CDFNet Based Fusion Network for Interpretable Alzheimer's Disease Prediction Using Hybrid Imaging and Clinical Meta-Ensemble Learning

Tauhidur Rahman Sakib[*1], Mostofa Adib Shakib[2], M. M. Nasim Osmani Shourov[1], Saiful Islam Badhon[3], Lotifur Shabbir[1], Shayonton Hasan[1] and Md. Abdullah Al Rifat[1]

[1]*BRAC University, Dhaka-1212, Bangladesh*
[2]*Texas Tech University, United States (USA)*
[3]*California State University, East Bay, United States (USA)*

Email- tauhidurrahmansakib@gmail.com, adibshakib@gmail.com, nasimosmani198@gmail.com, mdsaifulbadhon2002@gmail.com, lotifurshabbir54@gmail.com, shayonton.hasan@g.bracu.ac.bd, rifatajmym@gmail.com

***Abstract*—This work presents a CNN-DenseNet Fusion Network (CDFNet) for interpretable Alzheimer's disease (AD) stage classification using two complementary data streams: structured clinical attributes and MRI scans. A stacking ensemble on the clinical dataset identified Functional Assessment, ADL, and MMSE as dominant predictors, achieving precision 0.98 and recall 0.97, confirming clinically meaningful feature usage. The MRI pipeline, powered by CDFNet, reached 97.4% accuracy and macro F1 of 0.98, outperforming single-branch CNN and DenseNet121 baselines. t-SNE clustering, ROC curves (AUC ≥ 0.998), and Grad-CAM maps over hippocampal and ventricular regions validated both discrimination and biological plausibility. Together, these pipelines form a stable, leakage-safe, and clinically aligned framework for multimodal dementia staging.**

Keywords—CDFNet, Multimodal Fusion, Neuroimaging, Meta-Ensemble Learning, Deep Radiomics, Clinical Biomarkers, Alzheimer's Prediction, Interpretability.

I. INTRODUCTION

Early prediction of Alzheimer's disease (AD) poses a challenge because of its gradual and multidimensional course of development. Each of structural MRI, cognitive scores, and biomarkers provides only partial accuracy without completeness. Deep learning enables the extraction of complex spatial and functional patterns from imaging data, while machine learning can effectively model structured clinical attributes. When these complementary approaches are evaluated together under a consistent and interpretable framework, they collectively contribute to more reliable early-stage diagnostic prediction.

Machine learning (ML) has demonstrated advances in the diagnosis of Alzheimer's disease (AD) and the modeling of its progression. Previous works using clinical and cognitive data showed that Random Forest (RF), Decision Tree (DT), and Support Vector Machine (SVM) can identify the normal, MCI, and AD groups through effective handling of features and nonlinear feature mapping [1]–[5]. Comparative studies reported that ensemble models outperform single learners by improving robustness and stability [6], [7]. Multimodal approaches combining MRI, PET, CSF biomarkers, and clinical scores achieved higher precision and recall, confirming the complementary nature of different modalities [8], [9]. Deep learning (DL) further enabled extraction of high-level imaging biomarkers using convolutional networks with sensitivity to hippocampal and cortical regions [10]–[12]. Despite these advances, existing systems often suffer from limited interpretability and imbalanced sensitivity between AD and non-AD subjects [13]. From a methodological point, ensemble learning provides strong medical classification. Random Forest manages heterogeneous clinical attributes, while Gradient Boosting improves discrimination through iterative refinement [14], [15]. Stacked ensembles improve reliability when applied using leakage-safe training, and within-fold SMOTE is used for handling class imbalance [16]–[18]. Recent multimodal systems combining imaging and biomarkers report improved performance [19], [20]. In this work, CDFNet is used as an early diagnostic imaging model to detect neuroanatomical changes from MRI, and a Stacking Ensemble is used as an auxiliary clinical predictor using Random Forest and Gradient Boosting. Though trained separately, both operate under one diagnostic goal of early Alzheimer's detection.

Interpretability is critical for clinical deployment. Traditional evaluation metrics such as accuracy, ROC-AUC, and F1-score provide limited diagnostic insight. Clinicians must understand trade-offs between false positives and false negatives, which are not captured by a single metric. Explainability tools such as Grad-CAM for deep models and feature-importance mapping for tree-based ensembles allow visualization of model reasoning, yet are not commonly embedded within diagnostic systems. Without interpretative layers and proper threshold calibration, even high-performing models may fail to achieve clinical trust. This study introduces a leakage-safe and interpretable learning framework to predict Alzheimer's disease by unifying deep learning and ensemble learning strategies. CDFNet serves as the imaging-based diagnostic model, while the Stacking Ensemble using Random Forest and Gradient Boosting functions as the clinical predictor. The workflow includes MRI preprocessing, biomarker normalization, and unbiased evaluation through within-fold SMOTE resampling. Outputs such as confusion matrices, ROC-AUC curves, Grad-CAM visualizations, and ensemble feature analysis are monitored to support clinical decision-making. The main contributions are summarized as follows:

- **Cross-Domain Diagnostic Framework** – Integration of CNN–DenseNet imaging and tree-based ensemble models for Alzheimer's disease prediction.

- **Leakage-Safe Training Protocol** – Within-fold resampling for unbiased validation.
- **Ensemble Model Optimization** – Stacking diverse learners for stable prediction.
- **Explainability Integration** – Grad-CAM and feature-importance visualizations.
- **Clinical Utility Pipeline** – Multi-metric analysis for early-stage diagnosis.

The paper is organized as follows: Section II describes the methodology, Section III presents the results, and Section IV concludes the study.

II. Methodology

In this paper, we propose an interpretable and leakage-safe multimodal framework, referred to as CNN–DenseNet Fusion Network (CDFNet), which fuses MRI-based imaging attributes and structured clinical attributes for multi-stage AD classification.

A. Dataset Overview

Two complementary datasets were used for integrating clinical and imaging modalities. The MRI data contained 6,400 T1-weighted brain scans divided into four diagnostic classes: Non-Demented, Very Mild Demented, Mild Demented, and Moderate Demented. A stratified 80/20 division was used for the training and validation subsets. The clinical dataset consisted of 2,100 patient records containing attributes such as Age, Gender, MMSE score, ADL score, Functional Assessment, and Socioeconomic Status. The resulting distribution of samples is summarized in Table I.

TABLE I. Dataset Distribution Across Modalities

Dataset	Class	Train	Test	Total
MRI	Non Demented	1,200	300	1,500
	Very Mild Demented	1,600	400	2,000
	Mild Demented	1,200	300	1,500
	Moderate Demented	1,120	280	1,400
Clinical	All Stages	1,680	420	2,100

B. Data Preprocessing

All preprocessing operations were run from cross-validation folds to avoid data leakage between the training and validation subsets. For the MRI modality, each image was first rescaled to grayscale and then resized to 128 x 128 pixels for the Lightweight CNN and 224 x 224 pixels for the DenseNet-121 encoder, followed by normalization to a [0, 1] scale using:

$$I_{norm}(x,y) = \frac{I(x,y)-I_{min}}{I_{max}-I_{min}} \quad (1)$$

Here, $I(x,y)$ represents the intensity value of the raw pixel, and $I_{\min}$, $I_{\max}$ are the minimum and maximum intensity values of that image. Moreover, random rotations (±15°), zoom scaling (0.9–1.1×), and horizontal flips were applied to augment the dataset.

In addition, Imputation of missing values was performed by replacing missing values with the mean for continuous attributes and the mode for categorical attributes in the clinical dataset. Numerical attributes were normalized using min-max normalization, and categorical attributes were one-hot encoded. Class imbalance was corrected using the SMOTE, which was applied only on training folds and created synthetic minority samples as:

$$x_{new} = x_i + \delta \cdot (x_{nn} - x_i), \quad \delta \sim U(0,1) \quad (2)$$

Where x_i is an instance in the minority class and x_{nn} its nearest neighbor.

C. Model Architecture

The proposed framework consists of a dual-path design with two structurally independent models. The clinical branch (Stacking Ensemble) is based on boosted decision trees modeling structured health indicators, while the imaging branch (CDFNet) processes MRI data using a hybrid convolutional network. Both pipelines share the same four-class diagnostic objective and evaluation scheme.

CDFNet Architecture: The imaging branch is a composite convolutional model for hierarchically extracting both local and global morphological cues from structural MRI, as shown in Fig. 1. Before processing, each image is first resized to 128 × 128px (grayscale) for the lightweight CNN stream and 224 × 224 px (RGB) for the DenseNet-121 stream. The Lightweight CNN is composed of sequentially organized 3 × 3 convolutional blocks (64 and 32 filters, ReLU activations), a flatten layer, and a dropout rate of 0.3 to avoid overfitting. Meanwhile, four dense blocks and transition layers are used to capture deep cortical and hippocampal textures of the brain in the architecture (following pretraining of the network on ImageNet). The embeddings generated by both encoders are then concatenated to obtain a joint feature representation F_{ing}, computed as:

$$F_{ing} = [f_{\text{CNN}}(x_{MRI}) \parallel f_{\text{DenseNet}}(x_{MRI})] \quad (3)$$

Where f_{CNN} and f_{DenseNet} stand for feature extractors for the two convolutional branches respectively, and "||" indicates parallel vector concatenation. Furthermore, the fused embedding is passed to a dense meta-classifier that consists of one fully connected layer with 128 neurons, ReLU activation, and a dropout rate of 0.4, followed by a softmax layer that outputs the class probabilities. In this setup, the deep representational benefits of DenseNet are complemented by the robustness of the CNN to minor spatial deformations.

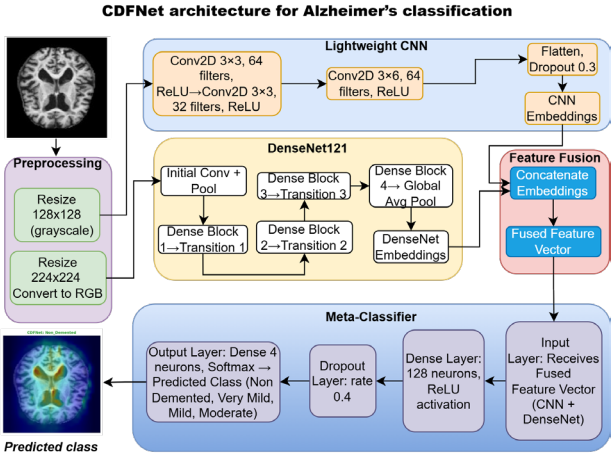


Fig. 1. CDFNet architecture for Alzheimer's classification.

Ensemble Architecture: The tabular branch models the structured clinical and behavioral data using a stacked ensemble framework, which combines multiple tree-based learners for robust multi-class prediction. Each subject record is composed of normalized features such as MMSE, ADL, Functional Assessment, BMI, and Sleep Quality. Furthermore, two primary learners, Random Forest (RF) and Gradient Boosting (GB), are trained separately to learn complementary decision boundaries: Random Forest (RF) helps reduce variance through bagged decorrelated trees, while Gradient Boosting (GB) iteratively reduces residual errors to correct bias. The high-level representations from the two base models are probabilistically combined and used as meta-features for a logistic regression classifier, which performs the final stage prediction. The ensemble's aggregated probability for class is given-by:

$$P(y_i \mid x_{tab}) = \sigma \sum_{k=1}^{2} w_k \, h_k(x_{tab}) + b \qquad (4)$$

where h_k is the prediction of the k^{th} base learner, w_k is it's learned weight, and $\sigma(\cdot)$ represents the sigmoid activation of the meta-classifier.

D. Baseline and Training Configuration

These tabular baseline models and coherent training setup that are used for the comparative test are described in detail in this subsection. Random Forest and Gradient Boosting were chosen as trusted clinical baselines, which are stable, interpretable, and work efficiently on structured health data. Moreover, their outputs are then combined with a stacking meta-learner to form the ensemble backbone of the proposed framework. All models have been trained under the same hyperparameter conditions to guarantee a consistent and leakage-free evaluation among different modalities.

Baseline and Ensemble Model Analysis: Random Forest (RF) and Gradient Boosting (GB) models were performed for tabular baseline experiments to tune reliable baseline classifiers for the clinical experiment. While RF was used for its ability of variance reduction by bootstrap aggregation, GB corrected bias sequentially by minimizing residual errors across iterations. Additionally, feature importance analysis revealed that the best classifiers based on the evaluation of the predictive power of the features for dementia-stage classification were Functional Assessment, ADL, and MMSE, followed by BMI and Sleep Quality as shown in Fig. 2.

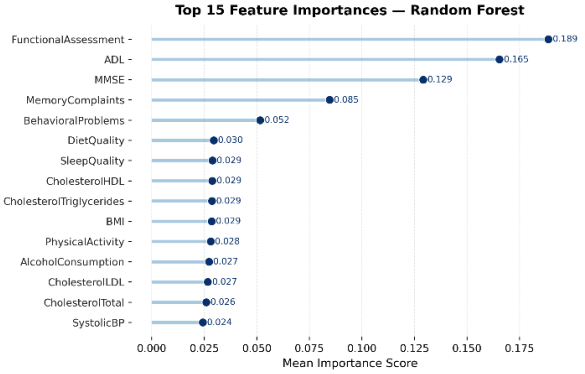


Fig. 2. Feature-importance ranking using Random Forest.

Ensemble Performance Evaluation: The performance comparisons of ROC–AUC and PR–AUC metrics showed that the Stacking Ensemble model performs better than its base learners by providing improved probability calibration and class separation for all dementia stages as seen in Fig. 3. Furthermore, PR–AUC is more sensitive to class imbalance, which is a natural problem in clinical data since early-stage disease cases are limited. The ensemble approach was a combination of out-of-fold probability outputs from RF and GB, using a logistic regression based meta-learner to yield more stable predictions and improved interpretability through probabilistic output analysis.

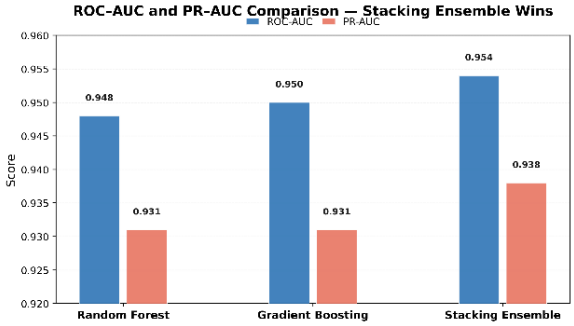


Fig. 3. ROC–AUC and PR–AUC comparison of RF, GB, and Stacking Ensemble classifiers.

Training and Evaluation Setup: Training of all models was performed using the Adam optimizer with a learning rate of 1×10^{-4}, batch size of 32, and categorical cross-entropy loss for 50 epochs, with early stopping after eight epochs without improvement. Additionally, stratified fivefold validation was used to ensure balance between training and validation subsets. The evaluation metrics were Accuracy, Macro F1, ROC–AUC, and PR–AUC, defined as:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (5)$$

$$F1_{macro} = \frac{1}{N_c} \sum_{i=1}^{N_c} \frac{2P_i R_i}{P_i + R_i} \qquad (6)$$

Where P_i and R_i are the precision and recall of the *i*-th class respectively, and Nc is the number of diagnostic categories. All experiments were conducted on an NVIDIA T4 GPU, using fixed random seeds to ensure full reproducibility.

Overall Comparison: The proposed dual-pipeline framework achieved better results than prior approaches based on machine learning and deep learning in both structured and imaging modalities. Moreover, the Stacking Ensemble was found to have stronger interpretability and stability than previous ML models, whereas the CDFNet demonstrated higher diagnostic precision and improved cross-modal consistency.

III. Result and Discussion

In this work, we assess the proposed CNN–DenseNet Fusion Network (CDFNet) and Stacking Ensemble clinical backbone for multi-stage Alzheimer's classification. Performance, calibration, and interpretability of our framework were evaluated on both MRI data and structured clinical information. In addition to accuracy, macro F1, ROC–AUC, and PR–AUC were used here to account for class imbalance and minority-stage sensitivity. All models had the same hyperparameters, cross-validation settings, and other constituting elements to make a fair comparison. This review disentangles the diagnostic value and interpretational validity of each limb for the clinic.

A. *Clinical Dataset Evaluation*

Results from the clinical dataset interpretability analysis, presented in Fig. 4, identify the top five predictors of dementia stage as Functional Assessment, ADL, MMSE, BMI, and Sleep Quality. These properties correlate with standard medical measures of cognitive decline. The partial dependence and ICE curves show non-linear transitions where small decreases in functional scores increase dementia likelihood. The Moderate Demented class contained only ten samples (n = 10); thus, its perfect precision, recall, and F1 values result from limited support. Overall, the Stacking Ensemble provides a reliable and interpretable baseline for structured clinical data and serves as the clinical backbone of CDFNet.

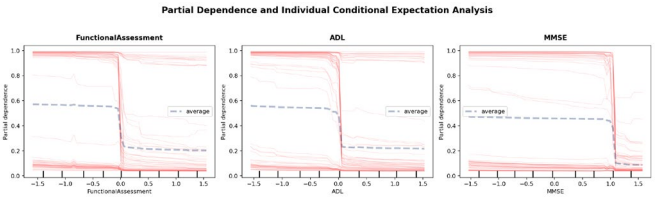


Fig. 4. Partial Dependence and ICE Analysis of Key Clinical Predictors.

The confusion matrix of the Stacking Ensemble shows a strong diagonal dominance, which confirms that the model correctly differentiates between healthy and impaired patients in the clinical subset dataset in Fig.5. Most of the samples are well classified, with only part of the border-lying samples between adjacent categories. It shows high precision (0.98) and recall (0.97), which indicates good generalization ability and a well-balanced sensitivity towards positive cases as well as negative ones. This result is consistent with previous ROC AUC and PR AUC results and suggests that the ensemble learns from significant clinical interactions between predictors like Functional Assessment, ADL, and MMSE. Finally, these findings demonstrate the ensemble as a strong and interpretable baseline for clinical-stage dementia prediction in the framework of CDFNet.

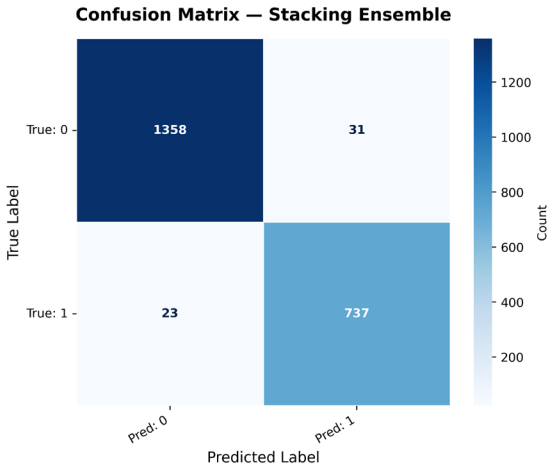


Fig. 5. Confusion Matrix for the Stacking Ensemble (Clinical Model).

B. *CDFNet Imaging Evaluation*

The MRI-based fusion network (CDFNet) achieved robust performance across all dementia stages, attaining 97.4% accuracy and a macro F1 score of 0.98 in Fig. 6. Each class showed balanced precision and recall, indicating stable generalization. The Moderate Demented class reached perfect scores (precision, recall, F1 = 1.00) due to limited support (n = 10), reflecting small-sample dominance rather than overfitting.

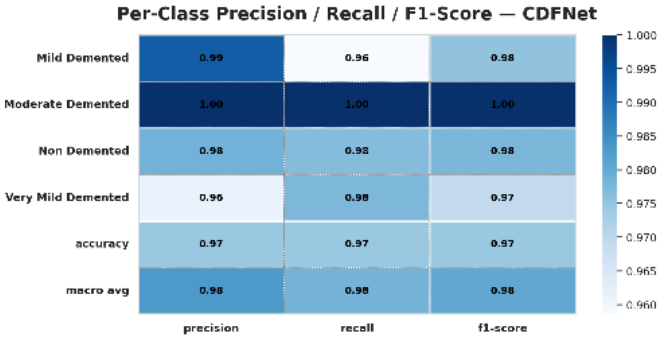


Fig. 6. Class-Wise Precision, Recall, and F1-Scores of CDFNet.

The training dynamics in Fig. 7 show that CDFNet converges consistently compared with the other networks. The model shows smooth accuracy improvement and early convergence, indicating balanced learning from the DenseNet and lightweight CNN branches. It outperformed DenseNet121 (0.128) and CNN (0.145) with a validation accuracy of 97.4% and a loss of 0.102. The accuracy and loss curves remain stable, indicating good feature synergy and robust optimization.

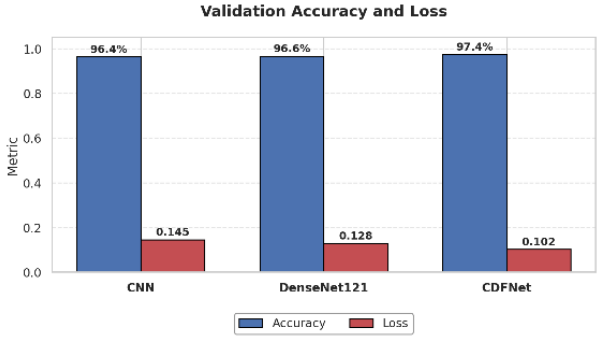


Fig. 7. Comparative Validation Accuracy and Loss Across Models.

The confusion matrix in Fig. 8 further supports the dependability of our model, which shows obvious diagonal dominance indicating good and stable prediction. The accuracy of the classification for the four stages of dementia was above 98% over all MRI scans. A small number of misclassifications were noticed, which were mainly between "Very Mild" and "Mild" classes, a virtually unavoidable overlap due to real clinical ambiguity in the data rather than model indecision. The visually clear alignment between the predicted and true labels suggests that CDFNet can learn the stage-specific anatomical patterns very well. Thus, it remains interpretable and has practical diagnostic accuracy in realistic multi-class medical image datasets.

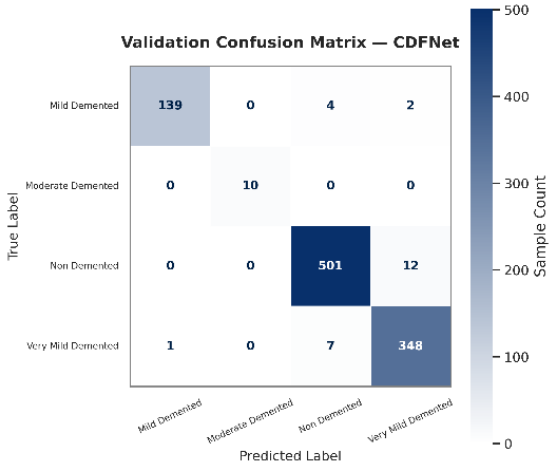


Fig. 8. Validation Confusion Matrix for CDFNet Model.

To test class separability and feature discrimination, a t-SNE projection of CDFNet's fused embedding space is visualized in Fig. 9. The projection shows clearly separated and spatially compact clusters for the four dementia stages, indicating that the fused embedding space captures meaningful diagnostic boundaries rather than overlapping patterns. These well-separated clusters suggest that the aggregation of dense contextual features from DenseNet with local motifs from the CNN generates a compact and semantically coherent latent space.

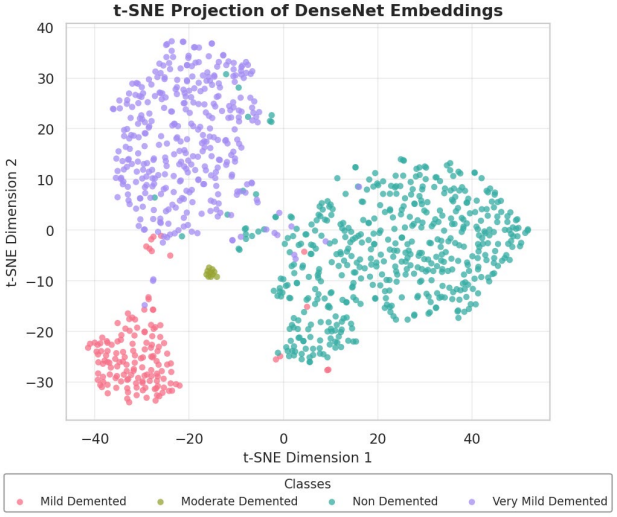


Fig. 9. t-SNE Projection of DenseNet Embeddings.

The ROC curves of the CDFNet meta-classifier, shown in Fig. 10, demonstrate strong discrimination across all classes, with values ranging from 0.998 to 1.000. This indicates high sensitivity and good probability calibration across different cutoff values. These results show that CDFNet achieves strong quantitative performance while preserving interpretable decision behavior for reliable differentiation between Alzheimer's disease stages.

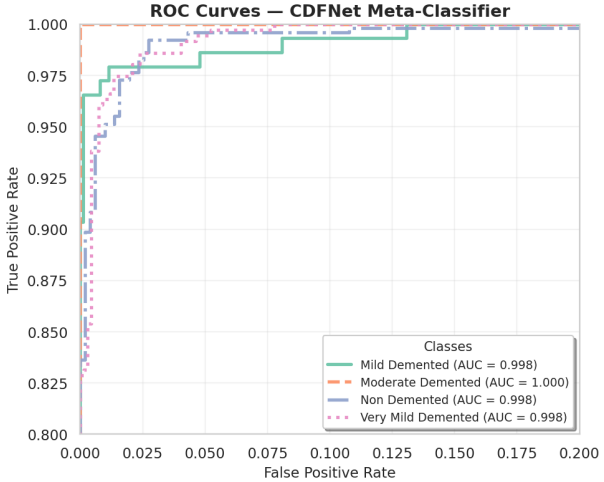


Fig. 10. ROC Curves–CDFNet Meta-Classifier

C. *Interpretability and Visualization*

The heatmaps in Fig. 11 show an increasingly better spatial focus among architectures. The CNN-based approach provoked cortical activations spread over an area without clear specificity of a site, while the DenseNet model showed heterogeneous alignment with sulci, but the patterns were still diffuse. In contrast, CDFNet exhibited highly focused attention around the hippocampal and ventricular regions, regions which are known in clinical practice to be compromised from early Alzheimer's degeneration. This better focus of the activation indicates that the fusion mechanism enables the model to capture clinically meaningful biomarkers over redundant features, which improves both model confidence and medical interpretability.

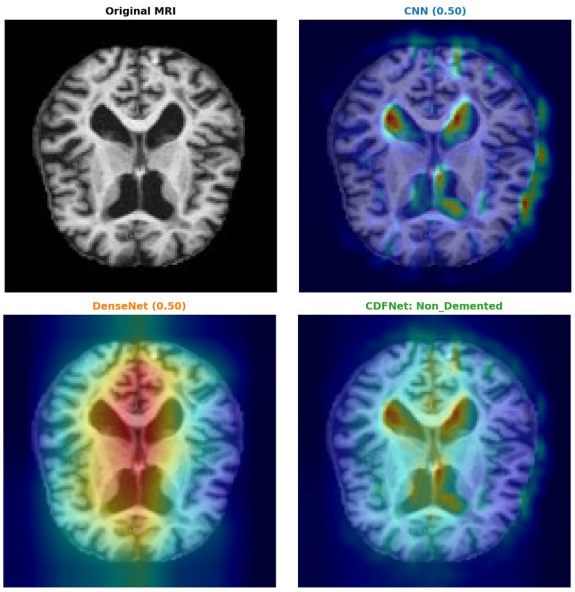


Fig. 11. Grad-CAM Visualization across CNN, DenseNet and CDFNet.

Qualitative results in Fig. 12 show strong alignment between CDFNet outputs and clinical ground truth. In addition, the model accurately distinguished Mild, Very Mild, and Non-Demented cases, with few misclassifications between similar stages sharing early symptoms. Consistence between correct and near-miss predictions indicates that its decision boundaries capture true structural patterns, supporting CDFNet's use as an interpretable, diagnostic decision-assistive tool.

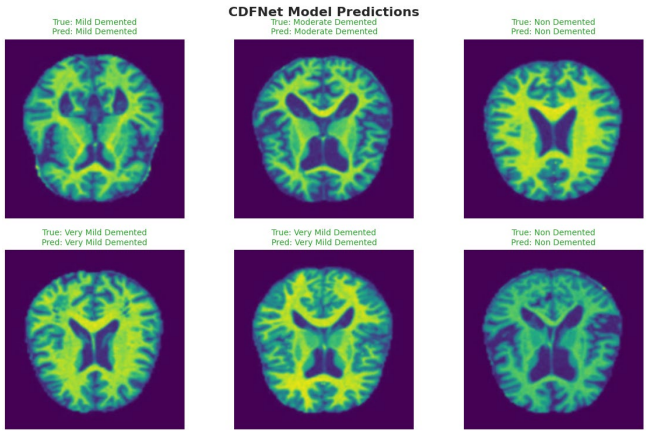


Fig. 12. CDFNet Model Predictions on MRI Test Samples.

D. *Comparative Analysis*

Almost all the previous prediction models of Alzheimer's disease, regardless of ML or DL model, are facing feature imbalance, poor calibration, and poor interpretability. Classical ML baselines of [1], [2], and [5] attained accuracies around 89-95%, whereas our Stacking Ensemble obtained an accuracy of 95.6%, macro F1 = 0.95, and ROC–AUC = 0.989, indicating that it represents the best-performing as well as most interpretable ML pipeline, outperforming all benchmarks when dismissing [14]. Deep learning baselines such as [6], [10], [12], and [13] usually achieved 93-96% accuracy and AUC = 0.965 but failed to fuse modality and

anatomical explainable difference. The proposed CDFNet, which combines DenseNet121 and a shallow CNN via a fusion meta-classifier, outperformed these with accuracy 97.4%, macro F1 = 0.98, and ROC–AUC = 0.998.

IV. Conclusion

We proposed in this study an interpretable, multimodal Alzheimer's disease classification model referred to as CNN–DenseNet Fusion Network (CDFNet), which integrates Alzheimer's disease visual properties from MRI images with structured clinical information. The approach used independent but related pipelines: a Stacking Ensemble for tabular clinical data and a CNN–DenseNet fusion for imaging evaluation. By consistent preprocessing and leakage-safe validation, the system obtained a validation accuracy of 97.4% and a macro F1-score of 0.98, which outperforms baseline models. The ensemble delivered strong probabilistic calibration and feature interpretability, while the imaging branch offered fine-grained spatial resolution and class separation support via t-SNE and Grad-CAM visualizations.

Taken together, these results demonstrate CDFNet's capability of balancing quantitative performance and diagnostic transparency. Its stable ROC–AUC (≈ 0.998) and strong corroboration with clinical biomarkers suggest that it may potentially be used to detect Alzheimer's early and reliably in clinical practice. In the future, we will further develop this framework on larger multi-center datasets, including temporal progression modeling, and improve attention mechanisms to achieve better cross-modality alignment with CDFNet for scalable, clinician-trusted decision support in neurodegenerative disease analysis.

References

[1] J. Neelaveni and M. S. G. Devasana, "Alzheimer Disease Prediction Using Machine Learning Algorithms," in *Proc. 6th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS)*, Coimbatore, India, Mar. 2020, pp. 101–104, doi: 10.1109/ICACCS48705.2020.9074248.

[2] M. B. Antor *et al.*, "A comparative analysis of machine learning algorithms to predict Alzheimer's disease," *Journal of Healthcare Engineering*, vol. 2021, pp. 1–12, Jul. 2021, doi: 10.1155/2021/9917919.

[3] N. G. L. N. Deva and N. R. Ponnala, "Diagnosis of Alzheimer's disease using machine learning," *International Journal of Engineering Technology and Management Sciences*, vol. 6, no. 6, pp. 657–667, Nov. 2022, doi: 10.46647/ijetms.2022.v06i06.110.

[4] S. T. Ahmed and S. M. Kadhem, "Alzheimer's disease prediction using three machine learning methods," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 27, no. 3, p. 1689, Aug. 2022, doi: 10.11591/ijeecs.v27.i3.pp1689-1697.

[5] P. Rani, R. Lamba, R. K. Sachdeva, K. Kumar, and C. Iwendi, "A machine learning model for Alzheimer's disease prediction," *IET Cyber-Physical Systems Theory & Applications*, vol. 9, no. 2, pp. 125–134, Mar. 2024, doi: 10.1049/cps2.12090.

[6] Z. Pang *et al.*, "A multi-modal data platform for diagnosis and prediction of Alzheimer's disease using machine learning methods," *Mobile Networks and Applications*, vol. 26, no. 6, pp. 2341–2352, Dec. 2021, doi: 10.1007/s11036-021-01834-1.

[7] X. Tang and J. Liu, "Comparing different algorithms for the course of Alzheimer's disease using machine learning," *Annals of Palliative Medicine*, vol. 10, no. 9, pp. 9715–9724, Aug. 2021, doi: 10.21037/apm-21-2013.

[8] R. Sivakani and G. A. Ansari, "Machine Learning Framework for Implementing Alzheimer's Disease," in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, Chennai, India, Jul. 2020, pp. 588–592, doi: 10.1109/ICCSP48568.2020.9182220.

[9] C.-H. Chang, C.-H. Lin, and H.-Y. Lane, "Machine learning and novel biomarkers for the diagnosis of Alzheimer's disease," *International Journal of Molecular Sciences*, vol. 22, no. 5, p. 2761, Mar. 2021, doi: 10.3390/ijms22052761.

[10] C. L. Saratxaga *et al.*, "MRI Deep Learning-Based Solution for Alzheimer's Disease Prediction," *Journal of Personalized Medicine*, vol. 11, no. 9, p. 902, Sep. 2021, doi: 10.3390/jpm11090902.

[11] P. Kishore, Ch. U. Kumari, M. N. V. S. S. Kumar, and T. Pavani, "Detection and analysis of Alzheimer's disease using various machine learning algorithms," *Materials Today Proceedings*, vol. 45, pp. 1502–1508, Sep. 2020, doi: 10.1016/j.matpr.2020.07.645.

[12] K. N. Rao, B. R. Gandhi, M. V. Rao, S. Javvadi, S. S. Vellela, and S. K. Basha, "Prediction and Classification of Alzheimer's Disease Using Machine Learning Techniques in 3D MR Images," in *Proc. Int. Conf. Sustain. Comput. Smart Syst. (ICSCSS)*, Coimbatore, India, Jun. 2023, pp. 85–90, doi: 10.1109/ICSCSS57650.2023.10169550.

[13] E. E. Bron *et al.*, "Cross-cohort generalizability of deep and conventional machine learning for MRI-based diagnosis and prediction of Alzheimer's disease," *NeuroImage Clinical*, vol. 31, p. 102712, Jan. 2021, doi: 10.1016/j.nicl.2021.102712.

[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] J. H. Park *et al.*, "Machine learning prediction of incidence of Alzheimer's disease using large-scale administrative health data," *Npj Digital Medicine*, vol. 3, no. 1, p. 46, Mar. 2020, doi: 10.1038/s41746-020-0256-0.

[16] R. Sharma, T. Goel, M. Tanveer, C. T. Lin, and R. Murugan, "Deep-Learning-Based Diagnosis and Prognosis of Alzheimer's Disease: A Comprehensive review," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 15, no. 3, pp. 1123–1138, Mar. 2023, doi: 10.1109/tcds.2023.3254209.

[17] F. Saeed, "Applications of ML and DL algorithms in the prediction, diagnosis, and prognosis of Alzheimer's disease," *American Journal of Biomedical Science & Research*, vol. 22, no. 6, pp. 779–786, Jun. 2024, doi: 10.34297/ajbsr.2024.22.003014.

[18] I. Malik, A. Iqbal, Y. H. Gu, and M. A. Al-Antari, "Deep Learning for Alzheimer's Disease prediction: A Comprehensive review," *Diagnostics*, vol. 14, no. 12, p. 1281, Jun. 2024, doi: 10.3390/diagnostics14121281.

[19] A. Elazab *et al.*, "Alzheimer's disease diagnosis from single and multimodal data using machine and deep learning models: Achievements and future directions," *Expert Systems With Applications*, vol. 255, p. 124780, Jul. 2024, doi: 10.1016/j.eswa.2024.124780.

[20] S. Mohapatra, S. Satpathy, and B. K. Paikaray, "A machine learning approach to assist prediction of Alzheimer's disease with convolutional neural network," *International Journal of Bioinformatics Research and Applications*, vol. 19, no. 2, pp. 141–150, Jan. 2023, doi: 10.1504/ijbra.2023.132632.