Federation University

Federation University ResearchOnline

https://researchonline.federation.edu.au

Copyright Notice

© 2018. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/

Ali, Imran, M., Alsulaiman, M., Shoaib, M., & Ullah, S. (2018). Chaos-based robust method of zero-watermarking for medical signals. *Future Generation Computer Systems, 88*, 400–412.

Which has been published in final form at:
https://doi.org/10.1016/j.future.2018.05.058

See this record in Federation ResearchOnline at:
http://researchonline.federation.edu.au/vital/access/HandleResolver/1959.17/181805

CRICOS 00103D RTO 4909

CHAOS-BASED ROBUST METHOD OF ZERO-WATERMARKING FOR MEDICAL SIGNALS

Zulfiqar Ali[1], Muhammad Imran[2], Mansour Alsulaiman[1], Muhammad Shoaib[2], Sana Ullah[3]

[1]Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia.

[2]Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia.
[3]Gyeongsang National University, South Korea

Corresponding authors: xxxx

ABSTRACT

The growing use of wireless health data transmission via Internet of Things is significantly beneficial to the healthcare industry for optimal usage of health-related facilities. However, at the same time, the use raises concerns of privacy protection. Health-related data are private and should be suitably protected. Several pathologies, such as vocal fold disorders, indicate high risks of prevalence in individuals with voice-related occupations, such as teachers, singers, and lawyers. Approximately, one-third of the world population suffer from the voice-related problems during the life span and unauthorized access to their data can create unavoidable circumstances in their personal and professional lives. In this study, a zero-watermarking method is proposed and implemented to protect the identity of patients who suffer from vocal fold disorders. In the proposed method, an image for a patient's identity is generated and inserted into secret keys instead of a host medical signal. Consequently, imperceptibility is naturally achieved. The locations for the insertion of the watermark are determined by a computation of local binary patterns from the time–frequency spectrum. The spectrum is calculated for low frequencies such that it may not be affected by noise attacks. The experimental results suggest that the proposed method has good performance and robustness against noise, and it is reliable in the recovery of an individual's identity.

Key Words: Chaotic system, logistic map, healthcare, privacy protection, vocal fold disorders.

1. INTRODUCTION

Recent developments in the area of the Internet of things (IoT) and signal processing technologies have significantly assisted the healthcare industry in remote data collection and processing. In fact, rapid progress in IoT has made the building of smart homes and cities a reality [1]. With the smart facilities provided in smart homes and

cities, healthcare has become an important service not only for providing health-related facilities without visiting hospitals but also for the optimal use of health resources. In the coming decades, the continuous occupation of beds in hospitals and medical staff, including doctors and paramedics with other resources of healthcare, will be a significant challenge, especially in countries with a large population of senior citizens [2]. The risk of health-related predicaments grows with age. However, several ailments can develop at any stage during life, and vocal fold pathologies are examples of such ailments. Statistics from a recently developed voice pathology database shows that the average age for various vocal fold disorders is 12–48 years [3]. In another voice disorder database, the age of patients who are suffering from such disorders is 21–58 years [4, 5]. Any person can be affected by a vocal fold disorder. However, professional users of voice indicate more risks of the prevalence of such health concerns [6]. Vocal fold disorders can appear because of vocal misuse, including yelling, excessive talking, screaming, and crying. Other factors, such as poor hydration, medication, alcohol consumption, and smoking, also contribute to the development of vocal fold disorders [7]. These factors directly affect the vocal folds and cause abnormal vibrations in them during voice production, thereby affecting voice quality [8].

A person can evaluate their voice for the presence of vocal fold disorders by sending their voice recording to healthcare staff via wireless communication for offline diagnosis. The clinical appraisal, also known as subjective assessment, can be conducted with the help of an endoscopic examination of vocal folds and different acoustic and perceptual measurements. Different techniques, such as Consensus Auditory–Perceptual Evaluation of Voice and Grade, Roughness, Breathiness, Asthenia, Strain scales are used for perceptual assessment [9]. Clinicians use these techniques frequently in practice, but several limitations occur because of the subjective nature of assessment. The limitations may be practitioners' experience and area of expertise, type of dysphonia rating scale, and the degree of vocal fold disorder [10]. Moreover, human error cannot be neglected during subjective assessment. In addition, certain visualization tools, such as video laryngostroboscopy (LSC), are also used in the detection of vocal fold disorders by inspecting vocal fold vibration [11]. The subjective interpretation of LSC examination results strongly depends on the area of specialization and professional expertise of an examiner; thus, LSC examination presents limitations. Therefore, rating scales were introduced in [12, 13]. However, no standard approach is available for LSC examination because of dependency on various factors, such as pitch estimation, periodicity, and sustained phonation frequency [14]. Moreover, videoendoscopic high-speed imaging and videokymography have shown more accurate performance than video LSC [15]. The progress of such approaches is promising [16, 17]. However, expensive equipment and intensive labor in performing data analysis may be the reasons that practitioners and investigators hesitate to adopt these new methods [18]. These limitations of subjective methods can be overcome by an automatic assessment of vocal fold disorders.

Various methods for the automatic detection of vocal folds have been reported in the literature in response to the various limitations of subjective evaluation [19–22]. These automatic methods, together with cloud or edge computing, can also be used for real-time detection of vocal fold disorders. Automatic methods do not require huge or big data to generate the acoustic model of normal and disordered persons for the training of classifiers. Only a reasonable amount of data, such as a few hundred speech signals, are sufficient for training. Once the classifier is trained, these automatic methods indicate capabilities to process big data for the detection of vocal folds reliably and with high accuracy.

In both situations, diagnosis is either offline or in real time. The privacy of a person is a prime concern. The term privacy does not have a single interpretation, and its meaning usually depends on the scenario. Westin defines privacy as the claim of an individual to determine what information about himself or herself should be known to others [23]. Depending on usage, privacy was defined in six different manners in [24]. One of these definitions is "secrecy – the concealment of certain matters from others." The breach of the identity of a person not only affects their personal life but can also disturb their professional life. Such a breach may create severe circumstances and can ultimately result in job loss. This study aims to protect the privacy of persons suffering from a vocal fold disorder by proposing and implementing a new method of zero-watermarking for medical signals.

Watermarking has been widely used for data concealment to protect and authenticate audio [25]. Watermarking ensures the protection and authentication of the identity of a person that is embedded into the host audio. Although the use of watermarking in medical images has been extensively studied [26–30], no significant work has been reported for medical speech signals. Only a few studies have been conducted for privacy protection using medical speech signals. In one study [31], the authors claimed that imperceptibility was achieved after the insertion of watermark in a speech signal. However, the inserted watermark changed the characteristics of the speech signal, thereby affecting the diagnosis of a disorder. To avoid the effect of a watermark in a medical speech signal, the other possibility is the insertion of the watermark in a secret key instead of the host medical signal. This approach is called zero-watermarking. In this approach, a significant challenge is to extract the suitable features of the medical signal such that a secret key containing the watermark can be generated by their use. Given the extraction of such features from the signals, privacy protection in medical audio signals via zero-watermarking is a difficult task. A disorder detection system uses normal and disordered signals as input. Therefore, extracted features should be robust such that both types of signal may have them. Otherwise, insertion of a watermark will not be possible in both types of signals.

A zero-watermarking algorithm for medical speech signals was proposed in [32, 33]. The identity of a patient is encrypted via visual cryptography before watermarking. The encrypted identity is inserted using the features of the signal in the time domain. With the use of zero-watermarking, imperceptibility is naturally achieved, and the experimental results indicated that the insertion and extraction processes of the developed method are reliable. One of the limitations of these studies is that the dimensions of the secret share are generated via visual cryptography. The dimension of each share is equivalent to the original identity. Thus, double capacity is required if two secret shares are generated. In another study, a zero-watermarking method for privacy protection in telemedicine was provided [34]. The identity of a patient is directly embedded using the characteristics of the signal. The speech signal used for the study contains running speech. Therefore, a voice activity detection (VAD) module is implemented to detect unvoiced frames for reliable insertion of identity. However, VAD is itself a difficult task, especially in case of disordered signals [35].

In this work, a new zero-watermarking method for privacy protection is proposed and implemented. The technique inserts a watermark, an identity of a person, in an audio signal. The identity is in the form of an image. After watermark insertion, speech signals are not distorted because the identity is inserted in a secret key instead of the host signal. Therefore, identity insertion does not affect the result of the detection of vocal fold disorders. Usually, a watermarking algorithm is evaluated on the basis of imperceptibility and robustness. Imperceptibility describes that the perceived quality of a speech signal should not be degraded after insertion of the watermark. In the proposed method, imperceptibility is naturally achieved. Various experiments are performed to ensure that no clue of an individual's identity is present after embedding in the secret key and thus ensure the robustness of the proposed zero-watermarking technique. Experimental results indicate that the proposed method is reliable in the extraction of identity and only authorized healthcare staff can extract identities with the relevant secret information.

The remainder of the paper is organized as follows. Section 2 describes the procedure for the computation of the time–frequency spectrum for the insertion of the watermark and generation of a chaotic sequence via a logistic map. In addition, this section presents the embedding and extraction processes of the proposed zero-watermarking method. Section 3 provides details of the vocal fold disorder detection system and shows the baseline results. Section 4 reports the evaluation of the proposed zero-watermarking method and the robustness of the technique against noise. Finally, Section 5 presents the conclusions.

2. PROPOSED ZERO-WATERMARKING METHOD

The identification of suitable locations for the insertion of a watermark for identity protection is crucial. The locations should be stable and provide sufficient space for the watermark. To determine such locations, the time–frequency spectrum for low frequencies is computed in this study. Moreover, the patterns in the spectrum are analyzed via the local binary pattern (LBP) operator, and histograms are used to represent the occurrence of the obtained LBP codes. In addition, randomness in the proposed zero-watermarking method is created by using a chaotic system, which provides deterministic randomness and can be regenerated by using the same initial conditions. This section also provides the embedding and extraction steps of the watermark in detail.

2.1 Low-frequency Regions in Medical Signals

The privacy of an individual is protected by inserting the subject's identity using the zero-watermarking method. Watermark insertion and extraction are two vigorous processes in the proposed method, and they are important for accurate insertion and recovery of the watermark. The watermark in the proposed method is inserted in the low-frequency spectrum of audio. The low-frequency regions of an audio exhibit constant behavior and contain only small variations. Therefore, these regions are stable for watermark insertion.

The time–frequency spectrum that contains low frequencies is obtained by applying discrete Fourier transformation (DFT) on an audio *A*. The DFT is computed by using the fast Fourier transformation algorithm, thereby reducing the complexity of DFT from $O(n^2)$ to $O(n\ logn)$. Then, the obtained spectrum is passed through triangular Mel-spaced bandpass filters. The Mel-scale filter bank is defined as

$$\Upsilon(z,m)=\begin{Bmatrix} 0 & \text{when} \quad \rho^z < \rho_c^{m-1} \\ \dfrac{\rho^z-\rho_c^{m-1}}{\rho_c^m-\rho_c^{m-1}} & \text{when} \quad \rho_c^{m-1} \le \rho^z < \rho_c^m \\ \dfrac{\rho^z-\rho_c^{m+1}}{\rho_c^m-\rho_c^{m+1}} & \text{when} \quad \rho_c^m \le \rho^z < \rho_c^{m+1} \\ 0 & \text{when} \quad \rho^z \ge \rho_c^{m+1} \end{Bmatrix}, \tag{1}$$

where $z = 1, 2, 3, \ldots, L\text{-}1$; L denotes the length of the *ith* frame of audio *A*.

The frequencies for the bandpass Mel-spaced filters are calculated by using Eq. (2), where ρ represents the frequencies in Hertz (Hz) and ψ stands for the corresponding frequencies in the Mel scale.

$$\psi = 2595\log_{10}\left(\frac{\rho}{700}+1\right) \tag{2}$$

Then, center frequencies of the bandpass filters are computed by using Eq. (3).

$$\begin{aligned} &\psi_c^m = m\cdot\Delta\psi \quad \text{for} \quad m=1,2,3,\ldots,M \\ &\text{where} \\ &\Delta\psi = \frac{\psi_{\max}-\psi_{\min}}{M+1} \end{aligned}, \tag{3}$$

where ψ_{max} and ψ_{min} are maximum and minimum frequencies in the Mel scale, respectively, and correspond to the maximum and minimum frequencies ρ_{max} and ρ_{min} in Hz, respectively. In addition, the number of Mel-scale bandpass filters is denoted by M.

All signals of the voice disorder database are down-sampled to 25 KHz before finding the low-frequency regions. Therefore, the maximum usable frequency in each audio signal is 12.5 KHz. To obtain the low-frequency regions, the minimum and maximum frequencies, ρ_{min} and ρ_{max}, respectively, are adjusted to 100 and 500 Hz, respectively. Moreover, M=24; thus, the number of bandpass filters in the Mel-spaced filter bank is 24. The computed time–frequency spectrum that contains low-frequency regions are shown in Figs. 1(a) and 1(b) for normal and disordered persons, respectively.

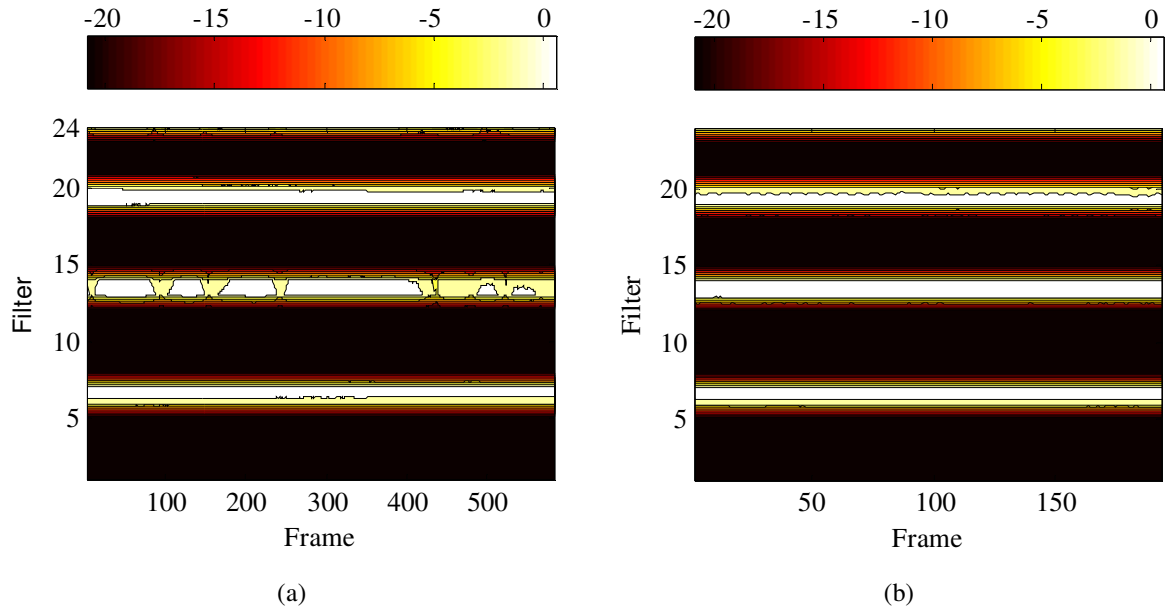
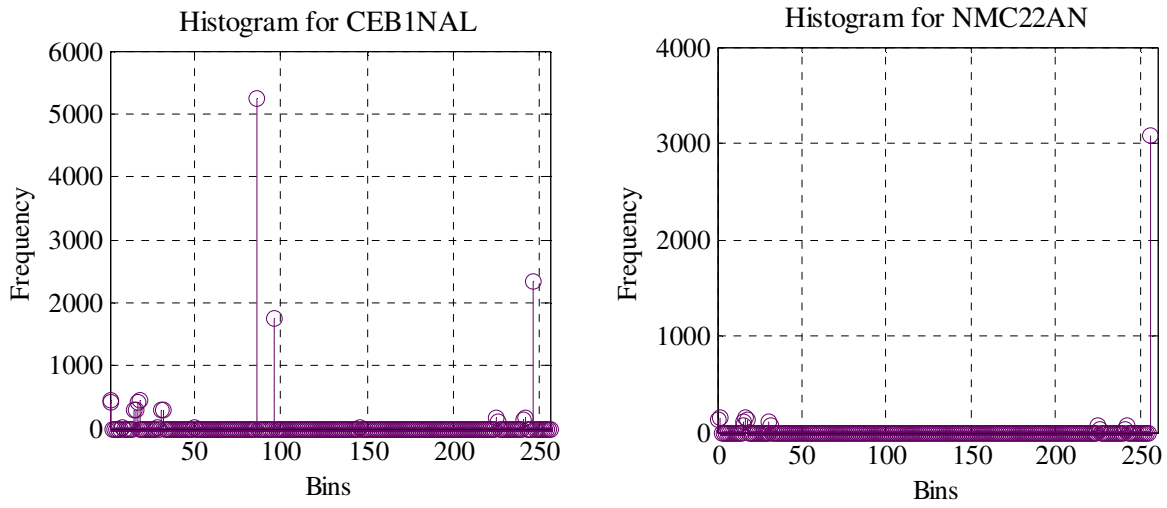


Figure 1: Low-frequency spectra for two different audio samples: (a) CEB1NAL - normal person, (b) NMC22AN - disordered person.

The patterns in the computed low-frequency region are investigated by applying a two-dimensional LBP operator. The LBP operator divides the region into 3 x 3 blocks. Each element of the time–frequency spectrum is considered a central element of the block, and it is replaced by the calculated LBP codes. To obtain the LBP code for a block, the center element is compared with all eight neighbors of the 3 x 3 block. If a neighbor is greater than or equal to the center element, then the neighbor is replaced by 1; otherwise, it is 0. This code is referred as sign LBP code [19]. Thus, an 8-bit binary digit is obtained, and the corresponding decimal number represents the required LBP code. The histograms obtained by applying the two-dimensional LBP operator for the time–frequency spectra are shown in Fig. 2. The figure shows that only few LBP codes are repeating. For instance, in Fig. 2 (a), occurrences of LBP codes 0, 85, 95, 245, and 255 are 445, 5238, 1746, 2328, and 0, respectively; by contrast, the same LBP codes in Fig. 2(b) are repeated 140, 0, 0, 0, and 3072 times, respectively. These codes provide sufficient space for watermark insertion. The occurrences of the total number of LBP codes vary in both histograms because of the varying durations of the audio signal. The duration of the recorded sample for the normal person is 3 s in the voice disorder database, while that for the disordered person is only 1 s.

To secure the identity of a person, the randomness in the insertion of the watermark is crucial. In this study, the randomness is attained via chaotic theory, and it generates random sequences deterministically.



(a) (b)

Figure 2: Histograms representing patterns for (c) CEB1NAL (normal person) and (b) NMC22AN (disordered person).

2.2 Chaos and Logistic Map

Chaos represents a state of randomness that is not merely a disarray. In fact, the randomness of a chaotic system is deterministic and can be obtained by using the same initial conditions. However, a slight change in the initial condition results in a highly different sequence [36, 37]. Despite the deterministic simplicity over time, chaos theory can produce wildly unpredictable and divergent behavior because of such sensitivity [38]. Defining chaos completely is difficult. According to Williams Garnett, "chaos is sustained and orderly-looking long-term evolution that satisfies certain special mathematical criteria and that occurs in deterministic nonlinear systems" [39].

In this study, a chaotic sequence is generated by using a logistic map, a degree 2 polynomial mapping given by Eq. (4).

$$R_{r+1} = 4\lambda R_r\left(1-R_r\right) \quad \text{or} \quad R_{r+1} = \mu R_r\left(1-R_r\right), \text{ where } \mu = 4\lambda \tag{4}$$

where system parameter $\mu \in [0, 4]$ and initial condition $R_0 \in (0,1)$. In the logistic function, a difference equation treats time as continuous. By contrast, in the logistic map, the difference equation examines discrete time steps [40, 41]. The logistics map is known as such because it maps the population value at the present time to the population value at the next time step.

The behavior of Eq. (4) varies for the different values of the μ. When μ is between 0 and 1, the population vanishes ultimately regardless of the value of the initial population. For $\mu \in (1, 2)$, the population approach to the value (μ-1)/ μ in a quick time duration is irrespective of the value of the initial population. By contrast, the logistic map behaves

chaotically for $\mu \in (3.5699456, 4]$ [42]. The logistic map with initial condition $R_0 = 0.5$ and system parameter $\mu = 3.7$ and 3.716 are plotted in Fig. 3(a) with 200 iterations; i.e., $r = 200$. This figure also emphasizes the difference in population for $\mu = 3.7$ and $\mu = 3.716$. A small difference of 0.016 in μ generates significantly diverse sequences. For further analysis, autocorrelation of both logistic maps, the first with $\mu = 3.7$ and the second with $\mu = 3.716$, are plotted

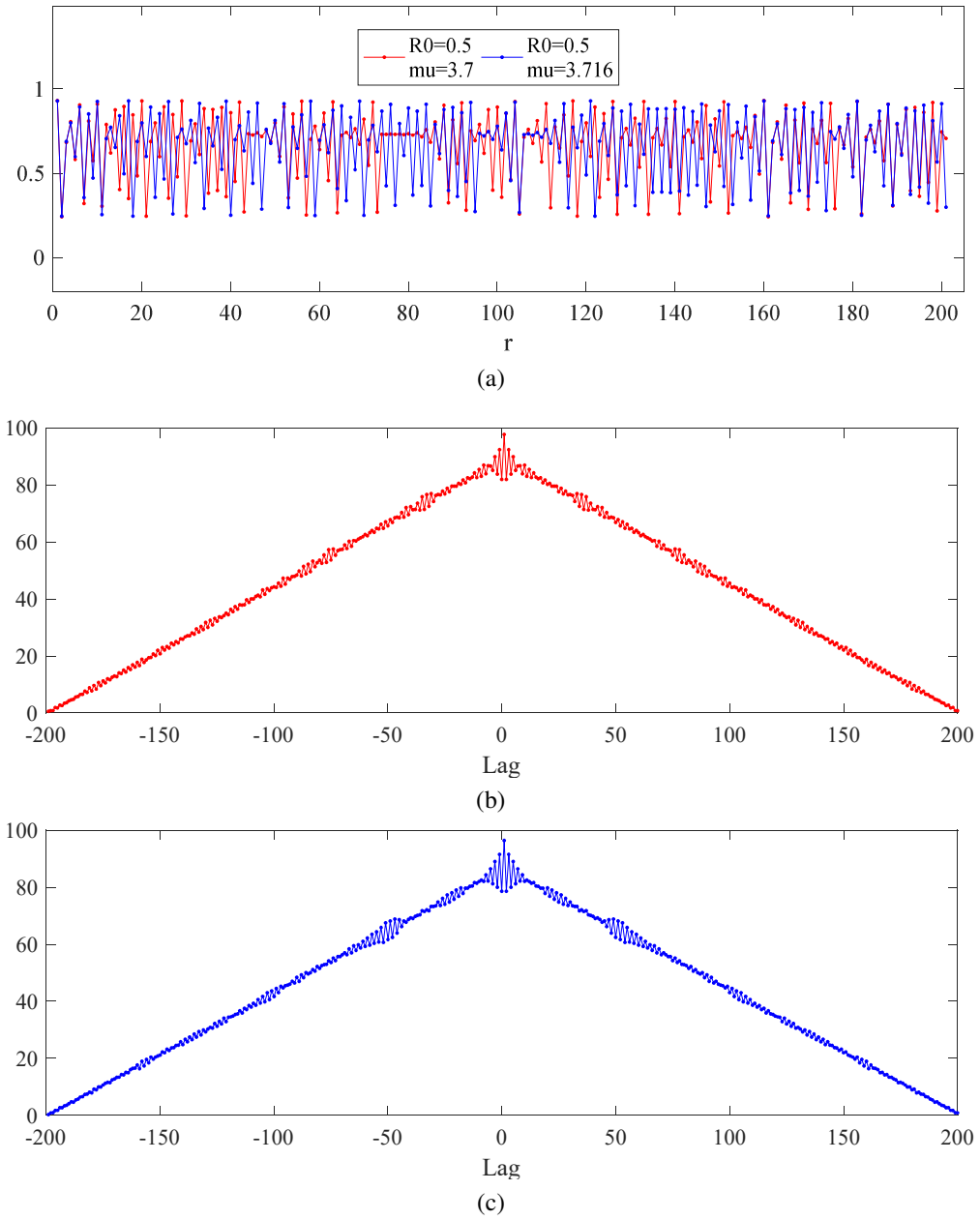


Figure 3: (a) Difference between two logistic maps (b) Autocorrelation of the first logistic map with parameters $R_0 = 0.5$ and $\mu = 3.7$ (c) Autocorrelation of the second logistic map with parameters $R_0 = 0.5$ and $\mu = 3.716$.

in Figs. 3(b) and 3(c), respectively. The figures show that the logistics maps generated by a small change in system parameter μ are statistically uncorrelated.

2.3 Embedding Process

The process for the insertion of the identity of a person is shown in Fig. 4 and implemented through the following steps.

1. We generate an image I for the identity of an individual with the dimension I_1 x I_2 to watermark the host medical signal A.
2. We divide the host signal A into overlapping frames $a_1,a_2,a_3,\ldots,a_f$, where the length of each frame is L. Then, we compute the time–frequency spectrum S in the range 100–500 Hz by using the Mel-spaced bandpass filter bank, as described in Section 2.1, to identify the locations for watermarking. The dimension of the obtained spectrum is t x q, which indicates that the spectrum has t frames and q bandpass filters.
3. We segment spectrum S into blocks $B_1, B_2, B_3, \ldots, B_l$, which are of size 3 x 3, to compute the LBP codes.
4. Using the computed LBP codes, we generate an index key X by using the criteria given in Eq. (5). The index key X contains the locations of following blocks, where (i) the center element is smaller than all

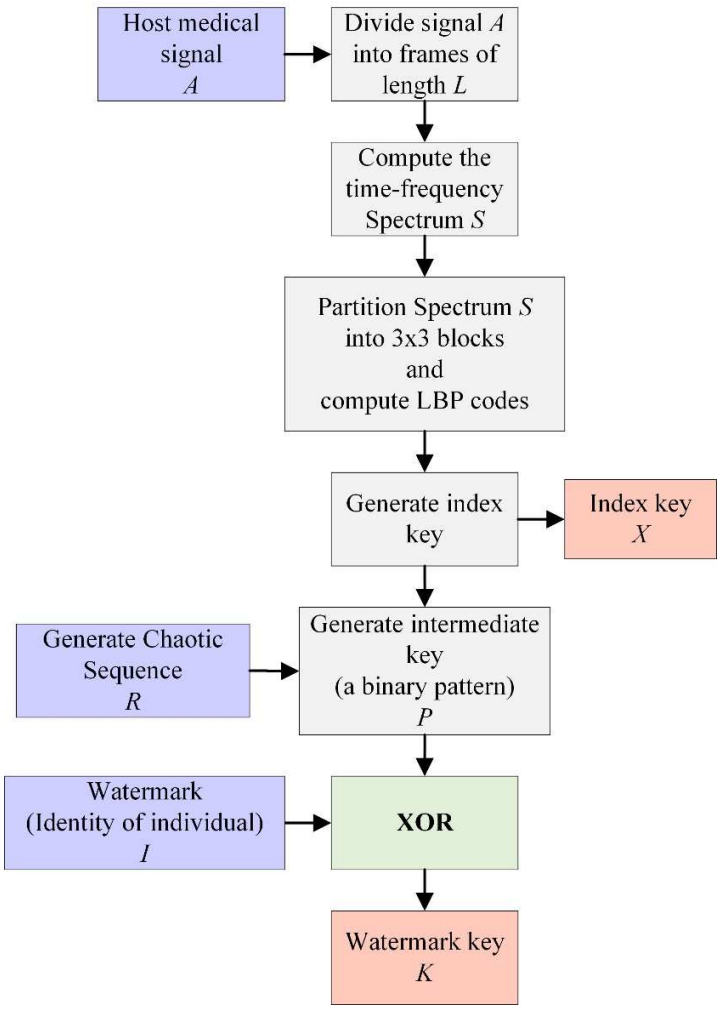


Figure 4: Process for watermark embedding.

neighbors, i.e., 11111111 (say $C1$); (ii) the center element is greater than all neighbors, i.e., 00000000 (say $C2$); (iii) the center element is smaller than odd neighbors and greater than even ones, i.e., 01010101 (say $C3$); (iv) the center element is smaller than the first four neighbors such that 01011111 (say $C4$); and (v) the center element is smaller the last four neighbors such that 111110101 (say $C5$). In each code, $C1$, $C2$, $C3$, $C4$, and $C5$, the bit at the extreme right represents the first neighbor.

$$X_i = \left\{ (T,Q) \mid LBP(b_{(T,Q)}) \in \left\{ C1, C2, C3, C4, C5 \right\} \right\} \quad \text{where} \quad T = 2,3,4,\ldots,t-1 \quad Q = 2,3,4,\ldots,q-1 \quad , \tag{5}$$

where $b_{(T,Q)}$ is the center element of a block.

5. We compute chaotic sequence R by using Eq. (4) with the initial conditions R_0 and μ. These conditions vary from signal to signal. The dimension of the sequence R is equivalent to the number of locations in X. The initial conditions of the generated sequence are known only to the sender and the intended receiver. Given that a small change in the initial conditions creates an entirely different sequence, no other individual will be able to generate the same sequence.

6. An intermediate key P is produced by obtaining the binary pattern P, which is given by Eq. (6).

$$P_i = \begin{cases} 1 & if \quad round\left(LBP(B(X_i)) \text{ divided by } R_i \right) \text{is Even} \\ 0 & if \quad round\left(LBP(B(X_i)) \text{ divided by } R_i \right) \text{is odd} \end{cases} \tag{6}$$

7. We create a watermark key K by using Eq. (3) and then conduct an XOR operation between the identity I of an individual and the intermediate key P.

$$K = I \oplus P \tag{7}$$

Two keys, the index key X and watermark key K, and the initial conditions of the chaotic system must be transmitted via a secure channel to the healthcare staff with a corresponding speech signal.

2.4 Extraction Process

A block diagram of the extraction process is shown in Fig. 5, and the identity of an individual is recovered via the following steps.

1. We partition the watermarked audio $\tilde{A}$ into frames of length L to compute the time–frequency spectrum $\tilde{S}$ of dimension t x q by using the Mel-spaced bandpass filter bank that is in the range of 100–500 Hz, as described in Section 2.1. The variables t and q represent the number of frame and bandpass filters in the spectrum $\tilde{S}$.

2. We compute the LBP codes for the blocks $\tilde{B}$ of size 3 x 3, whose locations (T,Q) are provided by the index key X; i.e., $LBP(\tilde{B}(X_i))$, where T = 2,3,4,…,t and Q = 2,3,4,…,q-1.

3. We generate chaotic sequence R with the same initial conditions and determine the binary pattern $\tilde{P}$ by using the relation provided in Eq. (8).

$$\tilde{P}_i = \begin{cases} 1 & if \quad round\left(LBP(\tilde{B}(X_i)) \text{ divided by } R_i \right) \text{is Even} \\ 0 & if \quad round\left(LBP(\tilde{B}(X_i)) \text{ divided by } R_i \right) \text{is odd} \end{cases} \tag{8}$$

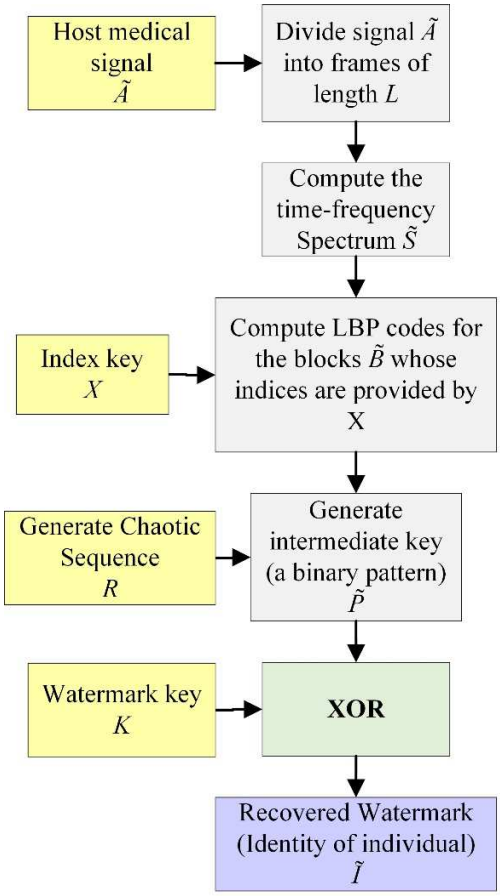


Figure 5. Process for watermark extraction.

4. We perform the XOR operation between the computed binary pattern $\tilde{P}$ and the watermark key K for the recovery of the individual's identity $\tilde{I}$ by using Eq. (9).

$$\tilde{I} = K \oplus \tilde{P} \tag{9}$$

After Step 4, the individual's identity $\tilde{I}$ is recovered from signal $\tilde{A}$, which was inserted in the medical signal A during the embedding process.

In the healthcare system, observing diagnosis accuracy is important. Therefore, a vocal fold disorder detection system is developed to obtain the baseline result for the detection of such disorders.

3. BASELINE RESULTS FOR DISORDER DETECTION

A reliable healthcare system may protect the identity of a person and provide correct decisions for the presence of an illness. In any type of healthcare system, the first preference is the accurate diagnosis of disease, which should not be compromised for privacy protection. In this work, an automatic vocal fold disorder detection (AVDD) system is implemented to discriminate between normal and disordered speech signals. The AVDD system aims to obtain

baseline results for the proposed privacy-protected healthcare system. The results will be used to observe the effect of zero-watermarking on a speech signal.

The AVDD system comprises two important components: (1) speech feature extraction method and (2) pattern-matching technique. Both components are crucial for the accurate differentiation of normal and disordered signals. Generally, speech feature extraction methods fall into two categories, namely, those based on the human voice production system and those based on the human auditory perception system. Linear prediction coefficients (LPC) and LPC-based cepstral coefficients (LPCC) belong to the first category. By contrast, Mel-frequency cepstral coefficients (MFCC) belong to the second category and have indicated great success over LPC and LPCC in various speech-related applications, including voice disorder detection [32]. Therefore, state-of-the-art speech features (MFCC) are applied for the extraction of speech features from both signal types. MFCC provides multi-dimensional feature vectors that cannot be interpreted by the human mind. Therefore, the pattern-matching phase is essential in determining the trend in a speech signal.

In this work, pattern matching is performed by implementing a machine learning technique, which performs better do statistical approaches. Machine learning techniques do not make strict assumptions about data and instead learn to represent complex relationships in a data-driven manner. A state-of-the-art machine learning technique, the Gaussian mixture model (GMM) [43, 44], is used to develop the AVDD system and has been successfully and widely used in different scientific areas. The basis for using GMM is that the distribution of feature vectors obtained from MFCC can be modeled by a mixture of Gaussian densities. GMM [45] uses a K-means algorithm to initialize the parameters. Moreover, these parameters are estimated and tuned by the well-known expectation maximization algorithm [46] to converge to a model, thereby providing a maximum log-likelihood value.

The implemented GMM is vital for both phases of the developed AVDD system. The first phase is the training phase, and the second is the testing phase. In the first phase, the computed MFCC features are inputted into the GMM for the generation of acoustic models for each type of speech signal, i.e., one acoustic model for normal signals and another for disordered signals. Once the models are generated, an unknown speech signal is compared with both generated models during the testing phase to compute the log-likelihood for decision. The unknown signal having maximum log-likelihood with a model is deemed the type of that signal. If the log-likelihood of an unknown signal is greater for the model of the normal signal, then the type of the signal is normal; otherwise, the type of signal is disordered.

To obtain the baseline results for the AVDD system, the normal and disordered speech signals are taken from the MEEI voice disorder database [4, 47] which is recorded at the Massachusetts Eye & Ear Infirmary laboratory. The MEEI database is recorded at sampling frequencies of 25 and 50 KHz. Therefore, all normal and disordered signals are down-sampled to 25 KHz to obtain a unique sampling frequency for all signals. In this study, a subset of the MEEI database represented by $MEEI_{sub}$ is used to perform the experiments, and it has been used in many voice disorder detection and classification systems [32, 48]. The distribution of normal and disordered speech signals in the $MEEI_{sub}$ is provided in Table 1.

One of the reasons of using $MEEI_{sub}$ is that the age range of both genders for normal and disordered subjects is nearly similar. Moreover, the age and gender of both types of subjects are evenly distributed, and various disordered are considered in it [5]. Overall, the $MEEI_{sub}$ contains 226 signals, namely, 173 disordered and 53 normal signals. Among the 173 disordered signals, 70 signals are recorded by male speakers and the remaining 103 are recorded by female speakers. In case of normal signals, the samples recorded by male and female speakers in $MEEI_{sub}$ are 21 and 32, respectively.

Table 1. Distribution of normal and pathological samples in $MEEI_{sub}$ [5]

Subjects	Gender	Number of Samples	Mean Age (Years)	Age Range (Years)	Standard Deviation (Years)
Pathological	Male	70	41.7	26–58	9.4
	Female	103	37.6	21–51	8.2
Normal	Male	21	38.8	26–59	8.5
	Female	32	34.2	22–52	7.9

The baseline results of the developed AVDD system are reported by using three performance measures. The first measure is sensitivity (SNY), which provides the detection rate of the accurately detected disordered signals. The second measure is specificity (SPY), which provides the detection rate of accurately detected normal signals. The third measure is accuracy (ACY), which provides the detection rate of all truly detected normal and disordered signals from the entire dataset. The relations used to compute these measures are provided in Eqs. (10)–(12).

$$SNY = \frac{\text{true positive}}{\text{true positive + false negative}} \times 100 \tag{10}$$

$$SPY = \frac{\text{true negative}}{\text{true negative + false positive}} \times 100 \tag{11}$$

$$ACY = \frac{\text{truely detected signals}}{\text{total number of signals}} \times 100 \tag{12}$$

Here, true positive means that the AVDD system detects a disordered signal as a disordered signal, true negative means that the AVDD system detects a normal signal as a normal signal, false negative means that the AVDD system detects a disordered signal as a normal signal, and false positive means that the AVDD system detects a normal signal as a disordered signal. All results provided in Table 2 are obtained by a five-fold cross-validation technique. The results listed in Table 2 are averaged over five folds, and the standard deviation (SD) is listed to show the variation of the measures among the folds.

Table 2. Baseline results of developed AVDD system

Number of Gaussians	Sensitivity SNY±SD	Specificity SPY±SD	Accuracy ACY±SD
2	89.56 ± 3.4	84.91 ± 5.1	88.49 ± 3.3
4	94.2 ± 2.9	86.91 ± 8.5	92.47 ± 1.2
8	94.22 ± 2	79.27 ± 8	90.71 ± 0.9
16	96.52 ± 1.3	73.82 ± 9.8	91.15 ± 2.7
32	97.68 ± 1.3	50.91 ± 10.3	86.72 ± 2.2

Various experiments are conducted for the detection of vocal fold disorders by using a different number of Gaussian mixtures, such as 2, 4, 8, 16, and 32. The maximum obtained accuracy is 92.47%, and it is achieved by using four Gaussian mixtures. The corresponding SNY and SPY is 94.2% and 86.91, respectively.

The obtained baseline results provide a platform from which to observe the effect of insertion and extraction of the watermark using the proposed method. To do so, the accuracy of the disordered detection system is computed after the insertion and extraction of the watermark from the host signal.

4. EVALUATION OF PROPOSED ZERO-WATERMARKING METHOD

Different experiments are performed to observe the performance of the proposed zero-watermarking method. Only authorized access with relevant secret keys can disclose the identity of an individual. Moreover, the proposed method is tested against noise attacks.

4.1. Embedding Reliability

During the embedding process, the proposed zero-watermarking method hides the identity of a person and generates secret keys, namely, index key X and watermark key K. These generated keys are transmitted to the healthcare staff with the corresponding speech signals for diagnosis of the vocal fold disorders. The identity of the person cannot be disclosed without access to the secret keys. For identity recovery, authorized staff computes the speech features. Then, transmitted keys help extract the identity of a person.

To observe the reliability of the embedding process of the proposed method, all samples of the $MEEI_{sub}$ given in Table 1 are used for identity insertion. In $MEEI_{sub}$, the signals of normal subjects are labeled with a string suffixed by NAL such as AXH1NAL and EDC1NAL. By contrast, the labels of disordered signals are suffixed by AN, such as AXT13AN and CAC10AN. In this study, black-and-white images are created by using the labels of signals, which are strings of seven alphanumeric characters. These created images are considered watermarks and embedded in medical signals by using the proposed zero-watermarking method.

In the proposed technique, the watermark is embedded in watermark key K instead of the host audio A. This key is produced by performing an XOR operation between the patient identity I and the generated binary pattern P during the embedding process. An identity of an individual (GZZ1NAL) and the embedded watermark by using the proposed method is shown in Fig. 6.

(a) (b)

Figure 6: (a) Identity of an individual and (b) embedded watermark.

Figure 6 shows that no information regarding the identity of an individual is visible through the embedded watermark. The difference between the two images, given in Fig. 6, can be found by calculating the bit-error rate (ERR) and peak signal-to-noise ratio (PSNR). The metrics ERR and PSNR are defined in Eqs. (13) and (14) and used to obtain the differences between two images objectively. ERR refers to the bits that are present in the embedded watermark but do not exist in the recovered watermark at the corresponding positions. Moreover, in Eq. (14), im_1 and im_2 represent the original identity of a person and embedded watermark, respectively, and BPS signifies the number of bits in each pixel of the image.

$$ERR = \frac{\text{erroneous bits}}{\text{total bits in the watermark}} \times 100 \ , \qquad (13)$$

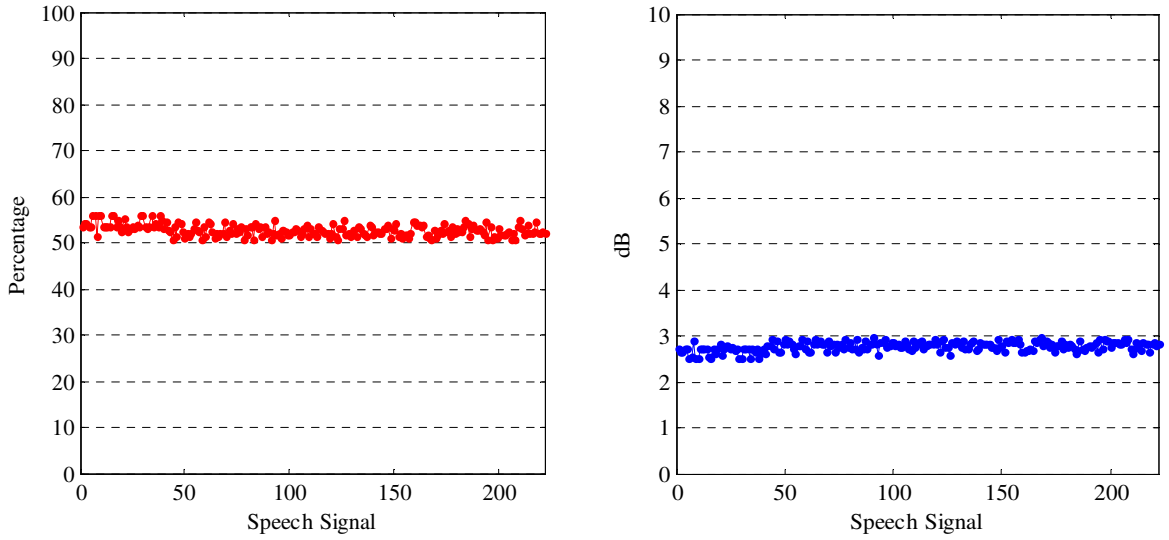
$$PSNR(im_1, im_2) = 20\log_{10}\left(\frac{2^{BPS}-1}{MSE}\right)$$

where . (14)

$$MSE = \frac{\sum_{\alpha=1}^{I_1}\sum_{\beta=1}^{I_2}\left(im_1(\alpha,\beta) - im_2(\alpha,\beta)\right)^2}{I_1 \times I_2}$$

The ERR between the original identity and the embedded watermark is 53.57%. Thus, more than 50% mismatch occurs between the bits of the original identity and the embedded watermark. Guessing the identity of an individual is impossible. The high percentage of the ERR ensures that the embedding process of the proposed method is reliable. Meanwhile, the value of the second metric PSNR used to measure the difference between the original identity and the embedded watermark is 2.71 dB. Two images are considered close to each other if the PSNR is above 20 dB. In the case of 50 dB, the images are nearly similar. For the images given in Fig. 6, the PSNR is close to zero; thus, the images are entirely different and identity breach is impossible. The second metric also strengthens the reliability of the embedding process.

Similarly, for all signals of $MEEI_{sub}$, ERR and PSNR are computed to observe the difference between the original identities and the embedded watermarks. The ERR and PSNR for all normal and disordered signals are shown in Figs. 7(a) and 7(b), respectively. The ERR for all signals is more than 50%, and the PSNR values are close to zero. Therefore,



(a) (b)

Figure 7: (a) ERR for all speech signals of $MEEI_{sub}$, (b) PSNR for all speech signals of $MEEI_{sub}$.

the proposed zero-watermarking method securely protects the identity of a person during the embedding process.

4.2 Extraction Reliability

The extraction of a patient's identity should be revealed only using the relevant secret keys and the initial conditions of the chaotic system. The extraction process cannot be considered reliable if the secret key of one patient discloses the identity of other patients. To observe the extraction reliability of the proposed zero-watermarking method, we attempt to reveal the identity of a person by using the secret information of the other patients. The EER and PSNR

between the original identity and the extracted identity computed using irrelevant secret information are depicted in Fig 8.

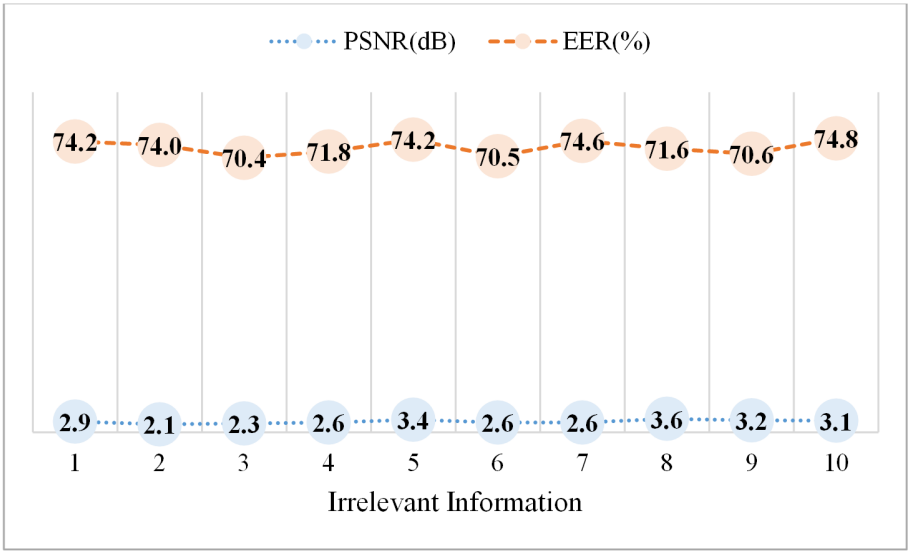


Figure 8: PSNR and EER with use of irrelevant information for identity disclosure of patient.

In Fig. 8, we attempt to reveal the identity of a person GZZ1NAL by using the information of 10 other persons. The experimental results show that when the information of 10 irrelevant people is used to access the identity of GZZ1NAl, the obtained PSNR is close to zero and the EER is above 70%. The low values of PSNR suggest that the original identity and the recovered identities with irrelevant information are significantly different. Similarly, the high EER values between the original identity and the recovered identities obtained using irrelevant information indicate that the recovered identities are entirely different from the original identities. Analysis of both metrics indicates that the proposed method reveals the identity of a person only when the secret information of the relevant person is available.

We must also investigate the proposed method for identity extraction with the relevant secret key. The recovered identities with the relevant secret information should be compared with the corresponding original identities to observe the similarity between them. Figure 9 shows the embedded and recovered identity of GZZ1NAL using the relevant secret keys and the initial conditions. The recovered identity is not distorted and is similar to the original identity shown in Fig. 6(a). The extraction of identity without distortion shows that the EER and PSNR between the original and recovered identities are 0% and infinity (Inf), respectively. These values indicate that the identity is recovered perfectly.

(a) (b)

Figure 9: Extraction of identity with relevant information: (a) embedded identity, (b) recovered identity.

Experiments are performed to extract the identity of all signals of $MEEI_{sub}$ with the relevant secret information generated during the embedding process, and both metrics are computed. The computed metrics for all signals are listed in Table 3.

The experimental results in Table 3 show that the computed metrics EER and PSNR are 0% and Inf, respectively, for all signals. Thus, the identities that use the relevant secret keys and initial conditions are recovered without any distortion. Hence, when secret keys of other speech signals are used to access the identity of an individual, the recovered identity does not disclose any information. The proposed method reveals the identity of a person only when

relevant secret keys are available. The proposed zero-watermarking method is reliable such that no unauthorized access to the individual's identity is possible.

Table 3: Comparison between original and recovered identities with relevant information for all signals in $MEEI_{sub}$

Metric	Comparison between Original and Recovered Identities
PSNR (dB)	Inf
EER	0%

The accuracy of the disorder detection system with the signal from which the identity is recovered must also be observed. Therefore, several experiments are performed to explore the effect of privacy protection on disorder diagnosis. The accuracy of the AVDD system with the recovered signals is listed in Table 4.

Table 4. Results of developed AVDD system with recovered signals

Number of Gaussians	**Sensitivity SNY±SD**	**Specificity SPY±SD**	**Accuracy ACY±SD**
2	88.56 ± 2.7	85.05 ± 4.6	88.23 ± 1.9
4	95.4 ± 3.1	86.36 ± 7.9	92.85 ± 1.1
8	94.78 ± 1.8	80.01 ± 6.9	91.10 ± 0.8
16	96.31 ± 1.3	74.23 ± 7.9	91.74 ± 2.2
32	97.16 ± 1.5	51.52 ± 9.5	87.09 ± 1.9

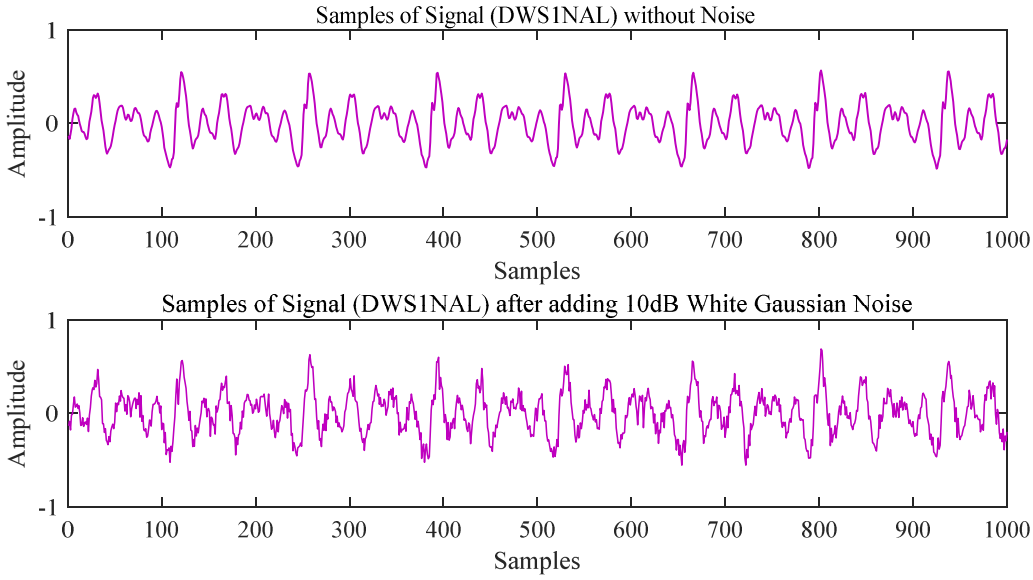
The setup used for the experiment is the same as that used to obtain the baseline results in Section 3. The acquired results are similar to those that are attained with the original signals. However, a comparison of Tables 2 and 4 indicates a small variation between the disorder detection results of the original signal and the signals used for recovery. This variation is the effect of the cross-validation approach, which assigns the testing and training signals randomly each time. In Tables 2 and 4, the maximum achieved accuracy is obtained with four Gaussian mixtures, and these accuracies are 92.47 ± 1.2 and 92.85 ± 1.1. Therefore, the proposed privacy protection method does not affect the accuracy of the diagnosis system.

The proposed method is investigated in the following section for protecting the identity of a person in case of a noise attack.

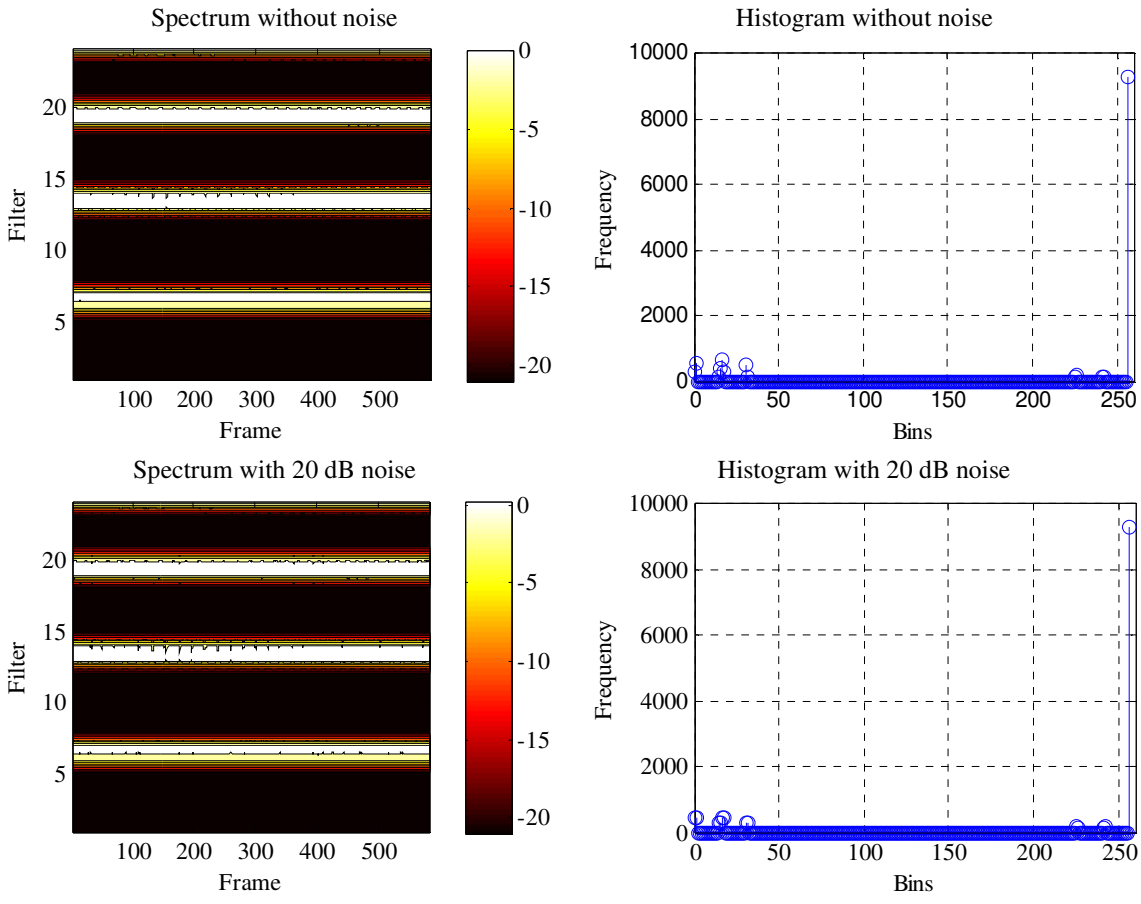
4.3 Robustness

A watermarking method is considered robust against noise if it can recover the identity of a person after the addition of noise in the watermarked audio. In case of a noise attack, the importance of the extracted feature can be realized. If extracted features are not robust against the attack, then the identity of a person will use the best method in this study such that they should not be affected by the noise attack. The features in the proposed method are extracted to determine the locations for insertion of the watermark. The features are vital in the embedding and recovery processes.

To observe the robustness of the proposed method, white Gaussian noise with SNR of 20 dB is added to the speech signal DWS1NAL. The speech signals without and with noise are shown in Fig. 10(a). Moreover, the time–



(a)



(b) (c)

Figure 10: (a) Speech signal (DWS1NAl) with and without noise. (b) Spectrum of DWS1NAl with and without noise. (c) Histogram of DWS1NAl with and without noise.

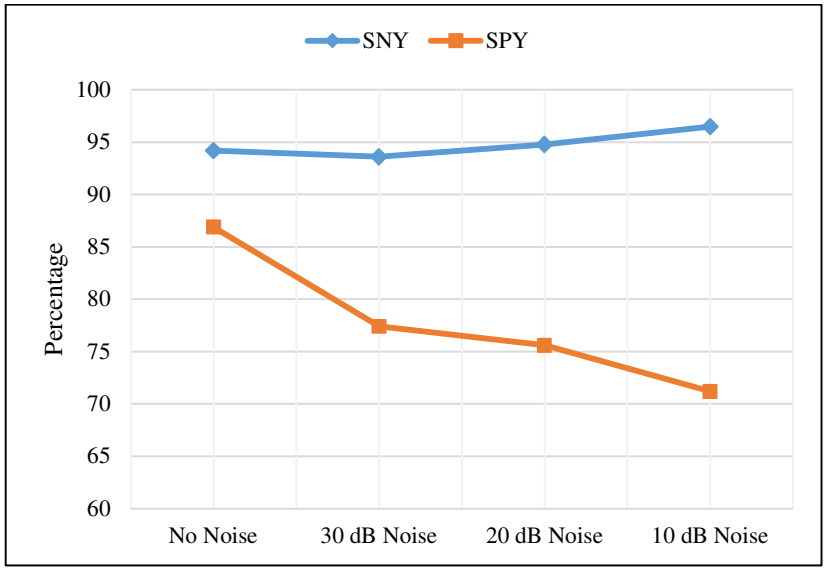


Figure 11: Comparison of vocal fold detection for clean and noisy signals.

frequency spectra with low-frequency regions are depicted in Fig. 10(b) for the signal with and without noise. Both spectra do not show any significant difference, thereby suggesting that the extracted features in the proposed method are robust against noise. Ultimately, the robustness of the proposed method against noise attacks is proven. Histograms for the noisy and clean signals are shown in Fig. 10(c). These histograms provide the frequency of the LBP codes, and they are unaltered for the clean and noisy signals. The identity of a person can be embedded and recovered reliably even in the case of a noise attack by using the proposed zero-watermarking method.

In addition, the effect on the detection of vocal fold disorder is analyzed by adding noise with SNR of 30, 20, and 10 dB in all normal and disordered speech signals of $MEEI_{sub}$. The results for the detection of vocal fold disorders without noise are presented in Table 2. The maximum obtained accuracy is obtained with four Gaussian mixtures. Therefore, for the noisy signals, the same number of Gaussian mixtures is used to obtain the SNY and SPY. A comparison between the SNY and SPY for the noisy and clean signals is shown in Fig. 11. SPY decreases with the increase in noise because the normal signals exhibit complex and transient behavior similar to disordered signals after the addition of noise. Therefore, the detection system AVDD detects the normal signals as disordered signals. Thus, the proposed method extracts the identity of an individual successfully from the signal even after a noise attack. By contrast, the detection system is unable to diagnose it correctly.

4.4 Comparison

The method in [32, 33] provides dual protection for the identity of a patient. The identity of the patient is encrypted by using visual cryptography prior to watermarking. The limitation of this method is the generation of secret shares with a large dimension. For instance, the dimension of the original identity is 20 × 126, and two secret shares of dimensions double those of the original identity, i.e., 40 × 252, are generated. This process increases the size of the watermark by 800% from 2520 bits to 20160 bits. Sometimes, such capacity to insert such a large watermark capacity is not available in the signals [49]. In the proposed study, the identity is not dual-protected. However, the highest obtained EER under the use of irrelevant keys to recover the identity is better than in [32, 33]. The highest EER in this study is approximately 75%, and that for [32, 33] is 63%, thereby suggesting that the proposed method is reliable in the recovery of a patient's identity.

In another study [34], a reliable method for the protection of patient information was presented. A zero-watermarking approach is implemented for embedding and recovery of the identity. The method requires unvoiced

frames for identity insertion. The study concluded that voiced frames are not suitable for identity insertion. Therefore, unvoiced frames, which can be considered a region of interest (ROI), are determined by applying a VAD module. The detection of voice activity is itself a challenging task, especially in the case of dysphonic patients, as mentioned by Umapathy et al. in [35]. The reason for this situation is that speech signals of dysphonic patients contain more noise than do the signals of normal persons, and VAD in noisy signals indicates many limitations [50]. The false calculation of ROI may reveal the identity of a person, which is not desirable [51]. Therefore, we do not use any VAD module in the proposed method to make it accurate and reliable.

5. CONCLUSIONS

A zero-watermarking method is proposed and implemented in this study. The proposed method hides the identity of a patient in normal speech signals and those affected by vocal fold disorders. Low-frequency regions in the time–frequency spectrum are computed for the insertion of the watermark because they are considered a stable part of a speech signal and are not considerably affected by noise. Given the use of the chaotic system, patient identity breach is impossible; a small change in the initial conditions generates an entirely different sequence. This characteristic is a strong and vital aspect of the proposed zero-watermarking method. Thus, the technique is reliable in the protection and recovery of identity. This method can be deployed confidently for the protection of patient identity in healthcare applications, in which data are transmitted via IoT, because the use of unauthorized access to disclose patient identity without relevant secret keys is not effective. In future work, the proposed method will be extended, and the identity of a patient can be embedded in the form of multiple secret shares. Patient identity will not be disclosed until all shares are available simultaneously. However, the generated secret shares should not increase the size of the watermark.

ACKNOWLEDGMENTS

The authors are grateful to the Deanship of Scientific Research, King Saud University, Riyadh, Saudi Arabia, for funding through the Research Group Project no. 1435-051.

REFERENCES

[1] T. K. L. Hui, R. S. Sherratt, and D. D. Sánchez, "Major requirements for building Smart Homes in Smart Cities based on Internet of Things technologies," *Future Generation Computer Systems,* vol. 76, pp. 358-369, 2017.

[2] T. Guelzim, M. S. Obaidat, and B. Sadoun, "Chapter 1: Introduction and overview of key enabling technologies for smart cities and homes," in *Smart Cities and Homes*, M. S. Obaidat and P. Nicopolitidis, Eds., 1 ed Cambridge, MA 02139, USA: Morgan Kaufmann, 2016, pp. 1-16.

[3] T. A. Mesallam, M. Farahat, K. H. Malki, M. Alsulaiman, Z. Ali, A. Al-nasheri*, et al.*, "Development of the Arabic Voice Pathology Database and Its Evaluation by Using Speech Features and Machine Learning Algorithms," *Journal of Healthcare Engineering,* vol. 2017, p. 13, 2017.

[4] Massachusette Eye & Ear Infirmry Voice & Speech LAB, "Disordered Voice Database Model 4337 (Ver. 1.03) ", ed. Lincoln Park, NJ: Kay Elemetrics Corp., 1994.

[5] V. Parsa and D. G. Jamieson, "Identification of Pathological Voices Using Glottal Noise Measures," *Journal of Speech, Language, and Hearing Research,* vol. 43, pp. 469-485, 2000.

[6] N. Roy, R. M. Merrill, S. D. Gray, and E. M. Smith, "Voice disorders in the general population: prevalence, risk factors, and occupational impact," *Laryngoscope,* vol. 115, pp. 1988-95, 2005.

[7] I. R. Titze, J. Lemke, and D. Montequin, "Populations in the U.S. workforce who rely on voice as a primary tool of trade: a preliminary report," *Journal of Voice,* vol. 11, pp. 254-259, 1997.

[8] P. L. Dhingra and S. Dhingra, *Diseases of ear, nose and throat*, 6 ed.: Elsevier, India, 2014.

[9] K. Nemr, M. Simoes-Zenari, G. F. Cordeiro, D. Tsuji, A. I. Ogawa, M. T. Ubrig*, et al.*, "GRBAS and Cape-V scales: high reliability and consensus when applied at different times," *J Voice,* vol. 26, pp. 812.e17-22, 2012.

[10] J. L. Sofranko and R. A. Prosek, "The Effect of Experience on Classification of Voice Quality," *Journal of Voice,* vol. 26, pp. 299-303, 2012.

[11] V. Uloza, A. Vegiene, and V. Saferis, "Correlation between the quantitative video laryngostroboscopic measurements and parameters of multidimensional voice assessment," *Biomedical Signal Processing and Control,* vol. 17, pp. 3-10, 2015.

[12] B. J. Poburka, "A new stroboscopy rating form," *Journal of Voice,* vol. 13, pp. 403-413, 1999.

[13] C. A. Rosen, "Stroboscopy as a Research Instrument: Development of a Perceptual Evaluation Tool," *The Laryngoscope,* vol. 115, pp. 423-428, 2005.

[14] S. Deguchi, Y. Ishimaru, and S. Washio, "Preliminary Evaluation of Stroboscopy System Using Multiple Light Sources for Observation of Pathological Vocal Fold Oscillatory Pattern," *Annals of Otology, Rhinology & Laryngology,* vol. 116, pp. 687-694, 2007.

[15] R. Patel, S. Dailey, and D. Bless, "Comparison of High-Speed Digital Imaging with Stroboscopy for Laryngeal Imaging of Glottal Disorders," *Annals of Otology, Rhinology & Laryngology,* vol. 117, pp. 413-424, 2008.

[16] C. Bohr, A. Kraeck, U. Eysholdt, A. Ziethe, and M. Döllinger, "Quantitative analysis of organic vocal fold pathologies in females by high-speed endoscopy," *The Laryngoscope,* vol. 123, pp. 1686-1693, 2013.

[17] C. Manfredi, L. Bocchi, G. Cantarella, and G. Peretti, "Videokymographic image processing: Objective parameters and user-friendly interface," *Biomedical Signal Processing and Control,* vol. 7, pp. 192-201, 2012.

[18] P. Woo, "Objective Measures of Laryngeal Imaging: What Have We Learned Since Dr. Paul Moore," *Journal of Voice,* vol. 28, pp. 69-81, 2014.

[19] Z. Ali, M. Talha, and M. Alsulaiman, "A Practical Approach: Design and Implementation of a Healthcare Software for Screening of Dysphonic Patients," *IEEE Access,* vol. 5, pp. 5844-5857, 2017.

[20] 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. PP, pp. 1-1, 2017.

[21] Z. Ali, I. Elamvazuthi, M. Alsulaiman, and G. Muhammad, "Detection of Voice Pathology using Fractal Dimension in a Multiresolution Analysis of Normal and Disordered Speech Signals," *Journal of Medical Systems,* vol. 40, p. 20, 2015.

[22] Z. Ali, I. Elamvazuthi, M. Alsulaiman, and G. Muhammad, "Automatic Voice Pathology Detection With Running Speech by Using Estimation of Auditory Spectrum and Cepstral Coefficients Based on the All-Pole Model," *Journal of Voice,* vol. 30, pp. 757.e7-757.e19, 11// 2016.

[23] A. F. Westin, "Social and Political Dimensions of Privacy," *Journal of Social Issues,* vol. 59, pp. 431-453, 2003.

[24] D. J. Solove, *Understanding Privacy*. Cambridge, Massachusetts, United States: Harvard University Press, 2008.

[25] G. Hua, J. Huang, Y. Q. Shi, J. Goh, and V. L. L. Thing, "Twenty years of digital audio watermarking—a comprehensive review," *Signal Processing,* vol. 128, pp. 222-242, 2016/11/01/ 2016.

[26] M. Arsalan, A. S. Qureshi, A. Khan, and M. Rajarajan, "Protection of medical images and patient related information in healthcare: Using an intelligent and reversible watermarking technique," *Applied Soft Computing,* vol. 51, pp. 168-179, 2017/02/01/ 2017.

[27] J. J. Garcia-Hernandez, W. Gomez-Flores, and J. Rubio-Loyola, "Analysis of the impact of digital watermarking on computer-aided diagnosis in medical imaging," *Computers in Biology and Medicine,* vol. 68, pp. 37-48, 2016/01/01/ 2016.

[28] A. Roček, K. Slavíček, O. Dostál, and M. Javorník, "A new approach to fully-reversible watermarking in medical imaging with breakthrough visibility parameters," *Biomedical Signal Processing and Control,* vol. 29, pp. 44-52, 2016/08/01/ 2016.

[29] F. Y. Shih and X. Zhong, "High-capacity multiple regions of interest watermarking for medical images," *Information Sciences,* vol. 367, pp. 648-659, 2016/11/01/ 2016.

[30] R. Thanki, S. Borra, V. Dwivedi, and K. Borisagar, "An efficient medical image watermarking scheme based on FDCuT–DCT," *Engineering Science and Technology, an International Journal,* 2017/07/01/ 2017.

[31] M. Alhussein and G. Muhammad, "Watermarking of Parkinson Disease Speech in Cloud-Based Healthcare Framework," *International Journal of Distributed Sensor Networks,* vol. 11, p. 264575, 2015.

[32] Z. Ali, M. Imran, M. Alsulaiman, T. Zia, and M. Shoaib, "A zero-watermarking algorithm for privacy protection in biomedical signals," *Future Generation Computer Systems,* vol. 82, pp. 290-303, 2018/05/01/ 2018.

[33] Z. Ali, M. Imran, W. Abdul, and M. Shoaib, "An Innovative Algorithm for Privacy Protection in a Voice Disorder Detection System," in *Biologically Inspired Cognitive Architectures (BICA) for Young Scientists*, 2018, pp. 228-233.

[34] Z. Ali, M. S. Hossain, G. Muhammad, and M. Aslam, "New Zero-watermarking Algorithm Using Hurst Exponent for Protection of Privacy in Telemedicine," *IEEE Access,* 2018. DOI: 10.1109/ACCESS.2018.2799604.

[35] K. Umapathy, S. Krishnan, V. Parsa, and D. G. Jamieson, "Discrimination of pathological voices using a time-frequency approach," *Biomedical Engineering, IEEE Transactions on,* vol. 52, pp. 421-430, 2005.

[36] A. Hastings, C. L. Hom, S. Ellner, P. Turchin, and H. C. J. Godfray, "Chaos in Ecology: Is Mother Nature a Strange Attractor?," *Annual Review of Ecology and Systematics,* vol. 24, pp. 1-33, 1993.

[37] D. Rickles, P. Hawe, and A. Shiell, "A simple guide to chaos and complexity," *Journal of Epidemiology and Community Health,* vol. 61, pp. 933-937, 2007.

[38] G. Boeing, "Visual Analysis of Nonlinear Dynamical Systems: Chaos, Fractals, Self-Similarity and the Limits of Prediction," *Systems,* vol. 4, p. 37, 2016.

[39] G. Williams, *Chaos Theory Tamed*: Joseph Henry Press: Washington, D.C., United States

1997.

[40] H. Pastijn, "Chaotic Growth with the Logistic Model of P.-F. Verhulst," in *The Logistic Map and the Route to Chaos: From The Beginnings to Modern Applications*, M. Ausloos and M. Dirickx, Eds., ed Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 3-11.

[41] S. H. Strogatz, *Nonlinear Dynamics and Chaos*, 2nd ed. Westview Press: Boulder, CO, USA, 2014.

[42] H.-O. Peitgen, H. Jürgens, and D. Saupe, *Chaos and Fractals: New Frontiers of Science*, 2nd ed.: Springer-Verlag New York, 2004.

[43] C. M. Bishop, *Pattern Recognition and Machine Learning*: Springer-Verlag New York, 2006.

[44] D. A. Reynolds, "Speaker identification and verification using Gaussian mixture speaker models," *Speech Communication,* vol. 17, pp. 91-108, 1995/08/01 1995.

[45] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: analysis and implementation," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 24, pp. 881-892, 2002.

[46] R. A. Redner and H. F. Walker, "Mixture Densities, Maximum Likelihood and the EM Algorithm," *SIAM Review,* vol. 26, pp. 195-239, 1984.

[47] Z. Ali, M. S. Hossain, G. Muhammad, and A. K. Sangaiah, "An intelligent healthcare system for detection and classification to discriminate vocal fold disorders," *Future Generation Computer Systems,* vol. 85, pp. 19-28, 2018/08/01/ 2018.

[48] G. Muhammad, G. Altuwaijri, M. Alsulaiman, Z. Ali, T. A. Mesallam, M. Farahat*, et al.*, "Automatic voice pathology detection and classification using vocal tract area irregularity," *Biocybernetics and Biomedical Engineering,* vol. 36, pp. 309-317, 2016/01/01/ 2016.

[49] Z. Xu and W. Yuan, "Watermark BER and Channel Capacity Analysis for QPSK-Based RF Watermarking by Constellation Dithering in AWGN Channel," *IEEE Signal Processing Letters,* vol. 24, pp. 1068-1072, 2017.

[50] S. Mousazadeh and I. Cohen, "Voice Activity Detection in Presence of Transient Noise Using Spectral Clustering," *IEEE Transactions on Audio, Speech, and Language Processing,* vol. 21, pp. 1261-1271, 2013.

[51] R. Eswaraiah and E. S. Reddy, "Robust medical image watermarking technique for accurate detection of tampers inside region of interest and recovering original region of interest," *IET Image Processing,* vol. 9, pp. 615-625, 2015.