

# Integrated Deep Learning System For Early Parkinson Identification From Neuro-Motor And Acoustic Data

Mrs M.Bhuvaneshwari<sup>1</sup>., Mrs P.Rohini<sup>2</sup>., Mrs C.Kowsalya<sup>3</sup>.,

<sup>1</sup>Student, Department of Computer Science and Engineering, Chendhuran College of Engineering and Technology, Pudukkottai.

<sup>2</sup>Assistant Professor&Head, Department of Computer Science and Engineering, , Chendhuran College of Engineering and Technology, Pudukkottai.

<sup>3</sup>Assistant Professor, Department of Computer Science and Engineering, , Chendhuran College of Engineering and Technology, Pudukkottai.

Email id:m.shreegokul@gmail.com<sup>1</sup>, rohini.ccet@gmail.com<sup>2</sup>, queenbbe9722@gmail.com<sup>3</sup>

**ABSTRACT** - Parkinson's disease (PD) is a progressive neurodegenerative disorder that affects motor functions, including handwriting, speech, and coordination, making early detection crucial for effective treatment and disease management. Traditional diagnostic approaches, such as clinical assessments and medical imaging techniques like MRI or PET scans, are often expensive, time-consuming, and dependent on specialized medical expertise. To address these limitations, this project proposes a multimodal deep learning framework that leverages EEG signals, handwritten spiral images, and audio recordings to predict Parkinson's disease accurately. By analyzing diverse datasets, the system captures a comprehensive view of neurological and motor function impairments associated with the disease, providing a more reliable and accessible diagnostic alternative. The proposed system utilizes advanced machine learning architectures tailored to each data modality: EEG data are processed with a hybrid CNN + LSTM model to extract both spatial and temporal features, handwritten images are analyzed using CNN to detect motor irregularities, and audio recordings are classified using XGBoost to recognize vocal biomarkers indicative of Parkinson's. Trained models are stored in a database for seamless retrieval and prediction, enabling the system to process new input data efficiently. Upon detection of Parkinson's disease, the system also provides precautionary measures and health recommendations to aid early intervention, ultimately enhancing patient care and supporting timely medical decision-making in a cost-effective and automated manner.

**Keywords---** Parkinson disease, deep learning, spectrum, speech recognition.

## I. INTRODUCTION

Deep learning which is a hot buzz nowadays and has firmly put down its roots in a vast multitude of industries that are investing in fields like Artificial Intelligence, Big Data and Analytics. For example, Google is using deep learning in

its voice and image recognition algorithms whereas Netflix and Amazon are using it to understand the behavior of their customer. In fact, you won't believe it, but researchers at MIT are trying to predict future using deep learning. Deep learning can be considered as a subset of machine learning. It is a field that is based on learning and improving on its own by examining computer algorithms. While machine learning uses simpler concepts, deep learning works with artificial neural networks, which are designed to imitate how humans think and learn. Neural Network is the biological neurons, which is nothing but a brain cell. Until recently, neural networks were limited by computing power and thus were limited in complexity. Deep learning has aided image classification, language translation, speech recognition. It can be used to solve any pattern recognition problem and without human intervention. Deep learning models are capable enough to focus on the accurate features themselves by requiring a little guidance from the programmer and are very helpful in solving out the problem of dimensionality. Deep learning algorithms are used, especially when we have a huge no of inputs and outputs.

## II. DEEP LEARNING ALGORITHMS

- Feed forward neural network
- Radial basis function neural networks
- Multi-layer Perceptron
- Convolution neural network (CNN)
- Recurrent neural network
- Modular neural network

## III. EXISTING SYSTEM

The existing systems for Parkinson's disease detection primarily rely on clinical assessments, medical imaging, and physician observation. Diagnosis is often based on visible motor symptoms such as tremors, rigidity, and slow movement, which usually appear only in the later stages of the disease. Medical imaging techniques like MRI, PET, or SPECT scans are sometimes used to assess brain activity and structure, but these procedures are highly expensive, time-

consuming, and not easily accessible to patients in rural or resource-limited areas. Additionally, the process requires specialized medical expertise, and variations in clinical interpretation can lead to inconsistencies in diagnosis. These limitations make early detection difficult, delaying treatment and reducing the chances of effective disease management. Existing research and automated approaches have attempted to use single-source data, such as voice recordings or handwriting samples, for early prediction. However, these methods often lack robustness since they rely on one type of biomarker and fail to capture the complex interactions between neurological and motor functions. Moreover, traditional models lack the ability to integrate multiple data types for a holistic analysis, resulting in lower prediction accuracy and limited generalization across different patient conditions. Most existing systems also do not provide precautionary measures or health advice after detection, restricting their practical use in real-world healthcare environments. Therefore, current methods remain inadequate for comprehensive and early Parkinson's disease prediction.

### A. SUPPORT VECTOR MACHINE (SVM)

Support Vector Machine (SVM) is one of the most widely used algorithms in existing Parkinson's disease detection models, especially for analyzing handwriting patterns and voice features. It works by identifying a hyperplane that best separates data points belonging to different classes — in this case, differentiating between Parkinson's and healthy subjects. When applied to biomedical data, SVM effectively handles high-dimensional feature spaces, such as time-series voice data or kinematic signals from handwriting. It uses kernel functions like linear, polynomial, or radial basis function (RBF) to map input data into higher dimensions, where the separation between classes becomes more distinct. Despite its strong classification performance, SVM has several limitations when applied to Parkinson's disease prediction. It performs best with linearly separable data and struggles with noisy or imbalanced datasets, which are common in real-world medical data. Moreover, SVM cannot easily capture temporal dependencies or non-linear relationships present in EEG or audio signals. It also lacks scalability when dealing with large datasets or multimodal data combinations. Thus, while SVM provides a baseline for classification, it is insufficient for comprehensive, multi-feature Parkinson's analysis.

### B. K-NEAREST NEIGHBORS (KNN)

K-Nearest Neighbors (KNN) is another commonly used algorithm in Parkinson's disease research, particularly for voice-based diagnosis. The algorithm classifies new data points based on the majority label of their nearest neighbors in the feature space. For example, if a voice sample shares acoustic similarities (like jitter or shimmer) with samples labeled as Parkinson's, it will likely be classified as such. KNN is easy to implement, non-parametric, and does not make assumptions about the underlying data distribution, making it suitable for smaller medical datasets. However, KNN's simplicity also introduces challenges. Its performance heavily depends on the choice of distance metrics (such as

Euclidean or Manhattan distance) and the value of  $k$ , which can affect classification accuracy. Furthermore, KNN suffers from high computational cost during prediction, as it must compare a new sample with every instance in the training dataset. It is also sensitive to noise and irrelevant features, which can mislead the model, especially in complex biomedical data. Therefore, while KNN provides a basic approach to pattern recognition, it lacks the advanced feature extraction capabilities needed for multimodal Parkinson's disease detection.

### C. RANDOM FOREST (RF)

Random Forest (RF) has been utilized in some existing Parkinson's detection studies for analyzing structured datasets like audio or sensor readings. It operates by constructing multiple decision trees during training and combining their outputs to form a final prediction. Each tree is trained on a random subset of the dataset and features, which helps improve generalization and reduce overfitting. In Parkinson's diagnosis, RF models are used to identify important biomarkers such as tremor frequency, vocal stability, or pen pressure variations in handwriting samples. Although Random Forest is more robust than single classifiers, it still has limitations when handling image and temporal data such as EEG signals. It cannot extract spatial or temporal dependencies, which are critical for detecting subtle neurological changes. Moreover, as the number of trees increases, the model becomes computationally expensive and difficult to interpret. While RF improves prediction stability compared to SVM or KNN, it remains a shallow learning technique incapable of capturing complex relationships between multimodal datasets.

### D. DISADVANTAGES

- Requires expert medical supervision and manual interpretation for diagnosis.
- High cost due to reliance on medical imaging and specialized equipment.
- Uses single-type data (voice or handwriting), leading to lower accuracy.
- Cannot detect early-stage symptoms effectively due to limited feature analysis.
- Lacks automation and preventive guidance, offering no real-time or precautionary support.

## IV. PROPOSED SYSTEM

### A. Hybrid CNN + LSTM (for EEG data)

The hybrid CNN + LSTM architecture combines convolutional neural networks' strength in spatial feature extraction with LSTM's ability to model temporal dependencies, making it well-suited for EEG signals that exhibit both spatial patterns across channels and temporal dynamics over time. In practice, the CNN layers act as trainable feature extractors that learn localized waveform and

spectral patterns (e.g., rhythms, channel correlations) from short windows or spectrogram representations of EEG, producing compact high-level feature maps. These feature maps are then fed into LSTM layers which capture the sequence-level evolution of those features, enabling the model to recognize temporal signatures and progression patterns that are characteristic of Parkinson's-related neural activity. During training, this hybrid model benefits from end-to-end learning: convolutional filters are tuned to emphasize the most informative spatial-spectral cues while LSTM units learn temporal dependencies and long-range context, which improves sensitivity to subtle, time-varying biomarkers. Practical considerations include proper windowing and overlap for EEG segments, normalization, and regularization (dropout, batch norm) to avoid overfitting, plus careful choice of sequence length so LSTM can capture meaningful patterns without exploding computational cost. While powerful, this hybrid approach requires sufficient labeled EEG data and more compute than shallow models, and its interpretability can be improved by attention mechanisms or saliency mapping to highlight which channels/times drive predictions.

### B. Convolutional Neural Network (CNN) (for handwritten images)

Convolutional Neural Networks are the go-to architecture for image analysis and are ideal for detecting motor-control abnormalities in handwritten spiral images or other pen-stroke samples. CNNs apply hierarchical convolutional filters that progressively learn low-level (edges, strokes) to high-level (shape irregularities, tremor patterns, pressure-induced texture) features from raw images; pooling and deeper layers provide translational invariance and compact representations helpful for classification. For handwriting analysis, preprocessing steps such as binarization, stroke-width normalization, resizing, and augmentation (rotation, scaling, elastic distortions) improve robustness; the learned CNN embeddings can distinguish subtle differences in curvature, smoothness, and continuity that correlate with Parkinsonian motor deficits. Training a CNN for this task typically involves transfer learning from large vision models or designing a compact architecture if dataset size is limited; metrics such as accuracy, precision, recall, and class-specific ROC curves guide model selection and hyperparameter tuning. CNNs are fast at inference and straightforward to deploy, but they may require explainability aids (visual saliency maps, Grad-CAM) to show clinicians which image regions influenced the decision. Care must also be taken to avoid bias from handwriting style variability (age, education, handedness), which can be mitigated by diverse training samples and augmentation.

### C. XGBoost (for audio/feature CSV data)

XGBoost is a high-performance gradient boosting decision-tree algorithm particularly effective on tabular feature sets an excellent choice when audio recordings are preprocessed into engineered features (e.g., MFCCs, jitter, shimmer, pitch statistics, harmonic-to-noise ratio) stored as CSVs. XGBoost builds an ensemble of decision trees in a stage-wise fashion, optimizing a differentiable loss function

while using regularization, shrinkage, and column/row subsampling to prevent overfitting; it handles heterogeneous feature scales, missing values, and non-linear interactions between acoustic biomarkers efficiently. For Parkinson's voice analysis, carefully engineered features capturing prosody, phonation stability, and spectral characteristics provide XGBoost with rich information to learn discriminative splits that separate affected and healthy subjects. Model tuning (learning rate, max depth, number of estimators, regularization terms) and feature selection or importance analysis are central to achieving robust performance; XGBoost also provides feature importance scores which aid interpretability by identifying the most influential vocal biomarkers. While XGBoost is less suited to raw time-series or waveform inputs than deep nets, it is computationally efficient, often performs well on moderate-sized datasets, and integrates easily into pipelines where audio is converted to a structured feature representation making it a pragmatic choice when labeled audio examples are limited or when explainability of features is desired.

## D. SYSTEM ARCHITECTURE

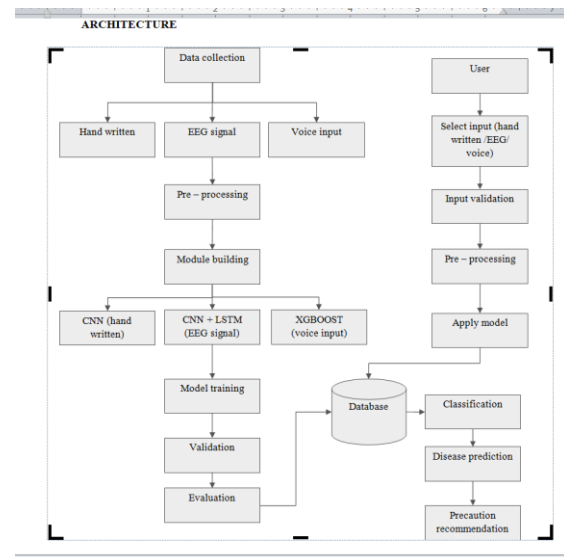


Fig 1. System architecture

## V. MODULES DESCRIPTION

### A. DATASET ACQUISITION

The dataset acquisition module is responsible for collecting and organizing the multimodal data used to train and evaluate the system. It includes three distinct datasets: handwritten image data, EEG signal data, and voice data. The handwritten dataset consists of spiral and wave drawings that reflect fine motor control impairments typical in Parkinson's patients. The EEG dataset contains electrical brain activity signals that reveal neurological irregularities associated with Parkinson's disease. The voice dataset includes recordings of speech patterns, capturing tremor and tone variations that indicate vocal impairment. All datasets are obtained from reliable sources or collected under controlled conditions. The acquired data are stored in a structured format (CSV or

image) for further processing. Proper labeling ensures supervised training of the models. This module serves as the foundation for subsequent preprocessing and model development.

### B.PREPROCESSING

The preprocessing module prepares the raw datasets for deep learning by cleaning, normalizing, and transforming the data into model-compatible formats. For handwritten images, preprocessing involves resizing, grayscale conversion, and noise removal to highlight essential stroke patterns. For EEG data, filtering techniques are used to remove artifacts such as muscle or eye movement noise, and the data are normalized to maintain uniform scale. Voice datasets undergo noise reduction, feature extraction (such as MFCC or pitch features), and segmentation for consistent input length. This module ensures that all datasets are standardized and free from unwanted variations. The primary goal is to enhance feature quality and improve model learning efficiency. After preprocessing, the clean and refined data are split into training and testing sets to ensure accurate model validation.

### C.MODEL BUILD – CNN MODEL

The Convolutional Neural Network (CNN) model is used for the handwritten image dataset to detect motor pattern irregularities linked to Parkinson’s disease. The CNN architecture automatically extracts hierarchical features such as edges, curves, and spatial relationships from the spiral or wave drawings. Multiple convolutional and pooling layers are employed to capture essential visual features, followed by fully connected layers for classification. The model effectively learns subtle variations in handwriting pressure, smoothness, and tremor intensity. The CNN’s robustness against noise and distortion makes it ideal for analyzing such medical images. Training is conducted on labeled handwritten datasets, and model optimization is achieved using activation functions and dropout layers to prevent overfitting. The output is a trained CNN model capable of distinguishing between Parkinson’s and healthy handwriting patterns.

### D.MODEL BUILD – CNN WITH LSTM MODEL

The hybrid CNN + LSTM model is designed for the EEG dataset, integrating spatial and temporal feature learning for enhanced accuracy. The CNN layers extract spatial features from EEG signal images, while the LSTM layers capture sequential dependencies and temporal dynamics of brain activity. This combination allows the model to analyze both localized signal characteristics and time-based fluctuations indicative of Parkinson’s disease. The architecture is trained on preprocessed EEG data, ensuring high adaptability to noise and complex signal variations. The hybrid model effectively recognizes neurological abnormalities and pattern shifts over time. During training, parameters such as learning rate, number of epochs, and hidden units are tuned for optimal performance. The resulting model can accurately predict Parkinson’s disease based on

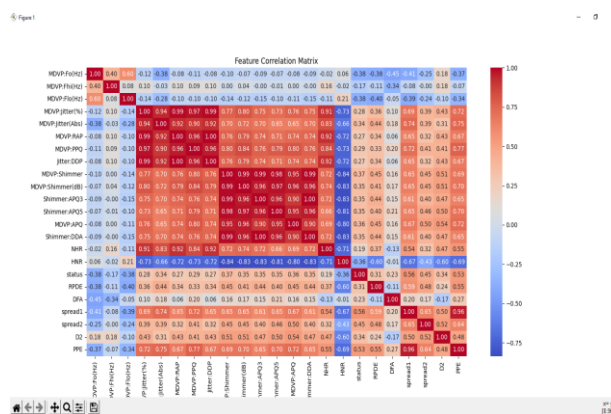
EEG signals, outperforming traditional single-model approaches.

### E.MODEL BUILD – XGBOOST ALGORITHM

The XGBoost (Extreme Gradient Boosting) algorithm is applied to the voice dataset for Parkinson’s disease classification based on vocal impairments. XGBoost is a powerful ensemble learning technique that uses gradient boosting for better accuracy and generalization. The pre-processed voice data are converted into numerical feature vectors, capturing characteristics like pitch, jitter, shimmer, and MFCC coefficients. The algorithm constructs multiple weak learners (decision trees) and combines them to form a strong classifier capable of distinguishing Parkinson’s-affected voices from normal ones. Regularization parameters help avoid overfitting while maintaining model efficiency. The model is trained and validated using labeled voice data to ensure reliability. Once trained, XGBoost provides fast and accurate predictions, making it suitable for real-time analysis and integration with the multimodal system.

### F.INPUT DATA CLASSIFICATION

The input data classification module serves as the final stage of the system, where new user inputs are analyzed to determine Parkinson’s disease presence. Depending on the type of data provided EEG, handwritten image, or audio the corresponding trained model (CNN + LSTM, CNN, or XGBoost) is automatically selected from the database. The input data undergo basic preprocessing to match the training format, ensuring consistency. The chosen model processes the input and outputs a classification label indicating whether Parkinson’s disease is detected. If the prediction is positive, the system generates precautionary advice and medical recommendations to assist the user. This module ensures seamless integration of multiple modalities under a unified prediction framework. It supports real-time diagnosis, enhances accuracy, and provides meaningful feedback for early intervention and disease management.



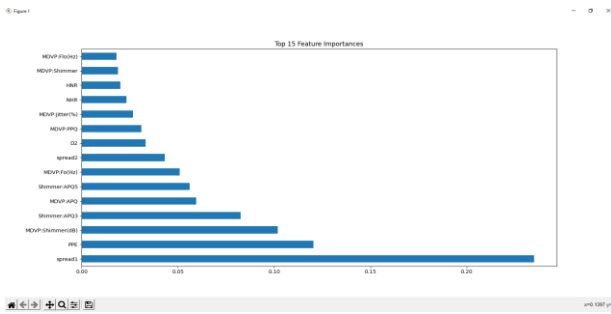


Fig 3. Tracking sheet

## VII. REFERENCES

- [1] Govindu, Aditi, and Sushila Palwe. "Early detection of Parkinson's disease using machine learning." *Procedia Computer Science* 218 (2023): 249-261.
- [2] Templeton, John Michael, Christian Poellabauer, and Sandra Schneider. "Classification of Parkinson's disease and its stages using machine learning." *Scientific reports* 12.1 (2022): 14036.
- [3] Rana, Arti, et al. "Imperative role of machine learning algorithm for detection of Parkinson's disease: review, challenges and recommendations." *Diagnostics* 12.8 (2022): 2003.
- [4] Majhi, Babita, et al. "An improved method for diagnosis of Parkinson's disease using deep learning models enhanced with metaheuristic algorithm." *BMC medical imaging* 24.1 (2024): 156.
- [5] Giannakopoulou, Konstantina-Maria, Ioanna Roussaki, and Konstantinos Demestichas. "Internet of things technologies and machine learning methods for Parkinson's disease diagnosis, monitoring and management: a systematic review." *Sensors* 22.5 (2022): 1799.
- [6] Quan, Changqin, Kang Ren, and Zhiwei Luo. "A deep learning based method for Parkinson's disease detection using dynamic features of speech." *IEEE access* 9 (2021): 10239-10252.
- [7] Trabassi, Dante, et al. "Machine learning approach to support the detection of Parkinson's disease in IMU-based gait analysis." *Sensors* 22.10 (2022): 3700.
- [8] Rehman, Amjad, et al. "Parkinson's disease detection using hybrid LSTM-GRU deep learning model." *Electronics* 12.13 (2023): 2856.
- [9] Shaban, Mohamed. "Deep learning for Parkinson's disease diagnosis: a short survey." *Computers* 12.3 (2023): 58.
- [10] Aljalal, Majid, et al. "Parkinson's disease detection from resting-state EEG signals using common spatial pattern, entropy, and machine learning techniques." *Diagnostics* 12.5 (2022): 1033.

## VI. CONCLUSION AND FUTURE ENHANCEMENT

### A. CONCLUSION

The proposed system for Parkinson's disease prediction using handwritten, EEG, and audio datasets demonstrates the effectiveness of integrating multimodal data with advanced deep learning and machine learning techniques. By combining CNN, LSTM, and XGBoost algorithms, the system can capture both spatial and temporal features across different data types, ensuring high prediction accuracy and reliability. The model efficiently identifies subtle motor, neurological, and vocal impairments that are early indicators of Parkinson's disease. This approach eliminates the dependence on costly medical imaging and expert-driven assessments, offering a more accessible, automated, and data-driven solution for early diagnosis. The integration of hybrid deep learning models and ensemble learning enhances system robustness and generalization across diverse input modalities. Furthermore, the system not only predicts the presence of Parkinson's disease but also provides precautionary measures to support patient care and early intervention. This makes it a valuable tool for healthcare practitioners and patients, especially in remote or low-resource settings. The storage of trained models in a centralized database enables scalability and real-time prediction capabilities. Overall, the project establishes a solid foundation for intelligent disease detection systems that leverage multimodal datasets and hybrid deep learning architectures, contributing significantly to the advancement of medical diagnosis and personalized healthcare technology.

### B. FUTURE ENHANCEMENT

In the future, this system can be enhanced by integrating real-time data acquisition through wearable sensors and mobile applications to continuously monitor neurological and motor activities. Incorporating advanced deep learning architectures such as Transformers and attention-based models could further improve multimodal feature fusion and prediction accuracy. The system can also be extended to include additional biomarkers, such as gait patterns and facial expressions, for a more comprehensive diagnosis. Moreover, implementing a cloud-based platform would allow remote access for patients and doctors, enabling personalized monitoring and continuous model updates.