IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 3 June 2024, accepted 13 June 2024, date of publication 19 June 2024, date of current version 28 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3416444

RESEARCH ARTICLE

A Visibility Graph Approach for Multi-Stage Classification of Parkinson's Disease Using Multimodal Data

K. DEEPA RAJ[1], G. JYOTHISH LAL[1], E. A. GOPALAKRISHNAN[2,3], V. SOWMYA[1], AND JUAN RAFAEL OROZCO-ARROYAVE[4,5]

[1]Amrita School of Artificial Intelligence, Amrita Vishwa Vidyapeetham, Coimbatore 641112, India
[2]Amrita School of Computing, Amrita Vishwa Vidyapeetham, Bengaluru 560035, India
[3]Amrita School of Artificial Intelligence, Amrita Vishwa Vidyapeetham, Bengaluru 560035, India
[4]GITA Laboratory, Faculty of Engineering, University of Antioquia UdeA, Medellín 050010, Colombia
[5]Pattern Recognition Laboratory, Friedrich-Alexander Universität Erlangen-Nürnberg, 91054 Erlangen, Germany

Corresponding author: G. Jyothish Lal (g_jyothishlal@cb.amrita.edu)

This work was supported in part by Amrita Vishwa Vidyapeetham (AMRITA) Seed under Grant ASG2022118, and in part by Comité para el Desarrollo de la Investigación (CODI) at Universidad de Antioquia (UdeA) under Grant PI2023-58010.

ABSTRACT Parkinson's disease (PD) is a neurodegenerative disorder characterized by several motor symptoms such as resting tremor, muscular rigidity, slowness of movement and different speech impairments. PD is a kind of singular, multi-system disorder that gradually worsen over the time. In this research, we classify the neurological state of the patients with Parkinson's disease (PwPD) according to the third section of the Movement Disorders Society - Unified Parkinson's Disease Rating Scale (MDS-UPDRS-III) using multimodal bio-signal data. As PD advances from low to advanced state, PwPD finds it difficult in their speech production and irregularities in their gait patterns. Monitoring the chaotic nature of time series data corresponding to speech and gait biomarkers can provide insights into the progression of the condition across different stages. This work for the first time analyze PD in a complex system perspective while representing the biomarkers as complex networks. The time-series corresponding to speech and gait signals are represented separately, as a complex network using the visibility graph algorithm. The characterization of the different stages of PD is explored for each modalities using different network features. Performance evaluation shows that the results obtained using the multimodal configuration of speech and gait left foot signals outperform the state-of-the art method. Moreover, performance comparison with the unimodal counterparts proves the need for multimodal assessment of PD severity. The configuration 'speech and gait left foot' outperforms (in terms of accuracy) that of the unimodal by 32% in speech, 3% in gait left foot, 19% in gait right foot, and 3% in gait both feet.

INDEX TERMS Complex network, complex systems, gait signal, graph, multi-modal data, network features, Parkinson's disease, speech signal, visibility network.

I. INTRODUCTION

A. PROBLEM DEFINITION AND MOTIVATION

Parkinson's disease (PD) is a progressive neurological condition of the central nervous system that primarily causes movement disorder. The symptoms begin with the degeneration of nerve cells in the region of brain called substantia nigra and the level of dopamine produced reduces [1]. Dopamine is an important chemical that facilitates the coordination of nerve and muscle cells involved in organism's movement. When enough dopamine is not produced the symptoms of PD starts to appear. The research study shows the signs of PD mostly develop after the age of 60 and only 5-10% of people experience onset before the age of 50 [2]. Early signs of PD motor symptoms include tremors, postural instability, slowness of movements, speech impairment and rigidity [3]. The patient with PD (PwPD) is as singular like any other people after all the symptoms varies in

The associate editor coordinating the review of this manuscript and approving it for publication was Juan A. Lara.

person to person over the period of time. Some people may experience changes in speech utterance like soft, quick or slurred speech, difficulties with handwriting, change in gait movements and thus making their daily activities difficult and time-consuming [4]. Currently, there is no known cure from PD [5]. However, upon early PD diagnosis and understanding the disease people can execute a proper medication strategy. This can ease the PD progression and PwPD can have a good quality of life.

The Parkinson's research community is still examining the PD with different bio-signals such as speech, handwriting, gait and neuro imaging data for a systematic diagnosis of disease. Gradually PD is affected with multiple motor impairment and the complexity involved in diagnosing the severity of the disease is also high. The clinical study indicates speech is affected at the beginning stages of PD and gradually the PwPD finds it difficult in their movements like gait impairments, hand tremors. For instance, Vásquez-Correa et al. [6] examined the multimodal data by integrating the information from start or stop movement transition in speech, handwriting and gait signals considering the deep learning architectures. Thereby, assessment of multiple motor impairment in PwPD provide unbiased information to clinicians to make right decisions about the treatment. Likewise taking into account the uniqueness and multimodal nature of symptoms in PD patients, the diagnosis of the neurological state of PD clearly does not benefit from solely relying on a single modality [7]. It is therefore imperative to have an investigation on information extraction from multiple modalities, speech and gait signals, to have more accurate assessment of PD.

B. NOVELTY

A complex system is a large network of relatively simple components with no central control, in which emergent behaviour is observed. Some real time examples are the human brain, the immune system, biological cells, metabolic networks, ant colonies, the Internet and World Wide Web, economic markets, and human social networks. The emergent complex behaviour is the fact that the system's global behaviour is not only complex but arises from the nonlinear interaction of the simple components. Most importantly, this emergent behaviour cannot be studied from its constituent components. For instance, the book by Newman et al. [8] emphasize the importance of "network thinking" in dealing with complex systems in the real world. The representation of complex systems in a complex network approach provides a comprehensive understanding of complex connectivity patterns in dynamical systems [9]. The characterization of network features such as the degree distribution, clustering coefficient, average path length is to better understand networks from a scientific point of view and to develop better technologies for designing and managing networks in desired ways.

Parkinson's Disease being a multi-system neurodegenerative disorder, the emergent behaviour from the interaction of many subsystems for speech production or gait movements can be viewed from complex system perspective. Precisely, we use the approach to detect the severity of Parkinson disease (PD). The perspective is motivated from the fact that the emergent behavior resulting from the interaction of many subsystems contribute to different stages in PD patients, whether it is speech produced or gait movements. For example, the dynamic variations occur during speech production can be attributed to the synchronization and desynchronization between the vocal folds and the sub/supraglottal system as reported in Zhang et al. [10], [11]. This process plays a crucial role in determining the vibratory characteristics and quality of the voice. This interplay between synchronization and desynchronization isn't unique to speech production system alone. This is a characteristic of many dynamical systems, which occurs due to the nonlinear interaction of many subsystems within the system. Consequently, dynamic variations occur in the timeseries data for both normal or PD patients in different stages, whether it is speech production or gait movements, is an emergent behavior (a characteristic of any complex system) from the interaction of many subsystems of human physical body. Moreover, PD evolves differently in every patient. Therefore, a reasonable approach to model PD progression is the one based on complex networks because the representation space of the disease condition would be very complex, and patient's specific patterns will be "embedded" differently in such a complex space.

The transformation of time-series data into a complex network domain supports the characterization of dynamical states involved in its production [12]. Some of related studies based on the complex network perspective are the dynamic variations of pseudo periodic time-series in cardiac behaviour [13], the different dynamic time series behaviour like chaotic, hyperchaotic, random and noisy periodic determined from a subset of networks in a dynamical system [14], recurrence network based analysis of complex system [15], [16], time-series forecast [17], thermoacoustic systems [18]. Nevertheless, there are no studies in literature (to the best our knowledge), which investigated the multiple stages (low, intermediate, and severe) of PD from a complex system perspective. Hence, the present work tries to investigate the problem from a complex network perspective using the visibility graph algorithm by Lacasa et al. [19], [20]. The choice of visibility algorithm for transformation of speech/gait signals into corresponding network is motivated from its success on other similar complex system studies. For instance, thermoacoustic systems [18] has explicitly used visibility algorithm to study the different regimes of dynamical transitions occurring in thermoacoustic system. Moreover, Bhaduri et al. [21] have employed visibility algorithm on speech data for characterization of emotions. Other applications include financial time series [22], electroencephalogram (EEG) signal [23], traffic data [24], and earthquake time series [25]. Thus, we conjecture that the use of visibility algorithm for network transformation of time-series corresponding to speech and gait movements can unravel the variation in the dynamics through its connection patterns. Hence, the novelty of the present work lies is in the use of visibility graph and its corresponding network features alone in the characterization/ classification of the variation in dynamics during the three stages of PD.

In the present study, we make the first attempt to investigate the different stages of PD using visibility network representation of different biomarker signals. Also, we intend to provide qualitative and quantitative characterization for the networks we created. After transforming the speech and gait signals into corresponding visibility networks, we try to identify dynamic regimes of PD severity by examining changes in network properties such as its connectivity measures, centrality measures, and network efficiency. These network properties can aid in early diagnosis, tracking disease progression, and assessing treatment efficacy. Precisely, a network transformation can give visual depiction for the underlying dynamics of the different dynamical stages (defined according to the MDS-UPDRS-III score [26]) in PwPD. Further, a multi-class classification is performed using machine learning models trained on network features to classify PwPD into different stages of disease severity such as low, intermediate and severe.

C. CONTRIBUTIONS

The major contribution of this research study are as follows:

- Characterization of the multiple stages in PD from a complex system perspective. We showcased the variations in the different stages of PD using network structure and network features.
- Developed a multimodal approach for the classification of PD stages using only network features to model speech and gait signals.
- Compared the performance of the proposed multimodal approach with state-of-the-art methods and its unimodal counterparts.
- Investigated the global feature importance and local feature importance of the proposed model in predicting the stages of PD severity.

The rest of the paper is organized as follows. Section II the literature review discussing the assessment of PD from speech, gait, and multimodal data in PwPD. The materials and methods employed in this study are detailed in section III. In section IV, initially we present the various experiments conducted, and later illustrate and discuss the results obtained for the multimodal assessment of PwPD performing the classification of different stages of PD. Section IV also explains the different ablation studies conducted. Finally, in section V we provide the conclusions of the study with limitations, and directions for future studies.

II. LITERATURE REVIEW

We extensively studied the articles corresponding to the assessment of PD with speech (section II-A, gait (section II-B), and multimodal data (section II-C). Primarily, we focused on the objective of multi-class classification of PD into different stages of severity according to the clinical motor impairment evaluation scale, MDS-UPDRS III. This supports the identification of relevant research gap as well as uniqueness of our research.

A. ASSESSMENT OF PD FROM SPEECH MODALITY

Among the different motor symptoms that cause PD, the research study found that speech impairment appears at the initial stages of the disease and worsens along the progression of disease [27]. When the dopamine level in brain is reduced, it in turn affects the communications between the brain cells. The speech impairments observed in PwPD includes the rigidity of the vocal folds, bradykinesia, and reduced muscular control of the larynx and other organs involved in speech production and are typically grouped and called *hypokinetic dysarthria* [28]. In the initial stages, the speech becomes slower and soft, mono-pitched, monoloudness which makes it difficult for others to understand. On progression of disease, in its mid stage PwPD may experience more pronounced slurring of speech and increased difficulty with voice projection. In advanced stages of PD, the speech becomes affected severely leading to significant communication challenges, dramatically reducing the patients' intelligibility. Patients may have very soft or hoarse voice quality, frequent pauses between words, the rapid repetitions of words and syllables and sudden deceleration or acceleration in speech [29]. All of these speech impairments are described in terms of four different dimensions such as phonation, articulation, prosody, and intelligibility [30]. Recently, a new dimension called Phonemic Identifiability has been introduced, which has shown high sensitivity in modeling motor and cognitive decline in PD patients [31]. The most common speech tasks the authors considered are the phonation of sustained vowels, reading isolated sentences or read story, diadochokinetic (DDK) exercises like rapid repetition of the syllables /pa-ta-ka/, and monologues.

Our research study mainly focused on the objective of multi-class classification of PD into different stages of severity according to the neurological score, MDS-UPDRS-III. In [32] the authors considered the phonation, articulation and prosody features to classify the PwPD in three levels of disease (initial, intermediate, and severe), and reported an unweighted average recall (UAR) score of 59.1%. Using the acoustic features of voice recordings, the voice disorders in PwPD is examined and classified into three different stages and healthy control (HC). To infer the severity of PD, the authors performed different learning strategies on speech tasks like sustained phonation, DDK and reading tasks. The combination of three tasks features, they resulted in a mean absolute error of 5.5 in [33]. The Interspeech 2015 Computational Paralinguistics Challenge (ComParE) [34] had estimated the neurological state of PD according to MDS-UPDRS-III score. The authors performed with the prediction of clinical scores with the gaussian process and deep neural networks (DNN) and reported spearman correlation of 0.649. Following this method, the authors in [35] evaluated the neurological state of PD with the articulation and intelligibility features and obtained a spearman's correlation of up to 0.72 with the linear support vector regression (SVR). In [36], the neurological state is predicted with the computation of cosine distance between the i-vector extracted from each features and a spearman correlation of -0.48 is reported for PD i-vectors. The authors in [37] used the similar set of features in ComParE challenge and explored the relationship between the features and the PD symptoms. With the support vector regression (SVR) and Gaussian mixture regression (GMR) methods they mapped the selected features to the MDS-UPDRS-III scores and reported the spearman correlation of 0.52.

In [38] and [39] the author investigated the dysphonia measures for the assessment of multi-class classification of PD into different stages using the Local Learning Based Feature Selection (LLBFS) algorithm (with an accuracy of 96.5%) with subspace discriminant and SVM classifier (with an accuracy of 92.5%) respectively. The most relevant features are selected from each sample in cepstral domains using the Mel-frequency cepstral coefficients (MFCCs) and Perceptual Linear Prediction (PLP) methods and therefore achieved an accuracy of 86.7% [40]. A model developed from the prosody features using the Speech Enhancement (SE) algorithm was used in [41] for the multi-class experiment following a SVM classification of a one vs all strategy. They obtained an UAR score of 67% for the training dataset and 51% for the test set data. In [6], the author classified the HC and the three stages of PD in a CNN based approach and achieved an UAR score of 37.8%, for the speech onset data. A multimodal architecture combining the feature learning and feature engineering approach was proposed by the authors in [42] to assess the severity level of PD. The sustained phonation vowels /a/, /i/ and /u/ are the speech data considered for the analysis and obtained a balanced accuracy (BACC) of 52% and Average Mean Absolute Error (AMAE) of 64% for the multimodal deep learning approach. In [43], the multi-class classification of PD using the phonetic characteristics of speech data performed 98.63% of accuracy for the KNN classifier. The relationship between the PD severity and the neurological state of PD is evaluated in [44]. Using the time frequency features from words and DDK speech tasks, a multi-class framework is developed with random Forest classifier along with an Error-Correcting Output Code (ECOC) approach. An unweighted average recall (UAR) score of 48.8% is reported for the word /petaka/. In [45], the author investigated in five speaking tasks such as vowel task, sentence task, DDK task, read text task, and monologue task and four speech impairment features and their fusion for the automatic classification of speech into three PD severity level classes (healthy, mild PD and severe PD) and obtained an accuracy of 58% for the monologue task using the articulation features and the Multi layer perceptron (MLP) classifer. In [46], the author considered experiments with all speech features extracted from phonation, prosody, articulation, openSMILE. phonological and recurrent autoencoder (RAE) for the multi-class classification of PD using the SVM classifer and the highest unweighted average recall score of 49.1% was obtained with the phonation features.

In short, speech being one of the early biomarker in PD, subtle changes are identified from these features of speech signals. Table 11 (refer in Appendix A) shows the summary of the multi-class classification of PD into different stages of severity from the speech data using diadochokinetic (DDK) tasks and according to MDS-UPDRS III score.

B. ASSESSMENT OF PD FROM GAIT MODALITY

In PwPD, the gait impairment is developed when nerve cells in basal ganglia start to damage or die [4]. The nerve cells residing in this region controls the movements happening in human body. One of the main symptoms visible in its initial stages is slowness in movements. The brain is less able to process what would normally be. Doctors evaluate the common automatic movements like swinging arms when walking or moving one foot after another for stage analysis. In mid-stages, PwPD may experience freezing of gait, where they suddenly feel as though their feet are glued to the ground and have difficulty in taking steps, especially when initiating movement (also known as akinesia) or navigating through tight spaces. Gait disturbances lead to increased risk of falls and loss of mobility towards the advanced stage of PD. Patients may also experience severe freezing episodes and difficulty maintaining balance while walking or standing. The stability decreases and the tremor appears to be more in the gait kinetic movements [5].

In this section we reviewed the work based on the analysis of gait impairment computed from the signals obtained from the inertial sensors attached to the foot. For example, the signals captured from the eGaIT system (embedded Gait analysis using Intelligent Technology) consisting of inertial sensors like accelerometers and gyroscopes attached to the lateral heels of the shoes [47]. The activities performed by the participants for the evaluation include walking straight for 10 meters and returning to the departure point (two times, namely 2×10 task, and four times, namely 4×10 task), heel-toe tapping, and circling foot movements. The common features extracted are kinematic, spectral, statistical, nonlinear dynamics (NLD) features and also spectrogram representation of raw signals for the assessment of PD from gait biomarker.

In [48], the authors collected gait data from patients examined under the study of movement disorder at the University Hospital, Erlangen, Germany. The 3D gyroscope and 3D accelerometer sensor units are attached to the lateral heel of both shoes to measure the angular velocity and acceleration. From these captured signals, the sequences of signals, step feature and frequency are analyzed to model gait impairments in patients. The authors in [47] computed the spectral and statistical features for the three stage (low, medium and high) classification of PD. The best performance result was yielded by the SVM classifier with linear kernal with an accuracy up to 89%. A deep learning approach was proposed in [6] using the spectrogram images of gait signals for multi-class classification of PD including HC. The authors analyzed the onset and offset regions of gait signals and achieved an UAR score of 48.6% and 50.9% respectively. Similarly in [49] using the raw gait signals from a multimodal corpus dataset, a multi-class classification was performed with the CNN approach. The reported result for the multi-class classification is UAR score up to 64.9% and F1 score up to 0.632. A summary of aforesaid studies is shown in Table 12 (Refer Appendix A).

In summary, most of the assessment of Parkinson's disease from gait is based on different classical features extracted from the gait patterns observed in patients. Here, we have discussed the studies related to the use of inertial sensors and the way the PD severity is classified using the different types of features extracted from the gait signals.

As the present study is focused on multi-class classification of PD stages, we excluded the studies on binary classification task 'PD Vs Healthy subjects' using speech and gait signals, separately.

C. ASSESSMENT OF PD FROM MULTIMODAL DATA

Although there are several studies considering unimodal data for the assessment of patients with PD, but combining the data from different modalities, multiple sensors is an approach to strengthen the performance of the PD diagnosis. The authors in [7] also proposed an interest in developing a system based on integrating the signals from multimodal data. The fusion of input modalities for example, speech recordings signals, data captured from different hand gestures using sensors and also exploiting multisensors for capturing signals from human movements can exhibit the advantage of extracting maximal information for getting more insight about the complexity involved in assessing the PD.

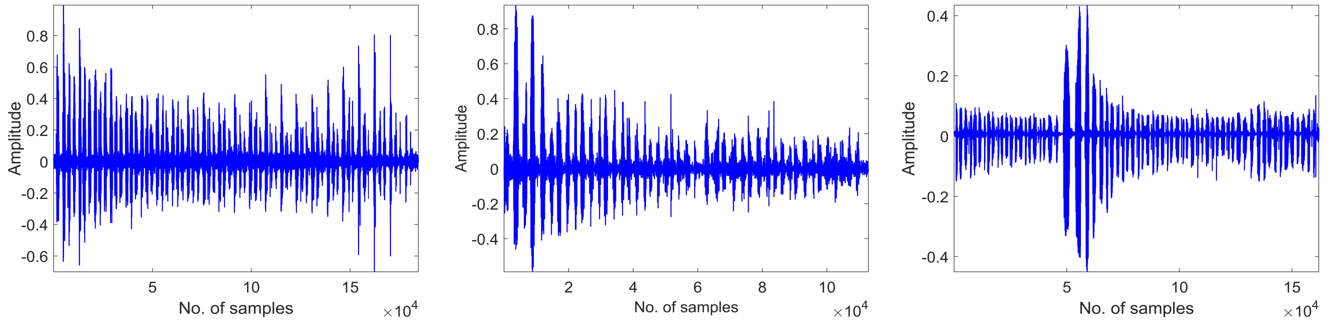
A multi-class classification of patients in different levels of motor severity (low, intermediate, and severe) according to the MDS-UPDRS-III score is reviewed in this section. In [50] the authors analysed the features extracted from the integration of signals from wearable motion sensors and audio sensor using an empirical wavelet transform (EWT) and an empirical wavelet packet transform (EWPT), respectively and achieved an accuracy of more than 90% for Extreme Learning Machine (ELM) classifier with RBF kernel. A deep learning approach proposed in [6] examined the evaluation of the neurological state of the patients with the performance of multi-class classification. The results obtained for the onset and offset data from the fusion of features from speech, handwriting, and gait modalities is UAR value of 55.6% and 45.2%, respectively. In [51], super-i-vectors is formed from the three bio-signals speech, handwriting and gait and score pooling. The evaluation of the neurological state of the patients are represented with the Spearman's correlation between the cosine distance and the MDS-UPDRS-III score for the young HC, elderly HC, and PD. Similarly, in [52] the performance with fusion of all set of features extracted using the user models based on Gaussian mixture models - universal background models (GMM-UBM) and i-vectors computed a Spearman's correlation of 0.634 and a mean absolute error of 10.5. In [53], the authors estimated the progression of PD using the data captured from smartphones. The severity of PD patients is determined by the Deep Multi-Layer Perceptron (DMLP) classifier from the behavioural data of speech and movement patterns captured via smartphones. All the above mentioned research studies is summarized in Table 13 (Refer Appendix A).

In [54] the authors considered the use of the generalized canonical correlation analysis (GCCA) to improve the prediction of the MDS-UPDRS-III score when only one modality is available in test (speech recordings), but multimodal information (speech, gait, and handwriting) is available in train. Improvements of up to 0.17 in the Spearman's correlation of speech representations with respect to the MDS-UPDRS-III were reported by the authors when the train model considers multimodal information.

III. MATERIALS AND METHODOLOGY

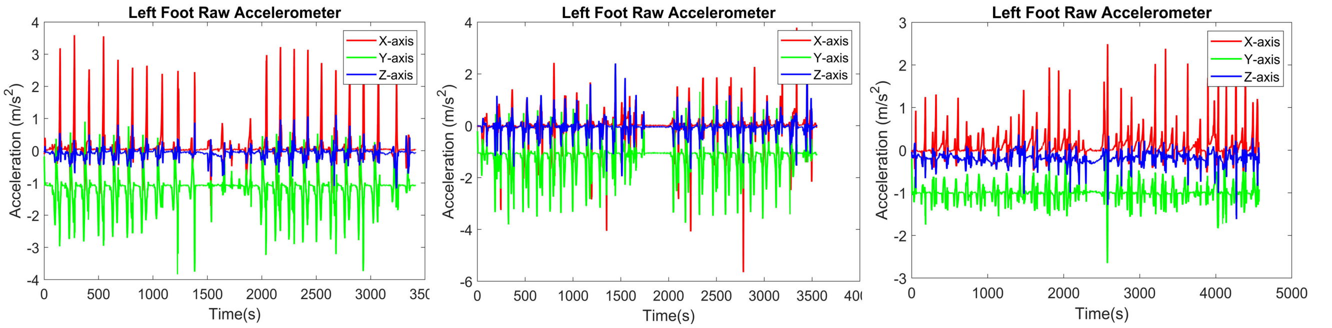
A. PROPOSED COMPLEX NETWORK FRAMEWORK

In this study, the severity of PD is identified using the information obtained from different biomarkers such as speech and gait signals. Figure 1 and Figure 2 respectively depicts the time series data of speech and gait signals captured from different sensors of three PD patients from different stages of PD severity. The dynamic behaviour evolved in the time series data of these biomarkers is analyzed with complex network approach. The methodology developed for the classification of different stages of PD in a complex network perspective is illustrated in figure 3. The raw time series data signals are normalized with min-max normalization in the data pre-processing step. The normalized signals are fed as input to the visibility graph algorithm [19], [55] and transformed into visibility graphs. The motive for this proposed methodology is to unveil the structural patterns hidden in the time series data. From visibility graphs, the network features are extracted to quantify the characteristics of the time series data. For a multimodal interface, the network features are fused at feature-level [7] to take the advantage of extracting the information of closely-coupled modalities. Here, we concatenated features estimated from the network corresponding to each modality as a single vector representation. Further, a multi-class classification of this feature vector is done using machine learning classifier, Random Forest [56]. The proposed model classifies the PwPD according to the PD severity (low, intermediate and severe state). The details of each step in the methodology are presented in the subsections below.



(a) Low (b) Intermediate (c) Severe

FIGURE 1. Speech signal of three PD patients in three different stages of disease. (a) Low state;(female, 63 years old, MDS-UPDRS III score = 19), (b) Intermediate state; (female, 58 years old, MDS-UPDRS III score = 35) and (c) Severe state; (female, 75 years old, MDS-UPDRS III score = 106). All time series corresponds to the syllable /pa-ta-ka/.



(a) Low (b) Intermediate (c) Severe

FIGURE 2. Gait Signal of three PD patients in three different stages of disease. (a) Low state;(female, 63 years old, MDS-UPDRS III score = 19), (b) Intermediate state; (female, 58 years old, MDS-UPDRS III score = 35) and (c) Severe state; (female, 75 years old, MDS-UPDRS III score = 106). All time series corresponds to x-axis, y-axis and z-axis of left foot captured from accelerometer sensor. Gait task considered: 2 × 10m walk with stop.

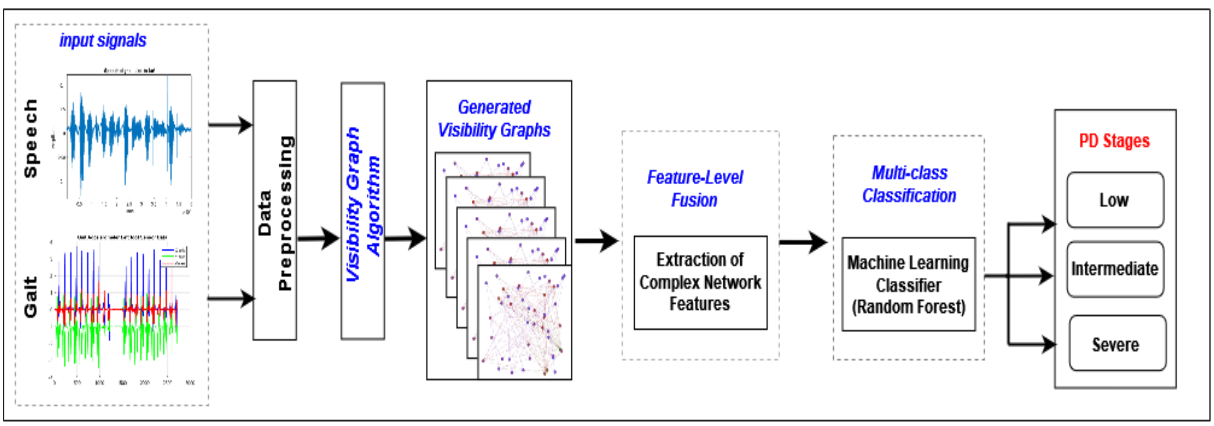


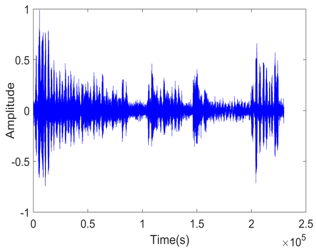
FIGURE 3. Illustration of the methodology followed for the classification of different stages of PD.

B. DATA DESCRIPTION

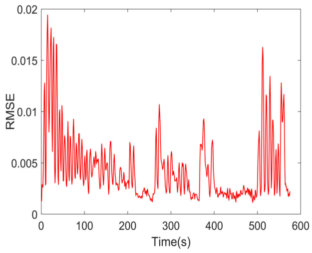
A multimodal corpus dataset procured from GITA laboratory is used in this research study. Information of 38 PD patients (including 13 males and 25 females) collected from two bio-signals such as speech and gait signals is shown in Table 1 [6]. The speech data recordings of the participants has a sampling frequency of 16 KHz and 16-bit resolution. We considered the speech recordings of DDK exercises which is the pronunciation of rapid repetition of the syllable /pa-ta-ka/. Gait signals were recorded with the eGait system [47]. The system uses a sampling frequency of 102.4Hz and 12 bit resolution and consists of a tri-axial accelerometer and a tri-axial gyroscope. We have considered the gait tasks performed by the PD patients which is a 2 × 10 meter walk in a straight line with a stop after 10 meter and then turns clockwise and comes back. The signals from accelerometer are: *accXLeft*, *accYLeft*, *accZLeft*, *accXRight*, *accYRight*, *accZRight* and

TABLE 1. General Information about Multimodal Data of three stages of PwPD.

	PD patients					
	Low		Intermediate		Severe	
	Male	Female	Male	Female	Male	Female
No. of Subjects	4	9	4	9	5	7
Age [in years] ($\mu \pm \sigma$)	59.8 ± 8.3	57.1 ± 12	62 ± 17.6	60.7 ± 6.9	66.8 ± 3.7	57 ± 15.9
Range of Age [in years]	48 - 65	29 -71	41 – 82	53 - 73	61 – 70	34 - 75
MDS-UPDRS – III score ($\mu \pm \sigma$)	13 ± 4.5	16.4 ± 4.4	27.8 ± 2.5	32.8 ± 4.1	57.4 ± 17	59.9 ± 22.1
Range of MDS-UPDRS – III score	8 - 19	9 - 21	25 – 31	26 - 38	40 – 82	43 – 106



(a) Speech signal



(b) RMS Energy of Speech signal

FIGURE 4. The compact representation of speech signal (corresponding to the syllables /pa-ta-ka/) of PD patient after RMS Energy computation.

from gyroscope: *gyrXLeft, gyrYLeft, gyrZLeft, gyrXRight, gyrYRight, gyrZRight*. For further details about the recording setting please see [47].

The MDS-UPDRS III score assigned to the patients are according to their level of motor impairment found during the clinical activities by the neurological experts. The range of MDS-UPDRS III score is defined as follows: 0 to 25 (low state), 26 to 40 (intermediate state) and 40+ (severe state). The range of the different stages of PD is defined in order to have balanced data in each of the three classes. Based on this MDS-UPDRS III score distribution, the multi-class classification of different stages of disease severity is performed. The number of samples in three stages of PD are as follows: low state (13 samples), intermediate state (13 samples) and severe state (12 samples), as shown in Table 1.

C. DATA PRE-PROCESSING

In data pre-processing step, the time series data are normalized with min-max normalization for speech and gait signals. Further, a two-step data pre-processing is done for the speech data signals: (1) zero padding technique [57] and, (2) compute the zero-mean Root Mean Square Energy (RMS -energy) [58].

Firstly, there exists a natural delay in the bio-signal data captured from PD patients due to the affect of different levels of motor impairments. The speech signals have varying length recordings for the rapid repetition of the /pa-ta-ka/ syllable. So in order to obtain a fixed length of signals for processing the data, zero-padding technique is done with the signals so that the length equals to the maximum length of the speech signal in each of the three stages of PD.

Secondly, considering the computational complexity involved while using the original speech data for the complex network transformation, the speech time series samples are converted into more compatible representation. This transformation is motivated from [59] where the author proposed of the method of applying sliding window for each signal with an overlap of 50% followed with the computation of zero-mean Root Mean Square Energy (RMS Energy) on each frames. The new compact time series representation of the original speech signal satisfies the condition for visibility graph algorithm which requires positive time series values. Figure 4 shows the original speech signal of PwPD with MDS-UPDRS-III score as 46 (patient in severe state) and its corresponding compact time series representation (RMS energy).

D. VISIBILITY GRAPH ALGORITHM

The transformation of time series data into complex network represents the subsystems/components of the complex system as nodes and interactions of the components as edges [60]. In this context, defining the rules of the transformation is important. In the proposed work, we adopted the *Visibility Graph Algorithm* from Lacasa et al. [19], which is a simple and fast method for network creation.

A network is represented as an adjacency matrix $A_{i,j}$ of size $N \times N$, where N is the total number of nodes and i, j are the pair of nodes. If a pair of nodes is connected, $A_{i,j}$ is set to 1; otherwise, it is set to 0. Therefore, it is a symmetric matrix with it's elements as zero's and one's. In the visibility graph algorithm, the nodes correspond to the positive data values in the time series. The edges between nodes are created based on the visibility condition. Reference [19] defined the condition as, any two nodes are linked when a straight line can be drawn between them, without intersecting any intermediate nodes. Mathematically, it can be expressed as follows.

$$A_{i,j} = \begin{cases} 1, & \text{if } A_i < A_j + (A_i - A_j)\dfrac{ts_i - ts_k}{ts_j - ts_i} \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

Precisely, the two nodes at timestamps ts_i and ts_j are connected if it satisfies the condition for all the points at ts_k with $ts_i < ts_k < ts_j$. As a result, the graph network derived from a time series data is always connected, undirected and constant that corresponds to its time series data. Figure 5

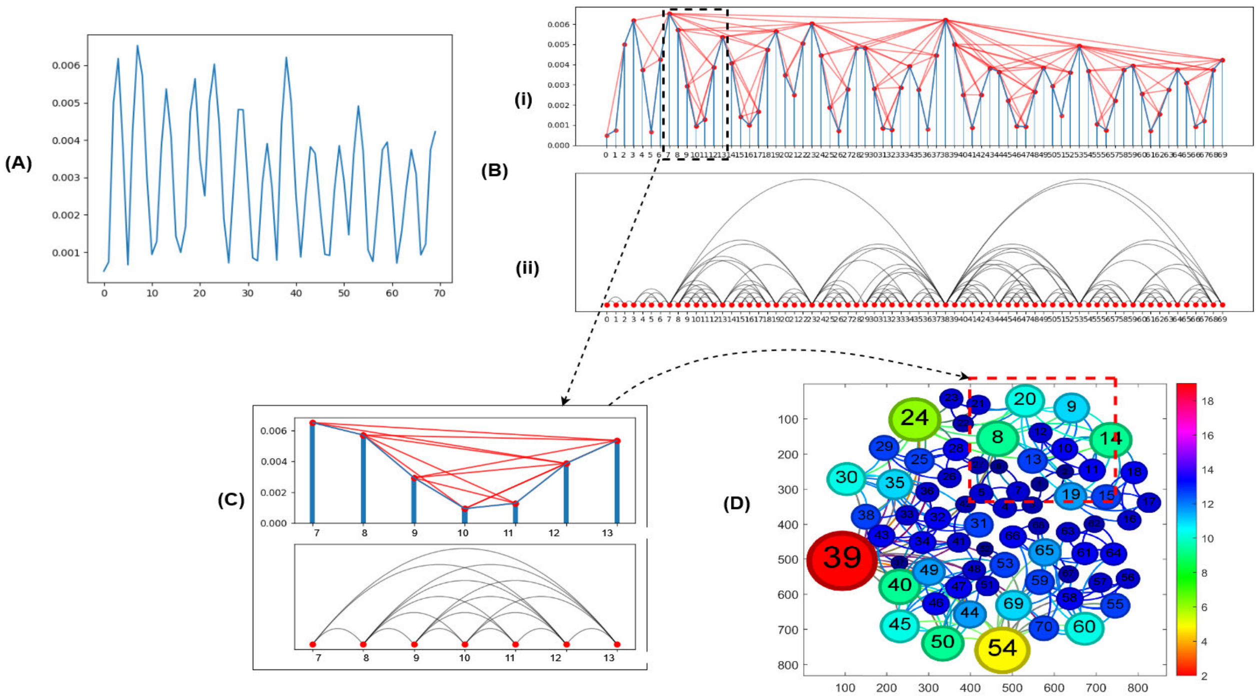


FIGURE 5. Illustration of transformation of time series data into a visibility Graph. (A) A sample time series data. The visibility condition of time series data points is represented as vertical bars in Figure 5-B(i) and the data points in time series data is converted into node which is represented as red dots and how each nodes are connected accordingly to the visibility condition is shown in Figure 5-B(ii). (c) Enlarged view of a segment of signal (which is shown in black dashed box in 5-B) and its corresponding graph is highlighted in red dashed box in 5-D. (D) The Visibility Network.

shows the illustration of the transformation of time series data into a visibility Graph.

E. COMPLEX NETWORK FEATURE EXTRACTION FROM VISIBILITY GRAPH

The adjacency matrix $A_{i,j}$ which is derived in accordance to the visibility condition mentioned in subsection III-D is used to extract relevant network features. Here, we extract 13 network features from visibility network corresponding to three stages of PwPD. This includes nodes of the network (N), edges of the network (E), characteristic path length (CPL), average clustering coefficient ($AvgCC$), diameter, degree centrality, mean betweenness centrality [61], [62], and other statistical measures such as median, kurtosis, skewness pertaining to degree centrality and betweenness centrality. Further details about the network measures can be found in Appendix B.

IV. EXPERIMENTS RESULTS AND DISCUSSION

In the present work, we perform the following major experiments to evaluate the performance of the proposed methodology.

1) Characterization of the three different Stages of PD patient using visibility network
2) Performance evaluation and comparison of the proposed multimodal approach with state-of-the-art method
3) Ablation studies to understand contribution of different components of the proposed methodology. This include an examination of:
    a) The classifier
    b) The modality
    c) Feature Normalization
    d) Age difference of patients
    e) Feature selection
    f) Performance of Deep Neural Network (DNN) on the proposed features

A. CHARACTERIZATION OF THREE DIFFERENT STAGES OF PD USING VISIBILITY GRAPH/NETWORK

To illustrate the characterization of the three different stages of PD, we took speech and gait samples of three different PD patients. Precisely, the time series corresponding to speech and gait signals are transformed into a visibility network. Figure 6 and 7, respectively, show the illustration of the transformation of time series of speech and gait signals of PwPD into their corresponding into a visibility network. The samples taken for the analysis have MDS-UPDRS III score as low (20 points), intermediate (31 points), and severe (43 points). Further, the quantitative characterization is done using standard network features such as number of nodes, number of edges, CPL, $AvgCC$, diameter, degree centrality, mean betweenness centrality and it is depicted in Figure 6 (g) and Figure 7 (g).

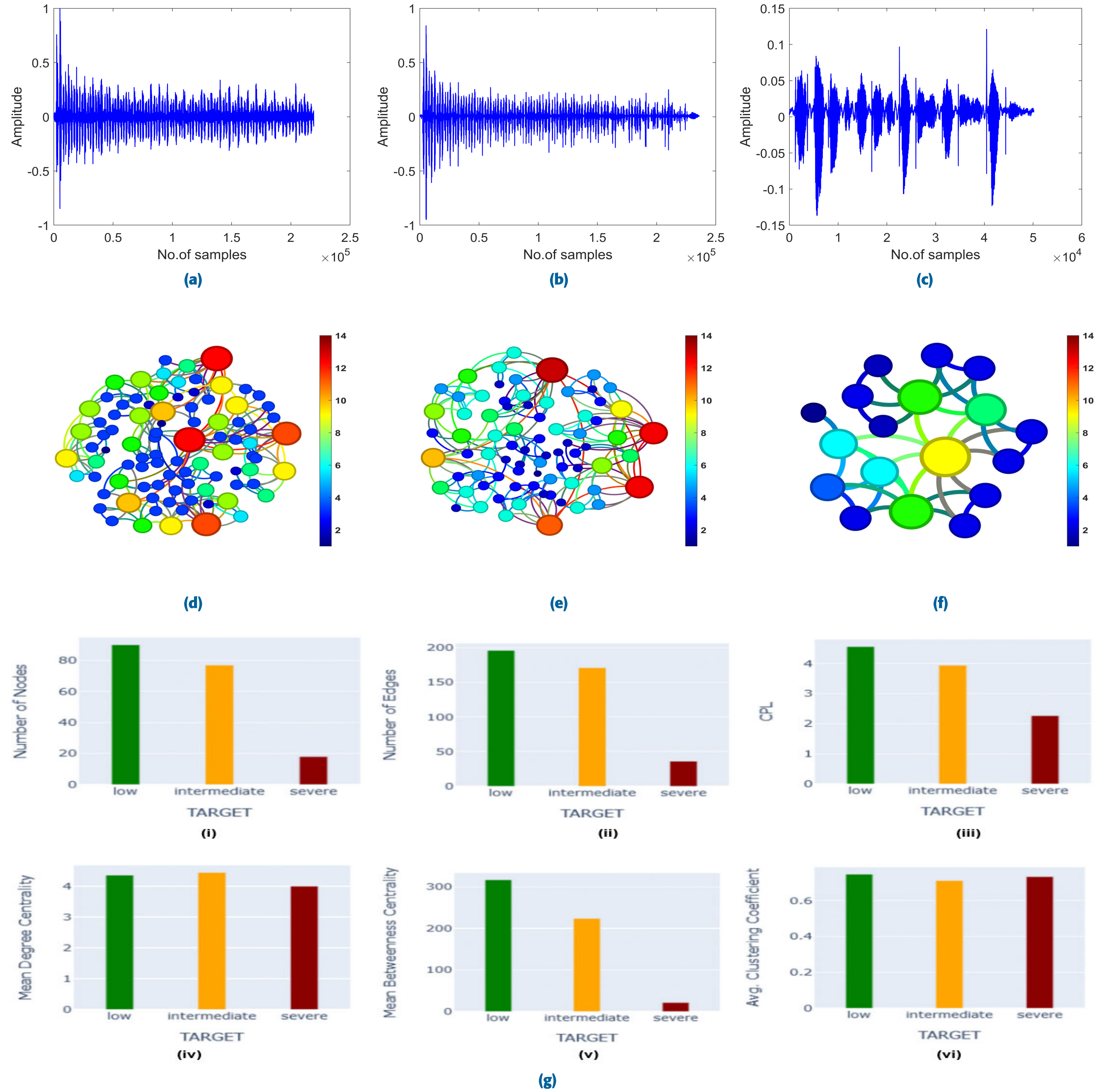


FIGURE 6. Illustration of time series data of speech signal from three different stages of PwPD transformed into visibility network. Figure 6 (a) and (d) refer to the low state of PwPD with MDS-UPDRS III score = 20. Figure 6 (b) and (e) refers to the intermediate state of PwPD with MDS-UPDRS III score = 31. Figure 6 (c) and (f) refer to the severe state of PwPD with MDS-UPDRS III score = 43. The visibility network is plotted using the Gephi software (available at gephi.org). Nodes color and size is fixed based on the degree of node. Figure (g) illustrate the bar plots of the trends of network measures, (i) number of nodes, (ii) number of edges, (iii) *CPL*, (iv) mean degree centrality, (v) mean betweenness centrality and (vi) *AvgCC* incorporated in the analysis of visibility graph for three stages of PD.

In the case speech signal, it is found that the *CPL* measure has a higher value in low state when compared to that of the intermediate and severe state. This is because, although there are many nodes and edges corresponding to the network of low state of PD, most of the nodes have a lower degree. Consequently, it takes more number of steps to reach from one node to another node of the network. Nevertheless, as there are a few nodes with many connections in network corresponding to low state, the mean degree centrality and betweenness centrality is observed to be high. In contrast, we observe that there are more number of hub nodes which is colored in yellow, green and cyan (refer Figure 6 (f)) in the network corresponding to the severe state. Hence, it leads to lower CPL. The significant loss of nodes and connections in the network corresponding to severe state results in fewer critical pathways. This has resulted in reduced

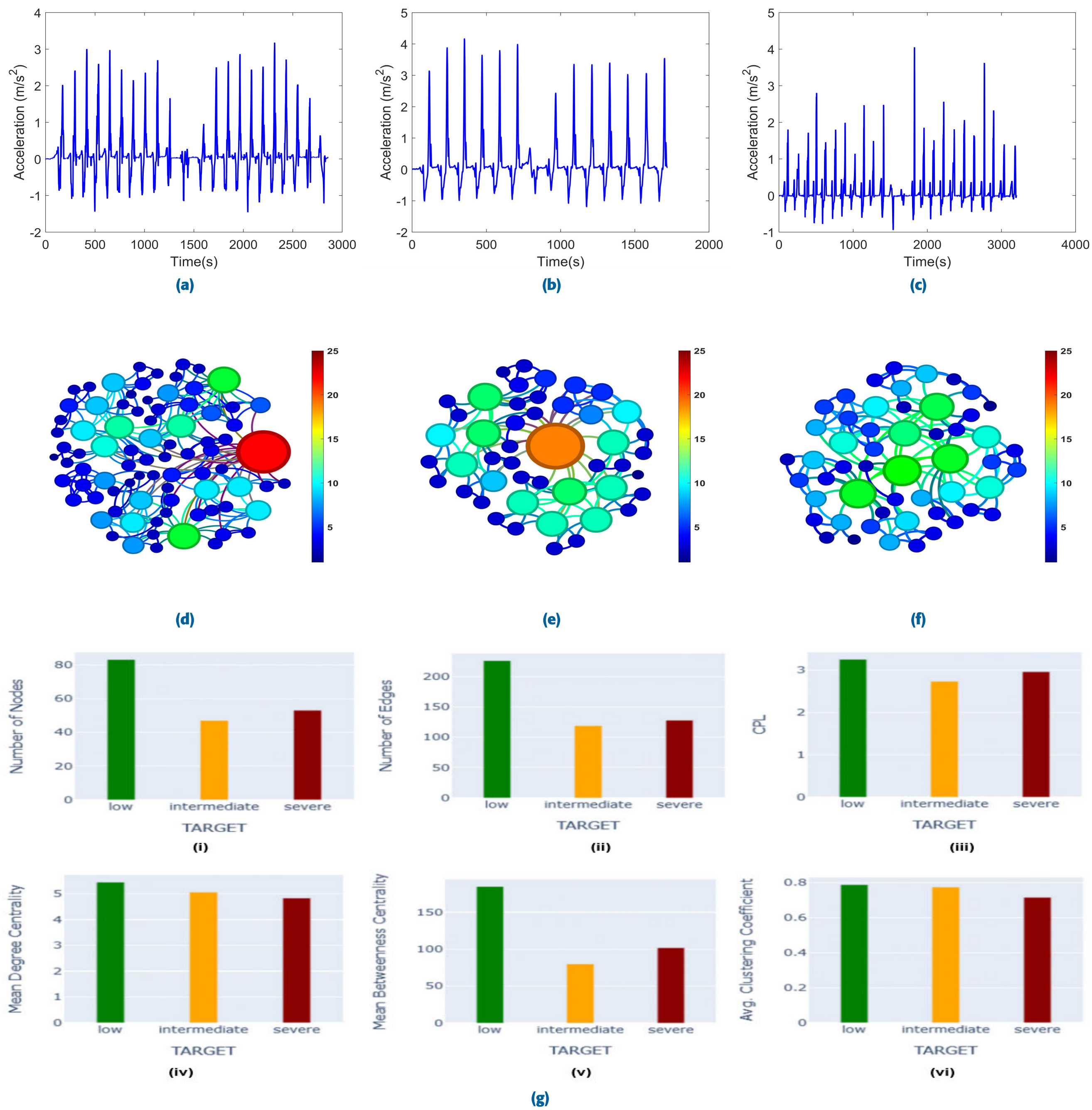


FIGURE 7. Illustration of time series data of Gait signal from three different stages of PwPD transformed into visibility network. Figure 7 (a) and (d) refer to the low state of PwPD with MDS-UPDRS III score = 20. Figure 7 (b) and (e) refers to the intermediate state of PwPD with MDS-UPDRS III score = 31. Figure 7 (c) and (f) refer to the severe state of PwPD with MDS-UPDRS III score = 43. The complex network is plotted using the Gephi software. Nodes color and size is fixed based on their degree of the node. Figure (g) illustrate the bar plots of the trends of network measures (i) number of nodes, (ii) number of edges, (iii) *CPL*, (iv) mean degree centrality, (v) mean betweenness centrality and (vi) *AvgCC* incorporated in the analysis of visibility graph for three stages of PD.

mean betweenness centrality and mean degree centrality in severe state.

Alike CPL, the *AvgCC* also show a similar trend between low and severe state of PD. Precisely, network corresponding to the low state PD show a higher value of *AvgCC*. However, intermediate state doesn't show a consistent behavior as in terms of its network properties. Hence, we computed the statistical variants such mean, median, skewness and kurtosis values of degree centrality and betweenness centrality to infer more variations within those standard measures.

The visibility network measures of gait signal from three different stages of PwPD shows a similar behaviour that found in speech biomarkers. The observed trend in the visibility network underscores the dynamic nature of PD progression, where the network features provide valuable insights for classifying the stages in PwPD.

B. PERFORMANCE EVALUATION OF THE PROPOSED MULTIMODAL APPROACH

For performance evaluation, the dataset is divided into training and testing set in 80:20 ratio and are stratified to get equal number of samples from three classes to have a balanced partition of data. Among the total 38 samples of PwPD, 30 samples are in the training set and the remaining 8 samples belong to the test set. The training dataset is further partitioned into a train-validation set, where the validation set is utilized for obtaining optimal hyper-parameters of the learning algorithm. Random Forest classification model considered in this work is validated and optimized, following a stratified 10-fold cross validation (CV) strategy. Range of hyper-parameters used for the optimization of classification model using grid search CV is shown in Table 14 of Appendix C.

It is ensured that the samples belonging to the same class are always in the same partition and they are not assorted in the train-test dataset. Thus, the optimal hyper-parameters obtained after the grid search CV are used to evaluate the test samples which gives valid results for the classification. The experiment is repeated 10 times with different test set to ensure that every patient is test only once.

Firstly in our experiments, the proposed multimodal architecture is investigated in the following feature set configurations.

1) speech and gait left foot data, with the feature set size of 91 features × 38 samples.
2) speech and gait right foot data, with the feature set size of 91 features × 38 samples.
3) speech and gait data from left and right foots with the feature set size of 169 features × 38 samples.

Table 2 shows the performance of the proposed approach on speech and gait left foot data, speech and gait right foot data, speech and gait both foot data, respectively. From the table, it is observed that the best results are obtained on the speech and gait left foot sensor data with an average accuracy of 85.3 %.

TABLE 2. Classification of PD patients in Three Stages of the Disease Using Multimodal Data. Test Acc.: Accuracy for the Test Data, Validation Acc.: Accuracy for the Validation Data. The result is shown as average accuracy ± standard deviation.

Multimodal Data	Test Acc.	Validation Acc.	F1 score
Speech+ Gait Left Foot	**85.3 ± 14.3**	**85.7 ± 3.9**	**82.9 ± 17.7**
Speech+ Gait Right Foot	65.9 ± 12.5	64.2 ± 5.1	63.6 ± 13.9
Speech+ Gait Both Foot	83.9 ± 12.0	83.7 ± 3.1	79.4 ± 16.7

Further, we compare the proposed method with the state-of-the-art method proposed by Vásquez-Correa et al. [6]. In this work, the authors try to investigate the transition regions of speech utterance and gait movements to model the difficulties of the patients to start/stop movement. Here, for the speech task, the transition from unvoiced to voiced segments (onset) are made according to the fundamental frequency f_0. Once the borders of the onset regions are detected, 80 ms of the signal to the left and right of the borders is extracted to form a chunk of signals with 160 ms of duration. Further, for gait modality (2 × 10 meter walking), the transition appears when the patient starts walking (onset). Similar to speech, a chunk of signal is extracted once the onset borders are detected. A frame of 3 s is considered to each side of the border while forming a signal of 6 s onset region. For both these modalities, the chunk of signals is transformed into a time-frequency representation using short time Fourier transform (STFT). The 2D spectrogram images of both speech and gait are considered as input to the convolutional neural network (CNN) model. The architecture and range of hyper-parameters used to train the CNN model is given in Appendix D. Table 3 shows the comparison results of our proposed model and the state-of-the-art method for multi-class classification of PD severity. As [6] used Unweighted average recall (UAR) score as the evaluation metric, we report the performance in terms of the same here, besides accuracy. From this it is evident that our proposed methodology gives better results than the CNN based model.

TABLE 3. Comparison of the of our proposed multi-class classification model with the state-of-the-art method. Test Acc.: Accuracy for the Test Data, Val Acc.: Accuracy for the Validation Data. Unweighted average recall (UAR) score. The result is shown as mean average accuracy ± standard deviation.

Classifier	Test Acc.	Val Acc.	UAR score
Speech + Gait (Both) (Proposed model)	83.9 ± 12.0	83.7 ± 3.1	82.3 ± 14.3
Speech + Gait (Left) (Proposed model)	**85.3 ± 14.3**	**85.7 ± 3.9**	**85.0 ± 14.6**
Speech + Gait (Right) (Proposed model)	65.9 ± 12.5	64.2 ± 5.1	68.6 ± 12.7
Vásquez-Correa et al (Multimodal data based on CNN model)	37.1 ± 10.2	43.6 ± 7.9	33.9 ± 4.7

A feature importance plot is created using Shapely Additive Explanations (SHAP) [63] to understand the global feature importance towards the model's prediction. Figure 8 shows the ranking of dominated features in predicting the classes of PD of our proposed model. We find that the kurtosis of betweenness centrality of gyroscope left foot sensor data is the crucial feature for predicting the three PD class. The other ranked features have less significance for predicting the classes compared to the kurtosis of betweenness centrality. This plot is demonstrated for a single instance of test data. Additionally, we used SHapley Additive exPlanations for Advanced Data Analysis (SHAPASH) tool for improving the interpretability of the model's predictions. Figure 11 shows the local feature importance in three classes of PD. Results of the study are provided in Appendix E.

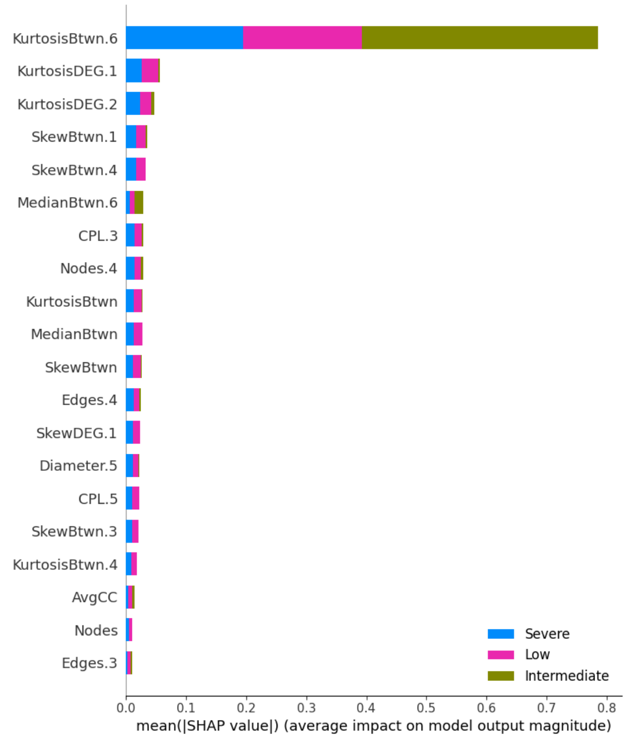


FIGURE 8. SHAP showing the global feature importance of the proposed model to predict the three PD classes for the multimodal (speech and gait left foot) data.

C. ABLATION STUDIES

In the present work, we have performed several ablation studies to understand contribution of different components of the proposed methodology. The details of each study are discussed below.

1) THE CLASSIFIER

Here, we have experimented with different classifiers such as Logistic Regression [64], Naive Bayes [65], KNN [66], Decision Tree [67], Ada Boost [68], XGBoost [69], SVM-rbf [70], as alternatives to the Random Forest classifier used in the proposed multimodal approach. The comparison results of our proposed modal performance with other machine learning classifiers is shown in Table 4, 5, 6 respectively for speech and gait left foot data, speech and gait right foot data and speech and gait data from both foots.

2) THE MODALITY

In this experiment, we compared the performance of the proposed multimodal approach with its unimodal counterparts. The network feature set size of speech signal is 13 features $\times$ 38 samples. Further, the feature set size when considering the individual foots, left and right, is 78 features $\times$ 38 samples. The combined gait (left and right foots) has the feature set configuration of 156 features $\times$ 38 samples. Table 7 shows the performance result of the individual modality for the multi-class classification of PD severity using the proposed network features and RF classifier. From the results, it is evident that the proposed multimodal approach shows almost 3% improvement in performance in terms of accuracy when compared with the best performing unimodal configuration.

TABLE 4. Performance evaluation of the proposed methodology for various classifiers in the classification of PD patients into three Stages of the Disease Using Multimodal Data (Speech and Gait Left Foot Sensor Data). Test Acc.: Accuracy for the Test Data, Validation Acc.: Accuracy for the Validation Data. The result is shown as mean average accuracy ± standard deviation.

Classifier	Test Acc.	Validation Acc.	F1 score
Random Forest	**85.3** $\pm$ **14.3**	**85.7** $\pm$ **3.9**	**82.9** $\pm$ **17.7**
Logistic Regression	56.3 $\pm$ 15.6	65.0 $\pm$ 4.7	54.4 $\pm$ 17.3
Naive Bayes	66.1 $\pm$ 13.5	65.6 $\pm$ 7.3	62.8 $\pm$ 15.5
Decision Tree	44.8 $\pm$ 22.1	58.2 $\pm$ 7.3	44.4 $\pm$ 22.8
KNN	51.3 $\pm$ 14.5	55.5 $\pm$ 3.6	49.20 $\pm$ 16.1
AdaBoost	81.3 $\pm$ 13.7	84.0 $\pm$ 3.8	77.9 $\pm$ 17.2
XGBoost	82.5 $\pm$ 18.0	87.0 $\pm$ 5.6	80.1 $\pm$ 19.9
SVM-rbf	37.6 $\pm$ 13.3	54.4 $\pm$ 7.9	33.7 $\pm$ 14.4

TABLE 5. Performance evaluation of the proposed methodology for various classifiers in the classification of PD patients into three Stages of the Disease Using Multimodal Data (Speech and Gait Right Foot Sensor Data). Test Acc.: Accuracy for the Test Data, Validation Acc.: Accuracy for the Validation Data. The result is shown as mean average accuracy ± standard deviation.

Classifier	Test Acc.	Validation Acc.	F1 score
Random Forest	**65.9** $\pm$ **12.5**	**64.2** $\pm$ **5.1**	**63.6** $\pm$ **13.9**
Logistic Regression	52.8 $\pm$ 11.2	52.4 $\pm$ 7.7	46.7 $\pm$ 14.2
Naive Bayes	37.0 $\pm$ 12.7	46.1 $\pm$ 15.2	32.2 $\pm$ 13.5
Decision Tree	49.5 $\pm$ 19.4	53.0 $\pm$ 6.8	46.5 $\pm$ 18.6
KNN	54.7 $\pm$ 14.8	60.5 $\pm$ 6.2	52.4 $\pm$ 14.7
AdaBoost	40.2 $\pm$ 16.1	44.1 $\pm$ 5.3	31.4 $\pm$ 12.5
XGBoost	57.7 $\pm$ 10.3	60.9 $\pm$ 5.5	54.3 $\pm$ 11.9
SVM-rbf	60.9 $\pm$ 12.3	59.3 $\pm$ 3.7	58.1 $\pm$ 15.3

TABLE 6. Performance evaluation of the proposed methodology for various classifiers in the classification of PD patients into three Stages of the Disease Using Multimodal Data (Speech and Gait Both Foot Sensor Data). Test Acc.: Accuracy for the Test Data, Validation Acc.: Accuracy for the Validation Data. The result is shown as mean average accuracy ± standard deviation.

Classifier	Test Acc.	Validation Acc.	F1 score
Random Forest	**83.9** $\pm$ **12.0**	**83.7** $\pm$ **3.1**	**79.4** $\pm$ **16.7**
Logistic Regression	62.4 $\pm$ 15.7	57.0 $\pm$ 10.9	58.6 $\pm$17.4
Naive Bayes	53.8 $\pm$ 19.5	60.8 $\pm$ 9.3	50.1 $\pm$ 19.9
Decision Tree	51.1 $\pm$ 20.6	69.3 $\pm$ 5.1	47.0 $\pm$ 19.2
KNN	52.6 $\pm$ 18.4	58.3 $\pm$ 5.9	51.0 $\pm$ 19.5
AdaBoost	82.6 $\pm$ 10.6	80.9 $\pm$ 6.3	77.8 $\pm$ 15.2
XGBoost	80.1 $\pm$ 19.7	87.1 $\pm$ 7.1	77.2 $\pm$ 22.8
SVM-rbf	41.4 $\pm$ 13.0	53.1 $\pm$ 6.0	35.9 $\pm$ 16.7

TABLE 7. Performance of speech and gait modality individually in the classification of PD patients into three stages of the disease. Test Acc.: Accuracy for the Test Data, Validation Acc.: Accuracy for the Validation Data. The result is shown as mean accuracy ± standard deviation.

Modality	Test Acc.	Validation Acc.	F1 score
Speech	52.6 $\pm$ 15.1	55.4 $\pm$ 5.8	47.1 $\pm$ 17.2
Gait (Left)	82.6 $\pm$ 13.7	85.6 $\pm$ 3.7	80.6 $\pm$ 15.6
Gait (Right)	66.1 $\pm$ 13.5	61.4 $\pm$7. 9	62.8 $\pm$ 16.9
Gait (Left + Right)	82.6 $\pm$ 12.2	86.3 $\pm$ 2.2	79.1 $\pm$ 16.2

Also, it is clearly evident while comparing the accuracy results of individual and combination of both gait foot data, the contribution of the left foot sensor data is more than the right foot sensor data. One reason considered here is that, most of the participants in this study are right handed and hence it shows the presence of contra-laterality effect in them, as explained in the Previc's neurodevelopmental theory [71]. The theory says that postural support on the left side of the body develops before voluntary motor control on the opposite (right) side and hence the left foot dominates in the task given [72].

3) FEATURE NORMALIZATION

In this experiment, we used the standardscaler method provided by scikit-learn library, which removes the mean and scale the features to unit variance. The experiment was conducted on the best multimodal configuration [Speech + Gait (Left)], which together had 91 network features. Table 8 shows the results obtained with normalization and without normalization of the extracted network features. It is observed that the feature normalization did not provide a significant improvement over the unscaled network features.

TABLE 8. Performance of the proposed method with and without normalizing the network feature of multimodal data. Test Acc.: Accuracy for the Test Data, Validation Acc.: Accuracy for the Validation Data. The result is shown as mean average accuracy ± standard deviation.

	Test Acc.	Validation Acc.	F1 score
With Normalization	81.4 ± 14.8	85.7 ± 3.9	79.8 ± 15.5
Without Normalization	85.3 ± 14.3	85.7 ± 3.9	82.9 ± 17.7

4) AGE DIFFERENCE OF PATIENTS

A notable difference is observed in the age group of male and female patients in severe state of PD. Therefore, we aimed to find whether there is a statistical difference in the network features of patients in severe group. To verify this, we evaluated the statistical difference in the network features with respect to gender using Kruskal-Wallis H-test [73]. This is a non-parametric method for testing whether the data is drawn from the same distribution. In this experiment, we evaluated the null hypothesis that the medians of the network features of the male and female patients in severe group are equal. The results are shown in Table 9. Among the network features of the speech signal, null-hypothesis is satisfied for almost all features, except the average clustering coefficient (p-value < 0.05). For the gait signal, none of the network features shows a significant difference between the male and female group of patients in severe state of PD. Hence, the network features are observed to be independent of age difference in male and female of severe group.

TABLE 9. Kruskal-Wallis H-Test to evaluate the statistical difference in network features of male and female PD patients in the severe group for speech and gait data.

Network features	Speech p-value	Gait p-value
Nodes	0.94	0.46
Edges	0.94	0.46
Mean Degree Centrality	0.17	0.22
Median Degree Centrality	0.44	0.80
Kurtosis Degree Centrality	0.94	0.81
Skewness Degree Centrality	0.94	0.68
Mean Betweenness Centrality	0.94	0.46
Median Betweenness Centrality	0.16	0.56
Kurtosis Betweenness Centrality	0.17	0.81
Skewness Betweenness Centrality	0.22	0.94
CPL	0.81	0.17
Diameter	0.19	0.27
Avg Clustering coefficient	$< \mathbf{0.05}$	0.29

5) FEATURE SELECTION

In this experiment, we investigate the performance of the model if only top 5-10 features are selected according to feature importance from the set of 91 features of speech and gait left signals (the configuration which provide best results in our study). The selection of features is implemented using the SelectKBest algorithm, a library provided by scikit-learn. This is one of the effective feature selection method for the optimal performance of the model. It relies on statistical measures to score and rank the features based on their relationship with the target variable. As aforesaid, the feature set size of this multimodal data is 91 features $\times$ 38 samples. Table 10 shows the performance of the proposed model with top 5-10 features selected.

TABLE 10. Ablation study on the effect of number of features in our proposed model in multi-class classification using multimodal data Test Acc.: Accuracy for the Test Data, Validation Acc.: Accuracy for the Validation Data. The result is shown as mean average accuracy ± standard deviation.

Number of features	Test Acc.	Validation Acc.	F1 score
Top 5 Features	50.0 ± 10.7	54.0 ± 8.7	50.7 ± 13.6
Top 6 Features	51.3 ± 12.7	56.0 ± 7.6	52.3 ± 13.3
Top 7 Features	47.6 ± 18.3	54.7 ± 6.3	48.1 ± 20.7
Top 8 Features	42.8 ± 8.0	54.6 ± 7.2	43.4 ± 9.4
Top 9 Features	41.4 ± 13.6	53.5 ± 7.3	41.9 ± 16.1
Top 10 Features	46.3 ± 11.0	55.3 ± 6.7	48.7 ± 11.3

6) PERFORMANCE OF DEEP NEURAL NETWORK (DNN) ON THE PROPOSED FEATURES

In addition to the machine learning classifier, we conducted the experiment using DNN model for classification. The DNN architecture is a sequential model with the input layer matching to the number of input features. The input to the model is the network features of size 1×91 considering the speech and gait left foot signal data. Here, we consider three hidden layers with 64, 32, and 16 neurons, respectively, followed by a dropout layer. The output layer consists of dense layer with three neurons representing three classes of PD severity (low, intermediate and severe state). The model is compiled with the sparse categorical cross-entropy loss, with an Adam optimizer, and a learning rate of 0.01 to update the parameters and to train it over 200 epochs. The number of samples in the training and testing set are in the ratio

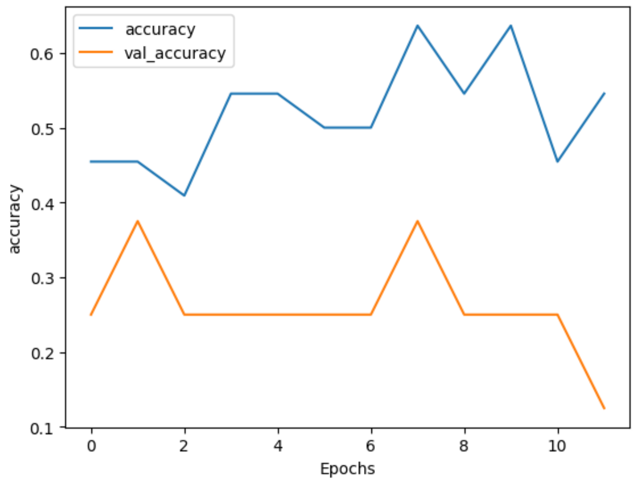


FIGURE 9. Training accuracy plot of DNN model.

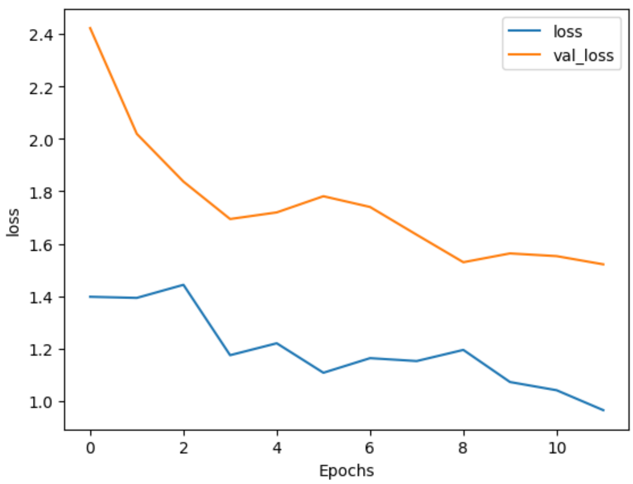


FIGURE 10. Training loss plot of DNN model.

of 80:20. Further, the training data is divided into training and validation set which is in the ratio of 75:25. After the model being trained on the training and the validation set, the training and loss plots shown in Figure 9 and 10. The learning curve shows a clear divergence between the training and validation loss it is evident that the model is over-fitting the data even after applying regularization techniques and early stopping. Furthermore, in comparison to our relatively small dataset the model led to poor generalization of data. Hence, the DNN model is not suitable for our dataset.

V. CONCLUSION AND FUTURE DIRECTIONS

In this research, we present a multi-modal assessment of Parkinson's Disease from a complex systems perspective with the aim to classifying the motor state severity of PwPD into three different stages (low, intermediate and severe) in accordance to the MDS-UPDRS III score. In PwPD there occurs natural delay in their motor movements due to the impairments in different biomarkers over the PD progression. We investigated two such biomarkers; speech and gait time series data and their combination as multimodal data. Precisely, the visibility network derived for the different modalities gave inferences about the structural differences in terms of nodes and connections (edges). Further, the network properties extracted from individual modalities are fused at feature-level to perform the multimodal analysis of PD. From the experimental results shown for the multi-class classification, the proposed model outperformed the state-of-the art method. Achieving a mean accuracy of 85.3%, the multimodal configuration (speech and gait left foot signals) also outperformed the unimodal counterparts, which showed mean accuracy of 52.6% on speech, 82.6% on gait left foot, 66.1% on gait right foot and 82.6% on gait both feet signals. This confirms that among the different configuration of multimodal data, the combination of speech and gait left foot signals provide robustness in classifying the different stages of PD. This can be further substantiated from the point of view that the participants in this study is mostly right handed which indicate the presence of contra-laterality effect. Furthermore, we have analyzed the significance of network features in order to quantify the impact of each feature on the model's predictions across the three different stages of PD.

The major limitation of our perceptive outlined in this study is the definition of node connections based on peaks from the time-series in visibility graph/network is artificial from a dynamic perspective. More precisely, the node connections are not motivated from the physics of speech productions or gait movements. Therefore, it is more advantageous to use network transformation focusing on physics of the system from which the time-series is generated, such as the recurrence network approach. Further investigations, perhaps considering global features, will be addressed to tackle this problem.

APPENDIX A
SUMMARY OF MULTI-CLASS CLASSIFICATION OF PD SEVERITY

In this section, we have shown a table summary of the section II-A: Speech in Table 11, section II-B: Gait in Table 12 and section II-C: Multi-modal data in Table 13.

APPENDIX B
COMPLEX NETWORK FEATURES AND THEIR STATISTICAL MEASURES

A. NODE OF THE NETWORK

A node (v) is a fundamental unit which represents an entity or a point in the graph or network.

B. EDGE OF THE NETWORK

An edge (e) is a connection between two nodes in a graph or network which represents the relationship or interaction between them.

C. CHARACTERISTIC PATH LENGTH

The length of the shortest path $L_{i,j}$ between a pair of nodes i and j of a graph is an important measure which tells how fast

TABLE 11. Summary of multi-class classification of PD into different stages of severity from speech data using DDK exercises according to MDS-UPDRS III score.

Study	Dataset Used	Features/ Representation	Classifier	PD Stage	Outcomes
Bocklet et al. [32]	176 German native speakers 88 PD, 88 HC	Articulation, Phonation, Prosody	SVM with Linear kernal & Sequential Minimal Optimization	Assess PD severity in UPDRS-III; U1, U2, U3	UAR: 59.1%
Bayestehtashk et al. [33]	English native speakers; 168 PD	Phonation, DDK & reading tasks	Ridge Regression Algorithm	Assess PD severity in UPDRS-III; clinic 1, clinic 2, clinic 3	MAE: 5.5
Grósz et al. [34]	Colombian, Spanish native speakers 50 PD, 50 HC	Pearson's correlation coefficient	DNN-based transfer learning	Assess clinical score;MDS-UPDRS-III	Spearman's correlation: 0.649
Orozco-Arroyave et al. [35]	Spanish, German Czech native speakers; 158 PD	Articulation, Intelligibility	SVR Regression	Assess neurological PD state	Spearman correlation up to 0.72
Garcia, N. et al. [36]	PC GITA 50 PD, 50 HC	Articulation, Phonation, Prosody	Cosine distance between test speaker i-vector & reference i-vector	Assess neurological PD state	For Phon./Art. Spearman's correlation: -0.48
Ramezani et al. [37]	Colombian, Spanish native speakers 50 PD, 50 HC	Acoustic features from openSMILE toolkit	Gaussian Mixture Regression & SVR	Assess PD severity	Spearman's correlation up to 0.52
Benmalek et al. [38]	PVA dataset 55 HC, 178 early, 118 intermediate, 24 severe	Dysphonia measures	LLBFS algorithm with subspace discriminant	Healthy, Early, intermediate, severe	Accuracy: 96.5%
Benmalek et al. [40]	PVA dataset 55 HC, 178 early, 118 intermediate, 24 severe	Ceptral coefficients of MFCC	LLBFS algorithm with discriminant analysis	Healthy, Early, intermediate, severe	Accuracy: 87.6%
Benmalek et al. [39]	PVA dataset 55 HC, 178 early, 118 intermediate, 24 severe	Dysphonia measures	Linear SVM classifier	Healthy, Early, intermediate, severe	Accuracy: 92.5%
Arias-Vergara et al. [41]	Colombian spanish native speakers 68 PD, 50 HC	Prosody	SVM classifier	MDS-UPDRS-III (3.1) Score 0 : Healthy Score 1: class 1 Score $\geq$ 2: class 2	Accuracy: 58% UAR: 51% Cohen's κ: 0.31
Vásquez-Correa et al. [6]	Colombian Spanish native speakers; 44 PD, 40 HC	Spectrograms	CNN	Healthy, Low, Intermediate, Severe	Speech onset: UAR: 37.8%
Arias-Londoño et al. [42]	Neurovoz, GITA dataset	Perturbation, Complexity Spectral and cepstral	CNN	Control, Mildly affected Affected	AMAE: 64% BACC: 52%
Bhardwaj et al. [43]	Parkinson's telemonitoring data 188 PD, 64 HC	HNR, DFA, PDE & RPDE	Decision tree classifier	Normal, Slight, Mild,Moderate	Accuracy: 98.63%
Karan et al. [44]	PC GITA 50 PD, 50 HC	Time-frequency features	Random Forest classifier	Mild, Moderate, Severe	/petaka/ UAR: 48.8%
Kodali M et al. [45]	PC-GITA, 50 PD, 50 HC	SFFCC & MFCC-SFF	SVM classifier	Healthy, Mild, Severe	Accuracy: For SFFCC: 51.5 $\pm$ 4 For MFCC-SFF: 49.8 $\pm$ 5
Vásquez Correa et al. [46]	Colombian, Spanish native speakers; 106 PD, 87 HC	Phonation, Prosody, Articulation, OpenSMILE, Phonological, RAE	SVM classifier	Healthy, Low, Intermediate, Severe	For Phonation: UAR: 49.1%

the information is transferred in a network. The characteristic path length (CPL) is defined as the average of all the shortest path lengths between the pair of nodes in a network. For the disconnected nodes, the length of the shortest

TABLE 12. Summary of multi-class classification of PD into different stages of severity from gait data using inertial sensors attached to the shoes. Task considered is n-meter walk and motor score according to MDS-UPDRS III score.

Study	Dataset Used	Features/ Representation	Classifier	PD Stage	Outcomes
Klucken et al. [47]	eGaIT system 42 PD, 39 HC	Spectral & Statistical Features	SVM with Linear Kernel	PwPD (low, mild & high)	Accuracy up to 89%
Vásquez-Correa et al. [6]	Multimodal corpus 44 PD, 40 HC	Spectrograms	2D CNN	Healthy & PwPD (low, intermediate, severe)	Gait onset: Accuracy: 48.6%
Pérez-Toro et al. [74]	eGaIT system 45 PD, 45 HC	NLD features	Random Forest classifier	Healthy & PwPD (low, intermediate, severe)	Accuracy: 64.4% Cohen's kappa, κ: 0.41
Vasquez-Correa et al. [49]	Multimodal corpus 106 PD, 105 HC	Raw gait signals	1D CNN	PwPD (mild, intermediate, severe)	UAR up to 64.9% F1 score up to 0.632

TABLE 13. Summary of multi-class classification of PD into different stages of disease severity from multimodal data and according to MDS-UPDRS III score.

Study	Modalities	Dataset Used	Features/ Representation	Classifier	PD Stage	Outcomes
Oung et al. [50]	Speech & motion data signals	50 PD, 15 HC	Wavelet energy & Entropy-based	ELM classifier	Healthy & PwPD (mild, moderate, severe)	Accuracy: 95.93%
Vásquez-Correa et al. [54]	Speech, Gait & Handwriting	PC-GITA 50 PD, 50 HC	Articulation, Prosody, Kinematics, Bio-mechanical	SVR with a linear kernel	Neurological state as per UPDRS score	Spearman correlation: /pa-ta-ka/ : 0.46
Vásquez-Correa et al. [6]	Speech, Gait & Handwriting	Multimodal corpus 44 PD, 45 HC	Spectrograms	2D CNN	Healthy & PwPD (low, intermediate, severe)	onset UAR: 55.6%
Wan S et al. [53]	Speech & movement patterns	UCI Dataset 20 PD, 20 HC	Accelerometer signals obtained obtained using using smartphone	DMLP classifier	Assess PD severity stage	Accuracy up to 97%
García N et al. [51]	Speech, gait & Handwriting	49 PD, 41 HC	i-vectors	SVM with Gaussian kernal	Assess PD severity stage	Spearman's correlation: young HC: 0.31, elderly HC: 0.20, PD: -0.41
Vásquez-Correa et al. [52]	Speech, gait & Handwriting	PC-GITA corpus 106 PD,87 HC	GMM-UBM and i-vectors	Linear regression	Assess PD severity stage	Spearman's correlation: 0.634, MAE: 10.5

path is ∞.

$$CPL = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} L_{i,j} \tag{1}$$

where N is the number of nodes in the network.

D. DIAMETER

Among all pair of nodes in a graph network, the maximum value of the shortest path defines the diameter of a network (D).

$$D = max(L_{i,j}) \tag{2}$$

E. CLUSTERING COEFFICIENT

The connectedness between the neighbouring nodes in a network is defined as clustering coefficient (CC) of a node v which is computed as:

$$CC = \frac{2N_v}{K_v(K_v-1)} \tag{3}$$

where N_v represents the density of connectedness between the neighbouring nodes and $K_v(K_v - 1)$ represents the existence of possible edges among them (K_v is the number of neighbours for a node). Here we compute the average of CC ($AvgCC$) of all the nodes in a network where its value ranges between 0 and 1.

$$Avg\ CC = \frac{1}{N} \sum_{v=1}^{N} CC \tag{4}$$

If all the neighbouring nodes are connected to each other, then the value of $CC = 1$.

F. DEGREE CENTRALITY

The degree centrality (C_D) of a node v quantifies the importance of a node in a graph network. It depends on the number of edges connected to it.

$$C_D(v_i) = \sum_j A_{i,j} \tag{5}$$

G. MEAN DEGREE CENTRALITY

To understand the connectivity of overall nodes in a network, the average degree centrality helps to find how well-connected are the nodes in a network. If the value of average degree centrality is high, it implies the nodes are well-connected and low average degree centrality implies less connected nodes in a network.

$$Avg\ C_D = \frac{1}{N}\sum_{i=1}^{N} C_D(v_i) \tag{6}$$

H. MEDIAN DEGREE CENTRALITY

The median degree centrality measures the significant variations in the degree centrality over the outliers and provides more stable values for the nodes with high degree centrality.

I. KURTOSIS DEGREE CENTRALITY

The kurtosis of degree centrality measures the node degree to which the degree centrality distribution of nodes deviates from a normal distribution. Lower the value of kurtosis implies an even distribution of degree centrality of nodes and the heterogeneity in centrality of all nodes in a network implies high kurtosis value.

J. SKEWNESS DEGREE CENTRALITY

The skewness of degree centrality measures the asymmetry of the degree centrality distribution which helps in understanding the structure of the network and the significance of all nodes in a network.

K. BETWEENNESS CENTRALITY

The betweenness centrality $C_B(v_i)$ of a node v measures the number of times a node lies on the shortest path between other nodes in a network. High value of $C_B(v_i)$ for a node implies how it acts between other nodes of the network. Also the significance of the node relies in maintaining the connectivity and efficiency of communication within the network. Contrarily, the removal of that node affects communications between other nodes of a network.

$$C_B(v_i) = \sum_{j<k} p_{j,k}(v_i)/p_{j,k} \tag{7}$$

where $p_{j,k}(v_i)$ is the number of shortest paths between j and k that pass v_i and $p_{j,k}$ is the number of shortest paths between j and k.

L. MEAN BETWEENNESS CENTRALITY

The mean betweenness centrality value provides an insight of overall nodes with strong connectivity for maintaining an effective communication in a network.

M. MEDIAN BETWEENNESS CENTRALITY

From the betweenness centrality distribution, higher the median betweenness centrality value provides the significant number of nodes acts for the effective communication. Conversly, lower the median value means less role in information transfer within the network.

N. KURTOSIS BETWEENNESS CENTRALITY

The kurtosis of the betweenness centrality distribution measures the how the node centrality is distributed around the mean betweenness centrality. High kurtosis value implies high betweenness centrality of the node in the network which maintain connectivity and communication.

O. SKEWNESS BETWEENNESS CENTRALITY

The skewness of betweenness centrality measures the asymmetry of the betweenness centrality distribution. The positive and negative value of skewness results in nodes with high betweenness centrality and low betweenness centrality respectively for identifying the connectivity of nodes in a network.

APPENDIX C HYPER-PARAMTER

The range of hyper-parameters used in the different classifier models are listed in Table 14.

TABLE 14. List of parameters for optimization.

Classifier	Optimized parameter

Random Forest	criterion = ['gini','entropy','log_loss'];
	max_features = ['sqrt','log2', None];
	max_depth = [2,5,10,20,30,50,100];
	n_estimators = [10,20,...90,100]

AdaBoost	n_estimators = [1,5,10,20,30,50,100]
Logistic Regression	C = np.logspace(-3,3,7); solver = ['newton-cg', 'lbfgs', 'liblinear']
Naive Bayes	Not Applicable

Decision Tree	criterion = ['gini', 'entropy'];
	max_depth = [2,5,10,20,30,50,100];
	min_samples_leaf = [5, 10, 20, 50, 100];
	max_features = ['auto', 'sqrt', 'log2'];

KNN	n_neighbors = [3,4,5,6]; metric = ['minkowski','euclidean','manhattan']

XGBoost	nthread = [4]; objective = ['multi:softmax'];
	learning_rate = [0.1, 0.01, 0.05];
	max_depth = range (2, 10, 1);
	n_estimators = range(5, 100, 5);

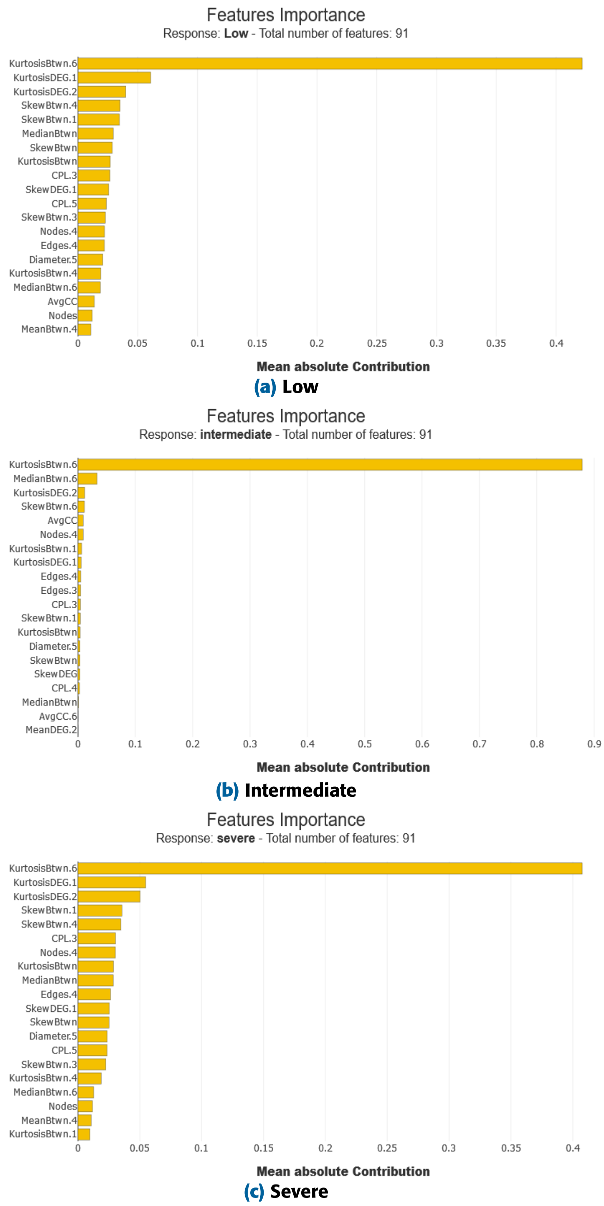
SVM-rbf	C = [0.0001,0.001,0.01,0.1,1,10,100,1000,10000]; kernel = ['rbf']; gamma = [0.0001,0.001,0.01,0.1,1,10,... 100,1000,10000];

APPENDIX D ARCHITECTURE AND HYPER-PARAMETERS OF CNN MODEL

See Table 15.

APPENDIX E SHAPASH: LOCAL FEATURE IMPORTANCE OF THE MODEL

The features are ranked according to their importance in predicting the state of PD by the Random Forest model is



(a) Low
(b) Intermediate
(c) Severe

FIGURE 11. SHAPASH, which shows the local feature importance by a single instance of multimodal (speech and gait left foot signal) data.

TABLE 15. CNN architecture for multi-class classification of PD using multimodal data.

Input (Spectrogram Images of Speech and Gait)
Convolutional layer 1 (ReLU unit)
Convolutional layer 2 (ReLU unit)
Max-pooling layer 1
Dropout 1
Convolutional layer 3 (ReLU unit)
Convolutional layer 4 (ReLU unit)
Max-pooling layer 2
Dropout 2
Fully connected hidden layer (ReLU unit)
Fully connected hidden layer (ReLU unit)
Output layer (Sigmoid unit)

as follows: (1) Low state: kurtosis betweenness centrality of gyroscope Y-axis of gait left foot, kurtosis degree centrality of speech, kurtosis degree centrality of accelerometer X-axis of gait left foot, (2) Intermediate state: kurtosis betweenness centrality of gyroscope Y-axis of gait left foot, median betweenness centrality of gyroscope Y-axis gait left foot, kurtosis degree centrality of accelerometer X-axis of gait left foot, and in (3) Severe state: kurtosis betweenness centrality of gyroscope Y-axis of gait left foot, kurtosis degree centrality of speech, kurtosis degree centrality of accelerometer X-axis of gait left foot.

TABLE 16. Hyper parameters used to train the CNN model.

Hyper parameter	Range of values
Filter size convolutional layers	3,5,7
Depth of convolutional layers	4,8,16,32,64
Hidden units in fully connected layers	16, 32, 64, 128
Learning rate	0.0001, 0.0005, 0.001
Probability of dropout	0.1, 0.2, 0.3,..... 0.9

DECLARATION OF COMPETING INTEREST

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: K. Deepa Raj and G. Jyothish Lal report that the financial support was provided by Amrita Vishwa Vidyapeetham.

DATA AVAILABILITY

The data that support the findings of this study is available upon reasonable request from Juan Rafael Orozco-Arroyave at the GITA laboratory, Faculty of Engineering, Universidad de Antioquía, Medellín, Colombia.

REFERENCES

[1] K. F. Tipton, "Biochemical aspects of drug-induced Parkinson's disease," *Biochem. Soc. Trans.*, vol. 18, no. 2, pp. 129–131, Apr. 1990.
[2] T. Pringsheim, N. Jette, A. Frolkis, D. Lam, and T. Steeves, "The prevalence of Parkinson disease: A systematic review and meta-analysis (P03.067)," *Neurology*, vol. 80, no. 7, pp. 1583–1590, Feb. 2013.
[3] S. Sveinbjornsdottir, "The clinical symptoms of Parkinson's disease," *J. Neurochem.*, vol. 139, pp. 318–324, Oct. 2016.
[4] *Medical News Today Parkinson's Disease Early Signs and Causes*. Accessed: Nov. 30, 2023. [Online]. Available: https://www.medicalnewstoday.com/articles/323396
[5] Nat. Inst. Neurological Disorders Stroke. *Parkinson's Disease*. Accessed: Nov. 30, 2023. [Online]. Available: https://www.ninds.nih.gov/health-information/disorders/parkinsons-disease
[6] J. C. Vásquez-Correa, T. Arias-Vergara, J. R. Orozco-Arroyave, B. Eskofier, J. Klucken, and E. Nöth, "Multimodal assessment of Parkinson's disease: A deep learning approach," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 4, pp. 1618–1630, Jul. 2019.
[7] Q. W. Oung, H. Muthusamy, H. L. Lee, S. N. Basah, S. Yaacob, M. Sarillee, and C. H. Lee, "Technologies for assessment of motor disorders in Parkinson's disease: A review," *Sensors*, vol. 15, no. 9, pp. 21710–21745, 2015.
[8] M. Newman, A.-L. Barabási, and D. J. Watts, *The Structure and Dynamics of Networks*. Princeton, NJ, USA: Princeton Univ. Press, 2011.
[9] M. Mitchell, "Complex systems: Network thinking," *Artif. Intell.*, vol. 170, no. 18, pp. 1194–1212, Dec. 2006.
[10] Z. Zhang, "Mechanics of human voice production and control," *J. Acoust. Soc. Amer.*, vol. 140, no. 4, pp. 2614–2635, Oct. 2016.
[11] G. Jyothish Lal, E. A. Gopalakrishnan, and D. Govind, "Glottal activity detection from the speech signal using multifractal analysis," *Circuits, Syst., Signal Process.*, vol. 39, no. 4, pp. 2118–2150, Apr. 2020.

[12] Y. Zou, R. V. Donner, N. Marwan, J. F. Donges, and J. Kurths, "Complex network approaches to nonlinear time series analysis," *Phys. Rep.*, vol. 787, pp. 1–97, Jan. 2019.

[13] J. Zhang and M. Small, "Complex network from pseudoperiodic time series: Topology versus dynamics," *Phys. Rev. Lett.*, vol. 96, no. 23, Jun. 2006, Art. no. 238701.

[14] X. Xu, J. Zhang, and M. Small, "Superfamily phenomena and motifs of networks induced from time series," *Proc. Nat. Acad. Sci. USA*, vol. 105, no. 50, pp. 19601–19605, Dec. 2008.

[15] G. J. Lal, E. A. Gopalakrishnan, and D. Govind, "A recurrence network approach for characterization and detection of dynamical transitions during human speech production," *Circuits, Syst., Signal Process.*, vol. 41, no. 12, pp. 6975–6998, Dec. 2022.

[16] R. V. Donner, M. Small, J. F. Donges, N. Marwan, Y. Zou, R. Xiang, and J. Kurths, "Recurrence-based time series analysis by means of complex network methods," *Int. J. Bifurcation Chaos*, vol. 21, no. 4, pp. 1019–1046, Apr. 2011.

[17] S. Mao and F. Xiao, "Time series forecasting based on complex network analysis," *IEEE Access*, vol. 7, pp. 40220–40229, 2019.

[18] M. Murugesan and R. Sujith, "Complex network approach for investigating thermoacoustic systems," Dept. Aerosp. Eng., Indian Inst. Technol. Madras, Chennai, India, 2016.

[19] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuño, "From time series to complex networks: The visibility graph," *Proc. Nat. Acad. Sci. USA*, vol. 105, no. 13, pp. 4972–4975, Apr. 2008.

[20] A. M. Nuñez, L. Lacasa, J. P. Gomez, and B. Luque, "Visibility algorithms: A short review," in *New Frontiers Graph Theory*. IntechOpen, Mar. 2012, ch. 6, pp. 119–152.

[21] S. Bhaduri and A. Chakraborty, "Speech emotion quantification with chaos-based modified visibility graph-possible precursor of suicidal tendency," *J. Neurol. Neurosci.*, vol. 7, no. 3, p. 100, 2016.

[22] B. Amin Goncalves and A. P. F. Atman, "Visibility graph combined with information theory: An estimator of stock market efficiency," *J. Netw. Theory Finance*, vol. 3, no. 2, pp. 1–15, 2017.

[23] M. Ahmadlou, H. Adeli, and A. Adeli, "New diagnostic EEG markers of the Alzheimer's disease using visibility graph," *J. Neural Transmiss.*, vol. 117, no. 9, pp. 1099–1109, Sep. 2010.

[24] Z. Zhang, A. Zhang, C. Sun, S. Xiang, and S. Li, "Data-driven analysis of the chaotic characteristics of air traffic flow," *J. Adv. Transp.*, vol. 2020, pp. 1–17, Sep. 2020.

[25] S. Kundu, A. Opris, Y. Yukutake, and T. Hatano, "Extracting correlations in earthquake time series using visibility graph analysis," *Frontiers Phys.*, vol. 9, Apr. 2021, Art. no. 656310.

[26] C. G. Goetz et al., "Movement disorder society-sponsored revision of the unified Parkinson's disease rating scale (MDS-UPDRS): Scale presentation and clinimetric testing results," *Movement Disorders, Off. J. Movement Disorder Soc.*, vol. 23, no. 15, pp. 2129–2170, 2008.

[27] A. Tsanas, M. A. Little, P. E. McSharry, and L. O. Ramig, "Accurate telemonitoring of Parkinson's disease progression by noninvasive speech tests," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 4, pp. 884–893, Apr. 2010.

[28] J. R. Duffy, *Motor Speech Disorders-E-Book: Substrates, Differential Diagnosis, and Management*. Amsterdam, The Netherlands: Elsevier, 2012.

[29] S. Yang, F. Wang, L. Yang, F. Xu, M. Luo, X. Chen, X. Feng, and X. Zou, "The physical significance of acoustic parameters and its clinical significance of dysarthria in Parkinson's disease," *Sci. Rep.*, vol. 10, no. 1, p. 11776, Jul. 2020.

[30] J. R. Orozco-Arroyave, *Analysis of Speech of People With Parkinson's Disease*, vol. 41. Berlin, Germany: Logos Verlag Berlin, 2016.

[31] A. M. García, T. Arias-Vergara, J. C. Vasquez-Correa, E. Nöth, M. Schuster, A. E. Welch, Y. Bocanegra, A. Baena, and J. R. Orozco-Arroyave, "Cognitive determinants of dysarthria in Parkinson's disease: An automated machine learning approach," *Movement Disorders*, vol. 36, no. 12, pp. 2862–2873, Dec. 2021.

[32] T. Bocklet, S. Steidl, E. Nöth, and S. Skodda, "Automatic evaluation of parkinson's speech—Acoustic, prosodic and voice related cues," in *Proc. Interspeech*, Aug. 2013, pp. 1149–1153.

[33] A. Bayestehtashk, M. Asgari, I. Shafran, and J. McNames, "Fully automated assessment of the severity of Parkinson's disease from speech," *Comput. Speech Lang.*, vol. 29, no. 1, pp. 172–185, Jan. 2015.

[34] T. Grósz, R. Busa-Fekete, G. Gosztolya, and L. Tóth, "Assessing the degree of nativeness and Parkinson's condition using Gaussian processes and deep rectifier neural networks," in *Proc. Interspeech*, Sep. 2015, pp. 919–923.

[35] J. R. Orozco-Arroyave, J. C. Vdsquez-Correa, F. Hönig, J. D. Arias-Londoño, J. F. Vargas-Bonilla, S. Skodda, J. Rusz, and E. Noth, "Towards an automatic monitoring of the neurological state of Parkinson's patients from speech," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2016, pp. 6490–6494.

[36] N. Garcia, J. R. Orozco-Arroyave, L. F. D'Haro, N. Dehak, and E. Nöth, "Evaluation of the neurological state of people with Parkinson's disease using I-vectors," in *Proc. Interspeech*, Aug. 2017, pp. 299–303.

[37] H. Ramezani, H. Khaki, E. Erzin, and O. B. Akan, "Speech features for telemonitoring of Parkinson's disease symptoms," *IEEE Eng. Med. Biol. Soc. (EMBC)*, pp. 3801–3805, Jul. 2017, doi: 10.1109/EMBC.2017.8037685.

[38] E. Benmalek, J. Elmhamdi, and A. Jilbab, "Multiclass classification of Parkinson's disease using different classifiers and LLBFS feature selection algorithm," *Int. J. Speech Technol.*, vol. 20, no. 1, pp. 179–184, Mar. 2017.

[39] E. Benmalek, J. Elmhamdi, and Jilbab, "Voice assessments for detecting patients with Parkinson's diseases in different stages," *Int. J. Electr. Comput. Eng. (IJECE)*, vol. 8, no. 6, pp. 4265–4271, 2018.

[40] E. Benmalek, J. Elmhamdi, and A. Jilbab, "Multiclass classification of Parkinson's disease using cepstral analysis," *Int. J. Speech Technol.*, vol. 21, no. 1, pp. 39–49, Mar. 2018.

[41] T. Arias-Vergara, J. C. Vasquez-Correa, J. R. Orozco-Arroyave, P. Klumpp, and E. Nöth, "Unobtrusive monitoring of speech impairments of Parkinson'S disease patients through mobile devices," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 6004–6008.

[42] J. D. Arias-Londoño and J. A. Gómez-García, "Predicting UPDRS scores in Parkinson's disease using voice signals: A deep learning/transfer-learning-based approach," in *Communications in Computer and Information Science*. Cham, Switzerland: Springer, 2020, pp. 100–123.

[43] S. Bhardwaj, D. Arora, B. Devi, V. G. Shankar, and S. Srivastava, "Machine learning assisted binary and multiclass Parkinson's disease detection," in *Intelligent Sustainable Systems*. Cham, Switzerland: Springer, 2022, pp. 191–206.

[44] B. Karan, S. Sekhar Sahu, and J. Rafael Orozco-Arroyave, "An investigation about the relationship between dysarthria level of speech and the neurological state of Parkinson's patients," *Biocybernetics Biomed. Eng.*, vol. 42, no. 2, pp. 710–726, Apr. 2022.

[45] M. Kodali, S. R. Kadiri, and P. Alku, "Automatic classification of the severity level of Parkinson's disease: A comparison of speaking tasks, features, and classifiers," *Comput. Speech Lang.*, vol. 83, Jan. 2024, Art. no. 101548.

[46] J. C. V. Correa, "Multimodal assessment of parkinson's disease patients using information from speech, handwriting, and gait," Dept. Electron. Comput. Eng., Universidad de Antioquia, Medellín, Antioquia, Colombia, 2023.

[47] J. Klucken, J. Barth, P. Kugler, J. Schlachetzki, T. Henze, F. Marxreiter, Z. Kohl, R. Steidl, J. Hornegger, B. Eskofier, and J. Winkler, "Unbiased and mobile gait analysis detects motor impairment in Parkinson's disease," *PLoS One*, vol. 8, no. 2, Feb. 2013, Art. no. e56956.

[48] J. Barth, M. Sünkel, K. Bergner, G. Schickhuber, J. Winkler, J. Klucken, and B. Eskofier, "Combined analysis of sensor data from hand and gait motor function improves automatic recognition of Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Aug. 2012, pp. 5122–5125, doi: 10.1109/EMBC.2012.6347146.

[49] J. C. Vasquez-Correa, T. Arias-Vergara, P. Klumpp, P. A. Perez-Toro, J. R. Orozco-Arroyave, and E. Nöth, "End-2-end modeling of speech and gait from patients with Parkinson's disease: Comparison between high quality Vs. Smartphone data," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 7298–7302.

[50] Q. W. Oung, H. Muthusamy, S. N. Basah, H. Lee, and V. Vijean, "Empirical wavelet transform based features for classification of Parkinson's disease severity," *J. Med. Syst.*, vol. 42, no. 2, pp. 1–17, Feb. 2018.

[51] N. Garcia, J. C. Vásquez Correa, J. R. Orozco-Arroyave, and E. Nöth, "Multimodal I-vectors to detect and evaluate Parkinson's disease," in *Proc. Interspeech*, Sep. 2018, pp. 2349–2353.

[52] J. C. Vasquez-Correa, T. Bocklet, J. R. Orozco-Arroyave, and E. Nöth, "Comparison of user models based on GMM-UBM and I-vectors for speech, handwriting, and gait assessment of Parkinson's disease patients," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 6544–6548.

[53] S. Wan, Y. Liang, Y. Zhang, and M. Guizani, "Deep multi-layer perceptron classifier for behavior analysis to estimate Parkinson's disease severity using smartphones," *IEEE Access*, vol. 6, pp. 36825–36833, 2018.

[54] J. C. Vásquez-Correa, J. R. Orozco-Arroyave, R. Arora, E. Nöth, N. Dehak, H. Christensen, F. Rudzicz, T. Bocklet, M. Cernak, H. Chinaei, J. Hannink, P. S. Nidadavolu, M. Yancheva, A. Vann, and N. Vogler, "Multi-view representation learning via gcca for multimodal analysis of Parkinson's disease," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 2966–2970.

[55] X. Lan, H. Mo, S. Chen, Q. Liu, and Y. Deng, "Fast transformation from time series to visibility graphs," *Chaos, Interdiscipl. J. Nonlinear Sci.*, vol. 25, no. 8, pp. 1–12, Aug. 2015.

[56] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, pp. 5–32, Oct. 2001.

[57] J. O. Smith, *Mathematics of the Discrete Fourier Transform (DFT): With Audio Applications*. W3K Publishing, 2008. [Online]. Available: https://ccrma.stanford.edu/~jos/st/

[58] J. Harrington and S. Cassidy, *Techniques in Speech Acoustics*, vol. 8. Cham, Switzerland: Springer, 1999.

[59] A. Pentari, G. Kafentzis, and M. Tsiknakis, "Investigating graph-based features for speech emotion recognition," in *Proc. IEEE-EMBS Int. Conf. Biomed. Health Informat. (BHI)*, Sep. 2022, pp. 01–05.

[60] A.-L. Barabási, "The network takeover," *Nature Phys.*, vol. 8, no. 1, pp. 14–16, Jan. 2012.

[61] M. Pósfai and A.-L. Barabási, *Network Science*. Princeton, NJ, USA: Citeseer, 2016.

[62] M. Newman, *Networks*. London, U.K.: Oxford Univ. Press, 2018.

[63] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–19.

[64] E. Bisong, "Logistic regression," in *Building Machine Learning and Deep Learning Models on Google Cloud Platform*. Berkeley, CA, USA: Apress, 2019, pp. 243–250.

[65] D. M. Farid, L. Zhang, C. M. Rahman, M. A. Hossain, and R. Strachan, "Hybrid decision tree and naive Bayes classifiers for multi-class classification task," *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1937–1946, 2014.

[66] M.-L. Zhang and Z.-H. Zhou, "A k-nearest neighbor based algorithm for multi-label classification," in *Proc. IEEE Int. Conf. Granular Comput.*, Jul. 2005, pp. 718–721.

[67] J. R. Quinlan, "Induction of decision trees," *Mach. Learn.*, vol. 1, no. 1, pp. 81–106, Mar. 1986.

[68] T. Hastie, S. Rosset, J. Zhu, and H. Zou, "Multi-class AdaBoost," *Statist. Its Interface*, vol. 2, no. 3, pp. 349–360, 2009.

[69] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 785–794.

[70] M. Sun, "A multi-class support vector machine: Theory and model," *Int. J. Inf. Technol. Decis. Making*, vol. 12, no. 6, pp. 1175–1199, Nov. 2013.

[71] F. H. Previc, "A general theory concerning the prenatal origins of cerebral lateralization in humans," *Psychol. Rev.*, vol. 98, no. 3, pp. 299–334, 1991.

[72] H. Sadeghi, P. Allard, F. Prince, and H. Labelle, "Symmetry and limb dominance in able-bodied gait: A review," *Gait Posture*, vol. 12, no. 1, pp. 34–45, Sep. 2000.

[73] W. H. Kruskal and W. A. Wallis, "Use of ranks in one-criterion variance analysis," *J. Amer. Stat. Assoc.*, vol. 47, no. 260, p. 583, Dec. 1952.

[74] P. Pérez-Toro, J. Vásquez-Correa, T. Arias-Vergara, E. Nöth, and J. Orozco-Arroyave, "Nonlinear dynamics and poincaré sections to model gait impairments in different stages of Parkinson's disease," *Nonlinear Dyn.*, vol. 100, pp. 3253–3276, Jul. 2020.

K. DEEPA RAJ received the B.Tech. degree in computer science and engineering from Mahatma Gandhi University, Kerala, India, in 2012, and the M.Tech. degree in computer science and engineering from Amrita Vishwa Vidyapeetham, Kollam, Kerala, in 2018. She is currently pursuing the Ph.D. degree in network science with the Amrita School of Artificial Intelligence, Amrita Vishwa Vidyapeetham, Coimbatore, Tamil Nadu, India. Her research interests include complex networks, complex systems, deep learning, and machine learning.

G. JYOTHISH LAL received the B.Tech. degree in electronics and communication engineering from Cochin University of Science and Technology and the M.Tech. degree in computational engineering and networking from Amrita Vishwa Vidyapeetham, Coimbatore, Tamil Nadu, India, and the Ph.D. degree in analyzing time-series data from the human speech production system for characterizing and predicting the dynamics of the system, in 2020. Since 2020, he has been an Assistant Professor with the Amrita School of Artificial Intelligence, Amrita Vishwa Vidyapeetham. His research interests include speech audio signal processing, nonlinear dynamics in complex systems, time series analysis, and machine learning deep learning for speech technology application.

E. A. GOPALAKRISHNAN received the Ph.D. degree from Indian Institute of Technology Madras, Chennai, India, in 2016. He is currently the Chairperson of the Amrita School of Computing, Amrita Vishwa Vidyapeetham, Bengaluru, India. From 2007 to 2011, he was an Assistant Professor with the Mechanical Department. From 2016 to 2023, he was an Associate Professor with the Center for Computational Engineering and Networking, Amrita Vishwa Vidyapeetham, Coimbatore, Tamil Nadu, India. He has published articles in prestigious journals, such as *Nature Scientific Reports*, *Chaos*, *Physical Review E*, *Proceedings of Combustion Institute*, IEEE TRANSACTIONS, *Expert Systems*, and *Journal of Fluid Mechanics*. His research interests include complex systems, data driven modeling and analysis, artificial intelligence, early warning signals for catastrophic transitions, and time series analysis. He serves as a Reviewer for journals, such as *PLOS One*, *Chaos*, *Physical Review E*, *Journal of Fluid Mechanics*, and *Expert Systems*.

V. SOWMYA received the M.Tech. degree in remote sensing and wireless sensor networks and the Ph.D. degree in artificial intelligence (AI) for natural scene analysis from the Amrita School of Engineering, Coimbatore, Tamil Nadu, India. Since 2011, she has been an Assistant Professor with the Amrita School of Artificial Intelligence, Coimbatore. She is currently an Assistant Professor with the Amrita School of Artificial Intelligence. She has published articles in IEEE TRANSACTIONS, *Artificial Intelligence Review*, *Multimedia Tools and Applications*, *Neural Computing and Applications*, *Computers and Security*, *Healthcare Technology Letters*. Her research interests include artificial intelligence for signal and image analysis, biomedical, agriculture, and ecology.

JUAN RAFAEL OROZCO-ARROYAVE received the B.Sc. degree in electronics engineering from Universidad de Antioquia, in 2004, the master's degree in marketing from EAFIT University, in 2009, the M.Sc. degree in telecommunications from Universidad de Antioquia, in 2011, the Ph.D. degree in electronics engineering from Universidad de Antioquia Medellín, Colombia, and the Ph.D. degree in computer science from Friedrich Alexander Universität Erlangen-Nürnberg, Erlangen, Germany, in a cotutelle program between the two universities. He is the Head of the signal processing line with GITA Research Group, Universidad de Antioquia, an Adjunct Researcher with the Pattern Recognition Laboratory, Friedrich Alexander Universität, and a Full Professor with Universidad de Antioquia. His research interests include speech processing, pattern recognition, digital communications, digital signal processing, and signal theory.