterministic forward pass, captures the model's full uncertainty (aleatoric and epistemic, as represented in the softmax distribution).

For reproducibility, per-sample results were stored in a CSV file. Method 1 columns: filename, true_class, predicted_class, confidence, entropy, ci_low, ci_high, prob_NonDemented, prob_VeryMild, prob_Mild, prob_Moderate.

For method 2, multiple stochastic forward passes were performed at test time to obtain an uncertainty estimate of the model by means of preliminary MC Dropout runs. These

exploratory runs were limited by the time needed for their completion and yet gave valuable variance-based uncertainty measures.

Method 2 columns: uncertainty (std. dev. of predictions) and class-wise average probabilities from MC Dropout.

IV. Results & Discussion

In this section, we present a comprehensive analysis of our proposed AD detection model based on several performance metrics. Accuracy, precision, recall, and F1-score are computed to assess the model performance for classifying the 4 cases. To handle the class distribution changes and minimize classification errors for different dementia stages, the confusion matrix and classification report were studied. Together these measures provide a complete evaluation of the diagnostic accuracy of the model and its performance in relation to each individual category.

TABLE I: Classification report of the proposed Alzheimer's disease detection model

Class	Precision	Recall	F1-Score	Support
Mild Dementia	0.975	1.000	0.987	501
Moderate Dementia	0.961	1.000	0.980	49
Non Demented	0.975	0.942	0.959	800
Very Mild Dementia	0.961	0.975	0.968	800
Accuracy			**0.969**	
Macro Avg	0.968	0.979	0.973	2150
Weighted Avg	0.969	0.969	0.969	2150

Table 1 summarizes the performance of the model. The test accuracy of 96.9%, while the training accuracy was 96.43%, indicating effective learning without overfitting. The system scored perfect recall rates for the classes of mild and moderate dementia, with non-demented and very mild dementia also showing consistent results. The high macro and weighted F1 scores further indicate that the model performs consistently across all four dementia stages. Method 1 used entropy-based predictive uncertainty, showing low entropy for correct predictions and higher uncertainty for misclassifications. All per-sample results were saved into a CSV file to perform detailed analysis, and initial runs of MC-Dropout (Method 2) also yielded variance-based uncertainty information.

Method 2 (MC-Dropout) achieved a test accuracy of 75.78% and a loss of 0.7554. Although this is lower than the performance of the main EfficientNetB3 model, the MC-Dropout runs still produced reasonably stable predictions across passes. This suggests that, even with the reduced accuracy, the model's behaviour under dropout was fairly consistent.

TABLE II: Comparison of Model Performance with Existing Studies

Author	Performance
Shagun Sharma [4]	Accuracy: 90.40%
Nivedhitha Mahendran [5]	Accuracy: 88.70%
Gurpreet Singh [23]	Accuracy: 91.63%
S. Baskar [14]	Accuracy: 87.36%
Shruti Pallawi [25]	Accuracy: 95.78%
Proposed Method	**Accuracy: 96.9%**

As shown in Table 2, the proposed algorithm reached 96.9% accuracy, which is superior to other work, reporting accuracies ranging from 87.36% to 95.78%. The results say that our model performs better than current state-of-the-art techniques reported in the literature.

Fig. 3: Training and validation accuracy across epochs.

The model shows steady improvement and stable convergence after approximately 40 epochs, indicating effective feature learning with minimal overfitting.

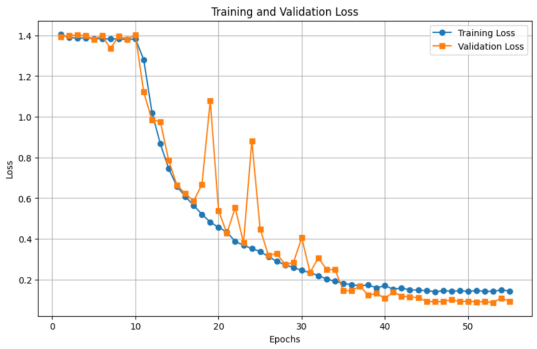


Fig. 4: Training and validation loss across epochs.

The consistent decline and convergence of loss values indicate effective error minimization and stable generalization throughout training.

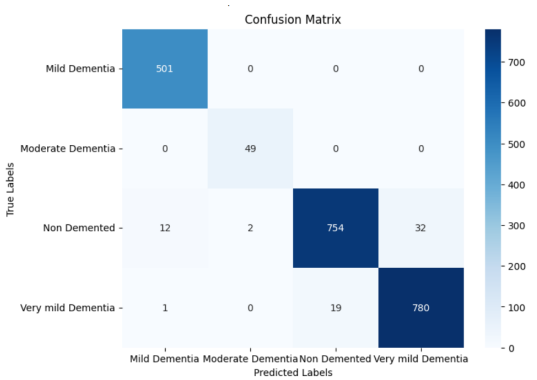


Fig. 5: Confusion matrix illustrating classification performance on the test dataset.

The matrix shows high true positive rates across all dementia stages with minimal misclassifications, indicating reliable and accurate differentiation between AD stages.

V. Conclusion

In this study, we explored the use of DL for early Alzheimer's detection. The model achieved a high classification accuracy of 96.9% with balanced F1-scores for each stage of dementia through the EfficientNetB3 model utilizing MRI brain scan images. The model achieved consistent and robust results for all classes due to appropriate data preprocessing, augmentation, and evaluation protocols, which allow it to be employed in clinical and real-world diagnosis. We applied uncertainty quantification to assess the reliability of the system in addition to the accuracy analysis. Where the predictive entropy function is a computationally simple method to identify uncertainty in predictions. The initial MC-Dropout experiments revealed new information about model uncertainty, but the research remained exploratory because of restricted computational resources. In the future, this work can be extended by generalizing uncertainty modeling, using 3D MRI data, explainable AI techniques, and accessing larger multicenter datasets, which will enhance transparency, scalability, and the reliability of Alzheimer's detection.

References

[1] A. Ashraf, S. Naz, S. H. Shirazi, I. Razzak, M. Parsad, and M. Parsad, "Deep transfer learning for Alzheimer neurological disorder detection," *Multimedia Tools and Applications*, vol. 80, no. 20, pp. 30117–30142, 2021, doi: 10.1007/s11042-020-10331-8.

[2] National Institute on Aging, "What is Alzheimer's disease?" 2025. [Online]. Available:https://www.nia.nih.gov/health/alzheimers-and-dementia/what-alzheimers-disease.

[3] World Health Organization, "Dementia," 2025. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/dementia

[4] S. Sharma, K. Guleria, S. Tiwari, and S. Kumar, "A deep learning-based convolutional neural network model with VGG16 feature extractor for the detection of Alzheimer disease using MRI scans," *Measurement: Sensors*, vol. 24, p. 100506, 2022, doi: 10.1016/j.measen.2022.100506.

[5] N. Mahendran and P. M. D. R. Vincent, "A deep learning framework with an embedded-based feature selection approach for the early detection of Alzheimer's disease," *Computers in Biology and Medicine*, vol. 141, p. 105056, 2022, doi: 10.1016/j.compbiomed.2021.105056.

[6] S. Wang, H. Wang, Y. Shen and X. Wang, "Automatic Recognition of Mild Cognitive Impairment and Alzheimers Disease Using Ensemble based 3D Densely Connected Convolutional Networks," *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Orlando, FL, USA, 2018, pp. 517-523, doi: 10.1109/ICMLA.2018.00083.

[7] A. Nareswari and D. Utari, "Advancements in Alzheimer's diagnosis through MRI using Bayesian convolutional neural networks and variational inference," *BAREKENG: Journal of Mathematics and Its Applications*, vol. 18, no. 4, pp. 2423–2434, 2024.

[8] K. Dashtipour et al., "Detecting Alzheimer's disease using machine learning methods," in *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol. 420, Springer, 2022, pp. 89–100, doi: 10.1007/978-3-030-95593-9_8.

[9] S. S. Attur and M. E. Harikumar, "Detection of Alzheimer's Disease using Fuzzy Inference System," *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Tirunelveli, India, 2022, pp. 1235-1241, doi: 10.1109/ICSSIT53264.2022.9716523.

[10] E. Sujatha and M. Subiksha, "Predictive Precision: A Comparative Study of ML in Alzheimer's Diagnosis," *2024 1st International Conference on Trends in Engineering Systems and Technologies (ICTEST)*, Kochi, India, 2024, pp. 01-05, doi: 10.1109/ICTEST60614.2024.10576101.

[11] A. Khan and S. Zubair, "Development of a three tiered cognitive hybrid machine learning algorithm for effective diagnosis of Alzheimer's disease," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 10, pp. 8000–8018, Jul. 2022, doi: 10.1016/j.jksuci.2022.07.016.

[12] G. V. Moodely, S. M. I. H. Amiree, M. Thazin, R. R N, M. Habib and S. Kumar Singh, "A Comparative Study on Hybrid Machine Learning Voting Classifier Models for Alzheimer's Disease Prediction," *2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI)*, Raipur, India, 2023, pp. 1-6, doi: 10.1109/ICAIIHI57871.2023.10489772.

[13] A. Shah, D. Lalakiya, S. Desai, Shreya and V. Patel, "Early Detection of Alzheimer's Disease Using Various Machine Learning Techniques: A Comparative Study," *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)*, Tirunelveli, India, 2020, pp. 522-526, doi: 10.1109/ICOEI48184.2020.9142975.

[14] K. M. M. Uddin, M. J. Alam, N. Jannat-E-Anawar, M. A. Uddin, and S. Aryal, "A novel approach utilizing machine learning for the early diagnosis of Alzheimer's disease," *Biomedical Materials & Devices*, vol. 1, no. 2, pp. 882–898, Apr. 2023, doi: 10.1007/s44174-023-00078-9.

[15] S. G. Lakshmi, C. Kumar, and P. Kumar, "Progressive Early Alzheimer's diagnosis: Leveraging hard and soft ensemble voting techniques in machine learning," *Procedia Computer Science*, vol. 258, pp. 1899–1908, Jan. 2025, doi: 10.1016/j.procs.2025.04.441.

[16] A. H. Syed, T. Khan, A. Hassan, N. A. Alromema, M. Binsawad and A. O. Alsayed, "An Ensemble-Learning Based Application to Predict the Earlier Stages of Alzheimer's Disease (AD)," in *IEEE Access*, vol. 8, pp. 222126-222143, 2020, doi: 10.1109/ACCESS.2020.3043715.

[17] S. El-Sappagh, F. Ali, T. Abuhmed, J. Singh, and J. M. Alonso, "Automatic detection of Alzheimer's disease progression: An efficient information fusion approach with heterogeneous ensemble classifiers," *Neurocomputing*, vol. 512, pp. 203–224, Sep. 2022, doi: 10.1016/j.neucom.2022.09.009.

[18] J. Jeyalakshmi, J. Lokesh Kumar and S. Vishnu, "Neuroimaging Analysis for Early Diagnosis of ADNI Progression with EfficientNet," *2024 10th International Conference on Communication and Signal Processing (ICCSP)*, Melmaruvathur, India, 2024, pp. 1555-1560, doi: 10.1109/ICCSP60870.2024.10543702.

[19] L. Srividhya, S. Vishvanathan, V. Ravi, E. A. Gopalakrishnan, and K. P. Soman, "Deep learning-based approach for multi-stage diagnosis of Alzheimer's disease," *Multimedia Tools and Applications*, vol. 83, 2023, doi: 10.1007/s11042-023-16026-0.

[20] A. Ebrahimi-Ghahnavieh, S. Luo and R. Chiong, "Transfer Learning for Alzheimer's Disease Detection on MRI Images," *2019 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*, Bali, Indonesia, 2019, pp. 133-138, doi: 10.1109/ICIAICT.2019.8784845.

[21] B. Khagi, B. Lee, J. -Y. Pyun and G. -R. Kwon, "CNN Models Performance Analysis on MRI images of OASIS dataset for distinction between Healthy and Alzheimer's patient," *2019 International Conference on Electronics, Information, and Communication (ICEIC)*, Auckland, New Zealand, 2019, pp. 1-4, doi: 10.23919/ELINFOCOM.2019.8706339.

[22] A. W. Salehi, P. Baglat, B. B. Sharma, G. Gupta and A. Upadhya, "A CNN Model: Earlier Diagnosis and Classification of Alzheimer Disease using MRI," *2020 International Conference on Smart Electronics and Communication (ICOSEC)*, Trichy, India, 2020, pp. 156-161, doi: 10.1109/ICOSEC49089.2020.9215402.

[23] G. Singh, K. Guleria and S. Sharma, "A Deep Learning-based Convolutional Neural Network Model for Alzheimer's Disease Detection," *2024 Second International Conference Computational and Characterization Techniques in Engineering & Sciences (IC3TES)*, Lucknow, India, 2024, pp. 1-5, doi: 10.1109/IC3TES62412.2024.10877583.

[24] S. Baskar, M. L. Prasad, N. Sharma, N. I, T. Katale and P. C. S. Reddy, "An Accurate Prediction and Diagnosis of Alzheimer's Disease using Deep Learning," *2023 IEEE North Karnataka Subsection Flagship International Conference (NKCon)*, Belagavi, India, 2023, pp. 1-7, doi: 10.1109/NKCon59507.2023.10396132.

[25] S. Pallawi and D. K. Singh, "Detection of Alzheimer's Disease Stages Using Pre-Trained Deep Learning Approaches," *2023 IEEE 5th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA)*, Hamburg, Germany, 2023, pp. 252-256, doi: 10.1109/ICCCMLA58983.2023.10346730.