Voice Pathology Detection Using the Adaptive Orthogonal Transform Method, SVM and MLP

Fadwa Abakarim[✉], Abdenbi Abenaou
National School of Applied Sciences, Ibn Zohr University, Agadir, Morocco
`fadwa.abakarim@gmail.com`

Abstract—In this paper, an automatic voice pathology recognition system is realized. The special features are extracted by the Adaptive Orthogonal Transform method, and to provide their statistical properties we calculated the average, variance, skewness and kurtosis values. The classification process uses two models that are widely used as a classification method in the field of signal processing: Support Vector Machine (SVM) and Multilayer Perceptron (MLP). The proposed system is tested by using a German voice database: the Saarbruecken Voice Database (SVD). The experimental results show that the Adaptive Orthogonal Transform method works perfectly with the Multilayer Perceptron Neural Network, which achieved 98.87% accuracy. On the other hand, the combination of the Adaptive Orthogonal Transform method and Support Vector Machine reached 85.79% accuracy.

Keywords—signal processing, voice pathology detection, adaptive orthogonal transform, support vector machine, multilayer perceptron, neural network models

1 Introduction

Signal processing is a science that analyzes and interprets the information contained in a signal. It has a great effect on our daily life. Signal processing is used to solve many problems in various fields, such as medicine, industry and economy.

Every year, the number of people who suffer from voice problems increases, with the percentage being about 25% of the world population [1]. The most affected are people who smoke, suffer from climate change, or who use their voices excessively in their work as a singer, lawyer, teacher [1], [2].

Voice pathology detection is done by an otolaryngologist, also known as an ear, nose, and throat doctor, who performs a painless examination to visualize the vocal cords and the larynx, called indirect laryngoscopy [3]. The doctor puts a light source on the forehead and uses a small specific mirror, called a laryngeal mirror [4], to check the larynx and vocal cords. The doctor may ask the patient to make sounds to examine the mobility of the vocal cords, but depending on the suspected cause of hoarseness, other examinations may be prescribed to refine the diagnosis, such as a phoniatric assessment [5], an X-ray [6], or a CT scan.

The problem of these techniques is that they require several tests to detect the type of disease. In addition, they make patients uncomfortable and sometimes it is difficult for them to go to the hospital every time, especially in severe cases. To avoid these issues, many researchers have proposed various systems of voice pathology recognition.

Muhammad et al., [7] proposed a system of automatic voice pathology detection and classification using vocal tract area irregularity, where features are extracted from the vocal tract area and the classification part is done by the Support Vector Machine method (SVM). The realized system has been tested by using an American and a German database: the Massachusetts Eye and Ear Infirmary (MEEI) database and the Saarbruecken Voice Database (SVD). With MEEI the system reached 99.22% accuracy, and 94.7% accuracy is reached with SVD.

Mohammed et al., [8] realized a system of voice pathology detection and classification using a Convolutional Neural Network model (CNN), the aim of this paper was to improve accuracy by using CNN and ResNet34 layers. To test the performance of the system, they used the Saarbruecken Voice Database (SVD). The authors obtained a result of 95.41% accuracy.

Shia and Jayasree [9] presented a voice pathology detection system using the Discrete Wavelet Transform method (DWT) and Feed Forward Neural Network (FFNN). They used the Saarbruecken Voice Database (SVD) to test their work. The energy of wavelet sub-band coefficients is calculated and FFNN is applied as a classification method. The proposed system achieved 93.3% accuracy.

Hammami, Salhi, and Labidi [10] realized a voice pathology recognition system using Empirical Mode Decomposition (EMD) and Discrete Wavelet Transform (DWT). Features of the Higher Order Statistics (HOS) and the coefficients of the wavelet decomposition are extracted. The database used in this work is the Saarbruecken Voice Database (SVD). This study achieved 99.29% accuracy, and 100% accuracy for detection and classification of voice pathologies.

Srinivasan, Ramalingam, and Arulmozhi [11] proposed a method of voice pathology recognition using Mel-Frequency Cepstral Coefficients (MFCC) for feature extraction, and different models of Neural Network for classification: the Multilayer Perceptron Neural Network (MLPNN), the Generalized Regression Neural Network (GRNN) and the Probabilistic Neural Network (PNN). 20 samples are taken for analysis: 10 for training and 10 for testing. The system achieved 100% accuracy in the classification process using MLPNN.

From these studies, we found that there are several methods widely used for detection and classification of voice pathologies such as MFCCs, DWT, EMD, MLP, and SVM. However, the researchers did not test all the diseases provided by the database used in their work, either with SVD or with MEEI. They chose only a few diseases [1], [7], [10] such as dysphonia, polyps, cysts, or paralysis. Moreover, the number of the voice signals used in the test phase is too small [9]–[12].

In this work, we propose an automatic voice pathology detection and classification system using the Adaptive Orthogonal Transform method to extract the important features from the input signals with a small training dataset [13]–[15]. The classification

phase is done by two of the most well-known methods in the field of signal processing: Support Vector Machine (SVM) [16] and Multilayer Perceptron (MLP) [17], [18], in order to compare them and to see which one will give the best results. To test the performance of the proposed system, the Saarbruecken Voice Database (SVD) [19] is used.

This paper is organized as follows: section 2 presents the proposed method for a voice pathology recognition system based on the Adaptive Orthogonal Transformations. The obtained experimental results are presented in section 3. Finally, we conclude the paper and give an overview about our future study in section 4.

2 Research method

2.1 Proposed system

The proposed system goes through several steps as shown in Figure 1. The first step is the compression of the voice signal to improve the speed of the algorithm using Fourier Transform. Then, the informative features are extracted by the Adaptive Orthogonal Transformations and used as the input of the classification methods SVM and MLP, which identify if the voice is pathological or not.
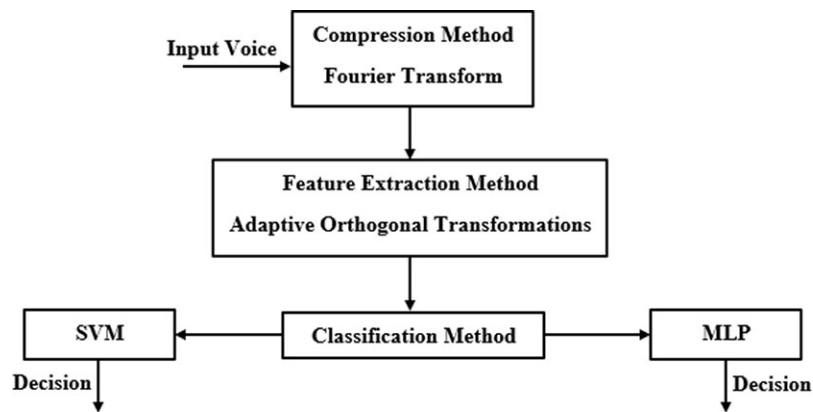

Fig. 1. The steps of the proposed system

Compression. Before starting to extract special features from a voice signal, it is first necessary to compress it by choosing the informative intervals without disturbing the information. In our study, the Fourier Transform is used as a compression method [20], [21].

Figure 2 shows the informative intervals of the pathological and normal voice signals by using the Fourier Transform:
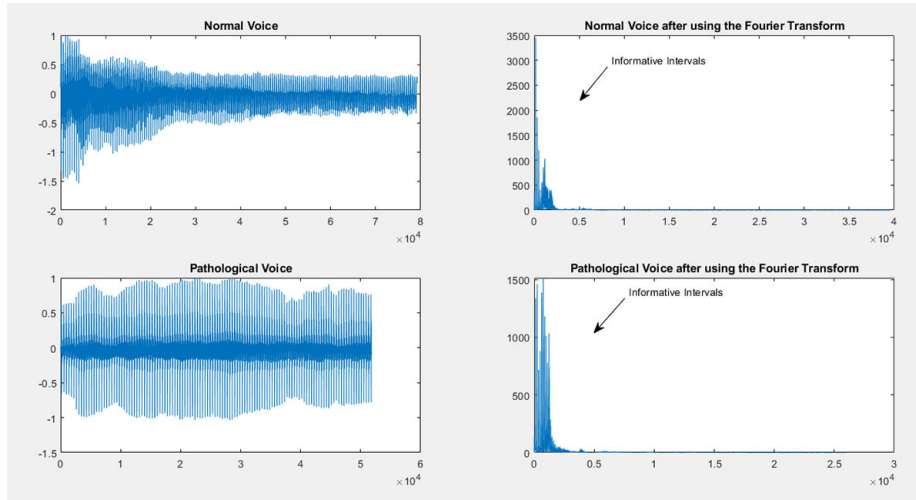
Fig. 2. The informative intervals of the pathological and normal voice signals by using the Fourier Transform

Feature extraction. Many researchers used several methods for feature extraction from the voice signal, such as Mel Frequency Cepstral Coefficients (MFCC) [11], Linear Prediction Cepstral Coefficients (LPPC) [22], Perceptual Linear Prediction Coefficients (PLP) [22], or Empirical Mode Decomposition (EMD) [10]. However, they have several disadvantages: they are sensitive to noise, take a long execution time and need a large training dataset to extract the special features from the input signal [15], [23]–[25]. Our goal is to solve these issues by using the Adaptive Orthogonal Transform method [13]–[15].

The proposed approach can extract the informative features of the voice signal with a minimal dimension by creating an operator H that will be adaptable to any input signal. First, it is necessary to calculate the average of the statistical features obtained at the compression phase to form the $\hat{R}_{sd}$ vector [15], [26], [27].

H is adapted to a class of signals represented by a standard vector $\hat{R}_{sd}$ when the following condition is verified:

$$\frac{1}{N} H_a \hat{R}_{sd} = Y_t = [y_{t1}, 0, \ldots, 0]^T \text{ with } y_{t1} \neq 0 \tag{1}$$

where Y_t is the target vector that constructs the adaptation criterion of the operator H_a to $\hat{R}_{sd}$.

Then the vector of the informative features is calculated by the following operation:

$$\frac{1}{N} H_{a,i} \hat{R}_{sd} = Y_i \tag{2}$$

where:

i is the class number. In our study, we worked with two classes: a pathological class and a normal class.

N is the number of class signals. In our study, the pathological class contains 200 voice signals (100 male and 100 female), and the normal class contains 200 voice signals (100 male and 100 female).

After extracting the vector of the informative features, we decompose it into several frames, and then the average, variance, skewness, and kurtosis of each frame is calculated to provide the statistical properties [28].

To find the number and size of frames corresponding to our vector, we used the following equations:

$$s = t * f \tag{3}$$

where:
s is the size of the frames.
t is the time per frame.
f is the frequency.

$$n = l/s \tag{4}$$

where:
n is the number of frames.
l is length of the vector.
s is the size of the frames obtained in equation (3).

The following equations show the formula of the average, variance, skewness and kurtosis:

$$\text{Average:} Avg = \frac{1}{N} \sum_{i=1}^{N} frm_i \tag{5}$$

$$\text{Variance:} Var = \frac{1}{N-1} \sum_{i=1}^{N} |frm_i - Avg|^2 \tag{6}$$

$$\text{Skewness:} Sk = \frac{\frac{1}{N} \sum_{i=1}^{N} (frm_i - Avg)^3}{(\sqrt{\frac{1}{N} \sum_{i=1}^{N} (frm_i - Avg)^2})^3} \tag{7}$$

$$\text{Kurtosis:} Kr = \frac{\frac{1}{N} \sum_{i=1}^{N} (frm_i - Avg)^4}{(\frac{1}{N} \sum_{i=1}^{N} (frm_i - Avg)^2)^2} \tag{8}$$

where N is the number of frames and frm_i is the *ith* frame.

In our case, the vector is decomposed into 4 frames with 16 different measures (4 frames * 4 values of average, variance, skewness and kurtosis).

Figure 3 shows the overall process of the Adaptive Orthogonal Transform method that goes through 7 steps [27]:

- Step 1: The input speech signals pass through the compression phase using the Fourier Transform.
- Step 2: We calculate the average of the statistical features obtained during the compression part.
- Step 3: The operator synthesis algorithm is applied to the average of the statistical features obtained in the previous step.
- Step 4: The output of the algorithm is the adaptive operator H.
- Step 5: The projection multiplication is applied between the operator H and the rest of the statistical features obtained after using the Fourier Transform as a compression method.
- Step 6: The result of the previous operation is a set of feature vectors that characterizes each signal of the class.
- Step 7: The average of feature vectors is calculated. The result is an informative feature vector with a minimum dimension that characterizes the whole class.
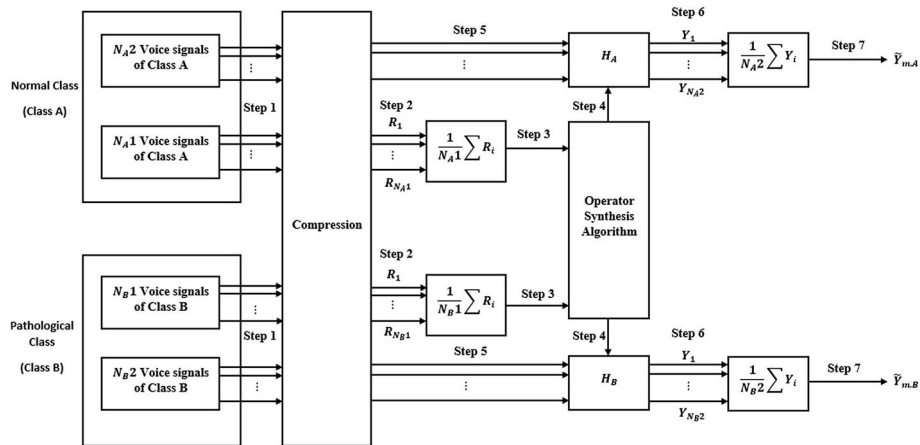
Fig. 3. The overall process of the Adaptive Orthogonal Transform method

Figure 4 shows the decomposition phase of the vector Y obtained as the output of the Adaptive Orthogonal Transformation method, and the calculation of the average, variance, skewness and kurtosis values:
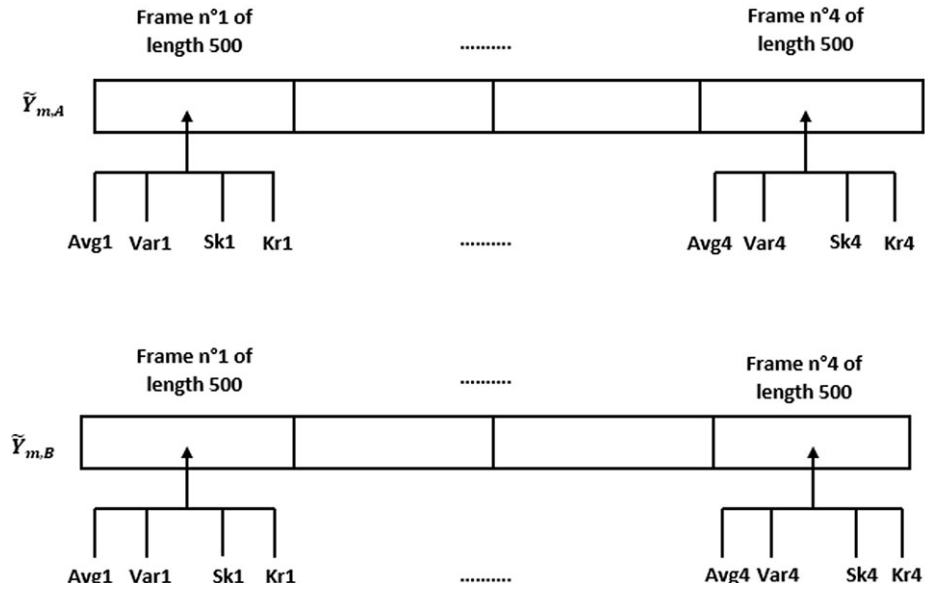
Fig. 4. The decomposition phase of the vector Y and the calculation of the average, variance, skewness and kurtosis values

Classification. After the extraction of the special features from the voice signals, we come to the phase of classification. In our study, we used two methods that are widely used by researchers: Support Vector Machine (SVM) and Multilayer Perceptron (MLP) [29]–[31].

- Support Vector Machine:
  Support Vector Machine (SVM) is very popular for data classification. To create it, we go through several steps. First, the creation of the input vector by assigning the informative features of the two classes: the normal class and the pathological class. Then, it is necessary to choose an efficient kernel, which is an important factor to obtain a good classification. In this work, we used the Radial Basis Function (RBF) kernel. Finally, for the optimization of the kernel parameters, c and gamma, we chose the grid search as an optimization method [7], [16], [32]–[34].
- Multilayer Perceptron:
  Neural Network models are numerous, and among them we find Multilayer Perceptron (MLP) which consists of three layers: the input layer, the hidden layer, and the output layer, as shown in the Figure 5:
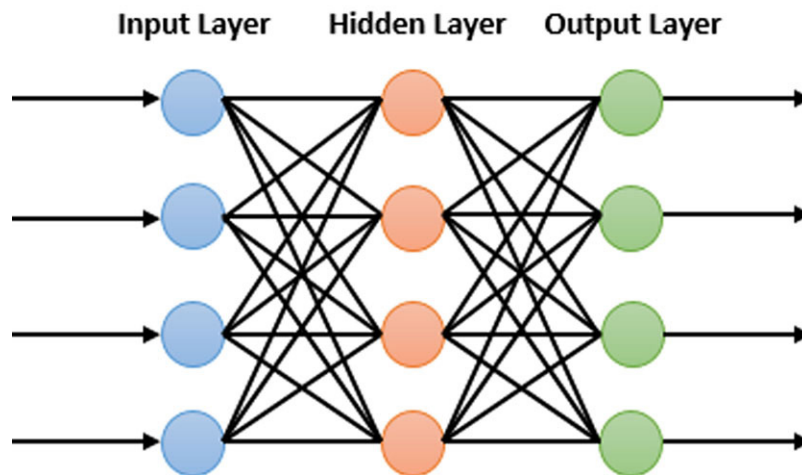
Fig. 5. The structure of MLP

The input layer receives the special features extracted in the previous phase, then an arbitrary number of hidden layers are placed between the input and output layer. Finally the classification and decision are made by the output layer [11], [17].

MLP neurons are trained by using the back-propagation learning algorithm, which consists of two steps: a forward pass and a backward pass through the network. The forward pass consists of passing the input layer by layer to the output neuron that will produce the true output of the network. The backward pass consists of detecting the error signal by deducting the desired output from the actual output [18], [34], [35].

2.2   Data

In this work, we used a German database, the Saarbruecken Voice Database (SVD), which is freely available online and produced by the Institute of Phonetics at Saarland University [19].

This database presents a collection of voice recordings from over 2000 people. It contains sustained vowels /a/, /i/, and /u/ in normal, high, low and low-high-low pitch, including voice recordings of the following sentence "Guten Morgen, wie geht es Ihnen?" which means "Good morning, how are you?" in English. The length of the signals varies between 1 and 4 seconds, and they are sampled at 50 kHz with 16 bits resolution. The number of pathologies provided by the database is 71. All these properties makes SVD very interesting for researchers. In our study, we used the sustained vowel /a/ in a normal pitch.

Table 1 gives more details about the Saarbruecken Voice Database (SVD):

Table 1. The number of records in the Saarbruecken Voice Database (SVD)

	Pathological Voice	Normal Voice	Age Range (years)	Average Age (years)
Male	629	433	06–89	47.5
Female	727	436	09–94	51.5

Table 1 presents the total number of the available records in SVD, but in our work we used only the sustained vowel /a/ in a normal pitch, therefore the values are slightly modified as shown in Table 2.

Table 2. The number of SVD records using in this study

	Pathological Voice	Normal Voice	Age Range (years)	Average Age (years)
Male	627	259	06–89	47.5
Female	727	428	09–94	51.5

Every system needs a training dataset to extract the special features from the input signals and a test dataset to evaluate the performance. Table 3 shows the number of voice recordings in our training and test dataset:

Table 3. Training and test datasets using in our system

	Pathological Voice		Normal Voice	
	Male	Female	Male	Female
Training dataset	100	100	100	100
Test dataset	527	627	159	328

3   Results and discussions

The results of our system performance are obtained by calculating the rate of successivity. The following equation presents the rate of pathological voice recognition:

$$RP = \frac{p}{ts} * 100 \tag{9}$$

where:
p is the number of the recognized pathological voice.
ts is the test dataset size.

The following equation presents the rate of normal voice recognition:

$$RN = \frac{n}{ts} * 100 \tag{10}$$

where:
n is the number of the recognized normal voice.
ts is the test dataset size.

From Table 4, we observe that the combination of our method and the MLP Neural Network model achieves higher results compared to the combination with SVM:

- 98.87% for pathological voice recognition with MLP, which means it successfully identified 1141 pathological voices among 1154.
- 98.36% for normal voice recognition with MLP, which means it successfully identified 479 normal voices among 487.
- 85.79% for pathological voice recognition with SVM, which means it successfully identified 990 pathological voices among 1154.
- 87.27% for normal voice recognition with SVM, which means it successfully identified 425 normal voices among 487.

The difference between the pathological and normal voice recognition rates between the two approaches is 13.08% for pathological voice recognition and 11.09% for normal voice recognition.

We can clearly see that the MLP model works perfectly with the Adaptive Orthogonal Transform method, and it shows the efficiency of our system for the detection and classification of voice pathology. On the other hand, the combination of our approach with the SVM classifier gives also good results, but they are not at the desired level. The execution time of the algorithm was fast with either SVM or MLP.

Table 4. The recognition rate obtained from this study

	Rate Obtained Using the Adaptive Orthogonal Transformations and MLP (%)	Rate Obtained Using the Adaptive Orthogonal Transformations and SVM (%)
Pathological voice	98.87	85.79
Normal voice	98.36	87.27

4 Conclusion

In this paper, our approach is evaluated with two classification models: MLP and SVM. The results obtained show the efficiency of our system by combining the Adaptive Orthogonal Transform method with MLP. The proposed method consists of extracting the informative feature vectors with minimal dimension to improve the speed of the algorithm, then decomposing them and calculating the average, variance, skewness and kurtosis values with the aim of providing their statistical properties. Finally, the obtained values are assigned as input for the SVM and MLP classification models.

Our future work is to combine other Neural Network models such as Probabilistic Neural Network (PNN) and Generalized Regression Neural Network (GRNN) to see if they will achieve more successful results than MLP and SVM or not.

5 References

[1] A. Al-Nasheri *et al.*, "Voice Pathology Detection and Classification Using Auto-Correlation and Entropy Features in Different Frequency Regions," *IEEE Access*, vol. 6, pp. 6961–6974, 2017. https://doi.org/10.1109/ACCESS.2017.2696056

[2] M. S. Hossain, G. Muhammad, and A. Alamri, "Smart healthcare monitoring: a voice pathology detection paradigm for smart cities," *Multimedia Systems.*, vol. 25, no. 5, pp. 565–575, 2019. https://doi.org/10.1007/s00530-017-0561-x

[3] S. R. Collins, "Direct and indirect laryngoscopy: Equipment and techniques," *Respiratory Care*, vol. 59, no. 6, pp. 850–864, 2014. https://doi.org/10.4187/respcare.03033

[4] J. F. Lapenã, "Mirrors and reflections: The evolution of indirect laryngoscopy," *Annals of Saudi Medicine.*, vol. 33, no. 2, pp. 177–181, 2013. https://doi.org/10.5144/0256-4947.2013.177

[5] E. Kankare, A. Geneid, A. M. Laukkanen, and E. Vilkman, "Subjective evaluation of voice and working conditions and phoniatric examination in kindergarten teachers," *Folia Phoniatrica et Logopaedica.*, vol. 64, no. 1, pp. 12–19, 2012. https://doi.org/10.1159/000328643

[6] S. G. Fletcher, R. L. Shelton, C. C. Smith, and J. F. Bosma, "Radiography in speech pathology," *Journal of Speech and Hearing Disorders.*, vol. 25, no. 2, pp. 135–144, 1960. https://doi.org/10.1044/jshd.2502.135

[7] G. Muhammad *et al.*, "Automatic voice pathology detection and classification using vocal tract area irregularity," *Biocybernetics and Biomedical Engineering.*, vol. 36, no. 2, pp. 309–317, 2016. https://doi.org/10.1016/j.bbe.2016.01.004

[8] M. A. Mohammed *et al.*, "Voice pathology detection and classification using convolutional neural network model," *Applied Sciences.*, vol. 10, no. 11, pp. 3723, 2020. https://doi.org/10.3390/app10113723

[9] S. E. Shia and T. Jayasree, "Detection of pathological voices using discrete wavelet transform and artificial neural networks," *2017 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS).*, pp. 1–6, 2017. https://doi.org/10.1109/ITCOSP.2017.8303086

[10] I. Hammami, L. Salhi, and S. Labidi, "Voice Pathologies Classification and Detection Using EMD-DWT Analysis Based on Higher Order Statistic Features," *Irbm*, vol. 41, no. 3, pp. 161–171, 2020. https://doi.org/10.1016/j.irbm.2019.11.004

[11] V. Srinivasan, V. Ramalingam, and P. Arulmozhi, "Artificial Neural Network Based Pathological Voice Classification Using MFCC Features," *International Journal of Science, Environment and Technology.*, vol. 3, no. 1, pp. 291–302, 2014.

[12] K. Ezzine and M. Frikha, "Investigation of glottal flow parameters for voice pathology detection on SVD and MEEI databases," *2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pp. 1–6, 2018. https://doi.org/10.1109/ATSIP.2018.8364517

[13] A. Abenaou, "Development of a Method for the Synthesis of Adaptive Spectral Operators for the Analysis of Random Signals," *International Journal of Computer Theory and Engineering.*, vol. 9, no. 4, 2017. https://doi.org/10.7763/IJCTE.2017.V9.1150

[14] A. Abenaou, F. Ataa Allah, and B. Nsiri, "Vers un système de reconnaissance automatique de la parole en amazighe basé sur les transformations orthogonales paramétrables," *Asinag.*, pp. 133–145, 2014.

[15] F. Abakarim and A. Abenaou, "Amazigh isolated word speech recognition system using the Adaptive Orthogonal Transform Method," *2020 International Conference on Intelligent Systems and Computer Vision (ISCV)*, pp. 1–6, 2020. https://doi.org/10.1109/ISCV49265.2020.9204291

[16] I. Hammami, L. Salhi, and S. Labidi, "Pathological voices detection using Support Vector Machine," *2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pp. 662–666, 2016. https://doi.org/10.1109/ATSIP.2016.7523162

[17] J. I. Godino-Llorente and P. Gómez-Vilda, "Automatic Detection of Voice Impairments by Means of Short-Term Cepstral Parameters and Neural Network Based Detectors," *IEEE Transactions on Biomedical Engineering.*, vol. 51, no. 2, pp. 380–384, 2004. https://doi.org/10.1109/TBME.2003.820386

[18] L. Salhi, T. Mourad, and A. Cherif, "Voice Disorders Identification Using Multilayer Neural Network," *International Arab Journal of Information Technology.*, vol. 7, no. 2, pp. 177–185, 2010.

[19] Barry WJ, Putzer M. Saarbruecken voice database. Available from: http://www.stimmdatenbank.coli.uni-saarland.de/

[20] J. Obuchowski, A. Wyłomańska, and R. Zimroz, "Selection of informative frequency band in local damage detection in rotating machinery," *Mechanical Systems and Signal Processing*, vol. 48, no. 1–2, pp. 138–152, 2014. https://doi.org/10.1016/j.ymssp.2014.03.011

[21] R. Mankar, M. J. Walsh, R. Bhargava, S. Prasad, and D. Mayerich, "Selecting optimal features from Fourier transform infrared spectroscopy for discrete-frequency imaging," *Analyst*, vol. 143, no. 5, pp. 1147–1156, 2018. https://doi.org/10.1039/C7AN01888F

[22] A. Alatwi, S. So, and K. K. Paliwal, "Perceptually motivated linear prediction cepstral features for network speech recognition," *2016 10th International Conference on Signal Processing and Communication Systems (ICSPCS)*, 2016. https://doi.org/10.1109/ICSPCS.2016.7843309

[23] M. Cutajar, E. Gatt, I. Grech, O. Casha, and J. Micallef, "Comparative study of automatic speech recognition techniques," *IET Signal Processing.*, vol. 7, no. 1, pp. 25–46, 2013. https://doi.org/10.1049/iet-spr.2012.0151

[24] D. Gupta, P. Bansal, and K. Choudhary, "The state of the art of feature extraction techniques in speech recognition," *Speech and Language Processing for Human-Machine Communications.*, vol. 664, pp. 195–207, 2018. https://doi.org/10.1007/978-981-10-6626-9_22

[25] R. K. Shukla, S. Gupta, R. S. Shukla, and R. K. Shukla, "Literature survey and review of techniques used for automatic assessment of Stuttered Speech," *International Journal of Management Technology And Engineering.*, vol. 9, pp. 239–237, 2019.

[26] M. Azergui, A. Abenaou, and H. Bouzahir, "Bearing fault classification based on the adaptive orthogonal transform method," *International Journal of Advanced Computer Science and Applications.*, vol. 9, no. 1, pp. 375–380, 2018. https://doi.org/10.14569/IJACSA.2018.090151

[27] F. Abakarim and A. Abenaou, "Comparative study to realize an automatic speaker recognition system," *International Journal of Electrical and Computer Engineering.*, vol. 12, no. 1, pp. 376–382, 2022. https://doi.org/10.11591/ijece.v12i1.pp376-382

[28] J. Lee and M. Hahn, "Automatic assessment of pathological voice quality using higher-order statistics in the LPC residual domain," *EURASIP Journal on Advances in Signal Processing.*, vol. 2009, pp. 1–8, 2010. https://doi.org/10.1155/2009/748207

[29] M. L. Al-dabag, H. T. S. ALRikabi, and R. R. O. Al-Nima, "Anticipating Atrial Fibrillation Signal Using Efficient Algorithm," *International Journal of Online and Biomedical Engineering.*, vol. 17, no. 2, pp. 106–120, 2021. https://doi.org/10.3991/ijoe.v17i02.19183

[30] A. Bourouhou, A. Jilbab, C. Nacir, and A. Hammouch, "Heart Sound Signals Segmentation and Multiclass Classification," *International Journal of Online and Biomedical Engineering.*, vol. 16, no. 15, pp. 64–79, 2020. https://doi.org/10.3991/ijoe.v16i15.16817

[31] M. K. Moridani and S. Marjani, "A Review of the Methods for Sudden Cardiac Death Detection: A Guide for Emergency Physicians," *International Journal of Online and Biomedical Engineering.*, vol. 16, no. 9, pp. 137–158, 2020. https://doi.org/10.3991/ijoe.v16i09.14485

[32] Z. Ali *et al.*, "Voice pathology detection based on the modified voice contour and SVM," *Biologically Inspired Cognitive Architectures.*, vol. 15, pp. 10–18, 2016. https://doi.org/10.1016/j.bica.2015.10.004

[33] Z. Fan, J. Qian, B. Sun, D. Wu, Y. Xu, and Z. Tao, "Modeling Voice Pathology Detection Using Imbalanced Learning," *2020 International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD)*, pp. 330–334, 2020. https://doi.org/10.1109/ICSMD50554.2020.9261679

[34] W. Yuanbo, Z. Changwei, F. Ziqi, Z. Yihua, Z. Xiaojun, and T. Zhi, "Voice Pathology Detection and Multi-classification Using Machine Learning Classifiers," 2020 *International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD)*, pp. 319–324, 2020. https://doi.org/10.1109/ICSMD50554.2020.9261710

[35] W. Gevaert, G. Tsenov, and V. Mladenov, "Neural networks used for speech recognition," *Journal of Automatic Control*, vol. 20, no. 1, pp. 1–7, 2010. https://doi.org/10.2298/JAC1001001G

6   Authors

Fadwa Abakarim received in 2015 the University Diploma of Technology in Computer Networks at the Higher School of Technologies of Agadir, Morocco. In 2016, she obtained the Professionnel License degree in Networks, Security and Information systems at the Faculty of Sciences Semlalila of Marrakech, Morocco. In 2018, she received the Specialized Master degree in IT Offshoring at the Faculty of Sciences of Agadir, Morocco. She is currently Ph.D. student in Ibn Zohr University of Agadir at the National School of Applied Sciences, Agadir, Morocco. Her current research interests are methods and algorithms for the recognition of signals.

Abdenbi Abenaou graduated from The Bonch-Bruevich St. Petersburg, State University of Telecommunications in 1998. He obtained his Ph.D degrees in technical sciences from the St. Petersburg State Electrotechnical University "LETI" in 2005. Now he is a professor of the Department of Software Engineering at the National School of Applied Sciences of Agadir, Morocco. His research interests are methods and algorithms for the recognition of signals and images.