IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 22 November 2023, accepted 15 December 2023, date of publication 18 December 2023,
date of current version 27 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3344464

RESEARCH ARTICLE

Voice-Based SVM Model Reliability for Identifying Parkinson's Disease

NEMUEL D. PAH[1,2], (Senior Member, IEEE), VERONICA INDRAWATI[1], (Member, IEEE), AND DINESH K. KUMAR[2], (Senior Member, IEEE)
[1]Electrical Engineering Department, University of Surabaya, Surabaya 60293, Indonesia
[2]School of Engineering, RMIT University, Melbourne, VIC 3000, Australia

Corresponding author: Nemuel D. Pah (nemuelpah@staff.ubaya.ac.id)

This work was supported by Penelitian Dasar Unggulan Perguruan Tinggi (PDUPT) Kemendikbudristek 015/SP-Lit/LPPM-01/KemendikbudRistek/FT/V/2023 Research Grant.

This work involved human subjects or animals in its research. The authors confirm that all human/animal subject research procedures and protocols are exempt from review board approval.

ABSTRACT One of the possibilities for developing non-invasive computerized diagnostic tools for Parkinson's disease (PD) is to detect changes in the voice, known as Parkinsonian dysarthria. Numerous classification models have been developed to diagnose PD based on voice features. However, the performance of models developed and trained only using voice features extracted from people with PD and healthy people might be affected when tested on individuals with other voice-related pathological conditions. Therefore, we investigated the reliability of voice-based machine-learning models that were developed only using datasets of people with PD and healthy people for accurately identifying people without PD when they have other voice-related pathological conditions (i.e. dysphonia and laryngitis). Three different support vector machines (SVMs) were developed and tested on voice features extracted from healthy people and those with PD, dysphonia, and laryngitis. The results confirmed that a voice-based SVM classifier only trained on the dataset of people with PD and healthy people was equally reliable in classifying other voice-related pathological conditions, such as dysphonia and laryngitis, as non-PD cases.

INDEX TERMS Dysphonia, laryngitis, Parkinson's disease, support vector machine, voice features.

I. INTRODUCTION

Parkinson's disease (PD) ranks as the second most prevalent neurodegenerative condition following Alzheimer's disease [1], and its incidence is expected to increase owing to the aging population. PD is characterized by a range of motor and non-motor deficits [2]. The prevailing method for PD diagnosis involves a clinical assessment employing either the Unified Parkinson Disease Rating Scale (UPDRS) [3] or the Hoehn and Yahr (H&Y) scale [4]. These assessments examine the presence of symptoms such as tremors, rigidity, bradykinesia, or postural impairment, as well as non-motor symptoms such as dysarthria, functional limitations, and cognitive decline [5].

To create portable, automated, and intelligent diagnostic tools for PD, it is essential to rely on non-invasive PD biomarkers. One of the early indicators of PD is a change in speech known as Parkinsonian hypokinetic dysarthria [6]. This symptom is present in approximately 90% of PD patients and is characterized by reduced voice intensity, increased voice nasality, heightened acoustic noise, alterations in speech prosody, imprecise articulation, a narrower pitch range, monotonous loudness, longer pauses, vocal tremors, harsh and breathy voice quality, and speech disfluency [7].

Voice features extracted from sustained vowels have been evaluated as potential parameters for both the diagnosis and monitoring of PD in several studies [8], [9], [10], [11], [12], [13], [14], [15]. These features cover various aspects, including issues related to glottal vibration, the harmonics-to-noise ratio (HNR), the control of glottal pressure through

The associate editor coordinating the review of this manuscript and approving it for publication was Jon Atli Benediktsson.

the respiratory mechanism, and vocal tract control [16]. The effectiveness of these voice features has been assessed through a range of statistical analyses, including calculations of mean values, standard deviations, effect sizes, and p-values derived from statistical tests such as t-tests and the analysis of variance (ANOVA).

Many classification models have been developed to diagnose PD [17]. Pah et al. [16], [18] and Motin et al. [19] developed support vector machine (SVM) models to identify people with PD and distinguish the effect of levodopa in PD patients. Ali et al. [20] developed an intelligent system that uses linear discriminant analysis (LDA) for dimensionality reduction and a genetic algorithm (GA) to detect people with PD. Sakar et al. [11], [21] investigated machine-learning and wavelet-based models to diagnose PD. Ensemble random forest and SVM were developed by Tsanas et al. [22] to classify people with PD, obtaining significant accuracy.

The previous classification models aimed to distinguish people with PD from people without PD. However, the models were developed and trained on voice features extracted from people with PD and people without any voice-related pathological condition. The performance of the developed models was not tested considering people with other voice-related pathological conditions. Thus, the models may identify people with other voice-related pathological conditions as PD. Dysphonia [23], [24] is a pathological condition that alters the voice features. Teixeira and Fernandes [25] analyzed the voice features of people with four types of dysphonia and concluded that there were particular changes in the voice parameters such as jitter and shimmer due to dysphonia. In the other work, Teixeira et al. [26] analyzed commonly used voice features, such as jitter, shimmer, harmonic-to-noise ratio (HNR), noise-to-harmonic ratio (NHR), and autocorrelation, extracted from the sound of sustained vowels (/a/, /i/ and /u/) of people with chronic laryngitis to identified statistical differences due to the pathological condition.

The present study aimed to investigate the reliability of voice-based machine-learning models that were developed only using datasets of people with PD and healthy people for accurately identifying people who have other voice-related pathological conditions (i.e., dysphonia and laryngitis).

II. METHODS

A. PATHOLOGICAL CONDITIONS

In addition to PD patients and healthy subjects, two pathological conditions were included in this work, namely dysphonia [23], [24], and laryngitis [27]. These pathological conditions were selected because of their close relation to voice production mechanisms.

Dysphonia is a medical condition characterized by abnormal alterations in vocal quality, pitch, loudness, or vocal effort. This condition can lead to difficulties in communication and have a detrimental impact on an individual's overall voice-related quality of life [23]. The root cause of dysphonia can be attributed to irregular vocal fold oscillations, which may stem from factors such as abnormalities in muscle tone, hypertonicity, incomplete closure of the glottis during speaking, changes in vocal fold mass, or the presence of a vocal fold lesion or tumor [24], [28].

Laryngitis [27] is a medical condition characterized by inflammation of the larynx. It can present in different forms, either acute or chronic, and may have infectious or inflammatory causes. Laryngitis can occur independently or as a component of a broader systemic ailment. Hoarseness is a common symptom associated with laryngitis. This condition is frequently linked to upper respiratory tract infections and can have a substantial impact on an individual's physical health, quality of life, and even psychological well-being and ability to work if the symptoms persist.

TABLE 1. Participants' demographics.

Pathology	Dataset	# Subject (M/F)	Age (mean ± SD)
PD	PC-GITA	25/25	61.02 ± 9.44
Dysphonia	SVD	22/22	62.32 ± 8.30
Laryngitis	SVD	28/16	60.48 ± 6.08
HC	PC-GITA	25/25	60.98 ± 9.46
	SVD	17/25	61.74 ± 6.63
p-value of ANOVA			0.848

TABLE 2. SNR of PC-GITA and SVD dataset.

Phoneme	SNR in dB (mean ± SD)		p-value of t-test
	PC-GITA	SVD	
/a/	29.12 ± 11.68	32.31 ± 2.66	0.087
/i/	26.02 ± 10.73	28.73 ± 3.08	0.117
/u/	27.66 ± 11.58	29.25 ± 3.70	0.395

B. DATASETS

The voice features analyzed in this work were extracted from two datasets, namely the PC-GITA dataset provided by Orozco-Arroyave et al. [29] and the Saarbrücken Voice Database (SVD) [30], [31]. Both datasets were generated following approval from their respective ethics committees, ensuring compliance with the principles outlined in the Helsinki Declaration.

The PC-GITA dataset consists of recordings of sustained vowels (/a/, /e/, /i/, /o/, and /u/) obtained from 100 native Colombian-Spanish speakers. Among these participants, 50 were diagnosed with PD, whereas the remaining 50 served as healthy control (HC) subjects. The control group was matched with the PD group in terms of age and gender. These voice recordings were conducted in a noise-controlled soundproof booth and saved as.wav files with a sampling rate of 44.1 kHz and 16-bit data resolution. Each subject provided three repetitions of each sustained vowel, but for our analysis, we only used one recording from each participant.

The SVD is a freely downloadable database provided by the Institute of Phonetics at Saarland University in

Saarbrücken, Germany [30], [31]. The database contains voice recordings from more than 2000 people with 71 different pathological conditions, including recordings from healthy subjects. The dataset includes the recordings of sustained vowels (/a/, /i/, and /u/) at normal, high, and low pitch, and the recordings of a sentence in Germany. The vowels were recorded as.wav files in a noise noise-controlled environment for 1 to 3 seconds at a 50 kHz sampling rate with 16-bit resolution. For the present work, only recordings of the sustained vowels at normal pitch were considered.

The recordings of subjects with dysphonia and laryngitis were obtained from the SVD dataset. The sustained vowel recordings of individuals with PD were sourced from the PC-GITA dataset, owing to the limited availability of PD recordings in the SVD dataset. To ensure dataset compatibility, the SVD dataset was restricted to individuals between the ages of 51 and 90 years. The demographic information for the participants included in this study is presented in Table 1. ANOVA was conducted to confirm the similarity between the age ranges of the groups, yielding a p-value of 0.848.

Before extracting features from the recordings, the amplitudes of all the recordings were normalized to 1. Furthermore, the signal-to-noise ratio (SNR) of the two datasets was compared using a t-test to ensure compatibility, as detailed in Table 2. Notably, all the p-values from the t-test were above 0.05, indicating the similarity in SNR between the two datasets.

C. FEATURES EXTRACTION

Before the feature extraction process, all the recordings from the SVD dataset were down-sampled from 50 kHz to 44.1 kHz with a 16-bit resolution to match the sampling rate of the recordings from the PC-GITA dataset. The down-sampling process was performed using MATLAB. A voice segment of 1.5 seconds was extracted from each recording to be used as the input to the feature extraction process. A total of 18 voice features were extracted from the recordings as shown in Table 3. The features were related to three aspects of speech production controls in PD: (i) the glottal vibration control (jitter [32], shimmer [32], standard deviation (SD) of pitch frequency, and the harmonics features [33]), (ii) the lung control (voice intensity), and (iii) the vocal tract control (the mean and standard deviation of formants F1 to F4 [34] and the first five coefficients of MFCCs [35]). The first five MFCC coefficients represent vocal tract impulse responses in the range of low to medium frequency [36], [37]. The features were extracted using a code developed on Praat [38], a publicly available speech analysis software.

D. STATISTICAL ANALYSIS

Before developing the SVM classifiers, the statistical distribution of the features was calculated. The normality of the extracted features was examined using the Anderson-Darling test [39]. As the majority of the features were not normally distributed, the Wilcoxon Rank-Sum test [40] was used to compare the differences between the features of PD and the other three groups (HC, dysphonia, and laryngitis). The 95% confidence level was considered for the analysis and a p-value of less than 0.05 indicated a significant difference between the compared groups.

The probability density distributions of the features [41] were calculated and analyzed to examine the degree of overlap between the features. The two-dimensional ($d = 2$) probability density distribution function, $\hat{f}(x)$, was calculated using the 'kdensity' function in MATLAB 2022 that performs the scaled Gaussian kernel, K, density estimation as expressed in Equation (1).

$$\hat{f}_h(x) = \frac{1}{nh^d(n)} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h(n)}\right), x \in \mathbb{R}^d \tag{1}$$

E. EXPERIMENTAL SETTING

Three experimental settings were conducted to investigate the effectiveness of various machine-learning development

TABLE 3. List of extracted features.

No	Feature (unit)	Description
1	Jitter-abs (s)	Time and amplitude perturbation of glottal pulses
2	Shimmer-abs (dB)	
3	SD(Pitch) (Hz)	Standard deviation of pitch frequency
4	HNR (dB)	Harmonics-to-noise ratio
5	NHR	Noise-to-harmonics ratio
6	Mean(F1) (Hz)	Mean of formant frequency
7	Mean(F2) (Hz)	
8	Mean(F3) (Hz)	
9	SD(F1) (Hz)	Standard deviation of formant frequency
10	SD(F2) (Hz)	
11	SD(F3) (Hz)	
12	MFCC0	Mell Frequency Cepstral Coefficients
13	MFCC1	
14	MFCC2	
15	MFCC3	
16	MFCC4	
17	Mean(Int) (dB)	Mean of voice intensity
18	SD(Int) (dB)	Standard deviation of voice intensity

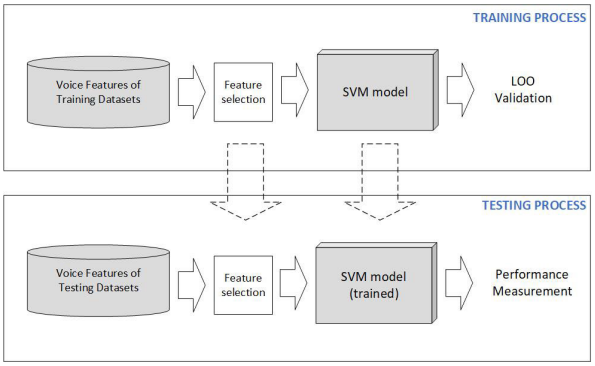

FIGURE 1. The experimental setting.

TABLE 4. Experimental settings.

	Training dataset	Testing dataset
Experiment 1 (binary SVM)	Class 0 = Non-PD (HC) Class 1 = PD	Class 0 = Non-PD (HC, Dysphonia, Laryngitis) Class 1 = PD
Experiment 2 (binary SVM)	Class 0 = Non-PD (HC + Dysphonia + Laryngitis) Class 1 = PD	Class 0 = Non-PD (HC, Dysphonia, Laryngitis) Class 1 = PD
Experiment 3 (multiclass SVM)	Class 0 = Non-PD (HC) Class 1 = Non-PD (Dysphonia) Class 2 = Non-PD (Laryngitis) Class 3 = PD	Class 0 = Non-PD (HC) Class 1 = Non-PD (Dysphonia) Class 2 = Non-PD (Laryngitis) Class 3 = PD

strategies in differentiating people with PD from people without PD (healthy people or people with other pathological conditions) based on voice features (Fig. 1).

In each experiment, an SVM [42] with a Gaussian kernel was developed and optimized based on the given training dataset. A total of 54 features (18 features from each of the three phonemes /a/, /i/, and /u/) were provided as inputs to the SVM. All input features were normalized to their mean and SD. The Relief-F algorithm [43] was applied as the feature selection process. The Relief-F algorithm ranks the features based on $k = 10$ nearest hits and misses and averages their contribution to the weights of each feature. The optimal number of ranked features was selected to achieve the highest classification (i.e., F1-score) using the leave-one-subject-out (LOSO) cross-validation method [44]. The trained model was then tested using the selected features extracted from the vowels of the testing dataset. The performance of the SVM in each experimental setting was evaluated based on its accuracy, recall, specificity, and F1-score. To avoid training bias toward the classes with a large number of recordings, each class was only represented by 44-50 randomly selected instances. A summary of the three experimental settings is presented in Table 4.

1) EXPERIMENT 1

The binary SVM was trained to differentiate two classes (PD and non-PD). During the training phase, the non-PD class was only represented by features from healthy people. The trained SVM was then tested with a non-PD class that contained healthy subjects and people with dysphonia or laryngitis. This setting was designed to investigate the ability of the SVM to recognize people with dysphonia and laryngitis as the non-PD class if the classifier was only trained with healthy people.

2) EXPERIMENT 2

In this experiment, the non-PD class consisted of healthy people and people with the two non-PD pathological conditions in both the training and testing phases. This setting was designed to determine whether there is any improvement in the binary SVM model if all the non-PD pathological conditions were included in the training dataset.

3) EXPERIMENT 3

The training and testing datasets in this experiment were similar to experiment 2. However, a multiclass SVM classifier was developed. This setting was designed to compare the effectiveness of binary and multiclass SVM models in classifying people with PD. The multiclass classifier comprised an ensemble of six binary SVM models (with a Gaussian kernel and normalized input) implemented using error-correcting output codes (ECOC) and the one-vs-one method [45], [46].

In all three experimental settings, the features of the PD recordings were extracted from the PC-GITA dataset, and the features of dysphonia and laryngitis were extracted from the SVD dataset. The HC features were extracted from the combination of PC-GITA and SVD datasets.

III. RESULTS

A. STATISTICAL ANALYSIS

Tables 5, 6, and 7 depict the statistical distribution and outcomes of the Wilcoxon Rank-Sum test for voice features associated with glottal vibration, vocal tract control, and lung control, respectively. Notably, there is a significant difference in the statistical distribution of PD features compared with the other three groups, particularly concerning features related to glottal vibration. Table 6 shows the effectiveness of vocal tract features related to frequency modulation (such as formants and MFCC) for distinguishing PD from dysphonia or laryngitis when examining /a/. Furthermore, Table 7 demonstrates that lung control features associated with /i/ and /u/ are effective for differentiating individuals with PD, dysphonia, or laryngitis.

To visualize the distribution of the features, Fig. 2 illustrates the probability density contour line of the four groups for Jitter-abs and Shimmer-abs. The 60.65% contour line represents the SD of the respective distribution. The distributions show that the features extracted from PD were significantly distinct from those extracted from HC, dysphonia, and laryngitis. In contrast, the distributions of the non-PD features were more overlapped.

B. SVM CLASSIFICATION

In the first experiment, the SVM was optimally trained with 34 selected features (Table 8). The SVM achieved an F1-score of 74.60%, with an accuracy of 77.46%, recall of 94.00%, and selectivity of 68.48%. Fig. 3 (a) shows the confusion matrix of the SVM model trained with PD and non-PD classes that only consisted of features from healthy people. The confusion matrices in Fig. 3 (b)-(c) show the performance of the trained

TABLE 5. The statistical analysis of features related to glottal vibration control.

Features	PD	Dysphonia	Laryngitis	HC	p-value of Wilcoxon Rank-Sum test		
					PD vs Dysphonia	PD vs Laryngitis	PD vs HC
Phoneme /a/							
Jitter-abs	6.04e-05(5.50e-05)	9.29e-05(1.07e-04)	8.90e-05(1.88e-04)	4.02e-05(6.91e-05)	0.780	0.550	0.000
Shim-abs	6.62e-01(3.08e-01)	5.48e-01(4.56e-01)	4.93e-01(3.91e-01)	3.96e-01(2.41e-01)	0.000	0.000	0.000
SD(Pitch)	1.12e+01(1.39e+01)	1.15e+01(2.52e+01)	6.54e+00(1.11e+01)	6.27e+00(8.79e+00)	0.000	0.000	0.000
HNR	1.55e+01(4.61e+00)	1.86e+01(6.93e+00)	1.95e+01(6.13e+00)	2.10e+01(4.65e+00)	0.000	0.000	0.000
NHR	9.43e-02(9.68e-02)	8.36e-02(1.49e-01)	5.72e-02(1.13e-01)	3.65e-02(5.95e-02)	0.000	0.000	0.000
Phoneme /i/							
Jitter-abs	4.30e-05(3.61e-05)	6.23e-05(7.71e-05)	6.40e-05(1.11e-04)	3.21e-05(4.19e-05)	0.820	0.570	0.000
Shim-abs	4.99e-01(2.05e-01)	3.39e-01(3.27e-01)	3.33e-01(3.91e-01)	3.17e-01(2.12e-01)	0.000	0.000	0.000
SD(Pitch)	1.05e+01(1.29e+01)	7.96e+00(1.66e+01)	5.60e+00(1.07e+01)	6.85e+00(1.19e+01)	0.000	0.000	0.000
HNR	1.96e+01(4.88e+00)	2.36e+01(7.13e+00)	2.32e+01(6.24e+00)	2.35e+01(5.09e+00)	0.000	0.000	0.000
NHR	4.72e-02(4.08e-02)	2.44e-02(5.24e-02)	3.05e-02(6.91e-02)	1.89e-02(2.36e-02)	0.000	0.000	0.000
Phoneme /u/							
Jitter-abs	4.16e-05(3.55e-05)	6.38e-05(8.01e-05)	4.14e-05(5.77e-05)	2.54e-05(2.63e-05)	0.650	0.160	0.000
Shim-abs	4.83e-01(2.10e-01)	3.89e-01(3.57e-01)	2.94e-01(3.09e-01)	2.87e-01(1.76e-01)	0.000	0.000	0.000
SD(Pitch)	1.22e+01(1.46e+01)	9.94e+00(1.57e+01)	6.11e+00(1.34e+01)	6.23e+00(9.58e+00)	0.010	0.000	0.000
HNR	2.41e+01(4.78e+00)	2.49e+01(6.68e+00)	2.69e+01(5.22e+00)	2.78e+01(4.75e+00)	0.130	0.000	0.000
NHR	2.81e-02(2.80e-02)	2.82e-02(5.65e-02)	1.45e-02(3.86e-02)	9.76e-03(1.43e-02)	0.000	0.000	0.000

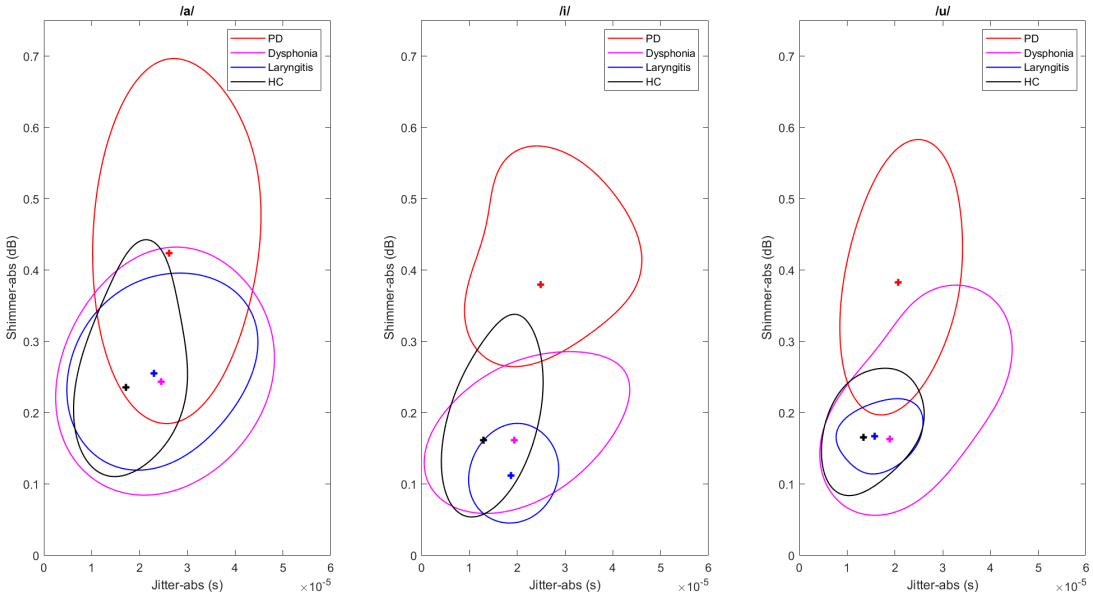

FIGURE 2. The 60.65% contour line of the probability density distribution of Jitter-abs and Shimmer-abs.

SVM when tested with non-PD classes using people who had dysphonia and laryngitis, respectively. Fig. 3 (d) shows the overall performance of the SVM by accumulating the results from Fig. 3 (a)-(c). The results show that although the non-PD class in the training set was only represented by healthy people, the SVM properly recognized people with dysphonia or laryngitis as non-PD, with a selectivity percentage of 77.27% to 79.55%.

In the second experimental setting, the optimally trained SVM achieved an F1-score of 64.34%, with an accuracy of 77.83%, recall of 92.00%, and selectivity of 73.89%. A total of 12 features were selected for the SVM model (Table 8). Fig. 4 (a) presents the confusion matrix of the SVM model when it was trained with PD and non-PD classes that contained features from individuals with HC, dysphonia, and laryngitis. Fig. 4 (b) to (d) display the testing confusion

TABLE 6. The statistical analysis of features related to vocal tract control.

Features	PD	Dysphonia	Laryngitis	HC	p-value of Wilcoxon Rank-Sum test		
					PD vs Dysphonia	PD vs Laryngitis	PD vs HC
Phoneme /a/							
Mean(F1)	7.77e+02(1.21e+02)	6.68e+02(1.27e+02)	6.41e+02(1.13e+02)	7.36e+02(1.40e+02)	0.000	0.000	0.070
Mean(F2)	1.41e+03(2.17e+02)	1.15e+03(1.44e+02)	1.14e+03(1.37e+02)	1.28e+03(1.93e+02)	0.000	0.000	0.000
Mean(F3)	2.73e+03(2.75e+02)	2.66e+03(3.44e+02)	2.68e+03(2.96e+02)	2.66e+03(2.60e+02)	0.290	0.440	0.150
SD(F1)	8.21e+01(9.65e+01)	6.30e+01(4.99e+01)	4.52e+01(4.01e+01)	4.26e+01(4.60e+01)	0.540	0.010	0.000
SD(F2)	1.18e+02(1.54e+02)	8.52e+01(7.14e+01)	7.50e+01(1.08e+02)	6.59e+01(7.62e+01)	0.980	0.020	0.010
SD(F3)	1.39e+02(1.31e+02)	1.59e+02(1.49e+02)	1.18e+02(9.88e+01)	1.07e+02(1.09e+02)	0.250	1.000	0.180
MFCC0	1.65e+03(2.39e+02)	1.80e+03(1.33e+02)	1.83e+03(1.38e+02)	1.71e+03(1.70e+02)	0.000	0.000	0.130
MFCC1	3.50e+02(8.38e+01)	4.24e+02(8.05e+01)	4.24e+02(7.22e+01)	4.10e+02(7.16e+01)	0.000	0.000	0.000
MFCC2	4.96e+00(6.97e+01)	6.55e+01(5.94e+01)	4.90e+01(5.80e+01)	1.03e+01(7.24e+01)	0.000	0.000	0.690
MFCC3	-2.81e+01(5.21e+01)	2.07e+01(5.97e+01)	3.41e+01(4.49e+01)	-1.57e+01(6.36e+01)	0.000	0.000	0.240
MFCC4	-2.08e+01(4.79e+01)	-7.41e+01(3.86e+01)	-7.03e+01(4.27e+01)	-5.52e+01(5.01e+01)	0.000	0.000	0.000
Phoneme /i/							
Mean(F1)	3.64e+02(9.64e+01)	3.05e+02(1.48e+02)	2.82e+02(4.16e+01)	3.42e+02(1.10e+02)	0.000	0.000	0.020
Mean(F2)	2.28e+03(2.72e+02)	2.42e+03(2.65e+02)	2.33e+03(2.49e+02)	2.36e+03(2.23e+02)	0.020	0.530	0.110
Mean(F3)	2.98e+03(2.34e+02)	3.15e+03(2.63e+02)	2.96e+03(3.00e+02)	2.99e+03(2.76e+02)	0.000	0.630	0.810
SD(F1)	5.65e+01(1.45e+02)	3.13e+01(1.36e+02)	2.15e+01(5.32e+01)	3.45e+01(1.25e+02)	0.000	0.000	0.000
SD(F2)	1.06e+02(1.29e+02)	1.05e+02(1.69e+02)	9.90e+01(9.36e+01)	1.05e+02(1.45e+02)	0.320	0.640	0.040
SD(F3)	1.44e+02(1.05e+02)	1.33e+02(1.12e+02)	1.51e+02(8.24e+01)	1.34e+02(9.31e+01)	0.750	0.270	0.720
MFCC0	1.58e+03(2.36e+02)	1.79e+03(1.44e+02)	1.84e+03(1.19e+02)	1.73e+03(1.88e+02)	0.000	0.000	0.000
MFCC1	2.52e+02(7.43e+01)	2.71e+02(7.54e+01)	2.74e+02(6.42e+01)	2.67e+02(6.35e+01)	0.080	0.090	0.170
MFCC2	4.51e+01(6.97e+01)	2.69e+01(5.40e+01)	9.99e+00(5.09e+01)	9.90e+00(5.20e+01)	0.120	0.010	0.000
MFCC3	1.30e+02(5.08e+01)	2.03e+02(4.88e+01)	1.92e+02(4.47e+01)	1.58e+02(5.86e+01)	0.000	0.000	0.010
MFCC4	1.06e+02(3.61e+01)	9.37e+01(4.66e+01)	8.81e+01(4.34e+01)	1.03e+02(5.11e+01)	0.190	0.050	0.800
Phoneme /u/							
Mean(F1)	4.12e+02(8.57e+01)	3.42e+02(8.49e+01)	3.57e+02(7.73e+01)	3.94e+02(8.22e+01)	0.000	0.000	0.160
Mean(F2)	1.12e+03(5.15e+02)	1.01e+03(4.34e+02)	1.04e+03(4.48e+02)	9.64e+02(4.45e+02)	0.080	0.360	0.000
Mean(F3)	2.81e+03(2.67e+02)	2.77e+03(3.01e+02)	2.75e+03(3.31e+02)	2.79e+03(2.72e+02)	0.320	0.140	0.500
SD(F1)	4.29e+01(2.78e+01)	5.42e+01(6.33e+01)	4.72e+01(5.90e+01)	4.17e+01(4.37e+01)	0.310	0.060	0.100
SD(F2)	2.41e+02(2.82e+02)	3.04e+02(3.04e+02)	2.70e+02(2.99e+02)	2.36e+02(2.88e+02)	0.330	0.980	0.570
SD(F3)	1.85e+02(1.96e+02)	2.48e+02(1.69e+02)	2.66e+02(2.18e+02)	1.77e+02(1.63e+02)	0.010	0.020	0.990
MFCC0	1.37e+03(1.82e+02)	1.53e+03(1.25e+02)	1.58e+03(1.10e+02)	1.49e+03(1.63e+02)	0.000	0.000	0.000
MFCC1	3.28e+02(5.15e+01)	3.95e+02(5.26e+01)	4.05e+02(5.70e+01)	3.82e+02(6.62e+01)	0.000	0.000	0.000
MFCC2	1.98e+02(5.68e+01)	2.13e+02(4.80e+01)	1.92e+02(4.76e+01)	1.88e+02(5.15e+01)	0.230	0.690	0.360
MFCC3	1.17e+02(4.88e+01)	1.58e+02(4.78e+01)	1.55e+02(5.14e+01)	1.30e+02(4.35e+01)	0.000	0.000	0.100
MFCC4	8.49e+00(4.47e+01)	-3.82e-01(4.84e+01)	-1.09e+01(5.72e+01)	5.47e+00(4.91e+01)	0.350	0.100	0.980

matrices for each of the three non-PD pathological groups. The results indicate that the SVM exhibited significantly high performance when trained with all possible non-PD conditions.

The highest F1-score of the multiclass SVM trained in the third experiment was significantly low (40.46%) as shown in Fig. 5 (a). A total of 4 features were optimized for the SVM model (Table 8). The results indicate that the SVM did not properly classify the four classes, especially the non-PD classes. Fig. 5 (b) presents the same result from a different perspective. The HC, dysphonia, and laryngitis classes were accumulated into one group of non-PD. The accuracy of the SVM from this perspective (PD versus non-PD) was 80.60% (recall of 86.00%, selectivity of 79.12%, and F1-score of 65.65%). This result shows that by accumulating the prediction of the three non-PD classes,

TABLE 7. The Statistical analysis of features related to lung control.

Features	PD	Dysphonia	Laryngitis	HC	p-value of Wilcoxon Rank-Sum Test		
					PD vs Dysphonia	PD vs Laryngitis	PD vs HC
Phoneme /a/							
Mean(Int)	8.23e+01(2.35e+00)	8.25e+01(2.67e+00)	8.22e+01(2.16e+00)	8.30e+01(1.76e+00)	0.540	0.650	0.050
SD(Int)	1.78e+00(1.26e+00)	1.42e+00(1.01e+00)	1.45e+00(6.86e-01)	1.51e+00(7.63e-01)	0.090	0.460	0.570
Phoneme /i/							
Mean(Int)	8.36e+01(2.11e+00)	8.54e+01(2.25e+00)	8.54e+01(1.80e+00)	8.50e+01(1.70e+00)	0.000	0.000	0.000
SD(Int)	2.04e+00(1.08e+00)	1.19e+00(9.35e-01)	1.07e+00(6.29e-01)	1.39e+00(8.67e-01)	0.000	0.000	0.000
Phoneme /u/							
Mean(Int)	8.46e+01(2.26e+00)	8.53e+01(2.58e+00)	8.58e+01(1.63e+00)	8.57e+01(1.68e+00)	0.070	0.000	0.000
SD(Int)	2.25e+00(1.25e+00)	1.21e+00(1.57e+00)	1.33e+00(1.85e+00)	1.30e+00(8.07e-01)	0.000	0.000	0.000

TABLE 8. The optimal features selected by relief-f algorithm for each experiment (marked with 'x').

No	Features	Experiment 1			Experiment 2			Experiment 3		
		/a/	/i/	/u/	/a/	/i/	/u/	/a/	/i/	/u/
1	Jitter-abs	x	x	x						
2	Shimmer-abs	x	x	x	x		x			
3	SD(Pitch)	x		x	x					
4	HNR	x			x			x		
5	NHR	x	x	x	x	x				
6	Mean(F1)		x	x						x
7	Mean(F2)	x	x	x						
8	Mean(F3)									
9	SD(F1)	x	x							
10	SD(F2)	x		x			x			
11	SD(F3)	x								
12	MFCC0	x	x	x	x	x	x		x	
13	MFCC1	x	x	x						
14	MFCC2		x							
15	MFCC3									
16	MFCC4									
17	Mean(Int)		x	x						
18	SD(Int)	x	x	x		x	x			x

the multiclass SVM performed as well as the previous two SVMs.

IV. DISCUSSION

Previous studies have developed machine learning models to identify people with PD based on the voice features extracted from sustained phonemes [17]. The models aimed to distinguish people with PD from people without PD. However, the models were mainly developed based on sustained phonemes recorded from people with PD and HC subjects. The ability of the models to selectively identify people with PD might be affected if the people without PD have other pathological conditions, such as dysphonia and laryngitis, that affect the voice production mechanism. In this work, the performance of SVM models developed based on datasets of PD and HC was evaluated considering voice features extracted from people with dysphonia or laryngitis.

The statistical analyses show that the glottal vibration features of PD were different from the non-PD groups (HC, dysphonia, and laryngitis). People with PD have much higher amplitude perturbation (shimmer) compared with people without PD, indicating the inability to control the glottal resistance that results in voice modulations correlated with the presence of noise emission and breathiness [33]. This condition was also demonstrated by the relatively higher NHR and lower HNR for PD compared with those for the other groups.

The ability to control vocal tract frequency modulation was significantly altered by PD compared with other voice-related pathological conditions. The formant frequencies of people with PD were higher than those for people with dysphonia

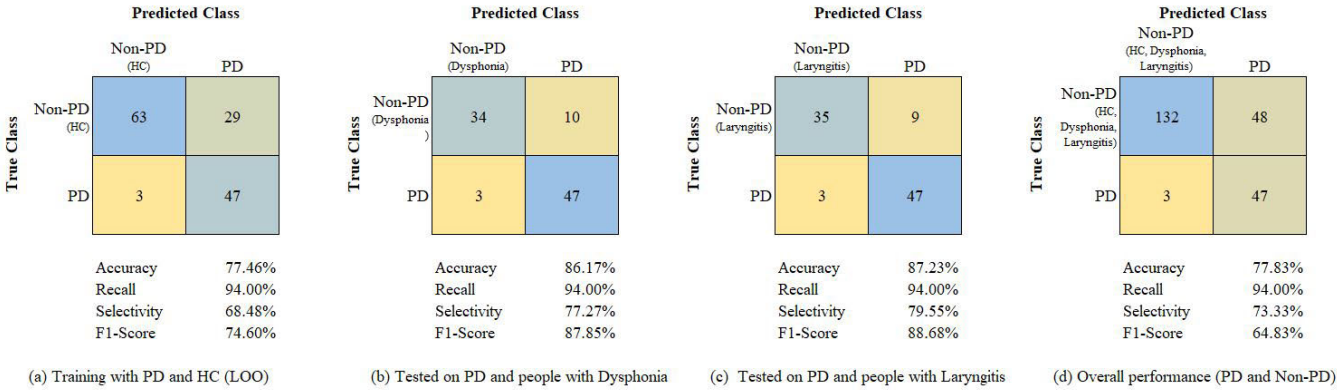

FIGURE 3. The confusion matrices obtained from the SVM created in experiment 1.

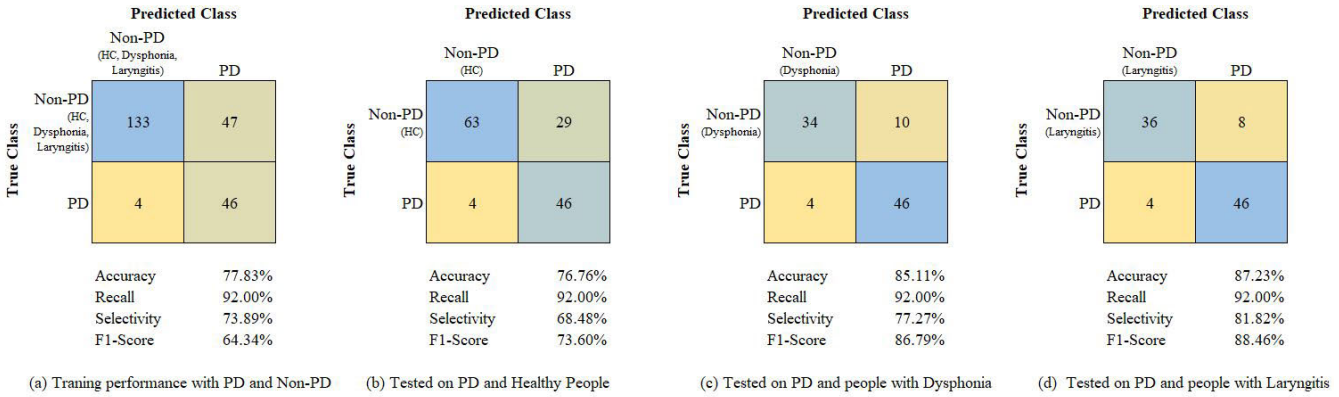

FIGURE 4. The confusion matrices obtained from the SVM created in experiment 2.

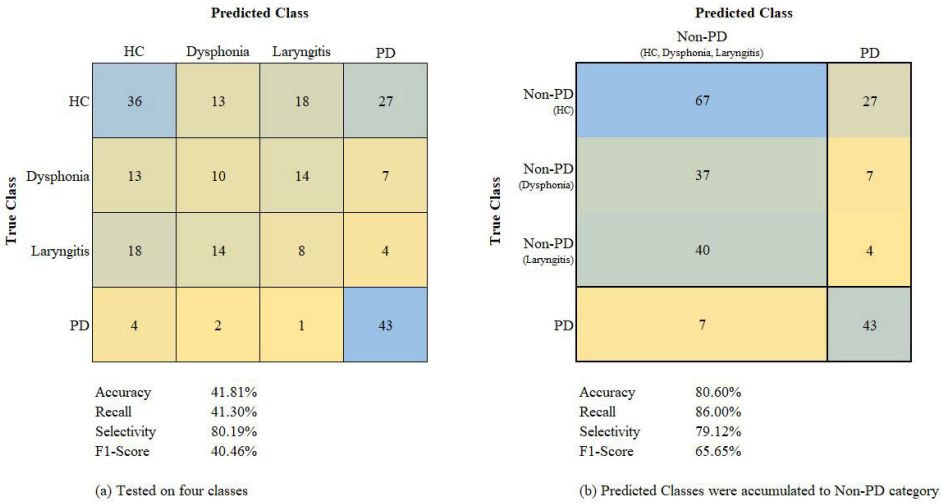

FIGURE 5. The confusion matrices obtained from the multiclass SVM created in experiment 3.

or laryngitis when pronouncing the vowel /a/. The MFCC coefficients of /a/ were lower for PD compared with those for the other groups. People with PD exhibited lower levels of voice intensity, resulting in low levels of voice loudness, compared with the other groups [7]. The results also show that people with PD had a much higher SD in voice intensity, which indicates higher instability in controlling the vocal tract and lung pressure while pronouncing the phonemes /i/ and /u/. The production of /i/ and /u/ requires precise control of the tongue, the shape/position of the lips, and the pressure of the airflow [47]. In contrast, the vowel /a/ is a back-open cardinal vowel that requires less precise control; therefore, PD affects the production of /i/ and /u/ more than the production of /a/. The effectiveness of these features was also consistent with

the optimal features selected by the Relief-F algorithm for SVM classification as shown in Table 8.

People with dysphonia showed higher jitter, shimmer, and NHRs compared with healthy people, indicating the effect of the pathological condition on the glottal vibration. These results were consistent with the previous findings reported by Teixeira and Fernandes [25], and similar results were observed in people with laryngitis [26]. The differences between dysphonia or laryngitis with HC, however, were not significant, as shown in Fig. 2. The probability density distribution of jitter and shimmer in people with dysphonia and laryngitis overlapped with the HC group.

The major finding of this study was that the SVM classifier developed based on voice features of the PD and HC groups was sufficient to recognize people with dysphonia or laryngitis as non-PD, with a significant selectivity of 77.27% to 79.55%, which was slightly higher than the selectivity of HC. This result suggests that the machine-learning models developed based on a dataset of PD and HCs were equally reliable when being implemented on people without PD who had other voice-related pathological conditions, particularly dysphonia or laryngitis. The probability density distribution of PD was significantly distinct from the other groups, whereas the distributions of dysphonia and laryngitis mostly overlapped with the HC distribution.

Additionally, this work demonstrates that including all the possible pathological conditions in the non-PD dataset might improve the performance of the SVM classifier. However, considering the limitations in collecting recordings from all the possible non-PD conditions, the SVM trained with only healthy people representing the non-PD class has a reliable performance. The multiclass classifier may also be used to distinguish PD and non-PD classes but is not effective in distinguishing the pathological conditions within the non-PD class.

V. CONCLUSION

This study confirmed that a voice-based SVM classifier trained only on the datasets of PD and HCs was equally reliable in classifying other voice-related pathological conditions, such as dysphonia and laryngitis, as non-PD, compared with the SVM classifier trained using datasets of additional pathological conditions. The statistical analyses reveal that the HC, dysphonia, and laryngitis distributions mostly overlapped and were significantly separated from the distribution of PD. The results also show that training the SVM with all possible non-PD conditions increased the overall performance of the SVM, but the multiclass classifier was not able to distinguish the non-PD classes. Overall, our findings confirm the reliability of the previously developed voice-based classifiers in cases where the non-PD group includes other voice-related pathological conditions.

REFERENCES

[1] L. M. de Lau and M. M. Breteler, "Epidemiology of Parkinson's disease," *Lancet Neurol.*, vol. 5, no. 6, pp. 525–535, 2006. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/16713924

[2] A.-M. Tăuţan, B. Ionescu, and E. Santarnecchi, "Artificial intelligence in neurodegenerative diseases: A review of available tools with a focus on machine learning techniques," *Artif. Intell. Med.*, vol. 117, Jul. 2021, Art. no. 102081.

[3] 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*, vol. 23, no. 15, pp. 2129–2170, 2008.

[4] J. M. Rabey and A. D. Korczyn, "The Hoehn and Yahr rating scale for Parkinson's disease," in *Instrumental Methods and Scoring in Extrapyramidal Disorders*. Berlin, Germany: Springer-Verlag, 1995, pp. 7–17.

[5] C. Simonet, A. Schrag, A. J. Lees, and A. J. Noyce, "The motor prodromes of Parkinson's disease: From bedside observation to large-scale application," *J. Neurol.*, vol. 268, pp. 1–10, Dec. 2019. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/31802219

[6] J. Rusz, R. Cmejla, H. Ruzickova, and E. Ruzicka, "Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease," *J. Acoust. Soc. Amer.*, vol. 129, no. 1, pp. 350–367, 2011. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/21303016

[7] 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, pp. 1–9, Jul. 2020, doi: 10.1038/s41598-020-68754-0.

[8] L. Parisi, N. RaviChandran, and M. L. Manaog, "Feature-driven machine learning to improve early diagnosis of Parkinson's disease," *Expert Syst. Appl.*, vol. 110, pp. 182–190, Nov. 2018.

[9] L. Moro-Velázquez, J. A. Gómez-García, J. I. Godino-Llorente, J. Villalba, J. R. Orozco-Arroyave, and N. Dehak, "Analysis of speaker recognition methodologies and the influence of kinetic changes to automatically detect Parkinson's disease," *Appl. Soft Comput.*, vol. 62, pp. 649–666, Jan. 2018, doi: 10.1016/j.asoc.2017.11.001.

[10] J. Goyal, P. Khandnor, and T. C. Aseri, "Classification, prediction, and monitoring of Parkinson's disease using computer assisted technologies: A comparative analysis," *Eng. Appl. Artif. Intell.*, vol. 96, Nov. 2020, Art. no. 103955, doi: 10.1016/j.engappai.2020.103955.

[11] C. O. Sakar, G. Serbes, A. Gunduz, H. C. Tunc, H. Nizam, B. E. Sakar, M. Tutuncu, T. Aydin, M. E. Isenkul, and H. Apaydin, "A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable Q-factor wavelet transform," *Appl. Soft Comput.*, vol. 74, pp. 255–263, Jan. 2019, doi: 10.1016/j.asoc.2018.10.022.

[12] D. Braga, A. M. Madureira, L. Coelho, and R. Ajith, "Automatic detection of Parkinson's disease based on acoustic analysis of speech," *Eng. Appl. Artif. Intell.*, vol. 77, pp. 148–158, Jan. 2019, doi: 10.1016/j.engappai.2018.09.018.

[13] B. Karan, S. S. Sahu, J. R. Orozco-Arroyave, and K. Mahto, "Hilbert spectrum analysis for automatic detection and evaluation of Parkinson's speech," *Biomed. Signal Process. Control*, vol. 61, Aug. 2020, Art. no. 102050, doi: 10.1016/j.bspc.2020.102050.

[14] B. Karan, S. S. Sahu, and J. R. 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, doi: 10.1016/j.bbe.2022.04.003.

[15] P. Warule, S. P. Mishra, and S. Deb, "Time-frequency analysis of speech signal using chirplet transform for automatic diagnosis of Parkinson's disease," *Biomed. Eng. Lett.*, vol. 13, no. 4, pp. 613–623, Nov. 2023, doi: 10.1007/s13534-023-00283-x.

[16] N. D. Pah, M. A. Motin, and D. K. Kumar, "Phonemes based detection of Parkinson's disease for telehealth applications," *Sci. Rep.*, vol. 12, no. 1, pp. 1–9, Jun. 2022, doi: 10.1038/s41598-022-13865-z.

[17] Q. C. Ngo, M. A. Motin, N. D. Pah, P. Drotár, P. Kempster, and D. Kumar, "Computerized analysis of speech and voice for Parkinson's disease: A systematic review," *Comput. Methods Programs Biomed.*, vol. 226, Nov. 2022, Art. no. 107133, doi: 10.1016/j.cmpb.2022.107133.

[18] N. D. Pah, M. A. Motin, P. Kempster, and D. K. Kumar, "Detecting effect of levodopa in Parkinson's disease patients using sustained phonemes," *IEEE J. Transl. Eng. Health Med.*, vol. 9, pp. 1–9, 2021.

[19] M. A. Motin, N. D. Pah, S. Raghav, and D. K. Kumar, "Parkinson's disease detection using smartphone recorded phonemes in real world conditions," *IEEE Access*, vol. 10, pp. 97600–97609, 2022.

[20] L. Ali, C. Zhu, Z. Zhang, and Y. Liu, "Automated detection of Parkinson's disease based on multiple types of sustained phonations using linear discriminant analysis and genetically optimized neural network," *IEEE J. Transl. Eng. Health Med.*, vol. 7, pp. 1–10, 2019.

[21] B. E. Sakar, M. E. Isenkul, C. O. Sakar, A. Sertbas, F. Gurgen, S. Delil, H. Apaydin, and O. Kursun, "Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings," *IEEE J. Biomed. Health Informat.*, vol. 17, no. 4, pp. 828–834, Jul. 2013. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/25055311

[22] A. Tsanas, M. A. Little, P. E. McSharry, J. Spielman, and L. O. Ramig, "Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 5, pp. 1264–1271, May 2012.

[23] S. R. Schwartz et al., "Clinical practice guideline: Hoarseness (dysphonia)," *Otolaryngol.-Head Neck Surg.*, vol. 141, no. S1, pp. 1–31, Sep. 2009.

[24] R. J. Stachler, D. O. Francis, S. R. Schwartz, C. C. Damask, G. P. Digoy, H. J. Krouse, S. J. McCoy, D. R. Ouellette, R. R. Patel, C. C. W. Reavis, L. J. Smith, M. Smith, S. W. Strode, P. Woo, and L. C. Nnacheta, "Clinical practice guideline: Hoarseness (dysphonia) (update)," *Otolaryngol.-Head Neck Surg.*, vol. 158, no. 1, pp. S1–S42, 2018.

[25] J. P. Teixeira and P. O. Fernandes, "Acoustic analysis of vocal dysphonia," *Proc. Comput. Sci.*, vol. 64, pp. 466–473, Jan. 2015, doi: 10.1016/j.procs.2015.08.544.

[26] J. P. Teixeira, J. Fernandes, F. Teixeira, and P. Fernandes, "Acoustic analysis of chronic laryngitis statistical analysis of sustained speech parameters," in *Proc. 11th Int. Conf. Bio-Inspired Syst. Signal Process.*, vol. 4, 2018, pp. 168–175.

[27] J. M. Wood, T. Athanasiadis, and J. Allen, "Laryngitis," *BMJ*, vol. 349, p. g5827, Oct. 2014.

[28] E. Van Houtte, K. Van Lierde, and S. Claeys, "Pathophysiology and treatment of muscle tension dysphonia: A review of the current knowledge," *J. Voice*, vol. 25, no. 2, pp. 202–207, Mar. 2011, doi: 10.1016/j.jvoice.2009.10.009.

[29] J. R. Orozco-Arroyave, J. D. Arias-Ledono, J. F. Vargas-Bonilla, and M. C. Gonzalez-Rativa, "New Spanish speech corpus database for the analysis of people suffering from Parkinson's disease," in *Proc. Int. Conf. Lang. Resour. Eval.*, Reykjavík, Iceland, May 2014, pp. 342–347.

[30] D. Martínez, E. Lleida, A. Ortega, A. Miguel, and J. Villalba, "Voice pathology detection on the Saarbrücken voice database with calibration and fusion of scores using multifocal toolkit," in *Advances in Speech and Language Technologies for Iberian Languages* (Communications in Computer and Information Science), vol. 328, Nov. 2014, pp. 99–109. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-84871500852&partnerID=tZOtx3y1

[31] M. Pützer and J. Koreman, "A German database of patterns of pathological vocal fold vibration," *Phonus*, vol. 3, pp. 143–153, Mar. 1997. [Online]. Available: http://www.coli.uni-saarland.de/publikationen/softcopies/Putzer:1997:GDP.pdf

[32] J. P. Teixeira and A. Gonçalves, "Accuracy of jitter and shimmer measurements," *Proc. Technol.*, vol. 16, pp. 1190–1199, Jan. 2014, doi: 10.1016/j.protcy.2014.10.134.

[33] J. P. Teixeira, C. Oliveira, and C. Lopes, "Vocal acoustic analysis—Jitter, shimmer and HNR parameters," *Proc. Technol.*, vol. 9, pp. 1112–1122, Jan. 2013, doi: 10.1016/j.protcy.2013.12.124.

[34] D. G. Childers, *Modern Spectrum Analysis*. Piscataway, NJ, USA: IEEE Press, 1978.

[35] A. Antony and R. Gopikakumari, "Speaker identification based on combination of MFCC and UMRT based features," *Proc. Comput. Sci.*, vol. 143, pp. 250–257, Jan. 2018, doi: 10.1016/j.procs.2018.10.393.

[36] N. D. Pah, V. Indrawati, and D. K. Kumar, "Voice features of sustained phoneme as COVID-19 biomarker," *IEEE J. Transl. Eng. Health Med.*, vol. 10, pp. 1–9, 2022.

[37] J. Lee, S. Shaiman, and G. Weismer, "Relationship between tongue positions and formant frequencies in female speakers," *J. Acoust. Soc. Amer.*, vol. 139, no. 1, pp. 426–440, Jan. 2016.

[38] B. P. Boersma and V. Van Heuven, "Speak and unSpeak with P RAAT," *Glot Int.*, vol. 5, nos. 9–10, pp. 341–347, 2001.

[39] L. Jäntschi and S. D. Bolboacă, "Computation of probability associated with Anderson–Darling statistic," *Mathematics*, vol. 6, no. 88, pp. 1–16, 2018.

[40] S. Datta and G. A. Satten, "Rank-sum tests for clustered data," *J. Amer. Stat. Assoc.*, vol. 100, no. 471, pp. 908–915, Sep. 2005.

[41] J. Jarnicka, "Multivariate kernel density estimation with a parametric support," *Opuscula Mathematica*, vol. 29, no. 1, p. 41, 2009.

[42] L. Hamel, *Knowledge Discovery With Support Vector Machines*. Hoboken, NJ, USA: Wiley, 2009.

[43] M. Robnik-Šikonja and I. Kononenko, "Theoretical and empirical analysis of ReliefF and RReliefF," *Mach. Learn.*, vol. 53, nos. 1–2, pp. 23–69, Oct. 2003. [Online]. Available: http://lkm.fri.uni-lj.si/xaigor/slo/clanki/MLJ2003-FinalPaper.pdf

[44] A. Geroldinger, L. Lusa, M. Nold, and G. Heinze, "Leave-one-out cross-validation, penalization, and differential bias of some prediction model performance measures—A simulation study," *Diagnostic Prognostic Res.*, vol. 7, no. 1, May 2023, doi: 10.1186/s41512-023-00146-0.

[45] Z. Yan and Y. Yang, "Performance analysis and coding strategy of ECOC SVMs," *Int. J. Grid Distrib. Comput.*, vol. 7, no. 1, pp. 67–76, Feb. 2014.

[46] S. Escalera, O. Pujol, and P. Radeva, "Separability of ternary codes for sparse designs of error-correcting output codes," *Pattern Recognit. Lett.*, vol. 30, no. 3, pp. 285–297, Feb. 2009, doi: 10.1016/j.patrec.2008.10.002.

[47] R. Ogden, *An Introduction to English Phonetics*. Edinburgh, U.K.: Edinburgh Univ. Press, 2009.

NEMUEL D. PAH (Senior Member, IEEE) received the B.E. degree in electronics from Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia, in 1992, and the M.Eng. and Ph.D. degrees in electronics engineering from RMIT University, Melbourne, VIC, Australia, in 1999 and 2003, respectively. He is currently an Associate Professor with the University of Surabaya, Indonesia, and a Visiting Associate Professor with RMIT University. His research interests include signal processing of bioelectrical signals, image processing, and the applications of AI in biomedical engineering.

VERONICA INDRAWATI (Member, IEEE) received the B.E. degree in electronics from STTS, Surabaya, Indonesia, in 1990, and the M.T. degree in electronics engineering from Institut Teknologi Bandung, Indonesia, in 1993. She is currently an Associate Professor with the University of Surabaya, Indonesia. Her research interests include control systems and signal processing.

DINESH K. KUMAR (Senior Member, IEEE) received the B.E. degree from the Indian Institute of Technology (IIT), Madras, India, and the Ph.D. degree from IIT Delhi, India, in 1990. He has been a Professor with RMIT University, Melbourne, since 1997, where he leads research in biosignals and his passion is to develop devices for affordable diagnostics. He has experience in two start-ups. He has authored more than 420 articles and seven books. He has been a Senior Professorial Fellow in Brazil and a Vajra Fellow in India. He is a fellow of the Australasian Institute of Digital Health (AIDH). He is on the TGA (Australia) Advisory Panel for Medical Devices.