A Novel Local Feature Generation Technique Based Sound Classification Method for Covid-19 Detection using Lung Breathing Sound

Türker Tuncer[1] [ID], Erhan Akbal[2] [ID], Emrah Aydemir[3*] [ID], Samir Brahim Belhaouari[4] [ID] and Sengul Dogan[5] [ID]

[1] Firat University, Department of Digital Forensics Engineering, Elazig, Turkey. (e-mail: turkertuncer@firat.edu.tr).
[2] Firat University, Department of Digital Forensics Engineering, Elazig, Turkey. (e-mail: erhanakbal@firat.edu.tr).
[3*] Sakarya University, Department of management information systems, Sakarya, Turkey. (e-mail: emrahaydemir@sakarya.edu.tr).
[4] Hamad Bin Khalifa University, College of Science and Engineering, Doha, Qatar. (e-mail: sbelhaouari@hbku.edu.qa).
[5] Firat University, Department of Digital Forensics Engineering, Elazig, Turkey. (e-mail: sdogan@firat.edu.tr).

ARTICLE INFO

ABSTRACT

Lung breathing sounds have been used to diagnose many diseases, including Covid-19. Nowadays, Covid-19 has affected daily life worldwide, and it has caused a global pandemic. Generally, computer vision methods have been presented to classify healthy, pneumonia, and Covid-19. They achieved high classification rates on datasets with a limited number of classes without taking into consideration other lung diseases. Our main hypothesis is to detect Covid-19 automatically among other lung diseases by using lung breathing sounds. Therefore, a dataset of lung breathing sound with ten classes has been collected, and a novel lung sounds classification method has been proposed in this paper. This method presents a novel local feature generation technique, and Substitution Box (S-Box) of the present lightweight encryption method is utilized as a pattern. A novel nonlinear pattern is presented based on S-Box, named Present-SBox-Pat (present S-Box pattern). A new pooling-based transformation (maximum tent pooling (MaTP)) is proposed to generate high, middle, and low levels features. It is considered as a preprocessing method of this work. ReliefF and iterative neighbourhood component analysis (RFINCA) selector is used to select the most discriminative and informative features. Two shallow classifiers are used to obtain results. The proposed Present-SBox-Pat and MaTP feature generation network and RFINCA feature selector-based method achieved 95.43% classification accuracy using the SVM classifier. These results demonstrated the success of techniques in generating and selecting features that facilitate the task of classifiers.

1. INTRODUCTION

Many methods/techniques have been presented to detect and cure the Covid-19 in the literature. One of them is machine learning [1, 2]. Machine learning methods have employed to detect Covid-19 automatically. Covid-19 affects the lungs and causes respiratory diseases. Respiratory sounds can provide information about human health to detect these diseases. Doctors and medical professionals also use lung breathing sounds in the diagnosis of many diseases. These are Fine Crackles (Rales), Wheezing (expiratory), Rhonchi, Stridor, Coarse Crackles (Rales), Bronchovesicular, Bronchial, and Coronavirus (COVID-19) [3, 4]. Covid-19 has become a disease that affects unprecedented public health in the world. The main symptoms in patients are cough, difficulty breathing, headache, and fever [5]. The disease exhibits behaviors similar to bronchitis and pneumonia. Therefore, lung breathing sounds of the patients have listened with a stethoscope, and then CT images are taken, and lung conditions are monitored. Lung and heart injuries are seen on CT images [5]. PCR test is performed to make a definitive diagnosis [6]. Diagnosing Covid-19 patients at an early stage increases the success of the treatment. The use of CT or RT-PCR kits extends diagnosis time [7]. The virus mostly affects the lung in infected patients, and therefore lung breathing sounds can be used to diagnose. Thus, rapid diagnosis and early diagnosis will be provided. Doctors use many different methods to identify and diagnose lung breathing sound features. The simplest and most common method used is the use of a stethoscope [8]. After auscultation of the sounds with a stethoscope, additional examinations are requested. Different factors may cause errors in diagnosing with a stethoscope. These are the

doctor's inability to distinguish between lung sounds, or the ear cannot hear low-frequency sounds. Thus, there are errors in diagnosis. Computer-aided techniques are patient-friendly and economical. It also minimizes errors caused by the human factor and helps the doctor to decide. An automated lung sounds classifier algorithm consists of pretreatment of signals, extraction of features and classification stages.

1.1. MOTIVATION

The widely known lung diseases are asthma, bronchitis, water collection in the chest, shortness of breath, hydatid cyst, pleurisy, tuberculosis, pneumonia, and Covid-19. The purpose of lung testing is to diagnose lung diseases, if possible, to determine their causes and evaluate the degree of severity. Many doctors want to scan lung or pulmonary function tests (PFTs), chest films or CT to examine the oxygen and carbon dioxide levels, to help diagnose and monitor lung function, and the anatomy of the lungs. Covid-19 is a very contagious disease, and many healthcare professionals have been infected. Also, early detection of the Covid-19 disease is crucial like other diseases. Therefore, automated Covid-19 detection systems should be developed to help healthcare professionals, and many X-ray and CT based Covid-19 disease detection methods have been proposed in the literature [9-11]. These methods are machine learning-based methods and used image processing/computer vision methods [12, 13]. However, these methods have been applied to the image datasets with two or three classes dataset. Since they cannot achieve high classification accuracies by using CT or X-ray image datasets with more classes, this paper motivates us to detect Covid-19 disease using lung breathing sounds gathered from stethoscopes. Therefore, we collected a novel lung breathing sound dataset with *ten* classes, and one of these classes is Covid-19. Our other motivations are to develop a novel intelligent lung disease detection method by using an intelligent stethoscope and show the effectiveness of the sound-based ML method to image-based methods.

1.2. LITERATURE REVIEW

Breathing, coughing, sneezing sounds have been used in automatic recognition systems and have a high accurate classification rate [10, 14-17]. In the literature, there are various studies conducted to diagnose artificial intelligence based on lung breathing sounds [18]. It is difficult to decide as a result of listening to the sounds of the lungs using a stethoscope. Analyzing sounds with artificial intelligence algorithms reduces the error rate of misdiagnosis. Therefore, various studies have been conducted in literature. Sankur et al. [19] used samples from 12 healthy 13 respiratory disease patients sound. These sounds are divided into two classes with the autoregression model as normal and abnormal. They obtained a 93.75% accuracy rate with the k-nearest neighbour method. Doyle [18] analyzed 35 patients' breathing sounds to train clinicians. These sounds include crackles, wheezes, pleural friction rub, and stridor. Histograms of sounds were extracted, and obtained signals were analyzed by artificial neural networks. An 83% classification rate was obtained. Lei,

Rahman, and Song [20] placed a microphone close to the mouth and nose of the patients, and breathing sounds were collected from 90 people. The number of healthy people in the dataset is 40, and the number of people with different respiratory disorders such as influenza, pneumonia, and bronchitis is 50. Breathing sounds were recorded for 1 minute in a comfortable position, and analyzes were performed with LSVM, SSVM, MLP, RBF, SOM, LVQ, KNN. The highest classification rate was achieved by 99% using LSVM and LVQ. Shokrollahi et al. [21] collected tracheal sounds of 9 patients. It is aimed to diagnose sleep apnea by segmenting snoring sounds. 95.9% accuracy rate was obtained. Aykanat et al. [3] collected 17930 lung sounds over 1630 people. The sounds were transferred to the computer using a specially developed electronic stethoscope. Rale, rhonchus, and normal sounds were analyzed using Mel frequency cepstral coefficient. CNN and SVM methods achieved an 86% successful classification rate. Huq and Moussavi [22] examined the breathing sounds of 93 individuals without any respiratory distress. They present a method using several breath sound parameters to differentiate between the two respiratory phases. 95.6% accuracy rate was obtained.

Scientists are working hard to facilitate the treatment and diagnosis phase of Covid-19. Therefore, many diagnostic methods based on artificial intelligence have been proposed. Han et al. analyzed the speech sounds of patients diagnosed with COVID-19. Disease severity, sleep quality, fatigue, and anxiety states were categorized with a 69% accuracy rate [23]. Jiang et al. [24] presented a deep learning-based method using thermal cameras and mobile phones. He proposed a device for analyzing respiratory sounds and determining the health of Covid-19. The device can detect with an accuracy of 83.7%. Maghdid et al. [25] proposed a method that uses sensors on the cell phone to diagnose low-cost Covid-19. [26] examined CT images of 108 patients diagnosed with Covid-19 and 86 patients with specific and viral pneumonia diseases. The results of AlexNet, VGG-16, VGG-19, SqueezeNet, GoogleNet, MobileNet-V2, ResNet-18, ResNet-50, ResNet-101 and Xception were compared using images with the diagnosis of Covid-19. As a result, successful classification procedures were performed between 86.27% and 99.51%. Li et al. F. Shi et al. [27] 78 patients who were positive for the Covid-19 test classified them in three groups as mild, normal, and critical. CT images were analyzed using the total severity score method and classified with 95% accuracy. Shi et al. [28] examined CT images of 1027 community-acquired pneumonia patients with 1658 Covid-19 cases. Random Forest method was applied to feature extraction images. 87.9% of successful classification results were obtained. Narin et al. [29] performed chest X-ray images of 50 Covid-19 patients in ResNet50, InceptionV3 and Inception-ResNetV2. As a result of the analysis, 98% accuracy for ResNet50, 97% for InceptionV3 and 87% for Inception-ResNetV2 were obtained.

1.3. OUR METHOD

In this work, we collected a novel lung breathing sound datasets to propose an intelligent hand-crafted feature-based multilevelled feature generation network is suggested to

extract high, middle, and low levels features. A pooling based decomposition method is considered to create levels for multilevel feature generation and improve routing ability of the pooling methods. This method is called as MaTP. The proposed feature generation network uses bot textural features (local binary pattern (LBP) like extractor feature) [30] and statistical features [31] together. In the LBP like feature extractors, linear patterns have been used for feature generation. This work uses a nonlinear pattern as feature extractors. S-Box of the present encryption method is considered as a pattern to design this textural feature extractor, and it is named as Present-SBox-Pat. The proposed Present-SBox-Pat and MaTP methods are used together to create a new generation hand-crafted feature generation network. RFINCA feature selector selects the most valuable features automatically. The selected features are utilized as an input of the k nearest neighbours (kNN) [32] and support vector machine (SVM) [33] classifiers, which are the most preferred shallow/conventional classifiers. This method aims to propose a highly accurate and automated breathing sound classification method.

1.4. CONTRIBUTIONS AND NOVELTIES

Novelties and contributions of the proposed Present-SBox-Pat and MaTP stable feature generator and RFINCA feature selector-based lung breathing sound classification method are;

- A new breathing sound dataset was collected from YouTube [34] to create a testbed for the sound classification method. These sounds are publicly published for educational purposes on YouTube. This experimental dataset is created to demonstrate the success of the sound classification method to diagnose lung diseases. Nowadays, Covid-19 is a disease that affects daily life the most. Therefore, this dataset includes Covid-19 breathing sounds.
- S-Boxes are the widely preferred nonlinear components of the symmetric encryption methods, and they have been generally used for substitution. This study uses an S-Box as a feature extractor for the first time as far as we know in the literature.
- Maximum pooling [35] has been widely used in deep learning networks, especially the convolutional neural network. A new improved version of maximum pooling is presented (MaTP) to improve the success of the maximum pooling.
- There are an optimum automated number of features selection problem for ReliefF [36] and NCA [37]. RFINCA method solves this problem and uses the usefulness of both ReliefF and NCA.
- We obtained a highly accurate model for lung breathing sound classification.

2. THE PROPOSED LUNG BREATHING SOUNDS CLASSIFICATION METHOD

This work proposes a new generation ML method to attain a high accurate breathing sound classification method. This method has four fundamental phases, and these are MaTP based preprocessing, feature generation, feature selection, and classification. MaTP method is utilized as

decomposition and level creation method to generate features deeply, and five levelled networks (five levelled networks is the best resulted according to experiments. Therefore, we used five levelled networks). To obtain high, low, and medium levels features, a fused feature generation network is presented by using the created levels by MaTP decomposition. In this network, statistical features and textural features are generated together. *Nineteen* most preferred statistical moments are employed for statistical feature generation. A new extractor is presented to extract textural features, and this method benefits the nonlinear structure of the symmetric ciphers, which is S-Box. Therefore, the S-Box of the present lightweight encryption method [38, 39] is employed as a pattern to show the success of the nonlinear structures in the feature generation. RFINCA method selects the most distinctive/illuminating features for the used classifiers, and these features selected are forwarded to classifiers. Graphical summarization of the proposed Present-SBox-Pat and MaTP based feature generator and RFINCA feature selector-based sound classification method is shown in Fig. 1.
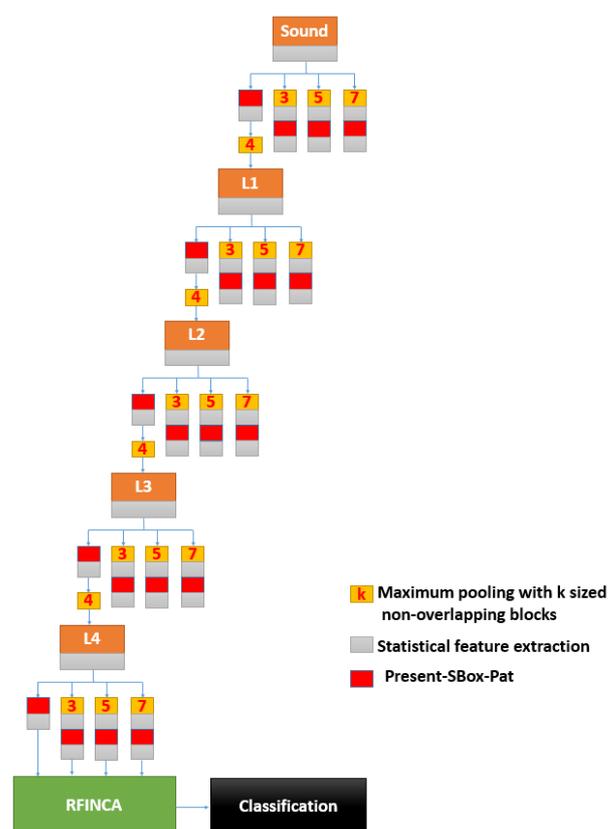
Figure 1. Graphical summarization of the proposed Present-SBox and RFINCA based breathing lung sounds classification method.

Details of the proposed Present-Sbox and RFINCA based methods are clarified in sub-sections.

2.1. PREPROCESSING

The preprocessing phase is the first phase of our method. Maximum pooling is one of the most preferred decomposition and preprocessing methods. Since it is

useful and has low computational complexity, however, it is not a good router because it routes the peaks value of a fixed-sized block. Therefore, we presented a novel maximum pooling, and it is named MaTP. Maximum Tent Pooling (MaTP) is utilized as preprocessing. MaTP is created by using variable pooling methods with variable sized non-overlapping blocks. This work uses *four* maximum pooling with 3,4,5, and 7 sized non-overlapping blocks. Mathematical explanations of the proposed MaTP method are defined in Eqs. 1-9.

$$T_1 = maxpool(Sound, 3) \tag{1}$$
$$L_1 = maxpool(Sound, 4) \tag{2}$$
$$F_1 = maxpool(Sound, 5) \tag{3}$$
$$S_1 = maxpool(Sound, 7) \tag{4}$$
$$L_{k+1} = maxpool(L_k, 4), k = \{1,2,3\} \tag{5}$$
$$T_{h+1} = maxpool(T_h, 3), h = \{1,2,3,4\} \tag{6}$$
$$F_{h+1} = maxpool(F_h, 5) \tag{7}$$
$$S_{h+1} = maxpool(S_h, 7) \tag{8}$$
$$maxpool(p_1, p_2) = \begin{cases} p_1, \ p_1 \geq p_2 \\ p_2, \ p_2 > p_1 \end{cases} \tag{9}$$

In Eqs. 1-9, the mathematical definition of the proposed MaTP based preprocessing method is shown. T_k, L_k, F_k and S_k are kth levelled of the maximum pooled sounds with 3, 4, 5, and 7 sized non-overlapping windows respectively. $maxpool(.,.)$ expresses maximum pooling function, p_1 and p_2 define first and second parameters.

By applying Eqs. 1-9, these pooled (decomposed) signals are calculated, and these decomposed signals and raw sound signal are utilized for feature generation. Details of the proposed feature generation are explained in Section 3.2.

2.2. FEATURE EXTRACTION

We used two feature extraction functions in the proposed feature generation network. These are statistical feature generation and the proposed Present-SBox-Pat based textural feature extraction. The main aim of this feature generation network is to use fused features to use the effectiveness of these functions together.

2.2.1. STATISTICAL FEATURE EXTRACTION

Our first feature extraction function used is statistical features. Statistical feature generation is one of the most useful feature generation methods. Therefore, they have been used for feature generation by researchers. We used 19 statistical moments as feature generation. These moments are shown in 10-28.

$$s1 = \frac{\sum_{i=1}^{L} sound_i}{L} \tag{10}$$

$$s2 = \sqrt{\frac{\sum_{i=1}^{L}(sound_i - s1)^2}{L}} \tag{11}$$

$$s3 = \sqrt{\frac{\sum_{i=1}^{L}(sound_i)^2}{L}} \tag{12}$$

$$s4 = \frac{\sum_{i=1}^{L} sound_i}{s3} \log(\frac{sound_i}{s3}) \tag{13}$$

$$s5 = \sum_{i=1}^{L} \frac{i * sound_i - s1}{\sigma_i} \tag{14}$$

$$s6 = \frac{\sum_{i=1}^{L} |sound_{i+1} - sound_i|}{L} \tag{15}$$

$$s7 = \frac{\sqrt{L(L-1)}}{L-2} \left(\frac{\frac{1}{L}\sum_{i=1}^{L}(sound_i - s1)^3}{\frac{1}{L}\sum_{i=1}^{L}(sound_i - s1)^2} \right) \tag{16}$$

$$s8 = \frac{L-1}{(L-2)(L-3)} \left[(L+1)\left(\left(\frac{\frac{1}{L}\sum_{i=1}^{L}(sound_i - s1)^4}{\frac{1}{L}\sum_{i=1}^{L}(sound_i - s1)^2} \right) - 3 \right) + 6 \right] \tag{17}$$

$$s9 = sound\left(\frac{L}{2}\right) \tag{18}$$
$$s10 = \{sound\} \tag{19}$$
$$s11 = \max\{sound\} \tag{20}$$
$$s12 = \frac{s1}{s2} \tag{21}$$

$$s13 = \sqrt{\frac{\sum_{i=1}^{L} sound_i^2}{L}} \tag{22}$$

$$s13 = \frac{\sqrt{\frac{1}{L}\sum_{i=1}^{L} sound_i^2}}{\frac{1}{L}\sum_{i=1}^{L} |sound_i|} \tag{23}$$

$$s15 = \frac{\max |sound|}{\sqrt{\frac{1}{L}\sum_{i=1}^{L} sound_i^2}} \tag{24}$$

$$s16 = \frac{\max |sound|}{\frac{1}{L}\sum_{i=1}^{L} \sqrt{|sound_i|^2}} \tag{25}$$

$$s17 = \frac{\max |sound|}{\frac{1}{L}\sum_{i=1}^{L} |sound_i|} \tag{26}$$

$$s18 = (\max(sound) - \min(sound)) \tag{27}$$
$$s19 = (\max(sound) - s1) \tag{28}$$

where $sound$ is the sound signal, L represents the length of the signal.

The procedure of the used statistical feature generation function is shown in Fig. 2.

Procedure: Statistical feature generation: $statistical(sound)$
Input: Sound signal or decomposed version of it with length of L ($signal$).
Output: Statistical features (sf) with length of 19
00: Load signal.
01: Apply Eqs. 10-18 to signal and obtain sf.

Figure 2. Statistical feature generation procedure.

2.2.2. THE PROPOSED PRESENT S-BOX BASED FEATURE EXTRACTOR

In this section, a novel textural feature extraction method is presented. This method includes information on both symmetric encryption and feature generation. This method aims to show nonlinear structures on the feature generation.

Therefore, the S-Box of the present encryption method is utilized as a pattern, and the signum function is employed as a binary feature generation method. The used present S-Box is shown in Fig. 3.

x	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
$S[x]$	C	5	6	B	9	0	A	D	3	E	F	8	4	7	1	2

Figure 3. S-Box of the present encryption method.

As seen in Fig. 3, the size of this S-Box is 16, and it has two values. These are the input value (x) and the output value ($S[x]$). Here, we used values indicated by the input values and output values in S-Box for bit generation. The proposed Present-SBox-Pat generates 16 bits in total. These bits are divided into two groups, and each group has 8-bits. These categories are named left and right, respectively. Two map values are generated by using these two categories. Histograms of these maps are concatenated, and feature vector with a size of 512 is obtained. Steps of the proposed Present-SBox-Pat are;

Step 1: Divide the signal into 16 sized overlapping windows.

$$b^i = signal(i: i + 15), i = \{1,2, \dots, L - 15\}$$

Step 2: Generate binary features from the block by using b^i and S-Box. The used bit generation function is signum.

$$bit(k) = Sg\left(b^i(k+1), b^i(S(k)+1)\right), k \quad (30)$$
$$= \{1,2, \dots, 16\}$$
$$Sg(p_1, p_2) = \begin{cases} 0, p_1 - p_2 < 0 \\ 1, p_1 - p_2 \geq 0 \end{cases} \quad (31)$$

where $Sg(.,.)$ expresses signum function.

Step 3: Calculate left and right map values.

$$left(i) = \sum_{k=1}^{8} bit(k) * 2^{k-1} \quad (32)$$

$$right(i) = \sum_{k=1}^{8} bit(k+8) * 2^{k-1} \quad (33)$$

Step 4: Extract histograms of the right and left map values.

$$histogram^{right}(j) = 0; j = \{1,2, \dots, 256\} \quad (34)$$
$$histogram^{left}(j) = 0; j = \{1,2, \dots, 256\} \quad (35)$$
$$histogram^{right}(right(i) + 1) \qquad (36)$$
$$= histogram^{right}(right(i) + 1) + 1$$
$$histogram^{left}(left(i) + 1) \qquad$$
$$= histogram^{left}(left(i) \qquad (37)$$
$$+ 1) + 1$$

Eqs. 34-37 define histogram extraction. Eqs. 34-35 express initial value assignment and Eqs. 36-37 define the histogram generation.

Step 5: Concatenate these histograms to obtain the final feature vector.

$$feat(j) = histogram^{right}(j) \qquad (38)$$
$$feat(j + 256) = histogram^{right}(j) \qquad (39)$$

where $feat$ is the final feature vector with the size of 512.

These steps (Step 1-5) define the Present-SBox-Pat feature generation procedure.

2.2.3. OVERVIEW OF THE PROPOSED FEATURE GENERATION NETWORK

This work presented a multilevelled and fused features-based feature generation network, as shown in Fig. 1. Our extracted features are divided into three types. These are textural features, statistical features, and statistical features of the textural features. The used statistical feature generation and textural feature generation functions extract 19 and 512 features, respectively. Fig. 4 summarized the feature generation process of this network.
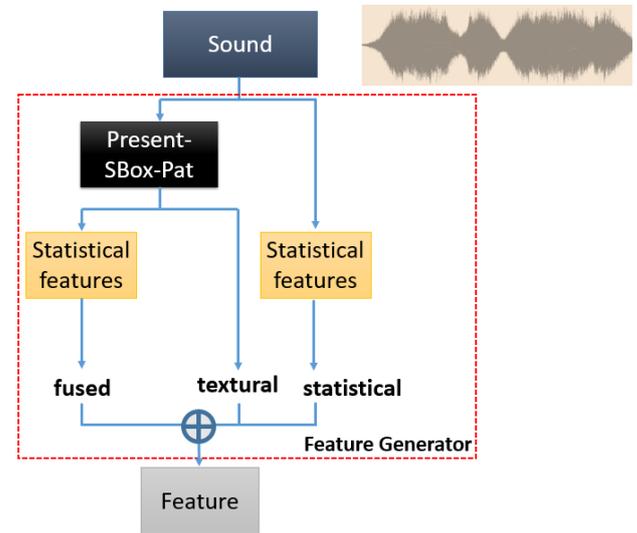
Figure 4. Our presented fused feature generation model.

Steps of our proposed feature generation network are;

Step 1: Extract statistical features of the raw sound signal and decomposed signals by using a statistical feature generation procedure, which is defined in Fig.4.

$$fs^1 = statistical(Sound) \qquad (40)$$
$$fs^k = statistical(L_{k-1}), k = \{2,3, \dots, 5\}, k \qquad (41)$$
$$= k + 4$$
$$fs^k = statistical(T_{k-5}), k = \{6,7, \dots, 10\}, k \qquad (42)$$
$$= k + 5$$
$$fs^k = statistical(F_{k-10}), k \qquad$$
$$= \{11,12, \dots, 15\}, k \qquad (43)$$
$$= k + 5$$
$$fs^k = statistical(S_{k-15}), k \qquad$$
$$= \{16,17, \dots, 20\} \qquad (44)$$

Eqs. 40-45 define statistical feature extraction and fs^k is k^{th} statistical feature vector with length of 19. 20

statistical feature vectors are obtained in this step. Therefore, 380 features are generated statistically in total.

Step 2: Extract textural features by using the proposed Present-SBox-Pat.

$$ft^1 = Present - SBox - Pat(Sound) \quad (45)$$

$$ft^k = Present - SBox - Pat(L_{k-1}), k = \{2,3,...,5\}, k = k + 4 \quad (46)$$

$$ft^k = Present - SBox - Pat(T_{k-5}), k = \{6,7,...,10\}, k = k + 5 \quad (47)$$

$$ft^k = Present - SBox - Pat(F_{k-10}), k = \{11,12,...,15\}, k = k + 5 \quad (48)$$

$$ft^k = Present - SBox - Pat(S_{k-15}), k = \{16,17,...,20\} \quad (49)$$

Eqs. 45-49 define textural feature extraction and ft^k is k^{th} textural feature vector with length of 512. *Twenty* textural feature vectors are obtained in this step. Therefore, 10240 features are generated texturally in total.

Step 3: Extract the statistical features of the textural features extracted.

$$ff^k = statistical(ft^k), k = \{1,2,...,20\} \quad (50)$$

Eq. 50 expresses statistical feature extraction of the textural features and ff^k is k^{th} fused feature vector with length of 19. *Twenty* fused feature vectors are obtained in this step. Therefore, 380 features are generated in total.

Step 4: Concatenate these features and obtain the final feature vector (X) with a size of 11,000.

$$X = concat(fs^k, ft^k, ff^k), k = \{1,2,...,20\} \quad (51)$$

In Eq. 51, $concat(.,.,.)$ expresses concatenation function

ReliefF and Iterative Neighbourhood Component Analysis feature selector: In this phase, a new generation 2-levelled feature selection method is presented, and this method is named as RFINCA. RFINCA uses the effectiveness of both ReliefF and NCA together [40]. ReliefF is used firstly. Since it generates both negative and positive weighted features, and negative weighted features express redundant features. Thus, to eliminate these redundant features, positive weighted feature selection is processed. In the second level of the RFINCA, the NCA feature selector is used. NCA is applied to selected positive weighted features. Then, the iterative NCA feature selection is processed. To calculated errors, each selected feature in this phase, the kNN classifier is chosen. The attributes of the used kNN classifiers are given as follows. k is selected one, and the distance metric is selected as Manhattan distance. The procedure of the RFINCA is shown in Fig. 5.

Figure 5. The procedure of the RFINCA feature selector.

RFINCA selects 289 the most informative features in this work.

Step 5: Select the most informative features by using RFINCA.

$$f^{final} = RFINCA(X) \quad (52)$$

where f^{final} is final feature.

2.3. CLASSIFICATION

Two shallow classifiers are used to obviously denote high capabilities of the Present-SBox-Pat and MaTP based fused features generation network and RFINCA feature selector. Our fused feature generation network generates 11.000 features, and RFINCA selects 289 the most valuable features from these 11.000 features. kNN and SVM classifiers are chosen to perform the classification process. The kNN used is also considered as an error value calculator. Attributes of the SVM used are given as follows. The kernel function is 3rd-degree polynomial, box constraint value (C) is chosen as one and coding type is selected as one-vs-all [41]. The validation and test strategy is employed as a 10-fold CV. The last step of the proposed Present-SBox-Pat and MaTP based fused features generation network, and RFINCA feature selector is classification, and this step is shown as below.

Step 6: Classify the selected 289 features (f^{final}) buy using kNN or SVM classifier.

3. EXPERIMENTAL RESULTS

3.1. COLLECTED DATASET

An experimental sound dataset is used in this work. This dataset uses YouTube breathing sounds videos, and these videos have been published for educational purposes. Therefore, there is no information about patients. This dataset was created by using 68 variable videos to prevent overfitting, and it includes 657 sounds with *ten* classes. The sampling rate of these sounds is 48 kHz. The characteristics of this experimental dataset are shown in Table 1. This dataset can be downloaded by using https://websiteyonetimi.ahievran.edu.tr/_Dosyalar/Genel/RespiratorySounds.rar URL. It is a public dataset.

TABLE I
Characteristics of the used breathing sound experimental dataset.

ID	Lung breathing sound type	Number of observations	Number of recordings
1	Vesicular	83	7
2	Fine Crackles (Rales)	57	7
3	Wheezing (expiratory)	70	8
4	Rhonchi	72	8
5	Stridor	63	7
6	Coarse Crackles (Rales)	82	7
7	Broncho vesicular	61	8
8	Bronchial	65	5
9	Coronavirus (COVID-19)	29	1
10	Healthy	75	10
	Total	657	68

3.2. EXPERIMENTAL SETUP

In Section 3, details of the proposed Present-SBox-Pat and MaTP based fused features generation network, and RFINCA feature selector-based sound classification method was given, and this method consists of preprocessing, feature generation, feature selection, and classification. To realize simulations of this method, MATLAB [42] programming environment was used on a system. This system has 32-GB main memory (RAM), Core i7-7700 (3.20 GHz) microprocessor (CPU). We did not use any graphics card components like GPU cores. MATLAB m files were utilized for coding MaTP preprocessing, feature generation, and feature reduction phases. MATLAB Classification Learner toolbox was considered to obtain results of the classifiers used, which are kNN and SVM. MATLAB source codes of kNN and SVM were generated to obtain more results, and these codes were related codes that were added to these m files.

3.3. RESULTS

Classification Accuracy (CA), Unweighted Average Recall (UAR), Unweighted Average Precision (UAP), F1-score ($F1$) and Geometrical Mean (GM) results were used to test the proposed Present-SBox-Pat and MaTP based fused features generation network and RFINCA feature selector based sound classification method comprehensively [43-45]. The number of true positives (tp), false positives (fp), true negatives (tn) and false negatives (fn) are employed to calculate these measurements as below.

$$CA = \frac{tp_j + tn_j}{tp_j + tn_j + fp_j + fn_j}, j = \{1,2,\dots,H\} \quad (53)$$

$$UAR = \frac{1}{H} \sum_{j=1}^{H} \frac{tp_j}{tp_i + fn_j} \quad (54)$$

$$UAP = \frac{1}{H} \sum_{j=1}^{H} \frac{tp_j}{tp_j + fp_j} \quad (55)$$

$$F1 = \frac{2 * UAP * UAR}{UAP + UAR} \quad (56)$$

$$GM = \sqrt[H]{\prod_{j=1}^{H} \frac{tp_j}{tp_j + tp_j}} \quad (57)$$

In Eqs. 53-57, H is number of classes.

The results of the used kNN and SVM classifiers were listed in Table 2.

TABLE II
Performances (%) of the used kNN and SVM classifiers.

Performance metric	kNN	SVM
CA	94.98	95.43
UAR	95.04	95.39
UAP	95.08	96.37
$F1$	95.06	95.86
GM	95.01	95.35

4. DISCUSSIONS AND CONCLUSION

Lung breathing sounds classification is one of the crucial topics for biomedical engineering and machine learning. Covid-19 is a new version of flu, and it causes lung disease. Nowadays, we have lived quarantine days of Covid-19. Therefore, proposing of automated Covid-19 disease detection methods is very important. Our main aim is to detect Covid-19 by using breathing sounds. However, there is no public breathing sound dataset includes Covid-19. Therefore, we collected an experimental breathing sound dataset. This dataset contains 657 breathing sound with *ten* classes (9 lung diseases and one healthy). An innovative hand-crafted features-based lung breathing sounds classification method for automated diagnosing lung diseases is presented in this work. We motivated to solve three problems, and these problems are to propose appropriate and effective preprocessing, feature generation, and feature selection methods. Therefore, a novel MaTP decomposition method is presented as preprocessing to improve the capability of the pooling-based decomposition methods. In the second phase, we combined structures of textural feature generation and symmetric encryption methods. S-Box of the present ultra-lightweight encryption method is utilized as a pattern for textural feature generation. Statistical features and statistical features of the extracted textural features were also used to obtain a comprehensive feature set. By using these methods, a novel fused feature generation network has been presented. RFINCA feature selector is employed to the extracted features, and it selected 289 most informative features. Two shallow classifiers were used to denote the classification capability of the extracted and selected features. 94.98% and 95.43% classification accuracies were achieved by

using kNN and SVM classifiers, respectively. Confusion matrices of these classifiers were also shown in Fig. 6 for apparent evaluation.

	1	2	3	4	5	6	7	8	9	10
1	78	1	1	2	0	0	0	1	0	0
2	0	53	1	0	0	2	1	0	0	0
3	0	4	64	0	0	1	0	0	0	1
4	0	1	1	67	0	1	1	0	0	1
5	0	0	0	1	61	0	1	0	0	0
6	0	3	0	0	0	79	0	0	0	0
7	0	0	0	0	0	0	59	1	1	0
8	3	0	0	0	0	0	0	61	0	0
9	0	0	0	1	0	0	0	0	28	0
10	0	0	0	1	0	0	0	0	0	74
	1	2	3	4	5	6	7	8	9	10

True Class (vertical) / Predicted Class (horizontal)

(a) kNN

	1	2	3	4	5	6	7	8	9	10
1	80	1	1	0	0	0	0	1	0	0
2	1	51	0	0	0	3	2	0	0	0
3	4	1	65	0	0	0	0	0	0	0
4	2	0	0	69	0	1	0	0	0	0
5	1	0	0	0	62	0	0	0	0	0
6	4	0	0	0	0	78	0	0	0	0
7	1	1	0	0	0	0	59	0	0	0
8	3	0	1	0	0	0	0	61	0	0
9	1	0	0	0	0	0	0	0	28	0
10	1	0	0	0	0	0	0	0	0	74
	1	2	3	4	5	6	7	8	9	10

True Class (vertical) / Predicted Class (horizontal)

(b) SVM

Figure 6. The calculated confusion matrices (a) kNN (b) SVM.

Fig. 6 evidently denotes the success rates of each class. The 9th class is Covid-19, and both classifiers achieved 96.55% classification accuracy for Covid-19 detection. Sharma et all [46] also used breathing and cough sounds for the diagnosis of Covid-19 and achieved 66.74% accuracy rate. On the other hand, Mugili et al [47] obtained the highest 75.17% AUC value from the speech voice by using different methods. In another study, detection was made from the sound of breathing with an accuracy rate of 80% [48]. In another article that tried to detect the presence of Covid-19 from respiratory sounds in binary, 80% AUC was obtained [49]. A cumulative accuracy of 99.8% was obtained using cough sounds collected from patients in Pakistan [50]. In this study, it is thought that only binary classification has an effect on obtaining a high success rate. In this study, 10 different diseases are detected. In another study conducted by analyzing respiratory sounds obtained over the phone, it was tried to determine whether there was Covid-19 by binary classification and 94.96% specificity was obtained [51]. Although this study has obtained values close to our study, it is thought the fact that it has made a binary classification is an issue that increases the success.

Overall, significant benefits of the proposed Present-SBox-Pat and MaTP based fused features generation network and RFINCA feature selector based sound classification method are;

- A novel experimental lung disease dataset was collected and published publicly.
- This work denoted the success of Covid-19 detection by using breathing sounds.
- A novel S-Box based feature generator was presented, and the proposed feature generator is shown the effectiveness of the S-Box for textural feature generation.
- A new maximum pooling-based decomposition method (MaTP) was proposed to extract high, medium, and low levels features.
- The proposed feature generation network used the effectiveness of both textural features and statistical features.
- The optimal distinguished feature selection problem was solved by using RFINCA.
- This method achieved high classification accuracies by using two conventional classifiers (see Table 2).
- The proposed method is a general ML method for one-dimensional data. Other one-dimensional data classification problems can be solved by using this method.
- The proposed lung sound classification method can run on a basic system with straightforward configurations.
- An intelligent lung disease monitoring system or stethoscope can be designed by using this method.
Overall, the disadvantages of this work are;
- The proposed Present-SBox-Pat and MaTP based fused features generation network and RFINCA feature selector-based sound classification method was tested on an experimental dataset. It can be tested on real, and more significant lung sounds datasets when they publicly published.
- The time complexity of RFINCA is high. Therefore, we used kNN (it is one of the fastest classifiers) as a loss calculator. Novel and more effective feature selectors can be used to select optimal number of features.

REFERENCES

[1] C. B. Kaya and K. Ebubekir, "A Novel Approach Based to Neural Network and Flower Pollination Algorithm to Predict Number of COVID-19 Cases," *Balkan Journal of Electrical and Computer Engineering,* vol. 9, no. 4, pp. 327-336, 2021.

[2] L. GÖKREM and M. ALTINDAŞ, "ENVIRONMENTAL NOISE TRACKING SYSTEM BASED ON WEB OF THINGS," *European Journal of Technique (EJT),* vol. 9, no. 2, pp. 175-185, 2019.

[3] M. Aykanat, Ö. Kılıç, B. Kurt, and S. Saryal, "Classification of lung sounds using convolutional neural networks," *EURASIP Journal on Image and Video Processing,* vol. 2017, no. 1, p. 65, 2017.

[4] U. Ozkaya, S. Ozturk, and M. Barstugan, "Coronavirus (COVID-19) Classification using Deep Features Fusion and Ranking Technique," *arXiv preprint arXiv:2004.03698,* 2020.

[5] C. Huang *et al.*, "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *The lancet,* vol. 395, no. 10223, pp. 497-506, 2020.

[6] L. Lan *et al.*, "Positive RT-PCR test results in patients recovered from COVID-19," *Jama,* vol. 323, no. 15, pp. 1502-1503, 2020.

[7] T. Ai *et al.*, "Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases," *Radiology,* p. 200642, 2020.

[8] H. Pasterkamp, S. S. Kraman, and G. R. Wodicka, "Respiratory sounds: advances beyond the stethoscope," *American journal of respiratory and critical care medicine,* vol. 156, no. 3, pp. 974-987, 1997.

[9]   J. Zhang, Y. Xie, Y. Li, C. Shen, and Y. Xia, "Covid-19 screening on chest x-ray images using deep learning based anomaly detection," *arXiv preprint arXiv:2003.12338,* 2020.

[10]   S. Amiriparian *et al.*, "CAST a database: Rapid targeted large-scale big data acquisition via small-world modelling of social media platforms," in *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2017: IEEE, pp. 340-345.

[11]   M. Li *et al.*, "Coronavirus disease (COVID-19): spectrum of CT findings and temporal progression of the disease," *Academic radiology,* 2020.

[12]   A. Alimadadi, S. Aryal, I. Manandhar, P. B. Munroe, B. Joe, and X. Cheng, "Artificial intelligence and machine learning to fight COVID-19," ed: American Physiological Society Bethesda, MD, 2020.

[13]   M. Barstugan, U. Ozkaya, and S. Ozturk, "Coronavirus (covid-19) classification using ct images by machine learning methods," *arXiv preprint arXiv:2003.09424,* 2020.

[14]   Z. Dokur, "Respiratory sound classification by using an incremental supervised neural network," *Pattern Analysis and Applications,* vol. 12, no. 4, p. 309, 2009.

[15]   R. X. A. Pramono, S. Bowyer, and E. Rodriguez-Villegas, "Automatic adventitious respiratory sound analysis: A systematic review," *PloS one,* vol. 12, no. 5, 2017.

[16]   J. Schröder, J. Anemiiller, and S. Goetze, "Classification of human cough signals using spectro-temporal Gabor filterbank features," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016: IEEE, pp. 6455-6459.

[17]   Z. Moussavi, A. Yadollahi, and S. Camorlinga, "Breathing sound analysis for detection of sleep apnea/popnea events," ed: Google Patents, 2009.

[18]   M. S. Doyle, "Analysis of lung sounds using neural networks," Vanderbilt University, 1994.

[19]   B. Sankur, Y. P. Kahya, E. Ç. Güler, and T. Engin, "Comparison of AR-based algorithms for respiratory sounds classification," *Computers in Biology and Medicine,* vol. 24, no. 1, pp. 67-76, 1994.

[20]   B. Lei, S. A. Rahman, and I. Song, "Content-based classification of breath sound with enhanced features," *Neurocomputing,* vol. 141, pp. 139-147, 2014.

[21]   M. Shokrollahi, S. Saha, P. Hadi, F. Rudzicz, and A. Yadollahi, "Snoring sound classification from respiratory signal," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016: IEEE, pp. 3215-3218.

[22]   S. Huq and Z. Moussavi, "Acoustic breath-phase detection using tracheal breath sounds," *Medical & biological engineering & computing,* vol. 50, no. 3, pp. 297-308, 2012.

[23]   J. Han *et al.*, "An Early Study on Intelligent Analysis of Speech under COVID-19: Severity, Sleep Quality, Fatigue, and Anxiety," *arXiv preprint arXiv:2005.00096,* 2020.

[24]   Z. Jiang *et al.*, "Combining Visible Light and Infrared Imaging for Efficient Detection of Respiratory Infections such as COVID-19 on Portable Device," *arXiv preprint arXiv:2004.06912,* 2020.

[25]   H. S. Maghdid, K. Z. Ghafoor, A. S. Sadiq, K. Curran, and K. Rabie, "A novel AI-enabled framework to diagnose Coronavirus COVID 19 using smartphone embedded sensors: Design study," *arXiv preprint arXiv:2003.07434,* 2020.

[26]   A. A. Ardakani, A. R. Kanafi, U. R. Acharya, N. Khadem, and A. Mohammadi, "Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks," *Computers in Biology and Medicine,* p. 103795, 2020.

[27]   K. Li *et al.*, "CT image visual quantitative evaluation and clinical classification of coronavirus disease (COVID-19)," *European Radiology,* pp. 1-10, 2020.

[28]   F. Shi *et al.*, "Large-scale screening of covid-19 from community acquired pneumonia using infection size-aware classification," *arXiv preprint arXiv:2003.09860,* 2020.

[29]   A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks," *arXiv preprint arXiv:2003.10849,* 2020.

[30]   T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE transactions on pattern analysis and machine intelligence,* vol. 28, no. 12, pp. 2037-2041, 2006.

[31]   M. Elangovan, N. Sakthivel, S. Saravanamurugan, B. B. Nair, and V. Sugumaran, "Machine learning approach to the prediction of surface roughness using statistical features of vibration signal acquired in turning," *Procedia Computer Science,* vol. 50, pp. 282-288, 2015.

[32]   S. A. Dudani, "The distance-weighted k-nearest-neighbor rule," *IEEE Transactions on Systems, Man, and Cybernetics,* no. 4, pp. 325-327, 1976.

[33]   J. A. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural processing letters,* vol. 9, no. 3, pp. 293-300, 1999.

[34]   YouTube, 11.05.2020. [Online]. Available: www.youtube.com.

[35]   D. Scherer, A. Müller, and S. Behnke, "Evaluation of pooling operations in convolutional architectures for object recognition," in *International conference on artificial neural networks*, 2010: Springer, pp. 92-101.

[36]   I. Belakhdar, W. Kaaniche, R. Djemal, and B. Ouni, "Single-channel-based automatic drowsiness detection architecture with a reduced number of EEG features," *Microprocessors and Microsystems,* vol. 58, pp. 13-23, 2018.

[37]   C. Qin, S. Song, G. Huang, and L. Zhu, "Unsupervised neighborhood component analysis for clustering," *Neurocomputing,* vol. 168, pp. 609-617, 2015.

[38]   A. Bogdanov *et al.*, "PRESENT: An ultra-lightweight block cipher," in *International Workshop on Cryptographic Hardware and Embedded Systems*, 2007: Springer, pp. 450-466.

[39]   V. Nandan and R. G. S. Rao, "Minimization of digital logic gates and ultra-low power AES encryption core in 180CMOS technology," *Microprocessors and Microsystems,* vol. 74, p. 103000, 2020.

[40]   T. Tuncer and F. Ertam, "Neighborhood component analysis and reliefF based survival recognition methods for Hepatocellular carcinoma," *Physica A: Statistical Mechanics and its Applications,* vol. 540, p. 123143, 2020.

[41]   U. Jain, K. Nathani, N. Ruban, A. N. J. Raj, Z. Zhuang, and V. G. Mahesh, "Cubic SVM Classifier Based Feature Extraction and Emotion Detection from Speech Signals," in *2018 International Conference on Sensor Networks and Signal Processing (SNSP)*, 2018: IEEE, pp. 386-391.

[42]   D. J. Higham and N. J. Higham, *MATLAB guide*. SIAM, 2016.

[43]   A. Rosenberg, "Classifying skewed data: Importance weighting to optimize average recall," in *Thirteenth Annual Conference of the International Speech Communication Association*, 2012.

[44]   T. Tuncer, S. Dogan, P. Pławiak, and U. R. Acharya, "Automated arrhythmia detection using novel hexadecimal local pattern and multilevel wavelet transform with ECG signals," *Knowledge-Based Systems,* vol. 186, p. 104923, 2019.

[45]   S. D. Kumar, S. Esakkirajan, S. Bama, and B. Keerthiveena, "A Microcontroller based Machine Vision Approach for Tomato Grading and Sorting using SVM Classifier," *Microprocessors and Microsystems,* p. 103090, 2020.

[46]   N. Sharma *et al.*, "Coswara--A Database of Breathing, Cough, and Voice Sounds for COVID-19 Diagnosis," *arXiv preprint arXiv:2005.10548,* 2020.

[47]   A. Muguli *et al.*, "DiCOVA Challenge: Dataset, task, and baseline system for COVID-19 diagnosis using acoustics," *arXiv preprint arXiv:2103.09148,* 2021.

[48]   U. Sait *et al.*, "A deep-learning based multimodal system for Covid-19 diagnosis using breathing sounds and chest X-ray images," *Applied Soft Computing,* p. 107522, 2021.

[49]   C. Brown *et al.*, "Exploring automatic diagnosis of COVID-19 from crowdsourced respiratory sound data," *arXiv preprint arXiv:2006.05919,* 2020.

[50]   M. U. Khan, S. Ali, K. Habib, H. Khan, F. Tariq, and S. Bibi, "A Novel Intelligent Model For COVID-19 Detection Using Cough Auscultations and Hjorth Descriptors," in *ICAME21, International Conference on Advances in Mechanical Engineering, Pakistan*, 2021, pp. 1-10.

[51]   M. Alkhodari and A. Khandoker, "Detection of COVID-19 in smartphone-based breathing recordings using CNN-BiLSTM: a pre-screening deep learning tool," *medRxiv,* 2021.

BIOGRAPHIES

Turker Tuncer, was born in Elazig, Turkey in 1986. He received the B.S. degree from the Firat University, Technical Education Faculty, Department of Electronics and Computer Education in 2009, M.S. degree in telecommunication science from the Firat University in 2011 and Ph.D. degree department of software engineering at Firat University in 2016. He works as research assistant Digital Forensic Engineering, Firat University. His research interests include data hiding, image authentication,

cryptanalysis, cryptography, feature extraction, machine learning and biomedical engineering.

Erhan AKBAL is currently working as an assistant professor at Digital Forensics Engineering Department of Firat University, He received his Ph.D. degree in electrical and electronics engineering in 2012, the M.S. degree in computer engineering in 2007, from Firat University, Turkey. His research interests include computer network security, wireless sensor network, intrusion detection and digital forensics.

Emrah Aydemir received the M.S. degrees in computer teaching from the University of Elazig Firat, in 2012 and the Ph.D. degree in informatics from Istanbul University, Turkey, TR, in 2017. From 2012 to 2015, he was an Expert with the Istanbul Commerce University. Since 2017, he has been an Associate Professor with the Sakarya University. He is the author of three books, more than 10 articles, and more than 40 conference presentation. His research interests include artificial intelligence, microcontroller, database and software.

Samir Brahim Belhaouari received the master's degree in telecommunications and network from the Institute National Polytechnique of Toulouse, France, in 2000, and the Ph.D. degree in mathematics from the Federal Polytechnic School of Lausanne-Switzerland, in 2006. He is currently an Associate Professor with the University of Hamad Bin Khalifa, Qatar foundation, in the Division of Information and Communication Technologies, College of Science and Engineering, HBKU. During last years, he also holds several Research and Teaching positions at Innopolis University-Russia, Alfaisal University-Saudi Arabia, the University of Sharjah-UAE, University Technology PETRONAS- Malaysia, and the EPFL Federal Swiss School-Switzerland. His research interests include vary areas from applied mathematics, statistics, and data analysis until artificial intelligence, image and signal processing (biomedical, bioinformatics, forecasting, etc.), due to both Mathematics and Computer Science backgrounds.

Sengul Dogan received the master's degree in bioengineering the Ph.D. degree in electrical and electronics engineering from the Firat University, Elazig, Turkey, in 2007 and 2011, respectively. She is currently an Associate Professor with the Digital Forensics Engineering, Technology Faculty, Firat University. Her main research interests include computer forensics, mobile forensics, image processing, and signal processing. She has been working actively on developing algorithms in machine learning for biomedical data.