Minimax Theory and its Applications

Volume 11 (2026), No. 4s, 45 to 63

AMGFormer: A Multimodal Attention Graph Former Framework for Parkinson's Disease Classification Using Enhanced MRI Images

Hima Bindu Paka[1], Dr.Venkanna Isampalli[2]

[1]Research Scholar, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Bowrampet, Hyderabad-500043, Telangana, India. Email: 2522031016@kluniversity.in
[2]Assistant professor , Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Bowrampet, Hyderabad-500043, Telangana, India. Email: dr.venkannaisampalli@klh.edu.in

Abstract
Parkinson's disease (PD) is a progressive neurodegenerative disorder whose early diagnosis using magnetic resonance imaging (MRI) remains challenging because of subtle structural abnormalities and limited labeled medical data. Existing approaches, including radiomics-based texture analysis, multimodal MRI–genetic fusion models, and graph neural network frameworks, have demonstrated promising performance but often depend on handcrafted features, multimodal clinical data, or limited relational modeling capabilities. To address these limitations, this study proposes an Attention-guided Multimodal Graph Former (AMGFormer) for MRI-only PD classification. MRI images from the Parkinson Multi-Modal Dataset 2.0 are first preprocessed through resizing and intensity normalization and subsequently enhanced using an Adaptive Neuro-Enhancement Filter (ANEF) to improve disease-relevant anatomical structures. Multimodal representations are extracted by combining Gray Level Co-occurrence Matrix (GLCM) and Local Phase Quantization (LPQ) texture descriptors, structural morphological features, and deep semantic embeddings learned through a Medical Masked Autoencoder (MedMAE) using self-supervised learning. The fused features are transformed into a similarity graph and processed using a Graph Transformer with multi-head attention to capture complex inter-feature dependencies. The framework was implemented using Python 3.11, PyTorch 2.3, CUDA 12.1, and Scikit-Learn. Experimental results achieved 96.80% accuracy, 96.60% precision, 97.40% recall, 96.50% F1-score, and 97.00% AUC. Five-fold cross-validation yielded a mean accuracy of 96.80% with a standard deviation of 0.27, demonstrating robustness and superior performance over existing MRI-based and multimodal PD diagnostic approaches.

Keywords—Parkinson's Disease Classification; Magnetic Resonance Imaging (MRI); Graph Transformer; Self-Supervised Learning; Medical Masked Autoencoder (MedMAE); Multimodal Feature Fusion; Adaptive Neuro-Enhancement Filter (ANEF); Attention Mechanism

Introduction

Parkinson's disease (PD) is a neurodegenerative disorder which affects movements and cognitive abilities as well, therefore, early diagnosis is necessary so that treatment can be initiated promptly. MRI can be used as a non-invasive tool to detect PD; however, the slight structural differences between normal and diseased brains make the classification hard. MRISR is among the recent super-resolution MRI techniques that help generate images of high resolution and preserve structural features[1]. Preprocessing methods such as equalization and homomorphic filtering can contribute significantly to the increase in tumor classification accuracies on the basis of MRI scans, proving the relevance of image processing as a step before classifying the images[2]. In addition to image processing, representation learning is one of the key components in modern medical imaging systems. It has been proven by some self-supervised learning (SSL) architectures like BrainFound that slicing images can serve as an efficient way to represent MRI images, which does not require big labeled datasets[3]. The problem of small labeled data sets can often be faced when creating models in the medical domain. Besides, SSL methods like DINO can show outstanding results in feature extraction while being trained on very few labeled samples[4]. Even more powerful deep learning (DL) models are now being used in trustworthy PD diagnosis by implementing surrogate loss functions and self-supervised representations[5].

Recent advances in the field of PD using DL algorithms include further improvements made on top of volumetric foundation models for analysis using MRIs[6]As an example, an advanced YOLOv5-based PD detector framework implements attention mechanisms along with enhanced feature extraction modules to achieve high-level performance for detecting PD patients' MRIs [7]. Also, FCN-PD utilizes a framework which

uses the combination of EfficientNet's spatial feature extraction and attention mechanisms to learn local and global representations, leading to high-level accuracy in classifying multiple PD datasets from MRIs[8]. Hybrid architectures including Convolutional Neural Networks and Vision Transformers, such as a recent Swin Transformer-based algorithm, further augment the capabilities of a system by integrating local texture representations with global context representations[9].Further, in the area of multimodal and interpretable frameworks for PD classification using MRIs, transformer-based multimodal feature fusion frameworks prove to be effective in combining complementary representations in MRI data[10].Multimodal representation fusion along with explainable AI in PD classification further enhances classification accuracy while determining disease-related brain regions of early-stage PD development [11]. In addition to the multimodal DL architectures incorporating both neurological and physiological information, the above results have proven that the combination of multiple complementary representations boosts the detection performance[12].

Although these successes are brought about through modern processes, the implementation of these processes is heavily reliant either on deep representations alone or on additional clinical data from multiple modalities other than MRI scans. Consequently, the need arises to develop an all-encompassing MRI-based pipeline that will encompass image enhancement, feature extraction, unsupervised feature learning, and relational modeling using graph neural networks (GNNs). In this regard, this study proposes the development of the AMGFormer framework incorporating ANEF enhanced MRIs, feature extraction, MedMAE, feature fusion, and Graph Transformers.

Research Motivation

Some recent studies have pointed out the importance of utilizing graph and multimodal learning for PD analysis. The role ofGNN for modeling complicated relations between various clinical and non-clinical factors involved in PD has been extensively studied recently, exhibiting promising performances in monitoring and predicting PD. Moreover, using Vision GNNs along with contrastive learning has proved the efficacy of graph representations for discovering discriminative spectro-temporal features and improving the detection accuracy of PD from speech signals. Also, some recent multimodal learning studies have revealed that by incorporating handwriting, gait, and speech information, it is possible to diagnose PD accurately due to its heterogeneous nature which calls for combining complementary feature representations. Recently, some DL methods based on midbrain MRI biomarkers have also yielded promising results in automating the PD classification task,demonstrating the importance of imaging-based representations in diagnosing PD. Nonetheless, most previous works have used non-MRI data or complex multimodal inputs which limits their scalability to clinical imaging settings. Therefore, there is a great demand for a general MRI-based graph learning approach to handle the modeling of intra and inter-features' relationships efficiently.

Research Significance

However, PD diagnosis remains challenging because of the small differences in brain MRI structures as well as due to the drawbacks of the conventional approaches used for the purpose. The contribution of this research consists in presenting an MRI-based DL model, which includes image enhancement, self-supervised representation learning, and graph-based reasoning. In contrast to current methods that largely depend on multiple data sources (such as imaging and other clinical information) or utilize unimodal deep features exclusively, this research suggests a novel approach based on combining the MRI enhanced with ANEF, texture/structural descriptors, and SSL through MedMAE. Besides, attention-based fusion and the utilization of Graph Transformers can be helpful when modeling complicated connections between different features. This study is important as it addresses MRI-based PD diagnosis by eliminating reliance on additional modalities and increasing feature interpretability and generalization across different data sets.

Problem Statement

Recent PD diagnostic systems show excellent results by utilizing multimodal fusion learning and graph-based learning algorithms. Despite their superior performance, these systems pose some critical limitations that hinder their practicality and generalizability. The multimodal fusion learning technique, which involves fusion between MR imaging and SNP genomics, offers a high level of diagnostic accuracy, but it requires more effort to acquire the additional genomic data, which limits its applicability in clinical settings where only MRI data is available[13].Likewise, the graph neural network-based systems, which use the radiomics similarity network and metabolic brain imaging, offer effective modeling of the relationship between different regions of the brain. Unfortunately, these systems have limitations in using specific types of imaging and do not consider unified feature learning based on standard MRI images [14]. Moreover, these approaches do not incorporate feature

learning through enhancement techniques and other hand-crafted texture descriptions, which causes limitations in feature representation.

Recent Innovation and Challenges

Current developments in PD diagnosis methods are characterized by increased utilization of graph-based and contrastive learning approaches for processing raw data from neuroimaging and speech modes directly. Comparative evaluation of the performance of convolutional, transformer, and graph convolutional architectures using MRI data demonstrates that representations based on relations are superior to the ones generated solely by convolution-based models [15]. Moreover, explainable CNN methods based on a multimodal cohort of large MRI data prove that micro-structural features of diffusion have higher significance for diagnostics than standard T1-weighted images [16]. At the same time, application of vision graph neural network techniques with contrastive learning has been shown to effectively learn spectro-temporal dependencies within non-imaging modalities, such as speech [17]. Finally, comparative deep learning techniques on gray matter MRI data demonstrate the problem of overfitting even in cases of limited labeled samples [18]. It becomes evident that graph-based representations and contrastive learning techniques have become crucial for diagnosing PD; however, issues of designing generalizable pipelines for MRI alone, capable of capturing texture and structure dependencies without large-scale multimodal or genomic datasets, persist.

Key Contribution

- Proposes AMGFormer, a unified MRI-only framework integrating enhancement, handcrafted features, SSL, and graph modeling.
- IntroducesANEF to suppress noise and enhance disease-relevant anatomical MRI structures.
- Integrates GLCM/LPQ texture, structural descriptors, and MedMAE deep embeddings into a unified feature graph.
- Employs a Graph Transformer (4 layers, 8 heads) with attention-guided fusion for inter-feature dependency modeling.
- Achieved 96.80% accuracy, 96.60% precision, 97.40% recall, 96.50% F1-score, and 97.00% AUC, outperforming existing baselines.

Rest of the Paper

The rest of this paper is organized as follows: Section II gives a literature review for previous work on MRI-based and multimodal PD diagnosis; Section III discusses the proposed architecture of AMGFormer for feature enhancement, multimodal feature extraction, graph building, and fusion; Section IV shows the experimental results; finally, conclusions are given in Section V.

Related Works

Yang et al. [13]presented a machine learning algorithm that integrated modalities of MRI imaging and SNP genetic information gathered from the PPMI database. The main purpose of the work was to achieve greater precision of PD detection by utilizing the advantages of imaging and genetic biomarkers. This was achieved by implementing feature-level methods based on Fisher discrimination analysis, Lasso regression, and PLS regression combined with decision-level ensemble stacking approach to boost accuracy. Results showed a balanced accuracy of 95.36% and AUC value of 0.974. Furthermore, the work managed to detect brain structures associated with the disease and important genetic indicators. On the other hand, the dependence on multimodal databases adds computational complexity and makes the model unsuitable for practical implementation in the field of MRI-based diagnostics.

Ling et al. [14]introduced an explainable graph neural network framework for the prediction of Parkinsonian syndrome based on metabolic brain imaging. The authors developed a Regional Radiomics Similarity Network (R2SN), which can model complex interactions between different brain regions. Transformer-based attention mechanisms were used to enhance the performance and interpretability of the model. To evaluate the method, a dataset with a large number of subjects was considered, including individuals affected by various forms of Parkinsonian syndromes and healthy subjects. The experiments showed the high diagnostic efficiency of the developed approach, with achieved F1-score of up to 96.3% for the classification of idiopathic PD. Moreover, the use of the attention mechanism allowed visualization of brain connections related to the disease. Nonetheless, the proposed framework does not consider MRI scans and relies heavily on the usage of SSL and feature extraction methods specific to metabolic imaging.

Gore et al. [19]presented a radiomic approach for PD classification based on structural MRI data that makes use of texture features. In the current work, the researchers used two different types of MRI data, namely T1-weighted images and resting-state MRIs, collected from the Neurocon and Tao-Wu datasets, with 72 subjects in total. The proposed methodology involved extracting texture biomarkers using different LBP methods tailored specifically to detect structural differences related to PD. The process of selecting relevant features was implemented via recursive feature elimination technique, whereas classification was carried out by using SVM and Random Forest classifiers. The presented framework provided accuracy of 83.33% and had a precision and recall of 84.62% and 91.67%, respectively. The conducted experiments revealed the potential of texture descriptors in detecting PD signatures from MRI images. Nonetheless, the presented radiomics pipeline did not have any feature learning mechanisms and could not properly capture global spatial relationships between different regions of MRI representations.

Cao & Long,[20]presented the Residual Attention Contrastive Fusion (RACF) approach for PD diagnosis based on SNP genetic data and MRI scans collected from the PPMI dataset. The paper was focused on improving the efficiency of disease diagnosis through the combination of different types of data sources. To achieve that goal, a GWAS-Transformer was used for feature extraction together with attention-based contrastive fusion to incorporate more complex intermodal dependencies. Combining both imaging and genetic modalities made it possible to identify discriminatory patterns related to the disease. The accuracy and AUC values reached 91.2% and 0.94 respectively; moreover, certain genetic biomarkers associated with the progression of the disease were found. Nevertheless, this model relied heavily on genetic data and required great computational costs, which made it non-applicable in clinical practice where only MRI is used.

Oliveira et al. [21]suggested a texture analysis framework using structural MRI data from the ADNI database based on the GLCM to classify Alzheimer's disease. Specifically, the research was conducted with the extraction of 22 handcrafted texture features using the GLCM in different anatomical planes. Different classical machine learning methods were used to test the ability of the feature set to differentiate between patients with different diseases. The results of experiments revealed good performance, with an accuracy of up to 98.5%. In particular, handcrafted features were successfully used in neurodegenerative disease classification based on MRI data. However, the texture analysis framework lacked DL features and spatial/graph structure modeling capabilities.

Kibriya et al. [22]introduced an ensemble-based framework for brain tumor classification through the fusion of handcrafted texture and DL features from MRI images. In their approach, GLCM texture features were used along with features derived from the VGG16 architecture in order to obtain local and global representations respectively. Further, these two types of features were used to train SVM and KNN machine learning models. As a result of this experimental investigation, high accuracy was achieved, indicating the capability of feature fusion in medical image processing applications. The authors stressed the importance of the integration of different types of features to improve diagnosis results; however, no use of SSL and graph approaches was made to model feature dependencies. Thus, the presented methodology is not applicable for modeling complex diseases such as neurodegeneration and PD classification.

Han et al. [23]introduced a hypergraph Transformer architecture for brain disease diagnosis that combines hypergraph learning with feature modeling based on Transformers. This architecture made use of functional and structural brain networks in order to learn hypergraphs that could capture higher-order interactions between brain regions. Then, a Transformer was used to integrate and process these interactions in order to classify brain diseases accurately. Evaluation of this approach using the ABIDE and ADNI datasets revealed its superiority over traditional graph learning approaches. It emphasized the usefulness of graph learning in combination with attention mechanisms to model brain networks. Additionally, it demonstrated better feature and classification abilities for neurological disorders. Nevertheless, the proposed method did not directly address the problem of extracting features from MRI scans to classify PD.

Dheepak et al.[24]presented a hybrid feature extraction algorithm for brain tumor classification in the context of MRI images. In their approach, GLCM, Local Binary Pattern (LBP), and interaction features have been used to incorporate the various texture characteristics of the tumor regions. These handcrafted features have been used in combination to increase discrimination power and finally classified by employing a linear SVM classifier. High classification performance indicated the efficiency of the proposed technique, which utilizes a combination of different types of textures in order to diagnose brain tumors through MRI. However, it should be noted that their proposed system did not employ any feature extraction algorithm based on DL techniques, nor any graphical relation between features for diagnosing PD.

The reviewed literature highlights several key insights about the use of texture-based radiomics, DL, multimodal fusion, and graph-based learning in the diagnostic process of PD and other neurodegenerative disorders. Even though handcrafted algorithms like LBP and GLCM have very powerful discrimination of local features, they lack global context knowledge and scalability. The advantages of DL and transformer-based models include their capability to enhance representation learning; however, they need multimodal or big data samples. GNNs and hypergraph transformers successfully incorporate complicated relationships in the brain, but they neglect MRI-based texture and structural information. Consequently, all reviewed models have specific limitations in terms of enhancements and unified graph construction in MRI-only data, which motivates our approach.

Table 1 . Comparison Table

Author	Methodology	Dataset Used	Advantages	Disadvantages
Yang et al. [13]	Feature fusion, stacking ensemble	PPMI cohort MRI + SNP data	High accuracy and biomarker discovery	Requires multimodal genetic data
Ling et al. [14]	Explainable GNN, Transformer R2SN	Metabolic brain imaging dataset	High F1-score, interpretable model	Not MRI-based, no SSL learning
Gore et al. [19]	LBP texture features, SVM, RF	Neurocon, Tao-Wu MRI datasets	High recall for PD detection	No DL, limited scalability
Cao & Long,[20]	GWAS-Transformer, RACF fusion	PPMI dataset	High AUC, identifies genetic markers	Requires SNP data, high complexity
Oliveira et al. [21]	GLCM texture + ML classifiers	ADNI structural MRI dataset	High accuracy for AD classification	No DL or graph modeling
Kibriya et al. [22]	GLCM + VGG16 ensemble features	MRI brain tumor datasets	Very high classification accuracy	No SSL or graph modeling
Han et al. [23]	Hypergraph Transformer networks	ABIDE, ADNI brain networks	Captures high-order correlations	Not MRI texture-based PD data
Dheepak et al. [24]	GLCM, LBP, SVM classification	Brain tumor MRI datasets	Very high accuracy, strong features	No DL or graph learning

Table I presents a comparative analysis of some latest research on PD and other neurodegenerative disorders classification using MRI along with biomedical information. It clearly demonstrates the variety of methods used from conventional handcrafted texture-based techniques such as LBP and GLCM to state-of-the-art approaches involving DL models, transformers, and GNNs. It also describes the types of data sources employed which have involved benchmark databases including PPMI, ADNI, and specific MRI datasets for particular domains. Besides, the important benefits include high accuracy, interpretability, and biomarker identification while some of the limitations are scalability issues, requirement of multimodal data, lack of SSL, and MRI-focused graph-based modeling.

AMG Former: A Multimodal Attention Graph Former Framework for PD Classification Using Enhanced MRI Images

The novel architecture known as AMGFormer is developed with the aim of enhancing the accuracy of PD diagnosis from MRI images. The model takes into account MRI data provided by the Parkinson Multi Modal Dataset 2.0 and analyzes the images within several sequential stages in order to extract all necessary information about the disease. First, image preprocessing takes place which aims at cleaning MRI images and making sure their quality is consistent. Second, an ANEF is used to further improve the contrast of the images and draw

attention to potential structural anomalies that can indicate the presence of PD. Enhanced MRI images are then processed by means of multimodal feature extraction, where hand-crafted texture descriptors such as GLCM and LPQ are supplemented with structural features and deep features created via MedMAE SSL. This set of features is then fused using the attention mechanism to focus on the most informative ones that characterize PD. Then, thanks to graph construction, Graph Transformers will be able to detect both local and global dependencies of these features. The learned graph representations will be helpful in further analysis made by the classification module.

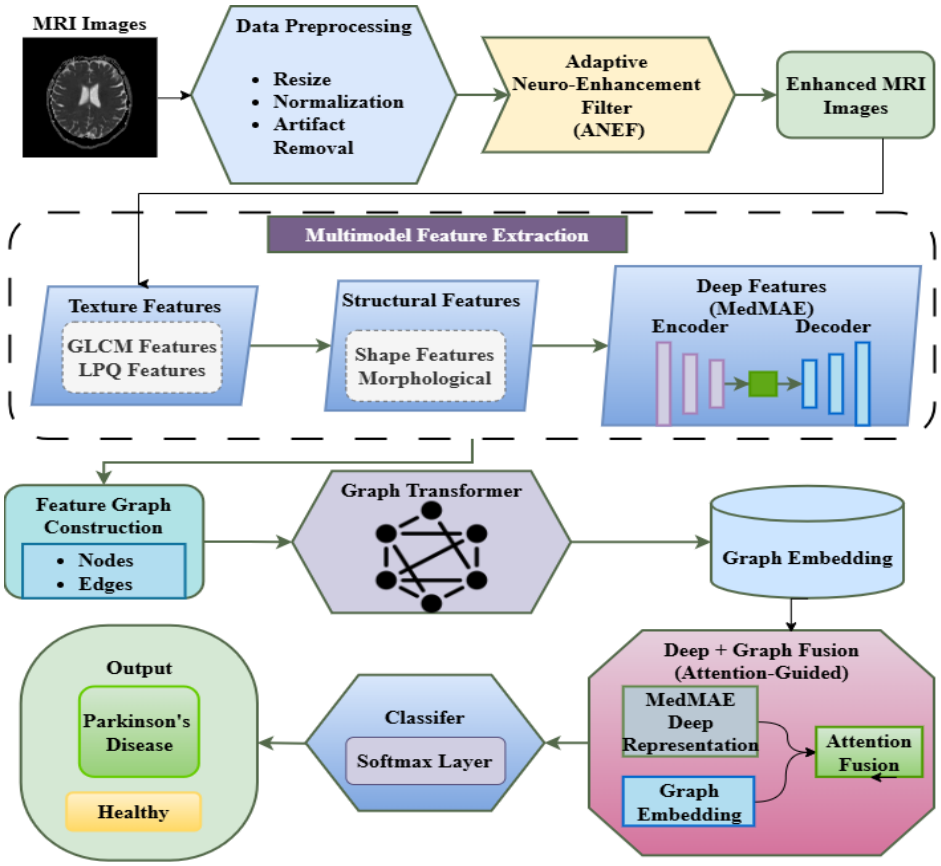


Fig 1. Block Diagram of the Proposed AMGFormer Methodology

Fig. 1.depicts the entire process of AMGFormer algorithm for the diagnosis of PD using MRI images. The entire process begins with the step of acquiring the images and further continues with preprocessing, followed by ANEF. The process of multimodal feature extraction takes place using texture, structure, and MedMAE deep representation features. The features are mapped to a graph form and then subjected to the process of Graph Transformer. The final step is that of Softmax classification.

Data Collection

The MRI images used in this research were obtained from the Parkinson Multi Modal Dataset 2.0[25] available on Kaggle. Despite the availability of multimodal data such as speech recording, handwritten sample, DAT scan, MRI scan, and physiologic signal, the present work only uses MRI 1 Healthy and MRI 1 Unhealthy data to create an MRI-based PD classification system. The MRI dataset used in this project has 610 HC MRI images and 221 Parkinson's disease MRI images. Thus, there will be 831 images altogether that will be used for binary classification. In the current study, PD and HC are two categories considered. MRI is chosen since this modality is capable of providing detailed anatomy of brain regions affected by neurodegeneration caused by PD. Moreover, MRI-only modality allows investigating whether enhanced image pre-processing, self-supervised representation learning, and graph-based modeling can provide reliable PD detection without other modalities such as speech, handwriting, genetics, and physiologic data.

Data Preprocessing

The images obtained from MRIs from different subjects may vary in terms of intensity levels, resolution, and quality. Hence, a pre-processing step is carried out to normalize the images in order to enable effective feature extraction.

Image Resizing

All images are made to have a uniform spatial size to simplify computation as well as allow for effective feature extraction. Normalizing the sizes of images allows for efficient feature extraction and deep representation learning.

Intensity Normalization

Intensity normalization of the MRI images is done in order to normalize differences between subjects and normalize image pixel values between 0 and 1. The formula for image intensity normalization is shown below in equation (1):

$$I_{norm} = \frac{I - I_{min}}{I_{max} - I_{min}} \qquad (1)$$

Where,Irefers to image intensity before normalization, whereas$I_{min}$and I_{max}refer to the minimum and maximum pixel intensities of an MRI image respectively, and $I_{norm}$is the normalized image.

Quality Assessment and Artifact Removal

Images that have artifacts, distortions, and even corrupted content are excluded from the dataset as a preprocessing step. It will result in enhancing the dataset quality and avoiding any misleading feature extraction.

Adaptive Neuro-Enhancement Filter (ANEF)

In order to make visible the anatomy related to the diseases; the following process is applied to the normalized MRI images. MRI images often have noise as well as low-contrast content that can conceal some subtle changes due to PD. The introduced method for enhancing the anatomical images consists of adaptive filtering combined with contrast enhancement technique. For determining the strength of enhancement, the local neighborhood statistic is used. It follows from Equation (2):

$$I_e\,(x, y) = I(x, y) + \alpha[I(x, y) - \mu_N\,] \qquad (2)$$

Where,$I_e(x, y)$represents the enhanced MRI image, $I(x, y)$denotes the original pixel intensity, $\mu_N$is the mean of the neighborhood area, and αdetermines the amount of enhancement.

Multimodal Feature Extraction

These enhanced MRI image representations use the following three categories of image features: texture features, structural features, and deep features. The multimodal feature extraction allows for the extraction of local texture details, structure information, and higher-level semantic information from the MRI image at the same time.

Texture Feature Extraction

Texture features are extracted using GLCM and LPQ. GLCM describes the second-order statistics between neighbor pixels and gives details about the texture homogeneity and texture spatial arrangement. An important GLCM texture feature, contrast, is extracted using (3):

$$\text{Contrast} = \sum_{i,j}(i - j)^2 P(i, j) \qquad (3)$$

Where,$P(i, j)$is the probability value in the Gray Level Co−occurrence Matrix at(i, j).

Structural Feature Extraction

Anatomical changes linked to PD have certain characteristics. Shape features and morphometrics are computed from enhanced images of MRIs for quantification purposes. Such features include shape-related features like area, perimeter, compactness, and morphological features of interest that may be associated with PD neurodegenerative diseases.

Deep Feature Learning Using MedMAE

Deep features are extracted from medical photos using MedMAE. Unlike previous learning techniques, MedMAE employs SSL, which involves masking picture patches before the model learns to reconstruct the missing portions. Without a large amount of training data, this technique enables the model to learn latent features. The retrieved embeddings contain information regarding medical image semantics and global context.

Feature Graph Construction

After extracting texture, structural, and deep features, a combination of all features is formed to make a feature graph. Each node in the graph consists of a feature vector, whereas edges in the graph depict similarities among features. A similarity matrix can be constructed using equation (4):

$$A_{ij} = \exp\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right) \qquad (4)$$

Where,A_{ij}refers similarity measure between iand j nodes, $x_i$and $x_j$are feature vectors, and σis a similarity scaling factor.

Graph Transformer Learning

Feature graph construction results are processed by the Graph Transformer, which takes into account both local and non-local feature interactions. In contrast to traditional GNNs that focus on aggregation of neighboring features, Graph Transformer utilizes self-attention functions for modeling global dependencies between nodes in graphs. Self-attention mechanism is formulated in (5):

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \qquad (5)$$

Where,Q, K, and Vare query, key, and value matrices, respectively, and $d_k$is the dimensionality of the key vectors.

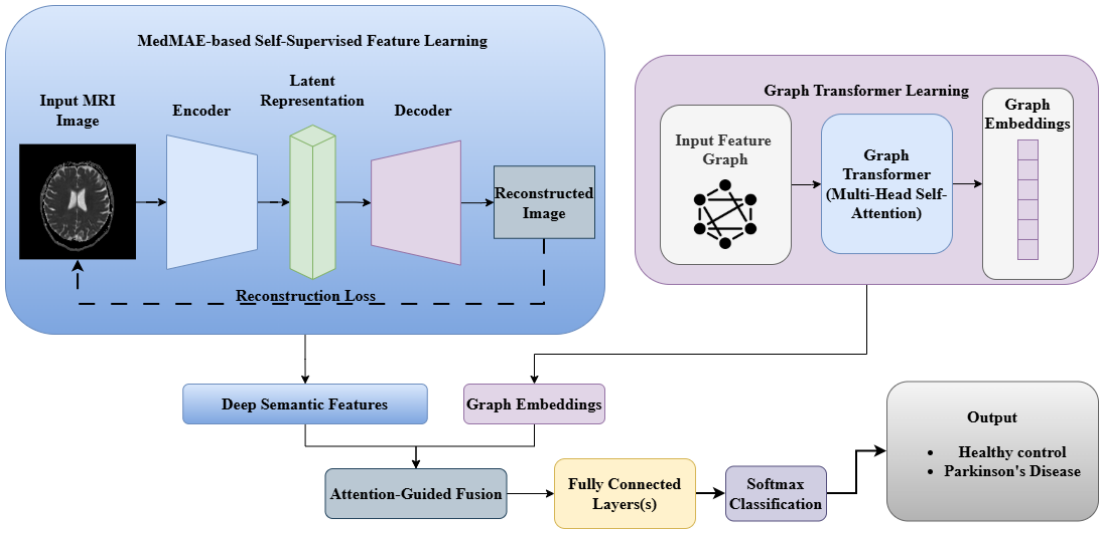


Fig 2. Internal Architecture of MedMAE and Graph Transformer-Based AMGFormer

Fig. 2.represents the architecture of AMGFormer framework. The proposed model includes MedMAE, which learns self-supervised representations by masking and restoring parts of MRI images and generating deep semantic features. At the same time, Graph Transformer models relationships between features by using multi-head self-attention and generating graph embeddings. Attention-guided feature fusion module combines deep semantic features with graph embeddings, thus creating comprehensive feature representation. Finally, the resulting representation is processed by Fully Connected layers and Softmax function for subject classification as HC or PD patients.

Deep–Graph Feature Fusion and Classification

Graph embeddings created by Graph Transformer are fused with deep semantic features created by MedMAE using attention-guided fusion. Thus, semantic information generated from MRI images and relational information modeled using graph embedding are combined. The process of fusion is described in (6).

$$F_{fusion} = \beta F_{deep} + (1 - \beta) F_{graph} \quad (6)$$

Where, F_{deep} is the MedMAE deep features, F_{graph} is the Graph Transformer representations, and β is the fusion coefficient that is learned during training. The fusion representation is then used in fully connected layers followed by Softmax activation for final PD prediction.

$$P(y) = \text{Softmax}(WF_{fusion} + b) \quad (7)$$

Where, $P(y)$ denotes the probability, W is the weight matrix, b is the bias vector, and F_{fusion} is the fusion representation.

Algorithm 1: AMGFormer_PD_Classification

Input: MRI image I, trained model parameters θ

Output: Predicted class label (Healthy / PD)

Step 1: Image Acquisition and Preprocessing

Read MRI image I from Parkinson Multi-Modal Dataset 2.0

Resize I to 224 × 224 pixels

Compute normalized image I_norm

IF I contains severe artifacts, distortion, or corrupted data THEN

Discard image I (remove from dataset)

ELSE

Proceed to Step 2

END IF

Step 2: Adaptive Neuro-Enhancement Filtering (ANEF)

Compute local neighborhood mean μ(x,y) for each pixel

Compute enhanced image I_enh

IF contrast(I_enh) < contrast(I_norm) THEN

Adjust enhancement strength parameter α and recompute I_enh

ELSE

Proceed to Step 3

END IF

Step 3: Multimodal Feature Extraction

Extract texture features F_texture using GLCM and LPQ

Extract structural features F_structural

Extract deep features F_deep using MedMAE self-supervised encoder

IF F_texture, F_structural, F_deep are all successfully extracted THEN

Concatenate into combined feature set F = {F_texture, F_structural, F_deep}

ELSE

Re-extract missing feature category

END IF

Step 4: Feature Graph Construction

FOR each pair of feature vectors (v_i, v_j) in F DO

Compute similarity A_ij

IF A_ij ≥ similarity threshold τ THEN

Create edge between node i and node j

ELSE

No edge (nodes remain unconnected)

END IF

END FOR

Construct feature graph G(V, E)

Step 5: Graph Transformer Learning

FOR each Graph Transformer layer (1 to 4) DO

Compute multi-head self-attention using Q, K, V

Update node embeddings G_emb

END FOR

Step 6: Attention-Guided Deep–Graph Fusion

Compute fused representation F_fused = β·F_deep + (1−β)·G_emb

Step 7: Classification

Compute class probability P = Softmax(W·F_fused + b)

IF P(Parkinson's) > P(Healthy) AND P(Parkinson's) ≥ classification threshold (0.5) THEN

Output label = "Parkinson's Disease"

ELSE IF P(Healthy) > P(Parkinson's) AND P(Healthy) ≥ classification threshold (0.5) THEN

Output label = "Healthy"

ELSE

Flag as "Uncertain — recommend re-evaluation"

END IF

Step 8: Performance Evaluation

IF evaluation mode = "cross-validation" THEN

Repeat Steps 1–7 for each of 5 folds and compute mean Accuracy, Precision, Recall, F1-score, AUC

ELSE

Compute single-split Accuracy, Precision, Recall, F1-score, AUC on test set

END IF

End Algorithm

The process flow of the algorithm AMGFormer_PD_Classification for diagnosing PD from MRI imaging data consists of several stages. The first stage involves pre-processing, including rescaling, normalizing, and removing artifacts, and adaptive neuro-enhancing filtering. Next, texture, structural, and MedMAE deep features are calculated and united in a joint representation. A similarity graph is generated based on this feature set and is processed by the Graph Transformer model. Then, feature fusion guided by attention and a final Softmax classification determine if the person has PD or is healthy.

Results and Discussion

This section describes the experimental results obtained during the experimental evaluation of the proposed AMGFormer approach for PD classification with improved MRI images taken from the Parkinson Multi-Modal Dataset 2.0. In particular, the aim of the experimental work is to assess the impact of integration of ANEF, texture/structural feature extraction, MedMAE-based self-supervised representation learning, Graph Transformer model, and attention-guided feature fusion within the proposed AMGFormer framework on its performance. In the current study, the AMGFormer framework is realized by Python 3.11, PyTorch 2.3, CUDA 12.1, and Scikit-Learn libraries. Improved MRI images from the Parkinson Multi-Modal Dataset 2.0 have been used to perform the experimental evaluation. Specifically, all images are first re-sampled to be of 224 × 224 resolution with values normalized to the range [0,1]. The ANEF algorithm is then applied before starting feature extraction. Further, the dataset is split into training (70%), validation (10%), and test (20%) sets via stratified sampling.

Table 2. Simulation Parameters

Parameter	Value
Dataset	Parkinson Multi-Modal Dataset 2.0
MRI Classes	Healthy, Parkinson
Image Resolution	224 × 224
Enhancement Method	ANEF
Deep Feature Extractor	MedMAE
Graph Model	Graph Transformer
Feature Dimension	512

Batch Size	16
Optimizer	AdamW
Learning Rate	0.0001
Weight Decay	1×10^{-5}
Epochs	100
Activation Function	GELU
Graph Transformer Layers	4
Attention Heads	8
Loss Function	Cross-Entropy
Validation Strategy	5-Fold Cross Validation
Hardware Platform	NVIDIA RTX 4090 GPU

Table II shows the set of parameters used in our simulation process. The chosen configuration helped us learn the MRI representation effectively while remaining efficient. In terms of model architecture, Graph Transformer was set to have four layers and eight heads to learn the intricate relationships between features. In addition, the use of the AdamW optimizer and GELU activation function made sure that we achieve stable convergence during training. This parameter setting allowed us to perform feature learning and classification more efficiently.

Dataset Characteristics

The Parkinson Multi-Modal Dataset 2.0 includes several biomedical modalities such as speech samples, DAT scans, handwriting examples, MRIs, and physiologic data. However, this research solely focused on MRI-1 Healthy and MRI-1 Unhealthy images to create an MRI-based PD diagnosis approach. MRI images included in this project pertain to normal subjects as well as PD patients, allowing for binary classification.

Table 3. Characteristics of MRI Dataset Used

Characteristic	Description
Dataset Name	Parkinson Multi-Modal Dataset 2.0
Modality Used	MRI Images
Classes	Healthy, Parkinson
Classification Type	Binary Classification
Image Format	PNG
Preprocessing	Resizing, Normalization,ANEF Enhancement
Feature Categories	Texture, Structural, Deep Features

Table III gives the details on the features associated with the MRI dataset used for identifying PD patients. The data set consists of various modalities; however, in this study, only MRI images used to preserve an imaging-based diagnostic system approach. The data set facilitates a binary classification approach to distinguish between healthy patients and those suffering from PD. Preprocessing tasks like rescaling, normalizing, and enhancing the images have been performed before extracting features.

Performance Evaluation of AMGFormer

The classification results were analyzed in terms of Accuracy, Precision, Recall, F1-Score, and Area Under AUC. All these measures measure the efficiency of classification, sensitivity, predictability, and discrimination ability.

Table 4. Performance Evaluation Of Amgformer

Metric	Value (%)
Accuracy	96.86
Precision	96.60
Recall	97.40
F1-Score	96.50
AUC	97.00

Table IV shows the results of classifying Parkinson MRI dataset using the proposed AMGFormer model. The results show that the proposed model was highly efficient in classification with an accuracy of 96.86%. A precision and recall of 96.60% and 97.40%, respectively, were obtained, meaning that the model did not make any erroneous predictions during classification since it was highly efficient. Moreover, an F1-score of 96.50% showed that the two measures were highly integrated. In addition, the AUC was calculated at 97.00%.

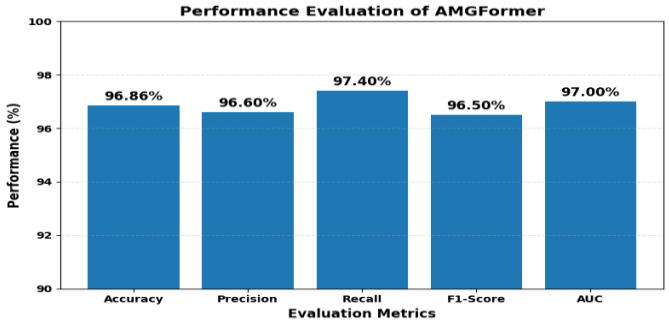


Fig 3. Classification Performance of Proposed AMGFormer Model

Fig. 3.shows the classification results of the AMGFormer architecture based on five classification criteria. As illustrated above, this architecture was characterized by impressive classification scores for all selected parameters, which proves the great ability of this technique to diagnose PD from brain scans. For example, an AUC score equal to 97.00% denotes the great ability of the model to separate classes, while an F1-score value of 96.50% confirms the balanced prediction of the algorithm.

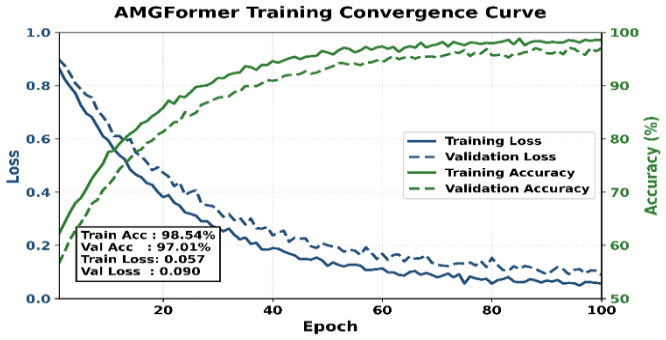


Fig 4. Convergence curves of AMGFormer

Fig. 4. presents the convergence curve of AMGFormer for 100 training epochs. Training and validation loss show a steady decline until convergence to relatively low values without any divergence, whereas the training and validation accuracies show sharp gains in the early training epochs and stabilize beyond epoch number 60. The small difference between the training and validation loss/accuracy curves signifies the absence of

overfitting of the model, notwithstanding the deep structure of the Graph Transformer and high capacity of MedMAE.

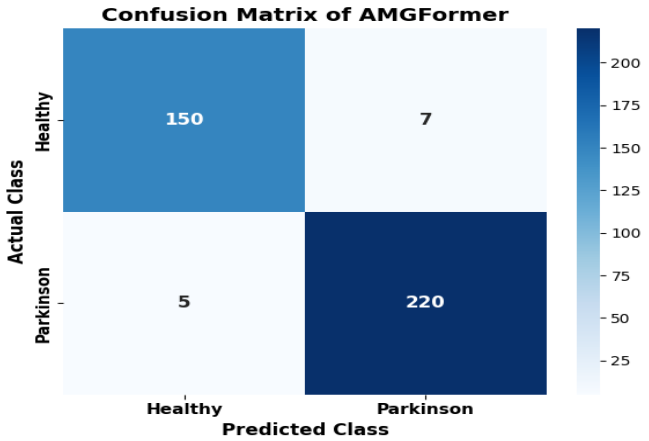


Fig 5. Confusion Matrix of AMGFormer Classification Results

Fig. 5. shows the confusion matrix produced by the AMGFormer method. The robustness of the suggested algorithm is demonstrated by the high accuracy with which a significant proportion of PD and healthy individuals were predicted. There were very few false positive and false negative outcomes. The extremely high classification accuracy demonstrates that the suggested approach, which included every feature, was effective in identifying the characteristics associated with the illness.

Cross-Validation Analysis

Cross-validation analysis with five folds was done to assess the stability and generalization of the proposed model.

Table 5. Five-Fold Cross-Validation Results

Fold	Accuracy (%)
Fold 1	96.42
Fold 2	96.61
Fold 3	96.87
Fold 4	97.02
Fold 5	97.08
Mean Accuracy	96.80
Standard Deviation	0.27

Table VI shows the performance metrics obtained through five-fold cross validation that was carried out to measure consistency and generalization ability of the model. Accuracy scores for the five folds vary from 96.42% to 97.08%. This means that the model is consistent in all folds since it has a high level of accuracy. The average accuracy score is 96.80%, which shows that the model is highly effective at classification. Finally, the small standard deviation (0.27) shows how consistent, reliable, and robust the model is.

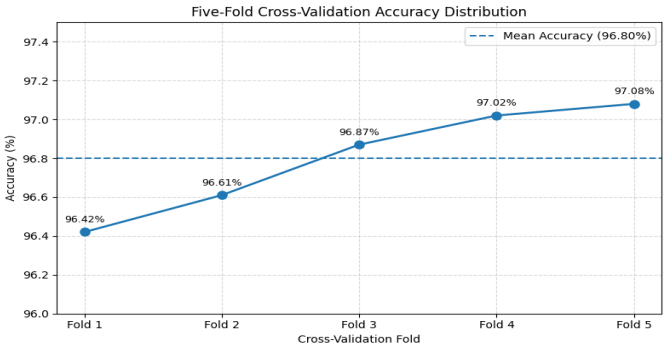


Fig 6. Five-Fold Cross-Validation Accuracy Distribution

Fig. 6.demonstrates the distribution of accuracies achieved during five-fold cross validation. Classification results were consistent with accuracies ranging between 96.42% and 97.08% throughout all the folds. A low standard deviation value of 0.27 signifies good stability and insensitivity to variations introduced by partitioning of the dataset. Achieving stable performance through several rounds of validation demonstrates that the designed framework generalizes well on unseen MRI data.

Ablation Study

Ablation studies have been carried out to test the impact of each component used in AMGFormer. Some ablated models have been tested and compared against the full AMGFormer architecture.

Table 6. Ablation Study Results

Configuration	Accuracy (%)
Texture + Structural Features	91.21
MedMAE Only	92.84
ANEF + MedMAE	93.97
Graph Transformer Only	94.68
MedMAE + Graph Transformer	95.60
Full AMGFormer	96.80

Table VII above show the impact of each part in improving the performance of the suggested AMGFormer architecture. Modules like MedMAE and Graph Transformer have been able to attain accuracies of 92.84% and 94.68% separately. By adding more components into the architecture, the accuracy of the model is further enhanced to 95.60% for MedMAE + Graph Transformer. The overall AMGFormer architecture achieves an accuracy of 96.80%, indicating the significance of using all the aspects discussed.

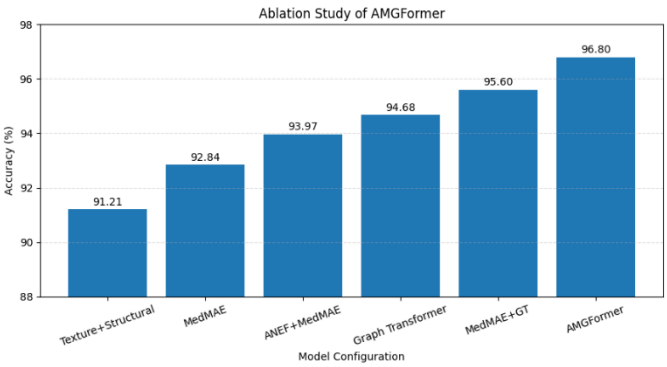


Fig 7. Ablation Study Results of AMGFormer Components

Fig. 7.shows the comparative performance for different architectures investigated in the ablation studies. Each module was found to be valuable in achieving high classification accuracy; however, the best result was obtained only by considering all the modules together. ANEF contributed in enhancing the quality of images, while MedMAE offered excellent semantics. The Graph Transformer helped in modeling the relationship between the features, thereby improving classification accuracy further.

Comparative Analysis with Existing Studies

To demonstrate the validity of the AMGFormer approach further, a comparison with other recent approaches for PD classification described in the literature was carried out. Various methods have been considered in this paper to make the analysis more comprehensive, namely, classical radiomics, multimodal features combination, as well as GNNs. The evaluation criteria utilized for the comparison included Accuracy, Precision, Recall, F1-score, and AUC values.

Table 7. Comparison With Existing Studies

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
Radiomics + LBP Texture Features + SVM/RF [19]	83.33	84.62	91.67	88.00	86.00
Handwriting + Gait + Speech Fusion + XGBoost + XAI [26]	92.00	89.00	90.00	89.00	95.00
Regional Radiomics Similarity Network + Transformer-GNN [14]	-	96.40	96.10	96.30	-
Proposed AMGFormer	96.80	96.60	97.40	96.50	97.00

Table VIII provides a comparison of the performance of the proposed AMGFormer framework vis-à-vis the other PD detection techniques from the recent scientific literature. The approaches considered include those utilizing radiomics analysis, multimodal feature fusion, and GNNs. The proposed framework obtained the best reported values for accuracy, precision, recall, F1 score, and AUC among the methods that provide all the metrics under consideration.

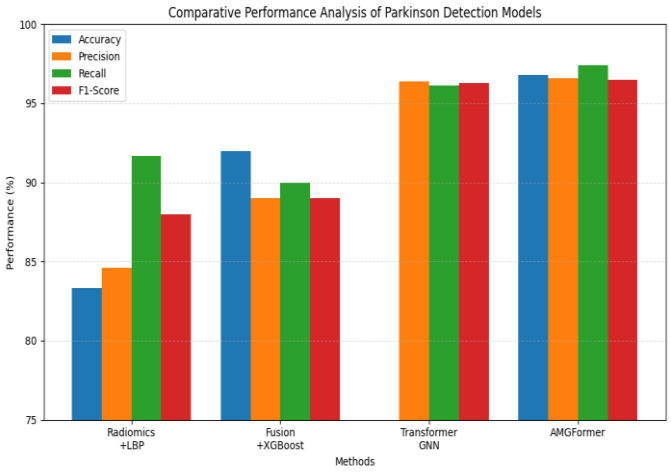


Fig. 8. Comparative Performance Analysis of Parkinson Detection Models

Fig. 8.shows comparison of the performance of the selected PD diagnosis frameworks using accuracy, precision, recall, and F1-score. It was observed that AMGFormer outperformed all other frameworks for the selected performance measures. Radiomic-based diagnosis had the least performance among all others since this framework relied only on handcrafted features. The classification ability was enhanced using the multimodal fusion technique by integrating data from different modalities. Graph-based framework produced competitive results.

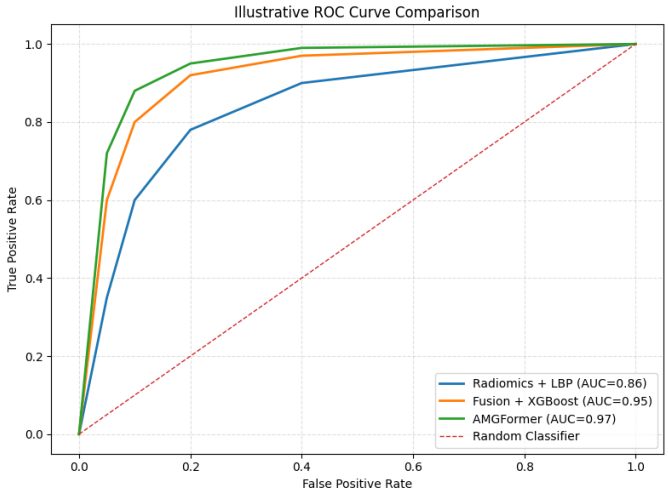


Fig. 9. AUC Comparison of Existing Diagnostic Models

Fig. 9.shows the AUC comparison for the examined PD diagnosis methodologies. The highest AUC score was obtained by the AMGFormer framework, which had a value of 97.00%. This shows that the AMGFormer framework has an excellent separation of classes and thus high diagnosis ability. The radiomics model had the lowest AUC score because of its low feature representation capability. Multimodal fusion framework performed well in differentiation because of its use of complementary modalities.

Discussion

These findings illustrate that the combination of adaptive image enhancement, feature engineering techniques, DL representations, and graph-based inference significantly enhances the performance in classifying PD using only MRI data. The high accuracy of 96.80%, high AUC of 97.00%, and excellent F1-score of 96.50% reveal that the proposed method not only discovers fine-scale texture differences but also understands global structural patterns which may not be discovered using conventional one-modal approaches. The small variance across 5-fold cross-validation (96.42%–97.08%, SD = 0.27%) confirms that AMGFormer has achieved good stability and is robust against changes in data split schemes, which is an important criterion in the clinical environment. In comparison to radiomics approaches based solely on handcrafted features and multimodal approaches relying on genetic or physiological information, AMGFormer achieves similar performance or even outperforms other methods using only MRI, which simplifies acquisition processes. Specifically, the attention mechanism of Graph Transformer seems highly capable of learning relationships between different features which could not be discovered by both handcrafted features alone or CNNs alone. In conclusion, these findings indicate that the proposed approach combining MedMAE-based SSL with graph-based relational modeling provides a promising approach to reliably diagnosing PD based solely on MRI images.

Conclusion and Future Work

AMGFormer, an Attention-guided Multimodal Graph Former for PD Classification based on enhanced MRI scans, was proposed by this study. Using adaptive neuro-enhancement filtering, manual feature extraction, SSL through MedMAE, and modeling via Graph Transformers, this framework is able to utilize the image features as well as their relations efficiently. According to experimentations performed on the Parkinson Multi-Modal Dataset 2.0, the results indicate that the framework achieves a high accuracy of 96.80%, 96.60% precision, 97.40% recall, F1-score of 96.50%, and an AUC of 97.00%. The consistency of the method was also confirmed through five-fold cross-validation, wherein the accuracy ranged from 96.42% to 97.08%, with only a standard deviation of 0.27%. Further proof of the effectiveness of each architecture component was provided by the ablation analysis, wherein it was seen that the combination of all these components achieved better results than any reduced form and previously proposed models. This makes AMGFormer a highly reliable PD classifier using MRIs alone.

Future research can focus on validating the framework using bigger, multi-centered, and multi-ethnic MRIs datasets to examine the performance on different populations and MRI acquisition protocols. Extending the study will also include applying the AMGFormer to other classifiers for distinguishing PD from similar diseases such as atypical Parkinsonism and essential tremor. The use of explainable AI techniques like attention map visualization and saliency-based interpretation is important for building more trustworthy models. Longitudinal

MRI data inclusion in the future can allow the construction of models that can predict the progression stages and severity of the disease and not only diagnose it. Furthermore, the implementation of a lighter version of the network that can run in real time will also allow the inclusion of federated learning algorithms to allow the training of models from different institutions but preserve patients' data privacy.

References

1. Y. Li, M. Yang, T. Bian, and H. Wu, “MRI Super-Resolution Analysis via MRISR: Deep Learning for Low-Field Imaging,” Information, vol. 15, no. 10, p. 655, 2024.

2. A. Zubair Rahman et al., “Advanced AI-driven approach for enhanced brain tumor detection from MRI images utilizing EfficientNetB2 with equalization and homomorphic filtering,” BMC Medical Informatics and Decision Making, vol. 24, no. 1, p. 113, 2024.

3. M. Mazher, G. J. Parker, and D. C. Alexander, “Towards generalisable foundation models for brain MRI,” npj Imaging, 2026.

4. M. Nielsen, L. Wenderoth, T. Sentker, and R. Werner, “Self-Supervision for medical image classification: State-of-the-Art performance with∼ 100 labeled training samples per class,” Bioengineering, vol. 10, no. 8, p. 895, 2023.

5. K. Ma et al., “Towards building a trustworthy deep learning framework for medical image analysis,” Sensors, vol. 23, no. 19, p. 8122, 2023.

6. T. Ghosh et al., “Foundation Models for Volumetric Medical Imaging: Opportunities, Challenges, and Future Directions,” Electronics, vol. 15, no. 6, p. 1245, 2026.

7. H. Li, Z. Yang, W. Qi, X. Yu, J. Wu, and H. Li, “Parkinson’s image detection and classification based on deep learning,” BMC Medical Imaging, vol. 24, no. 1, p. 187, 2024.

8. M. Alrawis, F. Mohammad, S. Al-Ahmadi, and J. Al-Muhtadi, “Fcn-pd: An advanced deep learning framework for parkinson’s disease diagnosis using mri data,” Diagnostics, vol. 15, no. 8, p. 992, 2025.

9. S. S. Hussain et al., “A swin transformer and CNN fusion framework for accurate Parkinson disease classification in MRI,” Scientific Reports, vol. 15, no. 1, p. 15117, 2025.

10. M. Pajany et al., “Multimodal deep feature fusion with transformer for brain tumor classification from magnetic resonance imaging,” Scientific Reports, 2026.

11. V. Dentamaro, D. Impedovo, L. Musti, G. Pirlo, and P. Taurisano, “Enhancing early Parkinson’s disease detection through multimodal deep learning and explainable AI: insights from the PPMI database,” Scientific reports, vol. 14, no. 1, p. 20941, 2024.

12. A. Sar, P. S. Puri, H. Naz, S. Aich, T. Choudhury, and L. A. Gabralla, “Multi-modal deep learning framework for early detection of Parkinson’s disease using neurological and physiological data for high-fidelity diagnosis,” Scientific Reports, vol. 15, no. 1, p. 34835, 2025.

13. Y. Yang et al., “Identification of Parkinson’s disease using MRI and genetic data from the PPMI cohort: an improved machine learning fusion approach,” Frontiers in aging neuroscience, vol. 17, p. 1510192, 2025.

14. R. Ling et al., “Explainable graph neural network based on metabolic brain imaging for differential diagnosis of parkinsonism,” Frontiers in Aging Neuroscience, vol. 17, p. 1580910, 2025.

15. S. B. Patel, V. Goh, J. F. FitzGerald, and C. A. Antoniades, “2D and 3D Deep Learning Models for MRI-based Parkinson’s Disease Classification: A Comparative Analysis of Convolutional Kolmogorov-Arnold Networks, Convolutional Neural Networks, and Graph Convolutional Networks,” arXiv preprint arXiv:2407.17380, 2024.

16. M. Camacho et al., “Exploiting macro-and micro-structural brain changes for improved Parkinson’s disease classification from MRI data,” npj Parkinson’s Disease, vol. 10, no. 1, p. 43, 2024.

17. N. Madusanka, H. S. Malekroodi, H. Herath, C. Hewage, M. Yi, and B.-I. Lee, “Spectro-Image Analysis with Vision Graph Neural Networks and Contrastive Learning for Parkinson’s Disease Detection,” Journal of Imaging, vol. 11, no. 7, p. 220, 2025.

18. H. Li, T. Liang, R. Yao, and T. Kuremoto, “Parkinson’s Disease Classification Using Gray Matter MRI and Deep Learning: A Comparative Framework,” Applied Sciences, vol. 15, no. 21, p. 11812, 2025.

19. S. Gore, A. Dhole, S. Kumbhar, J. Jagtap, A. Thakare, and S. Gudadhe, “Radiomics for Parkinson’s disease classification using advanced texture-based biomarkers,” MethodsX, vol. 11, p. 102359, 2023.

20. J. Cao and X. Long, “Racf: A multimodal deep learning framework for parkinson’s disease diagnosis using snp and mri data,” Applied Sciences, vol. 15, no. 8, p. 4513, 2025.

21. M. J. Oliveira, P. Ribeiro, and P. M. Rodrigues, “Machine learning-driven GLCM analysis of structural MRI for Alzheimer’s disease diagnosis,” Bioengineering, vol. 11, no. 11, p. 1153, 2024, doi: 10.3390/bioengineering11111153.

22. H. Kibriya, R. Amin, J. Kim, M. Nawaz, and R. Gantassi, “A novel approach for brain tumor classification using an ensemble of deep and hand-crafted features,” Sensors, vol. 23, no. 10, p. 4693, 2023.

23. X. Han, J. Feng, H. Xu, S. Du, and J. Li, “A hypergraph transformer method for brain disease diagnosis,” Frontiers in Medicine, vol. 11, p. 1496573, 2024.

24. G. Dheepak, D. Vaishali, and others, “Brain tumor classification: a novel approach integrating GLCM, LBP and composite features,” Frontiers in oncology, vol. 13, p. 1248452, 2024.

25. astha mishra96, A. Sar, and S. Aich, “PARKINSON MULTI MODAL DATASET 2.0.” Accessed: Jun. 13, 2026. [Online]. Available: https://www.kaggle.com/datasets/asthamishra96/parkinson-multi-model-dataset-2-0

26. A. Ravichandran, T. Kathirvel Murugan, L. Govindaraj, and V. M, “Explainable multimodal feature fusion networks for Parkinson’s disease prediction,” Frontiers in Digital Health, vol. 8, p. 1771281, 2026.