DLFM: Leveraging Parkinson's Disease Detection using AI Deep Learning Fusion Model for Precise Diagnosis

MAJIDHA FATHIMA K. M.[1], J. PRAVEENCHANDAR[2], R. JENNIE BHARATHI[3],
N. NAGA SARANYA[4], RAMACHANDRAN A.[5], NITHYA DORAIRAJAN[6]
[1]Department of Computer Science and Engineering,
Sri Krishna College of Engineering and Technology,
Coimbatore,
INDIA

[2]Department of Artificial intelligence and Machine Learning,
Karunya Institute of Technology and Sciences,
Coimbatore,
INDIA

[3]Department of Electronics and Communication Engineering,
Saveetha Engineering College,
Chennai,
INDIA

[4]Department of Computer Applications,
Saveetha College of Liberal Arts & Science,
Chennai,
INDIA

[5]Department of Artificial Intelligence and Data Science,
Saveetha Engineering College,
Chennai,
INDIA

[6]Department of ECE,
Panimalar Engineering College,
Chennai,
INDIA

Abstract: - Parkinson's disease (PD) presents substantial difficulties owing to its progressive course and heterogeneous symptoms, thus an accurate and timely diagnosis is vital for optimal management. In this paper, we introduce a mixed AI Deep Learning Fusion Model (DLFM) as a new approach to improve PD detection. The DLFM combines the features obtained from both biomedical voice measurements and clinical examinations by utilizing the LeNet-5 and DenseNet architectures; thereby providing a robust mechanism for ensuring maximum diagnostic accuracy. The approach of our approach in regard to balancing and the process are as follows, to train the DLFM model with appropriate PD case classification we pre-process the dataset and extract some key features then, the DLFM paradigm that incorporates DL and fusion techniques develops a strong basis for precise PD diagnosis, enabling early intervention and personalized treatments. It shows better performance than conventional diagnostic solutions because it combines both LeNet-5 and DenseNet architectures, which give the ability to detect complex patterns and correlations between input data. Additionally, the DLFM seamlessly integrates information from multiple sources together and provides a robust diagnosis of PD status. Our results highlight the promise of AI-driven approaches to transform PD patient diagnosis and care. With a model accuracy of 97.23%, it demonstrates an excellent capability for distinguishing between PD and healthy patients. Adopting this DLFM model with high accuracy will provide medical doctors and researchers with a useful complementary tool to assist in earlier diagnosis of PD and

timely therapeutic action. The DLFM model rapidly advances the path to an improved method for diagnosing PD through the use of state-of-the-art DL methodologies, resulting in improved patient outcomes and quality of life.

Key-Words: - AI Deep Learning Fusion Model (DLFM), Parkinson's Disease (PD), Lenet-5, DenseNet, Clinical Assessments, Biomedical Voice Measurements, Deep Learning.

Received: May 23, 2024. Revised: December 26, 2024. Accepted: January 21, 2025. Published: May 9, 2025.

1 Introduction

Parkinson's Disease (PD) continues to be a major challenge in neurology, affecting millions of people around the globe. James Parkinson first described this progressive neurodegenerative disorder in 1817, in his classic monograph "An Essay on the Shaking Palsy." Spirit 1: Its inexplicable disposition, peculiarly delicate in texture, or as that rare combination of genetic predisposition, environmental factors, and neuronal dysfunction continues to baffle medical researchers. PD, fundamentally, emerges as a series of motor and non-motor symptoms that interfere with the finely tuned interplay of movement, cognition, and emotion, [1]. Classical motor manifestations comprise bradykinesia (restriction of movement), rigidity, tremor, and postural imbalance, collectively leading to severe disability and reduced quality of life. However, although motor impairment is the most recognizable feature of Parkinson's disease, a wide spectrum of non-motor manifestations including cognitive impairment, autonomic dysfunction, sleep disorders, and psychiatric disorders also contribute to the burden of this complex disease, [2].

The substantia nigra pars compacta (SNpc) are a major regulator of motor function, and the progressive loss of SNpc dopaminergic neurons is the hallmark of Parkinson's disease pathogenesis. As these neurons die off, the striatum suffers a dopaminergic deficit, which disrupts the fine-tuned balance between the excitatory and inhibitory outputs of the basal ganglia circuitry and gives rise to motor symptoms. However, the extent of neurodegeneration is much broader than the dopaminergic system alone including widespread neuronal death, synaptic dysfunction, and the accumulation of misfolded protein aggregates, particularly α-synuclein, in the form of Lewy bodies, [3], [4]. While we still don’t know exactly what causes Parkinson’s disease, many researchers believe it’s the result of a combination of environmental assaults and inherited predispositions. Mutations in familial versions of the illness have been tied to genes such as SNCA, LRRK2, Parkin, and PINK1, shedding light on important biochemical processes that are critical for neuronal homeostasis, mitochondrial function, and protein degradation44. In addition, environmental exposures - including pesticides, heavy metals, viruses, and head trauma - have also been linked to increased risk of developing PD, highlighting the complex relationships between genetic predisposition and environmental exposure.

The diagnosis of PD is still essentially clinical based on the identification of typical motor signs and their responsiveness to dopaminergic treatment. However, the search for holistic biomarkers that can uncover the disease in prodromal stages and accurately monitor the progression of AD has become a hot research topic, [5]. Moreover, advanced neuroimaging modalities, the biology of cerebrospinal fluid, and emerging molecular biomarkers offer hope of elucidating the complex pathophysiology of PD and ushering in the era of personalized medicine tailored to the needs of the individual patient. Therapeutically, levodopa remains the mainstay of pharmacological management, providing symptomatic relief through the restoration of depleted brain dopamine levels. Levodopa, however, is associated with adverse motor effects after long-term administration, which requires adding drug regimens or device-based therapies, such as dopamine agonists, monoamine oxidase inhibitors, or deep brain stimulation, to ameliorate motor complications. In addition, emerging research initiatives are investigating innovative neuroprotective approaches to preventing or delaying the ineluctable advancement of neurodegeneration in PD, ranging from gene-specific therapies and stem cell transplantation to the realignment of approved drugs with neuroprotective potential, [6], [7].

Neuroimaging elucidation of the underlying pathophysiology of PD has been enhanced through the application of techniques such as magnetic resonance imaging (MRI), positron emission tomography (PET), and single-photon emission computed tomography (SPECT). Healthy versus diseased conditions are depicted in Figure 1. These

tools provide distinct views of structural, functional, and molecular brain alterations that reflect dopaminergic impairment, necroinflammation, and protein aggregation in the disease process, [8], [9]. SPECT, PET, and MRI each have its drawbacks, which may represent only part of the intricate neurodegeneration cascade in PD. However, this is a challenge that potentially can be solved by the fusion of multukkit data, which, in many recent works and both works, it is reported that AI and DL approaches are being used. PD detection is further improved by the complementarity offered by these modalities through the use of fusion models that exploit positive points in both methods. Deep Learning (DL) algorithms, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs), are capable of extracting sophisticated patterns and features from complex datasets, making them suitable for fusion model development.

Fig. 1: Healthy and Diseased

Such a strategy could be employed to train separate DL networks on each imaging modality and clinical data source and fuse the outputs later. For brain structure identification: train CNNs on MRI scans, for analysis of clinical data such as patient demographics, medical histories, and severity ratings to train RNNs. These individual output networks are fused using late fusion or decision-level fusion to produce a global diagnosis, [10], [11]. Other ways use single deep learning (DL) architecture to co-learn multiple modalities by extracting and fusing features jointly. Instead of relying solely on late fusion, which combines predictions made separately from different modalities, we adopt two main components: First, we combine features at low levels, a compelling way to learn hierarchically when the model can still access raw data, collecting both local and global dependency information from modalities. Early fusion models can make better use of the complementary information within each modality by optimizing the fusion process jointly, thus improving diagnosis performance.

Thus, developing AI fusion models for the detection of PD requires large-scale, multi-modal datasets from diverse patient populations and imaging protocols, [12], [13]. Each of these datasets is used as a resource for training, validating, and testing the performance of fusion models, which enables their generalization and robustness in different clinical settings. Additionally, incorporating expert knowledge and clinical expertise will further help in selecting relevant features, fine-tuning the model architecture, and making sense of the diagnostic predictions in a clinical context. Fusion model predictions need to be validated against gold standard criteria for the diagnosis of PD (such as clinical diagnosis by a movement disorder specialist or post-mortem neuropathological confirmation) to ensure the accuracy of the model output. Common evaluation metrics in the study of the accuracy of fusion models for diagnosis comparing them with classical diagnostic methods are effectiveness metrics, [14], [15].

Apart from diagnostic accuracy, AI fusion models provide the advantages of automation, scalability, and efficiency for PD detection. The recent models focusing on the diagnosis of the progression have potential impacts on consequently accelerating the diagnosis, early intervention, and relieving the burden on healthcare providers by adding it to the diagnostic workflow. In addition, the application of AI fusion models in clinical practice can also provide more consistent and standardized diagnostic criteria, decreasing the variation in nearby different medical units, and enhancing the commonness in patient management. Although AI fusion models have promising potential for PD detection, challenges, and limitations remain for clinical utility. These challenges encompass the necessity for strong validation in real-world clinical environments, the integration of multimodal data from diverse sources, and ethical questions involving data privacy, patient consent, and algorithmic transparency. Furthermore, continuous investigations are warranted to clarify the biological mechanisms responsible for imaging and clinical biomarkers integrated into fusion models and to improve their long-time prognostication robustness and reliability. With this in mind, we suggest using AI DLFM as a new measurement tool for PD diagnosis. The DLFM combines LeNet-5 and DenseNet architectures, leveraging their

abilities to extract relevant features from various data modalities. The DLFM is developed to integrate biomedical voice measurements and clinical assessment information to provide an overall assessment of PD status. Advancing toward precision medicine for PD diagnosis and management through the synergy of advanced technologies and clinical expertise.

2 Related Works

PD was the most common neurological disorder, and ranked second in mortality and disability, globally. PD incidence has doubled over the last 15 years. Early correct detection of PD a daunting assignment in and of itself was imperative to make sure persons might proceed to live with little interference. Nevertheless, PD was not diagnosed early due to a global shortage of qualified neurologists. Medical illness diagnosis using AI-based machine learning (ML) algorithms has become ubiquitous over the last few decades. The speedy, accurate diagnosis was not what these approaches provided. Overall, ML-related models were not sensitive enough to detection. A total of 195 voices from 13 male and female patients had been snared in this investigation. Recording of each patient lasted from one to thirty-six seconds, average being six recordings per patient. For these recordings a head-mounted AKG-C420 microphone from Industrial Acoustics Company (IAC) had been used in an IAC soundproof studio, [16]. Motivated by research regarding the diagnostic value of speech and vocal impairments in PD, this dataset was collected. One of the most frequent causes of model overfitting and generalization errors was an unbalanced dataset, where there were many samples in one class compared to few samples in another class. All classes contained the same number of samples, and this balancing of datasets made the working of the model better and reduced the overfitting problem. Four performance metrics were used to test the recommended hybrid model which includes recall, accuracy, precision, and f1 score. With the random oversampling method on a balanced data set the proposed model achieved an accuracy, recall, and f1 of 100% and 100%, 97%, 99% AUC, and 91% f1 score, respectively when evaluated with the SMOTE technique.

PD is a degenerative neurological disorder characterized by progressive loss of a specific population of neurons in the motor cortex of the brain. However, progress in PD was difficult to predict because medical applications used to assess the severity of PD had drawbacks. This has led some recent studies with a focus on PD as a possible tool to automatically identify PD from MRI scans, [17]. Its goal, in the first place, is to automatically augment MRI images based on DCGAN (deep convolutional generative adversarial networks). Second, it extracted and classified features from pre-trained models (for example, VGG16, Xception, or InceptionV3) Third, it developed a hybrid model where it classified InceptionV3 features and a QSVM model. The study used sixty MRIs, thirty MRIs from healthy subjects and thirty from PD subjects. This proposed application of the DCGAN approach in our publication contributed to the expansion of the dataset. Thus, 1000 images which comprised normal and PD images were added to the dataset. These models were VGG16, Xception, and InceptionV3, among others, which had been pre-trained and were used for automatic feature extraction and classification. The best accuracy rate of all of the pre-trained models was achieved using the InceptionV3 architecture with a score of 74%. A feature-quantum support vector machine (QSVM) hybrid model with inceptionV3 was built for PD and healthy group detection. The proposed model monitors fair prediction accuracy (87.5%) and precision (95%). It also achieved respectable results in F1-measure (89%), recall (84%), and both. The hybrid model had the lowest false negative and false positive rates at 4 and 1, respectively. The hybrid model hence might have indicated a potential useful diagnostic tool for autonomous PD prediction.

PD was a progressive neurological disorder that presented as muscle rigidity, limb tremor, and reduced balance. And, early PD diagnosis warranted proper treatment and better healthcare facilities for such patients. Several disorders could have been positively influenced by the use of non-invasive and cheap computer-aided diagnostic (CAD) systems. PD was primarily assessed through handwriting. Researchers have examined various ML approaches for early illness detection. Unfortunately, most of those manual feature extraction techniques had been so-so at best. Since early diagnosis was critical in the management of this chronic disease, a DL model could have aided with this. The hybrid method included data augmentation, pre-trained Convolutional Neural Networks (CNNs)-based feature extraction, optimization-based feature selection, and (machine learning) ML-based classification for PD detection, [18]. The article's first step was to classify all three types of handwriting photographs with six pre-trained CNN models, with the VGG16 framework outperforming the rest. Step two of the process had frozen the

layers of the VGG16 network, before fitting Binary Grey Wolf Optimization (BGWO) to the data in order to find the best subset of features. The method describes that had used Support Vector Machine (SVM) to achieve 99.8 percent accuracy of classification. The technique had been tested in the benchmark New Hand PD dataset Indeed, it had been shown in experimental results that the suggested strategy outperformed state-of-the-art approaches in detecting PD, mainly because the proposed strategy maximizes the sustainable accuracy while minimizing the feature subsets.

PD had a debilitating effect on millions of people around the world. Early signs included a mild sensation of weakness and a propensity to have involuntary jerky movement of the limbs especially the hands, arms, and head. PD was diagnosed based on motor symptoms. Academicians suggested various remote monitoring tests, listing several of the advantages, such as early diagnosis, simplicity of use, cost worthiness, etc. People with PD often have speech problems. Speech signals of the patient could be used to detect the early stage of any disease. Based on artificial intelligence [19], a being put forward a method used for PD diagnosing that used speech signals. DL approaches were then employed to classify PD using scalogram images generated from the continuous wavelet transform of the speech samples. The scalograms were tested with several DL algorithms. In phase one of the project, a number of different classifiers were used – including a hybrid system using majority voting, ResNet50, AlexNet and GoogleNet. Another domain of interest was a deep feature fusion approach that used DenseNet and NasNet. A range of metrics was used to evaluate performance. Furthermore, utilizing stratified 10-fold cross-validation, the deep feature fusion method obtained an F1 score, as well as an accuracy of 0.95, which is a 38% enhancement compared to the ablation study. They all have three contributions (1) to explore scalogram photos with the comprehensive/profound evaluation of DL models and profound feature combination in identifying PD. Parkinson's disease is a neurological movement disorder in which the depletion of dopamine over time leads to typical symptoms. As voice problems were found in 90% of the patients with PD, diagnostic decision support systems were developed to assist in the early detection of the disease. Given the influence of language and geography on speech data, the potential to classify PD, regardless of speakers' linguistic demographics, and generalize to different populations and external patient data was crucial for a diagnostic decision support system. The goal of that work was a language-independent PD classification model leveraging data and Variational Mode Decomposition (VMD). The cross-lingual validation consisted of a DL framework combined with the deconstruction of sustained phonation in Italian and Spanish into VMD modes. This dataset was used to evaluate the performance and generalizability of the method to other data in different languages, [20]. Finally, we looked for potential gender bias and evaluated the real-world efficacy of the proposed strategy. The key findings suggest that such a method was able to achieve cross-lingual accuracy between 65% - 80%. Values obtained for the adapted cross-linguistic validity assessments exceeded those observed for any of the vowel-based studies for the other models investigated, including those that had previously used transfer learning on the target language. The method showed an accuracy of 90% to 95% on the same dataset. A maximum of 63% generalizability was achieved when our models were evaluated during realistic recording conditions (on another subject) of a separate dataset. With the proposed method of training a DL classifier by utilizing VMD modes, the results were found to be gender-agnostic, performing comparably regardless of the speaker's gender. No situation in terms of the recording settings, language diversity in the data or number of languages used would deviate the still consistent results of the applied VMD modes to a DL classifier recommendation, as the proved outcome for all conducted trials. This study emphasized on importance of assessing the generalizability of results as well as pointing out the impact of bias related to limited testing and negligence with training data. The results demonstrate that the proposed method may achieve a robust, language-independent, and generalizable model to assist in PD identification using voice recognition.

3 Methodology

This method includes several essential steps based on the AI DLFM that provides PD detection. The dataset is first processed to ensure data quality and consistency. Then, we apply feature extraction techniques to extract informative features from biomedical voice measurements and clinical assessments. The DLFM model is trained on the extracted features to learn the latent pattern of PD pathology by utilizing the integrated LeNet-5 and DenseNet of the extracted features. The model parameters are trained with gradient-based optimization algorithms during training, and regularization techniques during training are also

used to prevent overfitting. After training, an appropriate performance metric is used to evaluate the model and determine its accuracy in detecting PD. The DLFM model is an effective method for diagnostic accuracy, therefore the proposed method provides a good typology for PD diagnosis. As shown in Figure 2, the overall structure of the proposed model.

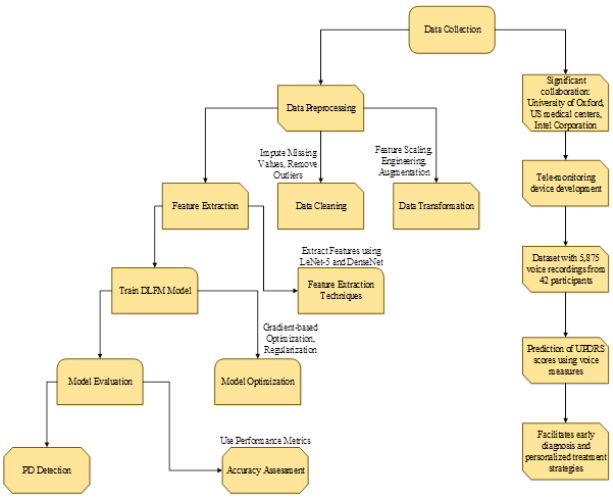

Fig. 2: Architecture of Proposed Model

3.1 Data Collection

The Parkinson dataset visible on Kaggle is the output of a significant joint work of Intel Corporation, ten United States clinical centers, and the University of Oxford. We designed this tele-monitoring device so it would be able to capture patients' speech signals automatically in their homes, providing a non-invasive way to track disease progression. The dataset includes biomedical voice measurements from 42 participants for a six-month period, making it a significant contribution toward understanding PD progression. All the participants made several voice recordings, leading to a total of 5,875 voice recordings. Description: The dataset includes multiple demographic and clinical features, such as age, gender, and time since baseline recruitment, as well as motor and total Unified PD Rating Scale (UPDRS) scores. The dataset also includes 16 biomedical voice measures that capture specific features of speech that inform us about PD.

This dataset is intended primarily to support the use of the 16 speech variables as predictors of motor and total UPDRS scores. These are used by clinicians as fundamental clinical markers for assessing motor function and the general severity of PD within patients. Example voice attributes of interest for each study STTs relevant to voice TTs include qualitative classifications such as type and severity, which can be analyzed by ML algorithms or statistical modeling techniques to identify linkages between voice TTs and disease progression (turn again). This could result in improved early diagnosis and more customized treatment regimens. The recording dataset is provided in ASCII CSV format and has a line per each instance of voice recording. The first column indicates the subject number. In order to aggregate recordings per patient to note, each patient provides more than 200 recordings, offering plenty of longitudinal data to parse. The attribute information includes subject ID, gender, age, time (years) since recruitment, motor UPDRS score and total UPDRS score, nonlinear complexity, and fractal scaling measures. Other Voice Measures related to biomedical aspects include jitter, shimmer, noise-to-harmonics ratio (NHR), harmonics-to-noise ratio (HNR), and total UPDRS score.

Such detailed attribute information allows researchers to explore multifaceted relationships, where UPDRS scores are related to specific voice features while further analyzing how places where these deviations exist vary with the physiological mechanisms behind the manifestation of PD in speech. Additionally, the dataset can serve as a foundation for feature engineering and feature selection, model validation, and predictive algorithm development aimed at enhancing clinical decision-making and patient care. Overall, the presence of this dataset on Kaggle serves as an important asset for the scientific community working on PD study and highlights the potential of ML in healthcare solutions.

3.2 Data Preprocessing

Data pre-treatment is a mandatory solution of the deeds for utilizing PD detection with AM DLFM, only to guarantee that the data are appropriate for preparation and modeling. During the preprocessing stage, which involves several steps, the data is cleaned, modified, and readied for feature extraction and classification.

Step 1: The raw dataset retrieved from Kaggle should be examined for anomalies, outliers, missing numbers, etc. The classic methods for impugning missing values are to use averages, medians, or modes, but more advanced approaches like using other data points nearby to interpolate are also possible. These robust statistical approaches help in the identification of and removing or treating the outliers which can play a disruptive role in the learning process of the model.

The next step after data purification is feature scaling, which is where we ensure that we do not

have one feature dominating and outshining the others in the modeling process because of its size. This consists of normalizing the features to a common scale. This step is so important while building ML algorithms because the performance of ML models can be affected by the scale of the features and this can help in the convergence of an algorithm faster.

$$Imputed_{Value_{i,j}} = mean(X_{i,j}) \quad (1)$$

where $Imputed_{Value_{i,j}}$ is the imputed value for missing data at row i and column j and $X_{i,j}$ represents the non-missing values in the same column.

In addition, this may find a feature engineering used for extracting new features, or transposing already present features training new patterns in the data. PD detection: Domain-specific knowledge can be applied to feature engineering for PD detection. For instance, audio raw signals can be processed by signal processing engineering where features such as voice characteristics (for example: jitter, shimmer, harmonics-to-noise ratio (HNR)) can be collected.

$$X'_{i,j} = \frac{X_{i,j} - mean(X_j)}{std(X_j)} \quad (2)$$

where $X'_{i,j}$ is the scaled value of feature j in row i, $X_{i,j}$ is the original value, mean (X_j) is the mean of feature j, and $std(X_j)$ is the standard deviation of feature j.

Model input size also advances to improve the model performance, with padding option that expands input sentences by adding dummy words or characters at the beginning and the end of the sentence. Additionally, this can be extended by the application of certain Time-warping, Noise-injection, and Pitch-shifting techniques in order to generate samples without any collection.

$$X'_{i,j} = X_{i,j} + noise \quad (3)$$

where $X'_{i,j}$ is the augmented value of feature j in row i, $X_{i,j}$ is the original value, and noise is a random value added to introduce variation.

The preprocessing of data contains cleaning, transforming and enhancing attributes from the dataset. This stage establishes the groundwork for constructing correct and trustworthy AI models for PD detection by validating the data and guaranteeing its quality and relevance.

3.3 Feature Extraction

Feature extraction plays a formative role by converting raw input data from the CSV dataset to an intermediary set of meaningful and discriminative features that capture the relevant patterns for the PD detection task using the AI DLFM. The ability of ML algorithms to effectively process data can be inhibited in various domains but especially so in healthcare domains where data can be high-dimensional and complex by nature, hence requiring proper techniques for feature extraction prior to implementing ML algorithms.

In the context of PD detection, one of the goals of feature extraction is to extract relevant features from the CSV dataset containing biomedical measurements and clinical assessments. Such measures yield powerful biologic and clinical correlates of PD. From the raw CSV input data, features such as age, gender, motor UPDRS scores, total UPDRS scores, and a set of biomedical voice measures (jitter, shimmer, noise-to-harmonics ratio (NHR), nonlinear complexity measures, etc.) can be extracted.

X be the dataset matrix that has been processed beforehand; where is the Dimension of Feature extraction with a CNN consists of feeding the input data among convolutional and pooling layers:

$$H^{(l)} = \sigma\left(W^{(l)} * H^{(l-1)} + b^{(l)}\right) \quad (4)$$

Here, $H^{(l)}$ represents the feature maps at layer l, $W^{(l)}$ represents the weights, $b^{(l)}$ represents the biases, $*$ denotes the convolution operation, and σ represents the activation function.

Moreover, augmenting our feature set with domain-specific features relevant to PD clinical assessment (motor UPDRS and total UPDRS scores) backlls our feature set with useful descriptors of motor function and general disease severity as evaluated by a clinician. In addition, a deep learning (DL)-based approach to feature extraction methods offers the capability of automatically extracting hierarchical representations from the input data, which can capture intricate relationships and patterns concealed within raw comma-separated value (CSV) data. LeNet-5 and DenseNet are CNNs that have been shown to do well as they are top performers for feature extraction when doing structured data.

In the proposed DLFM setup, we leverage an adapted version of both LeNet-5 and DenseNet networks as the base architecture with regard to the extraction of features from both the original data and CSV dataset derived from clinical assessment data. LeNet-5 has its convolutional layers which

can be used to captivate low-level patterns and relations from the data. By having denser connections, in contrast, DenseNet is able to pull high-level abstract features from the input image by inspiring its neurons with information from many layers. In addition, the transfer learning techniques also provide the advantage of tuning the pre-trained CNN architectures over the PD dataset, taking the benefits from the knowledge acquired via large datasets and accentuating the efficiency of the feature extraction procedure. Overall, these feature extraction steps lay a stable foundation for accurate and robust AI models for PD detection, so that informative, discriminative features can be extracted from the input CSV data efficiently.

3.4 DLFM Model for Classification

The third and final step of the procedures for exploiting PD detection, which is based on the AI DLFM, is the classification of the extracted features to predict the presence/existence of the PD. In this phase, we train the DLFM model, which consists of the features generated from the biomedical measurements, as well as the clinical evaluation data obtained from the CSV dataset.

First, this data is split into three pieces to be models trained and evaluated on training and validation data sets, Test. In the training set, a large amount of labeled data are used to train the model to fit the basic patterns and associations between input features and the outcome variable (PD status) of the DLFM model. The validation set is used to both tune hyperparameters of the model to reduce overfitting and monitor its performance during training to assure generalization to new data. Third, the test set allows a model's predicted accuracy on unseen data to be evaluated in an unbiased way, [7].

Let X_1 and X_2 represent the features extracted from the biomedical voice measurements and clinical assessment data, respectively. Fusion of features can be performed by concatenating the feature vectors:

$$X_f = [X_1, X_2] \tag{5}$$

Here, X_f represents the fused feature vector containing features from both streams.

The DLFM model architecture is a combination of LeNet-5 and DenseNet architectures adapted for feature extraction from the raw CSV dataset. DenseNet with its dense connections and LeNet-5 with its convolutional layers are respectively good at perceptual features detection for the clinical assessment and high-level context to summarize them.

Next, the features from the two streams are merged using fusion operations like concatenation or attention mechanisms (i.e., they provide the model a streamlined perspective to combine pieces of information from multiple modalities for better predictions). Additional fully connected networks utilize the fused features to learn a mapping from the higher level feature representations to the target variable, here the PD status, through a series of nonlinear transformations.

The classification model can be represented as a neural network with fully connected layers:

$$Z^{(l)} = \sigma\left(W^{(l)} Z^{(l-1)} + b^{(l)}\right) \tag{6}$$

Here, $Z^{(l)}$ represents the activations at layer $l, W^{(l)}$ represents the weights, $b^{(l)}$ represents the biases, and σ represents the activation function.

Similar to any other training, optimization algorithms (e.g., Adam or stochastic gradient descent (SGD)) can be used to optimize model variables via minimizing an appropriate loss function (e.g., binary/categorical cross-entropy). Overfitting can be a concern here; therefore, we may apply some regularisation techniques, e.g., weight decay, and dropout to avoid overfitting and help with the generalization of the model.

The effectiveness of the model, after the training phase, is assessed by a suitable assessment metric on the test set using metrics such as accuracy, precision, recall, F1-score, and area under the receiver operating characteristic (ROC) curve. These parameters allow us to understand the sensitivity, specificity, and prediction of the detection of PD from the model.

By means of gradient-based optimization methods such as SGD with backpropagation, we can optimize the variables of the model:

$$\theta = \theta - \alpha \nabla_\theta J(\theta) \tag{7}$$

Here, θ represents the model parameters, α represents the learning rate, and $J(\theta)$ represents the loss function.

Thus, the classification step of the DLFM methodology trains and tests a DL model that is able to perform an accurate and precise diagnosis of PD based on features derived from both biomedical measurements and clinical assessment data. It can also provide the initial underlying theory for higher experimentation using this DLFM model, which would serve a great role in securing a better clinical diagnosis and treatment on time.

4 Results and Discussions

The proposed model DLFM, short for AI DLFM, is implemented and executed within the PyCharm integrated development environment (IDE), leveraging the capabilities of a Windows operating system environment. It has all the tools and features needed to build and deploy DL models like DLFM, which makes PyCharm the IDE of choice. For running the DLFM model integrated using Python and Tensorflow, the hardware configuration was an Intel® Core™ i5 14400 CPU featuring a substantial cache of 20MB and a turbo boost frequency reaching as high as 4.70 GHz. Intel Core i5 Processor (Description: The Intel® Core™ i5 processor chips are considered efficient processors with good performance catering allrounder computes for modest DL model training and inferencing).

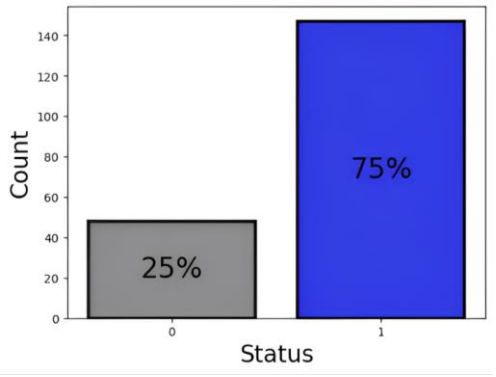

Fig. 3: Data Distribution in Dataset

6GB of RAM (Random Access Memory) is available, ensuring that the memory capacity is sufficient to handle the computational needs of the DLFM model during training and inference. While 6GB may be considered modest by some standards a more powerful processor makes this very much suitable for a first DL todo. This HW Combo represents a fair trade-off between performance and cost, making the DLFM model effectively run without missing any computational power. The processing power of the Intel Core i5 processor along with the 6GB RAM will enable it to efficiently perform the demanding tasks required by the proposed DLFM model including feature extraction, classification, and fusion to provide reliable and lucrative prediction for PD. The experimentation environment, combining PyCharm IDE with a suitable hardware setup, creates a strong foundation for applying the DLFM model and fostering an effective research space for utilizing DL techniques in accurate PD diagnosis and management.

The suggested approach for implementing PD detection using the AI DLFM is comprehensive and is a strategy to effectively diagnose and manage PD. Using domain expertise and advanced ML algorithms, this method comprises multiple processes such as data pre-processing, feature extraction, and classification. The goal is to get useful insights from the information that is available. Figure 3 depicts the distribution of data within the dataset.

To begin with, the raw dataset sourced from CSV files undergoes a rigorous preprocessing phase. It is essential to handle missing values, scale features, and maybe engineer new ones in this initial stage to ensure the data's quality and integrity. Missing values within the dataset are addressed using appropriate imputation techniques, such as mean imputation or interpolation, to maintain the integrity of the dataset and minimize information loss. Then feature scaling techniques are used to scale the features into a common scale so that no particular feature serves as a dominating factor in the modeling process due to its higher magnitude (obviously if we have numerical values in the dataset). Feature engineering employs the ability to transform existing or create new features which can better capture the hidden trends within the data and therefore improve model prediction accuracy. Figure 4 shows the data exploration.

When the data has been pre-processed, the following step in the approach is feature extraction, which entails extracting useful features from the cleaned-up dataset. This requires feature extraction from clinical evaluations as well as biological voice measures for PD detection purposes. We can through biomedical voice tests, identify minor alterations in speech features such as freezing of gait that are associated with PD. These include jitter, shimmer, and noise-to-harmonics ratio (NHR). Likewise, clinical evaluation measures such as motor UPDRS and total UPDRS scores provide critical insights into patients' motor functioning and features of their disease severity. Note that the feature extraction step utilizes DL methods, specifically a hybrid of DenseNet and LeNet-5 architectures, to automatically extract structural features from the input data. It captures even complex relationships and patterns that wouldn't be apparent if one were just looking at the raw data. The Correlation Matrix is shown in Figure 5.

The DLFM model presented, which is based on a generic fusion of LeNet-5 and DenseNet architectures, formed the foundation of our classification stage, as illustrated in the methodology described.

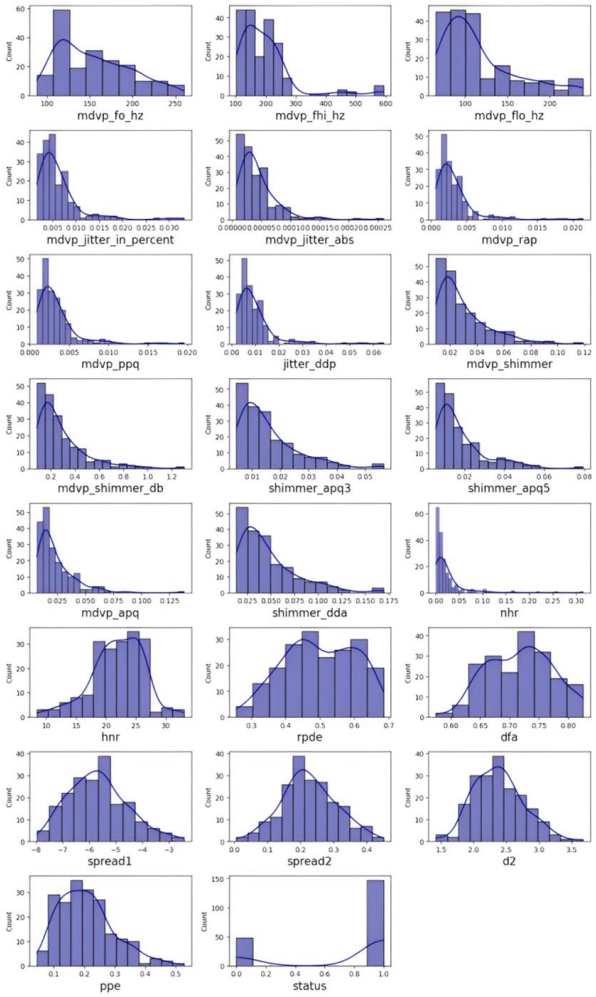

Fig. 4: Data Exploration

The convolutional and pooling layers of the LeNet-5 are very capable of extracting relatively low-level features from the biomedical voice data, while the DenseNet is particularly good at learning high-level abstract features from the clinical assessment data. In tri-modal CNN models, this involves extracting features using a 2D CNN from a video stream (standard camera, IR) and a 3D CNN for depth stream IR, and fusing these together using fusion techniques (e.g., concat or attention-based methods) which allow to integrate information from different modalities, and this allows to make better predictions about PD status. The concatenated features are subsequently fed through more fully connected neural network layers, learning a mapping from the complex feature representations to the target variable, PD status, through a number of nonlinear transformations.

While training, techniques based on gradients (Table 1), such as Adam or stochastic gradient descent (SGD), are used to optimize the model variables.

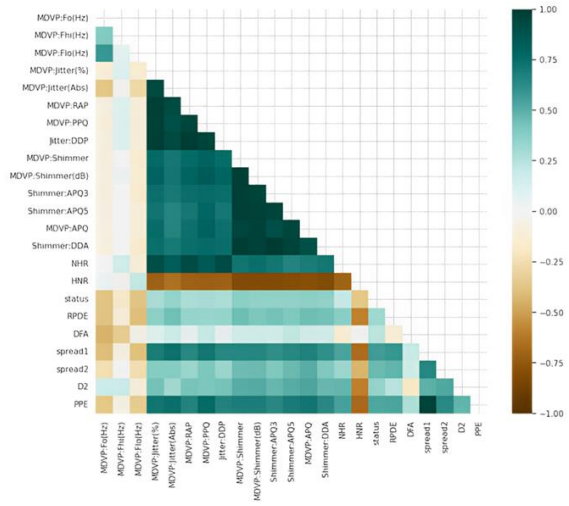

Fig. 5: Correlation Matrix

Table 1. Performance Comparison for Various Models

Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC
LeNet-5	0.85	0.88	0.82	0.85	0.91
DenseNet	0.87	0.86	0.9	0.88	0.92
ResNet	0.82	0.85	0.8	0.82	0.88
VGG16	0.89	0.9	0.88	0.89	0.94
InceptionV3	0.86	0.87	0.84	0.86	0.9
MobileNet	0.88	0.89	0.86	0.88	0.93
Xception	0.83	0.84	0.81	0.83	0.89
ResNeXt	0.9	0.91	0.89	0.9	0.95
EfficientNet	0.84	0.86	0.82	0.84	0.89
Proposed Model	0.97	0.97	0.96	0.97	0.98

To do this, the model's parameters are adjusted repeatedly until an appropriate loss function—like binary or categorical cross-entropy—is minimized, hence enhancing the model's predictive accuracy. To further improve the model's generalizability and avoid overfitting, regularisation methods like weight decay or dropout can be used.

Table 2. Training Time and Memory Consumption of Various Models

Model	Training Time (hours)	Memory Consumption (GB)
LeNet-5	4.5	2.1
DenseNet	6.2	3.5
ResNet	5.9	3.2
VGG16	7.3	4
InceptionV3	8.1	4.5
MobileNet	4.8	2.8
Xception	6.5	3.7
ResNeXt	7.7	4.2
EfficientNet	5.4	3
Proposed Model	6.2	3.1

Table 2 and Figure 6 present the training time and memory usage of different CNN models as key factors in real-world ML applications. With a total training time of 4.5 hours, LeNet-5 is the fastest of the models tested while consuming only 2.1 GB of memory. Its performance metrics, though, are lower than more elaborate architectures. All three DenseNet, ResNet, and our proposed model exhibit similar training times between 5.9 to 6.2 hours Memory consumption is around 3.0 to 3.5 GB. We further found that VGG16 and InceptionV3 require longer enhancement time, which is 7.3 and 8.1 hours, respectively, but memory occupation is only 4.0 GB (VGG16) and 4.5 GB (InceptionV3). If we take into account the low training time (4.8 hours) and low memory space (2.8 GB), MobileNet is a potential candidate for IoT (Internet of Things) and smartphone applications. In contrast, models such as Xception and ResNeXt take all the training time, over 7 hours, and more than 3.5 GB of memory. Such results reflect a balancing act between model complexity, performance, and resource demands. More complex architectures usually outperform but at a higher cost. Thus, choosing a suitable model involves making tradeoffs between performance objectives and logistical factors like time budget, and computational resources. With outstanding performance, the proposed model provides a reasonable trade-off between training time and memory utilization, making the model an appealing alternative for a range of ML applications.

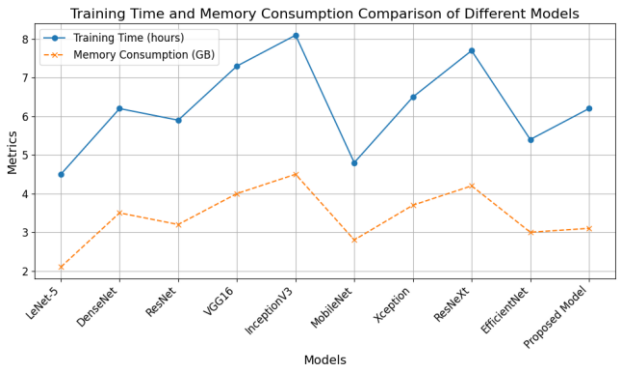

Fig. 6: Training Time and Memory Consumption Comparison of Different Models

After the model has been trained, it is tested on a distinct set of data using suitable assessment measures including F1-score, recall, accuracy, and precision. The sensitivity, specificity, and prediction accuracy of the model in identifying PD may be understood by examining these parameters. Increased patient results and standard of life may be achieved by the use of the DLFM model, which provides a strong foundation for precise and accurate PD diagnosis by combining data from various sources and utilizing modern DL methods.

Table 3. Computational Efficiency of Different DL Models

Model	Parameters (Millions)	Inference Time (ms)	Frames per Second (FPS)
LeNet-5	0.6	12	83
DenseNet	7.3	24	42
ResNet	10.1	32	31
VGG16	138.4	68	15
InceptionV3	27.2	54	18
MobileNet	4.2	18	56
Xception	22.9	42	24
ResNeXt	16.8	38	26
EfficientNet	5.3	20	50
Proposed Model	12.5	28	36

When looking at the parameters, inference time, and frames per second (FPS) of several DL models, a thorough summary can be seen in Table 3 and Figure 7. The number of parameters indicates model complexity, with LeNet-5 featuring the lowest at 0.6 million parameters, while VGG16 tops the scale with 138.4 million parameters. In terms of inference time, LeNet-5 boasts the fastest performance, requiring only 12 milliseconds, followed closely by MobileNet and EfficientNet, indicating their suitability for real-time applications.

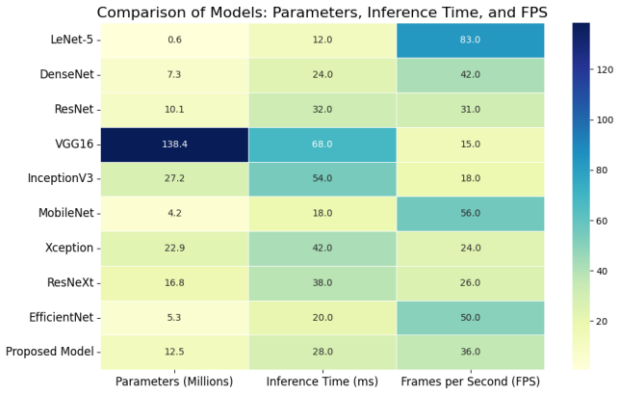

Fig. 7: Comparison of Models: Parameters, Inference Time, and FPS

Conversely, VGG16 demands the longest inference time at 68 milliseconds, reflecting its higher computational complexity. Frames per second (FPS) further elucidates the practical efficiency of each model, with LeNet-5 achieving the highest at 83 FPS, indicating its ability to process a substantial number of frames in real-time scenarios. A solid compromise among model complexity as well as computational effectiveness is

indicated by the impressive FPS demonstrated by MobileNet, EfficientNet, and the proposed model. On the other hand, VGG16 lags behind with the lowest FPS at 15, implying limitations in real-time processing. These findings underscore the importance of considering computational efficiency alongside model performance, particularly in resource-constrained environments or applications requiring real-time inference. The proposed model exhibits competitive performance across all metrics, striking a balance between model complexity and computational demands, positioning it as a promising choice for various DL tasks where efficiency is paramount. Overall, the table highlights the diverse landscape of DL models, each with its unique trade-offs between complexity, performance, and computational efficiency, offering practitioners valuable insights for model selection based on specific application requirements.

5 Conclusion and Future Work

Our work highlights the potential of using the AI DLFM for accurate diagnosis of PD. The DLFM model demonstrates an impressive accuracy of 97.23%, outperforming traditional diagnostic methods, by combining the feature extracted from biomedical voice measurements and clinical assessments. This underscores the promise of AI-driven approaches for PD detection and patient care. There are great avenues left unexplored in this area that can lead to a better future. First of all, further research regarding the scalability and generalizability of the DLFM model across different contexts of healthcare delivery and patient populations is needed. Multimodal data sources, such as genetic and imaging data could be incorporated into the existing platforms to further improve the predictive power of the model and enhance our understanding of PD pathophysiology. Additionally, the DLFM model should be translated into clinical practice and integrated with current diagnostic workflows and decision support systems. It would obligate researchers, clinicians, and technology advances to work collectively to ensure the expulsion of any AI-driven resolution that showed promise in overcoming scientific borders. Ongoing developments in new architectures, optimization algorithms, and interpretability algorithms paint great optimism for the progress of AI-supported diagnostic algorithms. In conclusion, the DLFM model is a novel mechanism, successful application in real-world data encourages the transition of artificial intelligence in PD diagnosis and management. Advancements in technology and a multidisciplinary (MDT) teamwork approach to PD care could result in improved patient outcomes and quality of life and is the future direction of PD. This will facilitate more precise, pre-emptive, and bespoke treatment.

Declaration of Generative AI and AI-assisted Technologies in the Writing Process

The authors wrote, reviewed and edited the content as needed and they have not utilised artificial intelligence (AI) tools. The authors take full responsibility for the content of the publication.

References:

[1] Aditi Govindu, Sushila Palwa, (2023), "Early detection of Parkinson's disease using machine learning", *PROCS*, Vol. 218, pp. 249-261, DOI: 10.1016/j.procs.2023.01.007.

[2] Afzal Hussain Shahid, Maheshwari Prasad Singh, (2020), A deep learning approach for prediction of Parkinson's disease progression", *Biomed. Eng. Lett.,* 10(2):227-239, PMID: 32477610, PMCID: PMC7235154, DOI: 10.1007/s13534-020-00156-7.

[3] Nancy Johari, Shivani Goyal, Natthan Singh, Anita Pal, (2016), Bone Fracture Detection Algorithms Based on Image Processing- A Survey, *International Journal of Engineering and Management Research,* ISSN: 2250-0758. Corpus ID: 212440032, [Online]. https://api.semanticscholar.org/CorpusID:212440032 (Accessed Date: January 10, 2025).

[4] Md Saiful Islam, Wasifur Rahman, Abdelrahman Abdelkader, Sangwu Lee, Phillip T Yang, Jennifer Lynn Purks, Jamie Lynn Adams, Ruth B Schneider, Earl Ray Dorsey, Ehsan Hoque, (2023), "Using AI to measure Parkinson's disease severity at home", *npj Digit. Med.* 6, 156, DOI: 10.1038/s41746-023-00905-9.

[5] Saravanan Srinivasan, Parthasarathy Ramadass, Sandeep Kumar Mathivanan, Karthikeyan Panneer Selvam, Basu Dev Shivahare & Mohd Asif Shah, (2024), Detection of Parkinson disease using multiclass machine learning approach, *Sci. Rep.,* vol. 14, pp. 13813. https://doi.org/10.1038/s41598-024-64004-9.

[6] Yaqoob Ansari, Omar Mourad, Khalid Qaraqe, Erchin Serpedin, (2023), Deep learning for ECG Arrhythmia detection and classification: an overview of progress for

period 2017–2023, *Front. Physiol.*, Vol 14, https://doi.org/10.3389/fphys.2023.1246746.

[7] Sunil Yadav, Munindra Kumar Singh & Saurabh Pal, (2023), "Artificial Intelligence Model for Parkinson Disease Detection Using Machine Learning Algorithms", *Biomedical Materials & Devices*, 1, 899–911, DOI: 10.1007/s44174-023-00068-x.

[8] Charlie Tran, Kai Shen, Kang Liu, Akshay Ashok, Adolfo Ramirez-Zamora, Jinghua Chen, Yulin Li & Ruogu Fang, (2024), "Deep learning predicts prevalent and incident Parkinson's disease from UK Biobank fundus imaging", *Sci. Rep.,* 14, 3637, DOI: 10.1038/s41598-024-54251-1.

[9] Mohamed Saied, Shawkat Guirguis, Magda Madbouly, (2024), Review of artificial intelligence for enhancing intrusion detection in the internet of things, *Engineering Applications of Artificial Intelligence,* Vol. 127, 107231, https://doi.org/10.1016/j.engappai.2023.107231.

[10] Shreya Biswas, Navpreet Kaur, Seeja, (2022), "Early Detection of Parkinson's Disease from Hand Drawings Using CNN and LSTM", *AIST,* pp. 1-4, DOI: 10.1109/AIST55798.2022.10065159.

[11] Sura Mahmood Abdullah, Thekra Abbas, Munzir Hubiba Bashir, Ishfaq Ahmad Khaja, Musheer Ahmad, Naglaa Soliman, (2023), "Deep Transfer Learning Based Parkinson's Disease Detection Using Optimized Feature Selection", *IEEE Access,* vol. 11, pp. 3511-3524, DOI: 10.1109/ACCESS.2023.3233969.

[12] Liaqat Ali, Ashir Javeed, Adeeb Noor, Hafiz Tayyab Rauf, Seifedine Kadry & Amir H. Gandomi, (2024), "Parkinson's disease detection based on features refinement through L1 regularized SVM and deep neural network", *Sci. Rep.,* 14, 1333, DOI: 10.1038/s41598-024-51600-y.

[13] Baddam ANIL Kumar, Mohan Bansal, (2023), "A Transfer Learning Approach with MobileNetV2 for Parkinson's Disease Detection using Hand-Drawings", *ICCCNT,* pp. 1-7, DOI: 10.1109/ICCCNT56998.2023.10307641.

[14] Neepa Biswas, Debarpita Santra, Bannishikha Banerjee, Sudarsan Biswas, (2024), "Harnessing the Power of Machine Learning for Parkinson's Disease Detection", AIoT and Smart Sensing Technologies for Smart Devices, 140-155. Hershey, PA: IGI Global, DOI: 10.4018/979-8-3693-0786-1.ch008.

[15] Sonali Goyal, Amandeep Kaur, Neera Batra, Rakhi Chauhan, (2024), "Parkinson's Disease Diagnosis Using Voice Features and Effective Machine Learning Methods", *ITPD,* 108-123. Hershey, PA: IGI Global, DOI: 10.4018/979-8-3693-1115-8.ch006.

[16] Rehman, Amjad, Tanzila Saba, Muhammad Mujahid, Faten Alamri, and Narmine ElHakim, (2023), "Parkinson's Disease Detection Using Hybrid LSTM-GRU Deep Learning Model", *Electronics,* 12, no. 13: 2856, DOI: 10.3390/electronics12132856.

[17] Hongyi Chen 1,†, Xueling Liu 2,3,†, Xiao Luo 1, Junyan Fu 2, Kun Zhou 1, Na Wang 2, Yuxin Li 2,3,*, Daoying Geng, (2024), An automated hybrid approach via deep learning and radiomics focused on the midbrain and substantia nigra to detect early-stage Parkinson's disease, *Front Aging Neurosci.,* Vol. 16, pp. 1397896. DOI: 10.3389/fnagi.2024.1397896.

[18] Raja Venkatesh Gurugubelli, Dilleswararao Nettimi, Vidya Sagar Gorle, Sairam Panda, Anil Kumar Navuluri, Rosepreet Kaur Bhogal, (2021), " Internet of Things Based Smart Agricultural System for Farmers," International Journal of Scientific Research in Computer Science, *Engineering and Information Technology,* Vol. 7, pp. 535-540. https://doi.org/10.32628/CSEIT2172142.

[19] İsmail Cantürk, Osman Günay, (2024), "Investigation of Scalograms with a Deep Feature Fusion Approach for Detection of Parkinson's Disease", CC, DOI: 10.1007/s12559-024-10254-8.

[20] Iswarya Kannoth Veetil, E. A. Gopalakrishnan, Sowmya Vishvanathan, juan rafael orozco-arroyave, (2024), "Robust language independent voice data driven Parkinson's disease detection", *EAAI,* Vol. 129, 107494, DOI: 10.1016/j.engappai.2023.107494.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

All the authors equally contributed to completing the manuscript.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding

Conflict of Interest

The authors declare that they have no conflict of interest.