Models and Algorithms for the Diagnosis of Parkinson's Disease and their Realization on the Internet of Things Network

By Uladzimir, Vishniakou & Yiwei, Xia

Belarusian State University of Informatics and Radioelectronics

Abstract- This article aims to investigate an innovative approach utilizing model, algorithms and IoT technology for early Parkinson's disease detection. It introduces the comprehensive IoT network that has the IoT platform, enabling the collection of voice data via mobile phones, extraction of relevant features and data processing. Within this process, a Fully Connected Neural Network (FCNN) model is employed to calculate the probability of Parkinson's disease, potentially providing healthcare professionals and patients with a convenient, accurate, and early diagnostic tool. The study delves into the structure, algorithms, and the integral role of the FCNN within the IoT network, emphasizing its potential impact on the healthcare sector.

Keywords: parkinson's disease, IoT technology, early detection, voice data, noise reduction, fully connected neural network, IT-diagnosis.

GJRE-J Classification: LCC: RC382

MODELSANDALGORITHMSFORTHEDIAGNOSISOFPARKINSONSDISEASEANDTHEIRREALIZATIONONTHEINTERNETOFTHINGSNETWORK

Strictly as per the compliance and regulations of:

Models and Algorithms for the Diagnosis of Parkinson's Disease and their Realization on the Internet of Things Network

Uladzimir, Vishniakou [α] & Yiwei, Xia [σ]

Abstract- This article aims to investigate an innovative approach utilizing model, algorithms and IoT technology for early Parkinson's disease detection. It introduces the comprehensive IoT network that has the IoT platform, enabling the collection of voice data via mobile phones, extraction of relevant features and data processing. Within this process, a Fully Connected Neural Network (FCNN) model is employed to calculate the probability of Parkinson's disease, potentially providing healthcare professionals and patients with a convenient, accurate, and early diagnostic tool. The study delves into the structure, algorithms, and the integral role of the FCNN within the IoT network, emphasizing its potential impact on the healthcare sector.

Keywords: parkinson's disease, IoT technology, early detection, voice data, noise reduction, fully connected neural network, IT-diagnosis.

I. Introduction

As society continues to evolve and science and technology advance, there is an increasing emphasis on early detection and diagnosis of health issues. Parkinson's disease, as a chronic neurological condition, profoundly impacts the quality of life for affected individuals. Early diagnosis is crucial in providing more effective treatment and care [1]. This paper focuses on exploring a novel approach using IoT technology for early detection of Parkinson's disease [2]. Authors [3] proposed approach to early Parkinson's disease detection on voice analythis base. We will introduce a comprehensive IoT network that collects voice data through mobile phones, extracts relevant features, facilitates data transmission and processing, and ultimately outputs the probability of Parkinson's disease. The implementation of this method holds the promise of providing a more convenient, accurate, and early diagnostic tool for both medical professionals and patients. In this paper, we will delve into the structure and algorithms of this IoT network and discuss its potential impact in the field of healthcare. It is our hope that this innovative approach will be a significant step forward in early intervention and treatment of Parkinson's disease.

Author α: Professor, Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus. e-mail: vish@bsuir.by
Author σ: Ph.D. Student, Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus.

Initially, voice data is collected and pre-processed using a mobile phone. This involves capturing speech data from Parkinson's disease (PD) patients for a duration of 5 seconds at a sampling frequency of 44.1 kHz. To enhance signal quality, the spectral subtraction algorithm [4] is employed to eliminate ambient noise.

Subsequently, features are extracted from the pre-processed speech data after noise reduction. The data is then transmitted to the voice channel an IoT platform. The data is fed into a Matlab analysis function.

The Matlab analysis module plays a pivotal role in interpreting the data by loading a 3-layer FCNN model deployed in the cloud. It processes the data and generates a probability value indicating the likelihood of a possible Parkinson's disease diagnosis.

Finally, the results are relayed from the IoT platform to a mobile phone and the outcomes are displayed on a screen for further examination and evaluation.

II. Data Collection and Pre-Processing

Use the spectrum subtraction algorithm for the voice data, the specific process is as follows:

1. Split the original sound signal into frames, the length of each frame is 256 samples, using 50% overlap to split the frames and get a series of frames of signal.

Let the original sound signal be denoted as $s[n]$, where n is the sample index. - Let the frame length be $L = 256$ samples and the overlap percentage be 50%.

Define the frame index k such that the start index of the k-th frame is given by:

$$n_k = (k-1) * \frac{L}{2}, k = 1,2,3 \qquad (1)$$

The k-th frame of the signal is then given by:

$$s_k[n] = s[n_k + n_{k+1}], 0 \le n < L \qquad (2)$$

2. Perform Fourier transform on each frame of signal to get the corresponding spectrum.

Let $s_k[n]$ be the k-th frame of the signal. Apply a window function $w[n]$ to the frame to reduce spectral leakage, such as a Hamming window.

Compute the Fourier transform of the windowed frame to get the complex spectrum $X_k[f]$:

$$X_k[f] = sum\left(w[n] * s_k[n] * exp\left(-j * 2 * pi * f * \frac{n}{L}\right)\right), 0 <= f <= \frac{L}{2} \tag{3}$$

Use spectral subtraction algorithm $X_{avg}[f]$ as the basis to calculate the noise spectrum.

3. Wiener filter the noise spectrum to obtain a more accurate noise estimate.

Let $X_{noise}[f]$ be the noise spectrum estimated from $X_{avg}[f]$.

Apply a Wiener filter to $X_{noise}[f]$ to obtain a more accurate estimate:

$$X_{noise}^{filtered}[f] = \frac{X_{noise}[f]}{(X_{noise}[f] + \alpha * X_s[f])} \tag{4}$$

where α is a smoothing parameter and $X_s[f]$ is the spectrum of the original signal.

4. Compare the spectrum of each frame with the Wiener filtered noise spectrum, calculate the signal-to-noise ratio and consider the frequency components with a signal-to-noise ratio lower than 10 dB as noise components [5].

Let $X_k[f]$ be the complex spectrum of the k-th frame of the signal.

Compute the signal-to-noise ratio (SNR) for each frequency component as:

$$SNR[f] = 10 * log10\left(\frac{|X_k[f]|^2}{|X_{filtered}[f]|^2}\right) \tag{5}$$

Consider the frequency components with SNR < 10 dB as noise components.

5. Subtract the noise frequency components by adjusting the coefficient to 0.5 to obtain the noise removed spectrum.

Let be the noiseless spectrum of the k-th frame of the signal.

For each frequency component f, if $SNR[f] <$ 10 dB, then set the magnitude of $X_k^{noiseless}[f]$ to:

$$\left|X_k^{noiseless}[f]\right| = 0.5 * \left|X_k[f] - X_{noise}^{filtered}[f]\right| \tag{6}$$

To obtain the complete signal after removing noise, we need to convert the noise-removed spectrum back to the time domain and superimpose each frame.

Let $X_k^{noiseless}[f]$ be the noiseless signal in the frequency domain of the k-th frame of the signal, and $X_k^{noiseless}[n]$ be the corresponding signal in the time domain. Similarly, let $Y_k[f]$ be the noise-removed spectrum of the k-th frame, and $Y_k[n]$ be the corresponding signal in the time domain.

To convert the noise-removed spectrum back to the time domain, we can apply the inverse Fourier transform to $Y_k[f]$, which gives us $Y_k[n]$:

$$Y_k[n] = IFFT(Y_k[f]) \tag{7}$$

Then, we can combine the noise-removed signal of each frame to obtain the complete signal without noise:

$$x_{noiseless}[n] = sum_k\left(x_k^{noiseless}[n] * w_k[n]\right) \tag{8}$$

where $w_k[n]$ is a window function applied to each frame, and the sum is taken over all frames.

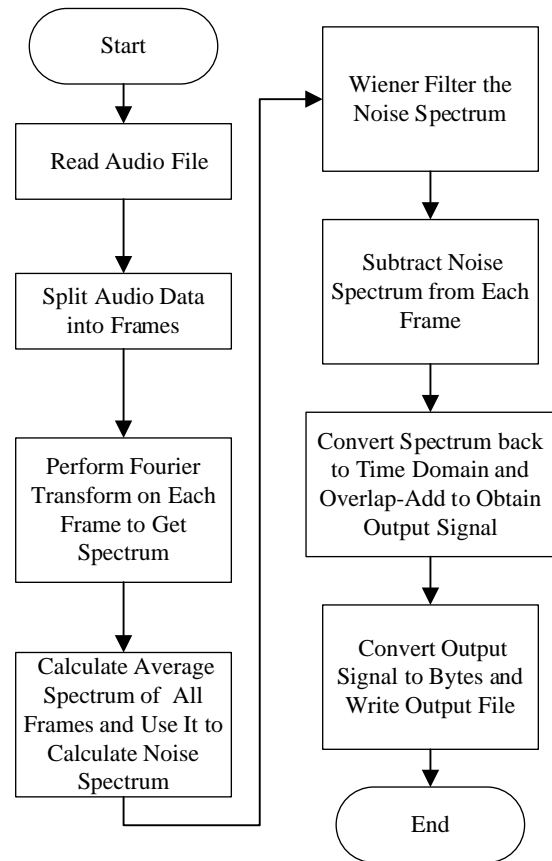The Figure 1 shows the algorithms of data pre-processing.

Figure 1: Algorithms of data pre-processing

The voice data after removing the noise is windowed. The main advantage of using the Hamming window [6] to extract the signal window is that it can reduce the oscillation effect at the edge of the window while retaining the main components of the signal inside the window. The window size is 1024 and the frequency of the voice data is 44.1 khz. The frequency of the voice data is 44.1 khz and the overlap rate of the window is 50%, so the speech time of a window is about 23 ms.

Let $x[n]$ be the original signal in the time domain with a sampling frequency of $fs = 44.1\,kHz$. Let $w[n]$ be the Hamming window of size $N = 1024$.

The windowed signal $x_w[n]$ is obtained by multiplying $x[n]$ with the window $w[n]$ and shifting the window by a hop size of $H = \frac{N}{2}$:

$$x_w[n] = x[n] * w[n - n_0] \tag{9}$$

where $n_0 = k * H$ for some integer k.

The Hamming window $w[n]$ is defined as:

$$w[n] = 0.54 - 0.46 * \cos\left(2 * pi * \frac{n}{N-1}\right), 0 \leq n \leq N-1 \tag{10}$$

The duration of each windowed segment is $T = \frac{N}{fs} = \frac{1024}{44100}, s = 0.023\sec(23\,ms)$, and the overlap between adjacent segments is $\frac{H}{T} = 2$.

III. FEATURE EXTRACTION

Feature extraction of the voice data [7] is performed within the specified window, as illustrated in Table 1 below.

Table 1: Extraction of all features

nNum	Feature Name	Description
1	MDVP:Fo(Hz)	Average vocal fundamental frequency
2	MDVP:Fhi(Hz)	Maximum vocal fundamental frequency
3	MDVP:Flo(Hz)	Minimum vocal fundamental frequency
4	MDVP:Jitter(%)	Measure of variation in fundamental frequency (percentage)
5	MDVP:Jitter(Abs)	Measure of variation in fundamental frequency (absolute value)
6	MDVP:RAP	Measure of variation in fundamental frequency (relative amplitude perturbation)
7	MDVP:PPQ	Measure of variation in fundamental frequency (pitch period perturbation quotient)
8	Jitter:DDP	Measure of variation in fundamental frequency (average of absolute differences of differences between adjacent periods)
9	MDVP:Shimmer	Measure of variation in amplitude (local variation in amplitude)
10	MDVP:Shimmer(dB)	Measure of variation in amplitude (local variation in amplitude in dB)
11	Shimmer:APQ3	Measure of variation in amplitude (amplitude perturbation quotient, 3-point method)
12	Shimmer:APQ5	Measure of variation in amplitude (amplitude perturbation quotient, 5-point method)
13	MDVP:APQ	Measure of variation in amplitude (average amplitude perturbation quotient)
14	Shimmer:DDA	Measure of variation in amplitude (average absolute difference of amplitudes between consecutive periods)
15	NHR	Ratio of noise to tonal components in the voice
16	HNR	Ratio of harmonics to noise in the voice
17	RPDE	Nonlinear dynamical complexity measure
18	D2	Nonlinear dynamical complexity measure
19	DFA	Signal fractal scaling exponent
20	spread1	Nonlinear measure of fundamental frequency variation
21	spread2	Nonlinear measure of fundamental frequency variation
22	PPE	Nonlinear measure of fundamental frequency variation

IV. DATA TRANSMISSION AND PROCESSING

To upload the 22 features to Thingspeak [8] IoT for analysis.

1. Creating a Channel. Figure 2 below shows the setup of the VOICE channel.

Figure 2: The voice channel settings

2. Getting the voice channel write/read API key and channel ID. The figure 3 below shows the voice channel write/read API key and channel ID.

Figure 3: The write/read API key and channel ID of the voice channel

3. Useing the API in the phone. The API library code uses the HTTP protocol to upload 22 voice features to the voice channel. The Figure 4 shows the ThingSpeak API data upload algorithm.
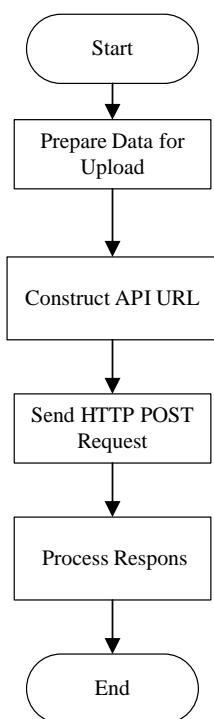
Figure 4: ThingSpeak API data upload algorithm

4. To load the pre-trained FCNN [9] model in Matlab Analysis module and to input the 22 data into the model for analysis to obtain the results. The model was trained using a publicly available dataset for Parkinson's Disease [10]. The Figure 5 shows the schematics of 3-Layer FCNN. Table 2 shows hyperparameters of 3-layer FCNN.
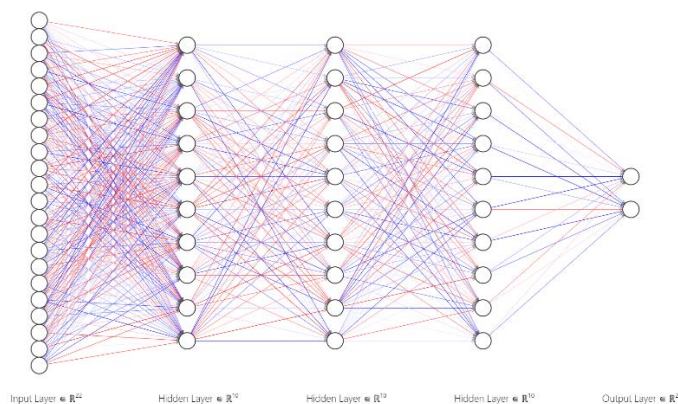

Input Layer $\in \mathbb{R}^{22}$ Hidden Layer $\in \mathbb{R}^{10}$ Hidden Layer $\in \mathbb{R}^{10}$ Hidden Layer $\in \mathbb{R}^{10}$ Output Layer $\in \mathbb{R}^{2}$

Figure 5: Schematics of 3-Layer FCNN

Table 2: Hyperparameters of 3-layer FCNN

Name	Hyperparameters
First Layer Size	10
Second Layer Size	10
Third Layer Size	10
Activation Function	ReLU
Iteration Limit	1000
Learning Rate	0.01
Learning Rate Update Algorithm	SGD
Regularization Strength (Lambda)	0
Standardize Data	Yes

In Thingspeak, store the result value to file1 in the voice channel, and then the phone reads the value of file1 in the voice channel.

V. Results and Discussion

The process begins with receiving the initial signal from the microphone, then the signal goes through a number of preprocessing stages, including high-frequency isolation, noise suppression and segmentation of the speech frame using a Hamming window. Key characteristics are then extracted from these processed data and compiled into datasets. These data sets are transmitted to a neural network for training and optimization of the model, visually represented in the figure as a multi-layered structure. The completion of training and optimization is the creation of a model file capable of classifying voice input data. The experiment was conducted on an international dataset [11]. The results of test experiments in the IV network for the diagnosis of PD in patients with speech changes are shown in Table 3.

Table 3: The data of test experiments for speech recognition

Набор данных/ Показатели	Средняя точность	Средняя чувствительность	Средняя F1 оценка	Точность тестирования
БП по речи_	92,95%	92,95%	92,95%	94,7%

The IoT network achieved 94.7% accuracy in diagnosing Parkinson's disease based on speech data and an F1 score of 92.95%. On the same data set, one of the best indicators of foreign studies is 95.8% [12], which indicates both good recognition results and the possibility of implementing an IV network for domestic PD diagnostics.

VI. Conclusions

In summary, the IoT network efficiently collects voice data from PD patients, processes it to remove noise, extracts essential features, and utilizes a 3-layer FCNN model to provide probability-based diagnostic outcomes, offering a viable solution for the timely detection of Parkinson's disease. Our work underscores the pivotal role of the IoT advancing healthcare. By seamlessly connecting devices and systems, IoT not only enables remote diagnostics but also promotes patient empowerment, personalized medicine, and enhanced healthcare delivery.

An approach using machine learning, neural networks, signal processing, and the Internet of Things (IoT) technologies for early detection of Parkinson's disease is described. A model and algorithms for processing audio signals from patients studied for the likelihood of Parkinson's disease are presented. The Internet of Things network collects voice data from PD patients, processes it to eliminate noise, extracts important characteristics, and uses a three-level FCNN model to obtain probability-based diagnostic results, offering a solution for timely detection of Parkinson's disease. 2. This work highlights the role of the Internet of Things in the development of IT diagnostics of patients. Thanks to the wireless connection of devices, the IoT network not only provides remote diagnostics, but also helps to empower patients and doctors, staff and improve the quality of medical care. As a result of experiments on an international dataset in the IV network, 94.7% accuracy was achieved in diagnosing Parkinson's disease based on speech data.

References Références Referencias

1. Whitehouse PJ. Ethical issues in early diagnosis and prevention of Alzheimer disease. *Dialogues in Clinical Neuroscience*. 2019, Mar 31;21(1):101-108. DOI:10.31887/DCNS.2019.21.1/pwhitehouse.
2. Giannakopoulou KM, Roussaki I, Demestichas K. Internet of things technologies and machine learning methods for Parkinson's disease diagnosis, monitoring and management: a systematic review. *Sensors*. 2022;22(5):1799. DOI:10.3390/s22051799.
3. Vishniakou UA, YiWei X. IT Diagnostics of Parkinson's Disease Based on the Analysis of Voice Markers and Machine Learning. *Doklady BGUIR*. 2023;21(3):102-110. DOI:10.35596/1729-7648-2023-21-3-102-110.
4. Upadhyay N, Karmakar A. Speech enhancement using spectral subtraction-type algorithms: A comparison and simulation study. *Procedia Computer Science*. 2015;54:574-84. DOI:10.1016/j.procs.2015.06.066.
5. Dendrinos M, Bakamidis S, Carayannis G. Speech enhancement from noise: A regenerative approach. *Speech Communication*. 1991;10(1):45-57. http://hdl.handle.net/123456789/10563
6. Al-Barhan HA, Elyass SM, Saeed TR, Hatem GM, Ziboon HT. Modified speech separation deep learning network based on Hamming window. In: IOP Conference Series: *Materials Science and Engineering*. 2021;1076(1).
7. Little M, Mcsharry P, Roberts S, Costello D, Moroz I. Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. *Nature Precedings*. 2007;1-11. DOI:10.1186/1475-925X-6-23.
8. Pasha S. ThingSpeak based sensing and monitoring system for IoT with Matlab Analysis.

International Journal of New Technology and Research (IJNTR). 2016;2(6):19–23. Https://ieeex plore.ieee.org/abstract/document/7178838/

9. Sainath TN, Vinyals O, Senior A, Sak H. Convolutional, long short-term memory, fully connected deep neural networks. In: 2015 *IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE. 2015:1-5. DOI:10. 1109/ICASSP.2015.7178838.

10. Little M, McSharry P, Hunter E, Spielman J, Ramig L. Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. Nature Precedings. 2008:1-27. DOI:10.1109/TBME.2008.20 05954.

11. Parkinson Data Set [Electronic resource]. − Access mode: https://archive.ics.uci.edu/ml/datasets/parkin sons. − Access date: 12.6.2024.

12. Sakar BE, Isenkul ME, Sakar CO, Sertbas A. Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. *IEEE Journal of Biomedical and Health Informatics*, 2013, vol. 17(4), pp. 828–834. DOI:10.1109/JBHI. 2013.2245674.