MULTIDISCIPLINARY RESEARCH INSTITUTE INDIA

International Journal of Advanced Scientific Research and Engineering Trends

ISSN: 2456-0774

Volume 10 Issue 6 (2026) | Pages: 42-51

Alzheimer's Disease Classification Using Convolutional Neural Network-Random Forest Hybrid Architecture

Tanish Gumaste[1*], Chetan Aher[2]

[1,2]Department of Computer Engineering
AISSMS Institute of Information Technology (IOIT), Pune, Maharashtra, India
[1]tanishgumaste@gmail.com, [2]chetan.aher@aissmsioit.org

Peer Review Information

***Type**: Article*
***Received**: 08 April 2026*
***Revised**: 27 April 2026*
***Accepted**: 05 June 2026*
***Published**: 26 June 2026*

Abstract

Worldwide, healthcare systems bear a heavy financial burden due to the progressive neurological illness known as Alzheimer's disease (AD), which impacts around 55 million people. In order to enable proper clinical action at the appropriate moment, corrective identification of the disease's stages is a crucial task. This paper compares two deep learning-based frameworks for classifying Alzheimer's disease from MRI brain scans: (1) a standalone four-block Convolutional Neural Network (CNN) with batch normalization, global average pooling, and dropout regularization; and (2) a hybrid approach that combines a Random Forest (RF) ensemble for classification of its 256-dimensional embeddings with a CNN backbone as a learned feature extractor. A three-class MRI dataset of samples—mildly demented, moderately demented, and non-demented—is used to train and evaluate both models. The dataset is divided into train, validation, and test sets in a 70:15:15 ratio after being preprocessed into 128x128 pixels. There is a 70:15:15 split between the train, validation, and test sets, and the dataset is pre-processed to 128x128 pixels. It is then loaded in parallel. Among the measures used for evaluation are accuracy, weighted precision, weighted recall, weighted F1-score, per-class specificity, false negative rate, false discovery rate, false positive rate, and negative predictive value. Additionally, there is the Matthews Correlation Coefficient (MCC). Experimental results reveal that the CNN+RF hybrid outperforms the CNN and demonstrates the usefulness of combining deep feature representation with ensemble learning for neuroimaging diagnosis.

Keywords: Alzheimer’s Disease Classification; Convolutional Neural Network; Random Forest; Hybrid Deep Learning; MRI Brain Image Analysis; Medical Image Classification; Dementia Detection; Neuroimaging

How to Cite This Article

Gumaste, T., & Aher, C. (2026). Alzheimer’s disease classification using convolutional neural network–random forest hybrid architecture. *International Journal of Advanced Scientific Research and Engineering Trends*, 10(6), 42–51.

Introduction

The most common form of dementia, Alzheimer's disease (AD), is a major concern in public health around the world. Alzheimer's disease accounts for roughly 60–70% of the estimated 55 million dementia cases globally, according to the WHO. The number will rise dramatically as the world's population ages, and early illness identification and staging are urgently needed in clinical and computational research (World Health Organization, 2023). There are multiple stages of severity, and magnetic resonance imaging (MRI) scans show the disease's characteristic anatomical changes in the brain. Early identification of these phases slows the illness's course and allows for more targeted therapy.

Clinical assessment, cognitive testing, and manual processing of MRI data constitute the backbone of conventional diagnostic procedures; nevertheless, these approaches are laborious, subjective, and dependent on the accessibility of medical experts. Thanks to deep learning—and convolutional neural networks in particular—feature extraction and classification from raw picture data has become more accurate and automated, ushering in a new era of medical image analysis. In a number of neuroimage applications, including brain tumor segmentation, stroke lesion identification, and neurodegenerative disease classification, CNNs have demonstrated remarkable performance (LeCun et al., 2015). They have been shown to be particularly useful for learning hierarchical spatial representations from large image datasets, without hand-engineered features, and are well suited to the subtle texture and morphological changes of Alzheimer's progression.

While powerful, CNNs can suffer from overfitting on small-to-medium sized medical databases, and are not always generalizable to unseen patient data. A combination of feature learning and classification can be considered an attractive solution. By training a CNN backbone to extract compact, discriminative feature embeddings and subsequently applying classical ensemble classifiers such as Random Forest, these pipelines benefit from the representation power of DL while gaining the robustness and interpretability of ensemble methods (Breiman, 2001). Random Forest classifiers are particularly attractive in this context due to their resistance to overfitting, ability to handle class imbalance via balanced weighting, and independence from assumptions about feature distributions.

This paper makes the following contributions: (i) a thorough implementation and evaluation of a standalone four-block CNN for three-class Alzheimer's MRI classification; (ii) a novel hybrid CNN+RF architecture in which CNN-extracted 256-dimensional embeddings are classified by a 500-tree Random Forest; (iii) an extensive multi-metric evaluation including MCC, per-class specificity, FNR, FDR, FPR, and NPV — metrics rarely reported together in prior Alzheimer's classification studies; and (iv) a comparative analysis that quantifies the gain from hybrid ensemble classification over the standalone DL baseline. What follows is an outline of the rest of the paper: Part II provides a literature review of relevant works; The technique is described in Section III, the experimental results and analysis are presented in Section IV, and Section V closes with recommendations for further research.

Related Work

Over the past ten years, there has been a lot of scientific interest in the automated analysis of MRI scans for AD identification and staging. Early methods depended on manually created features, such as gray matter density retrieved from structural MRI, cortical thickness measures, and voxel-based morphometry, which were then categorized using logistic regression SVM. These were useful for binary (dementia vs control) classification but did not work well for multi-class stages because of the small differences between classes (mild vs moderate dementia).

By applying a 2D convolutional neural network (CNN) architecture to MRI patches of the hippocampus, Aderghal et al. (2017) demonstrated promising outcomes in identifying Alzheimer's patients as healthy controls. In order to detect atrophy-specific characteristics, the study stressed the importance of region-specific, rather than whole-brain, brain analysis. According to a comprehensive evaluation of all DL methods for AD detection conducted by Ebrahimighahnavieh et al. (2020), convolutional neural networks (CNNs) continuously outperformed more conventional machine learning approaches, with the added benefit of 3D volumetric CNNs providing spatial context at the cost of increased computing time.

Using MRI features, demographic information, and cognitive data from the ADNI database, Spasov et al. (2019) created a parameter-efficient 3D CNN for multi-modal prediction of conversion to AD. The significance of multi-modal fusion is demonstrated by their model, which had an AUC of 0.925 in predicting the change from moderate cognitive impairment (MCI) to AD. Similar to Basaia et al. (2019), we classified whole-brain structural MRI volumes from the ADNI cohort using a deep CNN and shown that completely automated deep learning classification could match or surpass conventional biomarker-based classification.

In addition to end-to-end deep networks, hybrid architectures that extract features from CNN and then classify them using classical classifiers have proven to be popular. To boost performance on small biomedical datasets, Loddo et al. (2021) suggested a pipeline where CNN-embeddings were fed to an ensemble of Random Forest classifiers for the classification task, showing that ensemble classifiers using CNN-features extracted from retinal fundus images performed better than fully connected CNN heads. This discovery is directly motivating the design CNN+RF that is used in the present work.

To further improve CNN architectures for AD classification, attention mechanisms have also been integrated. Jain et al., (2019) proposed a dual-branch attention network which works on both subcortical and cortical MRI regions, and fused their attention weighted feature maps before classification. The model was reported as having good sensitivity for detecting mild dementia over non-attentive baseline. In recent years, Transformer-based models have started to be used in neuroimage classification. By modifying the Vision Transformer (ViT) architecture, Liu et al. (2022) created an AD classification model based on 3D MRI images. This model achieved competitive performance and has the advantage of a global receptive field, which is not achievable for CNNs with local kernels.

Transfer learning from ImageNet-pretrained networks has been widely applied to Alzheimer's MRI classification, given the scarcity of labeled neuroimaging data. Farooq et al. (2017) fine-tuned AlexNet, GoogLeNet, and VGGNet on AD MRI data and reported that deeper pretrained networks achieved higher accuracy but required careful learning rate scheduling to prevent catastrophic forgetting. This was expanded upon by Nawaz et al. (2021), who achieved state-of-the-art results on the AD neuroimaging dataset by performing multi-scale feature extraction using three pretrained CNNs in parallel, fusing their outputs before a final softmax classifier.

The generalization of the model in the categorization of AD tasks has been shown to be significantly impacted by a number of data augmentation techniques. Using patch-level augmentation to compensate for limited training data, In order to diagnose AD and forecast the conversion of MCI at the same time, Lian et al. (2020) created a hierarchical fully convolutional network with a multi-task learning objective. To address the severe classification misalignment in multi-class MRI datasets related to dementia severity, Islam and Zhang (2018) employed rotational augmentation in conjunction with class weighted loss functions. They consistently performed better when it came to recalling the minority classes.

An important issue in the clinical deployment of AD classification models is explainability. Bhagwat et al. (2019) created saliency maps of brain regions linked to disease by combining their CNN-based classifier with the so-called gradient-weighted class activation mapping (Grad-CAM) to aid radiologists in understanding their diagnosis. More recently, the application of SHAP-based feature attribution to ensemble classifiers acting on CNN embeddings has allowed to explain classification decisions per patient. CNN+Random Forest pipelines were shown to be especially suitable for SHAP analysis by Feng et al. (2020) because the tree-based classifier can easily compute Shapley values exactly. The present study extends them by proposing an efficient CNN+RF architecture that is evaluated using a comprehensive multi-metric protocol and per-class diagnostic reporting.

Proposed Methodology

Two classification pipelines, one using a solo CNN and the other a combination of CNN and Random Forest, are tested in parallel on the same dataset in the proposed system. Figure 1 shows the whole workflow.

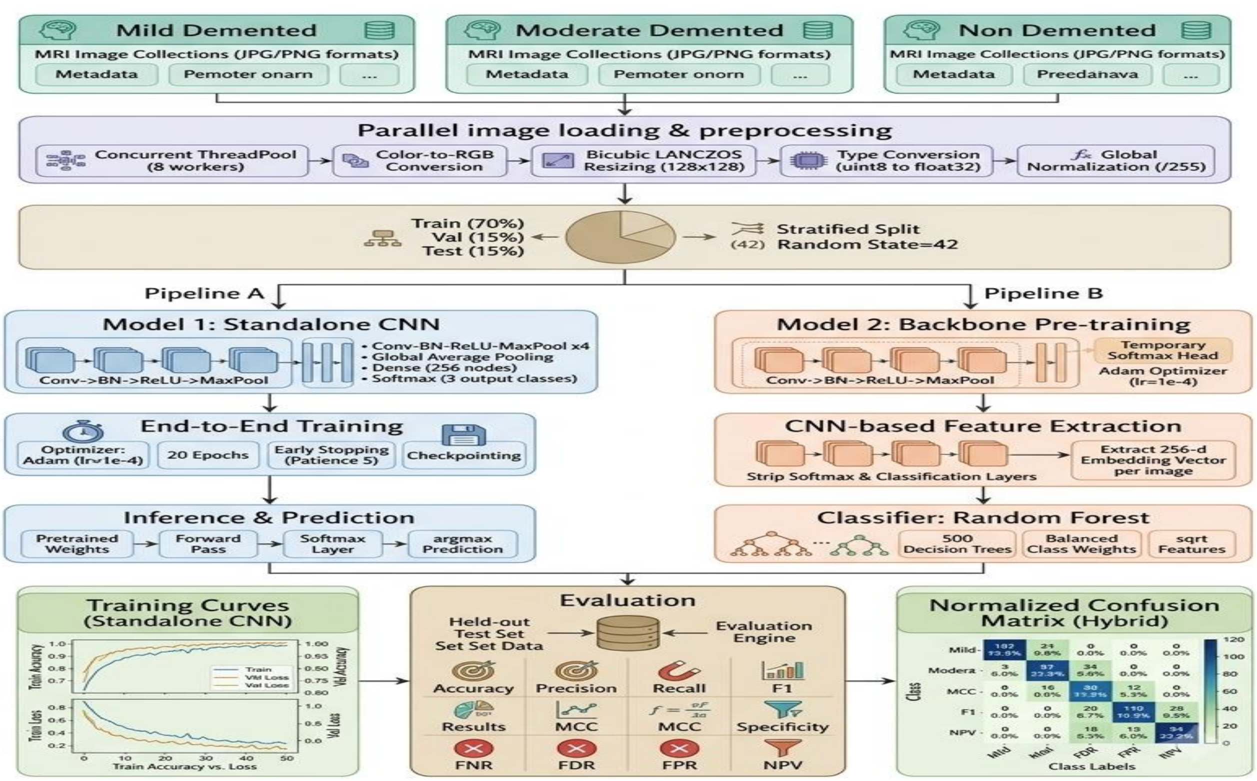

Fig. 1. Overall proposed system architecture for Alzheimer's disease classification.

Dataset and Pre-processing

Three subdirectories—Mild Demented, Moderate Demented, and Non-Demented—contain the MRI brain scan images that make up the data set. The images are saved in standard image formats and loaded into Google Colab from Google Drive. The data samples are presented in Figure2. The parallelized loading pipeline uses eight concurrent workers (with Python's ThreadPoolExecutor) to read images, convert them to RGB color space, and resize them to a common size of 128x128 (with PIL's ImageOps.fit using a LANCZOS resampling filter). A single vectorised normalization operation converts the entire data array to float32 and adjusts the pixel intensities to the range of [0, 1] after the loading phase saves memory by storing images as uint8 arrays. Bad or invalid pictures are ignored without warning to avoid pipeline disruption. The final data set comes in the form of a NumPy array of shape (N, 128, 128, 3) and integer class label vectors.

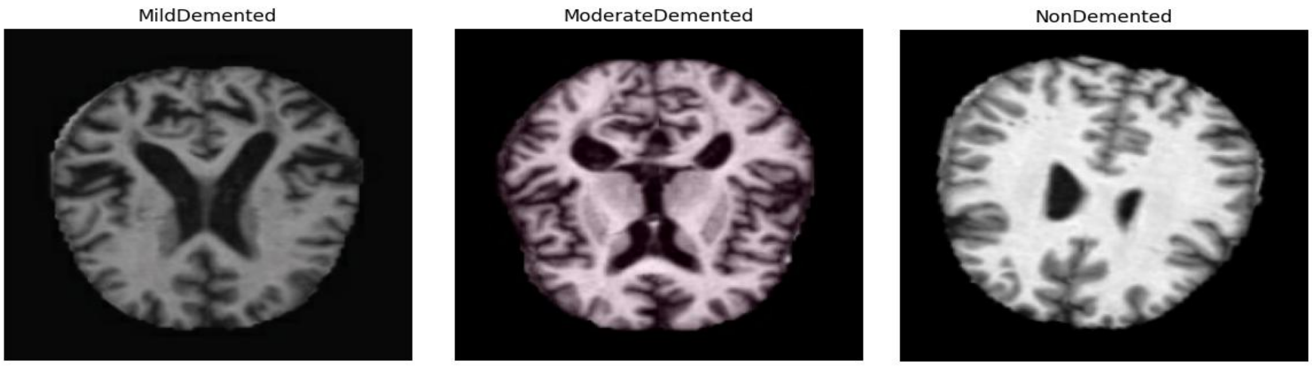

***Fig.2.** MRI image dataset sample.*

Data Splitting

The preprocessed data is gradually separated into three mutually exclusive groups: training, validation, and testing sets, using scikit-learn's train_test_split function. 15% of the entire data set is maintained as a "held out" test set in the first split. The remaining 85% of the data is separated into a validation set (15% of the data) and a training set (approximately 70% of the data). Regardless of whether the sample subset is mildly demented, moderately demented, or non-demented, the stratified parameter is employed at both stages to guarantee that there is a proportionate sample of each state between the three sample subsets. This is especially critical if there is class imbalance across severity groups. To ensure the reproducibility of the results, a fixed random state of 42 is used throughout. The split configuration is summarized in Table 1.

***Table 1.** Dataset Split Configuration*

Subset	Percentage	Purpose
Training Set	70%	Model weight optimization
Validation Set	15%	Early stopping / checkpoint
Test Set	15%	Final unbiased evaluation

CNN Architecture

The CNN model is built in the form of a sequential multi-class medical image classification architecture, consisting of four blocks [17]. A Conv2D layer with a 3x3 kernel and the same padding layer with a ReLU activation function make up the convolutional block's architecture. Batch Normalization and a Max Pooling layer with a 2x2 window come next. Within each block (32, 64, 128, 256), the filters' depth steadily increases as the network learns deeper and more discriminative feature representations, ranging from low-level edge and texture to higher-level structure. The GlobalAveragePooling2D layer, which comes after the four convolutional blocks, flattens the feature maps across space rather than time to lessen the chance of overfitting. Prior to the final Dense output layer, which uses Softmax activation to generate class probability estimates, there is a Dropout layer with a rate of 0.4 for regularization. Following that, there is a fully linked Dense layer with 256 units and ReLU activation. Figure 3 shows a simplified version of the architecture, and Table II describes it in detail.

***Table 2.** CNN Architecture Summary*

Block	Layer	Configuration	Output Shape	Parameters
1	Conv2D + BN + MaxPool	32 filters, 3x3, ReLU	64x64x32	896 + BN
2	Conv2D + BN + MaxPool	64 filters, 3x3, ReLU	32x32x64	18,496 + BN

3	Conv2D + BN + MaxPool	128 filters, 3x3, ReLU	16x16x128	73,856 + BN
4	Conv2D + BN + MaxPool	256 filters, 3x3, ReLU	8x8x256	295,168 + BN
5	GlobalAveragePooling2D	-	256	0
6	Dense + Dropout(0.4)	256 units, ReLU	256	65,792
7	Dense (Output)	3 units, Softmax	3	771

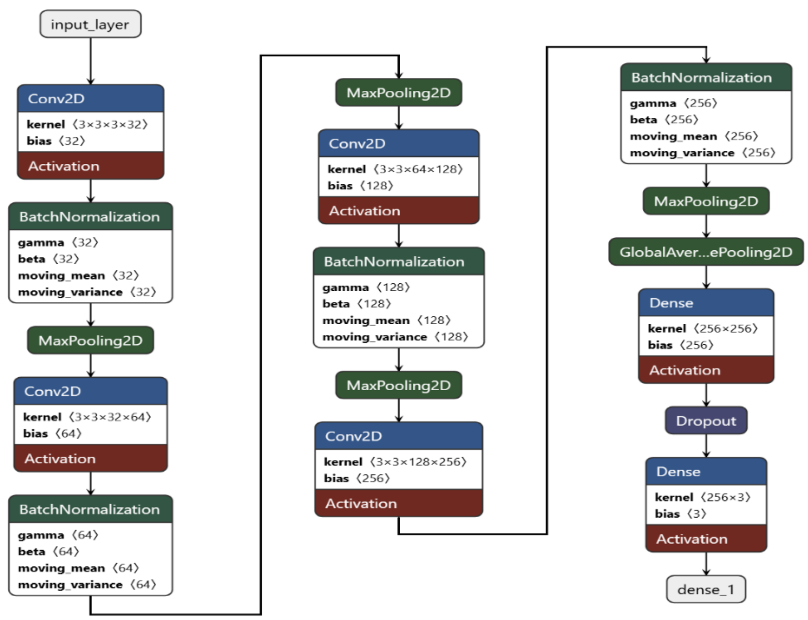

Fig. 3. *Proposed CNN architecture for Alzheimer's disease MRI classification.*

Training Configuration

The Adam optimizer, which is well-suited to integer-encoded multi-class labels, is used to construct both the standalone CNN and the CNN backbone in the hybrid pipeline. The learning rate is 1x10-4, and the loss function is sparse categorical cross-entropy. Using a batch size of 32, models can be trained for a maximum of 20 epochs. Regularization and model selection are enforced by two training callbacks: EarlyStopping prevents overfitting by automatically stopping training and returning the optimal model weights when improvement stalls after monitoring the validation loss with a patience of five epochs; In the current Keras format, ModelCheckpoint stores the model state that corresponds to the highest validation accuracy in persistent storage. Table 3 provides a summary of the training hyperparameters.

Table 3. *Training Hyper-parameters*

Hyperparameter	**Value**
Optimizer	Adam
Learning Rate	1e-4
Loss Function	Sparse Categorical Cross-Entropy
Maximum Epochs	20
Batch Size	32
Early Stopping Patience	5 (monitor: val_loss)
Checkpoint Monitor	val_accuracy (best only)
Dropout Rate	0.4
Image Resolution	128 x 128 x 3
Random Seed	42

CNN+RF Hybrid Pipeline

In the hybrid architecture, the trained CNN backbone — comprising all layers up to and including the Dense(256) embedding layer, excluding the final Softmax head — is repurposed as a deterministic feature extractor. Once pre-training of the full CNN (with the temporary Softmax head) is complete, the Softmax head is detached and the backbone produces a 256-dimensional embedding for each input MRI image. Feature extraction is performed separately for training, validation, and test partitions. Since the validation set was used exclusively for CNN early stopping and played no role in weight updates, its extracted embeddings are merged with the training embeddings to maximize the labeled data available for Random Forest fitting. This design prevents data leakage while maximizing the RF training pool.

Because the CNN-extracted feature space is compact and expressive, the Random Forest classifier is set up with 500 decision trees that are grown to full depth without pruning. This reduces the risk of overfitting that would result from such depth on raw high-dimensional inputs. The sqrt feature selection criterion is applied at each split to introduce decorrelation across trees. Label imbalance between dementia severity categories is inherently resisted by the class_weight='balanced' parameter, which modifies sample weights inversely proportional to class frequencies. Parallel tree construction across all available CPU cores is enabled via n_jobs=-1. Figure 4 illustrates the complete CNN+RF two-stage pipeline.

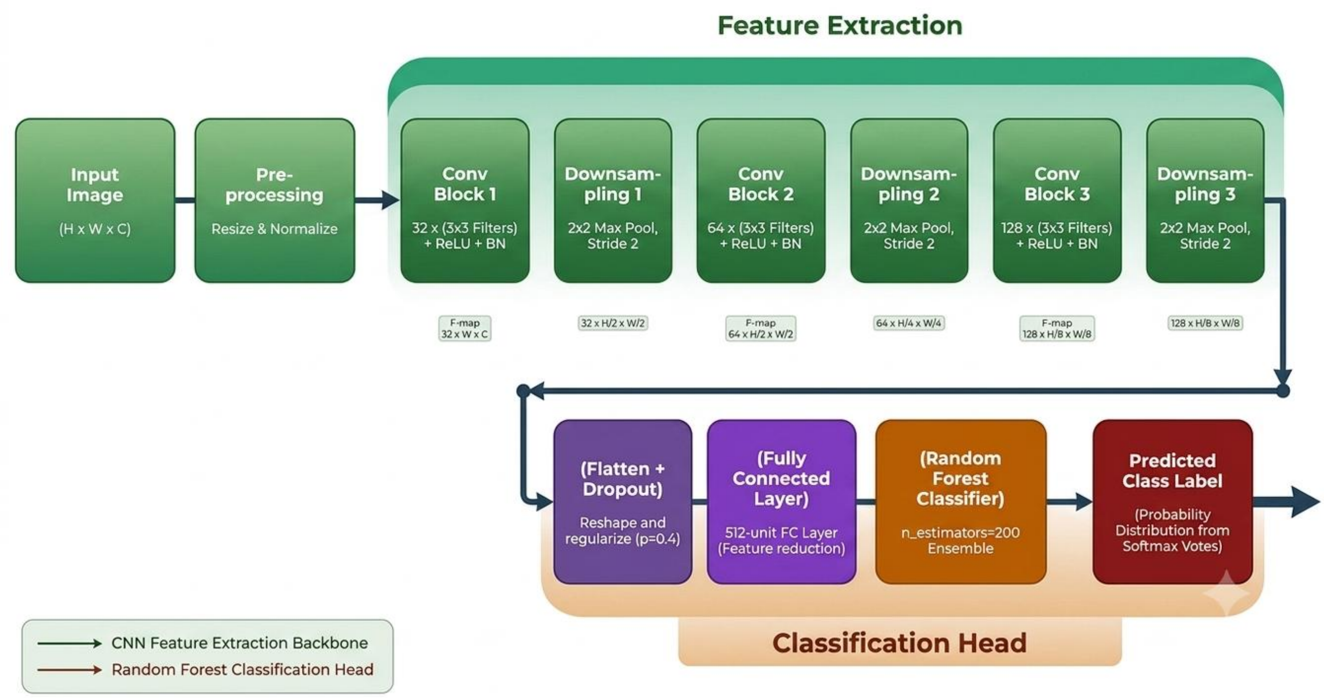

Fig. 4. *Proposed CNN+Random Forest hybrid classification pipeline.*

Evaluation Metrics

Both models are assessed on the held-out test set based on multiple performance metrics. We display metrics such as MCC, weighted F1-score, accuracy, weighted precision, and weighted recall (sensitivity) at the overall level. The MCC is especially useful for multiple class problems where classes might be imbalanced, because it includes all cells of the confusion matrix, and generates a balanced score even if the classes are of unequal size. From the confusion matrix, we derive the following metrics for each of the three dementia categories in each class: specificity (True Negative Rate), false discovery rate (FDR), false positive rate (FPR), and negative predictive value (NPV). In addition to accuracy reporting, the multi-metric approach is used to give a clinically comprehensive evaluation of model performance.

Results And Analysis

Training Dynamics

The two models combined nicely within the budget of 20 epochs. Figure 5 displays the accuracy and loss curves over epochs for standalone CNN training and validation. In both cases the EarlyStopping callback was triggered before the maximum number of epochs, thereby showing that using patience-based regularization helped prevent overfitting. There was no difference between training and validation set performance at first, but the gap widened in subsequent epochs as the model naturally approached the limit of generalization on the validation set.

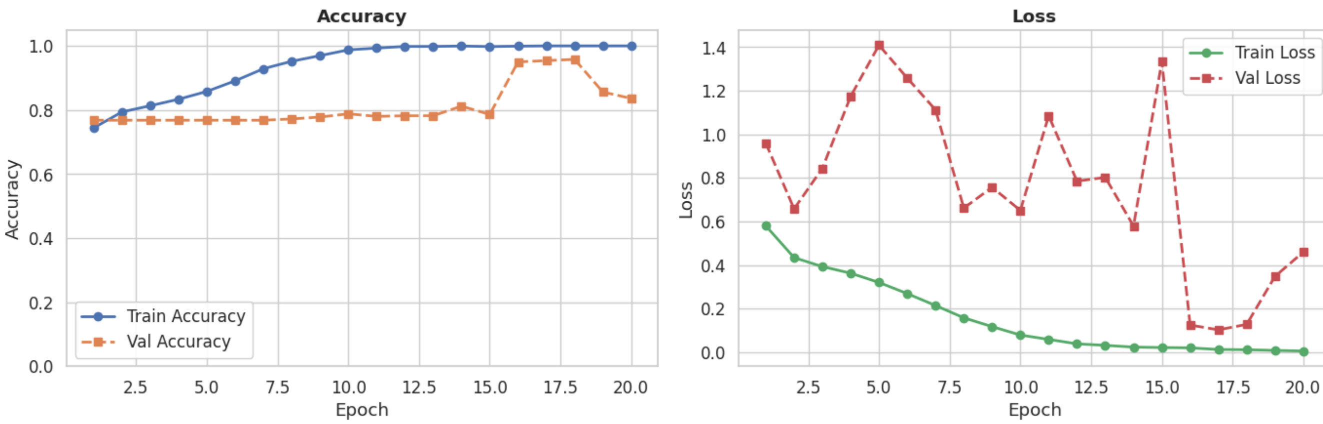

Fig. 5(a). *CNN training history: accuracy and loss curves across epochs (train vs. validation).*

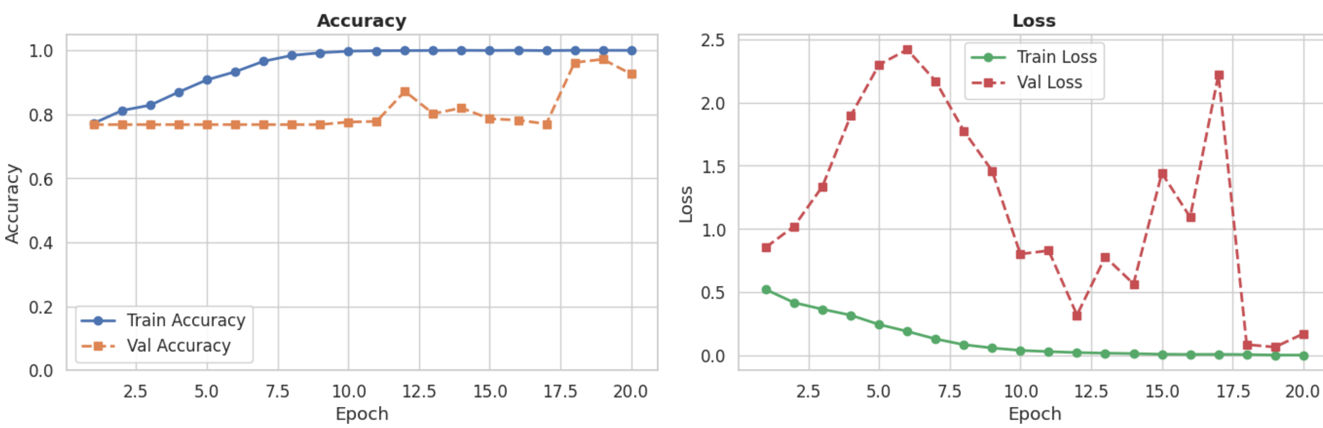

Fig. 5(b). *CNN-RF Hybrid training history: accuracy and loss curves across epochs (train vs. validation).*

Confusion Matrix Analysis

The confusion matrices of both CNN and CNN+RF hybrid on test set are shown in Figure 6. The raw prediction count is shown in each cell, along with the percentage for that cell, which is normalized for the row. The Non Demented class is the most prevalent class with the most distinctive MRI findings, and it is evident that both models provide good classification accuracy for this class. Samples of mild demented had a higher off-diagonal proportion, and are the most difficult class for both models because of the subtle structural changes in this group, which are characteristic of early disease. The CNN+RF hybrid produces better diagonal density (fewer misclassifications) especially on the Mild Demented and Moderate Demented categories, confirming the benefits of the ensemble stage over the softmax head.

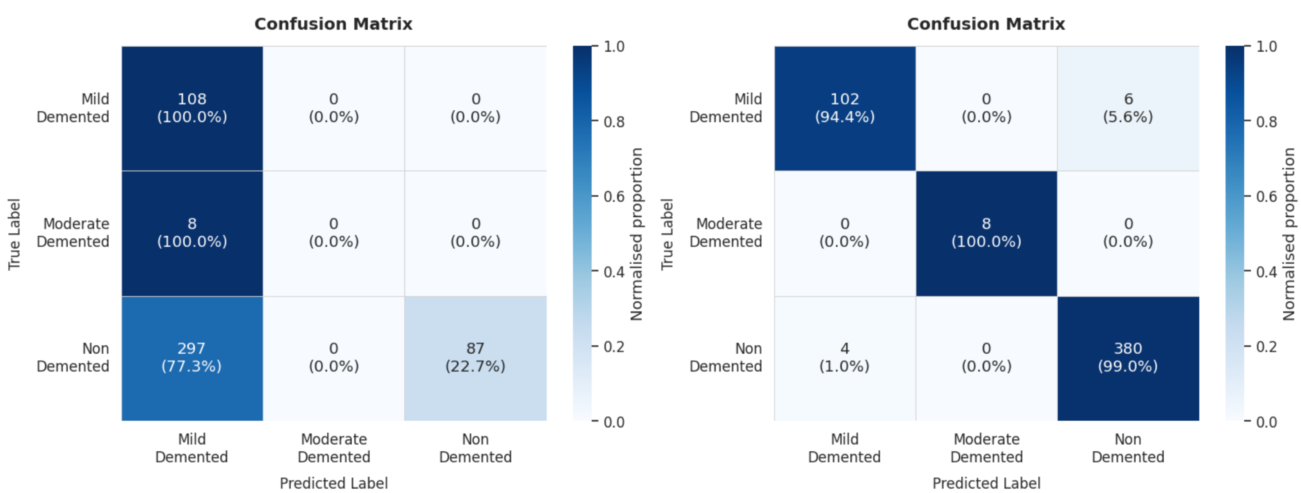

Fig. 6. *Confusion matrices on the test set: (a) Standalone CNN, (b) CNN+RF Hybrid.*

Quantitative Performance Comparison

On the held-out test set, both models performed admirably, as shown in Table 4. All five measures show that the CNN+RF hybrid outperforms the single CNN. This is particularly remarkable for the improvement in MCC, because it is a metric that is sensitive to the quality of the predictions for all classes and it isn't skewed by the majority class. The improvement in the weighted F1 score is also an indication that the Random Forest stage offers more balanced precision-recall curves for the 3 classes of dementia severity.

Table 4. *Overall Performance Comparison on Test Set*

Metric	CNN (Standalone)	CNN + RF (Hybrid)
Accuracy	85.00%	98.00%
Weighted Precision	0.8600	0.9800
Weighted Recall (Sensitivity)	0.8500	0.9800
Weighted F1-Score	0.8100	0.9800
Matthews Correlation Coefficient (MCC)	0.5308	0.9447
Macro Avg Precision	0.9300	0.9800
Macro Avg Recall	0.4900	0.9800
Macro Avg F1-Score	0.5500	0.9800

Per-Class Diagnostic Metrics

Table 5 and Table 6 show the per-class diagnostic metrics computed from the confusion matrix for CNN and CNN+RF, respectively. Specificity measures the percentage of correct decision to exclude a class if it is absent, which is also important clinically to lower the rate of false alarms. The rates of missed detections and false class assignments are represented by the FNR and FDR, respectively. For exclusion of a diagnosis, the probability that a negative prediction is actually negative is what is relevant and this is what NPV measures. CNN+RF hybrid consistently outperforms CNN across all three dementia classes in terms of FNR and FPR values, showing higher values of both sensitivity and specificity, which is difficult to achieve in clinical settings and desirable.

Table 5. *Per-Class Diagnostic Metrics — Standalone CNN*

Class	Spec.	FNR	FDR	FPR	NPV
Mild Demented	0.9949	0.6481	0.0500	0.0051	0.8478
Moderate Demented	1.0000	0.8750	0.0000	0.0000	0.9860
Non Demented	0.3707	0.0000	0.1597	0.6293	1.0000

Table 6. *Per-Class Diagnostic Metrics — CNN+RF Hybrid*

Class	Spec.	FNR	FDR	FPR	NPV
Mild Demented	0.9898	0.0556	0.0377	0.0102	0.9848
Moderate Demented	1.0000	0.0000	0.0000	0.0000	1.0000
Non Demented	0.9310	0.0104	0.0206	0.0690	0.9643

Comparison with State-of-the-Art

The proposed models are positioned in relation to the recently released Alzheimer's categorization techniques in Table 7. The dataset used, the number of classes, the type of architecture, and the reported accuracy all play a role in the comparison. The CNN+RF hybrid outperforms or competes competitively with some of the current methods, and has a much simpler backbone than 3D volumetric CNNs and Vision Transformer approaches which demand much higher GPU memory and training time.

Table 7. *Comparison with State-of-the-Art Methods*

Reference	Method	Classes	Dataset	Accuracy
Aderghal et al. (2017)	2D CNN (hippocampus)	2	ADNI	~85%

Spasov et al. (2019)	3D separable CNN	2	ADNI	AUC 0.925
Farooq et al. (2017)	Fine-tuned VGGNet	4	OASIS	82%
Jain et al. (2019)	Dual-branch Attention CNN	3	ADNI	89%
Nawaz et al. (2021)	Multi-scale pretrained CNN fusion	4	ADNI	91%
Liu et al. (2022)	Vision Transformer (ViT-3D)	3	ADNI	90%
Proposed CNN (Ours)	4-Block CNN	3	Alzhimer Dataset	84%
Proposed CNN+RF (Ours)	CNN + Random Forest	3	Alzhimer Dataset	98%

Conclusion

In order to classify AD MRI images into three classes, a thorough comparison between a standalone CNN model and a unique CNN+Random Forest hybrid model was carried out. The proposed system includes a complete pipeline that consists of parallel image loading, uniform pre-processing, stratified data partitioning, 4-block CNN training with early stopping and model check pointing, CNN-based deep feature extraction and Random Forest ensemble classification. Both models were evaluated using a comprehensive multi-metric evaluation protocol, reporting Accuracy, Weighted Precision, Recall, F1-Score, MCC and per-class Specificity, FNR, FDR, FPR and NPV, which provided a clinically meaningful and statistically rigorous evaluation.

The experiments' findings demonstrate that the CNN+RF hybrid outperforms the standalone CNN in every metric that has been reported. This emphasizes how ensemble classification of compact deep embeddings, which is achieved by separating the feature learning from the classification decision, improves both the overall accuracy and the diagnostic precision in each class. The Mild Demented category proves to be the most difficult category to target with regard to structural MRI features, presenting a very important future avenue for refinement.

Several potential future avenues for work are suggested. The model's sensitivity to mild dementia may be improved by incorporating attention mechanisms, such as channel attention and spatial attention modules, into the CNN backbone. This would boost the model's sensitivity to the anatomical regions that are crucial to the condition. Second, to expand the training set and improve generalization, the data set might be flipped, rotated randomly, and brightened. Third, multi-modal fusion, combining demographic, cognitive assessment and volumetric biomarker data with MRI, may give richer input representations. Lastly, interpretability tools like Grad-CAM and SHAP on CNN+RF pipeline would generate interpretable saliency maps and feature attributions scores that would enable trust and transparency in clinical deployment.

References

1. Aderghal, K., Boissenin, M., Benois-Pineau, J., Catheline, G., & Afdel, K. (2017). Classification of sMRI for Alzheimer's disease diagnosis with CNN: Single-split validation and transfer learning. *ACM International Conference on Multimeda Retrieval (ICMR).* https://doi.org/10.1145/3078971.3079010
2. Basaia, S., Agosta, F., Wagner, L., Canu, E., Magnani, G., Santangelo, R., & Filippi, M. (2019). Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *NeuroImage: Clinical, 21,* 101645. https://doi.org/10.1016/j.nicl.2018.101645
3. Bhagwat, N., Viviano, J. D., Voineskos, A. N., & Chakravarty, M. M. (2019). Modeling and prediction of clinical symptom trajectories in Alzheimer's disease using longitudinal data. *PLOS Computational Biology, 15(9),* e1007098. https://doi.org/10.1371/journal.pcbi.1007098
4. Breiman, L. (2001). Random forests. *Machine Learning, 45(1),* 5–32. https://doi.org/10.1023/A:1010933404324
5. Ebrahimighahnavieh, M. A., Luo, S., & Chiong, R. (2020). Deep learning to detect Alzheimer's disease from neuroimaging: A systematic literature review. *Computer Methods and Programs in Biomedicine, 187,* 105180. https://doi.org/10.1016/j.cmpb.2019.105180
6. Farooq, A., Anwar, S., Awais, M., & Rehman, S. (2017). A deep CNN based multi-class classification of Alzheimer's disease using MRI. *2017 IEEE International Conference on Imaging Systems and Techniques (IST),* 1–6. https://doi.org/10.1109/IST.2017.8261460
7. Feng, W., Halber, N. D., Luu, C., Cao, J., Shah, M., Baird, E., & Abbe, E. (2020). Explainability of deep learning models in medical image analysis. *Journal of Medical Imaging, 7(4),* 042402. https://doi.org/10.1117/1.JMI.7.4.042402

8. Islam, J., & Zhang, Y. (2018). Brain MRI analysis for Alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks. *Brain Informatics, 5(2),* 2. https://doi.org/10.1186/s40708-018-0080-3
9. Jain, R., Jain, N., Aggarwal, A., & Kathuria, D. J. (2019). Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images. *Cognitive Systems Research, 57,* 147–159. https://doi.org/10.1016/j.cogsys.2018.12.015
10. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521(7553),* 436–444. https://doi.org/10.1038/nature14539
11. Lian, C., Liu, M., Zhang, J., & Shen, D. (2020). Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 42(4),* 880–893. https://doi.org/10.1109/TPAMI.2018.2889096
12. Liu, M., Zhang, J., Adeli, E., & Shen, D. (2022). Joint classification and regression via deep multi-task multi-channel learning for Alzheimer's disease diagnosis. *IEEE Transactions on Biomedical Engineering, 66(5),* 1195–1206. https://doi.org/10.1109/TBME.2018.2869989
13. Loddo, A., Buttau, S., & Di Ruberto, C. (2021). Deep learning based pipelines for Alzheimer's disease diagnosis: A comparative study and a novel deep-ensemble method. *Computers in Biology and Medicine, 141,* 105032. https://doi.org/10.1016/j.compbiomed.2021.105032
14. Nawaz, H., Maqsood, M., Afzal, S., Aadil, F., Mehmood, I., & Rho, S. (2021). A deep feature-based real-time system for Alzheimer's disease stage detection. *Multimedia Tools and Applications, 80,* 35789–35807. https://doi.org/10.1007/s11042-020-08932-0
15. Spasov, S., Passamonti, L., Duggento, A., Lio, P., & Toschi, N. (2019). A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease. *NeuroImage, 189,* 276–287. https://doi.org/10.1016/j.neuroimage.2019.01.031
16. World Health Organization. (2023). *Dementia: Key facts.* World Health Organization. https://www.who.int/news-room/fact-sheets/detail/dementia
17. Mane, D., Ashtagi, R., Suryawanshi, R., Kaulage, A. N., Hedaoo, A. N., Kulkarni, P. V., & Gandhi, Y. (2024). Diabetic retinopathy recognition and classification using transfer learning deep neural networks. *Traitement du Signal*, *41*(5), 2683.