INTERSPEECH 2023
20-24 August 2023, Dublin, Ireland

An Analysis of Glottal Features of Chronic Kidney Disease Speech and Its Application to CKD Detection

Jihyun Mun[1], *Sunhee Kim*[1], *Myeong Ju Kim*[1], *Jiwon Ryu*[2], *Sejoong Kim*[1,2,3], *Minhwa Chung*[1]

[1]Seoul National University, Republic of Korea
[2]Seoul National University Bundang Hospital, Republic of Korea
[3]Seoul National University College of Medicine, Republic of Korea

jhhh_1202@snu.ac.kr, sunhkim@snu.ac.kr, 99528@snubh.org, bboddo5@gmail.com, sejoong@snubh.org, mchung@snu.ac.kr

Abstract

Chronic kidney disease (CKD) causes a continuous decline in kidney function and structural damage to the kidneys. The speech characteristics of CKD speakers will be different from those of non-CKD speakers because the typical characteristics of CKD, which are impairment of respiratory and laryngeal muscles, can affect respiration, the primary source of speech. In this paper, we identify the glottal characteristics of CKD speech and then investigate whether CKD can be automatically detected using the glottal features. Statistical analysis shows significant differences between groups in glottal source features, representing the breathy characteristic of CKD speech. Through the classification experiment, we compare the performance of solely using voice quality features (baseline) against additional glottal and spectral features. When glottal source features and voice quality features are used together, an F1-score of 0.88 with a 76% relative increase compared to the baseline is obtained.

Index Terms: chronic kidney disease, glottal source parameters, automatic detection, voice pathology

1. Introduction

Chronic kidney disease (CKD) is characterized by a continuous decline in kidney function and structural damage to the kidneys. CKD alters the function and structure of the kidney irreversibly over months and years, but there may not be any symptoms in the early stages of the disease [1]. To identify CKD, a test to measure the presence of albuminuria and estimated Glomerular filtration rate (eGFR) from the serum creatinine is essential, which requires blood and urine tests [2].

CKD affects various bodily systems, particularly the respiratory system, as well as the cardiovascular, neurological, musculoskeletal, immunological, endocrine, and metabolic systems [3]. In both healthy and ill conditions, lung and kidney functions are related to maintaining the body's acid-base balance. Any changes to the renal system will affect the respiratory system, and vice versa [3, 4]. The strength and endurance of the respiratory muscles are significantly reduced in CKD patients compared to non-CKD persons, and the potency of the laryngeal and respiratory muscles is also severely compromised [3, 5]. End-stage renal disease (ESRD) characteristics, including accumulation of uremic toxins, acid-base imbalance, and volume overloads, can lead to voice change owing to decreased lung function and edema of the vocal fold [6]. As respiration is the primary source of speech [4], CKD patients' voices might represent the presence and progression of the disease.

Previous studies investigated CKD patients' speech using voice quality, prosodic, pronunciation, aerodynamic, and glottal parameters [3, 4, 5, 7, 8, 9]. In common, voice quality (jitter, shimmer, harmonics-to-noise ratio (HNR)), pitch (F0), and aerodynamic (maximum phonation time, MPT) parameters were analyzed. All studies reported lower MPT values on CKD speakers, but those studies showed conflicting results on voice quality and pitch features. [3], [4], and [7] reported higher jitter values on CKD speakers, while [5] and [9] reported lower jitter values on CKD speakers. In terms of shimmer, [3], [4], [5], [7], and [8] reported higher shimmer values on CKD speakers, but [9] reported lower shimmer values on CKD speakers. Also, in HNR, [3] and [5] reported lower HNR values on CKD speakers, but [7], [8], and [9] reported higher HNR values on CKD speakers. In terms of pitch, [3], [4], [5], and [7] reported higher F0 values on CKD speakers, and [8] reported lower F0 values regardless of gender, while [9] reported higher F0 in males, and lower F0 in females. Additionally, [9] reported lower pronunciation accuracy, longer speech duration, lower articulation rate, and low H1-A3 value in CKD speakers' speech. As CKD affects the respiratory and laryngeal muscles, glottal parameters, which are highly related to respiration, should be investigated, but no studies investigated glottal parameters in depth.

[10] proposed an optimal methodology for automatically diagnosing and predicting the severity of CKD using CKD patients' utterances. They utilized a handcrafted feature set that contains spectral, voice quality, aerodynamic, glottal, and prosodic features, eGeMAPS, and CNN features extracted from sustained vowel, voiced sentence, and general sentence utterances, and SVM and XGBoost classifiers. Handcrafted features extracted from the general sentence and using XGBoost as a classifier obtained the best results on both the detection and severity prediction task. However, as mentioned in the paper, a significant class imbalance existed between non-CKD and CKD groups, making the detection result unreliable.

This paper analyzes CKD speakers' glottal characteristics and compares them with those of non-CKD speakers'. Then we compare an automatic CKD detection performance between solely using voice quality features analyzed in previous studies against using additional glottal and spectral features. The remainder of this paper is organized as follows: Section 2 describes the dataset used in this study. Section 3 outlines our comparative analysis and automatic detection tasks. Section 4 presents statistical and experimental results; the discussion and conclusions are outlined in the last section.

2. Dataset

We use Mun et al.[9]'s CKD dataset, which is built for research on pathological voice analysis, automatic illness identification, and severity prediction. This dataset includes utterances of sustained vowel, a sentence made up entirely of voiced sounds, and a paragraph of six phonetically balanced sentences that varied

10.21437/Interspeech.2023-2478

Table 1: *Sustained vowel data from [9]*

Stage	# of utterances (Total duration)	Speakers	
		# of male speakers (average age, std)	# of female speakers (average age, std)
Non-CKD	39 (8m 55s)	14 (55.7, 12.9)	20 (57.4, 15.7)
CKD stage 1	70 (13m 11s)	23 (52.9, 17.1)	32 (49.4, 14.6)
CKD stage 2	153 (29m 7s)	68 (58.9, 15.6)	49 (58.3, 12.1)
CKD stage 3	249 (47m 25s)	97 (64.7, 13.5)	65 (68.5, 12.5)
CKD stage 4	105 (17m 2s)	43 (66.2, 12.6)	19 (70.6, 9.8)

in length. We use only sustained vowel /a/ utterances to analyze glottal characteristics. Table 1 shows the sustained vowel utterances in Mun et al.'s dataset. Specific explanations about the stage of CKD are described in [9].

As can be seen in Table 1, non-CKD (control) group's data size is much smaller than the CKD group's data size. When there exists a significant difference in the number of data per group, samples belonging to small groups are more likely to be misclassified than samples belonging to large groups when machine learning or deep learning models are applied [11]. Thus to resolve the imbalance, we import control data from Saarbruecken Voice Database [12]. As the average age of each CKD group was between 52.9 and 70.6, we import neutral /a/ vowel utterances of the speakers over 50 at the time of the recording due to the aging effect on the voice [13]. We include utterances of 20 male speakers (59.7 ± 6.7 old) and 27 female speakers (60.3 ± 8.0 old). However, class imbalance still exists despite adding the control data because of the small number of speakers over 50. Thus, we randomly extract 20 utterances (ten from female, ten from male) from each CKD stage to balance the data amount of non-CKD and CKD classes.

3. Methods

3.1. Features

3.1.1. Glottal source features

We extract glottal source features using the Disvoice toolkit [14]. The Disvoice toolkit computes phonation features derived from the glottal source reconstruction from sustained vowels. These features are computed for every glottal cycle in segments with a 200ms length frame to measure the glottal flow's short-term perturbations [15]. Then, the nine descriptors' four statistical functionals (mean, standard deviation, skewness, kurtosis) are calculated, and 36 values are obtained per utterance. We use mean values among four functionals for the statistical analysis, identifying the differences between CKD and non-CKD speakers. For the automatic detection task, we use all four functionals. Nine descriptors are as follows.

- Variability of time between consecutive glottal closure instants (GCI) (var GCI)
- Average opening quotient (OQ) for consecutive glottal cycles: rate of opening phase duration / duration of glottal cycle (avg OQ)
- Variability of opening quotient (OQ) for consecutive glottal cycles (std OQ)
- Average normalized amplitude quotient (NAQ) for consecutive glottal cycle: ratio of the amplitude quotient and the duration of the glottal cycle (avg NAQ)
- Variability of normalized amplitude quotient (NAQ) for consecutive glottal cycles (std NAQ)
- Average H1H2: Difference between the first two harmonics of the glottal flow signal (avg H1H2)
- Variability H1H2 (std H1H2)
- Average of Harmonic richness factor (HRF): ratio of the sum of the harmonics amplitude and the amplitude of the fundamental frequency (avg HRF)
- Variability of HRF (std HRF)

GCIs are time instants that mark the completion of a glottal closure event regularly occurring across pitch cycles, once per cycle. Time variability between GCIs is related to the F0 value and variability of F0 [16]. OQ is known as inversely proportional to the intensity of the voice. Smaller OQ value refers to the higher intensity [17]. NAQ is also a time-domain parameter closely related to the closing quotient (CQ), the counterpart of OQ [18]. More recently, it has been demonstrated that NAQ is strongly correlated with voice quality variations and robust to noise and estimation errors [18]. Godin & Hansen (2015) revealed that NAQ could separate the type of phonation (breathy, normal, and pressed) effectively [18], and showed that in both females and males, breathy voice showed significantly higher mean and standard deviation value of NAQ than in normal voice. The difference between the amplitude of the first harmonic and the amplitude of the second harmonic indicates the relative length of the opening phase of the glottal pulse. H1-H2 is expected to be large and positive for breathy voices and small and positive or negative for creaky voices [19]. Lastly, HRF is higher in modal voicing than that for breathy voicing [18].

3.1.2. Voice quality features

We also extract voice quality features that are commonly used in previous studies. Three voice quality features, jitter, shimmer, and HNR, are extracted. It is known that those features can describe vocal traits and provide a pathological voice diagnosis [20]. Jitter represents changes in F0 over time, while shimmer, which is very similar to jitter, represents changes in amplitude. HNR is the proportion of harmonic to noise energy. All voice quality features are extracted using Praat [21], and the minimum and maximum pitches are set to 70 Hz and 625 Hz, respectively [9]. We use them as baseline features for the classification task.

3.1.3. Spectral features

We extract Mel-frequency cepstral coefficients (MFCCs) for the automatic detection task. A representation of a sound's short-term power spectrum used in sound processing is called a Mel-frequency cepstrum (MFC), which is based on a linear cosine transform of a log power spectrum on a nonlinear Mel scale of frequency. An MFC is made up of coefficients known as MFCCs. Not only applications for speaker identification and recognition, but medical areas also utilize MFCCs for speech quality evaluation [22]. Using the librosa [23] toolkit, we extract 12-dim MFCCs and log energy from each utterance.

3.2. Statistical analysis

We conduct statistical analysis to compare CKD and non-CKD speakers' glottal characteristics. First, we conduct the Kolmogorov-Smirnov normality test to test the data normality. Then the independent sample t-test is conducted on parameters that satisfy the data normality requirements. The Mann-

Table 2: *Kolmogorov-Smirnov normality test result*

	Non-CKD			CKD		
	Statistics	Degree of freedom	p-value	Statistics	Degree of freedom	p-value
var GCI	0.203	86	0.000	0.216	80	0.000
avg NAQ	0.243	86	0.000	0.100	80	0.047
std NAQ	0.251	86	0.000	0.103	80	0.034
avg OQ	0.076	86	0.200	0.073	80	0.200
std OQ	0.132	86	0.001	0.131	80	0.002
avg H1H2	0.112	86	0.009	0.084	80	0.200
std H1H2	0.068	86	0.200	0.103	80	0.036
avg HRF	0.448	86	0.000	0.275	80	0.000
std HRF	0.418	86	0.000	0.303	80	0.000

Whitney U test, the nonparametric equivalent of an independent sample t-test, is conducted on parameters that did not satisfy the requirements for data normality. All statistical analysis is performed using IBM SPSS Statistics 26 [24].

3.3. Automatic detection of CKD

3.3.1. Feature selection

We use feature selection methods to choose the optimal feature set for each task. Three primary feature selection methods for machine learning are filter, wrapper, and embedded methods [25]. The filter method is the most computationally efficient but cannot handle redundant features. The wrapper method tends to produce better classification accuracy, but it is computationally complex. The embedded method is computationally less complicated than the wrapper method, but it has the issue of generalizability. In this paper, we use two embedded feature selection methods, which are an L1-based (lasso) feature selection [26] and a tree-based ExtraTreesClassifier (ETC), and one wrapper method, which is Recursive Feature Elimination (RFE) [27] provided by the scikit-learn library [28]. Those three methods are frequently used algorithms in classification tasks.

3.3.2. Classification

CKD detection is implemented by training features on support vector machine (SVM) and extreme gradient boosting (XGBoost) classifiers. SVM is the most commonly used classifier for diagnosing speech disorders, and it is known to perform well in high-dimensional, small-scale data classification tasks [29]. XGBoost is known to solve real-world problems well with a small amount of data and shows good performance, especially for categorical data or small data sets [30]. The hyperparameters of SVM, C, and gamma are optimized between 10^{-4} and 10^{4} through grid search, using stratified k-fold cross-validation with ten splits. The depth of decision trees (3 to 6), number of estimators to generate (12, 24, 32), learning rate (10^{-4} to 10^{-1}), and gamma (0.5, 1, 2), which are the parameters of XGBoost, are also optimized through grid search as the SVM.

As described in Section 2, we use 86 utterances from non-CKD speakers and 80 from CKD speakers. Since multiple utterances exist from one speaker, the experiments are conducted for a speaker-independent scenario, where speakers for training and testing are separated. The ratio of training and test utterances is set to 8:2.

4. Results

4.1. Statistical analysis

As shown in Table 2, among the nine glottal source features, only one parameter satisfied the data normality requirements: the average value of OQ. Thus the independent t-test is conducted for the avg OQ, and the Mann-Whitney U test is conducted for the rest of the parameters. Table 3 shows the independent t-test results, and Table 4 shows the Mann-Whitney U test results. CKD speakers show significantly higher average and standard deviation values of NAQ and OQ than non-CKD speakers. CKD speakers show a higher average value of var GCI than non-CKD speakers, but the difference is insignificant. For H1H2, CKD speakers show significantly lower average values than non-CKD speakers. CKD speakers show a higher standard deviation than non-CKD speakers, but the difference is insignificant. Lastly, for HRF, CKD speakers show significantly lower average and standard deviation values of HRF.

Table 3: *Independent t-test result (M: mean, SD: standard deviation)*

	Non-CKD		CKD		t	p-value
	M	SD	M	SD		
avg OQ	0.396	0.078	0.449	0.085	-4.229	**0.000**

Table 4: *Mann-Whitney U test result (M: mean, SD: standard deviation)*

	Non-CKD		CKD		Mann-Whitney U	p-value
	M	SD	M	SD		
var GCI	4.87E-04	5.47E-04	5.01E-4	5.28E-04	3588.000	0.632
avg NAQ	0.006	0.005	0.010	0.004	5326.000	**0.000**
std NAQ	0.002	0.001	0.003	0.001	5406.000	**0.000**
std OQ	0.072	0.035	0.085	0.034	4313.000	**0.005**
avg H1H2	7.290	2.696	8.818	4.397	4135.000	**0.025**
std H1H2	5.785	1.996	6.078	2.642	3595.000	0.616
avg HRF	286.976	2080.985	-27.086	568.864	2402.000	**0.001**
std HRF	2666.371	12750.742	2342.591	4038.992	4329.000	**0.004**

4.2. CKD detection

Table 5 shows the detection results using various combinations of feature sets. The boldface shows the best performance. When the classification experiment uses only voice quality features as input, baseline F1-scores of 0.53 and 0.50 are obtained in SVM and XGBoost, respectively. Applying ETC and RFE feature selection increases performance, and both selection algorithms chose HNR among the three voice quality features.

Overall, using MFCCs or glottal features solely or adding those features on voice quality features leads to performance increases on both SVM and XGBoost classifiers. Using MFCCs as input leads to a performance increase, obtaining an F1-score of 0.79 on the SVM classifier and an F1-score of 0.73 on the XGBoost classifier. When glottal features are used, an F1-score of 0.82 on the SVM was obtained. When RFE feature selection is applied, an additional performance increase is obtained, with an F1-score of 0.85 using only 16 features. Using glottal features on the XGBoost classifier obtained an F1-score of 0.79. When ETC feature selection is applied, an F1-score is increased to 0.85 using only 15 features.

When glottal features and baseline features are used together, an F1-score of 0.73 is obtained on the SVM classifier, which is lower than the F1-score when using only glottal features. When the RFE feature selection is applied, an F1-score increases to 0.82, but it is still lower than the F1-score when applying RFE feature selection to glottal features. The same results as using glottal features are obtained when those features are used on the XGBoost classifier. Also, Lasso and the ETC feature selection lead to the same results. However, when RFE feature selection is applied, the highest overall score is obtained, with an F1-score of 0.88, leading to a relative increase of 76%.

When MFCCs and baseline features are used together, an

Table 5: *Classification results ('-' means that no feature selection is performed)*

Features	Feature selection		SVM				XGBoost			
	Algorithm	# of features	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
Voice quality (VQ)	- (Baseline)	3	0.53	0.53	0.53	0.53	0.50	0.50	0.50	0.50
	Lasso	3	0.53	0.53	0.53	0.53	0.50	0.50	0.50	0.50
	ETC	1	0.56	0.57	0.56	0.56	0.56	0.56	0.56	0.56
	RFE	1	0.56	0.56	0.56	0.55	0.56	0.56	0.56	0.56
MFCCs	-	13	0.79	0.86	0.79	0.79	0.74	0.75	0.74	0.73
Glottal	-	36	0.82	0.84	0.82	0.82	0.79	0.79	0.79	0.79
	Lasso	24	0.82	0.83	0.82	0.82	0.76	0.77	0.76	0.76
	ETC	15	0.82	0.83	0.82	0.82	0.85	0.85	0.85	0.85
	RFE	16	0.85	0.85	0.85	0.85	0.76	0.76	0.76	0.76
VQ + Glottal	-	39	0.74	0.74	0.74	0.73	0.79	0.79	0.79	0.79
	Lasso	27	0.71	0.71	0.71	0.70	0.76	0.77	0.76	0.76
	ETC	16	0.74	0.74	0.74	0.74	0.85	0.85	0.85	0.85
	RFE	17	0.82	0.84	0.82	0.82	**0.88**	**0.89**	**0.88**	**0.88**
VQ + MFCCs	-	16	0.82	0.87	0.82	0.82	0.74	0.75	0.74	0.73
	Lasso	16	0.82	0.87	0.82	0.82	0.74	0.75	0.74	0.73
	ETC	14	0.85	**0.89**	0.85	0.85	0.74	0.75	0.74	0.73
	RFE	14	0.79	0.86	0.79	0.79	0.74	0.75	0.74	0.73
Glottal + MFCCs	-	49	0.85	**0.89**	0.85	0.85	0.79	0.81	0.79	0.79
	Lasso	37	0.79	0.81	0.79	0.79	0.85	0.85	0.85	0.85
	ETC	28	0.79	0.80	0.79	0.79	0.82	0.82	0.82	0.82
	RFE	29	0.82	0.87	0.82	0.82	0.82	0.83	0.82	0.82
VQ + MFCCs + Glottal	-	52	0.85	0.87	0.85	0.85	0.82	0.84	0.82	0.82
	Lasso	43	0.85	0.87	0.85	0.85	0.85	0.85	0.85	0.85
	ETC	28	0.85	0.87	0.85	0.85	0.82	0.83	0.82	0.82
	RFE	49	0.85	0.87	0.85	0.85	0.82	0.84	0.82	0.82

F1-score of 0.82 is obtained on the SVM classifier, which is higher than solely using baseline or MFCCs. When the ETC feature selection is applied, an F1-score of 0.85 is obtained. However, when those features are used on the XGBoost classifier, the performance increase is not obtained compared to using MFCCs only, and also, feature selection does not increase the performance.

When MFCCs and glottal features are used together, an F1-score of 0.85 and the highest precision of 0.89 are obtained on the SVM classifier, which is higher than solely using glottal features or MFCCs. In this case, feature selection does not work well, leading to no performance increase. When those features are used on the XGBoost classifier, an F1-score of 0.79 is obtained, and when the Lasso feature selection is applied, F1-score increases to 0.85.

When all the features are used, an F1-score of 0.85 is obtained on the SVM classifier, and feature selection does not increase performance. When all features are used on the XGBoost classifier, an F1-score of 0.82 is obtained, and the Lasso feature selection increases an F1-score to 0.82.

5. Discussion and Conclusion

This paper examines CKD speakers' glottal source features and compares them with non-CKD speakers'. Then we investigate whether CKD can be automatically detected using glottal source features. We analyze nine glottal source features through statistical analysis. CKD speakers show significantly higher average NAQ, OQ, H1H2, and lower average HRF values than non-CKD speakers. As described in section 3.1.1., those results indicate CKD speakers' breathy phonation. It might be caused by the effects of CKD on the respiratory and laryngeal muscles. Compromise to respiration-related muscles can cause CKD speakers to have difficulty controlling the muscles needed for phonation and may be unable to maintain stable phonation. Also, due to the vocal cord edema, incomplete vocal cord contact might be caused, which leads to a breathy voice. The classification experiment results suggest that glottal features can be used to detect CKD automatically. Moreover, the best performance is obtained when glottal features are used in addition to voice quality features, with a relative increase of 76% compared to the baseline.

This paper is the first to analyze the glottal characteristics of CKD speakers and conduct a classification experiment using them. Future work includes analyzing selected features from the feature selection algorithms which led to the best performance. Additionally, we will conduct a severity classification experiment with features used in this paper. Lastly, further experiments with more CKD data should be conducted.

6. Acknowledgements

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2023-2018-0-01833) supervised by the IITP(Institute for Information & Communications Technology Planning Evaluation).

7. References

[1] A. C. Webster, E. V. Nagler, R. L. Morton, and P. Masson, "Chronic kidney disease," *The lancet*, vol. 389, no. 10075, pp. 1238–1252, 2017.

[2] S. H. Kwon and D. C. Han, "Diagnosis and screening of chronic kidney disease," *The Korean Journal of Medicine*, vol. 76, no. 5, pp. 515–520, 2009.

[3] E. S. Hassan, "Effect of chronic renal failure on voice: an acoustic and aerodynamic analysis," *The Egyptian Journal of Otolaryngology*, vol. 30, pp. 53–57, 2014.

[4] R. B. Kumar and J. S. Bhat, "Voice in chronic renal failure," *Journal of Voice*, vol. 24, no. 6, pp. 690–693, 2010.

[5] F. M. Abd El-gaber, Y. Sallam, and H. Mohammed Eid El Sayed, "Acoustic characteristics of voice in patients with chronic kidney disease," *International Journal of General Medicine*, pp. 2465–2473, 2021.

[6] S. Y. Jung, J.-H. Ryu, H. S. Park, S. M. Chung, D.-R. Ryu, and H. S. Kim, "Voice change in end-stage renal disease patients after hemodialysis: correlation of subjective hoarseness and objective acoustic parameters," *Journal of Voice*, vol. 28, no. 2, pp. 226–230, 2014.

[7] E. A. Zaky, H. Mamdouh, O. Esmat, and Z. Khalaf, "Voice problem in a patient with chronic renal failure," *The Egyptian Journal of Otolaryngology*, vol. 36, pp. 1–8, 2020.

[8] W. A. Mudawwar, E. S. Alam, D. S. Sarieddine, Z. A. Turfe, and A.-L. H. Hamdan, "Effect of renal failure on voice." *ENT: Ear, Nose & Throat Journal*, vol. 96, no. 1, 2017.

[9] J. Mun, S. Kim, M. J. Kim, J. Ryu, S. Kim, and M. Chung, "A speech corpus for chronic kidney disease," in *2022 25th Conference of the Oriental COCOSDA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA)*. IEEE, 2022, pp. 1–6.

[10] J. Mun, S. Kim, M. J. Kim, J. Ryu, S. Kim, M. Chung *et al.*, "Automatic detection and severity prediction of chronic kidney disease using machine learning classifiers," *Phonetics and Speech Sciences*, vol. 14, no. 4, pp. 45–56, 2022.

[11] Y. Sun, A. K. Wong, and M. S. Kamel, "Classification of imbalanced data: A review," *International journal of pattern recognition and artificial intelligence*, vol. 23, no. 04, pp. 687–719, 2009.

[12] W. Barry and M. Pützer, "Saarbrucken voice database," *Institute of Phonetics*, 2007.

[13] H.-J. Choi and J.-Y. Lee, "Comparative study between healthy young and elderly subjects: Higher-order statistical parameters as indices of vocal aging and sex," *Applied Sciences*, vol. 11, no. 15, p. 6966, 2021.

[14] E. A. Belalcázar-Bolanos, J. R. Orozco-Arroyave, J. F. Vargas-Bonilla, T. Haderlein, and E. Nöth, "Glottal flow patterns analyses for parkinson's disease detection: acoustic and nonlinear approaches," in *Text, Speech, and Dialogue: 19th International Conference, TSD 2016, Brno, Czech Republic, September 12-16, 2016, Proceedings 19*. Springer, 2016, pp. 400–407.

[15] J. C. Vásquez-Correa, J. Fritsch, J. R. Orozco-Arroyave, E. Nöth, and M. Magimai-Doss, "On modeling glottal source information for phonation assessment in parkinson's disease." in *Interspeech*, 2021, pp. 26–30.

[16] Y.-R. Chien, M. Borskỳ, and J. Gunason, "F0 variability measures based on glottal closure instants." in *INTERSPEECH*, 2019, pp. 1986–1989.

[17] M. Kohler, M. M. Vellasco, E. Cataldo *et al.*, "Analysis and classification of voice pathologies using glottal signal parameters," *Journal of Voice*, vol. 30, no. 5, pp. 549–556, 2016.

[18] K. W. Godin and J. H. Hansen, "Physical task stress and speaker variability in voice quality," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2015, pp. 1–13, 2015.

[19] B. R. Kumar, J. S. Bhat, and P. Mukhi, "Vowel harmonic amplitude differences in persons with vocal nodules," *Journal of Voice*, vol. 25, no. 5, pp. 559–561, 2011.

[20] J. P. Teixeira, C. Oliveira, and C. Lopes, "Vocal acoustic analysis–jitter, shimmer and hnr parameters," *Procedia Technology*, vol. 9, pp. 1112–1122, 2013.

[21] P. Boersma, "Praat, a system for doing phonetics by computer," *Glot. Int.*, vol. 5, no. 9, pp. 341–345, 2001.

[22] A. Benba, A. Jilbab, A. Hammouch, and S. Sandabad, "Voiceprints analysis using mfcc and svm for detecting patients with parkinson's disease," in *2015 International conference on electrical and information technologies (ICEIT)*. IEEE, 2015, pp. 300–304.

[23] B. McFee, C. Raffel, D. Liang, D. P. Ellis, M. McVicar, E. Battenberg, and O. Nieto, "librosa: Audio and music signal analysis in python," in *Proceedings of the 14th python in science conference*, vol. 8, 2015, pp. 18–25.

[24] I. Statistics, "Ibm corp. released 2013. ibm spss statistics for windows, version 22.0. armonk, ny: Ibm corp," *Google Search*, 2013.

[25] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Computers & Electrical Engineering*, vol. 40, no. 1, pp. 16–28, 2014.

[26] R. Muthukrishnan and R. Rohini, "Lasso: A feature selection technique in predictive modeling for machine learning," in *2016 IEEE international conference on advances in computer applications (ICACA)*. IEEE, 2016, pp. 18–20.

[27] A. Hernandez, S. Kim, and M. Chung, "Prosody-based measures for automatic severity assessment of dysarthric speech," *Applied Sciences*, vol. 10, no. 19, p. 6999, 2020.

[28] O. Kramer and O. Kramer, "Scikit-learn," *Machine learning for evolution strategies*, pp. 45–53, 2016.

[29] S. Hegde, S. Shetty, S. Rai, and T. Dodderi, "A survey on machine learning approaches for automatic detection of voice disorders," *Journal of Voice*, vol. 33, no. 6, pp. 947–e11, 2019.

[30] M. S. Darouiche, H. El Moubtahij, M. B. Yakhlef, and E. B. Tazi, "An automatic voice disorder detection system based on extreme gradient boosting classifier," in *2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*. IEEE, 2022, pp. 1–5.