Parkinson's Disease Detection using Machine Learning Techniques

Radha N, Sachin Madhavan RM, Sameera holy S

Abstract
Most Parkinson Disease (PD) patients suffer from vocal cord disorders. Speech impairment is an early indicator of PD. This paper focuses on development of Parkinson's disease detection system using acoustic features such as chroma STFT, RMS, spectral centroid and bandwidth, MFCC, Roll-off and Zero crossing rate which is derived from speaker's sound unit. The extracted features are modelled by various machine learning techniques. In this paper, a classification method based on Convolutional Neural Network (CNN), Artificial Neural Network (ANN) and Hidden Markov Model (HMM) are used to distinguish the PD patient's samples from healthy people. The CNN network is trained using spectrogram of speaker information and with the acoustic features. ANN and HMM models are trained only with acoustic features. The proposed method is tested on the data obtained by the voice recordings of spontaneous dialogues and passage from participants (both PD patients and healthy controls). The recognition accuracy of CNN with spectrogram is 88% and acoustic features gives 93.5%. ANN classifier enhances the classification ability of voiceprint features such as intensity, frequency, formants and MFCC. ANN based PD achieves with the recognition rate of 96%. The HMM based Parkinson detection-based system achieves 95.2% the recognition accuracy. Hence, ANN based Parkinson detection achieves better performance compared to HMM and CNN based Parkinson detection system. The experimental results of PD based detection system prove that the proposed detection method achieves a higher accuracy of the diagnosis of PD patients from healthy people. Therefore, the above approaches can provide a solid solution for the detection of PD in the preliminary or initial stages.
Keywords: *Parkinson Disease, Convolution Neural Network, Artificial Neural Network, Hidden Markov Model, Healthy Controls*

1. Introduction

Parkinson's disease is a neurodegenerative disease that develops as a result of the loss of dopamine, a neurotransmitter. PD is more common in the elderly population, causing changes in gait and posture that can increase the risk of falling and cause mobility problems. As a result, it has an effect on everyday activities and lowers patient and family quality of life. Parkinson's disease is mainly a motor disorder. Inability to move freely, decreased and sluggish movement, increased muscle tonus, and shaking movement in the resting posture are all symptoms of this movement disorder. Other characteristics include a lack of facial expression, difficulty with coordination, and distinctive shifts in speech and voice. During the rapid eye movement sleep period, people with PD can lose their sense of smell and experience sleep disturbances. Around 1% of the population over the age of 60 is thought to be affected by Parkinson's disease.

The cause of Parkinson's disease is unknown in the majority of cases. The most prominent characteristics of this disorder have been discovered to be pathological changes in dopaminergic neurons and neurochemical dysfunction results. The substantia nigra is a black substance formed by the majority of dopamine-producing neurons in the brainstem. This anatomical site has strong connections to other deep brain structures and aids in the production of normal body movement. Lack of dopamine synthesis in the substantia nigra's dopaminergic neurons reduces range of motion and influences voluntary motion. There is currently no cure for Parkinson's disease. The disease develops at varying rates and has a variable path. Various drugs can be used to treat the symptoms of Parkinson's disease. The discussion of current work that is applicable to PD is shown below.

In [1], natural speech analysis and support vector machines were the models used herein the diagnosis of PD from the database containing 43 recordings of patients reading “rainbow passage” evaluated with UPDRS and the accuracy was found to be 81.8%. Parkinson's disease diagnosis can be done using the models, random forest, tensor flow Keras deep neural networks containing voice activity algorithm, where they have used mpower voice dataset containing demographic questions and a voice saying aaah for ten seconds, having the accuracy as 85% [2]. The classification model, logical regression and extreme boosting (Xboost) were also used in the diagnosis of PD, the database used here was from the UCI machine learning repository having 21 features having XGboost accuracy of 96% and LR accuracy of 79% [3]. Acoustic signals with relevant PD can be processed by using an SVM classifier along with a LOSO scheme from Hlavnička et al. database containing 30 PWP and 50 HC and the accuracy as77.3% [4].

For detecting PD, various machine learning models such as logistic regression, naive Bayes, KNN, and forest decision tree were used, with the features used here being minimum-redundancy maximum-relevance and recursive feature elimination. The accuracy obtained was 95.3% using data from the UCI machine learning repository [5]. PD can also be detected from speech articulation neuro mechanics using ANN classifier and RLFSN and the accuracy obtained is 99.4% [6]. Detection of PD with sustained phonation is also done by pre-processing audio signals, feature extraction and training datasets with multiple classifiers from a dataset that contains 99 subjects of both male and female, the accuracy obtained here is 94.5% [7]. The use of genetic algorithms, clustering, and binary classification in the diagnosis of Parkinson's disease has been demonstrated in a database of 195 continuous vowel phonation of 31 men and women, with an accuracy of 85% to 95%. [8].

CNN model based on spectrograms was used in the diagnosis of PD from a database containing 268 subjects having 194 healthy and 74 PD patients, here the accuracy was 95.9% by splitting the sentences and without splitting the sentences the accuracy obtained was 93.2% [9]. A multiple classifier framework having KNN, naive Bayes, SVM were used in the diagnosis of PD, KNN - 77.5%, SVM - 87.5%, Discriminant analysis - 80% Naive Bayes - 82.5% and the training accuracy obtained was 92.5% and the testing accuracy obtained was 86.7% [10]. In the detection of PD, a back propagation learning algorithm based on the Levenberg-Marquardt back propagation algorithm was used, with 195 samples classified using 22 features. The training accuracy obtained was 92.5%, and the testing accuracy was 86.7% [11]. CNN classifier is used in the classification of MRI images of Parkinson’s and healthy patients, the data were taken from the PPMI database which has an accuracy of 96% [12].

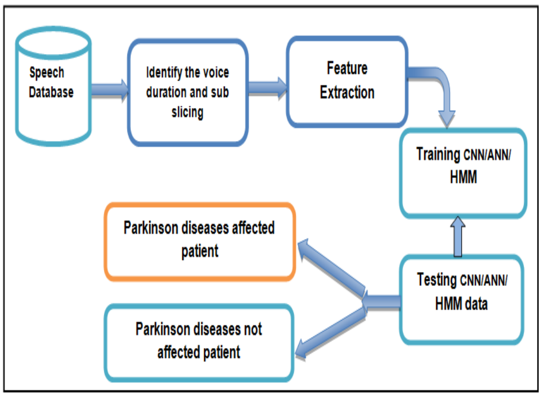


Figure 1 **Proposed PD framework block diagram**

A Parkinson's disease system focused on multiple types of sustained phonations and Linear Discriminant Analysis (LDA) and genetically optimised neural networks has been suggested. This system uses LDA model containing voice samples of 20 men and 20 women, which provides an accuracy of 91.4% [13]. The proposed ensemble machine learning method for voice-based identification of Parkinson's disease uses a support vector machine classifier with 756 instances and 754 attributes and has a recognition accuracy of 91% [14]. Voiceprint recognition of Parkinson’s patients based on deep learning which also uses SVM classifier and WMFCC approach which gives the accuracy 84.5% [15]. The existing work uses different classifier for disease detection and shows the more studies required for PD detection with acoustic features. Hence, to improve the recognition of PD detection system, this work proposes new set of acoustic features. The proposed method is a less expensive and more precise quantitative method of detecting Parkinson's disease in its early stages.

Figure 1 shows the block diagram of proposed PD detection system. The aim of this research is to create a deep learning model that incorporates two machine learning techniques. It is based on two distinct determining factors: spectrogram and acoustic features. The dataset featured has been taken from the Kings College London database. The speech dataset was collected initially and the voice duration of the people with Parkinson’s and healthy controls was identified. After that, the audio samples are sub sliced. The next step involves the conversion of audio samples into spectrogram images and feature extraction. The research delves into all of the important characteristics and uses some of them as feedback to the proposed machine learning model for disease prediction. In this paper, we suggest using a CNN to learn features. A CNN will process information across a series of layers, each of which is responsible for learning a different and finer representation. The data is partitioned into training data and testing data. The final step here involves the application of the classification algorithm. Here CNN, ANN and HMM as the classifiers. After applying these classification algorithms, the training dataset were identified as people affected by Parkinson's and people not affected i.e. the healthy controls.

The following is how the rest of the paper is organised: The study of acoustic measurements of speech signal such as pitch, intensity, spectrogram, and formant frequency is presented in Section 2. Section 3 explains the structure of our proposed model, as well as the dataset and the CNN, ANN, and HMM models that were used. The findings of the experimental study of the PD detection system are presented in Section 4. Section 5 concludes the work.

2. Acoustic analysis of Parkinson system

This section discusses in detail about acoustic analysis of the Parkinson disease patient sound units compared with healthy patient. The acoustic analysis is carried out using the following representation such as waveform, spectrogram, intensity and formant frequency. The participants were asked to read a passage, where the audio clip has been sub sliced to a time duration of 20 seconds which contains only the voice of PD patients and healthy people. Figure 2 shows the waveform representation of Parkinson's disease and healthy patient audio samples represent acoustic measures such as pitch (in blue) and intensity (in yellow). Figure 2a shows the acoustic waveform of the sound unit "The North Wind and the Sun were disputing which was the stronger, when a traveller came along wrapped in a warm cloak" which shows that the some of the peaks are flattened when its compared with the normal patient acoustic sound units (refer figure2b). This replicates that the similar kind of muffling state of throat microphone sound system. There is a pause between the words spoken in that 20 seconds duration for PD patients and for healthy controls, we see a uniform flow of words while speaking a sentence is shown as representation of spectrogram. The amplitude of the frequency variable is expressed by the darkness of a point on a spectrogram. The amplitudes of very dark points are high, while the amplitudes of light points are small.

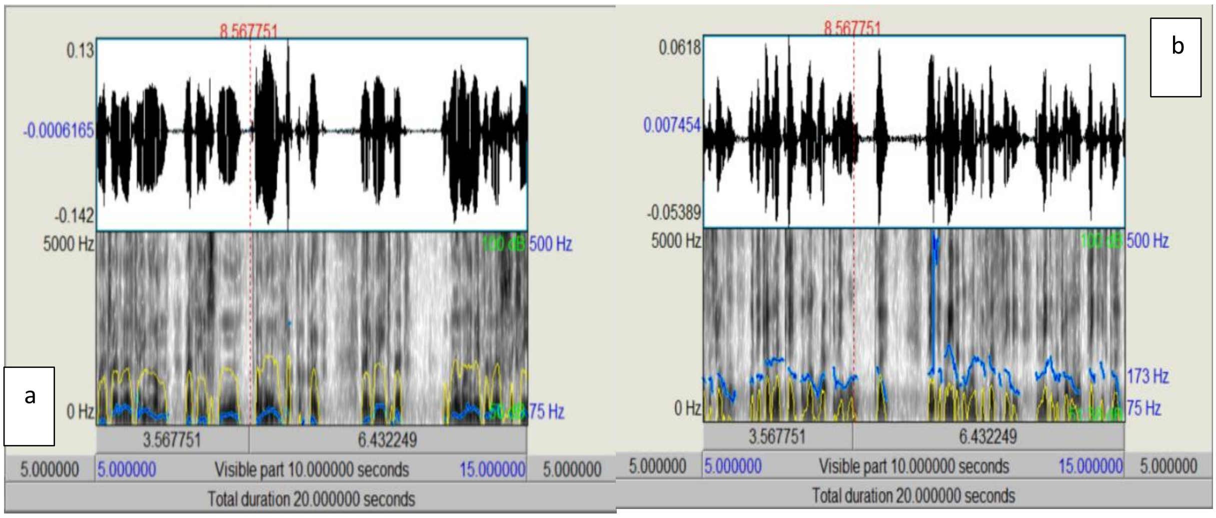


Figure 2. **(a) The waveform of PD patient sound units (b) The waveform of healthy patient of sound units**

Figure 3 shows the variations intensity of Parkinson patient sound unit and healthy conditioned one. Here we find that PD patients pronounce words with less intensity, i.e., below 58.1 dB, hence most of the words were not audible clearly whereas healthy people pronounce words with much higher intensity. Figure 4 shows the formant frequency of the audio sample of both PD patients and healthy individuals. Here, we see that the voice of healthy people indicates the frequency spectrum with clear formant, whereas the formant is not clear in the case of PD patients where the frequency is found to be lower. This analysis gives us motivating to use features such as related to intensity and formant features. The next section discusses in detail about the proposed system design

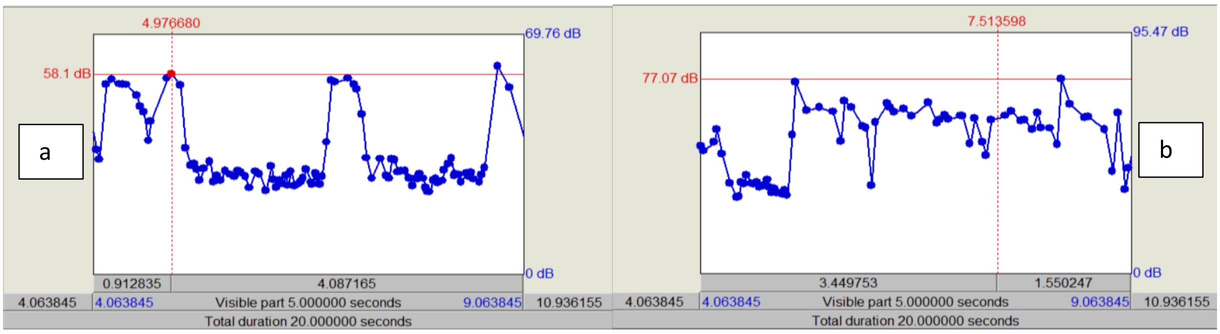


Figure 3. **(a) Intensity plot of PD audio sample (b) Intensity plot of HC audio sample**

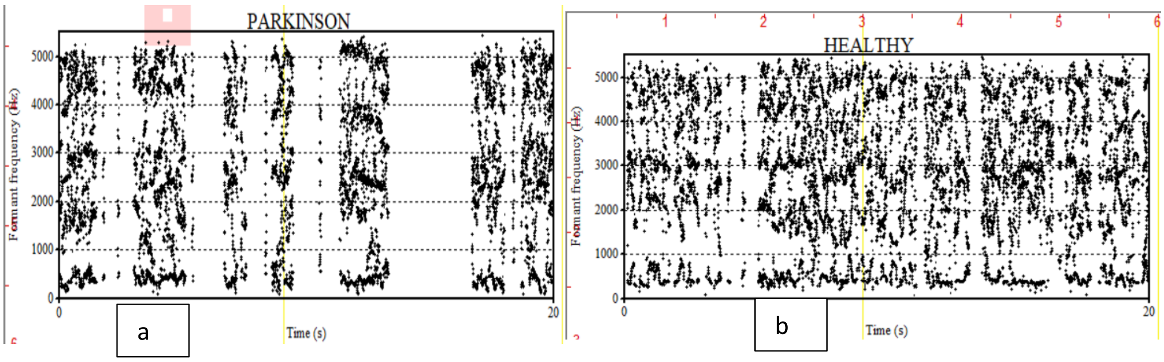


Figure 4. **(a) Formant speckles plot of PD audio (b) Formant speckles plot of HC audio**

3. Proposed System Design

The first section about dataset used for PD detection system is discussed in detail. The various feature extraction methods used for detecting features from the trained data is briefed in the second sub section. The final section shows the various machine learning algorithm used such that Convolution Neural network, Hidden Markov Model and Neural networks.

3.1 Dataset

The dataset used in this analysis was recorded in September 2017 at King's College London Hospital in Denmark Hill, London. Since the voice recordings were made in the practical situation of making a phone call, it is reasonable to conclude that all recordings were made within the reverberation radius and hence are “clean”. The participants were asked to read a passage from “The North Wind and the Sun” and “BNC” in the dataset, and some were asked to start a casual conversation with the participant, in which the test executor asks random questions about places of interest, local traffic, or personal interests, if that was appropriate. The dataset consists of a total of 37 sound recordings from the participants who read out the passage to the test executor and another 37 voice samples of the same participants who were made to engage in a spontaneous dialogue with the test executor. Out of those participants, 21 are healthy controls (HC) and 16 are affected by PD. Every audio sample collected has average time duration of 2 minutes. Each audio sample was sub sliced which involves filtering out the participant’s voice by ignoring the speech of the test executor. This is done to get more accuracy after analysing the samples. The main content i.e., the participant’s voice alone was extracted and further every audio sample was cut into 20 seconds duration using the python package. For making all the samples of equal length, silence is padded at both ends in samples that are lesser than 20 seconds. After sub slicing, the dataset increases with 191 HC audio samples and 131 PD audio samples which are used for the study.

3.2 Feature extraction

Each acoustic sample was split into time duration of 10 seconds and 20 seconds to get maximum accuracy. Then after sub slicing, the features are extracted of dimension 26 from the samples $x(n)$. The first, chroma features are extracted. Human perception of pitch is periodic in the sense that if two pitches differ by an octave, they are perceived as having the same hue. Pitch is broken down into two parts: tone height (octave number) and chroma. It's an efficient representation in which the entire spectrum is projected into 12 bins, resulting in a 12-dimensional chroma vector C as follows.

$$C \triangleq \{\ldots\ldots, C0, C1, C2, C3, \ldots..\} \quad (1)$$

The harmonic progression and normalisation of features are correlated with the sequence of these chroma vectors $(x(n) = x(n)|\|x(n)\|)$ and thus makes the invariant to changes in dynamics. Chroma STFT, which is having a dimension of (12x646), for given a signal $x(n)$, a hop size h, and window W which is a signal of length **N**, the STFT is defined as follow

$$[F\mathcal{T}[k] = \sum Nn = 0X[h\mathcal{T} + n]W\,n]e - j2\pi kn/N \quad (2)$$

Secondly, the Root Mean Square (RMS) having single dimension feature representation which is a measure of an acoustic speech signal and detecting changes in loudness are also important cues for PD detection. This method detects the boundaries based on the dissimilarity measure of the amplitude distributions. Given input sample, the signal is split into short non overlapping frames and the RMS is calculated for each frame. The window size in our implementation is 646 samples, i.e., with a sampling frequency of 16000 Hz, these windows are approximately 10/20ms long. The RMS computation of a given input signal as follows

$$RMS = \sqrt{\left(\sum_{n=1}^{N} x^2(n)\right)} \quad (3)$$

Third, a one-dimensional spectral centroid-based function is used to determine the brightness of a signal. The "centre of gravity" is determined using the frequency and magnitude knowledge from the Fourier transform. The average frequency weighted by amplitudes, divided by the sum of the amplitudes, is the individual centroid of a spectral frame.

$$Spectral\ Centroid = \frac{\sum_{m=1}^{N} k\,SC[m]\,.x(n)}{\sum_{m=1}^{N} SC[m]\,.x(n)} \quad (4)$$

$SC[k]$ is the amplitude in the DFT range that corresponds to bin m. Fourth, spectral bandwidth-based features are used to discriminate the lowest and highest frequency in the spectrum. Spectral bandwidth-based features having the dimension as (1x646), and calculated as follows

$$S.bandwidth = (sum\_kSC[m,t] * (freq[m,t] - centroid[t]) ** p) ** (1/p) \quad (5)$$

Mel frequency ceptstal features (MFCC) having the dimension as (20x646) delta and energy coefficients as fifth one which is calculated as follows

$$s2(n) = s(n) - a * s(n-1) \quad (6)$$

Roll-off based features as a sixth features having the dimension as (1x646), and calculated as follows,

$$[-(1 + ALpha)/2, (1 + ALpha)/2] \quad (7)$$

Last, the Zero-Crossing Rate (ZCR) indicates how much a signal crosses zero in a unit of time, which can happen if successive samples have different signs. A simple measure of the frequency content of an acoustic signal is the rate at which zero-crossings occur.

$$z_n = \frac{1}{2}\sum_n |s[x(n)] - s[x(n-1)]|\ w(n) \quad (8)$$

Where $s[x(n)] = \begin{cases} 1, & x(n) > 0 \\ -1, & x(n) < 0 \end{cases}$ $w(n)$ is a rectangle window of length N.

3.3 Machine learning techniques

The Convolutional Neural Network (CNN) is a version of the standard neural network that has demonstrated excellent pattern and object recognition efficiency. CNN's architecture is made up of consecutive pairs of convolution and pooling layers rather than completely linked layers. Internal representations are generated by convolving the input with a filter mask, which facilitates learning of local features and encourages weight sharing. Each convolutional layer has many kernels that are trained during training, and convolution creates a new feature map that is down sampled to a smaller size with the aid of a pooling layer. Input feature maps in CNN are normally arranged as a 3-dimensional array, with stacked input feature maps being flat 2D planes in the array. The VGGNet architecture, which has 144 million parameters and a stack of small-sized convolutional filters, is proposed in this paper. There are 13 convolutional layers with small-sized kernels, five max-pooling layers, three completely connected layers, and an output classifier layer with a sigmoid activation feature in each of the five convolution and pooling blocks.

Figure 5 shows the VGGNet pre-trained CNN architecture used for the identification of Parkinson's disease in an individual. Therefore, they are divided into 2 classes, one for PD and another for Healthy Control (HC). The spectrogram images of PD were reshaped to 224x224 dimensions and fed to the network. To train the spectrogram images, use the Keras package which has the pre-trained VGG model and loaded the 'ImageNet' weights. There are 5 blocks in the VGG model where convolution and pooling occur in every block. The input layer is 224x224x3 which is fed into the first convolution layer. Every layer gives an output dimension and also the number of parameters that are trained for each layer. The parameters are calculated as,

$$param_{number} = output_{channel_{number}} * (input\_channel\_number * kernel\_height * kernel\_width + 1) \quad (9)$$

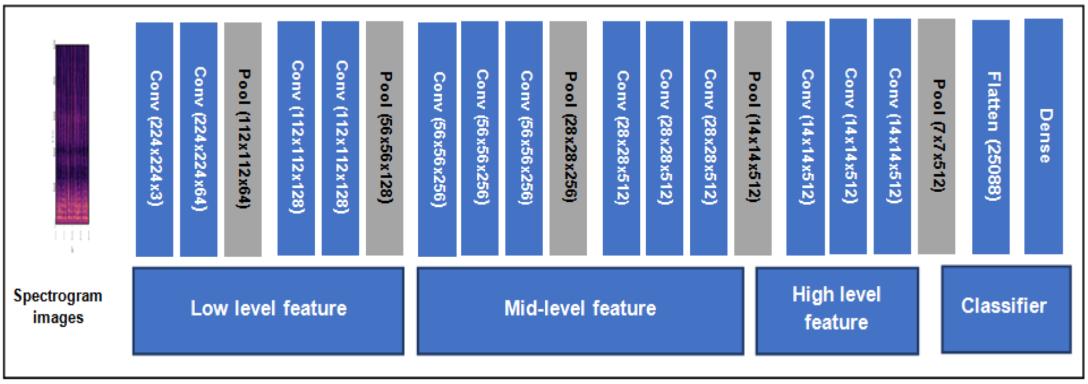


Figure 5. **VGG16-Convolutional Neural Network**

Here, kernel_width and kernel_height represent the size of the filter which is 3x3. The input channel number for the first convolution layer is 3 and the output channel number is 64. By using equation 9, the number of parameters as 64*(3*3*3+1) = 1792. They become the input for the subsequent layers. Every block encompassed 64, 128, 256, and 512 neurons respectively. Finally, the layers are flattened and fully connected which is followed by the sigmoid activation function which is also known as a binary classifier is used, as there are 2 classes for classification. The fully connected dense layer calculates the parameters using the following formula,

$$param\_number = output\_channel\_number * (input\_channel\_number + 1) \quad (10)$$

For classification using CNN, the acoustic samples were converted into spectrogram images as CNN works well for images. After sub slicing each audio sample into 20 seconds, the total dataset count was 322, for passage after sub slicing the audio samples into 10 seconds, the data set count was 150.

Figure 6 shows the second classifier such that Artificial Neural Networks (ANN) used for PD detection system. ANNs are made up of neurons, which are constitutive units that are interconnected by linking connections, each of which has a weight that is multiplied by the signal transmitted in the network. ANNs are data-driven self-adaptive methods that conform to the data without any clear definition of the underlying model's functional structure, and they can approximate any function with arbitrary accuracy. An ANN is made up of three layers: a node input layer, one or more hidden layers, and an output layer. In our case, the input layer is made up of neurons that represent various sound parameters. The secret layer is a set of neurons that act as an intermediary between the input and output layers. The neural network's hidden layer essentially maps the inputs into image space. The number of groups determines the number of neurons in the output layer. When using multilayer feed-forward neural networks to solve problems, one of the most critical considerations is the network architecture.

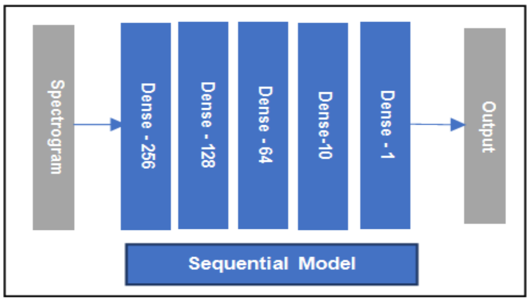


Figure 6. **Working flow of Artificial Neural Networks**

Hidden Markov Model (HMM) is a next classifier model used to capture the properties of acoustic speech signal. In this work, a simple left-to-right HMM model is used for modelling. For the given state models $m = \{n_1, n_2, \dots, n_5\}$ and the observed feature set $F = \{f_1, f_2 \dots . f_n\}$. The initial probabilities of model $\pi = \{\pi_1, \pi_2 \dots . . \pi_n\}$ and the $\pi_n = P(f_n = N_i)$.The transition probability is $T_i = P(f_{i+1} = N_i \,|f_t\, N_i)$. Each state in the HMM model represented as GMM and the model parameter of GMM is $Q = \{P_1, P_2 \dots . P_A, \mu_1 \mu_2 \dots . \mu_{n,} \varepsilon_1, \varepsilon_2 \dots . . \varepsilon_n\}$. The probability distribution for the given observed features $a_i(F)$ is computed by

$$a_i(F) = P(f = v|f_t S_i) = \sum_{k=1}^{M_i} \propto_{jk} P(F/\mu_{jk}, \varepsilon_{jk}) \quad (11)$$

Where $A = \{a_i(F)\}$. From this HMM model $\lambda = (T, A, \pi)$

A single HMM-Gaussian mixture model with two states and varying Gaussian mixtures used experimental purpose. The next section discusses the experimental analysis in detail.

4. Experimental Analysis

The datasets are used for experimental study categorized into three datasets namely passage Dataset1 (D1), spontaneous dialogues Dataset2 (D2) and combined dataset (C12). The number of training and testing dataset for CNN are 322 spectrogram images. In our proposed model, 80% of the training images are used to train the model, with the remaining 20% being used to validate the trained model. The accuracy metrics on the test set were used to assess the deep learning model's efficiency. In first, CNN model-based PD system recognition is discussed. The specification of the spectrogram images comprises of frequency range lies between 0-5000 Hz and the time range is 0-20 seconds. The three VGGNet-16 models were trained over the Parkinson spectrogram image dataset. Figure 7a shows the recognition accuracy with varying time per seconds of acoustic signal. The D1 dataset with 20 seconds duration achieves the recognition of 70% for PD patient identification.

The recognition accuracy of 78% achieved for dataset D2 shows the highest recognition among the entire recognition rate. The combined dataset with trained model gives the best recognition rate of 88%. We kept the threshold time as 10 seconds and 20 seconds, where the participants are asked to read a passage, to get the maximum accuracy. Using CNN classifier and keeping the threshold time as 10 seconds, the training and the testing accuracy obtained are 60% and 61% respectively. Whereas after increasing the threshold time to 20 seconds, the accuracy also increased, the training and testing accuracy obtained is 88% and 70% respectively. But the features trained CNN model gives the 93% recognition accuracy compared to spectrogram images.

In ANN classifiers, a sequential model is used which is a linear stack of layers. The next layers are fully connected layers. The first three layers use the ReLU activation function encompassing 256, 128, 64 neurons respectively. The next layer uses the SoftMax activation function. Finally, the proposed model's final layer used the sigmoid activation feature to divide the Parkinson data into two groups. Figure 7b shows the PD system performance recognition accuracy of ANN classifier.

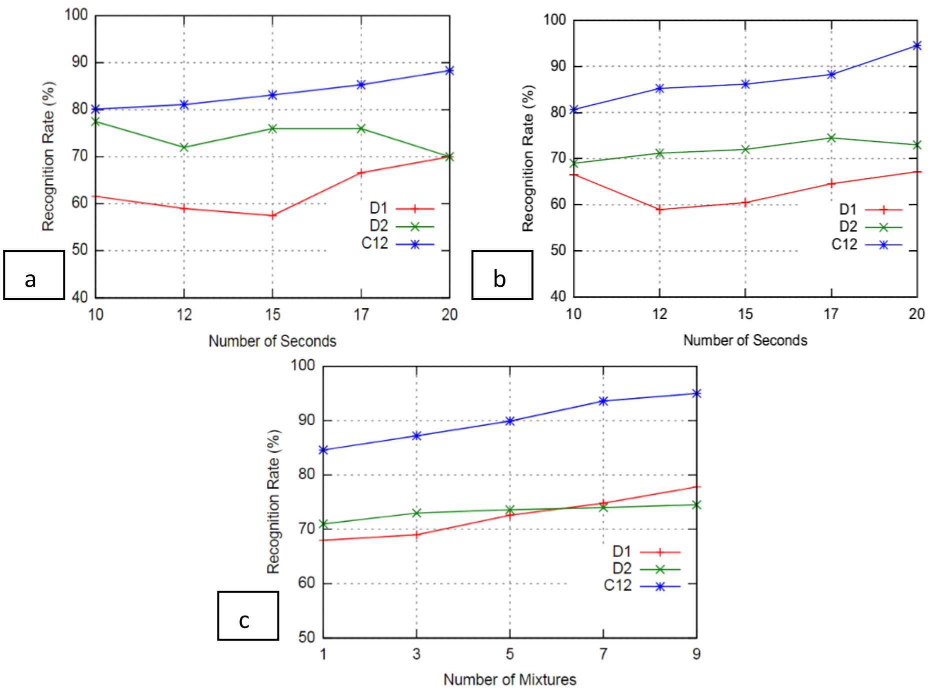


Figure 7. **The PD system performance recognition accuracy of a) CNN b) ANN c) HMM**

Using ANN classifiers and keeping the threshold time to 10 seconds, the recognition rate PD patient recognition rate obtained are 68% and 70% for D1 and D2 respectively, whereas after increasing the threshold time to 20 seconds, the accuracy obtained were 69% and 73% for D1 and D2 respectively. The combined dataset C12 shows the highest recognition rate of 96.2% for PD system recognition. Figure 7c shows the recognition rate of two state HMM with varying Gaussian mixtures. The Gaussian Mixture value stands of M=1 gives the lesser recognition rate compared while increasing the subsequent M values. A two state S=2 HMM with M=9 achieves the high recognition rate 74%, 78%, and 95% for D1, D2, and C12 dataset respectively. Table 1 shows the best recognition rate comparison of D1, D2, and C12 dataset with ANN, CNN and HMM modelling techniques.

Table 1. **Best recognition rate of ANN, CNN, and HMM**

	ANN	CNN	HMM
D1	67.2	70	77.8
D2	73	76	74.5
C12	96.2	88.33	95

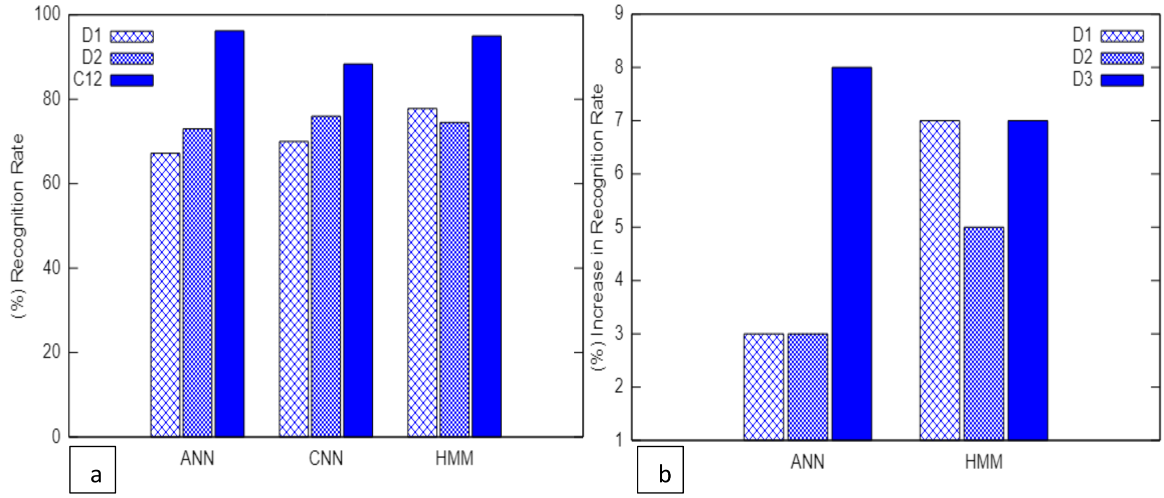


Figure 8. a) **The comparative PD system performance recognition accuracy (%) of b) Increase in recognition rate of best PD system (%)**

Figure 8a shows the comparative recognition system performance of different classifiers. We observed that, HMM based modelling gives the highest recognition rate of 79% for dataset1 which is the passage based one. The recognition rate of 80% achieved for spontaneous dialogue speech for PD patient of dataset 2 represents that small increase in recognition rate compared to dataset1. Hence, the combined dataset proposed with ANN based system gains the recognition accuracy of 96.2%. Figure 8b shows the increase in recognition of rate of best PD systems such as ANN and HMM are compared with CNN based system. The proposed PD system improves the recognition for dateset1 with 3% and 7.2%, dataset2 with 3.1% and 5% and combined dataset with 8.2% and 7% for ANN and HMM respectively. The experimental results shown that ANN based PD patient detection system performance is higher than compared to CNN and HMM model.

5. Conclusion

Using CNN, ANN, and HMM, this paper explores PD identification from a speech signal. Spectrograms and a variety of other short-term features were used as stacked 2D input maps for the CNN. The impact of each segment on PD detection efficiency was assessed and compared to the decision-level fusion of all segments in a speech recording. Despite the fact that deep learning outperforms machine learning models, it is difficult to conclude that the deep learning process is superior to the others. This is because we used a limited audio dataset to classify the deep learning methods. When compared to CNN, the ANN-based recognition produces better performance than the HMM-based recognition. As a result, the findings of this study can be seen as a first step toward applying advanced science to early disease detection. The future work includes the speech dataset, various other symptoms of the Parkinson's people can also be collected to detect PD in a very early stage. The dataset of various motor and non-motor symptoms could also be collected and analyzed in PD detection.

References

[1] Alex Frid, Hananel Hazan, Dan Hilu, Larry Manevitz, Lorraine O. Ramig, Shimon Sapir, “Computational Diagnosis of Parkinson's Disease Directly from Natural Speech Using Machine Learning Techniques”, IEEE, 2014.

[2] Timothy J. Wroge1, Yasin Ozkanca, Cenk Demiroglu, Dong Si, David C. Atkins and Reza

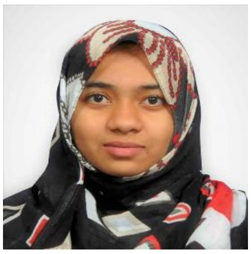
Hosseini Ghomi, “Parkinson’s Disease Diagnosis Using Machine Learning and Voice”, IEEE, 2018.
[3] V. Ulagamuthalvi, G.Kulanthaivel, G.Shri Nikhil Reddy and G.Venugopal, “Identification of Parkinson’s Disease Using Machine Learning Algorithms”, Vol-13, 2020.
[4] AchrafBenba, Abdelilah Jilbab, Sara Sandabad and Ahmed Hammouch, “Voice signal processing for detecting possible early signs of Parkinson’s disease in patients with rapid eye movement sleep behaviour disorder”, Vol-22, 2019.
[5] IqraNissar, Danish Raza Rizvi, Sarfaraz Masood and Aqib Nazir Mir, “Voice-based detection of Parkinson’s disease through ensemble machine learning approach”, EAI.EU, 2018.
[6] Pedro Gómez-VildaJ, Jiri Mekyska, José M. Ferrández, Daniel Palacios-Alonso and Andrés Gómez-Rodellar, “Parkinson Disease Detection from Speech Articulation Neuromechanics”, Vol-56, 2017.
[7] Jefferson Almeida, Pedro Pedrosa Rebouças Filho, Tiago Carneiro and Wei Wei, “Detecting Parkinson’s disease with sustained phonation and speech signals using machine learning techniques”, Vol-125, 2019.
[8] Abhishek M. S, Chethan C. R, Aditya C. R, Divitha D and Nagaraju T. R, “Diagnosis of Parkinson’s disorder through speech data using machine learning algorithms”, Vol-9, 2020.
[9] EvaldasVaiciukynas, AdasGelzinis, AntanasVerikas and MarijaBacauskiene, “Parkinson’s disease detection from speech using convolution neural networks”, Vol-223, 2018.
[10] MahnazBehroozi and Ashkan Sami, “A Multiple classifier framework for Parkinson’s disease detection based on various vocal tests”, 2016.
[11] Ms. Noor Afza, Mr. Manoj Challa and Dr.Jitendranath Mungara, “Speech processing algorithm for detection of PD”, Vol-2, 2013.
[12] Pir Masoom Shah, Adnan Zeb, UferahShafi, Syed Farhan Alam Zaidi, Munam Ali Shah, “Detection of Parkinsons disease in brain MRI using CNN”, 2018.
[13] Liaqat Ali, Ce Zhu, Zhonghao Zhang, and Yipeng Liu, “Automated Detection of Parkinson’s Disease Based on Multiple Types of Sustained Phonations Using Linear Discriminant Analysis and Genetically Optimized Neural Network”, 2019.
[14] IqraNissar, Danish Raza Rizvi, Sarfaraz Masood and Aqib Nazir Mir, “Voice-Based Detection of Parkinson’s Disease through Ensemble Machine Learning Approach: A Performance Study”, 2019.
[15] Zhijing Xu, Juan Wang, Ying Zhang, Xiangjian He, “Voiceprint recognition of Parkinson patients based on deep learning”, 2018.

Dr. N. Radha

Assistant Professor in the Department of Information Technology, SSN College of Engineering, Chennai. She has 14 years of teaching experience, including 1 year of industry experience. She received her Bachelor of Engineering degree in Computer Science and Engineering from Bharathiyar University, Coimbatore. She has completed her Master of Engineering in Computer Science and Engineering from Anna University, Chennai with distinction. Currently, she has completed her Ph. D degree (Part-Time Programme) from Anna University, Chennai in the area of Speech Recognition. She guided the projects in the area audio-visual speech recognition, speaker recognition, video processing and networks for Under Graduate and Post Graduate students. She has published about 20 research papers in her area of research both international journals and conferences.

Mr. R. M. SACHIN MADHAVAN

received his Bachelor of Technology degree in Information Technology from Sri Siva Subramaniya Nadar College of Engineering. He has completed Internal Funded Project on “Women’s Safety Wearable Device”. He has published 2 research papers in her area of research in international conferences. His research interest includes video processing, speech recognition, speech synthesis and machine learning



Ms. S. SAMEERA HOLY
received her Bachelor of Technology degree in Information Technology from Sri Siva Subramaniya Nadar College of Engineering. She has completed Internal Funded Project on “Women’s Safety Wearable Device”. She has published three research papers in her area of research in international conferences. Her research interest includes speech recognition, Audio-Visual speech recognition, speech synthesis and computer vision.