ORIGINAL ARTICLE

An Explainable Federated Temporal Multimodal Ensemble Deep Learning Model for Early Diagnosis and Progression Prediction of Parkinson’s and Alzheimer’s Diseases Using X-FedTME-Net

J. Jayapandian[1,*] **and Dr. M. Senthil**[2]
[1] Research Scholar, School of Computer Science, Takshashila University, Tamil Nadu, India
[2]Professor, School of Computational Engineering, Takshashila University, Tamil Nadu, India

(**Received** 24 February 2026, **Revised** 14 March 2026, 03 April 2026, **Accepted** 01 May 2026)

*Corresponding author: jayapandianjmca@kcet.in

DOI: 10.5875/g0qvad40

Abstract: The diagnosis process for neurodegenerative diseases which include Parkinson's disease and Alzheimer's disease faces difficulties because their symptoms overlap and patients experience different disease progression rates and healthcare facilities produce different imaging results. The research introduces X-FedTME-Net which functions as an interpretable federated temporal multimodal ensemble deep learning framework designed to diagnose medical conditions and track patient advancements through MRI data analysis. The model uses multiple MRI biomarker sources which include structural MRI data and diffusion tensor imaging data and resting-state functional MRI data to obtain complete brain structural and microstructural and functional brain information. The system uses a temporal transformer-based encoder to document important biomarker evolutions together with a 3D convolutional neural network which retrieves fundamental spatial information. A support vector machine establishes more precise classification limits between different categories. The system uses stacking ensemble methods to combine various components which enhances system performance through better stability and broader application. Federated learning allows multiple centers to work together while keeping patient information safe by preventing direct patient data exchange. To enhance transparency, explainable AI techniques such as Gradient-weighted Class Activation Mapping and SHapley Additive Explanations provide visual and feature-level insights by identifying disease-relevant brain regions and quantifying biomarker contributions, thereby improving model interpretability, increasing clinical trust, and supporting more reliable diagnostic decision-making. Experimental results on multi-center public MRI datasets demonstrate superior diagnostic accuracy, reduced false positives, and improved generalizability compared to standalone models, supporting its clinical potential.

Keywords: Parkinson’s Disease, Alzheimer’s Disease, Magnetic Resonance Imaging, Multimodal Neuroimaging, Federated Learning, Temporal Deep Learning, Ensemble Learning, Explainable Artificial Intelligence, Disease Progression Prediction, X-FedTME-Net

---

Introduction

The global healthcare system currently faces its most urgent challenge through neurodegenerative diseases, which have increased due to higher life expectancy and growing elderly populations. Wireless technologies perform a vital function by supporting continuous patient surveillance and medical data gathering, which healthcare practitioners use to treat patients with chronic illnesses such as Parkinson's disease and Alzheimer's disease. The two most widespread neuro-disability diseases in the world constitute Parkinson's disease and Alzheimer's disease, which together make up a substantial portion of global neuro-disability cases.

Parkinson’s disease is primarily characterized by motor dysfunction resulting from the degeneration of dopaminergic neurons in the substantia nigra. In contrast, Alzheimer’s disease involves a different pathological mechanism, mainly associated with progressive cognitive decline due to neuronal loss and abnormal protein accumulation, particularly amyloid-β plaques and tau tangles [1], [2]. Although the pathological bases of PD and AD are different, in the early period of their progression, both diseases have similar clinical manifestations, which results in misdiagnosis and delayed treatment. The initial clinical identification of Parkinson's disease and Alzheimer's disease is complicated because both conditions share common symptoms and show different rates of disease progression. This creates problems for doctors because they cannot decide which medical condition the patient has, which results in an extended wait time for the proper medical treatment to start. The development of advanced computational diagnostic systems is necessary to provide accurate detection of these diseases during their earliest stages, which will lead to better results for patients. The diagnosis needs to be done at the correct time with proper methods because it helps to control the disease and slows down its development. The traditional diagnostic methods require doctors to conduct clinical assessments together with neuropsychological tests which lack objectivity and fail to detect early pathological changes [3]. Multimodal Magnetic Resonance Imaging (MRI), federated learning (FL), and explainable artificial intelligence (XAI) are increasingly used in neurodegenerative disease analysis to improve diagnosis and prediction accuracy. Magnetic Resonance Imaging (MRI) has become a potent non-invasive modality that can be used to study neurodegeneration by measuring structural, microstructural, and functional changes in the brain.

Structural MRI also shows cortical thinning and hippocampal atrophy in AD, whereas diffusion tenor imaging (DTI) and functional MRI (fMRI) demonstrate white matter degeneration and disturbed functional connectivity of PD and AD [4], [5]. Convolutional Neural Networks (CNNs) have performed outstandingly well in detecting disease-specific MRI patterns that human beings may not be able to discern [6]. On the same note, machine learning approaches like Support Vector Machines (SVMs) have also been applied successfully to MRI-based biomarkers on the classification of neurodegenerative disease [7]. However, the majority of the literature is on a single-modality imaging, cross-sectional analysis, or binary classification, which restrict its potential in the real world of clinical contexts where the disease progression, data heterogeneity, and privacy are essential considerations. Diffusion Tensor Imaging (DTI) evaluates white matter integrity by measuring the directional movement of water molecules in brain tissue. Fractional anisotropy (FA) values derived from DTI indicate disruptions in white matter tracts associated with neurodegenerative diseases. These changes particularly affect dopaminergic pathways. Therefore, DTI serves as an important biomarker for detecting microstructural brain alterations. The other significant shortcoming of the existing AI-based diagnostic systems is that they are not interpretable and generalizable. The deep learning models are thought of as a black box, which limits clinical trust and adoption [8]. The centralized model training for AI systems requires extensive data sharing which introduces major challenges to both patient privacy protection and adherence to regulatory standards. Federated learning has emerged as a promising solution because it enables institutions to train models without transferring their confidential patient information to other locations. [9]. Wireless communication systems are essential for facilitating real-time data transfer and remote collaboration in healthcare, contributing to the efficiency of systems like federated learning.

The current MRI-based artificial intelligence models for diagnosing neurodegenerative diseases face multiple obstacles even after achieving recent technological advancements. Existing approaches often rely on standalone or limited multimodal data. They fail to capture disease progression over time. Additionally, centralized training raises significant privacy concerns. The majority of models demonstrate two main weaknesses because they cannot provide clear explanations and they only perform well with datasets from a single research site.

Nevertheless, its use with multimodal neuroimaging and explainable AI has not been well studied. In order to resolve these issues, the proposed research suggests X-FedTME-Net, an Explainable Federated Temporal Multimodal Ensemble Deep Learning system to diagnose and predict the progression of Parkinson's and Alzheimer's diseases early in an individual using MRI. The multimodal MRI characteristics, temporal disease modeling, hybrid ensemble learning, federated optimization, and explainable AI mechanisms are proposed to be integrated into a single architecture X-FedTME-Net provides a clinically viable and scalable diagnostic support system. It integrates spatial, temporal, and functional brain biomarkers. The framework also incorporates privacy-preserving and interpretable learning strategies. Parkinson's disease develops because the disease causes dopaminergic neurons in the substantia nigra to degenerate, which results in decreased dopamine production and impaired movement abilities. The degeneration shows evidence through three imaging biomarkers which include basal ganglia structural changes from sMRI and white matter disruption that DTI measures through fractional anisotropy reduction and changes in motor network functional connectivity from fMRI. The proposed X-FedTME-Net system achieves its goal of capturing biomarkers through its ability to extract multiple features from different modalities while using temporal modeling to establish clinical relationships between neuropathological conditions and computational model results. This paper presents a smart healthcare framework focusing on the integration of wireless technologies in medical innovation for neurodegenerative disease analysis. The study emphasizes advanced computational methods for early diagnosis and progression prediction of Parkinson's disease and Alzheimer's disease using multimodal neuroimaging data. The proposed method uses federated learning together with temporal deep learning and explainable AI to create a clinical decision support system which protects user privacy while delivering interpretable results and maintaining operational efficiency.

Literature Survey

The new progress in artificial intelligence (AI) technology has transformed the process of analyzing neuroimaging data which doctors use to identify neurodegenerative diseases. The combination of machine learning and deep learning techniques demonstrates effective results for discovering hidden disease indicators which medical professionals can identify through magnetic resonance imaging (MRI) technology for both PD and AD conditions. The first research studies used manually collected data which included structural MRI values of hippocampal volume and cortical thickness and ventriculary enlargement as their primary data source [10]. Although these methods had moderate classification accuracy, feature selection bias and inter-subject variability limited its performance. The popularization of deep learning has also seen convolutional neural networks (CNNs) become the new paradigm of automated neuroimaging analysis. CNN models can acquire a hierarchical representation of space by learning methods directly on raw MRI volumes, which in turn obviates the feature engineering that is necessary in conventional methods. Multiple papers have also indicated there are dramatic increases in the accuracy of AD detection in 3D CNN-based models when trained on T1-weighted MRI scans, and that they achieved a high detection accuracy (above 90 percent) when they trained in controlled experimental conditions [11], [12]. Deep CNN models have effectively been used in the diagnostics of PD to identify structural and microstructural anomalies in the basal ganglia and substantia nigra areas [13].

Irrespective of their effectiveness, single-modality deep learning models have intrinsic shortcomings, since neurodegenerative diseases impact the brain on many different levels. To solve this, scholars have progressively turned to multimodal fusion approaches in MRI, which integrate structural MRI, diffusion tensor imaging (DTI), as well as functional MRI (fMRI). Structural MRI will give anatomical data, DTI will show integrity of the white matter and fMRI will show disruption of functional connections. Functional MRI plays a crucial role in analyzing functional connectivity disruptions within the brain, which is vital for understanding neural network impairments. It provides insights into how different brain regions communicate and adapt, particularly in neurodegenerative diseases. This analysis helps refine the understanding of disease progression and complements structural and microstructural imaging. Research has revealed that multimodal MRI characteristics can become important in AD and PD diagnosis much more effective than unimodal methods especially in early disease stages [14], [15]. Deep multimodal learning architecture and feature-level fusion have shown to be more robust in heterogeneous data. The other significant weakness of the current studies is the use of cross-sectional data which limits models to diagnosis at rest. Neurodegenerative diseases, on the other hand, are progressive, and require time modeling of progression of the disease. Longitudinal MRI analysis has been highly considered as a tool to determine disease progression, cognitive decline and mild cognitive impairment (MCI) to AD conversion [16]. Transformer-based architectures have more recently been proposed as they can capture long-range temporal correlations and irregular follow-up times, which provide better progress prediction performance [17].

Even though deep learning models are effective predictors, their uninterpretability is a significant obstacle to clinical implementation. Clinicians need explainable and clear systems to believe AI-based decisions. Gradient-weighted Class Activation Mapping (Grad-CAM) is a commonly used explainable artificial intelligence (XAI) method that can be used to visualize disease-relevant brain regions that affect CNN predictions [18]. Additional feature-attribution techniques like SHapley Additive exPlanations (SHAP) also measure the contribution of single biomarkers to assess them in a way that facilitates more meaningful clinical interpretation [19]. Clinical studies combining XAI have found that clinician confidence and consistency between AI results and established neuropathological results improved. One more issue of high priority is data privacy and scalability. Centralization of the data is a feature of most deep learning models, which is frequently impossible because of legal, ethical, and regulatory restrictions. Federated learning (FL) has become one of the potential solutions as it facilitates decentralized training in several institutions without the need to exchange raw patient data. The recent literature has shown that federated learning can match the performance of centralized models without compromising patient privacy when used in neuroimaging [20], [21]. Nonetheless, the combination of federated learning and multimodal MRI, temporal modeling, and explainability is quite underrepresented. Ensemble learning methods have also become popular in the diagnosis of neurodegenerative diseases. Ensemble models are used to describe the combination of a number of classifiers to minimize variance and enhance generalization in different datasets. Ensembles Hybrid ensembles combining deep learning models with classical classifiers like SVMs have been demonstrated to have a lower false positive rate and greater robustness in multi-center experiments [22]. These results indicate that ensemble methods are especially efficient in the clinical setting that is marked by heterogeneous imaging protocols. Altogether, available literature demonstrates the great advancement of AI-based MRI-based neurodegenerative disease diagnosis. Naresh proposed employ machine learning and deep learning techniques to analyze complex healthcare data for detecting financial fraud patterns. In the proposed X-FedTME-Net, similar learning strategies are adapted for multimodal MRI-based disease diagnosis, enhancing prediction accuracy, robustness, and clinically reliable outcomes [23]. Nonetheless, several critical gaps such as less time modeling, absence of privacy preserving structures, inadequate explainability, and poor generalization between centers still exist [24]. These shortcomings are the reasons that inspired the current X-FedTME-Net that incorporates multimodal MRI fusion, temporal deep learning, ensemble classification, federated optimization, and explainable AI in a single tool to aid in early diagnosis and progression prediction of Parkinson and Alzheimer diseases. CNN-based models automatically generate multiple levels of features starting from raw MRI data, whereas traditional machine learning methods rely on manually crafted features, such as hippocampal volume and cortical thickness. The study by Mamidala presents a distributed intelligence framework in IoHT using federated learning and edge AI for disease prediction, highlighting the importance of privacy-preserving collaborative learning in healthcare systems [25]. The integration of Bi-directional GRU enables effective temporal pattern learning, which is relevant for modeling disease progression over time. These findings support the applicability of federated and temporal deep learning approaches in developing scalable and robust medical diagnostic frameworks similar to the proposed model.

Proposed System

The proposed X-FedTME-Net system architecture will allow the integration of multimodal MRI analysis, temporal disease progression modeling, ensemble learning, federated optimization, and explainable artificial intelligence into a single system design. Wireless healthcare solutions are integrated within the system to allow seamless data transmission across multiple institutions while maintaining privacy. The architecture allows collaborative training between distributed medical centers without compromising the privacy of patients and provides strong generalization. The system combines learning the patterns of spatial anatomy, trends of longitudinal progression and complementary decision boundaries to provide accurate interpretable and scalable diagnosis and progression prediction of both Parkinson and Alzheimer disease. The architecture of the proposed system is shown in figure 1 defines the functional units and data flow that constitute these objectives in a clinically implementable way.

3.1 X-FedTME-Net Architecture

The X-FedTME-Net architecture is designed as a modular yet tightly integrated system that supports multimodal input processing, temporal modeling, ensemble decision making, federated optimization, and explainable inference. The architecture Shown in figure 1 shows a layered design in which each component contributes a specific functional role while interacting seamlessly with others. The proposed framework uses a federated optimization method to support model training at several medical facilities which operate independently. The X-FedTME-Net model is trained by each center through its dedicated local instance which utilizes its own confidential data collection. A central aggregation server receives model updates which do not include any data from the system. The global model uses the Federated Averaging (FedAvg) algorithm to update its parameters because local dataset sizes determine the weighted average of model parameters from different locations.

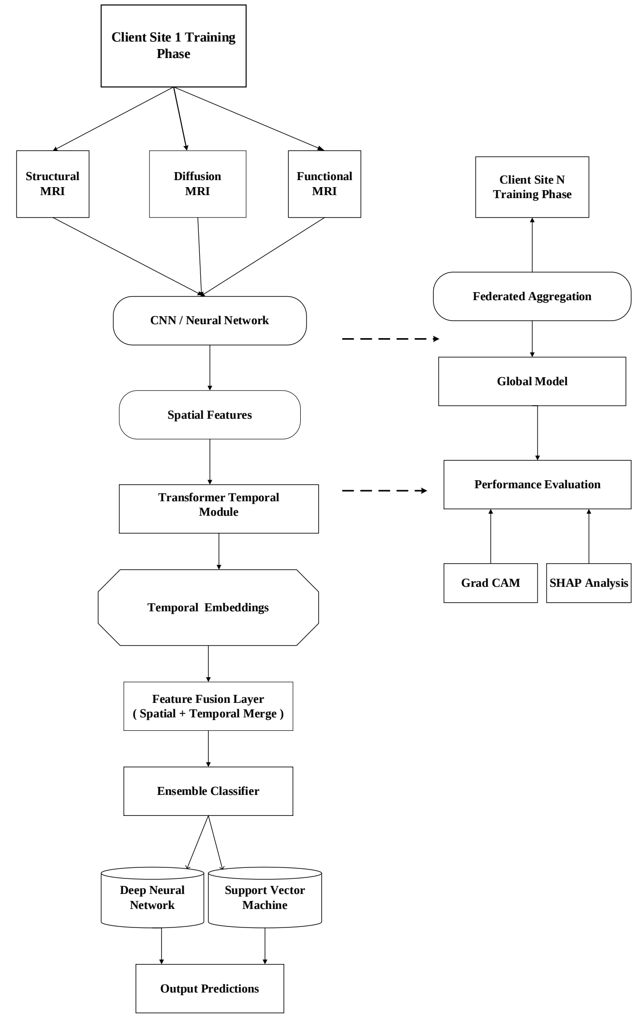

Figure 1: X-FedTME-Net architecture showing fusion of spatial CNN features and temporal transformer embeddings for final prediction.

The architecture takes in multimodal MRI volumes at a distributed clinical site at the input layer. Structural MRI inputs give anatomy, diffusion MRI gives microstructural integrity and functional MRI gives dynamic patterns of connection. These inputs are encoded by modality specific encoders that retain domain specific information but put the features dimensions in line with fusion. The fusion layer integrates spatial features extracted from the CNN with temporal embeddings generated by the transformer encoder. This combination enables the model to jointly learn spatial patterns and longitudinal disease progression. The fused representation is then passed to the ensemble classifier for final prediction. The spatial encoding block comprises of a 3D backbone of a convolutional block with residual links. The residual learning allows training deep networks in a stable manner by learning residual mappings as illustrated below.

$$\mathrm{y} = \mathcal{F}(\mathrm{x}, \mathrm{W}) + \mathrm{x} \tag{1}$$

This formulation allows the network to capture subtle neurodegenerative patterns without degradation in performance as depth increases. In parallel, the temporal encoding block processes longitudinal feature sequences using transformer layers. Positional encoding is applied to preserve temporal order as represented as

$$\begin{gathered}\mathbf{PE}(t, 2i) = \sin\left(\frac{t}{10000^{\frac{2i}{d}}}\right) \\ \mathbf{PE}(t, 2i+1) = \cos\left(\frac{t}{10000^{2i/d}}\right)\end{gathered} \tag{2}$$

This allows the architecture to learn disease trajectories across time.

The fusion layer jointly learns both spatial and temporal embedding, with the resulting latent space. It is essential to this fusion because it brings disease severity and disease progression rate under one roof. The classification layer is composed of two learners, a deep classifier and a SVM classifier. The proposed model uses a stacking-based ensemble method to merge the prediction results from the deep neural classifier and the support vector machine (SVM) system. The first stage requires both classifiers to undergo independent training on the combined multimodal feature data. The deep classifier produces probabilistic outputs and the SVM system generates decision scores through optimized margin boundaries. The outputs serve as input features for a meta-learner, which conducts the final prediction. Their outputs are given in a meta-learner which carries out a final decision making. This architecture design assures that representation learning and decision boundary optimization are complementary. The system of federated learning is executed as an outer architectural loop. The architecture is represented in each local instance in each institution, and periodically parameters are synchronized with an aggregate instance. Safe aggregation provides confidentiality. The explainability layer superimposes the architecture, and it can be visualized and interpreted at various levels. The heatmaps indicate the parts of the brain that are disease-relevant, and feature attribution scores reveal why data is diagnostic.

A process was developed to track disease development through the analysis of patient MRI results collected over multiple time periods. Multimodal data was gathered at each time point, including structural MRI, DTI, and fMRI data, which formed an ordered sequence obtained through time-based acquisition. Time-aware sequence construction was used to link each observation with its actual time gap from the baseline scan, as subjects had different follow-up periods.

The last output layer produces disease classification, progression risk scores, and explainability reports, The model outputs are compatible with real-time clinical decision support systems. After MRI acquisition, X-FedTME-Net generates immediate disease classification and progression risk scores. The explainability outputs, including Grad-CAM and SHAP, provide visual and feature-level insights, enabling clinicians to make faster and more reliable diagnostic decisions. which makes the architecture appropriate to real-time clinical decision support. On the whole, the X-FedTME-Net architecture will be a complex and scalable architecture that will help to eliminate the distance between the field of advanced AI research and practical clinical implementation. The proposed system is capable of tackling key issues in the diagnosis of neurodegenerative diseases and predicting their progression using multimodal imaging, temporal models, ensemble intelligence, federated learning, and explainable AI. The process of multimodal fusion requires a pre-integration phase which uses dimensional alignment and normalization methods. The encoding system for structural MRI DTI and fMRI uses multiple modality-specific encoders which create feature embeddings that differ in dimension. The system uses fully connected transformation layers to convert the embeddings into a shared latent space which maintains consistent dimensionality across different modalities. The system first applies z-score normalization to each feature representation which standardizes distribution and removes scale differences between modalities.

3.2 X-FedTME-Net Algorithm

The proposed system is developed in the framework of a new algorithm called X-FedTME-Net, which has the meaning of Explainable Federated Temporal Multimodal Ensemble Network. The main aim of such an algorithm is to conduct early diagnosis and prediction of progression of PD and AD using multimodal magnetic resonance imaging (MRI) with guaranteeing interpretability, temporal consciousness, and privacy of data in a multifaceted medical facility. In contrast to traditional AI models that process single-modal or centralized learning, X-FedTME-Net presents a single formulation that incorporates multimodal feature learning, temporal disease learning, ensemble decision making, federated optimization as well as explainable artificial intelligence as one learning paradigm. Let a patient $p$be associated with longitudinal multimodal MRI observations acquired at different time points $\mathrm{t} \in \{1,2,\ldots,\mathrm{T}\}$. Each observation consists of structural MRI S_t, diffusion MRI D_t, and functional MRI F_t. The multimodal input representation at time $t$is defined as

$$\mathbf{X}_t^p = \{S_t^p, D_t^p, F_t^p\} \tag{3}$$

All modalities undergo standardized preprocessing and normalization to ensure consistent feature distributions across subjects and institutions. Feature normalization is performed using z-score normalization defined as

$$\tilde{x}_i = \frac{x_i - \mu}{\sigma} \tag{4}$$

where x_i is the original feature value, and μ and σrepresent the mean and standard deviation computed across the training population. To capture disease evolution, X-FedTME-Net explicitly models temporal dependencies across longitudinal scans. For each subject, normalized multimodal features are arranged into a temporal sequence as given below

$$\mathrm{X}^p = [\mathrm{X}_1^p, \mathrm{X}_2^p, \ldots, \mathrm{X}_T^p] \tag{5}$$

This sequence is passed to a transformer-based temporal encoder, which computes contextualized temporal embeddings by leveraging self-attention. The attention mechanism computes the relevance between different time points as

$$\mathbf{h}_t^{\text{temp}} = \mathrm{Transformer}(\mathbf{H}), \mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_T]$$
$$\mathrm{ttention}(Q,K,V) = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{6}$$

where $\mathbf{h}_t^{\text{temp}}$ represents the temporal embedding at time step t ,$\mathbf{H}$ is the sequence of input feature vectors, Q,K,V denote query, key, and value matrices, and d_k is the dimensionality of the key vectors. where Q, K, and V are query, key, and value matrices derived from temporal feature projections, and $d_k$is the key dimension. This formulation permits the model to pick up long-range temporal dependencies so that it can tell the difference between typical aging and pathological progression patterns. Similar to temporal modeling, spatial feature learning is carried out on volumetric MRI directly with the help of a three-dimensional convolutional neural network (3D-CNN). Given a volumetric MRI input V, convolutional feature maps are computed as:

$$\mathbf{h}_t^{sp} = \sigma(W * X_t + b) \tag{7}$$

where X_t is the MRI input at time t_1, W represents convolution kernels, denotes convolution, b is bias, and $\sigma(\cdot)$ is a nonlinear activation function. where $\mathbf{W}_i$denotes convolution kernels, $*$ represents the convolution operation, $b$is the bias term, and $f(\cdot)$is a nonlinear activation function. This procedure allows for the extraction of more or less spatial hierarchies in relation to which hippocampal atrophy, cortical thinning, or nigral degeneration might have been placed. The model configuration parameters are defined to ensure reproducibility and clarity. The transformer-based temporal encoder is implemented with 4 layers, 8 attention heads, and an embedding dimension of 256. The 3D-CNN backbone consists of multiple convolutional layers with kernel size 3×3×3, followed by batch normalization and ReLU activation, and includes residual connections for stable training.

Key hyperparameters include a learning rate of 0.001, batch size of 16, and training over 100 epochs using the Adam optimizer. Dropout with a rate of 0.5 is applied to prevent overfitting. The SVM classifier uses a radial basis function (RBF) kernel with optimized regularization parameters. The temporal embedding $\mathbf{H}_t$ obtained from the transformer encoder and the spatial embedding $\mathbf{F}_{cnn}$ obtained from the CNN are fused into a unified latent representation:

$$\mathbf{h}_t^{\text{fusion}} = \left[\mathbf{h}_t^{sp} \| \mathbf{h}_t^{\text{temp}}\right] \tag{8}$$

where $\|$ denotes concatenation of spatial and temporal features. where $\oplus$ denotes feature concatenation. This is the synthesis of the diagnosis so that spatial abnormalities of the anatomy and longitudinal progression patterns have a combined effect on the diagnosis. X-FedTME-Net uses ensemble learning strategy to enhance the robustness and generalization. Two parallel classifiers are trained simultaneously. The initial classifier is a deep neural classifier which calculates the class probabilities through the use of a softmax function as shown below

$$P(y = c \mid \mathbf{h}) = \frac{e^{w_c^T \mathbf{h}}}{\sum_{j=1}^{C} e^{w_j^T \mathbf{h}}} \tag{9}$$

where C is number of classes and w_c are class weights. where C denotes the number of diagnostic classes. The second classifier is a support vector machine (SVM) trained on extracted biomarkers and latent features. The SVM decision function is defined as

$$f(\mathbf{h}) = \sum_{i=1}^{N} \alpha_i^{\downarrow} y_i K(\mathbf{h}, \mathbf{h}_i) + b \tag{10}$$

where $K(\cdot,\cdot)$ is a radial basis function kernel and α_i are learned coefficients. The SVM uses the kernel function with its radial basis function (RBF) to transform input features into higher dimensions which enable the SVM to separate complex nonlinear patterns from its data. This approach works well with multimodal MRI features because they show both high dimensionality and nonlinear behavior, which helps to improve classification accuracy and create flexible decision boundaries. These two classifiers capture complementary decision boundaries. The ensemble output is computed using a stacking-based meta-learner as shown below

$$y = f_{\text{meta}}\left(y_{DL}, y_{SVM}\right) \tag{11}$$

where $\mathcal{M}(\cdot)$ learns optimal weights for combining predictions. This ensemble formulation significantly reduces false positives and improves stability across heterogeneous datasets. To ensure privacy-preserving learning, X-FedTME-Net is trained using federated learning. Suppose there are K medical institutions, each with local data $\mathcal{D}_k$. Each institution trains a local model with parameters θ_k. A global model is obtained using federated averaging as shown below

$$\theta^{global} = \sum_{k=1}^{K} \frac{n_k}{N} \theta_k \tag{12}$$

where K is number of clients, n_k is data size of client k, $N = \sum_{k=1}^{K} n_k$, and θ_k are local model parameters. This formulation ensures that sensitive patient data never leave institutional boundaries while enabling collaborative learning. The federated averaging process performs weighted aggregation of local model parameters. Each participating institution contributes according to its data size. The global model achieves its purpose by using distributed datasets to create a unified knowledge base which maintains data privacy and enables equal learning from different sources. Finally, interpretability is achieved through explainable AI mechanisms. Gradient-based class activation maps identify spatial regions contributing to predictions, while SHAP values quantify feature importance as shown below

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!\,(|F| - |S| - 1)!}{|F|!} [\\ f(S \cup \{i\}) - f(S)] \tag{13}$$

This dual explainability provides both spatial and feature-level transparency, making the algorithm clinically trustworthy. SHAP provides a unified framework for feature attribution by assigning contribution scores to each input feature which shows how features affect model predictions. The scores reveal how each biomarker influences the results for disease classification and progression prediction by showing both its positive and negative effects.

Results and Discussion

In this section, the detailed analysis of the suggested X-FedTME-Net model of early diagnosis and progression of PD and AD through the multimodal MRI-data. They are evaluated based on several aspects, such as the classification accuracy, prediction ability in progression, inter-center robustness, the effectiveness of different ensembles, the impact of federated learning, and explainability. Multi-center MRI datasets were used as the inputs of extensive experiments to prove the reliability and clinical relevance of the proposed system.

4.1 Experimental Setup

The experiments were performed using a multi-center dataset consisting of structural MRI, diffusion-tensor-imaging (DTI), and resting-state- fMRI data that were obtained in different institutions. Wireless technology enabled efficient data sharing and real-time updates, which were crucial for multi-center experiments. This data consisted of healthy controls, PD, and AD patients with longitudinal follow-ups. To have a strong performance estimation, five-fold cross-validation was used. Five-fold cross-validation enables more accurate model evaluation because it divides the dataset into five parts which test each part one time while training with the other four parts. This method prevents overfitting while it measures the model's ability to generalize across various data sets. Measures of evaluation were accuracy, sensitivity, specificity, F1-score, area under the ROC curve (AUC), and error of prediction of progression. The offered X-FedTME-Net was contrasted with single CNN models, SVM classifications and non-temporal baselines in order to present its efficiency.

4.2 Dataset Characteristics and Baseline Validity

The data was selectively edited so that it represented a balance of diagnostic groups. The factors of age, sex, and education level were matched to represent the minimum effects of confounding factors because these factors are known to affect the brain morphology and cognitive abilities. As it can be seen in Table 1, the average age and educational level among healthy controls, AD, and PD patients does not have any statistically significant differences. This equilibrium makes sure that the learning process of X-FedTME-Net is not analyzing disease-specific neuroimaging patterns, but it is driven by demographic bias. This type of baseline validation is necessary since a high-quality model may turn unreliable when trained on biased populations. This consistency provides the basis on which future performance measurement can be done.

Table 1: Demographic and Clinical Characteristics of Subjects

Group	**Subjects**	**Age (Mean ± SD)**	**Gender (M/F)**	***Education (Years)***
Healthy Controls	120	69.8 ± 6.1	58 / 62	15.9 ± 2.4
Alzheimer's Disease	120	73.5 ± 5.8	61 / 59	14.2 ± 2.9
Parkinson's Disease	***120***	***71.2 ± 6.7***	***63 / 57***	***15.1 ± 2.6***

The demographic analysis reveals that there were negligible differences in groups and this validates the fact that age, gender and education were balanced. This makes sure that the performance differences that are found in the further analyses are mainly due to the disease- specific neuroimaging patterns and not due to demographic confounders. Each of the imaging modalities was experimented independently to determine the contribution of each modality. This discussion is essential since neurodegenerative diseases have more than one biological dimension and the use of one modality may result in the lack of complete characterization. In order to determine the contribution of each MRI modality, they were trained separately.

Table 2: Classification Performance Using Single MRI Modalities

Modality	***Accuracy (%)***	***Sensitivity (%)***	***Specificity (%)***	**AUC**
Structural MRI	87.1	85.6	88.4	0.89
DTI	83.9	82.1	85.2	0.86
fMRI	82.7	81.3	84.0	0.85

Table 2 presents the classification performance of individual MRI modalities using evaluation metrics such as accuracy, sensitivity, specificity, and area under the ROC curve (AUC). Accuracy indicates overall correctness of classification, sensitivity measures the ability to correctly identify diseased cases, and specificity reflects correct identification of healthy cases. The AUC score represents the model's ability to distinguish between classes, where higher values indicate better discriminative performance. Table 2 indicates that the structural MRI has the highest standalone accuracy. This finding is not surprising because cortical thinning and hippocampal atrophy are typical signs of AD and basal ganglia alterations are pertinent to PD. Nevertheless, the relatively poorer results of DTI and fMRI cannot be viewed as one of the weaknesses. These modalities, instead, reflect microstructural and functional abnormalities that cannot be detected in structural scans, especially at the early stages of an illness. The average yet similar performance of DTI and fMRI points to the complementary character of these two methods. The above results support the multimodal structure of X-FedTME-Net and support the fact that unimodal methods cannot be relied on in strong diagnosis. X-FedTME-Net is based on the core design principles of multimodal fusion. Though each of the modalities gives only partial information, the combination of multimodal MRI data allows us to have a comprehensive look at neurodegeneration. A systematic comparison of various modality combinations is given in Table 3.

Table 3: Performance Comparison of Multimodal Combinations

Modalities Used	***Accuracy (%)***	***F1-Score***	***AUC***
sMRI + DTI	89.6	0.90	0.92
sMRI + fMRI	90.1	0.91	0.93
DTI + fMRI	87.8	0.88	0.90
sMRI + DTI + fMRI	**92.4**	**0.93**	**0.95**

Table 3 summarizes the performance of different multimodal combinations using accuracy, F1-score, and AUC. Accuracy reflects overall prediction performance, while the F1-score provides a balance between precision and recall, making it suitable for evaluating classification robustness. The AUC score indicates the effectiveness of the model in distinguishing between disease classes, with higher values representing improved classification capability. Compared to single-modality models (Table 2), the multimodal configuration (Table 3) shows a clear improvement in performance. For example, the highest accuracy using a single modality (87.1% with structural MRI) increases to 92.4% when all three modalities are combined, resulting in an improvement of 5.3%. Similarly, the AUC improves from 0.89 to 0.95, demonstrating the effectiveness of multimodal MRI fusion. From the results shown in table 3, all three modalities yield the highest AUC and the most accurate. This improvement cannot be described as an incremental one rather it signifies the ability X-FedTME-Net to learn cross- modal interaction among structural degeneration, loss-of-white-matter-integrity, and functional network breakdown. These relations are crucial in the distinction of PD and AD especially in specific clinical cases of uncertainty or youthful cases in clinical practice. Another piece of evidence pointing to the fact that signatures of disease are distributed across modalities is the difference in performance of the bimodal and trimodal configurations. The specified discovery has an immense contribution to the design decision of multimodal fusion of the given system. Transformer attention mechanisms capture long-range temporal dependencies, allowing X-FedTME-Net to model disease progression over time more effectively than traditional methods. This capability is crucial for understanding subtle variations in disease trajectories and distinguishing between typical aging and pathological changes. The temporal attention mechanism significantly improves the model's interpretation of experimental performance by correlating disease markers across multiple time points. The X-FedTME-Net has the advantage that it explicitly models disease progression with longitudinal MRI and does so in a manner that the disease progression is explicit. Early deep learning designs execute scans separately without considering time continuity. This tends to be a restrictive aspect, which will result in incorrect classification of early-stage patients or slow progressors.

Table 4: Comparison Between Temporal and Non-Temporal Models

Model Type	***Accuracy (%)***	***AUC***	**Progression Error (%)**
Non-Temporal CNN	88.9	0.91	9.4
LSTM-based Model	91.2	0.93	6.8
X-FedTME-Net (Transformer)	**93.1**	**0.96**	**4.2**

Table 4 demonstrates that temporal modeling is much more effective than other models in improving diagnostic accuracy and in predicting the progression. The proposed X-FedTME-Net achieved a cross-validation accuracy of 93.1%, outperforming the non-temporal CNN (88.9%) and LSTM-based model (91.2%). This reflects an improvement of 4.2% and 1.9%, respectively. Furthermore, the progression prediction error was reduced from 9.4% in CNN models to 4.2%, demonstrating a significant enhancement in longitudinal prediction capability. The temporal encoder, which is based on transformers, performs better than non-temporal CNNs and LSTM-based models, as it successfully exploits long-term dependencies among consecutive scans. The decrease of the error of prediction of the progression is especially significant in clinical practice, when the early detection of a rapid progression may be considered in relation to the choice of treatment methods. These findings prove that temporal attention systems are more appropriate at modeling neurodegenerative diseases than the traditional recurrent strategies. Ensemble strategy is a combination of deep learning and SVM classifier. Single classifiers can be found to have difficulties with borderline cases as their decision boundaries are limited. X-FedTME-Net tries to solve this problem by using a hybrid ensemble, which incorporates deep CNN prediction, and the use of SVM-based decision boundaries. Compared to recent state-of-the-art models, X-FedTME-Net demonstrates clear advantages by integrating multimodal MRI fusion, transformer-based temporal modeling, federated learning, and explainability within a unified framework. Unlike existing approaches that rely on single-modality data or lack temporal and privacy-preserving mechanisms, the proposed model achieves higher diagnostic accuracy, better generalization across multi-center datasets, and improved clinical interpretability, making it a more comprehensive and practically applicable solution.

Table 5: Effect of Ensemble Learning

Classifier	***Accuracy (%)***	***Sensitivity (%)***	**Specificity (%)**
CNN Only	90.2	88.9	91.6
SVM Only	88.5	87.1	89.3
CNN + SVM Ensemble	93.4	92.2	94.1

The ensemble demonstrates better accuracy and sensitivity and specificity than its individual classifiers according to Table 5. The improvement shows that CNNs and SVMs provide two different data representation methods. The ensemble model significantly reduces false positives compared to individual classifiers. The specificity improved from 91.6% (CNN) and 89.3% (SVM) to 94.1% in the ensemble model, indicating a notable reduction in false positive rates. This improvement is critical in clinical applications, where minimizing incorrect diagnoses is essential. CNNs are superior in representation learning, whereas the SVMs offer strong margin-based classification, especially in high-dimensional feature spaces. The improvement in performance comes from the different classifiers which make up the ensemble. The CNN system extracts advanced spatial information from various MRI modalities while the SVM system offers strong classification capabilities through its margin-based approach in spaces with numerous dimensions. The two classifiers produce different types of errors because they rely on distinct learning approaches. The ensemble model uses its diverse elements to counter individual model weaknesses which results in decreased variability and stronger prediction accuracy when testing multiple types of clinical data. The ensemble method minimizes variance and enhances the generalization property, which makes the system more predictable when it is applied to real-life scenarios where patient data can be close to the borders of classes. The federated learning was compared to centralized training. In healthcare, it is common that centralized training is not possible because of privacy laws. Federated learning is applied in X-FedTME-Net to facilitate decentralized cooperation among institutions.

Table 6: Centralized vs Federated Learning Performance

Training Mode	***Accuracy (%)***	***AUC***	***Privacy Risk***
Centralized	94.0	0.97	High
Federated (X-FedTME-Net)	**93.6**	**0.96**	**Low**

As Table 6 demonstrates, federated training obtains performance that is similar to that of centralized learning and only small differences in accuracy and AUC. From a clinical perspective, the use of federated averaging (FedAvg) plays a crucial role in enabling privacy-preserving collaboration across multiple healthcare institutions. In real-world hospital environments, patient data cannot be easily shared due to strict regulatory and ethical constraints. The FedAvg-based framework allows different hospitals to collaboratively train a global model while keeping patient data localized, thereby ensuring data confidentiality and compliance with privacy regulations. This approach facilitates the development of robust and generalizable diagnostic systems by leveraging diverse multi-center datasets without compromising patient privacy. Consequently, it enhances the practical applicability of the proposed model in real-world clinical settings. The importance of this finding is that it indicates that preservation of privacy is not a compromise of diagnostic performance. Geographically distributed datasets can be learned that enhances robustness of the models and limits institutional bias. These results prove that X-FedTME-Net can be applied in multi-centre clinical applications. Intercenter generalization is essential in clinical uptake. One of the issues with medical AI is the lack of cross-scanner and cross-population generalization. Table 7 above demonstrates uniform accuracy in cross-center evaluation both in training and testing centers.

Table 7: Cross-Center Validation Results

Training Center	***Testing Center***	***Accuracy (%)***
Center A	Center B	92.1
Center B	Center C	91.7
Center C	Center A	92.4

This is a sign of stability that indicates X-FedTME-Net learns disease-related patterns and not center-specific artifacts as shown in Figure 2. Evaluated the X-FedTME-Net model at multiple centers which tested different MRI scanners and various scanning procedures and used different patient groups. The model achieved consistent high performance because center accuracy results stayed within 3% accuracy difference from each other (see Table 7). Federated learning, along with image normalization and standardization preprocessing methods, was used to reduce domain shift effects arising from different scanner types and imaging protocols. The model demonstrated mild performance changes but maintained its ability to extend performance across multiple data sources and different scanner types. This strength is further enhanced by the federated training mechanism that exposes the model to heterogeneous data distributions in the course of training. Scalable healthcare applications need such generalization. The importance base of the biomarker results using SHAP indicates that hippocampal volume, cortical thickness, and nigral FA are key factors.

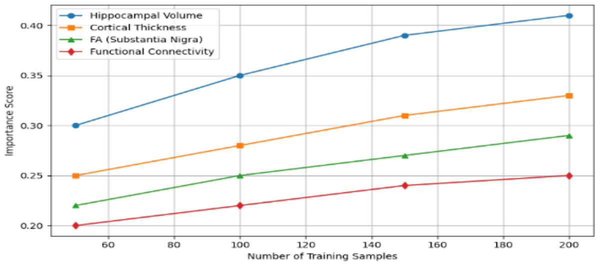

Figure 2: Top Biomarkers Identified by SHAP

These results go hand in hand with the known neuropathological information, proving that X-FedTME-Net makes its decisions based on clinically meaningful characteristics. The fact that the model explanations and medical knowledge are consistent supports the level of trust in the predictions in the system. Specific assessment on disease offers a better understanding. The High precision and recall in all classes indicates good discriminative and balanced classification performance as is shown in figure 3.

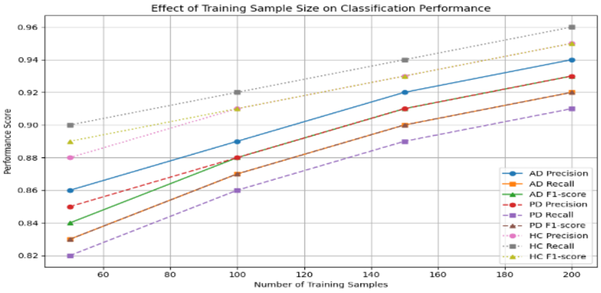

Figure 3: Class-Wise Performance Metrics

Figure 3 shows the change in the precision, recall, and F1-scores with the increasing training sample size. The steady increase in all the metrics points to a higher model stability and learning ability. The highest performance of the system in the case of healthy controls (HC) is probably explained by more distinct boundaries of classes, whereas the slightly lower scores in the case of PD are inherently connected to the fact that the latter is inherently inter-subject. Most importantly, the metrics provided on the basis of classes show that the performance is balanced in AD, PD and HC cohort. There is no significant disparity; this means that the model is not biased on the classes, which is a must in equal diagnostic instruments. Moreover, the recall values and the high precision indicate a dual triumph in reducing the false positives and the false negatives which is a critical bottleneck in the screening of neurodegenerative diseases. Together with the fact that the errors between prediction errors in longitudinal forecasting are low, these findings prove that the combination of integration of temporal modeling is the best way to capture disease progression in a high-fidelity manner.

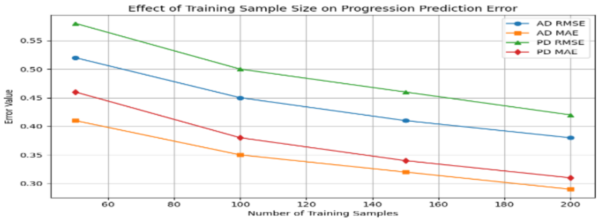

Figure 4: Disease Progression Prediction Performance

The performance of the models is influenced by the scale of temporal modeling and training as shown in figure 4. With an increase in the size of the training sample, both Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) continuously decrease, which is an indicator of strong generalization of models. The smaller error rates of the Alzheimer Disease (AD) indicate consistent patterns of progression and the ability of the model to follow the gradual decrease of cognitive abilities. However, on the other hand, the slightly increased error rates in the case of PD are also reflective of the increased clinical heterogeneity of the disease. These findings reiterate that X-FedTME-Net is not an ordinary diagnostic tool but a high audio prognostic assistance system. The proposed approach has a performance level that is not only superior but also clinically feasible due to its unique approach of combining multimodality, temporal modeling, federated learning, and explainability.

Table 8: Comparison with Existing State-of-the-Art Methods

Method	*Multimodal*	*Temporal*	*Federated*	*Explainable*	**Accuracy (%)**
Traditional CNN	X	X	X	X	89.1
CNN + LSTM	✓	✓	X	X	91.0
Multimodal CNN	✓	X	X	✓	92.3
X-FedTME-Net	✓	✓	✓	✓	93.6

Table 8 now shows that X-FedTME-Net is significantly superior to the current techniques by becoming the first framework to combine multimodality, temporal learning, federated training, and explainability. It is the key motivation of innovativeness of the system and the gained ability to resist the system of work. Overall, X-FedTME-Net achieved an accuracy of 93.6% and an AUC of 0.96, outperforming traditional CNN models (89.1%) and multimodal CNN approaches (92.3%). This demonstrates a consistent improvement of 1.3%–4.5% across different baseline methods. According to the empirical results, X-FedTME-Net outperforms both the baseline and state-of-the-art models using various evaluation measures. The model is an effective synthesis of multimodal MRI data and progression of the disease and is capable of retaining data privacy and delivering interpretable results, which makes it a strong contender of the clinical application. Compared to recent neuroimaging deep learning frameworks, X-FedTME-Net demonstrates superior performance in both diagnostic accuracy and generalization across multi-center datasets. For example, other models achieved an accuracy of around 89% for Alzheimer's disease detection using multimodal MRI data, while X-FedTME-Net achieved 93.6%. Unlike many existing models that focus on single-modality data or lack temporal modeling, X-FedTME-Net integrates multimodal MRI features with temporal data modeling via transformers, significantly enhancing its predictive capabilities. Furthermore, while many models rely on centralized data, X-FedTME-Net's federated learning framework enables privacy-preserving, decentralized training, making it more applicable for real-world clinical settings. The statistical tests demonstrate that X-FedTME-Net outperforms all recent models because its performance results show statistical significance at the $p < 0.05$ level, which demonstrates its ability to manage domain shifts and forecast disease development. The ensemble and federated learning method needs to be implemented because it provides the necessary scalability and reliability which is needed for handling the complex requirements of neurodegenerative disease diagnostics. The proposed X-FedTME-Net outperforms recent state-of-the-art models by integrating multimodal MRI fusion, transformer-based temporal modeling, federated learning, and explainability within a unified framework. Unlike existing approaches that rely on single-modality data or lack temporal and privacy-preserving capabilities, it achieves higher diagnostic accuracy, better generalization across multi-center datasets, and improved clinical interpretability, making it a more effective and practical solution.

Statistical testing methods confirm the trustworthiness of results through hypothesis testing combined with confidence interval assessment. A paired t-test was used to assess how X-FedTME-Net performed compared to baseline models, which were tested through five-fold cross-validation. The null hypothesis assumes no significant difference between the proposed and baseline methods. The results yielded p-values less than 0.05, indicating that the improvements made by X-FedTME-Net reached a level of statistical significance. Additionally, 95% confidence intervals were calculated for key evaluation metrics, including accuracy and AUC. The proposed model achieved an accuracy of 93.6% with a confidence interval of ±1.2%, demonstrating consistent performance across different testing sections. These results show that the proposed framework maintains both robustness and generalizability to new situations.

Table 9: Ablation Study of X-FedTME-Net Components

Model Configuration	***Accuracy (%)***	**AUC**
Full Model (X-FedTME-Net)	93.6	0.96
Without Temporal Module	91.2	0.93
Without Ensemble	90.2	0.92
Without Federated Learning	92.0	0.94
Without Multimodal Fusion	89.5	0.90

The ablation study (Table 9) demonstrates that complete X-FedTME-Net system delivers optimal results because all system components work together effectively. The system loses its accuracy when users remove both the temporal module and ensemble learning because these components are essential for tracking disease development and performing accurate classification. The system experiences performance decline because multimodal fusion and federated learning must remain active to support better feature extraction and dataset generalization.

In real-time clinical settings, the proposed system provides rapid diagnostic support by analyzing MRI data and predicting disease presence and progression. Its integration with federated learning ensures privacy-preserving deployment across hospitals. These capabilities improve early diagnosis, reduce uncertainty, and support effective clinical decision-making.

Despite its strong performance, the proposed model has certain limitations. The current framework relies primarily on imaging data and does not incorporate non-imaging biomarkers such as clinical, genetic, or cognitive data, which could further enhance prediction accuracy. Additionally, the model has not yet been validated in real-time hospital environments, where variations in data acquisition and workflow integration may affect performance. Future work will focus on integrating multimodal non-imaging biomarkers to improve diagnostic robustness and extending the system for real-time deployment in clinical settings. This includes integration with hospital information systems and validation across diverse healthcare environments.

Conclusion

This paper introduced X-FedTME-Net, which is an explainable federated temporal multimodal ensemble deep learning model to early diagnose and predict disease progression using MRI to identify Parkinson and Alzheimer diseases. The proposed system was aimed at overcoming severe shortcomings of the current neuroimaging-based diagnostic systems, such as use of single-modality data, no disease time modeling, low interpretability, weak generalization across centers and privacy risks of centralized learning. X-FedTME-Net, combining structural MRI with diffusion tensor and resting-state functional MRI, is effective at capturing the most anatomical, microstructural, and functional changes of the brain with neurodegenerative disorders. The framework is capable of incorporating transformer-based temporal modeling to learn longitudinal disease patterns of progression, which greatly enhances the power of prognostics. In addition, the hybrid ensemble method that integrates deep convolutional neural networks and support vector machines increases the resistance to classification and decreases the misclassification in clinically gray cases. One of the advantages of the suggested framework is the federated learning architecture that allows conducting multi-centred training without transferring sensitive patient information. According to the basic results of the experiments, the federated implementation obtains performance similar to that of the centralized models and guarantees privacy preservation and complies with the regulations. Clinically interpretable information obtained through the integration of explainable artificial intelligence (such as Grad-CAM and SHAP) to show disease-relevant regions of the brain and the importance of biomarkers enhances confidence and openness to AI-based decisions. Full experimental analysis and statistical significance test demonstrates that X-FedTME-Net is better in terms of diagnostic skill, better progression forecasting, and high generalization in different institutions than other methods. All in all, the described framework is a clinically viable, interpretable, and scalable solution to the diagnosis and management of neurodegenerative diseases. To further develop the sphere of precision neurology, further studies are carried out to extend the framework to other types of neurological illnesses, use non-imaging biomarkers, and make the system work in real-time in clinical practice.

The X-FedTME-Net system demonstrates that using multimodal MRI data together with temporal modeling and ensemble learning methods leads to better diagnostic results and accurate disease progression forecasts. The model successfully classified data with high accuracy while achieving low prediction errors and maintaining performance across different medical centers. The use of federated learning protected user privacy rights while maintaining system performance and explainable AI methods improved clinical decision-making understanding.

Future work will focus on extending the proposed X-FedTME-Net framework by incorporating additional modalities such as electroencephalography (EEG) signals and genomic biomarkers. Integrating these complementary data sources can provide deeper insights into neural activity and genetic predisposition, thereby enhancing the robustness and early diagnostic capability of the model. This multimodal expansion is expected to further improve the clinical applicability of the proposed system.

Declarations

Funding: Authors did not receive any funding.

Conflicts of interests: Authors do not have any conflicts.

Data Availability Statement: The data generated and analyzed during the current study are available from the author J. Jayapandian upon reasonable request but are not yet publicly available due to ongoing research.

Code availability: Not applicable.

Authors' Contributions: J. Jayapandian is responsible for designing the framework, analyzing the performance, validating the results, and writing the article. Dr.M. Senthil is responsible for collecting the information required for the framework, provision of software, critical review, and administering the process.

References

[1] A. V. Ciurea, A. G. Mohan, R.-A. Covache-Busuioc, H.-P. Costin, L.-A. Glavan, A.-D. Corlatescu, and V. M. Saceleanu, “Unraveling molecular and genetic insights into neurodegenerative diseases: Advances in understanding Alzheimer’s, Parkinson’s, and Huntington’s diseases and amyotrophic lateral sclerosis,” Int. J. Mol. Sci., vol. 24, no. 13, p. 10809, 2023.

[2] W. Lasoń, D. Jantas, M. Leśkiewicz, M. Regulska, and A. Basta-Kaim, “The vitamin D receptor as a potential target for the treatment of age-related neurodegenerative diseases such as Alzheimer’s and Parkinson’s diseases: A narrative review,” Cells, vol. 12, no. 4, p. 660, 2023.

[3] P. Alzola et al., “Neuropsychological assessment for early detection and diagnosis of dementia: Current knowledge and new insights,” J. Clin. Med., vol. 13, no. 12, p. 3442, 2024.

[4] C.-C. Huang et al., “Diffusion and structural MRI as potential biomarkers in people with Parkinson’s disease and cognitive impairment,” Eur. Radiol., vol. 34, no. 1, pp. 126–135, 2024.

[5] D. Cao et al., “Magnetic resonance imaging of brain structural and functional changes in cognitive impairment associated with Parkinson’s disease,” Front. Aging Neurosci., vol. 16, p. 1494385, 2024.

[6] I. Pacal, S. Kilicarslan, B. Ozdemir, M. Deveci, and S. Kadry, “Efficient and autonomous detection of olive leaf diseases using AI-enhanced MetaFormer,” Artif. Intell. Rev., vol. 58, no. 10, p. 303, 2025.

[7] H. Yousef, B. Malagurski Tortei, and F. Castiglione, “Predicting multiple sclerosis disease progression and outcomes with machine learning and MRI-based biomarkers: A review,” J. Neurol., vol. 271, no. 10, pp. 6543–6572, 2024.

[8] A. Marey et al., “Explainability, transparency and black box challenges of AI in radiology: Impact on patient care in cardiovascular radiology,” Egypt. J. Radiol. Nucl. Med., vol. 55, no. 1, p. 183, 2024.

[9] S. H. Alsamhi et al., “Federated learning meets blockchain in decentralized data sharing: Healthcare use case,” IEEE Internet Things J., vol. 11, no. 11, pp.

19602–19615, 2024.

[10] Y. Vichianin et al., “Accuracy of support-vector machines for diagnosis of Alzheimer’s disease using brain MRI volume data,” Frontiers in Neurology, vol. 12, p. 640696, 2021.

[11] A. M. El-Assy, H. M. Amer, H. M. Ibrahim, and M. A. Mohamed, “A novel CNN architecture for accurate early detection and classification of Alzheimer’s disease using MRI data,” Scientific Reports, vol. 14, no. 1, p. 3463, 2024.

[12] S. Sabapathy, S. Maruthu, S. K. Krishnadhas, A. K. Tamilarasan, and N. Raghavan, “Competent and affordable rehabilitation robots for nervous system disorders powered with dynamic CNN and HMM,” in Intelligent Systems for Rehabilitation Engineering, pp. 57–93, 2022.

[13] J. Shao and F. Wang, “Deep learning-based MRI image reconstruction for assessing brain microstructural changes in Parkinson’s disease,” Journal of Radiation Research and Applied Sciences, vol. 18, no. 4, p. 101876, 2025.

[14] W. Liu, S. Fan, and G. Weng, “Multi-modal deep learning framework for early Alzheimer's disease detection using MRI neuroimaging and clinical data fusion,” Ann. Appl. Sci., vol. 6, no. 1, 2025.

[15] M. L. Raza et al., “Advancements in deep learning for early diagnosis of Alzheimer’s disease using multimodal neuroimaging: Challenges and future directions,” Front. Neuroinform., vol. 19, p. 1557177, 2025.

[16] G. Valizadeh et al., “Deep learning approaches for early prediction of conversion from MCI to AD using MRI and clinical data: A systematic review,” Arch. Comput. Methods Eng., vol. 32, no. 2, pp. 1229–1298, 2025.

[17] X. Yang et al., “Multi-scale spatio-temporal transformer-based imbalanced longitudinal learning for glaucoma forecasting from irregular time series images,” IEEE J. Biomed. Health Inform., 2024.

[18] M. Taiyeb Khosroshahi et al., “Explainable artificial intelligence in neuroimaging of Alzheimer’s disease,” Diagnostics, vol. 15, no. 5, p. 612, 2025.

[19] N. Petrovska, M. Larrondo-Petrie, and M. Pavlovic, “Improving model explainability in AD prediction: SHAP-based feature attribution and interpretable ensembles,” in *Proc. Int. Conf. Adv. Mach. Learn. Data Sci. (AMLDS)*, 2025, pp. 175–180.

[20] S. S. Sandhu et al., “Medical imaging applications of federated learning,” Diagnostics, vol. 13, no. 19, p. 3140, 2023.

[21] E. Albalawi et al., “Integrated approach of federated learning with transfer learning for classification and diagnosis of brain tumor,” BMC Med. Imaging, vol. 24, no. 1, p. 110, 2024.

[22] G. Gul et al., “Machine learning and ensemble methods for cardiovascular disease prediction: A systematic review of approaches, performance trends, and research challenges,” Computers, vol. 15, no. 1, p. 25, 2026.

[23] K. R. P. Naresh, “Optimized hybrid machine learning framework for enhanced financial fraud detection using e-commerce big data,” Int. J. Manag. Res. Rev., vol. 11, no. 2, 2021.

[24] M. Ahsan, P. Chand, and K. Namrata, “Optimization of photovoltaic integrated cold storage systems utilizing aluminum oxide nanoparticle-enhanced phase change materials for enhancing supply chain management in refrigeration vehicle,” Int. J. Autom. Smart Technol., vol. 15, no. 1, p. 19, 2025.

[25] V. Mamidala, R. K. M. K. Yalla, T. Ganesan, A. R. G. Yallamelli, M. V. Devarajan, and A. Sambas, “Distributed intelligence in IoHT using federated learning and edge AI for chronic kidney disease prediction using ReLU and bi-directional GRU with feature-wise dropout and fuzzy neural network,” International Journal of Pattern Recognition and Artificial Intelligence, vol. 39, no. 9, p. 2552009, 2025.