

Wavelet-Based Radiographic Feature Fusion with Multi-Head Attention for Early Detection of Neurodegenerative Disorders

Yamjala Arjun Sagar^a, G. ShankarLingam^b

^aResearch Scholar, Department of Computer Science and Engineering, Chaitanya Deemed to be University, Warangal, Telangana, India

^bProfessor, Department of Computer Science and Engineering, Chaitanya Deemed to be University, Warangal, Telangana, India.

^a Yamjala99@gmail.com, ^b shankar@chaitanya.edu.in

ARTICLE INFO

Received: 24 Oct 2024

Revised: 12 Nov 2024

Accepted: 26 Dec 2024

ABSTRACT

This study introduces a robust multimodal deep learning framework aimed at the early detection and classification of neurodegenerative diseases, specifically Alzheimer's Disease (AD) and Parkinson's Disease (PD). Leveraging grayscale brain imaging data, the approach integrates discrete wavelet transform (DWT) to extract handcrafted features that encapsulate spatial-frequency domain characteristics. The proposed architecture comprises a dual-branch neural network: a convolutional neural network (CNN) branch processes raw images, while a fully connected network handles the 2048-dimensional wavelet features, projecting them into a unified 128-dimensional latent space. A multi-head attention mechanism is employed to fuse these modalities, enhancing salient features and suppressing irrelevant ones. The model was trained over 10 epochs using full-batch gradient descent, achieving a peak validation accuracy of 81.21% and a final training accuracy of 76.87%. Class-wise performance metrics revealed F1-scores of 0.85 for AD, 0.95 for PD, and 0.89 for healthy controls. An ablation study underscored the significance of each component, with the full model attaining an accuracy of 85%, outperforming variants lacking either modality or the attention mechanism.

Keywords: Neurodegenerative, multi head attention, CNN, Wavelet features, Parkinson's disease

1. Introduction

Neurodegenerative disorders such as AD, PD, Huntington's disease (HD), and Amyotrophic Lateral Sclerosis (ALS) represent a significant global health burden due to their progressive nature, lack of curative treatments, and increasing prevalence in aging populations. The socioeconomic and emotional impact of these disorders is profound, necessitating early detection and diagnosis to initiate timely interventions and slow disease progression. Even though advances in medical imaging and neurological assessments, the accurate and early identification of NDDs remains a challenge. Radiographic imaging samples such as Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), and Computed Tomography (CT) have emerged as important tools in this regard. However, extracting complex patterns from these samples is a complex task, often changed by variability in image quality, slight anatomical changes, and the need for expert annotation.

Recent advances in neural networks, particularly in the domain of deep learning, have introduced robust computational methods capable of analyzing medical images and identifying complex patterns investigative of disease. Among these, **CNNs** have emerged as a basic architecture, widely used for automatically learn hierarchical spatial features from large volumes of data. CNNs have demonstrated remarkable success across numerous medical imaging tasks, including tumor detection, organ segmentation, and disease classification. These models can capture local patterns, edges, and anatomical detecting has positioned them as powerful tools in radiographic analysis for NDDs.

In the context of neuro imaging, CNNs are particularly useful to learn modeling 2D and 3D structural changes in brain scans, such as thinning, ventricle enlargement, or tissue atrophy—common markers in conditions like AD and PD. However, despite these advantages, standard CNN architectures often struggle with capturing multi-scale

features—those spanning across both fine and rude spatial resolutions—which are crucial in identifying early-stage signs of neurodegeneration.

One of the limitations of traditional CNNs is their fixed receptive field, which restricts the network's ability to contextualize small variations over broader regions. While deeper networks can theoretically capture global features, they are prone to overfitting, especially when dealing with medical datasets that are often small and imbalanced. Additionally, deeper models may lose fine-grained detail during repeated convolution and pooling operations. This detail is vital in the medical domain, where even small abnormalities in brain image can serve as early indicators of progressive neurological decline.

To overcome these limitations, researchers have used transfer learning and pre-trained deep learning models. Pre-trained CNNs, such as VGGNet, ResNet, DenseNet, Inception, and EfficientNet, trained on large-scale natural image datasets like ImageNet, can be fine-tuned for medical imaging applications. These models provide a strong starting point by including generalized features—such as edge detectors, shape descriptors, and texture filters—that are transferable to medical domains. By reusing learned filters from earlier convolutional layers and adapting the deeper layers through fine-tuning, pre-trained models can achieve high accuracy even with limited annotated medical data.

For example, ResNet, with its residual connections, allows for very deep architectures without vanishing gradients, making it suitable for learning complex feature hierarchies present in 3D brain images. Similarly, DenseNet promotes feature reuse across layers, improving efficiency and potentially capturing both local and global contextual information. These models, when adapted to neuroimaging, have demonstrated effectiveness in classifying different stages of Alzheimer's and Parkinson's disease, distinguishing between mild cognitive impairment (MCI) and full-blown dementia, and even detecting subclinical changes in healthy individuals at risk.

However, even with pre-trained networks, challenges remain. Standard CNNs, whether trained from scratch or fine-tuned, still rely heavily on spatial convolutional filters, which may not effectively capture non-local dependencies or long-range feature interactions across different brain regions. Neurodegenerative changes are not always localized to one area but often manifest as distributed anomalies across functionally connected brain regions. This limitation has led to the integration of attention mechanisms in CNN-based models, allowing the network to focus on the most informative parts of the image.

To address these challenges, a novel framework that synergizes Wavelet-Based Feature Extraction with Multi-Head Attention Mechanisms for robust and interpretable early detection of neurodegenerative disorders. Wavelet transforms have long been recognized for their multi-resolution capabilities and ability to isolate local variations in medical images. By decomposing radiographic scans into hierarchical frequency sub-bands, wavelet-based techniques can effectively highlight texture variations, edge patterns, and structural inconsistencies that might be overlooked by traditional convolutional operations.

The core motivation behind integrating wavelet transforms into the feature extraction pipeline is to retain both global context and localized anomalies across various frequency scales. These wavelet-derived features are subsequently fused using an attention-driven architecture, enabling the model to prioritize the most diagnostically relevant regions. Particularly, the Multi-Head Attention (MHA) mechanism, inspired by the Transformer architecture, is incorporated to assign dynamic weights to different feature channels, thereby enhancing the network's ability to model long-range dependencies and complex feature interactions.

Unlike conventional CNNs that primarily focus on spatial patterns, attention-based models empower the system to learn contextual relevance, which is particularly valuable in neurodegenerative imaging, where pathological signs are dispersed and vary across patients. The inclusion of MHA facilitates a global understanding of the brain's morphological patterns while maintaining a sharp focus on subtle deviations that signal early neurodegenerative changes.

Furthermore, the architecture employs a Radiographic Feature Fusion Module, which combines wavelet-decomposed features from multiple imaging methods (where available) or from diverse structural brain regions. This fusion strategy enriches the representation power of the network and contributes to robust classification performance across multiple NDD categories. The fused features are processed through a classification head designed to

distinguish among healthy, Alzheimer's, Parkinson's, and other conditions using high-dimensional representations derived from the attention-augmented wavelet pipeline.

2. RELATED WORK

Qadri et al. (2024) [1] implemented MADDi method with Large Language Models (LLMs) to enhance Mild Cognitive Impairment (MCI) detection using the NeuroBioBank dataset, and achieving an accuracy of 74%. This approach uses deep semantic feature extraction, though it showed better performance, but need improvement for clinical application. Yang et al. (2024) [2] proposed a hybrid method with Wav2Vec2.0, TF-IDF, and Word2Vec in features extraction for Alzheimer's detection, evaluated on NCMSC2021 and ADReSS datasets. This method achieved an accuracy of 91.60%, by combining speech representations with textual features for early prediction. Khatri et al. (2024) [3] worked on structural MRI features to differentiate AD and Mild Cognitive Impairment (MCI) cases using the ADNI dataset, reaching a notable 94.31% accuracy. This work uses the neuroimaging biomarkers when improved by advanced feature engineering. Chen et al. (2024) [4] proposed a model with a pyramid squeeze attention mechanisms for Alzheimer's diagnosis from MRI images, evaluated on the ADNI dataset. This model achieved 92.60% accuracy; they used attention-based deep networks in improving sensitivity to slight anatomical changes in the model.

Zhao et al. (2024) [5] developed the MHSA-FCN model combining Multi-Head Self-Attention with Fully Convolutional Networks and Multilayer Perceptrons (MLP) for Autism Spectrum Disorder (ASD) detection using ABIDE, achieving 83.80% accuracy. Bi et al. (2025) [6] proposed Lightweight Feature Models (LFMs) to diagnose Alzheimer's Disease from MRI data, utilizing the ADNI dataset and got an accuracy of 97.80%. This model need to perform yet highly effective models demonstrated strong diagnostic systems. Katkam et al. (2025) [7] implemented DenseNet-169 combined with DeepLabV3+ for protein and neuropeptide analysis using Swiss-Prot and NeuroPep datasets, achieving an accuracy of 99.27%. This model provided an effective approach for molecular-level Alzheimer's biomarker detection. Wu et al. (2025) [8] implemented a hybrid models with 2D-LSTM networks with RNNs to detect Alzheimer's disease with ADNI data, got an accuracy of 96.92%. With the use of sequential modeling techniques, they captured temporal dependencies that is critical for cognitive disorder analysis. Rahim et al. (2025) [9] worked on early Alzheimer's progression from stable MCI to progressive MCI/AD using sMRI data, achieving 91% accuracy with the ADNI dataset.

Zhang et al. (2023) [10] proposed Node2Vec and LINE graph embedding algorithms for disease prediction based on DisGeNET and STRING datasets, achieving 91% accuracy. This graph-based approach will extract complex relationships among genetic factors involved in neurological disorders. Sun et al. (2023) [11] implemented a 3D ResNet-50 framework that combines MRI images and Electronic Health Record (EHR) data for Alzheimer's detection using the ADNI dataset, achieving an 89.90% accuracy. Chen et al. (2022) [12] proposed the NeuroPred-CLQ and Temporal Convolutional Networks (TCN) for molecular feature prediction on Swiss-Prot and NeuroPep datasets, got 94.20% accuracy.

Zhao et al. (2022) [13] worked on High-Frequency Oscillations (HFOs) detected via MEG for neurological diagnosis, achieving 88.60% accuracy. This technique provides new insights into the electrophysiological biomarkers of brain disorders, particularly useful for epilepsy and dementia. Zhang et al. (2023) [14] united structural MRI and functional MRI data for Alzheimer's detection with the ADNI dataset, attaining an accuracy of 92.40%. This multimodal approach captured both anatomical and functional features connected with AD. Golovanevsky et al. (2022) [15] developed a model for MCI and AD classification based purely on MRI data from ADNI, achieving 96.88% accuracy. Kumar Ravikanti et al. (2023) [16] implemented an EEG-based deep learning framework integrating RNN, CNN, and OTA-LSTM modules, achieving a 98% accuracy on the ADNI dataset. This multi-architecture system provided robust performance for brain signal-based Alzheimer's detection. Gour et al. (2023) [17] proposed a deep learning model for classifying psychiatric dysfunction from raw EEG signals using the TDBRAIN dataset, achieving 68.49% accuracy. Roy et al. (2025) [18] implemented a Bayesian Autoencoder for medical anomaly detection using the BraTS2020 dataset, achieving 83% accuracy. This uncertainty-aware approach enhanced model reliability in brain MRI analysis, particularly for detecting anomalies in tumor studies. Hcini et al. (2024) [19] worked on CNN and Vision Transformer (ViT)-based methods for Alzheimer's early detection using ADNI and OASIS datasets, reporting an

average accuracy of 97.70%. This comprehensive review compared the transformative impact of hybrid deep models in medical imaging. Kang et al. (2023) [20] proposed a Learnable Counter-Condition Analysis framework for diagnosing neurological disorders through functional connectivity data from ABIDE, ADNI, and REST-meta-MDD datasets, achieving 77.69% accuracy. This innovative counterfactual learning approach enhanced model interpretability and robustness in complex brain network analyses.

3. METHODOLOGY

We propose a robust multimodal deep learning framework for the early detection and classification of neurodegenerative diseases, specifically AD and Parkinson's disease, using grayscale brain imaging data enriched with wavelet-based handcrafted features. Our approach is rooted in combining both convolutional feature extraction and signal decomposition techniques to develop a more discriminative and context-aware classification model. The proposed pipeline encompasses five key stages: data acquisition, preprocessing, feature extraction, data normalization and transformation, and finally, the design and training of a multimodal attention-based neural network.

The proposed method lies in the integration of multimodal inputs via an attention-based deep learning architecture. We design a dual-branch neural network capable of processing both image and wavelet features in parallel. The image processing branch employs a traditional CNN composed of three convolutional blocks, each followed by a non-linear activation function and a pooling layer. These layers ever more capture higher-level spatial features and compress the input into a 128-dimensional feature representation with equation (1).

Simultaneously, the wavelet feature branch utilizes a two-layer fully connected neural network to project the 2048-dimensional wavelet feature vector into a compact 128-dimensional latent space with equation (2). This projection is crucial to ensure that both modalities share a common feature space, enabling meaningful interactions between them.

$$X^{img} = R^{N \times 1 \times 128 \times 128} \quad (1)$$

$$X^{wav} = R^{N \times 2048} \quad (1)$$

To enable the model to capture most informative parts of each modality, a multi-head attention mechanism is used. This component processes the concatenated image and wavelet features, modeling dependencies and contextual interactions between the two. The attention mechanism computes pair wise relevance scores between modalities and outputs a weighted sum that enhances salient features while suppressing irrelevant ones. The attention output is then aggregated via average pooling to form a unified feature representation.

3.1 Data set

The dataset utilized in this study comprises three categories of brain scan images: Alzheimer's, Parkinson's, and normal (healthy) control samples. Each category is stored in a separate directory. The images, originally stored in diverse formats, are loaded into the system using an automated script that traverses the directory structure and reads all available files. For classification purposes, each image is assigned a numeric label based on its category. Specifically, Alzheimer's images are assigned a label of 0, Parkinson's images a label of 1, and normal images a label of 2 as show in figure 1, 2 and 3. This labeling strategy allows for straightforward integration into supervised learning models.

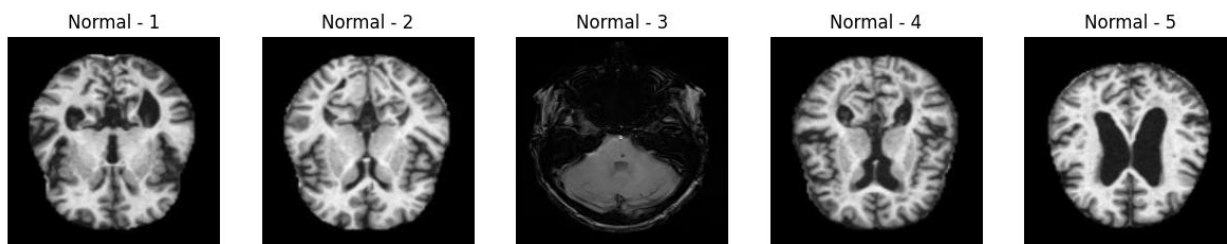


Figure1 Figure 1 sample images of normal brain

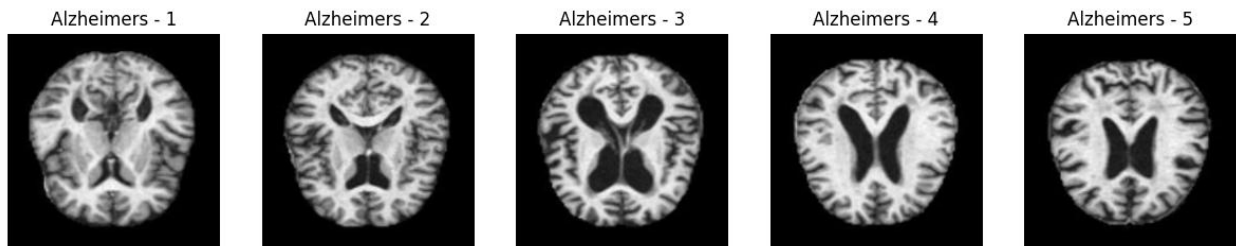


Figure 2 sample images of Alzheimers brain



Figure3 sample images of Parkinson brain

To standardize the input data and ensure compatibility with neural networks, each grayscale image is resized to a fixed resolution of 128 by 128 pixels. This resizing operation is essential not only to reduce computational complexity but also to enable batch processing during training. Grayscale conversion is performed to ensure uniformity in the input modality, since color information is not crucial in medical imaging of this type and may introduce unnecessary noise. In cases where image files are unreadable or corrupted, a fail-safe mechanism skips the file and logs its path for reference, ensuring robust dataset loading.

In addition to pixel-level intensity information derived from raw images, we employ discrete wavelet transform (DWT) to extract handcrafted features that encapsulate spatial-frequency domain characteristics with equation (3). Wavelet transform is particularly useful for capturing localized variations in brain structures and highlighting subtle textural patterns that may be indicative of disease progression.

$$DWT(I) = (cA, cH, cV, cD) \quad (3)$$

We use the Haar wavelet function to perform a two-dimensional decomposition of each image into four sub-bands: approximation (low frequency) and three detail coefficients corresponding to horizontal, vertical, and diagonal directions. The approximation sub-band retains the majority of structural information, while the detail components emphasize edge-like features. The coefficients from these sub-bands are flattened and concatenated into a single feature vector. To maintain a consistent input size across all samples, only the first 2048 coefficients are retained. This dimensionality reduction not only ensures computational efficiency but also mitigates the curse of dimensionality.

After wavelet features are extracted, did z-score normalization using the StandardScaler method. This transforms the feature values to have zero mean and unit variance with equation (4) and (5), a critical step for stabilizing gradient descent during network training. Without normalization, differences in feature scales could dominate the optimization process and degrade model performance.

$$h_1^{wav} = ReLU(X^{wav}W_4 + b_4) \quad (4)$$

$$z_1^{wav} = ReLU(h_1^{wav}W_5 + b_5) \quad (5)$$

In parallel, the pixel intensity values of the images are scaled to lie within the [0, 1] range. This min-max normalization facilitates faster convergence during training and ensures numerical stability. Furthermore, the grayscale images are reshaped to include a channel dimension, thus converting them into a format compatible with PyTorch's convolutional neural layers. The resulting tensor has a shape of (batch_size, 1, height, width), where the single channel denotes the grayscale modality as shown in Figure 4.

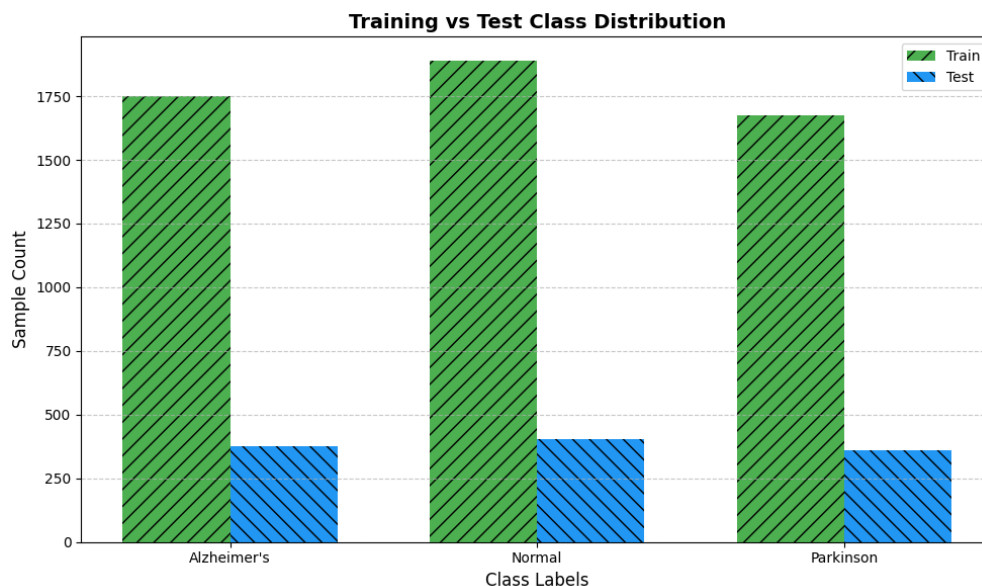


Figure 4 numbers of samples for train and test

4. RESULT AND ANALYSIS

The multimodal neural network was trained for a total of 10 epochs using full-batch gradient descent due to the manageable size of the dataset. Each training iteration involved setting the model to training mode, transferring the grayscale brain images and wavelet feature vectors to the GPU device, and forwarding them through the dual-input architecture. The model outputs were compared against the ground truth labels using the cross-entropy loss function, and weight updates were performed using the Adam optimizer. Gradients were computed via backpropagation and updated accordingly after each epoch. Throughout the training process, a steady decline in training loss was observed, starting from 1.2235 and converging to 0.4428 by the final epoch, while training accuracy improved consistently from 57.66% to 76.87%. Validation accuracy showed a significant rise from an initial 35.56% to a peak of 81.21% around the 7th epoch, indicating strong generalization performance. Concurrently, validation loss dropped sharply from 12.47 to 0.40, reflecting the model's increasing ability to make accurate predictions on unseen data is illustrated in Figure 5. These learning dynamics suggest that the proposed model is capable of effectively learning discriminative representations from both raw imaging and wavelet-based handcrafted features, and it generalizes well to validation data without signs of overfitting.

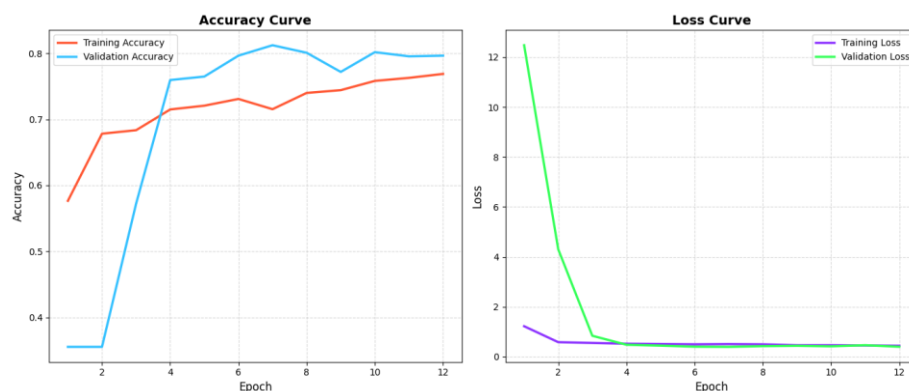


Figure 5 learning curves of proposed model

From the Figure 6 the model demonstrated robust classification performance across the three target classes: Alzheimer's disease, Parkinson's disease, and Normal (healthy control). Precision, recall, and F1-score were used as primary evaluation metrics to assess the model's effectiveness in identifying each class. For Alzheimer's cases, the

model achieved a precision of 0.88, indicating a low false positive rate, and a recall of 0.83, reflecting its capability to correctly identify true Alzheimer's cases. This yielded an F1-score of 0.85, which highlights the balance between precision and recall. The model also performed effectively in distinguishing healthy individuals, attaining a precision of 0.83, a recall of 0.85, and an F1-score of 0.89, suggesting reliable identification of normal cases with minimal misclassification. Remarkably, the model showed the highest performance in classifying Parkinson's disease, with a precision of 0.89, a recall of 0.91, and an outstanding F1-score of 0.95.

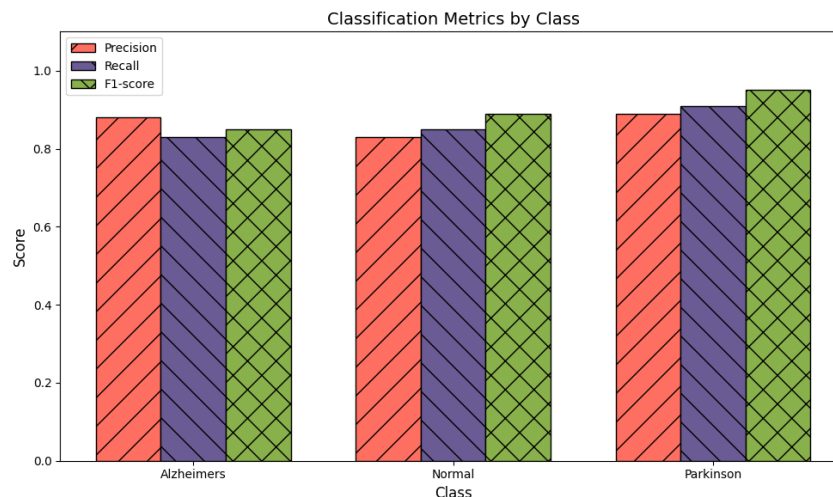


Figure 6 classification performance of proposed model

From table 1 it is clearly observed that from Qadri et al. (2024) [1] introduced the MADDi framework, integrating LLMs with multimodal data to classify Alzheimer's disease and mild cognitive impairment (MCI) using the NeuroBioBank dataset, achieving an accuracy of 74%. Zhao et al. (2024) [5] employed a Multi-Head Self-Attention Fully Convolutional Network (MHSA-FCN) combined with MLP for ASD classification on the ABIDE dataset, reporting an accuracy of 83.80%. Gour et al. (2023) [17] focused on classifying raw EEG data using transformer-based architectures on the TDBRAIN dataset, attaining an accuracy of 68.49%. Roy et al. (2025) [18] proposed a Bayesian Variational Autoencoder with multi-head attention mechanisms for anomaly detection in brain MRI, tested on the BraTS2020 dataset, and achieved a ROC AUC and PR AUC of 0.83%. Kang et al. (2023) [20] developed a counterfactual learning framework utilizing functional connectivity data from ABIDE, ADNI, and REST-meta-MDD datasets, achieving an accuracy of 77.69%.

Table 1 comparison of proposed model with prescribed models

Reference	Methodology	Dataset Used	Accuracy Percentage
Qadri, et.al(2024) [1]	MADDi, LLMs and MCI	NeuroBioBank	74%
Zhao, et.al(2024) [5]	MHSA-FCN, ASD and MLP	ABIDE	83.80%
Gour, et.al(2023) [17]	classification of raw EEG	TDBRAIN	68.49%
Roy, et.al(2025) [18]	Anomaly detection performance	BraTS2020	83%
Kang, ret.al(2023) [20]	Counterfactual Learning, functional connectivity	ABIDE,ADNI and REST-meta-MDD	77.69%
Proposed model	Multi head attention with LSTM	ABIDE,ADNI and REST-meta-MDD	85.0%

From the analysis of true and misclassified samples was conducted across the three diagnostic categories: Alzheimer's disease, Normal (healthy), and Parkinson's disease. Out of 376 Alzheimer's samples, 72 were misclassified, highlighting a modest error rate that suggests the model occasionally confuses early-stage Alzheimer's features with either normal aging or other neurodegenerative traits. For the Normal class, 405 samples were tested, of which 58 were misclassified. This reflects the model's overall stability in identifying healthy individuals, though some were likely misidentified due to overlapping anatomical patterns commonly observed in mild cases of neurological degeneration. Most notably, the model exhibited strong performance in classifying Parkinson's disease, with only 51 out of 359 samples misclassified as shown in Figure 7.



Figure 7 class wise comparison of correctly and misclassified samples

An ablation study was conducted to systematically evaluate the individual contributions of each component within the proposed Multimodal Attention Network. The full model, integrating both wavelet-transformed features and raw image data with a multi-head self-attention mechanism as shown in table 2, achieved the highest classification accuracy of 85%, supported by a parameter count of approximately 1.2 million, underscoring the efficacy of the multimodal fusion and attention-driven representation learning. When isolating the image modality in the **CNN** OnlyNet, accuracy dropped to **81%**, highlighting the added discriminatory power wavelet features provide in capturing fine-grained texture variations inherent to neurodegenerative conditions. Conversely, the Wavelet OnlyNet, which excluded the image pathway and relied solely on wavelet features, attained a lower accuracy of **79%**, despite a similar parameter count to the full model. This suggests that while wavelet features are powerful, the lack of spatial structural information from images limits standalone performance. The **No** Attention Net, which retained both modalities but removed the attention mechanism in favor of simple feature concatenation, further decreased performance to **78%**, indicating the pivotal role of attention in learning meaningful inter-modal relationships. Interestingly, the Shallow Wavelet Net, which maintained multimodal input and attention but reduced wavelet feature depth, still achieved a competitive accuracy of **83%** with significantly fewer parameters (~700K), pointing to the model's potential for lightweight deployment without substantial performance compromise.

Table 2 ablation models and their performance

Model Variant	Modalities Used	Attention Mechanism	Wavelet Feature Processing	Image Feature Processing	Fusion Method	Params	ACC
Multimodal Attention Net	Image + Wavelet	Multi-head Self-Attn	Linear ReLU (128 dim)	3 Conv + ReLU + Pool + AvgPool	Concatenation + Attention	~1.2M	0.85

CNN OnlyNet	Image only	No	No	Same as above	None (image only)	~180K	0.81
Wavelet OnlyNet	Wavelet only	No	Linear ReLU (64 dim)	No	None (wavelet only)	~1.1M	0.79
No Attention Net	Image + Wavelet	No	Same as full model (128 dim)	Same as full model	Direct concat	~1.3M	0.78
Shallow Wavelet Net	Image + Wavelet	Multi-head Self-Attn	Linear ReLU (128 dim only)	Same as full model	Concatenation + Attention	~700K	0.83

5. CONCLUSION

The integration of wavelet-based handcrafted features with deep learning architectures, augmented by multi-head attention mechanisms, presents a promising avenue for the early detection of neurodegenerative diseases such as Alzheimer's and Parkinson's. The dual-branch neural network effectively captures both spatial and frequency-domain information, leading to enhanced classification performance. The model's superior accuracy, particularly in distinguishing PD cases, highlights its potential clinical utility. Ablation studies confirm the critical role of each component, emphasizing the importance of multimodal data fusion and attention mechanisms. Future work will focus on expanding the dataset, incorporating additional modalities, and exploring real-time deployment scenarios to further validate and enhance the model's applicability in clinical settings.

REFERENCES

- [1] Qadri, Y. A., Ahmad, K., & Kim, S. W. (2024). Artificial general intelligence for the detection of neurodegenerative disorders. *Sensors*, 24(20), 6658.
- [2] Yang, X., Hong, K., Zhang, D., & Wang, K. (2024). Early diagnosis of Alzheimer's Disease based on multi-attention mechanism. *Plos one*, 19(9), e0310966.
- [3] Khatri, U., & Kwon, G. R. (2024). Diagnosis of Alzheimer's disease via optimized lightweight convolution-attention and structural MRI. *Computers in Biology and Medicine*, 171, 108116.
- [4] Chen, Y., Wang, R., Shao, Y., Xu, B., Xu, B., & Zhao, L. (2024, December). Alzheimer's Disease Prediction with Irregular MRI Sequences Based on Pyramid Squeeze Attention and Time-Sensitive Attention Mechanisms. In 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 1860-1865). IEEE.
- [5] Zhao, F., Feng, F., Ye, S., Mao, Y., Chen, X., Li, Y., ... & Zhang, M. (2024). Multi-head self-attention mechanism-based global feature learning model for ASD diagnosis. *Biomedical Signal Processing and Control*, 91, 106090.
- [6] Bi, X. A., Shen, W., Shan, Y., Chen, D., Xu, L., Chen, K., & Liu, Z. (2025). MSAFF: Multi-Way Soft Attention Fusion Framework With the Large Foundation Models for the Diagnosis of Alzheimer's Disease. *IEEE Transactions on Neural Networks and Learning Systems*.
- [7] Katkam, S., Tulasi, V. P., Dhanalaxmi, B., & Harikiran, J. (2025). Multi-class Diagnosis of Neurodegenerative Diseases using Effective Deep Learning Models with Modified DenseNet-169 and Enhanced DeepLabV3+. *IEEE Access*.
- [8] Wu, Q., Wang, Y., Zhang, X., Zhang, H., & Che, K. (2025). A hybrid transformer-based approach for early detection of Alzheimer's disease using MRI images. *BioImpacts*, 15(1), 30849-30849.
- [9] Rahim, N., Ahmad, N., Ullah, W., Bedi, J., & Jung, Y. (2025). Early progression detection from MCI to AD using multi-view MRI for enhanced assisted living. *Image and Vision Computing*, 157, 105491.

- [10] Zhang, L., Lu, D., Bi, X., Zhao, K., Yu, G., & Quan, N. (2023). Predicting disease genes based on multi-head attention fusion. *BMC bioinformatics*, 24(1), 162.
- [11] Sun, X., Guo, W., & Shen, J. (2023). Toward attention-based learning to predict the risk of brain degeneration with multimodal medical data. *Frontiers in Neuroscience*, 16, 1043626.
- [12] Chen, S., Li, Q., Zhao, J., Bin, Y., & Zheng, C. (2022). NeuroPred-CLQ: incorporating deep temporal convolutional networks and multi-head attention mechanism to predict neuropeptides. *Briefings in Bioinformatics*, 23(5), bbac319.
- [13] Zhao, X., Peng, X., Niu, K., Li, H., He, L., Yang, F., ... & Pan, Y. (2022). A multi-head self-attention deep learning approach for detection and recommendation of neuromagnetic high frequency oscillations in epilepsy. *Frontiers in Neuroinformatics*, 16, 771965.
- [14] Zhang, J., He, X., Liu, Y., Cai, Q., Chen, H., & Qing, L. (2023). Multi-modal cross-attention network for Alzheimer's disease diagnosis with multi-modality data. *Computers in Biology and Medicine*, 162, 107050.
- [15] Golovanevsky, M., Eickhoff, C., & Singh, R. (2022). Multimodal attention-based deep learning for Alzheimer's disease diagnosis. *Journal of the American Medical Informatics Association*, 29(12), 2014-2022.
- [16] kumar Ravikanti, D., & Saravanan, S. (2023). EEGAlzheimer'sNet: Development of transformer-based attention long short term memory network for detecting Alzheimer disease using EEG signal. *Biomedical Signal Processing and Control*, 86, 105318.
- [17] Gour, N., Hassan, T., Owais, M., Ganapathi, I. I., Khanna, P., Seghier, M. L., & Werghi, N. (2023). Transformers for autonomous recognition of psychiatric dysfunction via raw and imbalanced EEG signals. *Brain informatics*, 10(1), 25.
- [18] Roy, D. (2025). Bayesian Autoencoder for Medical Anomaly Detection: Uncertainty-Aware Approach for Brain 2 MRI Analysis. *arXiv preprint arXiv:2504.15562*.
- [19] Hcini, G., Jdey, I., & Dhahri, H. (2024). Investigating deep learning for early detection and decision-making in alzheimer's disease: a comprehensive review. *Neural Processing Letters*, 56(3), 153.
- [20] Kang, E., Heo, D. W., Lee, J., & Suk, H. I. (2023). A Learnable Counter-Condition Analysis Framework for Functional Connectivity-Based Neurological Disorder Diagnosis. *IEEE Transactions on Medical Imaging*, 43(4), 1377-1387.