Ana I. Pereira · Andrej Košir ·
Florbela P. Fernandes · Maria F. Pacheco ·
João P. Teixeira · Rui P. Lopes (Eds.)

Communications in Computer and Information Science 1754

Optimization, Learning Algorithms and Applications

Second International Conference, OL2A 2022
Póvoa de Varzim, Portugal, October 24–25, 2022
Proceedings

Springer

OL2A INTERNATIONAL CONFERENCE

Communications in Computer and Information Science 1754

Ana I. Pereira · Andrej Košir ·
Florbela P. Fernandes · Maria F. Pacheco ·
João P. Teixeira · Rui P. Lopes (Eds.)

Optimization, Learning Algorithms and Applications

Second International Conference, OL2A 2022
Póvoa de Varzim, Portugal, October 24–25, 2022
Proceedings

Springer

Editors
Ana I. Pereira
Instituto Politécnico de Bragança
Bragança, Portugal

Andrej Košir
University of Ljubljana
Ljubljana, Slovenia

Florbela P. Fernandes
Instituto Politécnico de Bragança
Bragança, Portugal

Maria F. Pacheco
Instituto Politécnico de Bragança
Bragança, Portugal

João P. Teixeira
Instituto Politécnico de Bragança
Bragança, Portugal

Rui P. Lopes
Instituto Politécnico de Bragança
Bragança, Portugal

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-3-031-23235-0 ISBN 978-3-031-23236-7 (eBook)
https://doi.org/10.1007/978-3-031-23236-7

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This CCIS volume 1754 contains the refereed proceedings of the Second International Conference on Optimization, Learning Algorithms and Applications (OL2A 2022), a hybrid event held during October 24–25, 2022.

OL2A 2022 provided a space for the research community on optimization and learning to get together and share the latest developments, trends, and techniques, as well as to develop new paths and collaborations. The conference had more than three hundred participants in an online and face-to-face environment throughout two days, discussing topics associated with optimization and learning, such as state-of-the-art applications related to multi-objective optimization, optimization for machine learning, robotics, health informatics, data analysis, optimization and learning under uncertainty, and Industry 4.0.

Five special sessions were organized under the following topics: Trends in Engineering Education, Optimization in Control Systems Design, Measurements with the Internet of Things, Advances and Optimization in Cyber-Physical Systems, and Computer Vision Based on Learning Algorithms. The OL2A 2022 program included presentations of 56 accepted papers. All papers were carefully reviewed and selected from 145 submissions in an single-blind process. All the reviews were carefully carried out by a scientific committee of 102 qualified researchers from 21 countries, with each submission receiving at least 3 reviews.

We would like to thank everyone who helped to make OL2A 2022 a success and hope that you enjoy reading this volume.

October 2022

Ana I. Pereira
Andrej Košir
Florbela P. Fernandes
Maria F. Pacheco
João P. Teixeira
Rui P. Lopes

Organization

General Chairs

Ana I. Pereira	Polytechnic Institute of Bragança, Portugal
Andrej Košir	University of Ljubljana, Slovenia

Program Committee Chairs

Florbela P. Fernandes	Polytechnic Institute of Bragança, Portugal
Maria F. Pacheco	Polytechnic Institute of Bragança, Portugal
João P. Teixeira	Polytechnic Institute of Bragança, Portugal
Rui P. Lopes	Polytechnic Institute of Bragança, Portugal

Special Session Chairs

João P. Coelho	Polytechnic Institute of Bragança, Portugal
Luca Oneto	University of Genoa, Italy

Technology Chairs

Alexandre Douplik	Ryerson University, Canada
Paulo Alves	Polytechnic Institute of Bragança, Portugal

Local Organizing Chairs

José Lima	Polytechnic Institute of Bragança, Portugal

Program Committee

Ana I. Pereira	Polytechnic Institute of Bragança, Portugal
Abeer Alsadoon	Charles Sturt University, Australia
Ala' Khalifeh	German Jordanian University, Jordan
Alexandre Douplik	Ryerson University, Canada
Ana Maria A. C. Rocha	University of Minho, Portugal
Ana Paula Teixeira	University of Trás-os-Montes and Alto Douro, Portugal
André Pinz Borges	Federal University of Technology – Paraná, Brazil
André R. da Cruz	Federal Center for Technological Education of Minas Gerais, Brazil

Andrej Košir	University of Ljubljana, Slovenia
Arnaldo Cândido Júnior	Federal University of Technology – Paraná, Brazil
António J. Sánchez-Salmerón	Universitat Politécnica de Valencia, Spain
António Valente	Universidade de Trás-Os-Montes e Alto Douro, Portugal
Bilal Ahmad	University of Warwick, UK
Bruno Bispo	Federal University of Santa Catarina, Brazil
Carlos Henrique Alves	CEFET/RJ, Brazil
Carmen Galé	University of Zaragoza, Spain
B. Rajesh Kanna	Vellore Institute of Technology, India
Carolina Gil Marcelino	Federal University of Rio de Janeiro, Brazil
Christopher E. Izquierdo	University of Laguna, Spain
C. Sweetlin Hemalatha	Vellore Institute of Technology, India
Damir Vrančić	Jozef Stefan Institute, Slovenia
Daiva Petkeviciute	Kaunas University of Technology, Lithuania
Dhiah Abou-Tair	German Jordanian University, Jordan
Diego Brandão	CEFET/RJ, Brazil
Dimitris Glotsos	University of West Attica, Greece
Diamantino Silva Freitas	University of Porto, Portugal
Eduardo Vinicius Kuhn	Federal University of Technology – Paraná, Brazil
Elizabeth Fialho Wanner	Federal Center for Technological Education of Minas Gerais, Brazil
Elaine Mosconi	Université de Sherbrooke, Canada
Esteban Clua	Federal Fluminense University, Brazil
Eric Rogers	University of Southampton, UK
Felipe N. Martins	Hanze University of Applied Sciences, Netherlands
Florbela P. Fernandes	Polytechnic Institute of Bragança, Portugal
Florentino F. Riverola	University of Vigo, Spain
Gaukhar Muratova	Dulaty University, Kazakhstan
Gediminas Daukšys	Kauno Technikos Kolegija, Lithuania
Glaucia Maria Bressan	Federal University of Technology – Paraná, Brazil
Glotsos Dimitris	University of West Attica, Greece
Humberto Rocha	University of Coimbra, Portugal
J. Marcos Moreno Veja	University of Laguna, Spain
João Paulo Carmo	University of São Paulo, Brazil
João P. Teixeira	Polytechnic Institute of Bragança, Portugal
Jorge Igual	Universidad Politécnica de Valencia, Spain
José Boaventura-Cunha	University of Trás-os-Montes and Alto Douro, Portugal
José Lima	Polytechnic Institute of Bragança, Portugal

Joseane Pontes	Federal University of Technology – Ponta Grossa, Brazil
Juan Alberto G. Esteban	Universidad de Salamanca, Spain
Juan A. Méndez Pérez	University of Laguna, Spain
Juani Lopéz Redondo	University of Almeria, Spain
Julio Cesar Nievola	Pontifícia Universidade Católica do Paraná, Brazil
João Paulo Coelho	Polytechnic Institute of Bragança, Portugal
Jorge Ribeiro	Polytechnic Institute of Viana do Castelo, Portugal
José Ramos	NOVA University Lisbon, Portugal
Kristina Sutiene	Kaunas University of Technology, Lithuania
Lidia Sánchez	University of León, Spain
Lino Costa	University of Minho, Portugal
Luis A. De Santa-Eulalia	Université de Sherbrooke, Canada
Luís Coelho	Polytecnhic Institute of Porto, Portugal
Luca Oneto	University of Genoa, Italy
Luca Spalazzi	Marche Polytechnical University, Italy
Maria F. Pacheco	Polytechnic Institute of Bragança, Portugal
Manuel Castejón Limas	University of León, Spain
Marc Jungers	Université de Lorraine, France
Marco A. S. Teixeira	Universidade Tecnológica Federal do Paraná, Brazil
Maria do R. de Pinho	University of Porto, Portugal
Marco A. Wehrmeister	Federal University of Technology – Paraná, Brazil
Markus Vincze	TU Wien, Austria
Martin Hering-Bertram	Hochschule Bremen, Germany
Mikulas Huba	Slovak University of Technology in Bratislava, Slovakia
Michał Podpora	Opole University of Technology, Poland
Miguel Ángel Prada	University of León, Spain
Nicolae Cleju	Technical University of Iasi, Romania
Paulo Lopes dos Santos	University of Porto, Portugal
Paulo Alves	Polytechnic Institute of Bragança, Portugal
Paulo Leitão	Polytechnic Institute of Bragança, Portugal
Paulo Moura Oliveira	University of Trás-os-Montes and Alto Douro, Portugal
Pavel Pakshin	Nizhny Novgorod State Technical University, Russia
Pedro Luiz de P. Filho	Federal University – Paraná, Brazil
Pedro Miguel Rodrigues	Catholic University of Portugal, Portugal
Pedro Morais	Polytechnic Institute of Cávado e Ave, Portugal
Pedro Pinto	Polytechnic Institute of Viana do Castelo, Portugal

Roberto M. de Souza	Federal University of Technology – Paraná, Brazil
Rui P. Lopes	Polytechnic Institute of Bragança, Portugal
Sabrina Šuman	Polytechnic of Rijeka, Croatia
Sani Rutz da Silva	Federal University of Technology – Paraná, Brazil
Santiago Torres Álvarez	University of Laguna, Spain
Sara Paiva	Polytechnic Institute of Viana do Castelo, Portugal
Shridhar Devamane	Global Academy of Technology, India
Sofia Rodrigues	Polytechnic Institute of Viana do Castelo, Portugal
Sławomir Stępień	Poznan University of Technology, Poland
Teresa P. Perdicoulis	University of Trás-os-Montes and Alto Douro, Portugal
Toma Roncevic	University of Split, Croatia
Uta Bohnebeck	Hochschule Bremen, Germany
Valeriana Naranjo-Ornedo	Universidad Politécnica de Valencia, Spain
Vivian Cremer Kalempa	Universidade Estadual de Santa Catarina, Brazil
Vitor Duarte dos Santos	NOVA University Lisbon, Portugal
Wynand Alkema	Hanze University of Applied Sciences, Netherlands
Wojciech Paszke	University of Zielona Gora, Poland
Wojciech Giernacki	Poznan University of Technology, Poland
Wolfgang Kastner	TU Wien, Austria

Artificial Intelligence

Smart Data Driven System for Pathological Voices Classification

Joana Fernandes[1](✉), Arnaldo Candido Junior[2], Diamantino Freitas[3], and João Paulo Teixeira[4]

[1] Research Centre in Digitaization and Intelligent Robotics (CeDRI), Instituto Politecnico de Braganca (IPB), Braganca 5300, Portugal, Faculdade de Engenharia da Universidade do Porto (FEUP), Porto 4200-465, Portugal
joana.fernandes@ipb.pt

[2] São Paulo State University, Institute of Biosciences, Language and Physical Sciences, São José do Rio Preto, SP, Brazil
arnaldo.candido@unesp.br

[3] Faculdade de Engenharia da Universidade do Porto (FEUP), Porto 4200-465, Portugal
dfreitas@fe.up.pt

[4] Research Centre in Digitaization and Intelligent Robotics (CeDRI), Applied Management Research Unit (UNIAG) and Laboratório para a Sustentabilidade e Tecnologia em Regiões de Montanha (SusTEC), Instituto Politecnico de Braganca (IPB), Braganca 5300, Portugal
joaopt@ipb.pt

Abstract. Classifying and recognizing voice pathologies non-invasively using acoustic analysis saves patient and specialist time and can improve the accuracy of assessments. In this work, we intend to understand which models provide better accuracy rates in the distinction between healthy and pathological, to later be implemented in a system for the detection of vocal pathologies. 194 control subjects and 350 pathological subjects distributed across 17 pathologies were used. Each subject has 3 vowels in 3 tones, which is equivalent to 9 sound files per subject. For each sound file, 13 parameters were extracted (jitta, jitter, Rap, PPQ5, ShdB, Shim, APQ3, APQ5, F0, HNR, autocorrelation, Shannon entropy and logarithmic entropy). For the classification between healthy and pathological, several classifiers were used (Decision Trees, Discriminant Analysis, Logistic Regression Classifiers, Naive Bayes Classifiers, Support Vector Machines, Nearest Neighbor Classifiers, Ensemble Classifiers, Neural Network Classifiers) with various models. For each patient, 118 parameters were used (13 acoustic parameters * 9 sound files per subject, plus the subject's gender). As pre-processing of the input matrix data, the Outliers treatment was used using the quartile method, then the data were normalized and, finally, Principal Component Analysis (PCA) was applied in order to reduce the dimension. As the best model, the Wide Neural Network was obtained, with an accuracy of 98% and AUC of 0.99.

Keywords: Speech pathologies · Machine learning · Speech features · Principal component analysis · Vocal acoustic analysis

A. I. Pereira et al. (Eds.): OL2A 2022, CCIS 1754, pp. 419–426, 2022.
https://doi.org/10.1007/978-3-031-23236-7_29

1 Introduction

Vocal pathologies interfere with vocal quality, altering the phonation process. These pathologies affect about 10% of the population [1] and exist in different stages of evolution and severity. Harmful habits lead to a drastic increase in these pathologies [2].

For the detection of pathologies associated with the voice, there are a several exams, which can be perfomed, but they are uncomfortable and invasive for patients, as well as time intensive.

The auditory acoustic analysis intends, in a non-invasive way, to quantify and characterize the voice sound, but it is not objective and will depend on the experience of the physician who performs the evaluation [3]. The development of automatic techniques for acoustic analysis saves time for the patient and gives support to the specialist to improve the accuracy of assessments [4].

Therefore, this article aims to understand which models provide better accuracy rates in the distinction between healthy and pathological.

This article is organized as follows: Sect. 2 describes the state of the art. Section 3 describes the database, the speech parameters and the feature extraction process. Section 4 presents the results obtained for the different classifier models. Finally, Sect. 5 presents the discussion and final conclusions.

2 State of the Art

The automatic recognition of voices with pathologies, using speech processing tools, has shown significant results. Most of the published works classify the sign as having the pathology or not, or by the classification of a specific pathology. Reid et al. 2022 [5], Zhang et al. 2021 [6] and Castellana et al. 2018 [7] are examples where the classification between healthy and pathological voice is made.

Recognition voice pathologies intends to identify the specific pathology that affects the patient's voice. However, this area has few contributions and the group of different pathologies is limited. These results do not allow these systems to be used in clinical evaluations, requiring a deeper investigation with different pathologies.

Ankışhan, 2018 [8] and Rabeh, et al., 2018 [9] obtained an accuracy of 100% in the classification between pathological and healthy and Chen and Chen 2022 [10] obtained an accuracy of 99.4%. Hammami, et al., 2020 [12] obtained an accuracy of 99.26% in the classification between pathological and healthy, however, when distinguishing between hyperfunctional dysphonia and recurrent paralysis, the accuracy is 90.32%. Ali, et al., 2018 [13] for the classification between pathological and healthy obtained an accuracy of 99.72%, however, to distinguish between polyp, keratosis, vocal cord paralysis, nodules and spasmodic dysphonia obtained accuracies of 97.54%, 99.08%, 96.75%, 98.65%, 95.83%, 95.83% and 99.13%. The classification of pathologies in the proposed system is binary in nature. Mojaly, et al., 2014 [14] in the binary classification between healthy and pathological obtained an accuracy of 99.98% with a standard deviation of 0.0263. For the multiclass classification (nodules, vocal cord polyps and keratosis), the accuracy was 99.85% with a standard deviation of 0.1657.

Published works demonstrate that it is possible to classify between healthy and pathological speech with a high degree of accuracy. However, these works use only a few pathologies and do not extend to a wider group of pathologies.

In studies that seek to classify different pathologies, they usually use a very small group of pathologies and with lower accuracy.

3 Materials and Method

This section describes the database, the speech parameters and the feature extraction process.

3.1 Database

The German Saarbrucken Voice Database (SVD) was used. This database is available online by the Institute of Phonetics at the University of Saarland [15]. The database consists of voice signals of more than 2000 subjects with several diseases and controls/healthy subjects. Each person has the recording of phonemes /a/, /i/ and /u/ in the low, normal and high tones, swipe along tones, and the German phrase "Guten Morgen, wie geht es Ihnen?". The size of the sound files is between 1 and 3 s and have a sampling frequency of 50 kHz.

For the analysis, 194 control subjects and 350 pathological subjects distributed across 17 pathologies were used. In the description of the curated database by Fernandes et al. 2019 [16], the total number of subjects is 901. However, since some sound files had some noise, these were excluded. Dysphonia was also excluded, as it is not a pathology, but a voice disorder, therefore, as a pathology, Hyperfunctional Dysphonia (89 subjects), Vocal Cord Paralysis (74), Functional Dysphonia (51), Psychogenic Dysphonia (31), Spasmodic Dysphonia (24), Chronic Laryngitis (23), Vocal Cord Polyp (14), Reinke's Edema (14), Hypofunctional Dysphonia (9), Carcinoma of Vocal Cord (7), Hypopharyngeal Tumor (4), Cyst (3), Granuloma (2), Hypotonic Dysphonia (2), Laryngeal Tumor (1), Intubation Granuloma (1), Fibroma (1).

3.2 Speech Parameters

This section describes the set of features extracted from each sound file.

Jitter is defined as the glottal variation between the vibration cycles of the vocal cords. Subjects who have difficulty controlling the vocal cords usually have higher jitter values [16,17]. The measures to be used of jitter will be Absolute Jitter (jitta), Relative Jitter (jitter), Relative Average Perturbation Jitter (RAP) and Five-point Period Perturbation Quotients Jitter (PPQ5).

Shimmer is defined as the variation in magnitude over the glottal periods. Higher shimmer values may be caused by reduced glottal resistance and injuries, leading to variations in glottal magnitude [16,17]. The shimmer measures to be used are Absolute Shimmer (ShdB), Relative Shimmer (shim), Three-point Amplitude Perturbation Quotient Shimmer (APQ3) and Five-point Amplitude Perturbation Quotient Shimmer (APQ5)

Fundamental Frequency (F0) in a speech signal, the vibration frequency of the vocal cords is considered to correspond to the fundamental frequency [18]. The F0 is determined using the Autocorrelation method with a frame window length of 100 ms and considering a minimum F0 50 Hz.

Harmonic to Noise Ratio (HNR) allows measuring the relationship between the harmonic and noise components of a speech signal, where the periodicity of the signal is provided, quantifying the relationship between the harmonic part and noise. The HNR value of a signal can vary as there are different configurations of the vocal tract, providing different amplitudes for the harmonics [19–21].

Autocorrelation measures the similar parts of the speech signal, repeated throughout the signal. The more similar repetitions there are along the signal, the greater the autocorrelation value [16,22].

Entropy it considers the amount of energy present in a complex system, as it allows quantitative assessment of the degree of randomness and uncertainty of a given data sequence. Entropy analysis makes it possible to accurately assess the nonlinear behavior characteristic of speech signals [23].

3.3 Feature Extraction Process

Each subject has 9 sound files per subject. For each sound file, 13 parameters were extracted (jitta, jitter, RAP, PPQ5, ShdB, Shim, APQ3, APQ5, F0, HNR, autocorrelation, Shannon entropy and logarithmic entropy), making 117 parameters per subject and the subject's sex was added. The input matrix is composed of 118 lines x N subjects.

Considering the work by Silva et al. 2019 [24], where an improvement of up to 13% points was obtained, the data of the input matrix were pre-processed using the treatment of Outliers by the quartile method. This process aims to identify outliers and change their value by a threshold value determined according to the method used.

Then the data were normalized using the resizing method between [−1, 1].

To reduce the dimension, Principal Component Analysis (PCA) was used. This technique uses mathematical concepts such as standard deviation, covariance of eigenvalues and eigenvectors. Starting from the covariance matrix, eigenvectors and eigenvalues are calculated, being necessary to decide the quantity of principal components. Then it is only necessary to calculate the accumulated percentage of the eigenvalues. Therefore, the first eigenvectors that will correspond to 90% or 95% of the accumulated percentage will be selected, meaning that 90% or 95% of the data are explained in the first eigenvectors. Finally, the fitted data are multiplied by the inverse of the selected eigenvector matrix [25]. This analysis was applied to the 118 parameters, with a variance of 95%, resulting in 21 new features.

4 Results and Discussion

In Table 1, it is possible to observe the results for the various classifiers used. For all classifiers, the cross-validation technique was applied, with cross-validation of 10 times.

Subjects were grouped into two groups, one group with control subjects and a second group with all subjects with some pathology. Therefore, the model is a binary classification process.

Table 1. Results of the various techniques used to classify.

Classifier	Model type	Accuracy (%)
Decision Tree	Fine Tree	88.1
	Medium Tree	88.1
	Coarse Tree	89.5
Discriminant Analysis	Linear Discriminant	96.1
	Quadratic Discriminant	94.3
Logistic Regression	Logistic Regression	96.5
Naive Bayes	Gaussian Naive Bayes	85.3
	Kernel Naive Bayes	86.8
Support Vector Machines (SVM)	Linear SVM	96.5
	Quadratic SVM	96.9
	Cubic SVM	97.2
	Fine Gaussian	92.1
	Medium Gaussian SVM	95.2
	Coarse Gaussian	64.9
Nearest Neighbor (KNN)	Fine KNN	89.7
	Medium KNN	84.9
	Coarse KNN	65.8
	Cosine KNN	91.7
	Cubic KNN	84.2
	Weighted KNN	84.2
Ensemble	Boosted Trees	87.1
	Bagged Trees	93.4
	Subspace Discriminant	96.1
	Subspace KNN	96.7
	RUSBoosted Trees	92.1
Neural Network	Narrow Neural Network	97.8
	Medium Neural Network	97.4
	Wide Neural Network	**98.0**
	Bilayered Neural Network	96.7
	Trilayered Neural Network	97.4

Analyzing the results presented in Table 1, it is possible to perceive that the best results are obtained with the Wide Neural Network, where an accuracy of 98% was obtained to distinguish between healthy and pathological subjects. For healthy individuals, in 194 cases it classified 7 incorrectly, and for pathological individuals in 350 it classified 4 incorrectly. However, it is possible to verify the good performance of this network, since it obtained an AUC of 0.99. This model has 100 node in the hidden layer and the activation function used was ReLU. Taking into account the state of the art these results are similar to what already exists, however these results are obtained taking into account more pathologies.

Teixeira et al., 2017 [26] presented hit rates of 100% to distinguish between female subjects and 90% for males, in the distinction between healthy and pathological, however they only used dysphonia as a pathology. Zakariah et al. 2022 [11] in the classification between healthy and pathological obtained an accuracy of 77.49%, as pathologies used Dysphonia, chronic laryngitis, dysody, functional dysphonia, vocal cord cordectomy and leukoplakia. Comparing these two studies with the results obtained, it is possible to perceive that the obtained ones are better, as they obtain a higher rate and consider more pathologies. If the results are compared with the state of the art, these results are similar to what already exists, however, these results are obtained taking into account more pathologies.

5 Conclusion

The objective of this work is to try to find a simple Artificial Intelligence model, which requires little computational load, and which achieves good accuracy rates in the distinction between healthy and pathological subjects, to later be implemented in a system for the detection of vocal pathologies. This system in a first experimental phase will be placed in hospitals and will allow to collect voices of subjects giving an indication if the subject is healthy or pathological. A second phase aims to distinguish the various pathologies.

Taking into account the state of the art, it is possible to see that there are several works where this topic is addressed and with high accuracy rates, however these works usually have few pathologies, even those that do not identify pathologies. In this work, 17 pathologies were considered, thus, 194 control subjects and 350 pathological subjects were used.

For each subject there are 9 sound files, which correspond to 3 vowels, and each vowel has 3 tones. For each sound file, 13 parameters were extracted, giving a total of 117 input features per subject, to which the subject's gender was also added. The input matrix is then composed of 118 lines x N number of subjects.

As pre-processing of the input matrix data, the Outliers treatment was used, where the quartile method was used. Then the data were normalized between [−1,1] and finally the Principal Component Analysis technique was applied in order to reduce the size of the data.

For the classification between healthy and pathological, several classifiers with different models were used. As the best model, the Wide Neural Network was obtained, with an accuracy of 98% and AUC of 0.99. This result is within the best values described in the state of the art.

As future work, we intend to identify the pathology and the degree of severity. However, there are certain difficulties as there are many pathologies with similar symptoms and voice characteristics. The database is still limited in size for the vast majority of pathologies, so an extension process of the database is already underway.

Acknowledgments. The work was supported by the Foundation for Science and Technology UIDB/05757/2020, UIDP/05757/2020 and 2021.04729.BD and by SusTEC LA/P/0007/2021.

References

1. Martins, R.H.G., Santana, M.F., Tavares, E.L.M.: Vocal cysts: clinical, endoscopic, and surgical aspects. J. Voice **25**(1), 107–110 (2011). https://doi.org/10.1016/J.JVOICE.2009.06.008
2. Godino-Llorente, J.I., Gomez-Vilda, P., Blanco-Velasco, M.: Dimensionality reduction of a pathological voice quality assessment system based on gaussian mixture models and short-term cepstral parameters. IEEE Trans. Biomed. Eng. **53**(10), 1943–1953 (2006). https://doi.org/10.1109/TBME.2006.871883
3. Fonseca, E.S., Guido, R.C., Junior, S.B., Dezani, H., Gati, R.R., Mosconi Pereira, D.C.: Acoustic investigation of speech pathologies based on the discriminative paraconsistent machine (DPM). Biomed. Signal Process. Control. **55**, 101615 (2020). https://doi.org/10.1016/J.BSPC.2019.101615
4. Hegde, S., Shetty, S., Rai, S., Dodderi, T.: A survey on machine learning approaches for automatic detection of voice disorders. J. Voice **33**(6), 947.e11-947.e33 (2019). https://doi.org/10.1016/J.JVOICE.2018.07.014
5. Reid, J., Parmar, P., Lund, T., Aalto, D.K., Jeffery, C.C.: Development of a machine-learning based voice disorder screening tool. Am. J. Otolaryngol. **43**(2), 103327 (2022). https://doi.org/10.1016/J.AMJOTO.2021.103327
6. Zhang, X.-J., Zhu, X.-C., Wu, D., Xiao, Z.-Z., Tao, Z., Zhao, H.-M.: Nonlinear features of bark wavelet sub-band filtering for pathological voice recognition. Eng. Lett. **29**(1), 49–60 (2021)
7. Castellana, A., Carullo, A., Corbellini, S., Astolfi, A.: Discriminating pathological voice from healthy voice using cepstral peak prominence smoothed distribution in sustained vowel. IEEE Trans. Instrum. Meas. **67**(3), 646–654 (2018). https://doi.org/10.1109/TIM.2017.2781958
8. Ankışhan, H.: A new approach for detection of pathological voice disorders with reduced parameters. Electrica **18**(1), 60–71 (2018)
9. Hamdi, R., Hajji, S., Cherif, A.: Voice pathology recognition and classification using noise related features. Int. J. Adv. Comput. Sci. Appl. **9**(11), 82–87 (2018). https://doi.org/10.14569/IJACSA.2018.091112
10. Chen, L., Chen, J.: Deep neural network for automatic classification of pathological voice signals. J. Voice **36**(2), 288.e15-288.e24 (2022). https://doi.org/10.1016/J.JVOICE.2020.05.029
11. Zakariah, M., Ajmi Alotaibi, Y., Guo, Y., Tran-Trung, K., Mamun Elahi, M.: An Analytical Study of Speech Pathology Detection Based on MFCC and Deep Neural Networks (2022). https://doi.org/10.1155/2022/7814952
12. Hammami, I., Salhi, L., Labidi, S.: Voice pathologies classification and detection using EMD-DWT analysis based on higher order statistic features. IRBM **41**(3), 161–171 (2020). https://doi.org/10.1016/J.IRBM.2019.11.004

13. Ali, Z., Hossain, M.S., Muhammad, G., Sangaiah, A.K.: An intelligent healthcare system for detection and classification to discriminate vocal fold disorders. Futur. Gener. Comput. Syst. **85**, 19–28 (2018). https://doi.org/10.1016/J.FUTURE.2018.02.021
14. Al Mojaly, M., Muhammad, G., Alsulaiman, M.: Detection and Classification of Voice Pathology Using Feature Selection (2014). https://doi.org/10.1109/AICCSA.2014.7073250
15. Pützer, M., Barry, W.J.: Saarbruecken Voice Database: Institute of Phonetics at the University of Saarland (2007). http://www.stimmdatenbank.coli.uni-saarland.de. Accessed 05 Nov 2021
16. Fernandes, J., Silva, L., Teixeira, F., Guedes, V., Santos, J., Teixeira, J.P.: Parameters for vocal acoustic analysis - cured database. Procedia Comput. Sci. **164**, 654–661 (2019). https://doi.org/10.1016/J.PROCS.2019.12.232
17. Teixeira, J.P., Gonçalves, A.: Algorithm for jitter and shimmer measurement in pathologic voices. Procedia Comput. Sci. **100**, 271–279 (2016). https://doi.org/10.1016/J.PROCS.2016.09.155
18. Hamdi, R., Hajji, S., Cherif, A., Processing, S.: Recognition of pathological voices by human factor cepstral coefficients (HFCC). J. Comput. Sci. (2020). https://doi.org/10.3844/jcssp.2020.1085.1099
19. Fernandes, J., Teixeira, F., Guedes, V., Junior, A., Teixeira, J.P.: Harmonic to noise ratio measurement - selection of window and length. Procedia Comput. Sci. **138**, 280–285 (2018). https://doi.org/10.1016/J.PROCS.2018.10.040
20. Teixeira, J.P., Fernandes, P.O.: Acoustic analysis of vocal dysphonia. Procedia Comput. Sci. **64**, 466–473 (2015). https://doi.org/10.1016/J.PROCS.2015.08.544
21. Teixeira, J.P., Fernandes, J., Teixeira, F., Fernandes, P.O.: Acoustic analysis of chronic laryngitis statistical analysis of sustained speech parameters. In: BIOSIGNALS 2018–11th International Conference on Bio-Inspired Systems and Signal Processing, Proceedings; Part of 11th International Joint Conference on Biomedical Engineering Systems and Technologies, BIOSTEC 2018, vol. 4, pp. 168–175 (2018). https://doi.org/10.5220/0006586301680175
22. Boersma, P.: Stemmen meten met Praat. Stem-, Spraak- en Taalpathologie **12**(4), 237–251 (2004)
23. Araújo, T., Teixeira, J.P., Rodrigues, P.M.: Smart-Data-Driven System for Alzheimer Disease Detection through Electroencephalographic Signals (2022). https://doi.org/10.3390/bioengineering9040141
24. Silva, L., et al.: Outliers treatment to improve the recognition of voice pathologies. Procedia Comput. Sci. **164**, 678–685 (2019). https://doi.org/10.1016/J.PROCS.2019.12.235
25. Teixeira, J.P., Alves, N., Fernandes, P.O.: Vocal acoustic analysis: ANN versos SVM in classification of dysphonic voices and vocal cords paralysis. Int. J. E-Health Med. Commun. **11**, 37–51 (2020)https://doi.org/10.4018/IJEHMC.2020010103
26. Teixeira, J.P., Fernandes, P.O., Alves, N.: Vocal acoustic analysis - classification of dysphonic voices with artificial neural networks. Procedia Comput. Sci. **121**, 19–26 (2017). https://doi.org/10.1016/J.PROCS.2017.11.004