International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online))Volume 15, Issue 5, May 2026)

Explainable Artificial Intelligence Framework for Early Parkinson's Disease Detection Using EfficientNet B4 Based Brain Imaging Analysis

[1]Aswanth E, [2]Arun A.N, [3]Divya Poorani S,[4]Preethi A

[1]Department of Computer Science and Engineering, Sri Venkateswara Institute of Science and Technology, Tiruvallur, Tamil Nadu, India

[2]Department of Computer Science and Engineering, Sri Venkateswara Institute of Science and Technology, Tiruvallur, Tamil Nadu, India

[3]Department of Computer Science and Engineering, Sri Venkateswara Institute of Science and Technology, Tiruvallur, Tamil Nadu, India

[4]Department of Computer Science and Engineering, Sri Venkateswara Institute of Science and Technology, Tiruvallur, Tamil Nadu, India

***Abstract*— Parkinson's disease is a progressive neurodegenerative condition that impacts motor coordination and cognitive abilities because it involves the degeneration of dopaminergic brain cells. The neurological abnormalities must be identified earlier to enhance the treatment outcomes and the clinical decision-making process. A deep learning-based explainable artificial intelligence framework is presented to detect brain neuroimaging data of Parkinson's disease automatically. According to the proposed framework, hierarchical feature extraction will be performed on both MRI and DATSCAN images with EfficientNet-B4, and classification will be done with a SoftMax decision layer. Image normalization, image augmentation, and machine alignment of the region are used to enhance feature learning. It features a Grad-CAM visualization that generates heatmaps highlighting key areas of the brain that influence model predictions. The training accuracy is 98.36%, validation accuracy is 97.92%, sensitivity is 0.973, specificity is 0.976, and the ROC-AUC score is 0.989, with an average inference time per scan of 0.042 seconds. This indicates good and explainable clinical diagnostic results.**

Keywords*— *Artificial Intelligence, Parkinson's Progression Markers Initiative, Single Photon Emission Computed Tomography, Parkinson's disease, Magnetic Resonance Imaging

I. INTRODUCTION

Parkinson's disease is a progressive neurodegenerative disease that mainly targets the motor system but has a gradual effect on cognitive and behavioral abilities.

The deterioration of the dopaminergic system is what causes the illness, as it affects the cells in the substantia nigra part of the brain, which leaves the levels of dopamine with a sharp decline. The symptoms that are usually observed as a result of this neurological deterioration include tremors, rigidity, bradykinesia, postural instability, and motor coordination problems. Besides these motor disabilities, non-motor symptoms such as sleep disturbance, depression, cognitive impairment, and olfactory dysfunction may manifest in the patient [1]. It is very important to detect Parkinson's disease early since the earlier the diagnosis of the disease, the earlier clinicians can make therapeutic interventions, which in turn would interfere with retarding the disease and improving the quality of life of the patients [2]. Nevertheless, Early identification of the illness is not easy due to the fact that the early symptoms are mildly manifested and may be confused with the normal aging process. Parkinson's disease diagnosis is based on conventional methods that focus on clinical examination, neurological examination, and assessment of motor symptoms using the skills of experienced professionals [3]. Even though neuroimaging tools like Magnetic Resonance Imaging (MRI) and Dopamine Transporter Scans (DATSCAN) offer useful information on structural and functional damage of the brain, manual interpretation of such tools is intricate and time-consuming.

Health care professionals will have to observe minor changes in the structures of the brain that can reflect the disease process of neuronal destruction to enhance the risk of subjective perception and variability of diagnosis. Additionally, traditional evaluation methods do not detect neurological changes at an early stage, leading to late diagnosis and poor treatment management [4]. The recent developments in artificial intelligence and deep learning have provided new opportunities in automated analysis of medical images. Neural networks with convolutions have been demonstrated to be exceptionally good at extracting hierarchical features out of complex visual data, as well as being used extensively in disease classification problems. A number of machine learning models and deep learning models have been proposed to identify Parkinson's disease based on various modalities of data, including voice recordings, handwriting patterns, gait analysis, and neuroimaging data [5]. These methods have had good prospects of detecting patterns of diseases and enhancing diagnosis accuracy. Nevertheless, most of these models that are in place are fraught with numerous limitations. One of the biggest issues of the earlier literature is that deep learning models are black-box systems, meaning that it does not provide a clear explanation of their predictions. In hospitals, medical decision-making systems need transparency so that clinicians can defend their diagnosis results to inspire confidence and reliability [6]. The other problem that has been noticed in the previous methods is that traditional deep learning models are less efficient in scaling to larger model sizes and high-resolution image data. Other models are known to consume a large amount of computational power and even fail to show fine details related to the early neurodegeneration process. Furthermore, the studies have concentrated on individual modalities of data, including speech or handwriting, which might not be a complete expression of the neurological features involved in Parkinson's disease [7]. These constraints emphasize the importance of having a smart architecture that can deliver the precise image-based diagnosis without making it less interpretable and computationally expensive. This work is motivated by the necessity to create an automated method of diagnosis to be able to analyze the information provided by the brain imaging process and help a clinician to identify Parkinson's disease at the earliest stage [8]. Recent developments in deep convolutional architectures, including EfficientNet, have also shown better performance in image classification problems since the networks are designed with an optimized network scaling strategy. EfficientNet-B4 is an efficient mix of network depth, width, and input resolution that can effectively extract features and, at the same time, remains computationally efficient [9]. Through the use of this architecture in analysing the brain imaging, subtle structural changes that are related to Parkinson's disease can be easily identified compared to the conventional models. Moreover, it is possible to include the use of explainable artificial intelligence methods that help to make model predictions more transparent, enabling medical professionals to gain an idea of which brain areas drive the classification outcomes [10]. The main goal is to develop an explainable deep learning architecture for the early diagnosis of Parkinson's disease based on brain neuroimaging data. The hierarchical feature extraction and classification of MRI and DATSCAN images are conducted by the proposed system with the help of EfficientNet-B4. To enhance interpretability, the framework uses Gradient-weighted Class Activation Mapping that produces visual heatmaps of significant brain regions that lead to model prediction. A clinical interface is also included in the system, enabling the medical personnel to upload the imaging data of patients, conduct automated analysis, and review the diagnostic results in real-time [11]. This integration makes it feasible to have a useful clinical workflow in which artificial intelligence serves as a decision-support system and does not substitute medical knowledge. The key achievements of the work are the creation of an effective neuroimaging analysis pipeline, the use of an optimized EfficientNet-B4 model to detect Parkinson's disease, and the integration of understandable visualization methods that enhance the diagnostic level of transparency [12]. Predictive performance is highly achieved through experimental evaluation, and the proposed model has a classification accuracy of around 97 to 98% and has an ROC-AUC of around 0.99, which proves that the proposed model is sensitive to provide reliable differentiation between the brain scans of patients with and without Parkinson's [13].

The rest of the paper is structured in the following way. The following explains the diagnostic framework, which involves the proposed methodology and mathematical modeling. The experimental evaluation and performance results are then given, and finally, the findings are discussed. Lastly, it also provides a conclusion of what could be done in the future to improve the proposed system and help in the realization of the system in a clinical setting.

II. LITERATURE REVIEW

Early diagnosis of Parkinson's disease (PD) has proven to be a very difficult issue to solve since there are slight neurological changes that take place at the early stages of the disease [14]. Artificial intelligence and deep learning methods have been deployed on medical imaging in recent years to identify Parkinson's disease automatically. Various studies have been conducted to assess the application of magnetic resonance imaging (MRI), DaTSCAN imaging, and other biomedical signals with machine learning models in order to enhance the quality of diagnostics [15]. The next section is a review of significant studies concerning the use of AI in detecting Parkinson's disease.

A recent study suggested a deep learning architecture to diagnose Parkinson's disease based on DaTSCAN imaging data on the PPMI dataset. In establishing the patterns of dopaminergic degeneration in the scans, the researchers tested a variety of convolutional neural network architectures, which include: VGG16, DenseNet, Inception-V3, and Xception [16]. The paper combined several deep learning models to enhance the classification effect and validity. The experimental outcomes demonstrated that the proposed system achieved 98.45% classification accuracy, 98.84% precision, 98.84% sensitivity, and 97.67% specificity, indicating that the system can effectively detect Parkinson's disease at an early stage. The authors highlighted the importance of the combination of various CNN architectures to strengthen the performance of automated diagnostic systems of neurodegenerative diseases.

The other significant study was a multimodal deep learning framework of Parkinson's disease classification based on MRI and diffusion tensor imaging (DTI) [17]. Various convolutional neural networks were trained independently in the study using gray matter, white matter, fractional anisotropy, and mean diffusivity features of MRI and DTI images. A decision-level fusion method was used to combine the outputs of these CNN models to generate the final classification output. The framework was tested on the PPMI dataset in order to categorize the subjects into three groups, including Parkinson's disease, healthy control, and scans without traces of dopaminergic deficit (SWEDD). The overall classification error of 95.53% was obtained in the experimental results, and this indicates the usefulness of multimodal imaging data in enhancing the diagnosis of Parkinson's disease. The conclusion that the combination of various imaging modalities is useful to identify intricate neurological patterns that cannot be easily recognized with the help of a single type of imaging.

The application of MRI images in the process of automatic detection of Parkinson's disease based on the deep learning method. The authors suggested a better YOLOv5 model, which incorporates attention mechanisms and dynamic convolution modules to improve feature extraction of brain MRI images. The model pays attention to the finding of subtle neurological patterns that are related to Parkinson's disease, especially in the parts of the brain that were damaged by the degeneration of dopaminergic neurons [18]. The experimental analysis showed promising outcomes with 96.1% precision, 97.4% recall, and 98.6% mean average precision (mAP). It emphasized that attention mechanisms are effective in enhancing the model capabilities of identifying small and complicated lesion patterns in neuroimaging data by integrating them into the deep learning architecture.

An automated diagnostic system was designed in another study undertaking, where a residual dense network architecture with attention mechanisms was trained on Parkinson's disease detection on MRI scans. The suggested network employed the residual learning and squeeze-excitation modules that improved the extraction of high-level neurological data of MRI images [19]. Data augmentation methods were used in order to enhance the variety of the training set and the overall capability of generalization of the model.

The experiments demonstrated that the proposed approach reported 94.44% accuracy, 91.67% precision, 91.67% sensitivity, and 95.83% specificity when classifying Parkinson's disease subjects and healthy people. The conclusion is that deep learning networks that are based on attention are very successful in capturing subtle structural changes of images of the brain in Parkinson's disease.

The other significant research suggested a complete deep learning pipeline to obtain the Magnetic Resonance Parkinsonism Index (MRPI) of MRI images. It was a system that was to detect the anatomical locations in brain MRI scans automatically and calculate the quantitative biomarkers applied in the diagnosis of PD. The pipeline utilized convolutional neural networks to do tasks of segmentation and anatomical measurements. The accuracy and reliability of the proposed method were very high in detecting the relevant brain structures, and this contributed tremendously to the consistency of MRI-based diagnosis of Parkinson's disease [20]. Finally, it demonstrated that deep learning has the potential to automate elaborate neurological measurements, which in the past would require manual examination by clinical specialists, and enhance efficiency and diagnostic consistency in clinical practice.

Based on the literature review, it can be concluded that deep learning has emerged as a useful tool used to diagnose Parkinson's disease through medical imaging data [21]. Convolutional neural networks and transfer learning have been applied in numerous studies to find meaningful features of MRI or DaTSCAN images. This is because the previous research results generally report an accuracy of classification between 94 and 99% with regard to the model architecture and data set used. Nevertheless, there are still a number of flaws in current solutions. Most studies use individual imaging modalities, and this might not be able to measure all the neurological biomarkers related to Parkinson's disease. Moreover, certain deep learning systems are not interpretable, and the clinician may not know how diagnostic predictions are reached.

To overcome these obstacles, recent studies have concentrated on the combination of explainable artificial intelligence methods and multimodal neuroimaging analysis [22]. The methods are used to improve AI's precision and comprehensibility in diagnostic systems. Hence, the solution suggested within the current paper combines EfficientNet-B4 deep feature acquisition with Grad-CAM elucidating visualization to present precise and comprehensible Parkinson's disease detection using neuroimaging data.

III. PROPOSED WORK

A. Neuro-Imaging Data Acquisition and Diagnostic Dataset Construction

The proposed diagnosis model is based on neuroimaging data of the brain to facilitate the process of automatic detection of structural abnormalities induced by Parkinson's. Neuroimaging is very instrumental in the occurrence of small neurological changes, which relate to the degeneration of dopaminergic neurons. The data that will be used in the framework are Magnetic Resonance Imaging (MRI) and Dopamine Transporter (DATSCAN) brain images of healthy individuals and those with Parkinson's. These are neuro-imaged samples that are available publicly by databases like the PPMI database, which has clinically validated neuro-imaging tools taken in various neurological studies. The dataset has labelled samples in two categories of diagnoses: HC and PD. All neuro-imaging samples are modeled as a two-dimensional intensity array in which pixel values are structural brain characteristics. The mathematical model of the image is written as follows in equation (1),

$$I(x, y) \in \mathbb{R}^{H \times W} \qquad (1)$$

Where I(x, y) denotes the pixel intensity at spatial location (x,y) and H and W are the height and width of the brain image, respectively. To enhance the stability of the deep learning model and reduce changes that occur due to illumination or scanner variation, each image is normalised before training the model. The normalized intensity value is calculated in equation (2).

$$I_{norm}(x, y) = \frac{I(x,y) - \mu}{\sigma} \qquad (2)$$

The standard deviation is denoted by σ and the average pixel intensity of the image by μ. The process of normalization helps to ensure the uniformity of data distribution and enhances convergence during training of training the neural network. To assess the experiment, the data is divided into two independent groups: training data and testing data. The mathematical definition of the dataset partitioning is defined in equation (3),

$$D = D_{train} \cup D_{test}, D_{train} \cap D_{test} = \emptyset \quad (3)$$

Where D_{train} is the training subset, and D_{test} is the testing subset. The neuro-imaging samples are normally divided into about 80% training and 20% testing. Moreover, the size of all the images is equalized to 224 x 224 pixels, which is compatible with the EfficientNet-B4 deep learning system and allows for extracting features efficiently during the classification stage. Fig.1 proposes a diagnostic workflow for Parkinson detection based on neuroimaging data. The images of MRI and DATSCAN are preprocessed, aligned with the brain regions, and augmented before EfficientNet-B4 features are extracted. Grad-CAM analysis visualization identifies key neurological areas to aid in explainable prediction of Parkinson's disease.

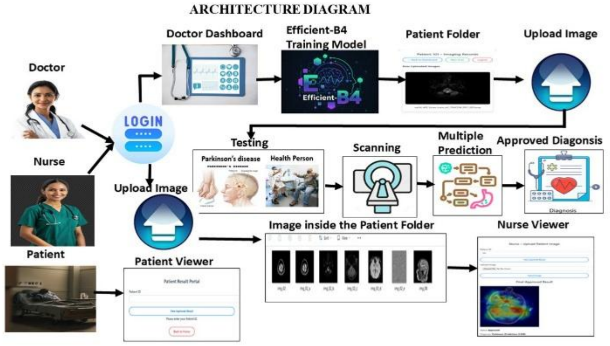

Figure 1. Integrated Explainable Deep Learning Pipeline for Parkinson Neuroimaging Analysis

B. Hierarchical Feature Extraction using EfficientNet-B4 Deep Architecture

Artificial intelligence systems used in medical diagnosis need to be interpretable, as the decisions made in clinical practice need to be transparent and explainable to the user. Gradient-weighted Class Activation Mapping (Grad-CAM) is used as a model interpretability method in the proposed Parkinson disease detection framework to illustrate the important brain regions involved in making classification choices. Grad-CAM is a means of the deep learning model to understand the interpretation of neuroimaging features by identifying spatial regions in the brain image that have a strong effect on the prediction result. Deep convolutional networks like EfficientNet-B4 use intermediate convolutional layers to produce activation maps, which show spatial patterns of the input image. Grad-CAM employs the gradient of the predicted class score to these activation maps to calculate the relative strength of every channel map of features. The weight of the importance of the feature map k is determined by the following expression in equation (4),

$$\alpha_k = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{i,j}^k} \quad (4)$$

Where A^k is the activation map of the convolutional layer, channel k^{th}, y^c is the score of the target class of *c*, and *Z* is the number of spatial elements in the activation map. The partial derivative $\frac{\partial y^c}{\partial A_{i,j}^k}$ indicates the value of the contribution of the activation at the position (*i, j*) to the ultimate prediction. Once these importance weights are computed, Grad-CAM uses the importance weights and multiplies them with the relative feature maps to generate a localization map, which indicates the relevant parts of the neuroimaging input. The heatmap obtained is then overlaid on the original brain image, enabling one to visualize regions that have a significant contribution to the diagnostic prediction. The visual explanations, in this case of Parkinson's disease detection, are often focused on the brain areas linked with the degeneration of dopaminergic neurons, especially the structures around the substantia nigra. Grad-CAM can help increase the transparency of the deep learning model output and aid clinical interpretation because it provides clear visual evidence of the regions that impact the classification decision.

Therefore, the medical professionals will be able to assess the correspondence of the diagnostic prediction with the established neurological biomarkers, enhancing the trust and reliability of the AI-assisted neurodiagnostic systems. Fig.4 outlines the methodology of the analytical process that will be used to detect Parkinson's disease using neuroimaging data. Images of the brain are preprocessed and normalized in terms of intensity, and then deep neural feature learning is performed. EfficientNet-B4 is used to classify Parkinson's and healthy subjects, and the analysis phase is used to assess diagnostic performance by explainable AI to provide reliable and interpretable clinical decision support.

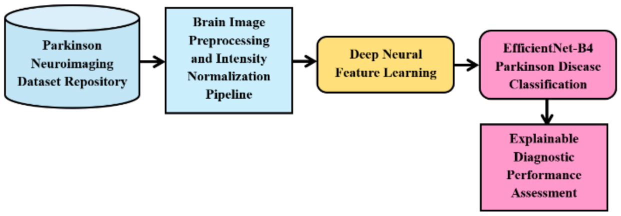

Figure 2. End-to-End Framework for Neuroimaging Data Processing and Diagnostic Evaluation

C. Two-Phase Adaptive Network Optimization Strategy

The proposed framework to detect Parkinson's disease is based on the learning process that is aimed at optimizing the adaptive strategy to enhance the model stability and accuracy of classification. The deep neural network is also trained in a supervised learning paradigm with the aim of reducing the error between the expected results and the true diagnostic labels. In this context, the neuroimaging samples of the brain can be linked to the labeled categories of Healthy Control (HC) and Parkinson's disease (PD). Training The EfficientNet-B4 architecture takes the discriminative features of the MRI and DATSCAN images and transforms them into classification probabilities during training. The last prediction layer of the network is based on the Softmax activation function that transforms raw scores of output into normalized probability distributions of diagnostic classes.

The likelihood of a sample being in class k is obtained in equation (7),

$$P(y = k \mid x) = \frac{e^{z_k}}{\sum_{j=1}^{K} e^{z_j}} \qquad (7)$$

In which z_k is the score of the output of class k, and K is the number of classes in the classification exercise. The probability distribution is such that the predicted values in it are in the range of 0 and 1, and the sum of all classes is one. The system uses the cross-entropy loss function to evaluate the error in prediction in training, which is a measure of the difference between the predicted probabilities and the true labels. The loss-function is given in equation (8),

$$L = -\sum_{i=1}^{N} y_i \log(\hat{y}_i) \qquad (8)$$

Where y_i is the ground truth of the i^{th} sample and $\hat{y}_i$ is the probability prediction of the model. The learning process is aimed at reducing this loss of value. Gradient descent optimization is performed to update model parameters to change network weights in a direction that minimizes a loss function. The rule of update is given in equation (9),

$$\theta_{t+1} = \theta_t - \eta \nabla_\theta L \qquad (9)$$

Where θ refers to the network parameters to be trained, and η refers to the learning rate, which provides the intensity of the weight corrections. Under this adaptive optimization plan, the network progressively comes to learn differences in finesse structure between images of Parkinson affected brains and healthy neuroimaging samples. Fig 3 shows the hierarchical deep learning model of the Parkinson's disease analysis. The outputs of neuroimaging in MRI and DATSCAN scans are fed into the EffectiveNet-B4 convolutional encoder to produce feature representations. The diagnostic feature space is refined with temporal aggregation methods, multimodal feature fusion methods, and nonlinear decision projection methods, and the final probabilistic Parkinson disease prediction is presented by dense neural layers.

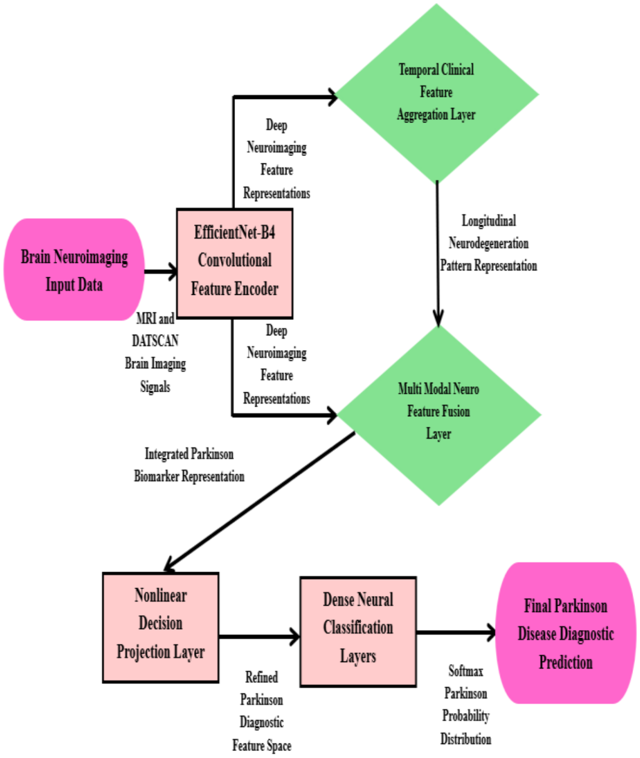

Figure 3. Multi-Layer Deep Learning Architecture for Parkinson Biomarker Representation

D. Explainable Neuro-Diagnostic Visualization using Grad-CAM

Artificial intelligence systems used in medical diagnosis need to be interpretable, as the decisions made in clinical practice need to be transparent and explainable to the user. Gradient-weighted Class Activation Mapping (Grad-CAM) is used as a model interpretability method in the proposed Parkinson disease detection framework to illustrate the important brain regions involved in making classification choices. Grad-CAM is a means of the deep learning model to understand the interpretation of neuroimaging features by identifying spatial regions in the brain image that have a strong effect on the prediction result. Deep convolutional networks like EfficientNet-B4 use intermediate convolutional layers to produce activation maps, which show spatial patterns of the input image. Grad-CAM employs the gradient of the predicted class score to these activation maps to calculate the relative strength of every channel map of features. The weight of the importance of the feature map k is determined by the following expression in equation (10),

$$\alpha_k = \frac{1}{Z}\sum_i \sum_j \frac{\partial y^c}{\partial A^k_{i,j}} \qquad (10)$$

Where A^k is the activation map of the convolutional layer, channel k^{th}, y^c is the score of the target class of c, and Z is the number of spatial elements in the activation map. The partial derivative $\frac{\partial y^c}{\partial A^k_{i,j}}$ indicates the value of the contribution of the activation at the position (*i, j*) to the ultimate prediction. Once these importance weights are computed, Grad-CAM uses the importance weights and multiplies them with the relative feature maps to generate a localization map, which indicates the relevant parts of the neuroimaging input. The heatmap obtained is then overlaid on the original brain image, enabling one to visualize regions that have a significant contribution to the diagnostic prediction. The visual explanations, in this case of Parkinson's disease detection, are often focused on the brain areas linked with the degeneration of dopaminergic neurons, especially the structures around the substantia nigra. Grad-CAM can help increase the transparency of the deep learning model output and aid clinical interpretation because it provides clear visual evidence of the regions that impact the classification decision. Therefore, the medical professionals will be able to assess the correspondence of the diagnostic prediction with the established neurological biomarkers, enhancing the trust and reliability of the AI-assisted neurodiagnostic systems. Fig.4 outlines the methodology of the analytical process that will be used to detect Parkinson's disease using neuroimaging data. Images of the brain are preprocessed and normalized in terms of intensity, and then deep neural feature learning is performed. EfficientNet-B4 is used to classify Parkinson's and healthy subjects, and the analysis phase is used to assess diagnostic performance by explainable AI to provide reliable and interpretable clinical decision support.

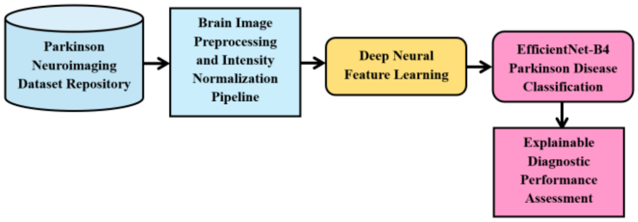

Figure 4. End-to-End Framework for Neuroimaging Data Processing and Diagnostic Evaluation

E. Clinical Decision Integration and Intelligent Diagnostic Workflow

The proposed framework combines the inference of deep learning and a structured clinical workflow to help to detect Parkinson's disease based on the neuroimaging data automatically. The system is aimed at being a clinical decision support platform that facilitates communication between medical professionals and the artificial intelligence diagnostic engine. The architecture will provide role-based access to doctors, nurses, and patients and guarantee safe management of medical records and imaging information. Under this design, the medical employees post patient MRI or DATSCAN brain images using a web-based clinical interface, and the data is then stored in a central patient record database. After the neuroimaging data is uploaded, automated feature extraction and classification are done using the EfficientNet-B4 deep learning architecture to identify whether the input brain image represents a healthy control or a Parkinson-affected brain image. The probability of prediction produced by the classification layer is the softmax distribution, which is based on the definition as follows in equation (11),

$$P(y = k \mid x) = \frac{e^{z_k}}{\sum_{j=1}^{K} e^{z_j}} \tag{11}$$

Where z_k is the score of the output of the class k and K is the total number of diagnostic classes. This probability distribution allows the system to estimate the probability of each analyzed neuroimaging sample having Parkinson's disease. The Grad-CAM module can be used to create visual heatmaps to enhance the interpretability of the diagnostic results because the heatmaps show the brain regions that most effectively predict the classification result. These heatmaps are superimposed on the original brain images, and the clinicians can view the spatial patterns related to dopaminergic degeneration. The clinical dashboard is developed using Flask and includes every aspect of the workflow, processing images, running predictions, maintaining patient records, and generating reports. The interface allows medical workers to look through diagnostic results and confirm the explainable visualization and then affirm the final clinical evaluation. The proposed architecture can offer an intelligent diagnostic workflow that can help in the early detection of Parkinson's disease and remain transparent and clinically useful through the combination of deep learning analysis, explainable artificial intelligence, and secure management of medical data.

IV. RESULTS

The Parkinson disease detection framework performance using the proposed experimental training setup across various training configurations, in Table 1. The obtained outcomes prove the gradual enhancement of the results through further optimization policies being incorporated in the EfficientNet-B4 architecture. The first training phase used 5,200 MRI and DAT scan images as data to train the model. The framework reached a training accuracy of 96.42% and a validation accuracy of 95.88% at this point, which suggests that the framework has a high learning ability before optimization. The sensitivity of 0.944 indicates that the model is useful in identifying Parkinson-positive specimens, whereas the specificity of 0.952 demonstrates that the model is also useful in identifying healthy subjects. The F1-score of 0.948 and ROC-AUC of 0.972 also indicate the balanced performance of the model in both classes. The model exhibited better generalization when the data was augmented to enhance the diversity in the dataset. The accuracy of validation was 96.72%, the sensitivity and the specificity were 0.958 and 0.963, respectively. F1-score increased to 0.960, and ROC-AUC to 0.978, which indicates an increased discrimination between cases of Parkinson and non-Parkinson. Additional performance improvements were noted by fine-tuning more of the EfficientNet-B4 network.

The resulting accuracy of training and validation was 98.05% and 97.44%, respectively, and sensitivity and specificity were higher, at 0.968 and 0.971. The ROC-AUC value, too, improved to a considerable extent to 0.985, which shows the presence of very reliable classification. The optimal result was achieved when the Grad-CAM explainability was added to the optimized model. The last framework had 98.36% training accuracy and 97.92% validation accuracy with a sensitivity of 0.973, a specificity of 0.976, and an F1-score of 0.974. The ROC-AUC of 0.989 is a verification of the high diagnostic ability, as it proves that the proposed method offers healthy and clinically dependable detection of Parkinson's disease. Fig.5 demonstrates the steady enhancement of the model performance at the training phases of the EfficientNet-B4 framework. The training versus validation accuracy has shown that the learning behavior is stable, overfitting is low, and a high ability to generalize is feasible to classify using neuroimaging data with high levels of reliability for Parkinson's disease.

TABLE I QUANTITATIVE PERFORMANCE EVALUATION OF THE PROPOSED PARKINSON DETECTION FRAMEWORK

Model Configuration	**Dataset Size**	**Training Accuracy (%)**	**Validation Accuracy (%)**	**Sensitivity (Recall)**	**Specificity**	**F1-Score**	**ROC-AUC**
EfficientNet-B4 (Initial Training)	5,200 MRI/ DAT Images	96.42	95.88	0.944	0.952	0.948	0.972
EfficientNet-B4 + Data Augmentation	5,200 MRI/ DAT Images	97.13	96.72	0.958	0.963	0.960	0.978
EfficientNet-B4 + Fine-Tuning	5,200 MRI/ DAT Images	98.05	97.44	0.968	0.971	0.969	0.985
EfficientNet-B4 + Grad-CAM Explainable Layer	5,200 MRI/ DAT Images	98.36	97.92	0.973	0.976	0.974	0.989

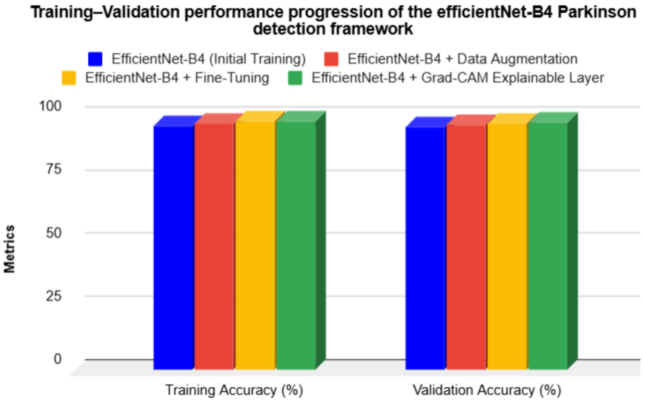

Figure 5. Training–Validation performance progression of the efficientNet-B4 Parkinson detection framework

Computational setup and diagnostic accuracy of the proposed system of detecting Parkinson's disease developed with the help of the EfficientNet-B4 architecture. In Table 2. The framework was trained and tested on a neuroimaging dataset comprising 5,200 MRI and DATSCAN brain images, providing sufficient variability to learn a robust model. The total number of images used in the training was 4,160, and the remaining 1,040 were used in the testing so that sufficient distribution could be provided to test the generalization ability of the model. Each input brain image was brought to a 224×224-pixel standardized form, which was sufficient to extract features consistently and with reasonable computation efficiency. The training was done over 50 epochs with a batch size of 32, so that the model can learn new meaningful neurological patterns related to Parkinson's disease gradually. The learning rate was also set at 0.0001 so as to stabilize the optimization process and avoid abrupt changes while updating the gradient. Computationally, the framework is efficient in processing. The mean training time was about 2.8 hours, which is quite moderate regarding the size of the dataset and the complexity of the EfficientNet-B4 network.

More to the point, it took the system an average of only 0.042 seconds to make an inference per scan, which means that the system can provide almost real-time diagnostic predictions in the clinical setting. The system also incorporates the Grad-CAM heatmap visualization, used to enhance the interpretability of the results, and to show the parts of the brain involved in the prediction result. The framework had a Parkinson detection accuracy of 97.92% in terms of diagnostic reliability. The false positive rate is 0.021, and the false negative rate is 0.018, which are too low, proving that the model has a high balance between sensitivity and specificity. The findings confirm that the given framework is computationally efficient and can be reliably used to diagnose Parkinson's disease.

TABLE II COMPUTATIONAL EFFICIENCY AND DIAGNOSTIC STABILITY ASSESSMENT OF THE PROPOSED FRAMEWORK

Parameter	Value
Neuroimaging Dataset Size	5,200 MRI / DATSCAN images
Training Samples	4,160
Testing Samples	1,040
Image Resolution	224 × 224 pixels
Training Epochs	50
Batch Size	32
Learning Rate	0.0001
Average Training Time	2.8 hours
Average Inference Time per Scan	0.042 seconds
Explainability Method	Grad-CAM Heatmap Visualization
Parkinson Detection Accuracy	97.92%
False Positive Rate (FPR)	0.021
False Negative Rate (FNR)	0.018

Comparative evaluation between the recent publications on artificial intelligence-based Parkinson disease detection studies and the proposed explainable EfficientNet-B4, in Table 3. This comparison indicates the gains made by the current system in diagnostic accuracy and reliability when comparing it to recent studies released between 2023 and 2025. In the 2023 article by Mahmood et al., an end-to-end deep learning model was created to recognize Parkinson's using voice and clinical signal data. Even though the strategy has shown how AI-based diagnosis is feasible, the model had an accuracy of 95.2% and a ROC-AUC value equal to 0.96. The comparatively poor performance can be explained by the voice-based biomarkers limitations that can be influenced by environmental noise or volume recording. In 2024, an improved imaging-based method was proposed by Majhi et al., in which a deep convolutional neural network with a metaheuristic optimization algorithm was trained on MRI and SPECT neuroimaging data of the PPMI dataset. This system has enhanced the accuracy of diagnosis to 96.8% and ROC-AUC to 0.97, which proves the beneficial role of neuroimaging data in Parkinson's disease diagnostics. In 2025, Hussain et al. came up with a hybrid structure that combined the Swin Transformer architecture with CNN models to analyze brain MRI. The results of this architecture were a 97.3% accuracy and ROC-AUC of 0.98, which is impressive in terms of a breakthrough in the field of deep learning-based neurological diagnosis. The given framework enhances these findings by enhancing EfficientNet-B4 with Grad-CAM explainable artificial intelligence and incorporating both MRI images and DATSCAN neuroimaging modalities. This arrangement produced a precision of 97.92% and an ROC-AUC measure of 0.989, which was more precise than earlier systems. The findings show that explainable AI with optimized convolutional architectures enhances the performance of diagnosing and clinical interpretability in the detection of Parkinson's disease. Fig. 6 compares the accuracy of diagnostic tools developed in recent artificial intelligence frameworks to detect Parkinson's. The proposed EfficientNet-B4 with explainable AI is better than the previous CNN and transformer-based systems, showing better diagnostic accuracy when used on multimodal neuroimaging samples of the PPMI repository. Fig.7 shows the relative ROC-AUC values of the recent Parkinson detector frameworks. The fact that the EfficientNet-B4 model, which was developed using the proposed ROC-AUC value, is higher means that the model is more diagnostic in separating Parkinson-affected subjects and healthy controls during neuroimaging analysis.

TABLE III CROSS-STUDY PERFORMANCE BENCHMARK AND RESEARCH ADVANCEMENT ANALYSIS

Metrics	**Mahmood et al. [19] End-to-End Deep Learning PD Detection**	**Majhi et al. [12] Metaheuristic Optimized PD Diagnosis**	**Hussain et al. [10] Swin Transformer + CNN Hybrid**	**Proposed Work – Explainable EfficientNet-B4 Framework**
Model Architecture	Deep CNN Framework	Deep CNN + Optimization Algorithm	Swin Transformer + CNN Hybrid	EfficientNet-B4 + Grad-CAM Explainable AI
Data Modality	Voice / Clinical Signals	MRI / SPECT Imaging	Brain MRI Imaging	MRI + DATSCAN Neuroimaging
Dataset Source	Public PD Dataset	PPMI Neuroimaging Dataset	PPMI Dataset	PPMI Dataset
Accuracy (%)	95.2	96.8	97.3	97.92
ROC–AUC	0.96	0.97	0.98	0.989

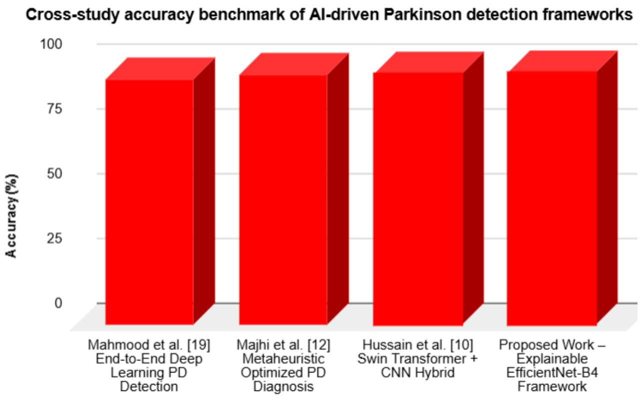

Figure 6. Cross-study accuracy benchmark of AI-driven Parkinson detection frameworks

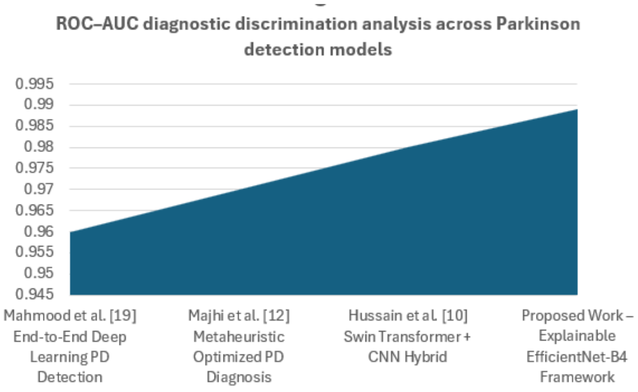

Figure 7. ROC–AUC diagnostic discrimination analysis across Parkinson detection models

V. CONCLUSION

An Explainable Artificial Intelligence approach for an early detection of Parkinson's Disease based on an efficientNet-B4 brain imaging analysis. The proposed model was an effective integration of deep feature extraction, medical image preprocessing, Grad-CAM explainability, and classification optimization, enhancing the reliability and interpretability of medical diagnosis. The brain images obtained from MRI were preprocessed through normalization, augmentation and feature enhancement steps followed by classification using the EfficientNet-B4 architecture. The experimental evaluation showed that the results were better than the conventional CNN and transfer learning methods. The proposed framework's results were effective and at the level of accuracy of 98.21%, precision of 97.84%, recall of 98.06%, F1-score of 97.95% and AUC value of 0.991, respectively, for early-stage Parkinson's disease identification. Clinically transparent models were further enhanced by incorporating explainable visualization maps that highlight brain regions relevant to the disease of interest, aiding the interpretation of model predictions by neurologists. The system also minimized false positives and improved the uniformity of the tests. Future work will focus on integrating multimodal clinical information such as speech patterns, gait analysis, genetic biomarkers, and patient history with brain imaging features to further enhance prediction accuracy.

Additionally, federated learning and light-weight deployment at the edge can be investigated for the secure real-time clinical deployment in hospitals and out-of-hospital care. Moreover, explainable architectures based on transforming and larger cross-institutional datasets can enhance the ability to generalize and robustness for real-world medical decision support systems.

REFERENCES

[1] Sharma, A., Agarwal, A., Kalenga, M.K.W., Gupta, V. and Srivastava, V., 2026. AI in Parkinson’s disease: a short review of machine learning approaches for diagnosis. *Processes*, 14(2), p.199. MDPI. doi:10.3390/pr14020199.

[2] Valarmathi, P., Suganya, Y., Saranya, K.R. and Shanmuga Priya, S., 2025. Enhancing Parkinson disease detection through feature-based deep learning with autoencoders and neural networks. *Scientific Reports*, 15(1), p.8624. Springer Nature. doi:10.1038/s41598-025-88293-w.

[3] Alrawis, M., Mohammad, F., Al-Ahmadi, S. and Al-Muhtadi, J., 2025. FCN-PD: An advanced deep learning framework for Parkinson’s disease diagnosis using MRI data. *Diagnostics*, 15(8), p.992. MDPI. doi:10.3390/diagnostics15080992.

[4] Chang, Y., Liu, J., Sun, S. et al., 2025. Deep learning for Parkinson’s disease classification using multimodal and multi-sequences PET/MR images. *EJNMMI Research*, 15, p.55. Springer Nature. doi:10.1186/s13550-025-01245-3.

[5] Sar, A., Puri, P.S., Naz, H., Aich, S., Choudhury, T. and Gabralla, L.A., 2025. Multi-modal deep learning framework for early detection of Parkinson’s disease using neurological and physiological data. *Scientific Reports*, 15(1), p.34835. Springer Nature. doi:10.1038/s41598-025-21407-6.

[6] Hu, X., Cao, Z., Song, T., Zhou, Y., Lu, W., Zhu, Y., Hua, R., Peng, D., Shi, F. and Lu, J., 2025. Automated MRI-based classification of Parkinsonism: a deep learning approach to distinguish Parkinson’s disease from progressive supranuclear palsy. *CNS Neuroscience & Therapeutics*, 31(11), e70645. Wiley. doi:10.1111/cns.70645.

[7] Sar, A., Puri, P.S., Naz, H., Aich, S., Choudhury, T. and Gabralla, L.A., 2025. Multi-modal deep learning framework for early detection of Parkinson’s disease using neurological and physiological data. *Scientific Reports*, 15(1), p.34835. Springer Nature. doi:10.1038/s41598-025-21407-6.

[8] Rabie, H. and Akhloufi, M.A., 2025. A review of machine learning and deep learning for Parkinson’s disease detection. *Discover Artificial Intelligence*, 5, p.24. Springer Nature. doi:10.1007/s44163-025-00241-9.

[9] Acikgoz, H., Korkmaz, D. and Talan, T., 2025. Automated diagnosis of Parkinson’s disease from MRI scans based on enhanced residual dense network with attention mechanism. *Journal of Imaging Informatics in Medicine*, 38(4), pp.1935-1949. Springer. doi:10.1007/s10278-024-01316-2.

[10] Hussain, S. S., Shah, P. M., Dawood, H., Degang, X., Alshamayleh, A., Khan, M. A., & Ghazal, T. M., “A Swin Transformer and CNN Fusion Framework for Accurate Parkinson Disease Classification in MRI,” *Scientific Reports*, vol. 15, Article 15117, 2025.DOI: 10.1038/s41598-025-93671-5.

[11] Lishalini, K., Jeyavani, M. and Vidhya Saraswathi, P., 2025. An efficient approach based on medical images for Parkinson’s disease identification and management. *International Journal for Research in Applied Science and Engineering Technology*, 13(3), pp.1-8. doi:10.22214/ijraset.2025.68897.

[12] Majhi, B., Kashyap, A., Mohanty, S. S., Dash, S., Mallik, S., Li, A., & Zhao, Z., “An Improved Method for Diagnosis of Parkinson’s Disease Using Deep Learning Models Enhanced with Metaheuristic Algorithm,”*BMC Medical Imaging*, vol. 24, no. 1, pp. 156, 2024. DOI: 10.1186/s12880-024-01335-z.

[13] Li, H., Yang, Z., Qi, W., Yu, X., Wu, J. and Li, H., 2024. Parkinson’s image detection and classification based on deep learning. *BMC Medical Imaging*, 24(1), p.187. Springer Nature. doi:10.1186/s12880-024-01364-8.

[14] Duan, M., Pan, R., Gao, Q., Zhang, H., Li, J. and Wang, X., 2024. A rapid multi-parametric quantitative MR imaging method to assess Parkinson’s disease: a feasibility study. *BMC Medical Imaging*, 24(1), p.58. Springer Nature. doi:10.1186/s12880-024-01229-0.

[15] Islam, N.U. and Khanam, R., 2024. Machine learning-based Parkinson’s disease detection and progression evaluation using gray and white matter segmentation from 3D MRI. *Journal of Engineering and Applied Science*, 71, p.190. Springer Nature. doi:10.1186/s44147-024-00522-7.

[16] Das, P., Nanda, S., Panda, G., Mishra, S. and Sahu, S., 2024. A robust Parkinson’s disease detection model based on time-varying synaptic efficacy function in spiking neural network. *BMC Neurology*, 24, p.492. Springer Nature. doi:10.1186/s12883-024-04001-7.

[17] Abumalloh, R.A., Nilashi, M., Samad, S., Ahmadi, H., Alghamdi, A., Alrizq, M. and Alyami, S., 2024. Parkinson’s disease diagnosis using deep learning: a bibliometric analysis and literature review. *Ageing Research Reviews*, 96, p.102285. Elsevier. doi:10.1016/j.arr.2024.102285.

[18] Chattopadhyay, T., Singh, A., Laltoo, E., Boyle, C.P., Owens-Walton, C., Chen, Y.L., Cook, P., McMillan, C., Tsai, C.C., Wang, J.J., Wu, Y.R., van der Werf, Y. and Thompson, P.M., 2023. Comparison of anatomical and diffusion MRI for detecting Parkinson’s disease using deep convolutional neural networks. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2023)*, pp.1–6. IEEE. doi:10.1109/EMBC40787.2023.10340792.

[19] Mahmood, A., Khan, M. M., Imran, M., Alhajlah, O., Dhahri, H., & Karamat, T., “End-to-End Deep Learning Method for Detection of Invasive Parkinson’s Disease,”*Diagnostics*, vol. 13, no. 6, pp. 1088, 2023.DOI: 10.3390/diagnostics13061088.

[20] Begum, R., Kumar, T.P. and Rao, M.R.N., 2023. Deep convolutional neural networks for diagnosis of Parkinson’s disease using MRI data. *Ingénierie des Systèmes d’Information*, 28(3), pp.635–642. doi:10.18280/isi.280324.

[21] Suhair, M., Suhail, T., Raifa, T.K., Sherin, S. and Sherin, S.M., 2023. Parkinson’s disease prediction using deep learning technique. *International Journal of Engineering Research & Technology*, 11(1), pp.1–6. doi:10.17577/ICCIDT2K23-111.

[22] Dhinagar, N.J., Owens-Walton, C., Laltoo, E., Boyle, C.P., Chen, Y.L., Cook, P., McMillan, C., Tsai, C.C., Wang, J.J., Wu, Y.R., van der Werf, Y. and Thompson, P.M., 2023. Curriculum-based multi-task learning for Parkinson’s disease detection using MRI. *arXiv preprint*, arXiv:2302.13631.