Vol. 06, No. 4 (2024) 1541-1550, doi: 10.24874/PES06.04.013

Proceedings on Engineering Sciences

www.pesjournal.net

A NOVEL VOICE-BASED SYSTEM FOR PARKINSON'S DISEASE DETECTION USING RNN-LSTM

Repudi Pitchiah[1]
T. Sasi Rooba
K. Uma Pavan Kumar

Received 24.04.2024.
Received in revised form 18.09.2024.
Accepted 30.09.2024.
UDC –004.032.26:616.858

Keywords:

Classification, Deep Neural Networks, Parkinson's disease, Deep Learning

A B S T R A C T

PD (Parkinson's Disease) is a neurological illness that develops with time and causes motor and non-motor symptoms. A lot of PD patients have trouble moving normally in the early phases of the illness. Vocal disorders are among the most prevalent symptoms. Latest PD detection investigations have concentrated on diagnostic methods on the basis of vocal problems, which are an excitingly new area of study with a lot of potential. For a range of prediction problems that are troubling medical practitioners, Deep Learning (DL) has gained popularity recently. In this study, RNN-LSTM is combined with numerous architectures to develop more accurate prediction models for the detection of PD on the basis of feature analysis of various patient speech samples. Importantly, Deep Neural Networks have become the best classification tool for PD detection even without the use of a feature selection strategy. RNN-LSTM was then fine-tuned, resulting in an accuracy of 98.772 percent.

1. INTRODUCTION

PD is a significant neurological disease that impacts1 percent of the population over 60 years old, or 1 to 2 people per 1000 (Wirdefeldt et al., 2011). According to age-standardized prevalence rates and a rise in the number of senior persons, the estimated number of individuals impacted with PD globally increased by more than double between 1990 and 2016 (from 2.5-6.1 million) (Rana et al., 2015). Planning, initiation, and execution of motions are all included in PD, which is a degenerative neurological condition having motor &non-motor symptoms (Naranjo et al., 2016). Before cognitive and behavioral problems, movement-related signs like tremors, stiffness, and initiation difficulties might be seen during its development (Cantürk & Karabiber, 2016). The QoL (Quality of Life) of PD patients is substantially impacted, as are their social interactions, family connections, and economic costs on both the individual and societal levels (Shanmugam et al., 2018; Delić et al., 2019; Chenausky et al., 2011).Conventionally, motor signs were utilized to make the PD diagnosis. Most of the rating scales applied to quantify the severity of the illness were not thoroughly determined and validated (Naranjo et al., 2016), despite the identification of cardinal indications of PD in clinical examinations. Although non-motor symptoms, like behavioral, cognitive, and sleep abnormalities, as well as sensory anomalies like olfactory dysfunction are frequent in patients before the PD onset (Naranjo et al., 2016; Sakar & Kursun, 2010), They lack specificity, are difficult to evaluate, or result

[1] Corresponding author: Repudi Pitchiah
Email: pitchaiah99@gmail.com

in variability across patients (Can, 2013). Thus, Non-motor signs have been employed as supporting diagnostic criteria, but they do not yet allow for an independent PD diagnosis (Almeida et al., 2019.

1.1. Different stages of PD

The following list describes the various stages of PD:

- **Stage 1 (Mildest Stage):** The least interruption with everyday activities is now experienced by PD patients. Other symptoms, including tremors, only affect one side of the body.
- **Stage 2 (Moderate Stage):** At the moment, both sides of the body are experiencing signs including stiffness, resting tremors, and shaking. Facial expressions could alter in PD patients as well.
- **Stage 3 (Mid-Stage):** Substantial changes in PD patients at this stage include loss of balance, decreased flexibility, and stage II signs. Occupational therapy and drugs can be used to reduce symptoms.
- **Stage 4 (Progressive Stage):** At this time, the PD patient's condition may worsen, making it hard for them to travel without using an aid like a walker.

Stage 5 (Advanced stage): The most debilitating and painful phase of PD occurs at this stage, affecting sufferers. Leg stiffness makes standing challenging. The inability to stand without falling is another issue for patients. The occasional paranoia and hallucination may occur in them.

Most common illnesses and disorders may be divided into types that are contagious and non-contagious, each with a range of prognosis and symptom severity. While some diseases have a treatment-responsive course of action and may be recovered from, some illnesses have no cure and continue to bother people for years or even their whole lives. Nearly 11 million people globally suffer from this ailment. In search of a more effective cure for the illness, several research have been conducted. The process of Machine Learning (ML) entails the examination of empirical data to develop a workable future solution (Can, 2013; Åström & Koker, 2011; Irfan et al., 2024).

1.2 Symptoms of Parkinson Disease

There are 2 major categories of symptoms associated with PD: Motor &Non-Motor Symptoms

1.2.1 Motor Symptoms

This is a medical disorder when a patient finds it difficult to engage in voluntary activity. Some signs of movement disorders include tremors, stiffness, freezing, Brady kinesia, or other spontaneous muscular contractions (Lahmiri, 2017).

1.2.2 Non-Motor Symptoms

It includes apathy, cognitive impairment, and complicated personality problems. They are generally classified by doctors as primary &secondary PD symptoms.

1.2.3 Primary Symptoms

It is the most significant symptom. Regular symptoms include tremors, rigidity as well as movement slowness (Lahmiri, 2017; Kollias et al., 2018).

1.2.4 Secondary Symptoms

The impact of this type of symptom on a person's life is significant. There are two types of these: motorized and non-motorized. Each person reacts to it differently. There are many different ways that PD expresses itself.

Vocal impairment affects 90 percent of persons with PD (Shanmugam et al., 2018). Vocal defects shouldn't appear out of the blue. They mark the end of a protracted phase that is sometimes disregarded at the beginning. Therefore, for patients and researchers, early tele-monitoring and detection tools on the basis of precise, efficient, and unbiased prediction models are essential. In recent studies, dysphonia characteristics (acoustic testing) have been utilized to identify speech issues (Nawir et al., 2019) using ML approaches. The underlying frequency variations or changes in vocal oscillation pitch (F0) are reflected in the diagnosed illnesses. Absolute sound pressure level, which represents the relative speech loudness, is another criterion. Jitter, for example, is a cyclic fluctuation in fundamental frequency, whereas shimmer is a cyclic change in voice loudness. The level of acoustic periodicity is represented by a criterion like harmonicity. To forecast the potential of PD, we have attempted to use BPDNN ("Back Propagation Deep Neural Networks") to the characteristics derived from various speech samples of tested individuals with various designs (Kollias et al., 2018). Through an unsupervised parameter pre-training approach, the DNN is trained with improved starting parameters. Based on this, the model further optimizes parameters using the supervised training method (Sadek et al., 2019; Lahmiri et al., 2018).

2. RELATED WORK

ML and DL approaches would be the greatest method for separating PD patients from healthy individuals. This section examines several ML & DL-based speech signal-based PD prediction systems. The pitch, glottal pulse, Shimmer, Jitter, and MFCC properties of speech signals may be extracted using a technique that Sharma & Giri (2014) suggested. TQWT (Tunable Q-Factor Wavelet Transform) approach was suggested by Sakar et al. (2019) to identify "PD patients with voice signals. This dataset has been examined with DL techniques known as CNN (Convolutional Neural Networks) (Gunduz, 2019).

Bouchikhi et al. (2013) suggested the relief-F feature selection model along with SVM classifier. Out of 22 features, this approach selected 10 features. Relief-F feature selection demonstrated 96.88 percent accuracy using SVM classifier with 10-fold cross-validation. 195 speech samples make up the experimental dataset. RBA (Relief- Based Algorithms) are vulnerable to closest neighbor" noise intrusion. A nonlinear SVM with a PCA on the basis of PD detection method was suggested by Hemmerling & Sztaho (2019). This model has an accuracy rate for PD categorization of 93.43 percent. The experimental dataset was small and the level of prediction accuracy was poor.

Parisi et al. (2018) suggested a multi-layer perceptron with a lagrangian SVM-based classifier for PD patient identification. The appropriate features were allocated by MLP using unique cost functions that included AUC score and accuracy. The 20 most significant and pertinent characteristics are extracted using MLP using the score value. Compared to other methods, the suggested model has 100% accuracy, hybrid preprocessing and PD classification technique. To lessen the variance of the dataset, SCFW ("Subtractive Clustering Features Weighting") was suggested as a preprocessing technique. They suggested the KELM ("Kernel-Based Extreme Learning Machine") as a classifier and supported the effectiveness of KELM in terms of specificity, accuracy,and sensitivity, as well as the value of the Rocand Kappa statistic curve.

Caliskan et al. (2017) suggested a DNN classifier for the identification of PD that uses a stacked autoencoder to extract vocal features. This paper compares DNN-based PD prediction to conventional ML models and finds that it secures good accuracy. Additionally, DNN requires more data during the training phase, and searching the parameter space requires additional training time.

In (Vásquez Correa et al., 2017), CNN was applied to the wavelet transform and STFT (Short-Time Fourier Transform) to extract the voice characteristics. The main challenge for CNN is to use dilated convolution layers to simulate long-distance contextual data (Funahashi & Nakamura, 1993). RNN (Recurrent Neural Networks) gets over this problem and can handle distant contextual input by preserving the results of past computations. Nevertheless, the gradient problem with RNN-based methods makes it challenging to fine-tune the network layer's parameters. This problem has been resolved using LSTM (Hochreiter & Schmidhuber, 1997).

In the present research, a classification model for the detection of PD with RNN-trained LSTM with graph structure was suggested. The MABO optimizer, based on numerous speech characteristics, considerably increases the categorization accuracy. Without increasing the model size, RNN-LSTM can handle larger datasets. LSTM is more efficient than typical time series models because it learns long-run dependencies that follow the prior time sequence and forward to subsequent layers. With the help of feature extraction and preprocessing techniques, the suggested model also gains additional advantages. It solves the drawbacks of earlier techniques, such as their use of feed-forward layers and small datasets, which reduce the accuracy of PD prediction. The accuracy of PD prediction was improved by RNN-usage LSTM's of loop networks both forward and backward.

3. PROPOSED MODEL METHODOLOGY

PD sufferers typically have a difficult time being classified since control efficiency is a pattern categorization issue. To correctly identify such patterns, the data is split into sub-datasets including tests of persons with a distinctive mode of expression, known as speech samples. Following that, elements from a speech sample are chosen, and their significance in the presence of PD is evaluated. The chosen characteristics are then fed into a classifier as input using the extracted data from every voice sample (m indicates the number of samples). A majority preference determines the final result after each classifier guesses the name of its class. Figure 1 displays a block schematic of the suggested approach.

3.1 PD Dataset

It included speech samples from UCI's ML archives that had previously been utilized in research to identify PDs (Ma et al., 2014). We included 266 healthy control (HC) participants and 160 PD subjects for this research, with the latter grouping into early (72 newly diagnosed subjects) and mid-advanced (88 subjects with moderate to severe disability) patients. Following patient evaluations, three copies of each person's vowel have been noted at a frequency of 44.1 kHz, along with their voices.

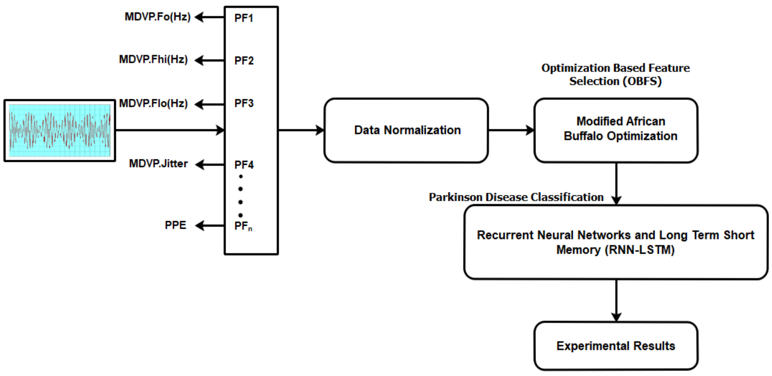

Figure 1. PD Patients Identification/Classification System Model

3.2 Normalization

At this phase, meander, spiral, voice, and speech-Sakar datasets from four PDs were normalized. Normalization has been carried out on the whole features of every dataset. Standard normalization is one of the fundamental techniques for normalization. Each feature has been normalized in an interval between minimum $MinX$ and maximum $MaxX$ in the suggested approach, and the interval was then transformed into a new interval $[NewMinX, NewMaxX]$ on the basis of eq. (1). As a result, every value in each feature has been normalized to a new one. The terms are employed to normalize the data, according to the equation below. As a consequence, the chosen dataset for analysis is the findings obtained.

Features are transformed to have a comparable scale using normalization or min-max scaling. The new point is determined as follows:

X_new=(X-X_min)/(X_max-X_min) (1)

This scale ranges between [0, 1] or sometimes [-1, 1].

3.3 Feature Extraction from Speech PD Data

Voiceprint recognition depends on the extraction of speech feature characteristics. Slower changes occur in the speech signal. The voice signal is typically thought to be steady at intervals between 10 and 30 ms when it is received quickly. Therefore, short-time spectrum analysis could be used for computation (Wang et al., 2020). Applying the Mel scale, which is computed by 1000 Hz equating to 1000 Mel, estimates the frequency perception of the human ear. To provide more accurate tests to identify speech problems, this study used the cepstrum, spectrum, and temporal speech quality domains (Shirvan & Tahami, 2011; Vaiciukynas et al., 2017). Basic measures include variables such as the harmonic noise ratio (HNR), jitter, shimmer, and fundamental frequency of F0 ("Vocal Cord Vibration").

Table1. Acoustic Analysis Findings of Healthy with PD

Condition	Sex	Age (Range)	F0(HZ)	Shimmer (%)	Jitter (%)	HNR(dB)
Person Healthy	F	11.8 ± 55.7	205.6 ± 37.8	0.36 ± 0.47	1.24 ± 1.27	11.2 ± 7.2
	M	12.6 ± 58.5	127.4 ± 17.4	0.24 ± 0.11	0.05 ± 0.37	14.9 ± 4.7
Person with Disease	F	61.9 ± 10.8	193.6 ± 16.6	0.69 ± 0.92	1.92 ± 1.34	8.2 ± 5.2
	M	62.3 ± 9.8	120.6 ± 20.7	0.38 ± 0.17	0.96 ± 0.78	10.5 ± 3.8

The distinctive parameters for analysis were derived on the basis of pronunciation features of PD patients. Nevertheless, the feature parameters' components each have unique speech characterization skills for various speech samples.

3.4 Feature Selection

The choice of feature subsets in this study is made using Modified African Buffalo Optimization (MABO), where the features of the samples are taken into account on the basis of impacts of feature existences within PDs. Then, classifiers employ such chosen properties from samples (signifies the number of voice samples). Classifiers project their class labels, and decisions are ultimately made based on assessments. The features with the greatest weights are selected as the original characteristics, and feature weights are assigned to them to highlight their importance to classifications.

The proposed MABO is presented in Appendix.

3.5 DNN Classification

Several features are extracted to train the LSTM for classification as follows.

In contrast to RNN, LSTM contains memory cells rather than each standard node in the hidden layer, allowing data to be stored and retrieved for an extended period to enhance learning and prevent the vanishing gradient problem. LSTM networks are often employed in time series prediction in applications including voice recognition, air pollution forecasting, and machine translation (Caliskan et al., 2017). The feed-forward neural network with internal memory in its extended version is called an RNN. The RNN's output is based on earlier calculations and is returned to the recurrent network. The input series are operated on and a decision is made by the RNN using internal memory. Back propagation is the training method used for LSTM (Long Short-Term Memory). The 3 gates that make up an LSTM are the input, output, and forget gates. To choose the input values that change the memory, input gates employ sigmoid activation functions. The forget gate examines whether information from the previous state should be deleted, while the output gate controls the output. In contrast to conventional LSTM, graph LSTM presents a single LSTM unit in each tree node, as observed in Fig. 2.

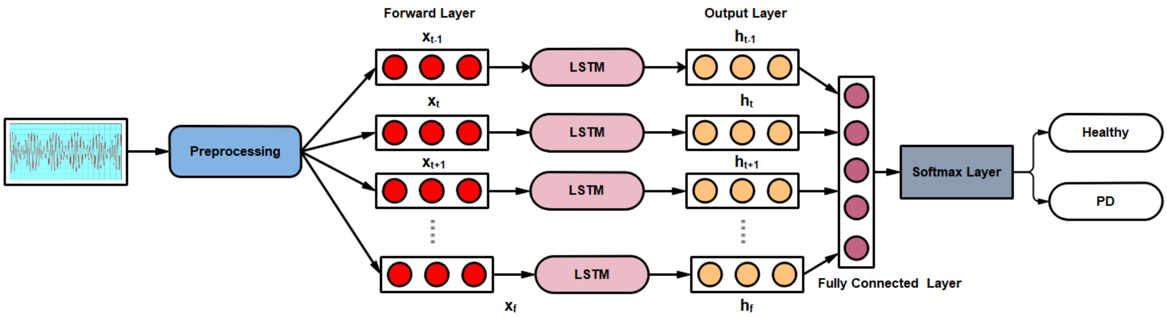

Figure 2. Proposed RNN-LSTM classifier for PD diagnosis

An input, 5 hidden, as well as an output layer, are among the seven layers in this model. An input layer made up of LSTM cells makes up a recurrent neural network. Each LSTM layer's input layer corresponds to a speech signal's PF (“Phonation Features”). Twenty-three neurons inside I/P layer of the LSTM cell, which is depicted in Figure 3, represent 23 features.

$$\textbf{(Hidden Forward)}\ \vec{h} = H(w_{PF\vec{h}} PF_t + w_{\vec{h}\vec{h}} \vec{h}_{t-1} PF_t + b_{\vec{h}}) \quad (2)$$

$$\textbf{(Hidden Backward)}\ \vec{h} = H(w_{PF\overleftarrow{h}} PF_t + w_{\overleftarrow{h}\overleftarrow{h}} \overleftarrow{h}_{t-1} PF_t + b_{\overleftarrow{h}}) \quad (3)$$

$$\textbf{(Output)}\ y_t = w_{\vec{h}y} \vec{h}_t + w_{\overleftarrow{h}} \overleftarrow{h}_t + b_y \quad (4)$$

Here, H signifies the hidden layer function of every feature, denotes the weight matrices, and B presents the bias vector.

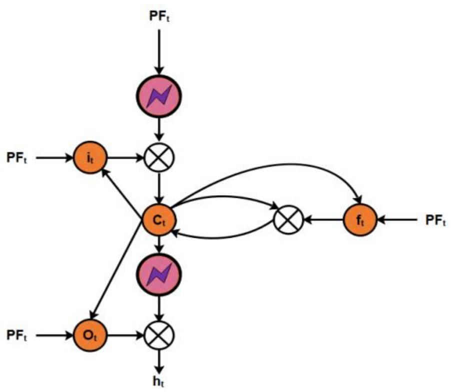

Figure 3. LSTM cell

The parameters in figure 3 for the forward LSTM, including input gate (i_t), forget gate (f_t), output gate (o_t), along with cell state (c_t) for a certain iteration t with an activation function (σ), are changed as in the subsequent equations (5) from to (9).

$$f_t = \sigma(w_{PFf} PF_t + w_{\vec{h}f} \vec{h}_{t-1} + w_{cf} c_{t-1} + b_f) \quad (5)$$

$$i_t = \sigma(w_{PFi} PF_t + w_{\vec{h}i} \vec{h}_{t-1} + w_{ci} c_{t-1} + b_i) \quad (6)$$

$$o_t = \sigma(w_{PFo} PF_t + w_{\vec{h}o} \vec{h}_{t-1} + w_{co} c_{t-1} + b_o) \quad (7)$$

$$c_t = f_t c_{t-1} + i_t \tanh(w_{PFc} PF_t + w_{\vec{h}c} \vec{h}_{t-1} + b_c) \quad (8)$$

$$\vec{h}_t = o_t \tanh c_t \quad (9)$$

The following updates are made to the settings for backward GLSTM in Equations (10) to (14).

$$f_t = \sigma(w_{PFf} PF_t + w_{\overleftarrow{h}f} \overleftarrow{h}_{t-1} + w_{cf} c_{t-1} + b_f) \quad (10)$$

$$i_t = \sigma(w_{PFi} PF_t + w_{\overleftarrow{h}i} \overleftarrow{h}_{t-1} + w_{ci} c_{t-1} + b_i) \quad (11)$$

$$o_t = \sigma(w_{PFo} PF_t + w_{\overleftarrow{h}o} \overleftarrow{h}_{t-1} + w_{co} c_{t-1} + b_o) \quad (12)$$

$$c_t = f_t c_{t-1} + i_t \tanh(w_{PFc} PF_t + w_{\overleftarrow{h}c} \overleftarrow{h}_{t-1} + b_c) \quad (13)$$

$$\overleftarrow{h}_t = o_t \tanh c_t \quad (14)$$

4. EXPERIMENTAL RESULTS

The following system specifications were used in the implementation: 1 TB hard drive,4.00 GB RAM,64-bit operating system, Windows 8.1 Pro, and Intel (R) Core^{TM}i3-4160T CPU@3.10 GHz 3.09 GHz processor.

4.1 Evaluation Metrics

The results obtained from the proposed technique were calculated with the assessment criteria explained below. The definition of accuracy is the proportion of accurate forecasts to the total number of predictions. The letters t_p, t_n f_p & f_n in the confusion matrix mean t_p (“True

Positive"), t_n ("True Negative"), f_p ("False Positive"), and f_n ("False Negative")correspondingly. Accuracy, F-measure, recall, and precision were computed with formulae on the basis of counts

$$\textbf{Accuracy} = \frac{t_p + t_n}{t_p + t_n + f_p + f_n} \quad (15)$$

$$\textbf{F-Measure} = \frac{2 * t_p}{2 * (t_p + f_p + f_n)} \quad (16)$$

$$\textbf{Recall} = \frac{t_p}{t_p + f_n} \quad (17)$$

$$\textbf{Precision} = \frac{t_p}{t_p + f_p} \quad (18)$$

Scikit-Learn 0.22.1 for Python has been used to develop the suggested PD diagnostic model RNN-LSTM optimized with ADAM. The parameters for conventional ML models are displayed in Table 2.

Table 2. ML Parameter Settings

ML Models	Parameters
Multilayer Perceptron	Sizes of Hidden Layer =100,200,300
Proposed RNN-LSTM	Hidden Layer=5, RNN, Min-Max Normalization, 23 Neurons in Hidden Layer, Modified African Buffalo Optimization (MABO), Dropout Rate=0.2.

Table 3. Voice Dataset Segmentation on the basis of Different Features

S. No.	Features	Number
1	Band width + Formant	8
2	Vocal Fold	22
3	Baseline	26
4	MFCC (Mel-Frequency cepstral coefficients)	84
5	Wavelet transform applied to F0	182
6	TQWT	432

4.2 Results Comparison

Three different types of features were used in experimental assessments of classifiers to assess recall, precision, F-measure, and accuracy. The accuracy rate for the combination of MFCCs+ Wavelets+ Concated features for SVM was 88.1294 percent while the accuracy rate of these features for CNN was 94.1752 percent. The recommended RNN-LSTM classifier with combinations of above-mentioned features had the greatest accuracy rates, coming in at 98.7720 percent (F-measure rate of 71.400 and 98.5010 percent for MCC).

Table 4. Accuracy findings of feature level combination versus classifiers.

Feature Combination	SVM	MLP	CNN	Proposed
TQWT + MFCC + Wavelet	85.1035	87.5696	93.047	96.6381
TQWT + MFCC + Concat	83.664	91.9315	93.0854	98.0244
TQWT + Wavelet + Concat	87.4662	88.0694	93.1261	97.3457
MFCC + Wavelet + Concat	88.1294	94.1752	95.1557	98.772

Figure 4 shows the accuracy of the x-axis as measured with feature-level combinations on classifiers. RNN-LSTM obtained 98.772% accuracy when compared to CNN, MLP, and SVM which attained 95.1557%, 94.1752%, and 88.1294%, accuracy at the final feature level com bination.

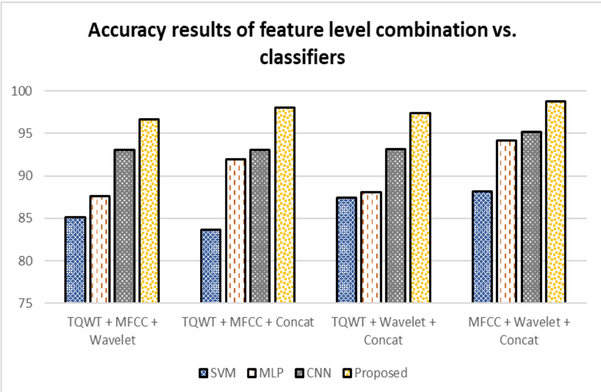

Figure 4. Results of feature level combinations vs. classifiers in terms of accuracy

Table 5. F-measure results of feature level combination versus classifiers

Feature Combination	SVM	MLP	CNN	Proposed
TQWT + MFCC + Wavelet	82.9150	85.8697	94.2258	98.3100
TQWT + MFCC + Concat	80.7960	90.2315	91.5250	96.5900
TQWT + Wavelet + Concat	86.3590	86.3695	93.4200	97.5200
MFCC + Wavelet + Concat	87.4510	92.4752	91.6921	98.5010

Fig. 5 compares the F-measure outcomes of four distinct feature-level combinations employing different classifiers. With the initial feature level combination, the proposed RNN-LSTM generated a greater F-measure value of 98.3100 percent, outperforming CNN, MLP, and SVM, which attained F-measures of 94.2258 %, 85.8697 %, and 82.9150 %, correspondingly.

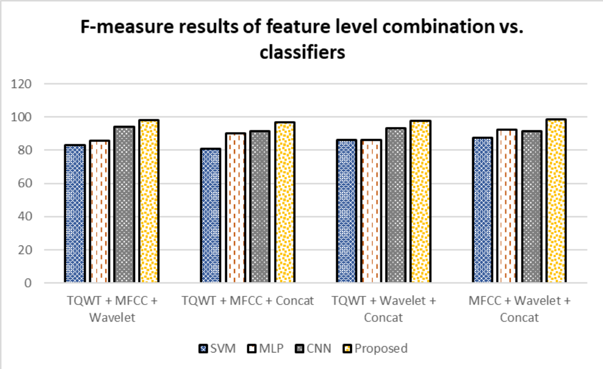

Figure 5. F-measure outcomes for a combination of feature levels versus classifiers

5. CONCLUSION

PD is the 2nd most common neurological condition, causes severe damage, diminishes quality of life, and has no therapy. To give a multiclass classification problem for PD analysis in this study, a feature selection is applied. A PD diagnosis is then made using the RNN-LSTM classifier. It is a good and reliable method for correctly detecting the condition at an early stage, which may assist medical professionals in helping PD sufferers get better and be cured. Accuracy, F-measure, recall, and precision were used to evaluate the performance of classification algorithms evaluated with UCI's ML libraries. The data demonstrate that the recommended model's accuracy is greater than the other present methods when the findings are compared to other methods currently in use.

References:

Almeida, J. S., Rebouças Filho, P. P., Carneiro, T., Wei, W., Damaševičius, R., et al. (2019). Detecting PD with sustained phonation and speech signals using machine learning techniques. *Pattern Recognition Letters, 125*, 55–62.

Åström, F., & Koker, R. (2011). A parallel neural network approach to prediction of Parkinson's disease. *Expert Systems with Applications, 38*(10), 12470–12474.

Bouchikhi, S., Boublenza, A., Benosman, A., & Chikh, M. (2013). PD detection with SVM classifier and relief-F features selection algorithm. *South East European Journal of Soft Computing, 2*(1), 1–4.

Caliskan, A., Badem, H., Basturk, A., & Yuksel, M. E. (2017). Diagnosis of Parkinson's disease by using deep neural network classifier. *IU-Journal of Electrical & Electronics Engineering, 17*(2), 3311–3318.

Can, M. (2013). Neural networks to diagnose the Parkinson's disease. *Southeast Europe Journal of Soft Computing, 2*(1), 15467–15479.

Cantürk, F., & Karabiber, M. (2016). A machine learning system for the diagnosis of PD from speech signals and its application to multiple speech signal types. *Arabian Journal for Science and Engineering, 41*(12), 5049–5059.

Chenausky, K., MacAuslan, J., & Goldhor, R. (2011). Acoustic analysis of PD speech. *Parkinson's Disease, 2011*, 1–13.

Delić, V., Perić, Z., Sečujski, M., Jakovljević, N., Nikolić, J., Mišković, D., ... & Delić, T. (2019). Speech technology progress based on new machine learning paradigm. *Computational intelligence and neuroscience*, *2019*(1), 4368036.

Funahashi, K. I., & Nakamura, Y. (1993). Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Networks*, *6*(6), 801–806.

Gunduz, H. (2019). Deep learning-based PD classification using vocal feature sets. *IEEE Access, 7*, 115540–115551.

Hemmerling, D., & Sztaho, D. (2019). PD classification based on vowel sound. *Journal of Speech, Language, and Hearing Research, 56*(5), 1485–1502.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780.

Irfan, N., Zafar, S., Hussain, I., & Biswas, S. S. (2024). Enhancing maternal psychological health during pregnancy through e-health informatics. *Proceedings on Engineering Sciences,* *6*(3), 1379-1390. doi: 10.24874/PES.SI.25.03A.006

Kollias, D., Tagaris, A., Stafylopatis, A., Kollias, S., & Tagaris, G. (2018). Deep neural architectures for prediction in healthcare. *Complex & Intelligent Systems*, *4*(2), 119–131.

Lahmiri, S. (2017). PD detection based on dysphonia measurements. *Physica A: Statistical Mechanics and its Applications, 471*, 98–105.

Lahmiri, S., Dawson, D. A., & Shmuel, A. (2018). Performance of machine learning methods in diagnosing PD based on dysphonia measures. *Biomedical Engineering Letters, 8*(1), 29–39.

Ma, C., Ouyang, J., Chen, H. L., & Zhao, X. H. (2014). An efficient diagnosis system for PD using kernel-based extreme learning machine with subtractive clustering features weighting approach. *Computational and Mathematical Methods in Medicine, 2014*, 1–14.

Naranjo, L., Perez, C. J., Campos Roca, Y., & Martin, J. (2016). Addressing voice recording replications for PD detection. *Expert Systems with Applications, 46*, 286–292.

Nawir, M., Amir, A., Yaakob, N., & Lynn, O. B. (2019). Effective and efficient network anomaly detection system using machine learning algorithm. *Bulletin of Electrical Engineering and Informatics, 8*(1), 46–51.

Parisi, L., Ravi Chandran, N., & Manaog, M. L. (2018). Feature-driven machine learning to improve early diagnosis of Parkinson's disease. *Expert Systems with Applications, 110*, 182–190.

Rana, A. Q., Ahmed, U. S., Chaudry, Z. M., & Vasan, S. (2015). Parkinson's disease: A review of non-motor symptoms. *Expert Review of Neurotherapeutics, 15*(5), 549–562.

Sadek, R. M., Mohammed, S. A., Abunbehan, A. R. K., Ghattas, A. K. H. A., Badawi, M. R., et al. (2019). PD prediction using artificial neural network. *International Journal of Academic Health and Medical Research, 3*(1), 1–8.

Sakar, C. O., & Kursun, O. (2010). Telediagnosis of PD using measurements of dysphonia. *Journal of Medical Systems, 34*(4), 591–599.

Sakar, C. O., Serbes, G., Gunduz, A., Tunc, H. C., Nizam, H., et al. (2019). A comparative analysis of speech signal processing algorithms for PD classification and the use of the tunable Q-factor wavelet transform. *Applied Soft Computing, 74*, 255–263.

Shanmugam, M., Nehru, S., & Shanmugam, S. (2018). A wearable embedded device for chronic low back patients to track lumbar spine position. *Biomedical Research, 29*, S118–S123.

Sharma, A., & Giri, R. N. (2014). Automatic recognition of PD via artificial neural network and support vector machine. *International Journal of Innovative Technology and Exploring Engineering, 4*(3), 2278–3075.

Shirvan, R. A., & Tahami, E. (2011). Voice analysis for detecting PD using genetic algorithm and KNN classification method. In *Proceedings of the 18th Iranian Conference of Biomedical Engineering (ICBME)*, Tehran, Iran (pp. 278–283). IEEE.

Vaiciukynas, E., Verikas, A., Gelzinis, A., & Bacauskiene, M. (2017). Detecting PD from sustained phonation and speech signals. *PLOS ONE, 12*(10), e0185613.

Vásquez Correa, J. C., Orozco Arroyave, J. R., & Nöth, E. (2017). Convolutional neural network to model articulation impairments in patients with Parkinson's disease. In *INTERSPEECH* (pp. 314–318). Stockholm, Sweden.

Wang, C., Deng, C., & Wang, S. (2020). Imbalance-XGBoost: Leveraging weighted and focal losses for binary label-imbalanced classification with XGBoost. *Pattern Recognition Letters, 136*, 190–197.

Wirdefeldt, K., Adami, H. O., Cole, P., Trichopoulos, D., & Mandel, J. (2011). Epidemiology and etiology of Parkinson's disease: A review of the evidence. *European Journal of Epidemiology, 26*, 1–58.

Repudi Pitchiah
Dept. of Computer Science Engineering,
Annamalai University,
Chidambaram, Tamil Nadu
pitchaiah99@gmail.com
ORCID 0000-0002-3092-914X

T. Sasi Rooba
Dept of Computer Science; Engineering,
Annamalai University,
Chidambaram, Tamil Nadu
sasiruba@gmail.com
ORCID 0000-0001-9544-4496

Uma Pavan Kumar
Dept of Computer Science and Engineering,
MRIT,
Hyderabad,
Telangana
dr.kethavarapu@gmail.com
ORCID 0000-0003-1448-685X

Appendix

Modified African Buffalo Optimization (MABO)

1. Set the algorithm control parameters and objective function(s).
2. Create a random but workable population of buffaloes and distribute them at random around the search space.
3. Compute each buffalo's fitness value.
4. Sort the buffalo population based on each individual's fitness values.
5. Update the location and fitness of buffalo i; where $i \in$ {Sorted upper half buffalo

Population, $i = 1, 2, 3, \ldots, \frac{n_p}{2}$ } by employing

$$m_{i+1} = m_i + lp_1(h_{best} - w_i) + lp_2(s_{best,i} - w_i)$$

Here, m_i and w_i indicates the exploitation and exploration moves of i^{th} buffalo $(i = 1, 2, 3, \ldots, n_p)$, respectively. lp_1 and lp_2 represents the learning factors varying from 0.1 to 0.6. Moreover, h_{best} and $s_{best,i}$ denotes the best fitness value of the herd and the individual best fitness of buffalo i respectively.

$$w_{i+1} = \frac{w_i + m_i}{\pm\ 0.5}$$

5. Now, update the fitness and buffalo location j; where, $j \in$ {lower half of the sorted

population $j = (\frac{n_p}{2} + 1), \ldots, n_p$ } as follows.

Generate a random number, $r_1 \in [0,1]$,

a) If $r_1 \geq 0.5$ then update the buffalo fitness j, from the lower half population, as suggested below

$$m_{j+1} = m_j + lp_1(h_{best} - w_r) + lp_2(s_{best,j} - w_r)$$

here, r indicates the randomly chosen buffalo from the upper half population $r \in \{i = 1, 2, 3, \ldots, \frac{n_p}{2}\}$, termed as local legislators. The above equation's recommended alteration will make it easier for the buffaloes in the upper half of the population to lead those in the bottom half of the population. The change will also aid the herd's buffaloes who are looking for assistance.

b) Update the jth buffalo location $w_{i+1} = \frac{w_i + m_i}{\pm\ 0.5}$, as follows

$$w_{j+1} = \frac{w_j + m_j}{\pm 0.5} \quad \forall j = (\frac{n_p}{2} + 1), \ldots, n_p$$

c) If $r_1 \leq 0.5$ then randomly update the location of buffalo j in the search space, as recommended in the equation given.

$$w_{j+1} = b_{min} + (b_{max} - b_{min})r_2$$

here, b_{min}, b_{max} ,and $r_2 \in [0,1]$ are the minimum and maximum permissible location limits of buffaloes, and random numbers respectively.

6. Repeat steps 3 to 6 until the stopping criteria" are attained.