International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 / https://doi.org/10.14445/22315381/IJETT-V70I9P222

Volume 70 Issue 9, 214-221, September 2022

Original Article

Parkinson's Syndrome Diagnosis Applying Perceptual Linear Prediction Cepstral Analysis on Several Speech Recordings

Sara Sandabad[1], Achraf Benba[2], Hasna Nhaila[3]

[1] *Laboratory of Control and Mechanical Characterization of Materials and Structures, ENSEM, Hassan II University, Casablanca, Morocco.*
[2] *Electronic Systems Sensors and Nanobiotechnologies team, ENSAM, Mohammed V University in Rabat, Morocco.*
[1,3] *Electrical Engineering and Intelligent Systems, ENSET, Hassan II University, Mohammedia, Morocco.*

[1]*Corresponding Author : s.sandabad@isem.ac.ma*

Received: 27 June 2022 Revised: 16 September 2022 Accepted: 22 September 2022 Published: 30 September 2022

Abstract *- This study's objective is to distinguish between two populations: 20 patients with Parkinson's disease (PD) and 20 healthy volunteers. Each person's three sustained vowel types (/a/, /o/, and /u/) were recorded, and then cepstral analysis was performed on these voice signals. In this study, we will use Perceptual Linear Prediction (PLP) to generate voiceprint from every voice sample (PLP). In order to collect the voiceprint from every voice recording, the obtained PLP cepstral coefficients were compressed by calculating their average value. Furthermore, a classification method used in our work is obtained by combining the Support Vector Machines classifier and a com the Leave One Subject Out validation scheme (LOSO). We did an independent test to evaluate our results using another database with 28 PD. According to the research result, combining the linear kernel of SVM and LOSO on the sustained vowel /o/, the best classification accuracy on the first dataset was 80%. And adopting the hybridization of two sustained vowels, /a/ and /o/, with the MLP kernel of the SVM, the maximum classification accuracy using the independent test was 87.50 percent.*

Keywords *- Voice analysis, Parkinson's disease, Perceptual Linear Prediction, Voiceprint, Leave One Subject Out, Support Vector Machine.*

1. Introduction

Parkinson's disease (PD) is a degenerative disease caused by the gradual and cumulative degeneration of brain neurons. Actually, Alzheimer's disease is the second most prevalent neurological disease. As it progresses, PD develops a range of symptoms and damages the system to maintain the execution of learned motor plans, including walking, talking, or carrying out other routine tasks. [1] [2] [3]. Clinicians and speech pathologists' primary concerns now revolve around evaluating speech quality and diagnosing the reasons why it is degrading in the context of Parkinson's disease via phonological and auditory signals. It is well-recognized that PD often affects the voice in around 90% of patients [4] and affects adults over 50, making physical visits for diagnostics, monitoring, and therapy very problematic [5] [6]. To discriminate between various disease states in PD patients, clinicians and speech pathologists have developed techniques based on acoustic signals. In an attempt to develop more objective analyses to identify vocal abnormalities, recent research use studies of voice quality in the temporal, spectral, and cepstral domains [7]. These characteristics include harmonicity, jitter, shimmer, and the fundamental frequency of voice oscillation (FO) [1] [8] [9].

Little et al. [1] aimed to classify healthy and PD patients by identifying dysphonia in this context. For 31 participants, the sustained vowel "a" phonations were recorded, 23 of whom had PD diagnoses. Then, using a kernel Support Vector Machine, they determined four out of ten highly uncorrelated metrics that, when combined, resulted in overall correct classification performance of 91.4%. Multiple types of sound recordings taken from persons with Parkinson's disease were studied by Betul et al. [6]. By adopting a leave-one-subject-out (LOSO) cross-validation technique and summarized Leave-One-Out, the collected data were put in SVM and k-Nearest Neighbor (k-NN) classifiers for PD diagnosis. Most research employs SVM classification to discriminate between healthy patients and PWP [1] [6] [22]. True positive (TP), true negative (TN), false positive (FP), and false negative (FN) values are used to evaluate the efficiency of the diagnostic methods. In this paper, we used Perceptual Linear Prediction (PLP) to extract voiceprint from each voice sample. PLP was first presented forth by Hynek Hermansky [10] and has long been used in applications for speaker identification and recognition. In order to collect the voiceprint from each voice recording, the acquired PLP cepstral coefficients were compressed by calculating their average value. Later, a classification technique based on the LOSO validation methodology and Support Vector Machines was carried out. To evaluate our proposed

technique and validate the accuracy of our results, we calculate Accuracy, Sensitivity, Specificity, Matthews' correlation coefficient and Probability Excess.
The structure of this article is as follows: Section II describes the relevant database. Section III presents the PLP techniques. Section IV of this article presents the research's methodology. Section V presents the results that were collected, and Section VI concludes.

2. Data Acquisition

The data used in our work was acquired and used in [6]. There are 20 PD patients, combining 6 women and 14 males and 20 healthy individuals (10 women and 10 males). Patients with PD range in age from 43 to 77 (mean: 64.86, standard deviation: 8.97), and the time from diagnosis is between 0 and 6 years. Healthy individuals range in age from 45 to 83 (mean: 62.55, standard deviation: 10.79). A Trust MC-1500 microphone with a resonant frequency of 50 Hz to 13 kHz was used for all recordings. The subjects were positioned 15 cm away from the microphone set at 96 kHz, 30 dB. Each sample was recorded in stereo and saved as a WAV file.

Each of the 40 participants (20 with Parkinson's disease and 20 in great health) was asked to say the three sustained vowels /a/, /o/, and /u/ at a comfortable volume for three different types of recordings. This provided us with a database of 120 voice samples, and these samples were used for the analysis.

The same recording equipment and experts were used to build a supplementary dataset. The sustained vowels /a/ and /o/ were asked to be pronounced by 28 PD three times, and we then only chose one sample of each vowel. The age of these individuals spans from 39 to 79, and the period from diagnosis extends from 0 to 13 years (mean: 62.67, standard deviation: 10.96). This dataset was utilized to verify and test the outcomes from the initial dataset.

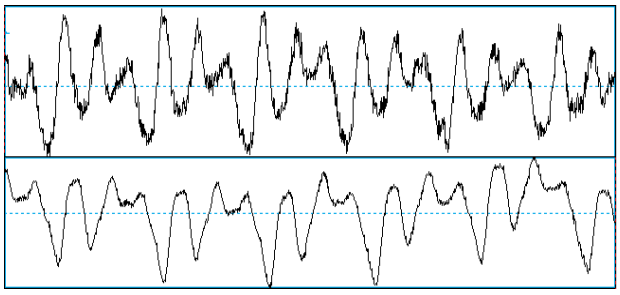

Fig. 1 Two voice signals: top a patient with Parkinson's disease and bottom a healthy person.

3. PLP Technique

Our primary objective was to convert the speech waveform into a parametric representation for advanced processing and analysis [11]. Speech is a slow time-varying signal that is referred to as quasi-stationary. It tends to be relatively stable when studied over a brief time [11]. The voice signal's waveform does, however, modify with time. As a result, a short-time spectrum analysis should be used to characterize it. [11]. Figure 1 illustrates the computation of the PLP, and the following paragraphs describe it.

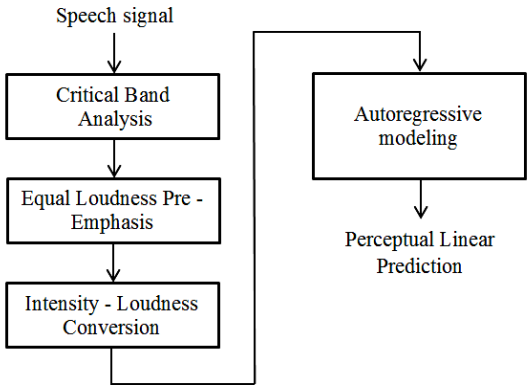

Fig. 2 Perceptual Linear Prediction PLP Coefficients Block Diagram

3.1. Spectral Analysis

The voice signal can only be treated on a finite number of samples even though it is a real signal with a definite period [12]. To do this, the speech segment is filtered by the Hamming window as the first step of the PLP process [10]. The goal is to smooth out the edges of the signal, so they connect with the origins while reducing signal discontinuities [12]. It was obtained by introducing the following formula to the samples and applying the Hamming window to taper the signal to zero at the start and finish of each frame [10]:

$$W(n) = \left\{0{,}54 - 0{,}46\ cos\left\{\frac{2\pi n}{N-1}\right\}\right\} \qquad (1)$$

N: the Hamming window's length.
Length around 20 ms

The next treatment phase requires using the Fast Fourier Transform (FFT) to transform each frame of N samples from the time domain into the frequency domain [11]. We would use the FFT since implementing the Discrete Fourier Transform (DFT) is easy and fast [11].

As is common knowledge, the DFT is formulated as having on the set of N samples (Sn). [11]:

$$S_n = \sum_{k=0}^{N-1} s_k e^{-2\pi jkn/N} \qquad (2)$$

$$n = 0,1,2,3, \dots \dots, N-2, N-1$$

The short-term power band is determined by multiplying the squares of the short-term speech spectrum's real and imaginary components, as follows [10]:

$$P(\omega) = Re[S(\omega)]^2 + Im[S(\omega)]^2 \qquad (3)$$

3.2. Critical-band Spectral Resolution

The following equations warp the short-term power spectrum *P(ω)* into Bark frequency **Ω** [10]:

$$\Omega(\omega) = 6 \ln\left\{\frac{\omega}{1200\pi} + \sqrt{\frac{\omega}{1200\pi}^{2} + 1}\right\} \quad (4)$$

$$\Omega(f) = 6 \ln\left\{\frac{f}{600} + \sqrt{\frac{f}{600}^{2} + 1}\right\} \quad (5)$$

$$\Omega(f) = 6\, sinh^{-1}\left(\frac{f}{600}\right) \quad (6)$$

ω : angular frequency [rad/s],
f : frequency [Hz].

The purpose of the subsequent step is to convolve the output warped power with the power spectrum of the simulated critical-band masking curve Ψ(Ω) Hynek Hermansky [10] :

$$\Psi(\mathrm{C}) = \begin{cases} 0 & \text{when} \quad \Omega < -1{,}3 \\ 10^{2{,}5(\Omega+0{,}5)} & \text{when} \quad -1{,}3 \le \Omega \le -0{,}5 \\ 1 & \text{when} \quad -0{,}5 \le \Omega \le 0{,}5 \\ 10^{-1(\Omega-0{,}5)} & \text{when} \quad 0{,}5 \le \Omega \le 2{,}5 \\ 0 & \text{when} \quad \Omega > 2{,}5 \end{cases} \quad (7)$$

The structure of acoustic filters is approximated here in a very simple way.

Then doing the discrete convolution of Ψ(Ω) with P(ω), by using the following equation, samples of the critical-band power spectrum are generated [10]:

$$\theta(\Omega_i) = \sum_{\Omega=-1,3}^{2,5} p(\Omega - \Omega_i)\Psi(\Omega) \quad (8)$$

In comparison to the original *P(ω)*, the spectral resolution $\theta(\Omega)$ of the convolution between the rather large critical-band masking curve Ψ(Ω) and the short-term power spectrum *P(ω)* is even less.

3.3. Equal-loudness Preemphasis

The next procedure in this approach is to pre-emphasize the samples Θ[Ω(ω)] by adopting the following equation to the simulated equal-loudness curve [10]:

$$\hat{\theta} = \Theta[\Omega(\omega)] \times \mathrm{E}(\omega) \quad (9)$$

E(ω) approximates the human ear's non-equal sensitivity to perception at variable frequency. The practical approximation utilized in this study was first suggested by Makhol and Cosell [13] and taken by Hynek-Hermansky [10]. It is represented by the equation below:

$$E(\omega) = \frac{(\omega^2 + 56{,}8\ 10^6)\omega^4}{(\omega^2 + 6{,}3\ 10^6)^2 \times (\omega^2 + 0{,}38\ 10^9)} \quad (10)$$

$$E(f) = \left[\frac{f^2}{f^2 + 1{,}6\ 10^5}\right]^2 \times \left[\frac{f^2 + 1{,}44\ 10^6}{f^2 + 9{,}6\ 10^6}\right] \quad (11)$$

3.4. Intensity-loudness Power Law

The cubic-root amplitude compression comes last before the all-pole modeling. The next formula models the non-linear connection between sound intensity and perceived loudness and approaches the power law of sound waves.

$$\Phi(\Omega) = \hat{\theta}^{0,33} \quad (12)$$

3.5. Autoregressive Modeling

Perceptual Linear Prediction (PLP) [10] [14] is a technique where an all-pole spectrum models the signal spectrum. It is the final step of the Perceptual Linear Prediction process, where Φ(Ω) is approximated by the spectrum of an all-pole model using the autocorrelation method of all-pole spectral modeling.

In this paper, we generated the autoregressive model using spectral magnitude samples adopting the Linear Predictive Coefficient (LPC) algorithm. By transforming the LPC of 'n' coefficients into frames of cepstra, the autoregressive coefficients are translated to cepstral coefficients of the all-pole model.

3.6. Liftering

Their main benefit is cepstral coefficients have an uncorrelated structure [12] [16]. However, as seen in Figure 3, their challenge is that the cepstral coefficients of higher order are comparatively modest [12] [16]. As a result, it is crucial to rescale these cepstral coefficients so that their magnitudes are relatively similar seen in Figure 4. The following equation was used to generate the cepstral coefficients to do this.

$$c_n{}' = \left(1 + \frac{L}{2}\, sin\left(\frac{\pi n}{L}\right)\right) \times c_n \quad (13)$$

L : Cepstral sine lifter parameter (we use L=22) [16].

4. Methodology

Building a dataset containing voice samples of healthy individuals and PD patients was the initial stage of this work. Finally, we made use of PD datasets provided in [6]. 17 voices from each group are included in this collection, giving us 34 records. All participants were asked to comfortably articulate the sustained vowel (/a) [16].

The PLP's multi-cepstral coefficients were then retrieved from each voice sample. The extracted coefficients were in the range of 2 to 20. We determined the ideal number of coefficients required for the best classification accuracy.

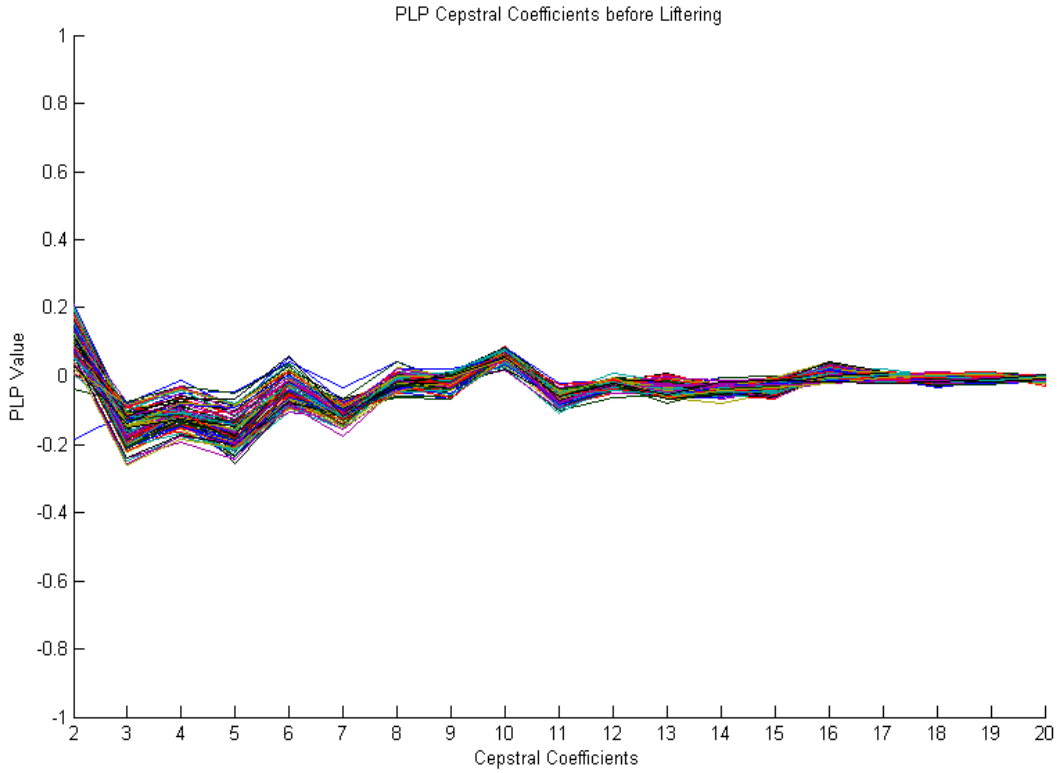

Fig. 3 Despite considering the initial cepstral coefficient before Liftering, the first 20 Perceptual Linear Prediction Cepstral Coefficients were retrieved from PD patients.

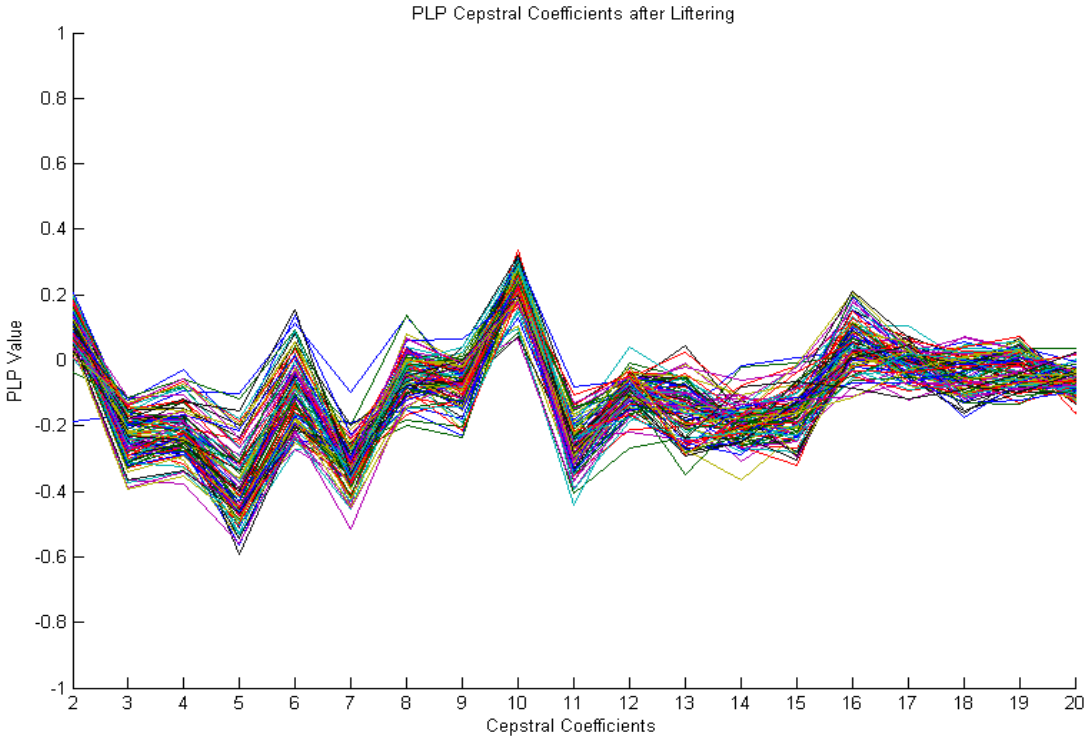

Fig. 4 Without using the first cepstral coefficient after Liftering, the first 20 Perceptual Linear Prediction Cepstral Coefficients were obtained from PD patients

We collected the PLP cepstral coefficients from each voice sample to obtain each individual's voiceprint. The PLP has many frames, which takes a long time to classify and makes it impossible to make the proper diagnostic determination. To solve this issue, we computed the average value of these frames to obtain the voiceprint of each individual, as shown in Fig. 5.

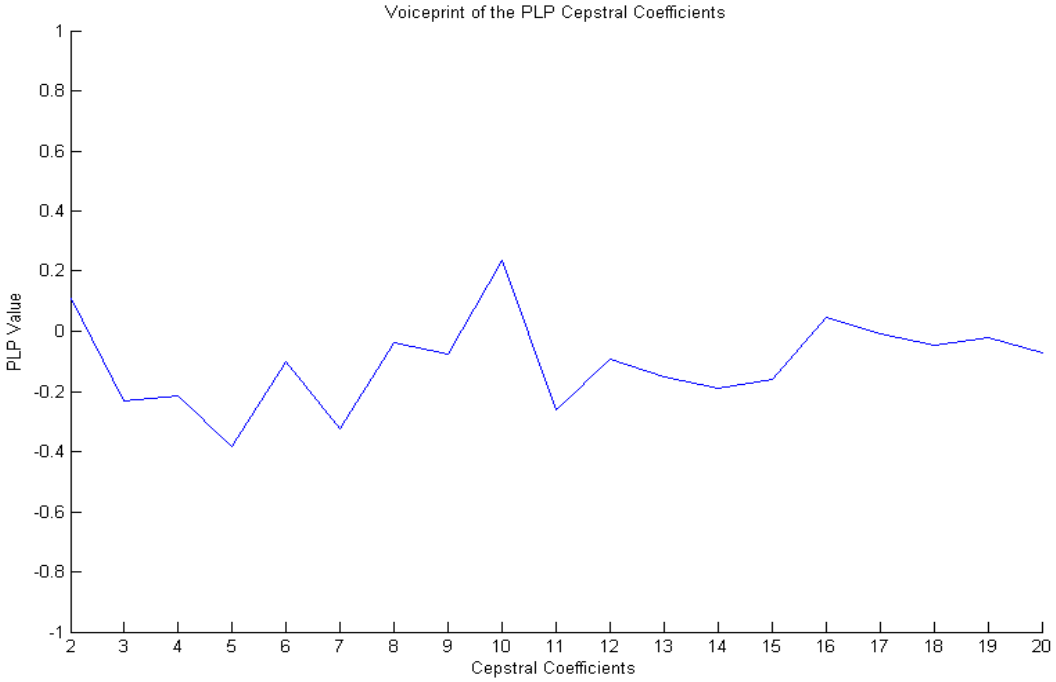

Fig. 5 The first 20 perceptual linear predictions' voiceprints Despite applying the initial cepstral coefficient, cepstral coefficients were extracted from PD patients

We used a classification process known as LOSO to create and verify our classifier. With this technique, we trained a classifier on the remaining compressed frames of other persons while leaving out all the compressed frames of the PLP cepstral coefficients of one person to be used for validation as if it were an unobserved person [6]. Up until all 20 coefficients per subject, we employed the LOSO iteratively for each coefficient per subject [16]. In our case, the RBF, Linear, and Polynomial kernel types of the SVM classifier were used.

4.1. Evaluation Metrics

We assessed the accuracy, sensitivity, and specificity to assess how well our classifiers distinguished between PD patients and healthy individuals [16] [17].

$$Spe. = \frac{TN}{TN + FP} \tag{14}$$

$$Sen. = \frac{TP}{TP + FN} \tag{15}$$

$$Acc. = \frac{TP + TN}{TP + TN + FP + FN} \tag{16}$$

Sensitivity is how well the classifier detects healthy individuals, whereas specificity measures how well it detects other PD patients. Accuracy measures how well the classifier distinguishes between the two categories. [18] [19] [20] [21].

FP stands for false positive: PD patients who were mistakenly diagnosed, TP stands for true positive: healthy people who were correctly classified, TN stands for the true negative: PD patients who were correctly classified, and FN stands for false positive: healthy people who were incorrectly classified.

Matthews' correlation coefficient MCC, which illustrates the effectiveness of binary classification in machine learning, and the Probability Excess PE will be implemented to get a good evaluation of our results.

$$MCC = \frac{(TP \times TN) - (FN \times FP)}{\sqrt{(FN + TP)(FP + TN)(FP + TP)(FN + TN)}} \tag{17}$$

$$PE = \frac{(TP \times TN) - (FN \times FP)}{(FN + TP)(FP + TN)} \tag{18}$$

5. Results

Each type of voice sample from the first dataset, which contains the sustained vowels /a/, /o/, and /u/, was tested separately at the initial stage of this investigation. Each speech sample's voiceprint was extracted for these tests. Then, using LOSO and SVM, we evaluated the results with Accuracy, Sensitivity, Specificity, MCC, and PE.

Table 1 displays the classification accuracy for the database's 20 patients with Parkinson's disease and 20 healthy individuals using solely sustained vowels (/a/). As can be seen, a linear SVM kernel with the top 12 PLP cepstral coefficients was used to reach a maximum classification accuracy of 75%.

Results for specificity represent the categorization metrics for people with Parkinson's disease;
Results for sensitivity represent the classification metrics for healthy people.

These findings lead us to the conclusion that 30 individuals were correctly identified, while 10 persons were incorrectly classified. Maximum sensitivity and specificity results of 75% were obtained using identical conditions.

These findings lead us to the conclusion that the classification of 15 PD patients and 15 healthy individuals were accurate. Using the SVM kernel as well, a maximum MCC and PE of 0.5 were attained.

Table 1. Results of vowel /a/

SVM Kernel	Specificity	Sensitivity	Accuracy	MCC	PE	Coef
MLP	70.00 %	70.00 %	70.00 %	0.4001	0.4000	4
Linear	75.00 %	75.00 %	75.00 %	0.5010	0.5009	12
Polynomial	65.00 %	80.00 %	72.50 %	0.4551	0.4500	12
RBF	60.00 %	80.00 %	70.00 %	0.4082	0.4000	3

The classification accuracy utilizing the sustained vowel /u/ from the same dataset is shown in Table 2. As can be observed, a linear kernel of SVM with the first 2PLP cepstral coefficients was used to reach a maximum classification accuracy of 77.50 percent.

Using the same conditions, a maximum sensitivity result of 85% and a specificity result of 70% were also attained.

These findings lead us to the conclusion that 14 PD patients, 17 healthy individuals, and 9 more participants were incorrectly diagnosed. Using the same SVM kernel, a maximum MCC and PE of 0.55 were attained.

Table 2. Results of vowel /u/

SVM Kernel	Specificity	Sensitivity	Accuracy	MCC	PE	Coef
MLP	55.00 %	95.00 %	75.00 %	0.5545	0.5000	2
Linear	70.00 %	85.00 %	77.50 %	0.5563	0.5500	2
Polynomial	70.00 %	75.00 %	72.50 %	0.4506	0.4500	9,10
RBF	60.00 %	90.00 %	75.00 %	0.5241	0.5000	2

Table 3 summarizes the research findings. The best classification accuracy was 80%. The linear SVM kernel with the initial 7 PLP cepstral coefficients was used to produce this result.

In the same conditions, a maximum sensitivity of 75% and a maximum specificity of 85% were attained.
These findings lead us to conclude that 8 participants were incorrectly classified, while 17 PD patients and 15 healthy individuals were accurately identified.

Using the SVM kernel, a maximum MCC 0.6030 and PE 0.6 were attained.

Sustained vowel /o/ voice samples give the highest discriminative information compared to other voice samples.

Table 3. Results of vowel /o/

SVM Kernel	Specificity	Sensitivity	Accuracy	MCC	PE	Coef
MLP	70.00 %	75.00 %	77.50 %	0.4020	0.4000	3, 13
Linear	85.00 %	75.00 %	80.00 %	0.6030	0.6000	7
Polynomial	80.00 %	60.00 %	70.00 %	0.4082	0.4000	7
RBF	65.00 %	85.00 %	70.00 %	0.5563	0.5500	2

At such a step, we needed to check our classifiers applying the dataset's different speech recordings simultaneously. These recordings include the sustained vowels /a/, /o/, and /u/ from 40 individuals, giving us 120 voice samples in all (3*40).

Table 4 displays the results from all multiple sorts of voice recordings. As can be observed, the greatest accuracy of 78.33 percent was obtained using the linear SVM kernel and the first 12PLP cepstral coefficients.

Using the same kernel, a maximum value of 0.5667 for Matthews' correlation and Probability Excess was discovered.

The highest Sensitivity and Specificity of 78.33 percent, achieved using the same settings, indicate that 26 participants were incorrectly classified, leaving 47 PD patients and 47 healthy individuals to be accurately identified.

Table 4. Results of vowel /a/, /o/ and /u/

SVM Kernel	Specificity	Sensitivity	Accuracy	MCC	PE	Coef
MLP	66.67 %	73.33 %	70.00 %	0.4009	0.4000	7
Linear	78.33 %	78.33 %	78.33 %	0.5667	0.5667	12
Polynomial	71.67 %	70.00 %	70.83 %	0.4167	0.4167	7
RBF	55.00 %	85.00 %	70.00 %	0.4193	0.4000	3

We used the independent dataset containing 28 PD patients to test and validate our findings after performing the initial test on two participant groups (patients with PD and healthy individuals).

First, we used sustained vowels from 20 PD patients and 20 healthy individuals to train our classifiers. The sustained vowels from the independent dataset were then used to test the findings. The best classification accuracy, as shown in Table 5, was achieved utilizing the Polynomial kernel of SVM and the top three PLP cepstral coefficients. It indicates that 9 additional PD patients were incorrectly classified, while 19 PD patients were effectively classified.

Table 5. Test Results of vowel /a/

SVM Kernel	Accuracy	Coefficients
MLP	35.71 %	19
Linear	7.143 %	12 - 20
Polynomial	67.85 %	3
RBF	7.143 %	9, 10, 11

Similarly, we used sustained vowels from 20 PD patients and 20 healthy individuals to train our classifiers. The independent dataset's sustained vowels (/o/) were used to test the findings. Table 6 shows that utilizing the MLP kernel of SVM and the first 8PLP cepstral coefficients, the best classification accuracy was 53.57 percent. It indicates that 13 additional PD patients were incorrectly classified, while 15 PD patients were effectively classified.

Table 6. Test Results of vowel /o/

SVM Kernel	Accuracy	Coefficients
MLP	53.57 %	8
Linear	35.71 %	16 - 19
Polynomial	35.71 %	16
RBF	35.71 %	6

The results employing the hybridization of the independent dataset's two vowels, /a/ and /o/, are shown in Table 7. As shown, utilizing the first 4PLP cepstral coefficients with the MLP kernel of SVM, the greatest accuracy obtained was 87.50 percent.

This result indicates that 49 PD patients were accurately classified, while only 7 individuals were incorrectly classified.

Table 7. Test Results of vowel /a/ and /o/

SVM Kernel	Accuracy	Coefficients
MLP	87.50 %	4
Linear	1.786 %	4
Polynomial	42.86 %	3
RBF	3.571 %	11

6. Conclusion

Voice troubles and PD symptoms do not happen suddenly. The early phases of this gradual process could go unrecognized. So we gathered a variety of voice recordings from various people as they pronounced the sustained vowels /a/, /o/, and /u/ in order to improve the assessment of PD. We compressed the frames of the PLP cepstral coefficients by determining their average value to recover each individual's voiceprint.

It has been proven to be a good criterion for diagnosing voice dysfunction in the context of PD to use voiceprint to differentiate between PD patients and healthy individuals. Our findings led us to the conclusion that compared to other voice recordings, prolonged vowel /o/ contains more discriminative analysis. Using the initial and independent datasets, the linear kernel of SVM was the most successful kernel in revealing the best classification accuracy.

References

[1] Little, And Al, "Suitability of Dysphonia Measurements for Telemonitoring of Parkinson's Disease," *Biomedical Engineering, IEEE Transactions on* , vol.56, no.4, pp.1015-1022, 2009.

[2] L. Ishihara, And C. Brayne, "A Systematic Review of Depression and Mental Illness Preceding Parkinson's Disease," *Acta Neurologica Scandinavica* , vol.113, no.4, pp.211-220, 2009.

[3] J. Jankovic, "Parkinson's Disease: Clinical Features And Diagnosis," *Journal of Neurology, Neurosurgery & Psychiatry*, vol.79, no.4 , pp. 368-376, 2008.

[4] S. B. O'sullivan, T. J. Schmitz, "Parkinson Disease," *Physical Rehabilitation, 5th Ed.* Philadelphia, Pa, Usa: F. A. Davis Company, pp. 856–894, 2007.

[5] Huse, Daniel M., Et Al, "Burden of Illness in Parkinson's Disease," *Movement Disorders* , vol.20, no.11, pp.1449-1454, 2005.

[6] S. Betulerdogdu, Et Al, "Collection and Analysis of a Parkinson Speech Dataset With Multiple Types of Sound Recordings," *Biomedical And Health Informatics, IEEE Journal of* , vol.17, no.4, pp.828-834, 2013.

[7] U. K. Rani, M.S. Holi, "Automatic Detection of Neurological Disordered Voices Using Mel Cepstral Coefficients and Neural Networks," *2013 IEEE Point-of-Care Healthcare Technologies (PHT)*, Bangalore, India, 2013.

[8] M. A. Little, P. E. Mcsharry, S. J. Roberts, D. A. Costello, I. M. Moroz, "Exploiting Non-Linear Recurrence and Fractal Scaling Properties for Voice Disorder Detection," *Biomedical Engineering Online*, 2007.

[9] D. A. Rahn, M. Chou, J. J. Jiang, Y. Zhang, "Phonatory Impairment In Parkinson's Disease: Evidence from Non-Linear Dynamic Analysis And Perturbation Analysis," *Journal of Voice.* vol.21, pp.64-71, 2007.

[10] H. Hynek, “Perceptual Linear Predictive (PLP) Analysis of Speech,” *The Journal of the Acoustical Society of America* , vol.87, no.4 , pp.1738-1752, 1990.

[11] Ch. S. Kumar, P. R. Mallikarjuna, "Design of an Automatic Speaker Recognition System Using Mfcc, Vector Quantization And Lbg Algorithm," *International Journal on Computer Science And Engineering*, vol.3, no.8, 2011.

[12] S. Young, G. Evermann, T. Hain, D. Kershaw, X. Liu, G. Moore, J. Odell, D. Ollason, D. Povey, V. Valtchev, P. Woodland, "The Htk Book (For Htk Version 3.4)," *Cambridge University Engineering Department*, Copyright. 2001-2006.

[13] J. Makhoul, And L. Cosell, “Lpcw: An Lpc Vocoder With Linear Predictive Spectral Warping,” *Acoustics, Speech, And Signal Processing, IEEE International Conference on Icassp'76*, IEEE vol. 1, 1976.

[14] J. Makhoul, "Spectral Linear Prediction: Properties and Applications," *Acoustics, Speech And Signal Processing, Ieee Transactions on* , vol.23, no.3, pp. 283-296, 1975.

[15] Ayush Jain, Prathamesh Patil, Ganesh Masud, Prof. Sunantha Krishnan, Prof. Vijaya Bharathi Jagan, “Detection of Sarcasm Through Tone Analysis on Video and Audio Files: A Comparative Study on Ai Models Performance," *SSRG International Journal of Computer Science And Engineering*, vol. 8, no. 12, pp. 1-5, 2021. Crossref, https://doi.org/10.14445/23488387/Ijcse-V8i12p101

[16] A. Benba, A. Jilbab And A. Hammouch, "Voice Analysis for Detecting Persons with Parkinson's Disease Using MFCC and VQ," *The 2014 International Conference on Circuits, Systems and Signal Processing*, Saint Petersburg State Polytechnic University, Saint Petersburg, Russia, September 23-25, 2014.

[17] A. Benba, A. Jilbab And A. Hammouch, "Hybridization of Best Acoustic Cues For Detecting Persons With Parkinson's Disease," *2nd World Conference on Complex System (Wccs'14), IEEE*, Agadir November 10-12, 2014.

[18] A. Benba, A. Jilbab And A. Hammouch, “Voiceprint Analysis Using Perceptual Linear Prediction And Support Vector Machines For Detecting Persons With Parkinson's Disease,” *The 3rd International Conference on Health Science And Biomedical Systems (Hsbs '14)* Florence, Italy, November 22-24, 2014.

[19] A. Benba, A. Jilbab, And A. Hammouch, "Voice Analysis for Detecting Persons With Parkinson's Disease Using PLP and VQ," *Journal of Theoretical and Applied Information Technology* , vol.70, no.3, 2014.

[20] A. Benba, A. Jilbab, And A. Hammouch, “Detecting Patients With Parkinson's Disease Using PLP and VQ,” The *7th International Conference on Information Technology*, 12-15 Mai, Amman, Jordan, 2015.

[21] A. Benba, A. Jilbab, And A. Hammouch, “Detecting Patients With Parkinson's Disease Using Mel Frequency Cepstral Coefficients And Support Vector Machines,” *International Journal on Electrical Engineering and Informatics* , vol.2, no.7, pp. 297 – 307, 2015.

[22] S. Malathi, D. Kannan, "Prediction of Autism Spectrum Disorder Using Reliable Ant Colony Optimisation Based Relevant Vector Machine," *International Journal of Engineering Trends and Technology*, vol. 70, no. 9, pp. 57-63, 2022. Crossref, https://doi.org/10.14445/22315381/IJETT-V70I9P206.

[23] J. Martinez, H. Perez, E. Escamilla, M. M. Suzuki, "Speaker Recognition Using Mel Frequency Cepstral Coefficients (MFCC) And Vactor Quantization (VQ) Techniques," *IEEE Electrical Communications and Computers*, Cholula, Puebla, pp. 248-251, 2012.