IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-4, April 2019 E-ISSN: 2347-2693

# Perspective Analysis of Voice Disorder Detection using various Approaches

**P. Kokila[1*], G. M. Nasira[2]**

[1]Dept. of Computer Science, Chikkanna government Arts college, Tirupur-641602, Tamil Nadu, India
[2]Dept. of Computer Applications, Chikkanna Government Arts College, Tirupur-641602, Tamil Nadu, India

[*]*Corresponding Author: kohilasairam@gmail.com*

DOI: https://doi.org/10.26438/ijcse/v7i4.12081212 | Available online at: www.ijcseonline.org

Accepted: 20/Apr/2019, Published: 30/Apr/2019

***Abstract***— Automatic voice pathology detection and classification systems effectively contribute to the assessment of voice disorders, which helps clinicians to detect the existence of any voice pathologies and the type of pathology from which patients suffer in the early stage. This paper performs detailed study on various methodologies like feature extraction techniques, pattern recognition using machine learning, artificial intelligence, data mining, etc., used by various researches to detect the voice disorder using signal processing and voice recordings. The identification of an algorithm that discriminates between pathological and healthy voices with more accuracy is necessary to realize a valid and precise health system. The key contribution of this study is to investigate the performance of several machine learning techniques useful for voice pathology detection. This work provides detailed survey and comparison of the existing works pros and cons. This study also highlights the drawbacks in the existing methods and outlines the important factors to be considered while performing

***Keywords***— voice pathology, voice disorder, signal processing, machine learning, data mining and feature extraction

## I. INTRODUCTION

Pathological voice recognition has been received a great attention from researchers in the last decade. Speech processing has proved to be an excellent tool for voice disorder detection. Among the most interesting recent works are those concerned with Parkinson's Disease (PD), Multiple Sclerosis (MS) and other diseases which belong to a class of neurodegenerative diseases that affect patient's speech, motor, and cognitive capabilities [1, 2]. The speech production is a complex motor act that implies a big number of muscles, of physiological variables and a neurological control implying different cortical and under cortical regions. They distinguished three systems contributing to the production of the speech: the respiratory system, the laryngeal system and the supra-laryngeal system (the articulators) [3, 4]. The nervous system also controls the prosody. This one schematically covers the variations of height (intonation, melody), the variations of intensity (accentuation) and the temporal progress (pauses, debit, and rhythm). Voice pathologies affect the vocal folds during the phonation process. They make vocal folds producing irregular vibrations due to the malfunctioning of different factors contributing to vocal vibrations. Vocal folds are differently affected by vocal fold pathologies resulting in variation in the vibratory cycle of vocal folds because their ability to be closed properly is decreased.

The paper is organized in such a way that the section II discusses about the related work done by researchers in the field of voice disorder detection, the table summarizes the complete details of discussed prevailing works. The section III deals with the findings of this research problem in voice disorder detection with suggested improvements. The final section concludes the entire work summary and the importance of voice disorder detection using machine learning algorithms.

## II. RELATED WORK

Among several machine learning techniques existing in literature, Support Vector Machine (SVM) has been widely used in voice signal processing. Godino et al. [5], focused on the classification of pathological and healthy voices based on MFCC to train and test an SVM classifier. These have obtained a good accuracy (92%). The SVM technique was also used in [6] to estimate the presence of dysphonia, investigating four types of pathology: chronic laryngitis, cysts, Reinke's edema and spasmodic dysphonia. The authors proposed an algorithm based on the use of MFCC and Linear Discriminant Analysis (LDA) as a dimensionality reduction method. This algorithm identifies the presence of a pathology with a discrete accuracy (86%). However, it was tested on a very limited dataset. In fact, only 70 pathological and 40 healthy voices were selected by the Saarbruechen Voice Database (SVD) [7].

El Emary et al. [8], instead, classified the speech signal by estimating not only the MFCC but also jitter and shimmer. The detection of voice suffering from neurological disorders was performed using the GMM algorithm on a very small subset of the SVD database containing only 38 pathological and 63 healthy voices.

In [9], an algorithm based on Daubechies discrete wavelet transform, linear prediction coefficients and last squares support vector machine (LS-SVM) was used to identify laryngeal pathologies. The experiments were carried out using a private database.

Another private dataset collected in the Busan National University Hospital was used in the study described in the work of Wang et al [10] they classified pathological voices using Hidden Markov Models (HMM), GMM and SVM. The voice disorders considered in their study included vocal polyps, vocal cord palsy, nodules, cysts, edema, laryngitis and glottic cancer.

Henríquez *et al.[11]* studied the usefulness of six nonlinear chaotic measures based on nonlinear dynamics theory in the discrimination between two levels of voice quality: healthy and pathological. The studied measures are first and second order Rényi entropies, the correlation entropy and the correlation dimension. The values of the first minimum of mutual information function and Shannon entropy were also studied. Two databases were used to assess the usefulness of the measures: a multi-quality and a commercial database (MEEI Voice Disorders). A classifier based on standard neural networks was implemented in order to evaluate the measures proposed. Global success rates of 82.5% (multi-quality database) and 99.7% (commercial database) were obtained.

In Forero *et al [12]*, several parameters of glottal signal were used to identify nodule, unilateral paralysis or healthy voices. The database, obtained from a speech therapist, was composed of records of voices from 12 speakers with nodule, 8 speakers with vocal fold paralysis and 11 speakers with normal voices. Eight voice records were taken of each speaker making a total of 248 records. Three different classifiers were used, an Artificial Neural Network, a Support Vector Machine (SVM) and Hidden Markov Model. The best accuracy, 97.2%, was reached using glottal signal parameters and MFCC's with a SVM as classifier.

Markaki *et al* [13] explored the information provided by a joint acoustic and modulation frequency representation, referred to as modulation spectrum, for detection and discrimination of voice disorders. The initial representation is first transformed to a lower dimensional domain using higher order singular value decomposition (HOSVD). For voice pathology detection an accuracy of 94.1% was achieved using SVM as classifier.

In Panek *et al* [14], a vector made up of 28 acoustic parameters is evaluated using Principal Component Analysis (PCA), kernel principal component analysis (kPCA) and an auto-associative neural network (NLPCA) in four kinds of pathology detection (hyperfunctional dysphonia, functional dysphonia, laryngitis, vocal cord paralysis) using the /a/, /i/ and /u/ vowels, spoken at a high, low and normal pitch. The results shows a best efficiency levels of around 100%.

Al-Nasheri *et al.*[15] investigated different frequency bands using correlation functions. The authors extracted maximum peak values and their corresponding lag values from each frame of a voiced signal by using correlation functions as features to detect and classify pathological samples. Three different databases were used, Arabic Voice Pathology Database (AVPD), Saarbruecken Voice Database (SVD) and Massachusetts Eye and Ear Infirmary (MEEI). A Support Vector Machine was used as classifier. For detection of pathology an accuracy of 99,8%, 90.9% and 91.1% was achieved for the three databases respectively. In classification of the pathology task an accuracy of 99.2%, 98,9% and 95.1%, respectively, was achieved for the three databases

In Sellam et al.[16], an attempt is made to analyze and to discriminate pathological voice from normal voice in children using different classification methods. The classification of pathological voice from normal voice is implemented using Support Vector Machine (SVM) and Radial Basis Functional Neural Network (RBFNN). Several acoustic parameters were extracted such as the signal energy, pitch, formant frequencies, mean square residual signal, reflection coefficients, Jitter and Shimmer. The best accuracy results was obtained by RBFNN with, 91%, and for the SVM 83%. The artificial neural networks are among the most used classifiers for this kind of task.

Hugo Cordeiro [17] presented a set of experiments to identify the best set of features from the vocal tract (MFCC, Line Spectral Frequencies (LSF), Mel-Line Spectral Frequencies (MLSF) and first peak of the spectral envelop) and the best classifiers amongst SVM and Gaussian Mixture Models (GMM) for the identification of pathologic voices. He achieved an accuracy of 84,4% for the identification between 3 groups (healthy subjects, subjects with physiological larynx pathologies - vocal fold nodules and edemas, and subjects with neurological larynx pathologies - unilateral vocal fold paralysis). He also used Regression Trees to the pathological voice recognition based on formant analysis and harmonic-to-noise ratio with 95% of recognition rate.

In research work [18], the speech signal is analyzed by the acoustic parameters like Signal Energy, pitch, Silence removal, Windowing, Mel frequency Cepstrum, and Jitter. At the end, the classification technique i.e Support Vector Machine is used to classify the normal and pathology voice based on the features extracted in the previous phase.

Laura Verde, Giuseppe De Pietro [19] in their work focused on dysphonia, an alteration of the voice quality that affects about one person in three at least once in his/her lifetime. Voice disorders are rapidly spreading, although they are often underestimated. Mobile health systems can be an easy and fast support to voice pathology detection. The identification of an algorithm that discriminates between pathological and healthy voices with more accuracy is necessary to realize a valid and precise mobile health system. The key contribution of this study is to investigate and compare the performance of several machine learning techniques useful for voice pathology detection Joao Paulo [20] in their work developed a novel vector model which best predictors/parameters for diagnose of dysphonia were experimented. A vector made up of 4 Jitter parameters, 4 Shimmer parameters and Harmonic to Noise Ratio (HNR), determined from 3 different vowels at 3 different tones, in a total of 81 features, was used. Variable selection and dimension reduction techniques such as hierarchical clustering, multilinear regression analysis and principal component analysis (PCA) was applied.

Zhijian Wang et al [21] in their work developed a balancing automatic voice valuation system by using multidimensional acoustical trials based on the well-known GRBAS perceptual rating scale. In this work nearly 65 features are measures using conventional acoustic methods, Nonlinear dynamical analysis, Glottal to Noise excitation approaches and dynamical analysis are used to construct a feature matrix. To decrease duplication in features four different feature extraction methods were applied. RBF kernel SVM and extreme learning machines are used for multiclass classification. The result of classification was moderately correlated with GRBAS ratings of severity with the best accuracy of around 77.55 and 80.58 % correspondingly. The table 1 shows the summarized overview of the existing methods on voice disorder detection with their methodology used and its remark.

Table 1: Summary of the Survey on Existing Methods of Voice Disorder Detection

| S.No | Author name | Methodology Used | Remarks |
|---|---|---|---|
| 1 | Godino et al. [5] (2005) | MFCC to train and test an SVM classifier | classification of pathological and healthy voices |
| 2 | Souissi & Cherif [6] | MFCC and Linear Discriminant | Estimates four different types |
| 3 | M. El Emary et al [8] (2014) | MFCC and GMM Algorithm | classification of pathological and healthy voices |
| 4 | E. S. Fonseca [9] (2007) | discrete wavelet transforms, linear prediction coefficients and last squares support vector machine | laryngeal pathologies |
| 5 | J. Wang & C. Jo [10] (2007) | Hidden Markov Models (HMM), GMM and SVM | Classification of voice disorders |
| 6 | J. Wang & C. Jo [11] (2009) | Non-linear dynamic theory with Neural Network | Voice Quality classification as pathological and healthy voices |
| 7 | Forero et al[12] (2015) | Artificial Neural Network, Support Vector Machine (SVM) and Hidden Markov Mode | nodule, unilateral paralysis or healthy voices |
| 8 | Markaki & Stylianou[13] (2011) | Higher order singular value decomposition for feature Selection and SVM Classifier | voice pathology detection |
| 9 | Panek [14] (2015) | Principal Component Analysis (PCA), kernel principal component analysis (kPCA) and an auto-associative neural network (NLPCA) | four kinds of pathology detection |

 

| | | | |
|---|---|---|---|
| 10 | Al-nasheri [15] (2016) | Support Vector Machine was used as classifier | classification of pathological and healthy voices |
| 11 | Sellam & Jagadeesan [16] (2014) | Support Vector Machine (SVM) and Radial Basis Functional Neural Network (RBFNN) | discriminate pathological voice from normal voice in children |
| 12 | Hugo Cordeiro & Carlos Meneses Ribeiro [17] (2016) | Feature Extraction MFCC, LSF, MLSF and SVM and GMM are used as classifiers | identification of pathologic voices. |
| 13 | Punitha & Sheela Selvakumari [18] (2017) | SVM Classifier | classification of pathological and healthy voices |
| 14 | Laura Verde [19] (2018) | several machine learning techniques | voice pathology detection |
| 15 | Joao Paulo Teixeiraa [20] (2017) | hierarchical clustering, multilinear regression analysis and PCA with ANN classifier | diagnose of dysphonia |
| 16 | Zhijian Wang [21] | RBF kernel SVM classifier | automatic voice valuation |

## III. FINDINGS

This review reports the detailed study about the feature extraction and different techniques for classification model. The chief aim of this research work is to realize the process of feature extraction methods and classification models to develop an efficient voice pathology system with great accuracy.

It is observed from the above study that the following research of interest in lacking more in voice disorder detection they are

- Providing optimized feature extraction method to handle the voice pathology detection in presence of high noise ratio
- While handling such voluminous and complex structure of voice dataset most of the research work fails to put forth the importance of potential feature selection
- In Real time voice dataset there will be presence of uncertainty in classifying the voice which is not properly defined either as healthy or pathological voice.

Thus, this research work will be focused to develop optimal feature extraction on voice signal by filtering the presence of noise and extracting optimal features which will result in enhancement of the voice classification. The feature selection method will be preprocessed as audio signal preprocessing using heuristic feature selection methods. The uncertainty will be greatly handled by introducing generalization of fuzzy logic in artificial neural network.

## IV. CONCLUSION

Voice disorders can have a significant negative impact on the social and professional life of those affected. Although such disorders are often underestimated, their early detection and accurate diagnosis are necessary to reduce serious consequences. Computer-based systems, such as m-health solutions, provide an opportunity to improve and support the main medical techniques necessary to diagnose the presence of these disorders. This paper investigated various existing approaches done on voice disorder detection using machine learning approaches, signal processing and mining approaches. From the result it is observed that still there is no proper proof on selecting optimal features and determining and handling uncertainties in the voice disorder prediction and classification. The potential feature subset selection plays an important role to improve the accuracy of the classification model. The issues in real time voice recognition dataset fails to handle uncertainty for discovering whether the voice belongs to healthy or pathology due to vague information. Thus, this research work aims at developing the models in near future by focusing feature selection and uncertainty handling using classification models.

## REFERENCES

[1]. Davis B., “Acoustic Characteristics of Normal and Pathological Voices,” in Proceedings of Speech and Language: Advances in Basic Research and Practice, Orland, pp. 271-335, 1979.
[2]. Parsa V. and Jamieson G., “Interactions between Speech Coders and Disordered Speech,” Computer Journal of Speech Communication, vol. 40, no. 7, pp. 365-385, 2003.
[3]. Wang J. and Cheolwoo J., “Performance of Gaussian Mixture Model as a Classifier for Pathological Voice,” in Proceedings of the ASST in Auckland, Australian, pp. 165-169, 2006.
[4]. Yu P., Ouaknine M., Revis J., and Giovanni A., “Objective Voice Analysis for Dysphonic Patients: A Multiparametric Protocol Including Acoustic and Aerodynamic Measurements,” Computer Journal of Voice, vol. 15, no. 4, pp. 529-542, 2001
[5]. J. I. Godino-Llorente, P. G´omez-Vilda, N. S´aenz-Lech´on, M. Blanco Velasco, F. Cruz-Rold´an, and M. A. Ferrer-Ballester,

"Support vector machines applied to the detection of voice disorders," Lecture notes in computer science, vol. 3817, p. 219, 2005.
[6]. N. Souissi and A. Cherif, "Dimensionality reduction for voice disorders identification system based on mel frequency cepstral coefficients and support vector machine," in Modelling, Identification and Control (ICMIC), 2015 7th International Conference on. IEEE, 2015, pp. 1–6.
[7]. W. Barry and M. P¨utzer, "Saarbrucken voice database," Institute of Phonetics, Universit¨at des Saarlandes, http://www. stimmdatenbank.coli. uni-saarland. de, 2007.
[8]. M. El Emary, M. Fezari, and F. Amara, "Towards developing a voice pathologies detection system," Journal of Communications Technology & Electronics, vol. 59, no. 11, p. 1280, 2014.
[9]. E. S. Fonseca, R. C. Guido, P. R. Scalassara, C. D. Maciel, and J. C. Pereira, "Wavelet time-frequency analysis and least squares support vector machines for the identification of voice disorders," Computers in Biology and Medicine, vol. 37, no. 4, pp. 571–578, 2007.
[10]. J. Wang and C. Jo, "Vocal folds disorder detection using pattern recognition methods," in Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE. IEEE, pp. 3253–3256, 2007.
[11]. Henríquez, P., Alonso, J. B., Ferrer, M. A., Travieso, C. M., Godino-Llorente, J. I., Díaz-di-María, F. Characterization of Healthy and Pathological Voice Through Measures Based on Nonlinear Dynamics. IEEE Transactions on Audio, Speech, and Language Processing, Vol. 17, No. 6, 1186-1195, August 2009.
[12]. Forero, L. A., Kohler, M., Vellasco, M., Cataldo, E. Analysis and Classification of Voice Pathologies Using Glottal Signal Parameters. Journal of Voice, 30(5):549-556, 2015.
[13]. Markaki, M., Stylianou, Y. Voice Pathology Detection and Discrimination Based on Modulation Spectral Features. IEEE Transactions on Audio, Speech, and Language Processing, Vol. 19, No. 7, 1938-1948, 2011.
[14]. Panek, D., Skalski, A., Gajda, J., Tadeusiewicz, R. Acoustic Analysis Assessment in Speech Pathology Detection. Int. J. Appl. Math. Comput. Sci., 2015, Vol. 25, No. 3, 631–643.
[15]. Al-nasheri, A., Muhammad, G., Alsulaiman, M., Ali, Z. Investigation of Voice Pathology Detection and Classification on Different Frequency Regions Using Correlation Functions. Journal of Voice, 31(1):3-15, 2016.
[16]. Sellam, V., Jagadeesan, J. Classification of Normal and Pathological Voice Using SVM and RBFNN. Journal of Signal and Information Processing, 2014, 5, 1-7.
[17]. Hugo Cordeiro, Carlos Meneses Ribeiro, Speaker Characterization with MLSFs, The Journal for Nurse Practitioners. Elsevier, 2016
[18]. S. C. Punitha¸ N. A. Sheela Selvakumari, Voice Pathology Identification System using SVM Classifier, International Journal of Advance Research in Computer Science and Management Studies, Volume 5, Issue 1, January 2017
[19]. Laura Verde, Giuseppe De Pietro, Voice Disorder Identification by using Machine Learning Techniques, IEEE. Translations and content mining, pp 1-11, 2018.
[20]. Joao Paulo Teixeiraa, Paula Odete Fernandesa, Nuno Alvesa, Vocal Acoustic Analysis – Classification of Dysphonic Voices with Artificial Neural Networks, ScienceDirect, Procedia Computer Science 121 (2017) 19–26
[21]. Zhijian Wang, Ping Yu, Nan Yan, Lan Wang, Manwa L. Ng, Automatic Assessment of Pathological Voice Quality Using Multidimensional Acoustic Analysis Based on the GRBAS Scale, Sign Process Syst (82:241–251), 2016