Received October 5, 2016, accepted October 24, 2016, date of publication November 8, 2016,
date of current version November 28, 2016.

Digital Object Identifier 10.1109/ACCESS.2016.2626316

Healthcare Big Data Voice Pathology Assessment Framework

M. SHAMIM HOSSAIN[1], (Senior Member, IEEE), and GHULAM MUHAMMAD[2]
[1]Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia
[2]Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia

Corresponding author: M. S. Hossain (mshossain@ksu.edu.sa)

This work was supported by the Deanship of Scientific Research at King Saud University, Riyadh, Saudi Arabia, through the Research Group Project under Grant RGP-228.

ABSTRACT The fast-growing healthcare big data plays an important role in healthcare service providing. Healthcare big data comprise data from different structured, semi-structured, and unstructured sources. These data sources vary in terms of heterogeneity, volume, variety, velocity, and value that traditional frameworks, algorithms, tools, and techniques are not fully capable of handling. Therefore, a framework is required that facilitates collection, extraction, storage, classification, processing, and modeling of this vast heterogeneous volume of data. This paper proposes a healthcare big data framework using voice pathology assessment (VPA) as a case study. In the proposed VPA system, two robust features, MPEG-7 low-level audio and the interlaced derivative pattern, are used for processing the voice or speech signals. The machine learning algorithms in the form of a support vector machine, an extreme learning machine, and a Gaussian mixture model are used as the classifier. In the experiments, the proposed VPA system shows its efficiency in terms of accuracy and time requirement.

INDEX TERMS Healthcare big data, voice pathology, classification, feature extraction.

I. INTRODUCTION

An increasing number of sensors, smart phones, wearable or portable/implantable devices are being used to provide healthcare services at home and in hospitals. These services produce a huge amount of disparate healthcare data, including structured (e.g., electronic health records; EHRs), semi-structured (physician-to-patient, and patient-to-patient communication through email, social media, and web) and unstructured (clinical notes, claims and informal texts) data. This so-called 'big data' comes in different forms such as text, images, audio, and video. Being so broad and complex, traditional healthcare data analysis or processing techniques, algorithms or frameworks, may be unable to handle big data [1] properly to provide useful insights into the healthcare industry.

The development of an increasing number of healthcare technologies and services, healthcare big data [2]–[13], [15] has attracted attention from academia and industry to understand patients, to predict diseases, to make clinical decisions and to conduct research into better patient care. Chen et al. [1] discussed some open issues in big data, including the standardization and format of big data, big data transfer and processing, searching and mining of big data, and management of services. Andreu-Perez *et al.* [2] presents an overview of big data in healthcare [3]. A review of healthcare big data can be found in [4]; mobile cloud computing and big data analytics [5] have been envisioned. Trustworthiness [6] is a concern while processing healthcare big data over the cloud. Chen *et al.* [7] describes a framework in which the cloud and big data analytics are intended for use in health monitoring. Some healthcare data (e.g., voice data and other multimedia) need to be processed and classified as a stream in real-time, in this case by a powerful algorithm [14] and effective analytics are required. Hu *et al.* [8] proposes a diagnostic model for outpatient doctors using healthcare data analytics; support vector machines (SVMs) and neural networks are used for classifications. There are few studies of voice pathology assessment in a big data platform. For example, a health monitoring system has been proposed using the Internet of Things [9]; however, the proposed system is not entirely based on voice pathology assessment. A throat polyp detection system based on compressed big data has also been developed [10]; however, the influence of heterogeneous healthcare big data remains unclear.

To handle a large amount of heterogeneous healthcare data, utmost care and a sophisticated algorithm are required, to process the data properly without loss of information. There can be thousands of hospitals and medical centers connected to a large medical network; however, the data structures used by the institutions may vary, and data can easily be lost. In this paper, we study the case of voice pathology assessment. Whereas typical patient information can be stored in a text format, voice data is captured using microphones, video cameras or high-speed cameras attached to smart devices or personal computers. Voice or speech data (known as the 'speech signal') of the patients can be individual phonemes (e.g., /a/, /i/ or /o/), a combination of phonemes, isolated words or continuous speech. Managing all these variables to develop a reliable system needs special attention for feature extraction and classification. We, therefore, propose a voice pathology assessment (VPA) system using big data, to be deployed in a health monitoring system. MPEG-7 low-level audio features [16] and the interlaced derivative pattern (IDP) [17] are non-linearly combined to produce a set of robust features in the proposed system. At the classification stage, a combination of stochastic, shallow, and deep learning methods is used. In particular, we utilize the Gaussian mixture model (GMM), SVM, and extreme learning machine (ELM) for classification. The proposed system makes two key contributions. First, non-linear combination of two complementary sets of features, and second, making a decision based on the majority voting on three state-of-the-art classification techniques. The system can easily manage our data, because the features are not affected by recording equipment or environment. These features can also extract relevant information even from unstructured data. A combination of the three machine learning algorithms is used to remove some uncertainty at the classification stage.

The remainder of this paper is organized as follows. Section II describes the proposed big healthcare data framework that uses the voice pathology assessment (VPA) as a case study. Section III presents the experiments, results, and discussion Finally, Section IV concludes the paper.

II. PROPOSED HEALTHCARE BIG DATA FRAMEWORK

A. HEALTHCARE BIG DATA SOURCE ECO SYSTEM

Healthcare big data is a revolutionary tool in the healthcare industry, and is becoming vital in current patient-centric care. Owing to the massive growth of data in the healthcare industry, diverse data sources have been aggregated into the healthcare big data ecosystem. These data sources are is used by a healthcare provider to enable him or her to make decisions and provide appropriate care. Major data sources, along with the challenges involved, are discussed below:

1) *Physiological data.* These data are huge in terms of volume and velocity. Regarding data volume, a variety of signals is collected from heterogeneous sources to monitor patient characteristics, including blood pressure, blood glucose, and heart rate. Sources include electroencephalogram, electrocardiogram, and electroglottogram. Data velocity can be observed from the growing rate of data generation from continuous monitoring, especially for patients in a critical condition, requires these signals to be processed in real-time, for decision making. These signals need to be extracted efficiently and processed with the suitable machine learning algorithm to provide meaningful data for effective patient care. Efficient and comprehensive methods are also required to analyze and process the collected signals to provide useable data to the healthcare professionals and other related stakeholders. The combination of EHR and physiological signals may increase the precision of data based on the surrounding context of the patient.
2) *EHRs/EMRs*. EHRs or electronic medical records (EMRs) are digitized structured healthcare data from a patient. The EHRs are collected from and shared among hospitals, research centers, government agencies, and insurance companies. Security, integrity and privacy violations of these data can cause irremediable damage to the health, or even death, of the individual and loss to society. Thus, big healthcare data security is now a key topic of research.
3) *Medical images.* These images generate a huge volume of data to assist healthcare professionals for identifying or detecting disease, treatment, predicting and monitoring of patients. Medical imaging techniques such as X-ray, ultrasound, or computed tomography scan play a crucial role in diagnosis and prognosis. Owing to the complication, dimensionality and noise of the collected images, efficient image processing methods are required to provide clinically suitable data for patient care.
4) *Sensed data* from patients are collected using different wearable or implantable devices, environment-mounted devices, ambulatory devices, and sensors and smart phones from home or in hospitals [11]. The sensed data forms a key part of healthcare big data, as these sensors are used to capture critical events or provide continuous monitoring. However, sensed data must be collected, pre-processed, stored, shared and delivered correctly in a reasonable time to be of use to healthcare providers when making clinical decisions. Owing to the enormous volume of data collected, automated algorithms are required to reduce noise and to allow for the deployment with big data analytics so that computation time can be reduced. Moreover, it is a challenge to collect and collate multimodal sensed data from multiple sources at the same time.
5) *Clinical notes*. The clinical notes, claims, recommendations, and decisions constitute one of the largest unstructured sources of healthcare big data. Owing to the variety in format, reliability, completeness, and accuracy of the clinical notes, it is challenging to ensure the health care provider has the correct information. Efficient data mining and natural language processing techniques are required to provide meaningful data.

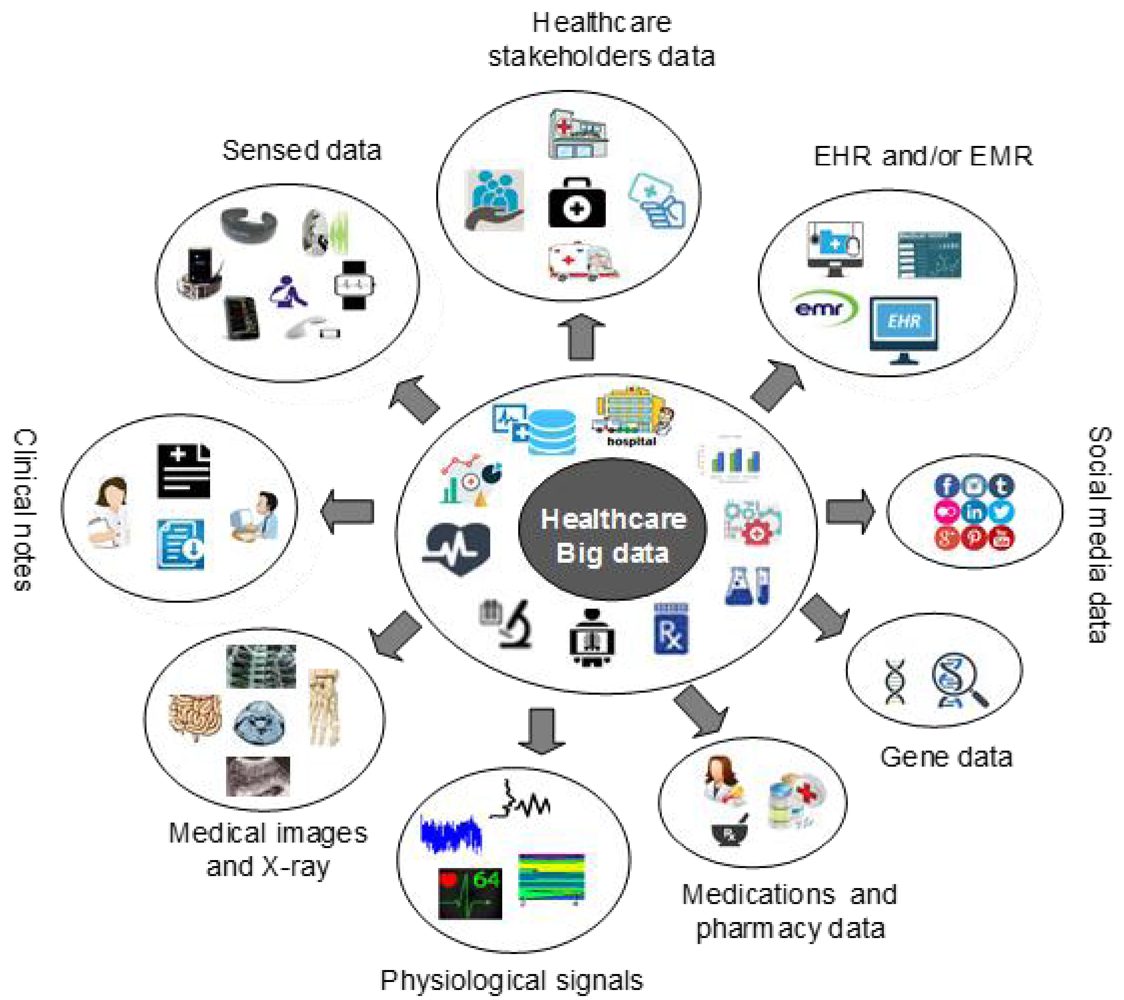

FIGURE 1. **Big healthcare data source ecosystem.**

6) *Gene data*. The genome data makes a major contribution to healthcare big data. The human genome has a huge number of genes; collecting, analyzing, and classifying data on these genes has taken years. These gene data have now been integrated from the genetic level to physiological level of a human being.
7) *Social media data.* Social media connect healthcare professionals and patients outside their clinics, hospitals, and homes through machine-to-machine, physician-to-patient, physician-to-physician, and patient-to-patient communications.

Patients with similar symptoms and diseases can share their experiences through social media to get ad-hoc counselling, which constitutes a big data problem. Based on study [18], 80% of unstructured data comes social media. Messages, trending medical images, location information, and other features of social media contribute to healthcare big data. For example, social media has recently been used for investigating depression and suicide rates using real-time emotional state analysis from Twitter [19]. Because of the heterogeneous nature of social healthcare media data, it is difficult to conduct data analysis and provide meaningful data to healthcare big data stakeholders. Thus, this data needs to be appropriately mined [20], analyzed and processed to improve the quality the healthcare services in healthcare providers.

B. HIGH-LEVEL SYSTEM ARCHITECTURE

Figure 2 shows the high-level system architecture and data flow of healthcare big data. A typical healthcare big data solution utilizes only selected parts of this architecture, based on their desired functionality. Big healthcare data is contributed to by various data sources. From these data sources, relevant information is sent to the healthcare cloud datacenters for analysis by a big data application. The data is delivered through intermediate communication and processing, where it has been pre-processed for noise reduction, unreliability, inconsistency, and analog-to-digital conversion. Sometimes this pre-processing is done based on the opinion of the healthcare professionals.

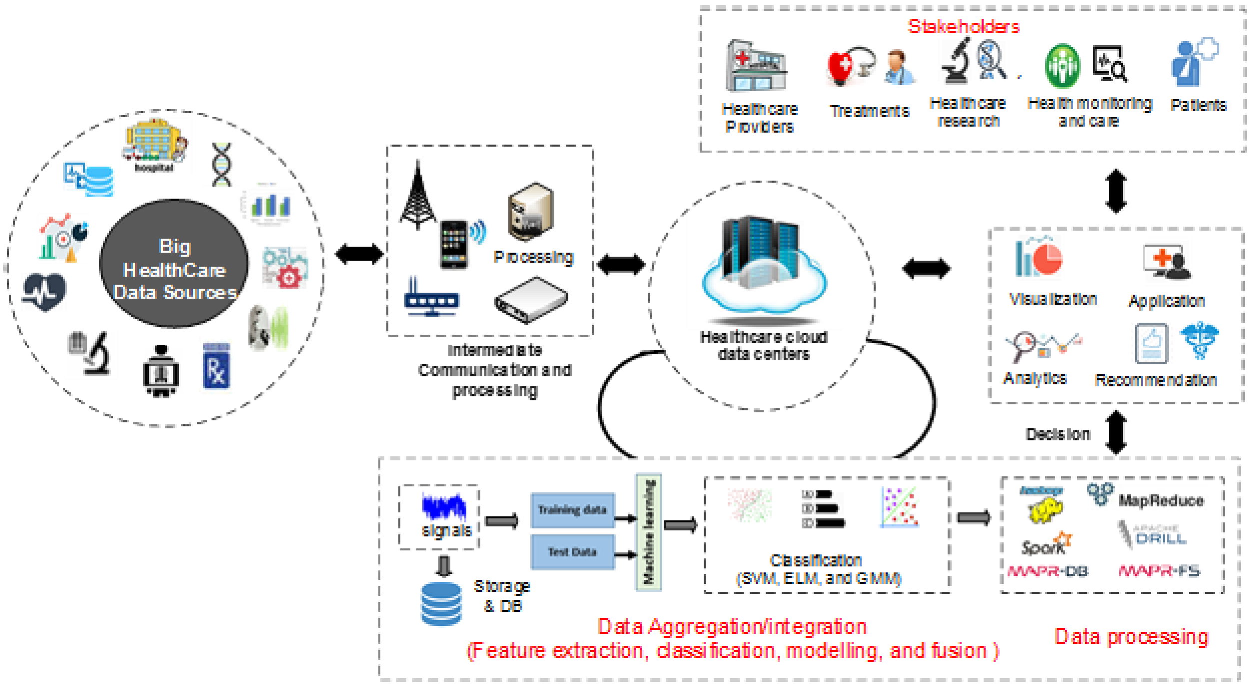

FIGURE 2. High-level system architecture and data flow of big healthcare data.

With the increasing volume of collected data extracted from heterogeneous sources, it is necessary that this large dataset should be stored in a database or a distributed file system for aggregation or processing by different healthcare stakeholders. In the aggregation phase, at first, the feature is extracted from the original data (e.g., signal) and then further processed by classification, normalization, fusion, or modelling. The features can be extracted based on the advice of healthcare professionals or domain experts. Then processed data (as training) and test data are fed through the machine learning algorithm. Training is sometimes based on previous historical data. Classification algorithms have the potential to detect abnormalities, such as with voice pathology detection, or to assist with diagnosis as a reference for physicians.

A popular platform for parallel programming is Google's MapReduce, which is characterized by its sophisticated techniques of load balancing and fault tolerance. Apache Hadoop's big data platform can implement the MapReduce algorithm for healthcare big data analysis [21]. MapReduce is used in processing healthcare data where response time is not so critical. For large healthcare data, the MapR (MapReduce version 2) can be used for implementing machine learning algorithms, to provide useful and accurate data for improved patient care. However, for the fast processing as well as for continuous streaming of data, Apache Spark [22] may be used. It has a set of Application program interfaces (APIs) for machine learning. After data processing, the meaningful healthcare data recommendations are sent to the relevant stakeholder.

C. PROPOSED VPA SYSTEM

Figure 3 shows a block diagram of the proposed VPA system. A cloud server receives the medical data from many different hospitals and clinics. In this paper, we mainly focus on speech signals from patients with or without vocal fold pathology. The speech signals are processed on the cloud server. For feature extraction, we adopt two methods, MPEG-7 low-level audio and IDP, which are complementary in nature. These two methods are then normalized and fused to produce a robust set of features. These features are fed into a classification unit in the cloud. In the classification unit, three machine learning algorithms, SVM, GMM, and ELM, are used. A ranking system based on the scores from these algorithms is introduced to provide a classification. The major components of the system are described below.

1) HOSPITAL DATA

The hospital data for the VPA may vary in many ways. As we are concerned with only speech signals, we focus on the diversity of these signals captured in many hospitals. The speech signals may contain only sustained vowels /a/, /i/, or /o/, or a combination of them. Alternatively, they may contain more complex speech such as whole words, phrases, or sentences from a conversation, spontaneous speech, or reading aloud. Therefore, speech signals differ in both content and length. Speech signals may be recorded by different media and can be captured in different environments, such as in a sound-treated room or in a normal office room. The sampling frequencies may also differ. For each patient, there is information about gender, age, smoking or non-smoking, weight, diagnosis,

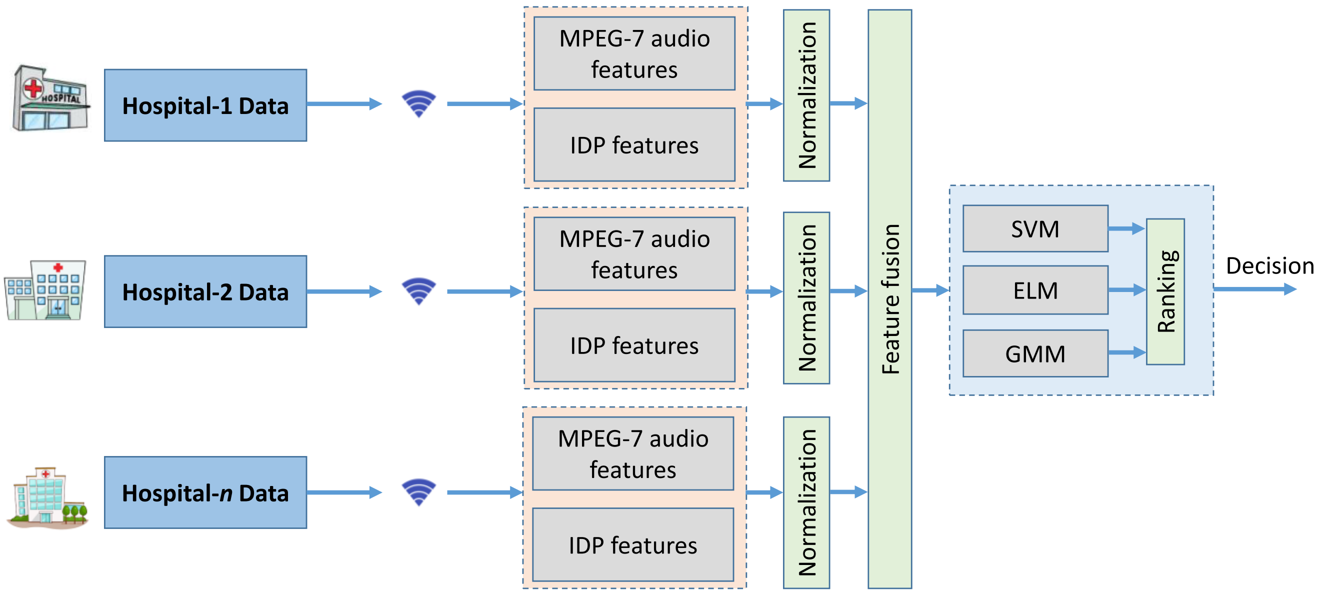

FIGURE 3. Block diagram of the proposed big data processing for voice pathology assessment.

severity, and so on; however, some fields in this information may be missing. This results in a very large volume of data that are then transferred to the cloud for processing. At the processing stage, we process only voice or speech signals; EHR data are kept in a separate index. Due to the unstructured nature of the data, we also have to filter out the poor quality or incorrectly labeled samples; this should be done at least for the training data.

2) FEATURE EXTRACTION

As the nature of the data is heterogeneous, we need to be careful for feature extraction. There are many feature extraction techniques involved in the VPA. These include Mel-frequency cepstral coefficients (MFCC) [23], multi-dimensional voice program (MDVP) [24], MPEG-7 low-level audio descriptors [25], IDP [26], and glottal noise parameters [27]. Each of them has its own advantages and disadvantages; however, after a careful consideration, we find that both the MPEG-7 audio features and the IDP features provided good results in the literature, and they are not too much affected by the diversity of the recorded signals. Therefore, we adopt these two features in the proposed assessment. It can be noted that these two feature extraction techniques provide different types of information to the classifier, and hence, can be complementary to each other.

a: MPEG-7 LOW-LEVEL AUDIO FEATURES

The low level audio features of MPEG-7 were originally designed to detect events in an audio signal [16]. Since then they are used in many audio and speech processing applications including environment detection [28], voice pathology detection, and speaker identification. Some of the features are scaler, and the rest are vector. A list of scaler and vector (with size) MPEG-7 low-level audio features [16] is given in Table 1. Audio waveform, audio power, and temporal centroid are temporal features, while the rest are spectral features.

As the voice pathology occurs due a malfunction in the vocal folds, more information about the pathology can be found in a voice signal or in the voiced part of a speech signal. Therefore, we apply a voice activity detection module to identify the voiced part in the speech signal. Features are extracted only in the voiced parts. In this way, the features will be more meaningful for distinguishing normal voices from pathological voices. For example, a normal voice will have more harmonic components than a pathological voice will have. Therefore, audio harmonicity (AH) feature will return a high value for normal voices compared with pathological voices. Similarly, audio spectrum spread (ASS) will be more in pathological voices than in normal voices.

The MPEG-7 features are calculated frame by frame. There are several parameters that need to be set while extracting the features. Table 2 gives these parameters values in the proposed system. The frame size is set to 40 milliseconds (ms) and the overlapping between the successive frames is 50%. The filters are spaced in a logarithmic scale, where the lowest edge is 250 Hz, and the highest edge is 12.5 kHz. The frequency resolution is 4 octaves per band. The mode 'instantaneous' means that the features are calculated at frame level.

Figure 4 shows the discrimination capability, in terms of scatter plot, of two MPEG-7 features at a time. For example, (a) shows a scatter plot between ASC and ASS for normal and pathological samples. Most of the normal samples can be separated from most of the pathological samples; however, some samples overlap or cannot be separated.

TABLE 1. MPEG-7 audio features and their size.

Scaler Features	Vector Features	
Name	Name	Size
Audio Waveform (AW) (min. and max.)	Audio Spectrum Flatness (ASF)	22 features
Audio Power (AP)	Audio Spectrum Basis (ASB)	2 features
Audio Spectrum Centroid (ASC)	Audio Spectrum Projection	2 features
Audio spectrum Spread (ASS)	Audio Spectrum Envelope	3 features
Audio Harmonicity (AH)		
Audio Fundamental Frequency (AFF)		
Temporal Centroid (TC)		
Spectral Centroid (SC)		
Harmonic Spectral Centroid (HSC)		
Harmonic Spectral Deviation (HSD)		
Harmonic Spectral Spread (HSS)		
Harmonic Spectral Variation (HSV)		

TABLE 2. Parameter values set for MPEG-7 audio feature extraction.

Parameters	Value	Parameters	Value
Hop size	20 ms	Resolution	4 Octave / band
Frame size	40 ms	Low limit	40 Hz
Low edge	250 Hz	High limit	500 Hz
High edge	12.5 kHz	Mode	Instantaneous

b: IDP FEATURES

The IDP features were first introduced in image processing applications [29]. They have been successfully applied to speech recognition [17] and voice pathology detection [26]. The IDP features have several interesting characteristics, which include a compact representation of second-order derivatives on four directions (time, frequency, time-frequency positive direction, and time-frequency negative direction), and a good encoding of time-frequency variations. The steps of calculating IDP features are as follows:

Step 1: Divide the signal into frames (we use frame size of 40 ms) and multiply the frames by a Hamming window.

Step 2: Apply a Fourier transform to convert the signal into the frequency domain.

Step 3: Apply a 24 band-pass filters, whose center frequencies are spaced on a Bark scale. Take logarithms on the outputs of the filters.

Step 4: Measure the IDP for each frame [26].

Figure 5 shows the time-domain signal and the IDP representation of a normal sample and a pathological sample. From the figure, we see that, for the normal sample, the higher filters (representing higher frequency) have low energy. On the other hand, for the pathological sample, the energy is distributed throughout the filters, even high energy can be found in higher filters. This demonstrate the discrimination capability of the IDP features for voice pathology assessment.

c: FEATURE FUSION

As the MPEG-7 features and the IDP features are complementary in nature, we fuse them non-linearly to obtain a maximum performance. First, we normalize the features, because they are obtained from signals captured at different types of environments and by various types of microphones. Second, we introduce a parameter, alpha, as a weight to the MPEG-7 features and the IDP features. If the value of alpha is set to zero, there will be no contribution of the MPEG-7 features, while if it is set to one, there will be no contribution of the IDP features. Therefore, the value of the alpha ranges between zero and one. We increment the value of alpha by 0.1.

3) CLASSIFICATION

We use three machine learning algorithms, each of which is of different characteristics, in the proposed system. The classifiers that we use are the SVM, the ELM, and the GMM. A simple description of these algorithms are provided below.

a: SVM

The SVM has been intensely used in the literature for many two-class problems because of its flexibility for determining the threshold for separating the classes, good generalization of out-of-samples, producing a unique solution contrary to other forms of the neural networks. In the SVM, a kernel function is used to map the features from a lower to a higher-dimensional space to find an optimal hyperplane to separate the samples of the classes. There are three basic kernel functions in the SVM, which are linear, polynomial, and radial basis function (RBF). The polynomial kernel of order n can be written as follows.

$$K(x_a, x_b) = \left(\gamma x_a^T x_b + 1\right)^n, \quad \gamma > 0$$

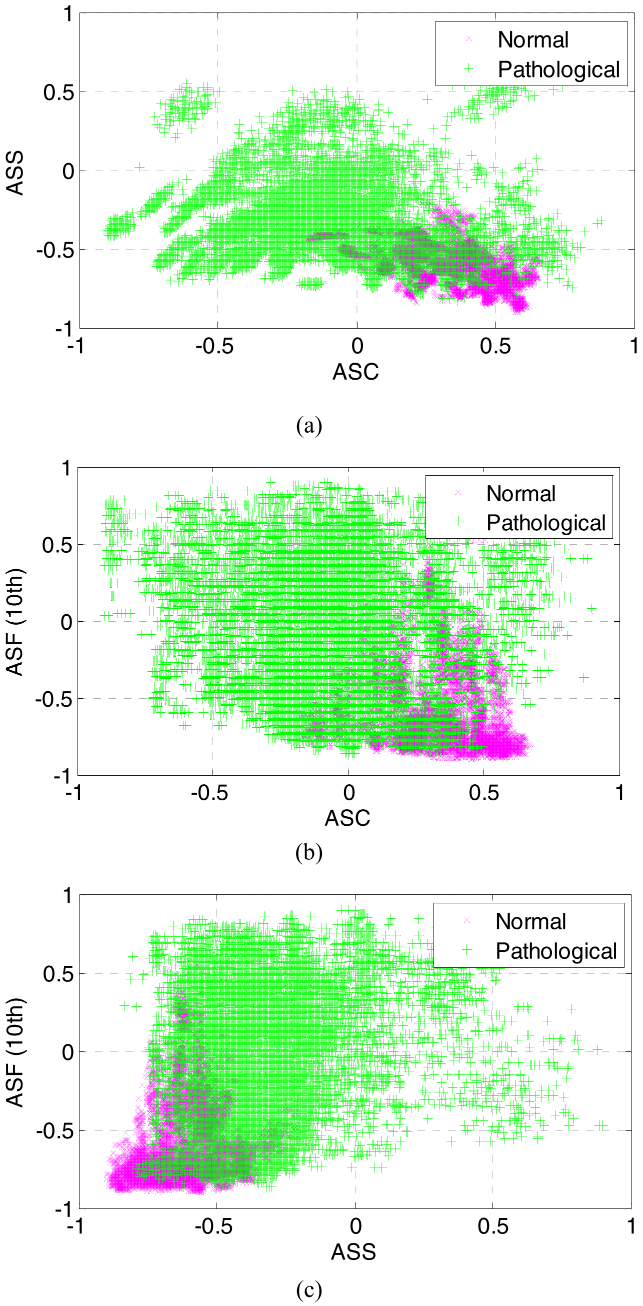

FIGURE 4. Scatter plot of (a) ASC vs ASS, (b) ASC vs ASF (10^{th} filter), and (c) ASS vs ASF (10^{th} filter), showing the distribution of normal and pathological samples.

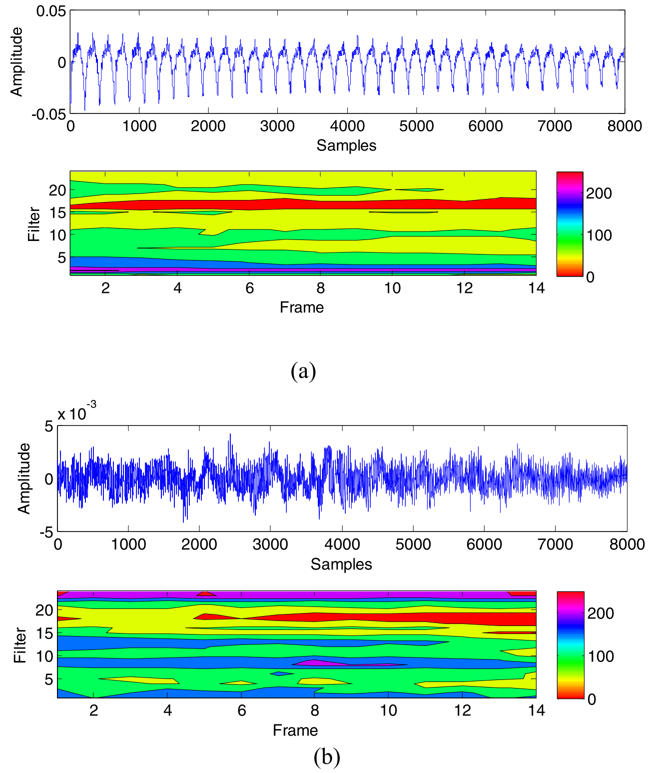

FIGURE 5. The IDP representation of (a) a normal sample, and (b) a pathological sample. In each sub-figure, the upper one shows the time-domain signal, and the lower one shows the IDP representation.

Where, x_a and x_b are the samples of the two classes a and b, and γ is the parameter to be tuned using training samples. In our system, we use the order to be 3.

The RBF kernel can be expressed as follows.

$$K(x_a, x_b) = \exp\left(-\gamma ||x_a - x_b||^2\right), \quad \gamma > 0$$

b: ELM

The ELM is a powerful machine learning algorithm [30], and it has several important properties including (i) it is fast, (ii) the solution is dense, and (iii) the feature mapping can be in an infinite or a finite space. If there are M number of hidden neurons in the ELM, bias b_j of the j-th hidden neuron, input weight is $\mathbf{a}_j$ and the output weight is θ_j, then the output function $o(\mathbf{x})$ of $\mathbf{x}$ training samples can be calculated as follows.

$$o(\mathbf{x}) = \sum_{j=1}^{M} \Omega\left(\mathbf{a}_j, b_j, \mathbf{x}\right) \cdot \theta_j, \quad \mathbf{a}_j \in \Re^d, b_j \in \Re$$

$\Omega(.)$ is the mapping function of the ELM. There are three commonly used mapping functions, which are expressed as follows.

Sigmoid function: $\Omega\left(\mathbf{a}, b, \mathbf{x}\right) = \frac{1}{1+e^{-\mathbf{a}.\mathbf{x}+b}}$

Gaussian function: $\Omega\left(\mathbf{a}, b, \mathbf{x}\right) = e^{-b||\mathbf{x}-\mathbf{a}||}$

Fourier basis: $\Omega\left(\mathbf{a}, b, \mathbf{x}\right) = \cos\left(\mathbf{a} \cdot \mathbf{x} + b\right)$

c: GMM

The GMM is a stochastic modeling technique, which is frequently used in speech processing applications such as speech recognition, speaker recognition, and environment detection. The GMM-based approach is fast during testing, because the models are represented by a few number of parameters. In this approach, each class is modeled by a GMM having several Gaussian mixtures. The number of mixtures varies depending on the hidden variables. In our proposed method, we tried with different number of Gaussian mixtures.

d: RANKING

The classifiers are fused at the decision level by a majority voting criteria. If two or more classifiers' decisions are pathological, then the final decision is pathological. This will eliminate any unfortunate misdiagnosis of a single classifier.

III. RESULTS AND DISCUSSION

A number of experiments were carried out to check the performance of the proposed VPA system. For the experiments, we collected data from the Massachusetts Eye and Ear Infirmary (MEEI) database [31], and the Saarbruecken Voice Database (SVD) [32]. There are data from more than 600 patients and 53 non-patients in the MEEI database, and approximately 1800 of each in the SVD. We adopt a 5-fold cross validation approach, where the total data are divided into five groups. We have five iterations, and in each iteration, four groups are trained and the remainder tested. At the end of the fifth iteration, all the groups are tested. We resampled all the signals to 16 kHz. All the data in the MEEI database and the SVD are already filtered and labeled.

First, we tried with the three kernels in the SVM to find the best one in the system. In this case, we did not fuse the two types of the features (MPEG-7 and IDP), and we used only the SVM as the classifier. Figure 6 shows the accuracy. From the figure, we see that the RBF kernel performed the best for both the features. The highest accuracy (71.3%) was obtained with the IDP features and RBF kernel of the SVM.

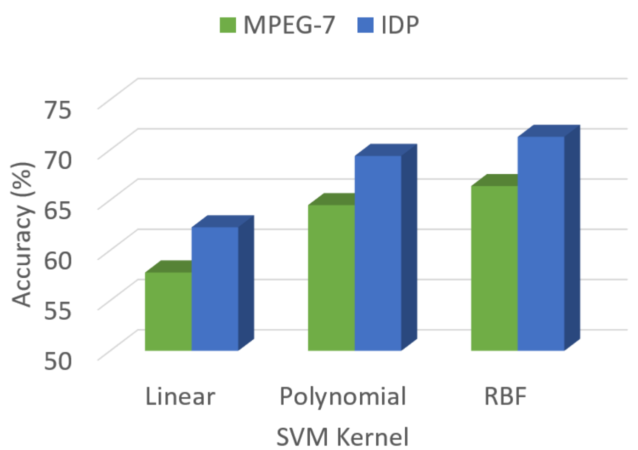

FIGURE 6. Accuracy of the system using different kernels of the SVM.

Next experiment was to check the performance of three different mapping functions of the ELM in the system. Only the ELM was used as the classifier. Figure 7 shows the accuracy of the systems using different mapping functions of the ELM. The Gaussian mapping function achieved the best accuracy, and with the IDP features and the ELM with the Gaussian mapping function the accuracy was 80.4%.

The last experiment to find the parameters in the classifier was with the GMM. We tried with different number of Gaussian mixtures, and the results are shown in Figure 8.

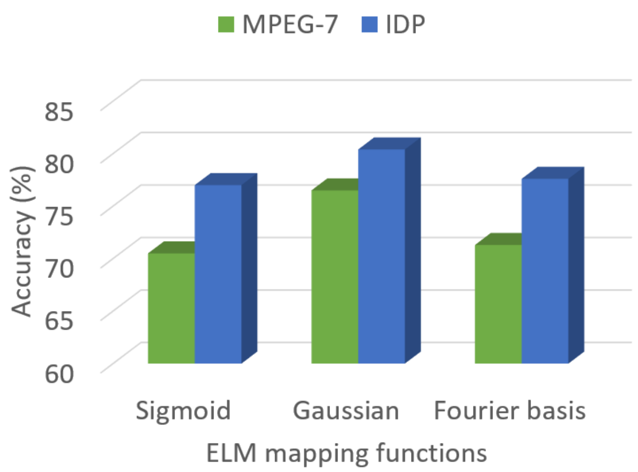

FIGURE 7. Accuracy of the system using different kernels of the ELM.

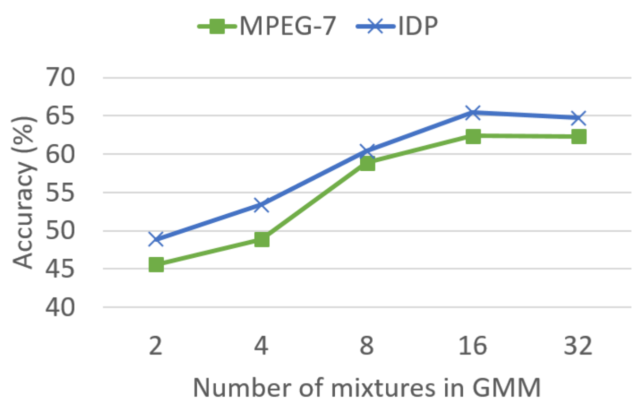

FIGURE 8. Accuracy of the system using different numbers of Gaussian mixtures of the GMM.

The best performance was obtained with 16 mixtures, and the accuracy in this case was 65.4% using the IDP features. In the subsequent experiments, we fixed the RBF as the kernel of the SVM, Gaussian mapping function as the mapping function of the ELM, and 16 mixtures for the GMM.

When fusing the two types of the features, we introduce a parameter, alpha, as described before. Figure 9 shows the effect of alpha on the accuracy of the system. For each of the classifiers, the optimal value of the alpha lies between 0.3 and 0.4. It means that the IDP features are contributing more than the MPEG-7 features towards a better accuracy. From the figure, we also find that the ELM performs better than the SVM and the GMM. With alpha = 0.4, the ELM achieved 80.4% accuracy, the SVM had 73.2%, while the GMM obtained 70.4% accuracy.

The overall accuracy of the proposed VPA system was 95.6%. In this case, the fusion of the features and the ranking of the classifiers were taken into account (see Figure 3). This accuracy is reasonable considering the heterogeneous nature of the big data available in different hospitals. The processing time during testing of each sample was 1.1 seconds in the

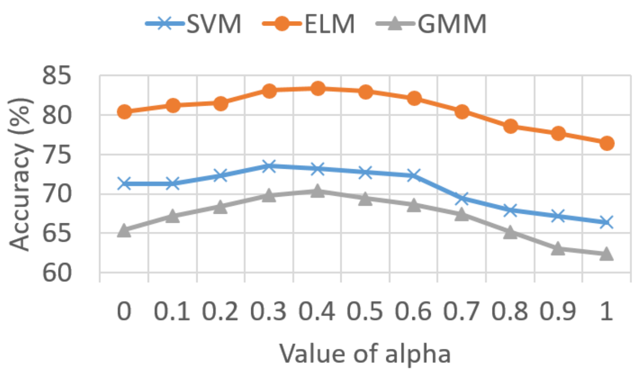

FIGURE 9. Changing of accuracies for different values of alpha, related to the fusion of the two types of the features.

server. This time is very much in line with the real-time applications.

IV. CONCLUSION

This paper presents a healthcare big data VPA framework, where data are generated from various data sources ranges from heterogeneous healthcare providers to consumers. The VPA uses voice or speech signals. The system includes the MPEG-7 audio features and the IDP features to process the signals. These two types of features are fused in a nonlinear way. We adopt a ranking criterion between the machine learning algorithms, the SVM, the ELM, and the GMM, to classify the signal as normal or pathological. In the experiments, we found that the combination between the IDP features and the ELM classifier was more contributing more than the other combinations. The overall system achieved more than 95% accuracy, and the processing time for each sample was just over a second. We focused on VPA using big data as a case study. Our methods can be extended to other types of assessments or predictions using electrocardiograms, mammograms, or other data. All these assessments can be made robust by incorporating machine learning algorithms as we did in this study.

REFERENCES

[1] M. Chen, S. Mao, and Y. Li, "Big data: A survey," *Mobile Netw. Appl.*, vol. 19, no. 2, pp. 171–209, 2014.
[2] J. Andreu-Perez, C. C. Y. Poon, R. D. Merrifield, S. T. C. Wong, and G. Z. Yang, "Big data for health," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 4, pp. 1193–1208, Jul. 2015.
[3] M. Viceconti, P. Hunter, and R. Hose, "Big data, big knowledge: Big data for personalized healthcare," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 4, pp. 1209–1215, Jul. 2015.
[4] J. Luo, M. Wu, D. Gopukumar, and Y. Zhao, "Big data application in biomedical research and health care: A literature review," *Biomed. Inform. Insights*, vol. 8, pp. 1–10, Oct. 2016.
[5] L. A. Tawalbeh, R. Mehmood, E. Benkhelifa, and H. Song, "Mobile cloud computing model and big data analysis for healthcare applications," *IEEE Access*, vol. 4, pp. 6171–6180, Sep. 2016.
[6] S. Nepal, R. Ranjan, and K. K. R. Choo, "Trustworthy processing of healthcare big data in hybrid clouds," *IEEE Cloud Comput.*, vol. 2, no. 2, pp. 78–84, Mar./Apr. 2015.
[7] M. Chen, Y. Ma, J. Song, C. Lai, and B. Hu, "Smart clothing: Connecting human with clouds and big data for sustainable health monitoring," *Mobile Netw. Appl.*, vol. 21, no. 5, pp. 825–845, 2016.
[8] Y. Hu, K. Duan, Y. Zhang, M. S. Hossain, S. M. M. Rahman, and A. Alelaiwi, "Simultaneously aided diagnosis model for outpatient departments via healthcare big data analytics," *Multimedia Tools Appl.*, Jun. 2016, doi: 10.1007/s11042-016-3719-1.
[9] M. S. Hossain and G. Muhammad, "Cloud-assisted industrial Internet of Things (IIoT)—Enabled framework for health monitoring," *Comput. Netw.*, vol. 101, pp. 192–202, Jun. 2016.
[10] W. Wang, Z. Chen, J. Mu, and T. Han, "Throat polyp detection based on compressed big data of voice with support vector machine algorithm," *EURASIP J. Adv. Signal Process.*, vol. 2014, no. 1, pp. 1–6, Dec. 2014.
[11] A. R. Maria, P. Sever, and V. Carlos, "Cloud computing for big data from biomedical sensors monitoring, storage and analyze," in *Proc. Conf. Grid, Cloud High Perform. Comput. Sci. (ROLCG)*, Cluj-Napoca, Romania, Oct. 2015, pp. 1–4.
[12] M. Chen *et al.*, "Disease prediction by machine learning over big healthcare data," *Sci. Rep.*, to be published.
[13] J. Archenaa and E. A. M. Anita, "A survey of big data analytics in healthcare and government," *Procedia Comput. Sci.*, vol. 50, pp. 408–413, Apr. 2015.
[14] M. S. Hossain, M. Zaman, G. Muhammad, A. Ghoneim, and A. Alamri, "Big data-driven services composition using parallel clustered particle swarm optimization in mobile environment," *IEEE Trans. Serv. Comput.*, vol. 9, no. 5, pp. 806–817, Sep./Oct. 2016.
[15] A. Suinesiaputra, P. Medrano-Gracia, B. R. Cowan, and A. A. Young, "Big heart data: Advancing health informatics through data sharing in cardiovascular imaging," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 4, pp. 1283–1290, Jul. 2015.
[16] *Information Technology-Multimedia Content Description Interface-Part 4: Audio*, document. Rec. ISO/IEC CD 15938-4, 2001.
[17] G. Muhammad, "Automatic speech recognition using interlaced derivative pattern for cloud based healthcare system," *Cluster Comput.*, vol. 18, no. 2, pp. 795–802, Jun. 2015.
[18] M. Nemschoff. (Sep. 2013). *Social Media Marketing: How Big Data is Changing Everything*, CMS WiRE. [Online]. Available: http://www.cmswire.com/cms/customer-experience/social-media-marketing-how-big-data-is-changing-everything-022488.php
[19] M. E. Larsen, T. W. Boonstra, P. J. Batterham, B. O'Dea, C. Paris, and H. Christensen, "We feel: Mapping emotion on twitter," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 4, pp. 1246–1252, Jul. 2015.
[20] A. Akay, A. Dragomir, and B.-E. Erlandsson, "Mining social media big data for health," *IEEE Pulse*, Nov./Dec. 2015, [Online]. Available: http://pulse.embs.org/november-2015/mining-social-media-big-data-for-health/
[21] E. A. Mohammed, B. H. Far, and C. Naugler, "Applications of the mapreduce programming framework to clinical big data analysis: Current landscape and future trends," *BioData Mining*, vol. 7, no. 22, pp. 1–23, 2014.
[22] C. MCdonald. *How Big Data is Reducing Costs and Improving Outcomes in Health Care*, MapR Technologies, accessed on Aug. 15, 2016. [Online]. Available: https://www.mapr.com/blog/reduce-costs-and-improve-health-care-with-big-data
[23] J. Godino-Llorente, P. Gomez-Vilda, and M. Blanco-Velasco, "Dimensionally reduction of a pathological voice quality assessment system based on Gaussian mixture models and short-term cepstral parameters," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 10, pp. 1943–1953, Oct. 2006.
[24] A. A. Nasheri *et al.*, "An investigation of multi-dimensional voice program parameters in three different databases for voice pathology detection and classification," *J. Voice*, 2016, doi: 10.1016/j.jvoice.2016.03.019.
[25] G. Muhammad and M. Melhem, "Pathological voice detection and binary classification using MPEG-7 audio features," *Biomed. Signal Process. Controls*, vol. 11, pp. 1–9, May 2014.
[26] G. Muhammad *et al.*, "Voice pathology detection using interlaced derivative pattern on glottal source excitation," *Biomed. Signal Process. Control*, vol. 31, pp. 156–164, Jan. 2017.
[27] V. Parsa and D. G. Jamieson, "Identification of pathological voices based on glottal noise measures," *J. Speech Hearing Res.*, vol. 43, pp. 469–485, Apr. 2000.
[28] G. Muhammad, Y. A. Alotaibi, M. Alsulaiman, and M. N. Huda, "Environment recognition using selected MPEG-7 audio features and mel-frequency cepstral coefficients," in *Proc. 5th Int. Conf. Digital Telecommun. (ICDT)*, Greece, Jun. 2010, pp. 11–16.
[29] A. Shobeirinejad and Y. Gao, "Gender classification using interlaced derivative patterns," in *Proc. 20th Int. Conf. Pattern Recognit. (ICPR)*, 2010, pp. 1509–1512.

[30] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 513–529, Apr. 2012.
[31] *Kay Elemetrics Corp., Disordered Voice Database, Version 1.03 (CD-ROM), MEEI*, Voice Speech Lab., Boston, MA, USA, Oct. 1994.
[32] W. J. Barry and M. Putzer, "Saarbrucken voice database," Inst. Phonetics, Univ. Saarland, Saarbrücken, Germany, Tech. Rep. BA 737/9-6, [Online]. Available: http://www.stimmdatenbank.coli.uni-saarland.de/

M. SHAMIM HOSSAIN (SM'09) received the Ph.D. degree in electrical and computer engineering from the University of Ottawa, Canada. He is currently an Associate Professor with King Saud University, Riyadh, Saudi Arabia. He has authored or co-authored around 120 publications, including the refereed IEEE/ACM/Springer/Elsevier journals, conference papers, books, and book chapters. His research interests include serious games, social media, Internet of Things, cloud and multimedia for healthcare, smart health, and resource provisioning for big data processing on media clouds. He is a member of the ACM and the ACM SIGMM. He has served as a member of the organizing and technical committees of several international conferences and workshops. He was a recipient of a number of awards including, the Best Conference Paper Award, the 2016 ACM *Transactions on Multimedia Computing, Communications and Applications* Nicolas D. Georganas Best Paper Award, and the Research in Excellence Award from King Saud University. He has served as a Co-Chair, the General Chair, the Workshop Chair, the Publication Chair, and a TPC for over 12 IEEE and ACM conferences and workshops. He currently serves as a Co-Chair of the 6th IEEE ICME Workshop on Multimedia Services and Tools for E-health MUST-EH 2016. He is on the Editorial Board of the IEEE ACCESS, the *Computers and Electrical Engineering* (Elsevier), the *Games for Health Journal*, and the International Journal of *Multimedia Tools and Applications* (Springer). He currently serves as a Lead Guest Editor of the *IEEE Communication Magazine*, the IEEE TRANSACTIONS ON CLOUD COMPUTING, the IEEE ACCESS and the *Sensors* (MDPI). He served as a Guest Editor of the IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE (currently JBHI), the International Journal of *Multimedia Tools and Applications* (Springer), the *Cluster Computing* (Springer), the *Future Generation Computer Systems* (Elsevier), the *Computers and Electrical Engineering* (Elsevier), and the *International Journal of Distributed Sensor Networks*.

GHULAM MUHAMMAD received the Ph.D. degree in electrical and computer engineering from Toyohashi University and Technology, Japan, in 2006. He is currently an Associate Professor with the Department of Computer Engineering, CCIS, King Saud University, Riyadh, Saudi Arabia. He has authored or co-authored over 120 publications, including the refereed IEEE/ACM/Springer/Elsevier journals, conference papers, books, and book chapters. He holds a U.S. patent on audio classification. His research interests include serious games, cloud and multimedia for healthcare, image and speech processing, medical signal processing, face recognition, and multimedia forensics.