Received July 30, 2019, accepted August 10, 2019, date of publication August 29, 2019, date of current version September 13, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2938265

Leveraging Artificial Intelligence to Improve Voice Disorder Identification Through the Use of a Reliable Mobile App

LAURA VERDE[1], GIUSEPPE DE PIETRO[1], (Member, IEEE), MUBARAK ALRASHOUD[2], (Member, IEEE), AHMED GHONEIM[2], (Member, IEEE), KHALED N. AL-MUTIB[2], AND GIOVANNA SANNINO[1], (Member, IEEE)

[1]Institute of High Performance Computing and Networking (ICAR-CNR), 80131 Naples, Italy

[2]Department of Software Engineering (SWE), College of Computer and Information Sciences (CCIS), King Saud University, Riyadh 11543, Saudi Arabia

Corresponding author: Giovanna Sannino (giovanna.sannino@icar.cnr.it)

The authors extend their appreciation to the Deanship of Scientific Research at King Saud University, Riyadh, Saudi Arabia for funding this work through the research group project no. RG-1440-135.

ABSTRACT The evolution of the Internet of Things, cloud computing and wireless communication has contributed to an advance in the interconnectivity, efficiency and data accessibility in smart cities, improving environmental sustainability, quality of life and well-being, knowledge and intellectual capital. In this scenario, the satisfaction of security and privacy requirements to preserve data integrity, confidentiality and authentication is of fundamental importance. In particular, this is essential in the healthcare sector, where health-related data are considered sensitive information able to reveal confidential details about the subject. In this regard, to limit the possibility of security attacks or privacy violations, we present a reliable mobile voice disorder detection system capable of distinguishing between healthy and pathological voices by using a machine learning algorithm. This latter is totally embedded in the mobile application, so it is able to classify the voice without the necessity of transmitting user data to or storing user data on any server. A Boosted Trees algorithm was used as the classifier, opportunely trained and validated on a dataset composed of 2003 voices. The most frequently considered acoustic parameters constituted the inputs of the classifier, estimated and analyzed in real time by the mobile application.

INDEX TERMS Voice disorders, smart healthcare monitoring, artificial intelligence algorithms, smart cities, security, privacy.

I. INTRODUCTION

Nowadays, smart cities have the new potential to improve the quality of life of the citizens. The main services offered, such as education, transport or healthcare, can benefit from interactions in real time between several stakeholders, the collection and transmission of a massive amount of data relating to the fundamental services of a smart city. If, on the one hand, intelligent management systems represent the foundations of a smart city, on the other, they also constitute the main source of potential vulnerability.

Although, in fact, the use of advanced tools and technologies contributes to support the collection, storage and transmission of data and to provide smart planning ideas, construction models and data management approaches [1], the great amount of heterogeneous data, either extracted from the environment or provided by the citizens, can be subject to security attacks or privacy violations [2]. The enforcement of security and privacy policies that preserve data confidentiality and authentication and access control are fundamental, particularly in the healthcare sector.

In recent years, a large number of wearable sensors have started to generate enormous amounts of data and several mobile health (m-health) applications have been developed for the continuous promotion and monitoring of healthy lifestyles, personal fitness and physical activity. Such applications contribute to improve the quality of care through better-personalized diagnosis systems and solutions: the patient's traditional regular appointment at the clinic can be substituted by video-conferencing, remote monitoring, electronic consultations and wireless communications, removing any potential barriers to care and reducing healthcare

The associate editor coordinating the review of this article and approving it for publication was Mehedi Masud.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/

costs [3]. This continuous monitoring permits patients to receive the right care, at the right place, at the right time, as in the solutions described in [4]–[9]. However, it is essential that the data exchange between the patient and medical specialist is safe and secure [10], and indeed, several secure systems for data trasmission for smart cities have been proposed in the literature [11]–[13].

Based on these considerations, this paper proposes a reliable mobile voice disorder detection system able to detect voice disorders by using a machine learning (ML) algorithm. A trained model was directly embedded in a mobile application, allowing the user to evaluate the health of his or her voice, anywhere and at any time, without the necessity of transmitting user data to or storing user data on any server. This avoids the possibility that the patient data could be subject to security attacks or privacy violations, undermining integrity, confidentiality and authentication. The capability of processing and analyzing the voice signal directly on a mobile device constitutes, at the time of writing, a significant innovation on account of the fact that none of the existing studies in literature has proposed a similar application. Most of these studies, in fact, limit the use of the mobile device to the tasks of acquiring the useful signal, transmitting it to an external server to be analyzed and visualizing and communicating the results obtained to the users [14], [15]. Unfortunately, the transmission of these patient data can be subject to asecurity attacks on security or privacy violations. The proposed mobile system, instead, is able to acquire in real time the voice signal, process it to extract the characteristic acoustic features and analyze these signs to classify the voice as healthy or pathological directly on the mobile device, without to transmitting any data, so limiting the probability of any security attack.

To define the classifier, we have used artificial intelligence techniques. Although the functioning of these techniques can be difficult to interpret, they can integrate, quickly and easily, different variables and produce reliable results in complex situations so as to support the correct diagnosis of a disorder [16]–[18] or the handling of large amounts of data [19], [20].

The remaining sections of the paper are organized as follows. Dysphonia and the main techniques for its detection are discussed in Section II. The proposed mobile system is introduced in Section III, while the experimental phase is described in Section IV. Finally, Section V presents our conclusions.

II. BACKGROUND AND RELATED WORK

Dysphonia is an alteration of the voice production due to anatomical and functional complications of the pneumo-phono articulatory apparatus. It is widespread, mainly among the so-called vocal professionals, people who routinely use their voice in their work, such as teachers, singers or lawyers. However, its symptoms are often underestimated, constituting a potential obstacle to appropriate treatment [21].

In the clinical practice, the medical expert follows an appropriate protocol, for example in Italy the SIFEL (Societá Italiana di Foniatria e Logopedia - the Italian Society of Logopedics and Phoniatrics) protocol [22] is used to diagnose the presence or not of dysphonia. Such a protocol provides a series of different examinations, such as the anamnestic evaluation to analyze the behavioral characteristics and vocal attitudes of the subject and the familial history of the disorder, the laryngo-videostroboscopic examination to detect any physiological or morphological laryngeal alteration, the acoustic analysis, and the subjective self-assessment of the voice. Among these examinatios, the acoustic analysis is used to give an objective quantification of the health of the speech signal through an evaluation of characteristic parameters calculated from a recording of the vowel /a/ of five seconds in length. The acoustic analysis is a useful instrument to evaluate the presence of a possible laryngeal alteration that can cause a deviation in vocal quality.

Currently, a rule-based approach is used in the acoustic analysis to evaluate the presence of a voice disorder. The acoustic parameters estimated are evaluated according to IF/ELSE rules. The clinical expert compares the values of the parameters with a fixed healthy range. The voice is healthy if these values are included within the healthy range, pathological if otherwise. Unfortunately, a standard healthy range does not exist but changes from laboratory to laboratory. <This influences the scale validity and reliability in the detection of a voice disorders.

The use of machine learning techniques allows us to overcome this limitation. In fact, these techniques not only allow the evaluation of all the vocal parameters, providing a global evaluation of vocal quality, but also avoid the use of a fixed healthy range whose choice can affect the reliability of the system.

Mobile healthcare frameworks capable of detecting dysphonia were proposed in [14] and [15]. The detection of the voice pathology was performed by a convolutional neural network (CNN) in both cases. Alternatively in [23], a Gaussian mixture model-based approach constitutes the classifier to distinguish healthy and pathological voices, analyzing features extracted from voice and electroglottographic (EGG) signals. Other automatic smart healthcare monitoring system for the detection of voice disorder is described in [24]. Additionally, in this case, a Gaussian Mixture Models (GMM) was used to disscriminate between pathological and healthy subjects.

Several studies existing in literature use GMM to distinguish between pathological and healthy voices, such as [25], [26]. However, different data mining approaches have been proposed to estimate the presence of a voice disorder [27]. Besides GMM, Support Vector Machine (SVM) and Decision Tree (DT) are widely used to discriminate between pathological and healthy voices [28]–[31].

The choice of the classifier is fundamental in order to distinguish correctly between healthy and pathological voices. An additional important task is the choice of features to

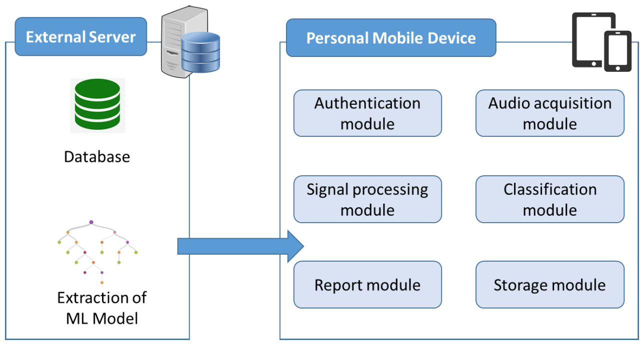

FIGURE 1. Voice disorder detection system.

extract from the signal to use as the input of the classifier. The most frequently used features in data mining approaches are the main acoustic parameters evaluated in clinical practice, such as jitter or shimmer, Mel-Frequency Cepstral Coefficients (MFCC), temporal derivatives and noise parameters [28], [30], [32].

It is important to note that the majority of the solutions presented use the mobile system only to capture the signals, but they do not involve the performance of the analysis on the mobile application. These signals are transmitted to an external machine (server) to be analyzed for an evaluation of the presence or not of a voice disorder. Unfortunately, this transmission can be a source of attacks of security attacks or privacy violations. To limit this possibility, a solution may be to analyze the voice signals directly on the mobile devices.

Certain other solutions, such as for example OperaVox [33] or the solution proposed by Van Leer et al in [34], are iOS-based apps capable of analyzing the signals on a mobile device. However, these solutions are only able to estimate certain acoustic parameters, including the fundamental frequency (F_0) and several perturbation measurements. They indicate to the user the values of the calculated parameters. However, these results cannot be easily interpreted without the support of an expert who can detect the presence of a voice disorder by comparing the value obtained for each parameter with a fixed healthy range. As indicated previously, the choice of a healthy range is not standard and, indeed, this deficiency may influence the correct detection of a voice disorder. The use of machine learning techniques allows us to overcome these limitations.

III. MOBILE VOICE DISORDER DETECTION SYSTEM

The proposed mobile voice disorder detection system is capable of distinguishing between pathological and healthy voices by using an artificial intelligence classifier on a mobile device. A valid machine learning algorithm, opportunely trained on an external server, has been embedded in a mobile application. This device is able to acquire a voice signal, process it to extract appropriate features and classify the new data by using the trained model.

The proposed system, illustrated in Figure 1, consists of two main components: an external server and a personal mobile device.

The former is necessary to define and fit the artificial intelligence model able to classify the voice signal as healthy or pathological. A Boosted Trees algorithm was used as the classifier [35], opportunely trained by using voice voices samples selected from three different databases. The classifier and dataset used to train the classifier are described, in detail, in the following sections. The trained model was saved and deployed to the mobile device by using the Simulink Support Package for Android devices of Matlab R2019a [36], in Intel(R) Core (TM) i7-4770 CPU@3.40GHz, 16 GB RAM based Windows 10. All the tests were performed using a Samsung S4 device, its operating system version being Android 5.0.1.

The personal mobile component is deployed as a mobile application running on an Android smartphone given to the patient, and is in charge of providing an interface between the user and the monitoring system, such as the one described in [37]. In detail, it is composed of the following modules:

- an Authentication module: this is responsible for collecting the personal data of the user. In detail, the user inserts her/his age and gender, information necessary for the subsequent analysis, due to the relationship between these data and the presence of a voice disorders [38]–[40];
- an Audio acquisition module: this allows the recording of a vocalization of the vowel /a/ of five seconds in length without any interruption of sound, necessary to perform the subsequent extraction of the acoustic features. An opportune filter was adopted to reduce the effects of noise accidentally added during the acquisition [41];
- a Signal processing module: this is responsible for processing in real time the acquired voice signal. The main acoustic parameters are estimated, such as the F_0, jitter, shimmer and Harmonic to Noise Ratio (HNR), used as inputs of the classifier;
- a Classification module: this classifies the voice sample as healthy or pathological, analyzing the characteristic features provided as the input;
- a Report module: this module is in charge of generating appropriate reports containing the analysis made with the goal of saving such reports in pdf documents stored directly on the mobile device. In this way, the patient and/or the medical personnel involved can visualize them at their convenience; and
- a Storage module: this module saves the results of the analyses and acquired audio recordings on the mobile device so that they can be visualized at any time and anywhere.

A. CLASSIFIER

Boosted Trees (BT) is the machine learning algorithm used to assess voice disorders. This is a supervised learning algorithm able to distinguish between a pathological and a healthy voice by combining several decision trees to obtain a better predictive performance than would be possible with a single decision tree [35]. Decision Tree is an accurate predictor for the classification problem, particularly used due to its simplicity, speed, ease of implementation, interpretation and explaination, and capability of working with missing values and categorical and continuous data, advantageous which are fundamental in health data analysis. It is composed of nodes and branches. In this case, the branches constitute the rules that the algorithm follows to predict a voice disorder, while the nodes are the predictors that influence the predictive path [42]. In detail, in this study we used Ensemble Decision Tree. The technique employed for its realization is Boosting, by means of the AdaBoost algorithm [43]. This algorithm gives a weight to each training example and iteratively weighted models are built. The algorithm focuses on the more "difficult" samples that are then weighed frequently. The idea is, in fact, to focus on that subset of data that cannnot be properly classified. The realized algorithm has a number of learners equal to 30, while the maximum number of splits is 20.

The inputs of the classifier are the F_0, jitter, shimmer and HNR. These features constitute the most commonly used measurements evaluated in the clinical diagnosis of dysphonia in order to make a quantitative voice evaluation according to the SIFEL protocol [22]. The use of these acoustic parameters provides a specific focus on several aspects of the pneumo-phono articulatory apparatus and voice production mechanism. Measurements of instabilities in the voice signal or the noise content, indices of voice alterations, are considered. In our study, the former were evaluated through the jitter (frequency pertubation) and shimmer (amplitude pertubation), estimated by using the methods indicated in [44], while the latter was estimated by using the HNR, calculated by using de Krom's algorithm [45]. Finally, the F_0 represents the rate of vibration of the vocal folds and its alteration indicates the presence of possible modifications of the laryngeal function. It was estimated using our personalized methodology described in [46]. Additionally, other inputs of classifier are the age and gender of the subject, data which are important due to their relationship with the occurrence of alterations of the pneumo-phono articulatory apparatus.

IV. EXPERIMENTAL PHASE

The capability of the Boosted Trees algorithm to distinguish correctly between a healthy and pathological voice was evaluated through an appropriate experimental phase. Several tests were carried out using voice samples extracted from opportune databases.

The goodness of the considered ML algorithm was evaluated according to its classification accuracy in distinguishing between pathological and healthy voices. In detail, the accuracy, sensitivity, specificity, precision and F-measure were calculated. We identified the number of true positives (TP) and true negatives (TN), these being the outcomes where the classifier correctly predicts, respectively, the presence and the absence of a voice disorder, and the number of false positives (FP) and false negatives (FN), these being the outcomes where the classifier incorrectly predicts, respectively, the presence and the absence of a voice disorder. These metrics were calculated as:

TABLE 1. Number of voice samples used in this study for each database.

Database	Category	Gender	No
MEEI [48]	Healthy	Female	32
		Male	21
	Pathological	Female	228
		Male	144
SVD [49]	Healthy	Female	428
		Male	257
	Pathological	Female	428
		Male	257
VOICED [50]	Healthy	Female	37
		Male	21
	Pathological	Female	98
		Male	52
Total	*Healthy*	*all*	*796*
	Pathological	*all*	*1207*

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (1)$$

$$Sensitivity = \frac{TP}{(TP + FN)} \quad (2)$$

$$Specificity = \frac{TN}{(TN + FP)} \quad (3)$$

$$Precision = \frac{TP}{(TP + FP)} \quad (4)$$

$$F - measure = \frac{2 * Precision * Sensitivity}{(Precision + Sensitivity)} \quad (5)$$

Additionally, we performed a receiver operating characteristic (ROC) analysis. The performance of the classifier was measured by calculating the area under the ROC curve (AUC). The maximum, namely AUC = 1, was obtained when the model correctly classified all the voice samples, while the minimum (AUC = 0), was achieved when the model incorrectly classified all the samples.

The reliability of the classifier was validated by using a 5-fold cross validation. This is also useful to limit the effects of overfitting, one of the major causes of a possible unreliable analysis [47]. The dataset was, randomly, divided into 5 subsets: 80% of the data is used as training set, the remaining 20% is used as test set. The tests were repeated until each subset had been used as the test set.

In the following subsections, the dataset used to evaluate the classification accuracy and the results achieved are described.

A. DATABASE

Voice samples selected from the Massachusetts Eye and Ear Infirmary (MEEI) [48], the Saarbruecken Voice Database (SVD) [49] and the VOice ICar fEDerico II (VOICED) [50] database constituted the dataset used to test the reliability of the classifier.

TABLE 2. Comparison of the results obtained with several classifiers.

	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F-measure (%)	AUC
BT [35]	**84.5**	82.9	**86.2**	**85.7**	**84.3**	**0.91**
SVM [52]	53.8	79.0	27.9	52.3	62.9	0.70
DT [42]	81.1	77.9	84.4	83.3	80.5	0.84
NB [53]	75.1	**85.7**	64.4	70.7	77.4	0.85
KNN [54]	55.4	77.4	33.4	53.7	63.4	0.67

The dataset consists of 2003 voice samples, 796 healthy voices and 1207 pathological voices. All the samples contain the recording of the vowel /a/, as indicated in the SIFEL protocol [22]. Pathological voices are defined as voices belonging to subjects suffering from one of several voice pathologies. Functional and organic disorders are considered, the former due to an inefficient and improper use of the vocal mechanism (e.g. vocal abuse), the latter, instead, being caused by alterations of the respiratory, laryngeal or vocal tract mechanisms.

Table 1 reports the number of pathological and healthy voices contained in the considered dataset for each database. In detail, the MEEI database contains voice samples of healthy subjects and of those suffering from several voice diseases, collected by the MEEI Voice and Speech Laboratory. All the samples were recorded with a sample frequency equal to 50 kHz and 25 Hz with a 32-bit resolution.

A resolution of 16-bit and a sample frequency equals to 50 kHz were used, instead, to record the voice samples contained in SVD database. Recordings of sustained /a/, /i/ and /u/ vowels were collected by the Institute of Phonetics of the University of Saarland in collaboration with the Department of Phoniatrics and Ear, Nose and Throat (ENT) at the Caritas clinic St. Theresia in Saarbruecken.

Finally, voice samples selected from the VOICED database were collected by the Institute of High Performance Computing and Networking of the National Research Council of Italy (ICAR- CNR) in collaboration with the Hospital University Hospital of Naples Federico II, available on the PhysioNet website [51].

B. PERFORMANCE

To validate and estimate the goodness of the proposed classifier, we compared the classification performance of the Boosted Trees algorithm with the most commonly used and cited machine learning classifiers.

An exhaustive comparison phase was carried out with different machine learning algorithms. These latter have been compared in terms of classification accuracy in distinguishing between pathological and healthy voices. It is important to note that we report only the performances of the classifiers that achieved the best classification accuracy, chosen as representative of a class of algorithms based on similar characteristics, namely:

- Support Vector Machine (SVM) [52];
- single Decision Tree (DT) [42];
- Naive Bayes (NB) [53]; and
- K-nearest neighbour (k-NN) [54].

SVM constitutes one of the most commonly used machine learning algorithms. The aim of this algorithm is to identify the class of belonging of certain data, determining the optimal hyperplane, equally distant from the support vectors from different classes. Its performance can be improved by choosing the opportune kernel function form. In this study, we used a polynomial kernel with a degree equal to 1. The Naive Bayes algorithm, instead, is based on a probabilistic approach to classifying data where nodes and strings represent a set of random variables and their conditional dependencies. According to this classifier the effect of an attribute is independent of the other attributes and, additionally, it considers that hidden attributes do not influence the prediction process. Finally, k-NN is an instance-based learning algorithm, where the classification is based on k nearest neighbours of a new instance. In this study, we considered $k = 10$ neighbours.

Table 2 shows the performances in terms of accuracy, specificity, sensitivity, precision, f-measure and ROC area (AUC). These results show that the best performance was, generally, achieved by using the Boosted Trees algorithm. Although, in fact, the best sensitivity (about 86%) was achieved by the NB classifier, the best accuracy classification, as well as the best results achieved in the other performance metrics, were obtained with the Boosted Trees algorithm. The classification accuracy obtained was, in fact, equal to about 85%, higher than that of the other ML algorithms, namely SVM and kNN, which achieved an accuracy, respectively, equal to 53.8% and 55.4%. The obtained values of sensitivity and specificity, respectively of about 83% and 86%, indicate that Boosted Trees has a lower number of false negatives and false positives (i.e. pathological voices erroneously classified as non-pathological or healthy voices erroneously classified as pathological) in comparison with the other algorithms.

Observing the results obtained, we can note that there is an improvement in the performance using Boosted Trees rather than the single DT classifier. The accuracy achieved by BT is higher than that obtained by DT, principally due to an improvement in the detection of pathological voices, as demonstrated by the higher sensitivity achieved by using BT rather than DT the algorithm.

V. CONCLUSION

Currently, smart healthcare has gained attention in the detection and monitoring of specific pathologies thanks to the possibility of offering healthcare services anywhere and at any time, improving the quality of life of the patients.

In this paper, a reliable mobile voice disorder detection system is proposed capable of detecting the presence of voice disorders. This system is able to acquire a voice signal, process it to extract the acoustic parameters and classify a voice as healthy or pathological through a machine learning algorithm, by using a mobile application. Boosted Trees was used as the classifier. The performances were opportunely tested on a dataset composed of voice samples selected from three databases. The results obtained have demonstrated that this classifier achieved the best classification accuracy compared with the performances of the main machine learning algorithms existing in literature.

In our future plans, we aim to improve the classification accuracy, optimizing performances of the classifier using appropriate voice features, opportunately selected form feature selectors. The classification accuracy obtained can be compared with that achieved by the main deep learning algorithms. Further tests on several Android mobile devices are necessary to evaluate the reliability of the proposed mobile system. Finally, we are planning a new usability study on this current version of the mobile application in order to collect users' opinions about itsefficiency, any difficulties encountered using this system, the perceived quality of the information provided and any preferences for different application features, as performed with preliminary version of the app described in [55].

REFERENCES

[1] V. Albino, U. Berardi, and R. M. Dangelico, "Smart cities–definitions, dimensions, and performance," in *Proc. IFKAD*, 2013, pp. 1723–1738.

[2] S. Sicari, A. Rizzardi, L. A. Grieco, and A. Coen-Porisini, "Security, privacy and trust in Internet of Things: The road ahead," *Comput. Netw.*, vol. 76, pp. 146–164, Jan. 2015.

[3] D. A. M. Budida and R. S. Mangrulkar, "Design and implementation of smart HealthCare system using IoT," in *Proc. Int. Conf. Innov. Inf., Embedded Commun. Syst. (ICIIECS)*, Mar. 2017, vol. 5, no. 3, pp. 1–7.

[4] S. Banka, I. Madan, and S. Saranya, "Smart healthcare monitoring using IoT," *Int. J. Appl. Eng. Res.*, vol. 13, no. 15, pp. 11984–11989, 2018.

[5] S. Kajornkasirat, N. Chanapai, and B. Hnusuwan, "Smart health monitoring system with IoT," in *Proc. IEEE Symp. Comput. Appl. Ind. Electron. (ISCAIE)*, Apr. 2018, pp. 206–211.

[6] D. C. Yacchirema, D. Sarabia-Jácome, C. E. Palau, and M. Esteve, "A smart system for sleep monitoring by integrating IoT with big data analytics," *IEEE Access*, vol. 6, pp. 35988–36001, 2018.

[7] M. S. Hossain, S. U. Amin, M. Alsulaiman, and G. Muhammad, "Applying deep learning for epilepsy seizure detection and brain mapping visualization," *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 15, no. 1s, 2019, Art. no. 10.

[8] M. S. Hossain and G. Muhammad, "Emotion recognition using deep learning approach from audio–visual emotional big data," *Inf. Fusion*, vol. 49, pp. 69–78, Sep. 2019.

[9] M. S. Hossain and G. Muhammad, "Healthcare big data voice pathology assessment framework," *IEEE Access*, vol. 4, pp. 7806–7815, 2016.

[10] A. Papageorgiou, M. Strigkos, E. Politou, E. Alepis, A. Solanas, and C. Patsakis, "Security and privacy analysis of mobile health applications: The alarming state of practice," *IEEE Access*, vol. 6, pp. 9390–9403, 2018.

[11] M. S. Hossain, G. Muhammad, W. Abdul, B. Song, and B. Gupta, "Cloud-assisted secure video transmission and sharing framework for smart cities," *Future Gener. Comput. Syst.*, vol. 83, pp. 596–606, Jun. 2018.

[12] Q. Huang, L. Wang, and Y. Yang, "Secure and privacy-preserving data sharing and collaboration in mobile healthcare social networks of smart cities," *Secur. Commun. Netw.*, vol. 2017, Aug. 2017, Art. no. 6426495.

[13] L. Wu, B. Chen, S. Zeadally, and D. He, "An efficient and secure searchable public key encryption scheme with privacy protection for cloud storage," *Soft Comput.*, vol. 22, no. 23, pp. 7685–7696, 2018.

[14] M. Alhussein and G. Muhammad, "Voice pathology detection using deep learning on mobile healthcare framework," *IEEE Access*, vol. 6, pp. 41034–41041, 2018.

[15] G. Muhammad, M. F. Alhamid, M. Alsulaiman, and B. Gupta, "Edge computing with cloud for voice disorder assessment and treatment," *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 60–65, Apr. 2018.

[16] T. Panch, P. Szolovits, and R. Atun, "Artificial intelligence, machine learning and health systems," *J. Global Health*, vol. 8, no. 2, 2018, Art. no. 020303.

[17] G. Fiumara, A. Celesti, A. Galletta, L. Carnevale, and M. Villari, "Applying artificial intelligence in healthcare social networks to identity critical issues in patients' posts," in *Proc. HEALTHINF*, 2018, pp. 680–687.

[18] L. Carnevale, R. S. Calabrò, A. Celesti, A. Leo, M. Fazio, P. Bramanti, and M. Villari, "Toward improving robotic-assisted gait training: Can big data analysis help us?" *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1419–1426, Apr. 2018.

[19] M. Masud, M. S. Hossain, and A. Alamri, "Data interoperability and multimedia content management in e-Health systems," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 6, pp. 1015–1023, Nov. 2012.

[20] J. Wan, J. Li, Q. Hua, A. Celesti, and Z. Wang, "Intelligent equipment design assisted by cognitive Internet of Things and industrial big data," *Neural Comput. Appl.*, pp. 1–10, Sep. 2018. doi: 10.1007/s00521-018-3725-5.

[21] R. H. G. Martins, H. A. do Amaral, E. L. M. Tavares, M. G. Martins, T. M. Gonçalves, and N. H. Dias, "Voice disorders: Etiology and diagnosis," *J. Voice*, vol. 30, no. 6, pp. 761.e1–761.e9, Nov. 2016.

[22] A. R. Maccarini and E. Lucchini, "La valutazione soggettiva ed oggettiva della disfonia. Il protocollo sifel," *Acta Phoniatrica Latina*, vol. 24, nos. 1–2, pp. 13–42, 2002.

[23] M. S. Hossain, G. Muhammad, and A. Alamri, "Smart healthcare monitoring: A voice pathology detection paradigm for smart cities," *Multimedia Syst.*, pp. 1–11, Jul. 2017. doi: 10.1007/s00530-017-0561-x.

[24] Z. Ali, G. Muhammad, and M. F. Alhamid, "An automatic health monitoring system for patients suffering from voice complications in smart cities," *IEEE Access*, vol. 5, pp. 3900–3908, 2017.

[25] J. Wang and C. Jo, "Vocal folds disorder detection using pattern recognition methods," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2007, vol. 31, no. 1, pp. 3253–3256.

[26] R. T. Ritchings, M. McGillion, and C. J. Moore, "Pathological voice quality assessment using artificial neural networks," *Med. Eng. Phys.*, vol. 24, nos. 7–8, pp. 561–564, Sep. 2002.

[27] L. Verde, G. De Pietro, and G. Sannino, "Voice disorder identification by using machine learning techniques," *IEEE Access*, vol. 6, pp. 16246–16255, 2018.

[28] J. I. Godino-Llorente, P. Gómez-Vilda, N. Sáenz-Lechón, M. Blanco-Velasco, F. Cruz-Roldán, and M. A. Ferrer-Ballester, "Support vector machines applied to the detection of voice disorders," in *Proc. Int. Conf. Nonlinear Anal. Algorithms Speech Process.*, vol. 3817. Berlin, Germany: Springer, 2005, pp. 219–230.

[29] N. Souissi and A. Cherif, "Dimensionality reduction for voice disorders identification system based on Mel frequency cepstral coefficients and support vector machine," in *Proc. 7th Int. Conf. Modelling, Identificat. Control (ICMIC)*, Dec. 2015, pp. 1–6.

[30] F. Amara, M. Fezari, and H. Bourouba, "An improved GMM-SVM system based on distance metric for voice pathology detection," *Appl. Math. Inf. Sci.*, vol. 10, no. 3, pp. 1061–1070, 2016.

[31] H. Cordeiro, J. Fonseca, and C. Meneses, "Spectral envelope and periodic component in classification trees for pathological voice diagnostic," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2014, pp. 4607–4610.

[32] I. M. M. El Emary, M. Fezari, and F. Amara, "Towards developing a voice pathologies detection system," *J. Commun. Technol. Electron.*, vol. 59, no. 11, pp. 1280–1288, 2014.

[33] M. M. Baki, G. Wood, M. Alston, P. Ratcliffe, G. Sandhu, J. S. Rubin, and M. A. Birchall, "Reliability of OperaVOX against multidimensional voice program (MDVP)," *Clin. Otolaryngol.*, vol. 40, no. 1, pp. 22–28, 2015.

[34] E. Van Leer, R. C. Pfister, and X. Zhou, "An iOS-based cepstral peak prominence application: Feasibility for patient practice of resonant voice," *J. Voice*, vol. 31, no. 1, pp. 131.e9–131.e16, 2017.

[35] T. Chen, "Introduction to boosted trees," *Univ. Washington Comput. Sci.*, vol. 22, p. 115, Oct. 2014.

[36] *MATLAB and Statistics Toolbox Release 2017B*, The MathWorks, Natick, MA, USA, 2017.

[37] U. Cesari, G. De Pietro, E. Marciano, C. Niri, G. Sannino, and L. Verde, ''Voice disorder detection via an m-Health system: Design and results of a clinical study to evaluate Vox4Health,'' *BioMed Res. Int.*, vol. 2018, Aug. 2018, Art. no. 8193694.

[38] E. J. Hunter, K. Tanner, and M. E. Smith, ''Gender differences affecting vocal health of women in vocally demanding careers,'' *Logopedics Phoniatrics Vocol.*, vol. 36, no. 3, pp. 128–136, 2011.

[39] E. T. Stathopoulos, J. E. Huber, and J. E. Sussman, ''Changes in acoustic characteristics of the voice across the life span: Measures from individuals 4–93 years of age,'' *J. Speech, Lang., Hearing Res.*, vol. 54, no. 4, pp. 1011–1021, 2011.

[40] M. Brockmann, M. J. Drinnan, C. Storck, and P. N. Carding, ''Reliable jitter and shimmer measurements in voice clinics: The relevance of vowel, gender, vocal intensity, and fundamental frequency effects in a typical clinical task,'' *J. Voice*, vol. 25, no. 1, pp. 44–53, 2011.

[41] L. Verde, G. De Pietro, P. Veltri, and G. Sannino, ''A noise-aware methodology for a mobile voice screening application,'' in *Proc. 6th Int. Workshop Adv. Sensors Interfaces (IWASI)*, Jun. 2015, pp. 193–198.

[42] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*. Burlington, MA, USA: Morgan Kaufmann, 2016.

[43] T. G. Dietterich, ''Ensemble methods in machine learning,'' in *Proc. Int. Workshop Multiple Classifier Syst.* Berlin, Germany: Springer, 2000, pp. 1–15.

[44] M. Farrús, J. Hernando, and P. Ejarque, ''Jitter and shimmer measurements for speaker recognition,'' in *Proc. 8th Annu. Conf. Int. Speech Commun. Assoc. (InterSpeech)*, 2007, pp. 778–781.

[45] F. Severin, B. Bozkurt, and T. Dutoit, ''HNR extraction in voiced speech, oriented towards voice quality analysis,'' in *Proc. 13th Eur. Signal Process. Conf.*, Sep. 2005, pp. 1–4.

[46] L. Verde, G. De Pietro, and G. Sannino, ''A methodology for voice classification based on the personalized fundamental frequency estimation,'' *Biomed. Signal Process. Control*, vol. 122, no. 5, pp. 2960–2961, 2018.

[47] A. Y. Ng, ''Preventing 'overfitting' of cross-validation data,'' in *Proc. ICML*, vol. 97, 1997, pp. 245–253.

[48] M. Eye and E. Infirmary, ''Voice disorders database, version 1.03 (CD-ROM),'' Kay Elemetrics Corp., Lincoln Park, NJ, USA, Tech. Rep., 1994.

[49] M. Pützer and J. Koreman, ''A german database of patterns of pathological vocal fold vibration,'' Phonus, Tech. Rep., 1997, pp. 143–153, vol. 3.

[50] U. Cesari, G. De Pietro, E. Marciano, C. Niri, G. Sannino, and L. Verde, ''A new database of healthy and pathological voices,'' *Comput. Elect. Eng.*, vol. 68, pp. 310–321, May 2018.

[51] PhysioNet. (2018). *Voiced (Voice ICar fEDerico II) Database*. [Online]. Available: https://physionet.org/physiobank/database/voiced/

[52] B. Schölkopf, C. J. C. Burges, and A. J. Smola, *Advances in Kernel Methods: Support Vector Learning*. Cambridge, MA, USA: MIT Press, 1999.

[53] G. H. John and P. Langley, ''Estimating continuous distributions in Bayesian classifiers,'' in *Proc. 11th Conf. Uncertainty Artif. Intell.* Burlington, MA, USA: Morgan Kaufmann Publishers, 1995, pp. 338–345.

[54] D. W. Aha, D. Kibler, and M. K. Albert, ''Instance-based learning algorithms,'' *Mach. Learn.*, vol. 6, no. 1, pp. 37–66, 1991.

[55] L. Verde, G. De Pietro, and G. Sannino, ''Vox4Health: Preliminary results of a pilot study for the evaluation of a mobile voice screening application,'' in *Proc. Int. Symp. Ambient Intell.* Cham, Switzerland: Springer, 2016, pp. 131–140.

LAURA VERDE received the master's degree in biomedical engineering from the University of Naples Federico II, Italy, in 2011, and the Ph.D. degree in information engineering, in 2019. She is currently a Research Fellow with the Institute of High-Performance Computing and Networking (ICAR), National Research Council of Italy (CNR). Her research interests include biomedical signal processing, voice signal analysis, m-health systems, and artificial intelligence for healthcare.

GIUSEPPE DE PIETRO is currently the Director of the Institute for High Performance Computing and Networking, CNR, and an Adjunct Professor with the College of Science and Technology, Temple University, Philadelphia. He has been actively involved in many European and national projects, with industrial co-operations. He is the author of over 200 scientific articles published in international journals and conferences. His current research interests include cognitive computing, clinical decision support systems, and software architectures for e-health. He is also a KES International Member. He is also involved in many program committees and journal editorial boards.

MUBARAK ALRASHOUD received the Ph.D. degree in electrical and computer engineering from Ryerson University, Ottawa, Canada. He is currently an Assistant Professor and the Head of the Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia.

AHMED GHONEIM (M'10) received the M.Sc. degree in software modeling from the University of Menoufia, Egypt, in 1999, and the Ph.D. degree in software engineering from the University of Magdeburg, Germany, in 2007. He is currently an Associate Professor with the Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia. His research interests include software evolution, service-oriented engineering, software development methodologies, net-centric computing, and human–computer interaction.

KHALED N. AL-MUTIB is currently an Associate Professor with the Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia, where he is currently an Associate Professor. His research interests include robotics, computation intelligence, and healthcare.

GIOVANNA SANNINO received the bachelor's degree in computer engineering and the master's degree (European Master on Critical Networked Systems) from the University of Naples Federico II, in 2008 and 2009, respectively, and the master's degree *(cum laude)* in telecommunications engineering and the Ph.D. degree in information engineering from the Parthenope University of Naples, in April 2011 and April 2015, respectively. She is currently a Researcher with the Institute for the High-Performance Computing and Networking (ICAR-CNR) and an Adjunct Professor of informatics with the University of Naples Federico II. Her research interests include the areas of mobile health, pervasive computing, pattern recognition, signal processing, and artificial intelligence for healthcare. She is also an IEEE 11073 Personal Health Device Working Group Member.