Parkinson's Disease Prediction Using Deep Learning Technique

Prof .Muhammed Suhair K
Assistant professor
Computer science and engineering
Mea engineering college Perinthalmanna, India
suhair@meaec.edu.in

Muhammed Suhail T
Computer science and engineering
Mea engineering college
Perinthalmanna, India
19mcs04@meaec.edu.in

Raifa TK
Computer science and engineering
Mea engineering college
Perinthalmanna, India
19ncs11@meaec.edu.in

Shadiya Sherin
Computer science and engineering
Mea engineering college
Perinthalmanna, India
19gcs08@meaec.edu.in

Shahana Sherin ME
Computer science and engineering
Mea engineering college
Perinthalmanna, India
19mcs40@meaec.edu.in

***Abstract*—This project's goal is to forecast Parkinson's disease patients using voice dataset and spiral sketching inputs. A Deep Neural Network (DNN) algorithm is used in the predictive analytics framework to analyze the data and make predictions. Voice data is obtained from Kaggle, and spiral drawing inputs include pressure, grip angle, and timestamp, which are analyzed and extracted using the DNN algorithm. The extracted values are then compared to a trained database using deep learning techniques, resulting in accurate predictions. The integration of both voice and spiral drawing data provides a comprehensive analysis and the use of DNN and deep learning techniques allows for quick and accurate results. It is anticipated that the outcomes of this suggested approach will enhance clinical surveillance of older patients and raise their chances of living longer.**

***Keywords* —Parkinson's disease, predictive analytics, Deep Neural Network (DNN) algorithm, voice data, spiral drawing inputs, pressure, grip angle, timestamp, deep learning techniques, comprehensive analysis, clinical monitoring .**

I. Introduction

Parkinson's disease is a chronic condition that affects both the neurological system and the bodily components that are under the control of the nervous system. Symptoms emerge gradually. The initial sign could be a slight tremor in just one hand. Although tremors are typical, the disease might also make you stiff or move more slowly. A few cases of Parkinson's disease can be linked to particular genetic mutations, and some cases of the condition appear to be hereditary. Although Parkinson's disease is known to have a genetic component, the condition rarely runs in families. Many scientists now think that a mix of genetic and environmental factors, including exposure to chemicals, causes Parkinson's disease. Early detection of disease would help to control the symptoms in patients. This method uses both voice and spiral drawing datasets for the prediction of disease.Both are collected from the Kaggle.And these datas are This is the system which will predict the disease easily and efficiently.

II. Related Works

A. The Parkinson's Disease detection using machine learning techniques

In the experiment described here, a dataset comprising of both healthy individuals and patients with the condition is designed to estimate the presence of Parkinson's disease using a collection of 24 symptoms. The dataset is divided into two sets to train and test the machine learning algorithms: a training set and a testing set, each of which accounts for 60% of the total. The status column of the dataset indicates whether a patient has Parkinson's disease (represented by "1") or not (represented by "0") [1].

B. Intelligent Parkinson disease prediction using machine learning algorithms

The dataset used in the experiment was downloaded from the UCI Machine Learning repository and contains a voice that has been damaged by Parkinson's disease. In addition to 22 voice measurement-related attributes, the dataset also includes a "status" feature that shows whether or not the patient has Parkinson's disease. The biggest range is shown in the parallel coordinates of the data. Support Vector Machines (SVM) had the best accuracy of 88.9 percent, Random Forest was second with 90.26 percent, and Naive Bayes algorithm had the lowest accuracy of 69.23 percent, according to an analysis of various machine learning algorithms. Clustering algorithms like Hierarchal clustering and Self-Organizing Maps revealed differences in the number of clusters between healthy and unwell patients (SOM) [2].

C. Machine learning for the diagnosis of Parkinson's disease

The study includes studies on the use of machine learning for diagnosis of Parkinson's disease . The only studies that accessible through IEEE Xplore and PubMed were those that were published up until February 14, 2020. A total of 209 research were examined, with information obtained regarding their goals, data sources, types, and machine learning strategies, as well as their findings. The review concluded that machine learning and new biomarkers have great potential to improve the diagnosis of PD and result in a more informed and systematic approach to clinical decision making [3].

D. Machine Learning Technique Based Parkinson's Disease Detection From Spiral And Voice Inputs

The study provided instructions for a technique that uses spiral drawings and speech records to identify Parkinson's illness. it said RStudio is used to analyse the dataset, and machine learning techniques like k-means clustering and decision trees are used. a spiral using Python, it is possible to analyse drawings and extract features using the PCA technique. Using the SVM machine learning technique, the retrieved characteristics are then contrasted with a trained database. With the help of clinical surveillance, the study hopes to identify diseases early and thereby enhance the quality of life for seniors [4].

E. Diagnosis of Parkinson's disease using machine learning and datamining systems from voice dataset

This study uses UCI repository to obtain the Parkinson's disease dataset. The study tested the efficiency and value of several classification algorithms on the dataset. The statistical classification of Parkinson patients and healthy participants using Sieve Multigram and Survey Graph on the speech data allowed for accuracy. According to the study, KStar and NNge are two accurate categorization techniques. The nodes Fhi, Flo, Jitter, JitterAbs, RAP, and PPQ were connected in the sieve multigram, while Shimmer and ShimmerDB were linked to KStar and NNge. While SimpleCART had 13 and 7 leaves, the ADTree had 21 and 31 leaves. DBScan and SimpleKMeans both discovered that the majority of the clusters were different, with 25 and 38 percent of the clusters, respectively, indicating Parkinson's illness [5].

F. Machine Learning techniques based Parkinson's disease disease detection from spiral and voice input

The method described in the study combines voice and spiral drawing recordings to detect Parkinson's illness. The analysis of the audio data uses RStudio and machine learning techniques like decision trees and k-means clustering. The spiral drawings are analyzed with Python and features are extracted using the PCA algorithm. The extracted values are compared to a trained database using the SVM machine learning technique. The goal of the system is to provide early disease detection and support clinical monitoring of elderly patients to improve their quality of life [6].

G. The personalized Parkinson's project-examining disease progression through broad biomarkers in early Parkinson's disease

The Personalized Parkinson Project aims to create biomarkers for Parkinson's disease through long-term measurement of multiple biomarkers. The project intends to link established and new biomarkers to disease progression and treatment response through hypothesis-driven analysis. The ultimate objective is to provide a dataset that can be used widely to identify new Parkinson's disease biomarkers and treatment targets [7].

H. Parkinson's Disease-Etiopathogenesis and treatment

The study focuses on how different clinical subtypes, genetic variations, and environmental factors have changed how people view Parkinson's disease as a singular entity. The study underlines that non-motor symptoms, including as insomnia, anosmia, constipation, and depression, start out as early signs of the disease and proceed along with cognitive decline and dysautonomia. The paper also describes the primary molecular causes of PD, including misfolded -synuclein, oxidative stress, poor protein clearance, mitochondrial dysfunction, and neuroinflammation. According to the study, PD involves a number of neurotransmitter pathways, including dopaminergic, noradrenergic, glutamatergic, serotonergic, and adenosine, which together contribute to the disease's wide range of clinical symptoms [8].

I. Eye Tracking and Machine Learning significance in Parkinson's disease symptoms prediction

This study used eye movements in conjunction with standard neurological and cognitive examinations to forecast the progression of Parkinson's disease. Using the Scikit Learn and Pandas libraries, the study's machine learning models were created in Python. The study's prediction accuracy for patients who were in advanced stages and responded to treatment was 75%, whereas it was 70% for all classes combined. The research suggests that eye movements could be a mmeasurable indicator for assessing Parkinson's disease symptoms [9].

J. Analysis of vertical eyemovements in Parkinson's disease and its potential for diagnosis

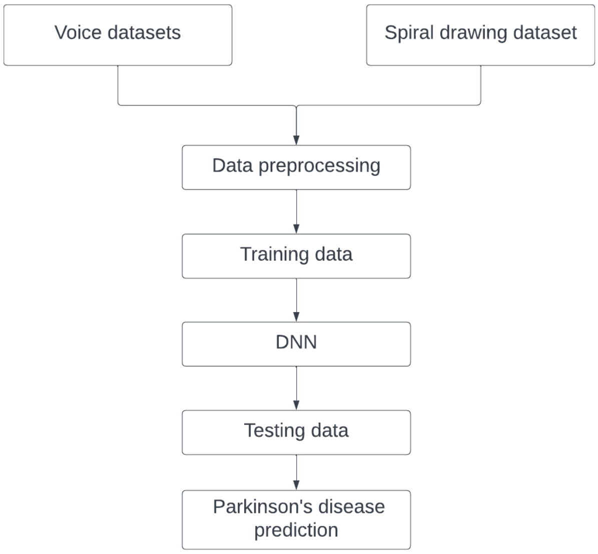
This study used vertical electrooculography (VEOG) to measure eye movements when subjects were at rest, both with their eyes open and closed. Features from VEOG time series in the time, frequency, and time-frequency domains were used to conduct the study. The most crucial features for a support vector machine classifier that corrects output codes were chosen using p-values. The results showed that the VEOG data had a discriminating accuracy for OFF-medication conditions of 69.10% and for ON-medication states of 87.27%. The study discovered that levodopa prescription impacts vertical eye movement amplitude variation but not movement rates (frequency content) [10].

III. METHODOLOGY

The proposed methodology for Parkinson's disease prdiction uses datasets of voice and spiral drawings of both healthy and diseased persons. Both datasets were collected from Kaggle. MFCC feature extraction technique is used to filter speech dataset to avoid noise and clean the data. Then trained the system by providing 70% of preprocessed data. About 30%

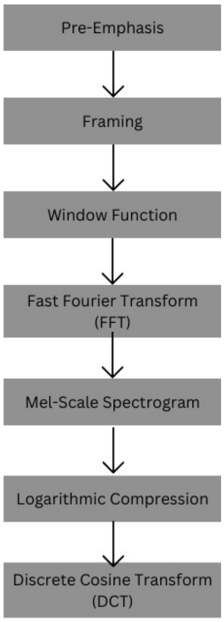
of data is taken for tasting. By using the trained data a DNN model created. This model can predict the Parkinson disease using speech data clearly. This model can recognize whether the given voice is of a Parkinson's patient's or not. The spiral drawings of Parkinson's patients and healthy persons are analyzed and those images are processed through different layers of CNN in order to extract and understand pattern for classification. Transfer learning method approach is made to create the model.So Pretraind model VGG16 is used since it has more accuracy.This neural network is trained with the spiral drawing dataset and the model compare with testing dataset. When an image is uploaded the model will predict whether the person has Parkinson or not. The above two models are connected and an accurate model of Parkinson's disease prediction system is created.The whole system can clearly predict whether a person has Parkinson's or not by using both voice and spiral drawing of that particular person.

A. System Architecture



1) Voice Dataset : It will require a sizable and varied dataset of speech recordings from healthy people as well as people with Parkinson's disease in order to create a deep learning model for voice recognition-based Parkinson's disease prediction. The dataset should include audio samples that cover a range of speaking styles and emotions, as well as demographic information about the speakers. It is important to ensure that the dataset is well-curated and balanced, with an equal number of samples from individuals with and without Parkinson's disease . Our Voice Dataset collected from the Kaggle used as primary input. Then the dataset is divided for training and testing data.

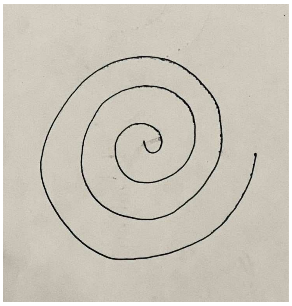
2) MFCC Feature Extraction: MFCC (Mel-Frequency Cepstral Coefficient) is a commonly used feature extraction technique in speech processing and audio analysis. The MFCCs are a set of coefficients that represent the spectral envelopeope of a speech signal in a compact form. MFCC feature extraction from a voice dataset, follow these steps:



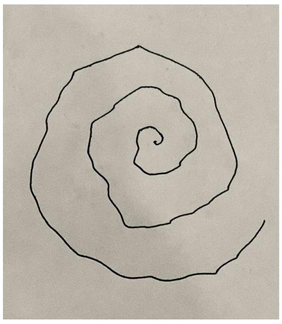
- Pre-Emphasis: The input audio signal is filtered using a high-pass filter to emphasize high-frequency components.
- Framing: The input audio signal is divided into overlapping or non-overlapping frames to create a sequence of short-time windows.
- Window Function:Each frame is given a window function (like the Hamming window), which reduces spectral leakage.
- Fast Fourier Transform (FFT): For each frame, the Fourier Transform is employed to convert the time-domain data into the frequency domain.
- Mel-Scale Spectrogram: The magnitude of the FFT coefficients is transformed into the Mel scale, which approximates the human auditory system's non-linear frequency resolution.
- Logarithmic Compression: The Mel-scale spectrogram is transformed into the logarithmic domain to reduce the dynamic range.
- Discrete Cosine Transform (DCT): As a last step, the logarithmic Mel-scale spectrogram is subjected to the DCT to produce the MFCC coefficients, which serve as a condensed representation of the spectral envelope of the audio source.

Once the MFCC features have been extracted, they can be used as input to a machine learning model for speech recognition, speaker identification, or other speech-related tasks. The MFCCs are a compact representation of the spectral information in the speech signal, and they capture the important spectral variations that are relevant for speech analysis.

3) Spiral drawing Dataset: A dataset of spiral drawings is a collection of spiral drawings that have been photographed or recorded. This type of dataset is frequently used for creating and testing motion analysis and pattern recognition algorithms in the domains of computer vision and machine learning. A spiral drawing dataset may include different variations of spirals, such as clockwise or counter-clockwise spirals, spirals with different curvatures and sizes, and spirals drawn with different tools (e.g., pencil, pen, stylus). The dataset may also include demographic information about the participants, such as age, gender, and handedness. The images or recordings in the dataset can be used to train machine learning models to recognize and classify spirals based on various attributes, such as the direction of the spiral, the size of the spiral, and the speed of the drawing. The models can then be used to analyze new spiral drawings and make predictions about the attributes of the spirals. It is important to ensure that the spiral drawing dataset is diverse, balanced, and representative of the population of interest. This will help to ensure that the machine learning models developed using the dataset are accurate and generalizable to new spiral drawings.



Spiral Drawing of healthy Person



Spiral Drawing of parkinson's patient

4) Data Preprocessing: Data preprocessing is the step in the machine learning pipeline where raw data is cleaned, transformed, and organized into a format that can be used to build a predictive model.Preparing the data so that it may be utilised to create a model that will generate reliable predictions is the aim of data preprocessing.

5) Train Data: The input data used to train a deep learning model is referred to as "training data" in this context.To minimise the discrepancy between the model's predictions and the actual target values in the training examples, the optimum set of parameters must be identified. The model's parameters are typically updated iteratively using labelled training data and an optimization approach like gradient descent. A deep learning model's effectiveness is significantly influenced by the number and quality of the training data. It is important to have a large and diverse dataset that is representative of the problem you're trying to solve, so that the model can generalize well to new data.

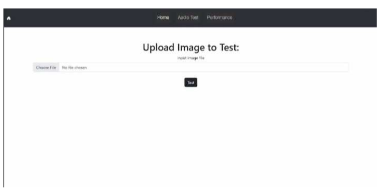
6) DNN: Deep Neural Network, also known as DNN, is a kind of artificial neural network created to include numerous hidden layers. The network can learn and simulate complicated interactions between inputs and outputs thanks to the architecture, which makes it suitable for a variety of tasks like image categorization, audio recognition, and natural language processing. For voice data, the DNN could be trained to recognize and categorize different types of speech sounds, such as vowels and consonants, or to perform speaker recognition. The inputs used here is audio features extracted from the raw sound wave, such as mel-frequency cepstral coefficients (MFCCs). For spiral data, the DNN could be trained to recognize and classify different types of spirals, such as clockwise or counterclockwise, or to perform image classification. The inputs used to the network is image features extracted from the raw image data, such as intensity values or gradient information. Both types of DNNs are trained using a labeled dataset, where the network is provided with input-output pairs to learn from.During the training phase, the network's weights and biases would be adjusted to minimise the gap between predictions and actual labels.

7) Test data: Testing is an important step in the development of a DNN model as it provides a measure of the model's ability to generalize to new data and make accurate predictions. The results of the testing process are used to make decisions about how to improve the model or choose between different models.

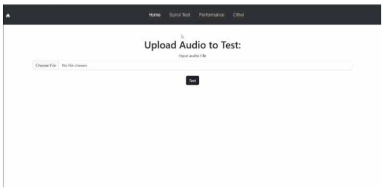
8) Parkinson's disease prediction: The prediction stage of the model is now underway for Parkinson's disease. The technology will determine if a person has Parkinson's disease when new data becomes available.

IV. EXPERIMENTAL RESULTS

The system home page will lead to two separate web pages, one is for uploading audio file and other is for uploading the spiral image.



Web page to upload spiral image



Web page to upload audio file

V. CONCLUSION

The proposed method for predicting Parkinson's disease uses a combination of voice and spiral drawing data, analyzed through a Deep Neural Network (DNN) algorithm. The system is expected to show improved early disease identification and increased chances of a longer lifespan for elderly patients. The use of deep learning techniques results in quick and accurate predictions, providing a comprehensive analysis through the integration of both voice and spiral drawing inputs.

REFERENCES

[1] Mathur, R., Pathak, V., Bandil, D. (2019). Parkinson Disease Prediction Using Machine Learning Algorithm. In: Rathore, V., Worring, M., Mishra, D., Joshi, A., Maheshwari, S. (eds) Emerging Trends in Expert Applications and Security. Advances in Intelligent Systems and Computing, vol 841. Springer, Singapore.

[2] Sriram TV, Rao MV, Narayana GS, Kaladhar DS, Vital TP. Intelligent Parkinson disease prediction using machine learning algorithms. Int. J. Eng. Innov. Technol. 2013 Sep;3:212-5.

[3] Mei J, Desrosiers C, Frasnelli J. Machine learning for the diagnosis of Parkinson's disease: a review of literature. Frontiers in aging neuroscience. 2021 May 6;13:633752.

[4] Jaichandran R; Leelavathy S; Usha Kiruthika S; Goutham Krishna; Melvin John Mathew; Jomon Baiju. "MACHINE LEARNING TECHNIQUE BASED PARKINSON'S DISEASE DETECTION FROM SPIRAL AND VOICE INPUTS". European Journal of Molecular Clinical Medicine, 7, 4, 2020, 2815-2820.

[5] Sriram, T.V.S., Rao, M.V., Narayana, G.V.S., Kaladhar, D.S.V.G.K. (2015). Diagnosis of Parkinson Disease Using Machine Learning and Data Mining Systems from Voice Dataset. In: Satapathy, S., Biswal, B., Udgata, S., Mandal, J. (eds) Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2014. Advances in Intelligent Systems and Computing, vol 327. Springer, Cham.

[6] Mathew, Mevin John, and Jomon Baiju. "Machine learning technique based parkinson's disease detection from spiral and voice inputs." Eur. J. Mol. Clin. Med 7 (2020): 2815-2819.

[7] Bloem BR, Marks WJ, Silva de Lima AL, Kuijf ML, Van Laar T, Jacobs BP, Verbeek MM, Helmich RC, Van De Warrenburg BP, Evers LJ, Inthout J. The Personalized Parkinson Project: examining disease progression through broad biomarkers in early Parkinson's disease. BMC neurology. 2019 Dec;19(1):1-0.

[8] Jankovic J, Tan EK. Parkinson's disease: Etiopathogenesis and treatment. Journal of Neurology, Neurosurgery Psychiatry. 2020 Aug 1;91(8):795-808.

[9] Chudzik A, Szymański A, Nowacki JP, Przybyszewski AW. Eye-Tracking and Machine Learning Significance in Parkinson's Disease Symptoms Prediction. InAsian Conference on Intelligent Information and Database Systems 2020 Mar 23 (pp. 537-547). Springer, Cham.

[10] Farashi S. Analysis of vertical eye movements in Parkinson's disease and its potential for diagnosis. Applied Intelligence. 2021 Nov;51(11):8260-70.