# Improved and Secured Electromyography in the Internet of Health Things

Muhammad Usman, *Member, IEEE*, Mohsin Kamal, *Senior Member, IEEE*, and Muhammad Tariq, *Senior Member, IEEE*

***Abstract*—Physiological signals are of great importance for clinical analysis but are prone to diverse interferences. To enable practical applications, biosignal quality issues, especially contaminants, need to be dealt with automated processes. For example, after processing surface electromyography (sEMG), fatigue analysis can be done by looking into muscle contraction and expansion for clinical diagnosis. Contaminants can make this diagnosis difficult for the clinician. In real scenarios, there is a possibility of the presence of multiple contaminants in a biosignal. However, most of the work done until now focuses on the presence of a single contaminant at a time. This paper proposes a new method for the identification and classification of contaminants in sEMG signals where multiple contaminants are present simultaneously. We train a 1D convolutional neural network (1D-CNN) to classify different contaminant types in sEMG signals without prior feature extraction. The network is trained on simulated and real sEMG signals to identify five types of contaminants. Additionally, we train and test 1D-CNN to identify multiple contaminants when present simultaneously. Furthermore, to securely and accurately transfer the data to the clinician, we also present experimental results to securely route the data in a proposed Internet of health things (IoHT) by using received signal strength indicators (RSSI) to generate link fingerprints (LFs). The results show higher accuracy of the classification system at low signal-to-noise ratios (SNR) and witness lightweight security of the IoHT.**

***Index Terms*—Light weight security, electromyography, biomedical signal processing, Internet of Health Things (IoHT).**

## I. Introduction

DUE to the rapidly increasing population around the world, biosignal quality analysis has been given immense importance in industry as well as in academia. The significance of biosignal quality analysis is rising with the development of remote monitoring applications, which are essential for the comprehensive tracking of routine activities. This analysis ensures adequate signal quality for future processing by providing a link between acquisition and processing.

Manuscript received March 12, 2021; revised August 19, 2021 and September 13, 2021; accepted September 26, 2021. Date of publication October 8, 2021; date of current version May 5, 2022. *(Corresponding author: Muhammad Tariq.)*

The authors are with the Department of Electrical Engineering, National University of Computer and Emerging Sciences, Islamabad 44000, Pakistan (e-mail: usman.abbasi@nu.edu.pk; mohsin.kamal@nu.edu.pk; mtariq@princeton.edu).

Digital Object Identifier 10.1109/JBHI.2021.3118810

The analysis of biosignals include electromyography (EMG), electroencephalography (EEG), and electrocardiography (ECG). Quality analysis of physiological signals is typically done by semi-automated processes where an operator visually rejects contaminated data [1]. As the number of remote applications increase, sifting the data manually becomes impractical as it is more prone to human errors and also time-consuming. Therefore, automated applications and processes are required for biosignal quality analysis to deal with the growing quantity of data. Big data, cloud computing, Internet of Things (IoT) and multi-core high-speed microprocessors have also paved a way in making automated biosignal quality a reality. In the scenario of smart health monitoring, several false alarms can occur. In literature, methods are there to tackle single contamination of physiological signals while work on multiple contaminants still needs to be done [2].

There are multiple noise sources that can contaminate EMG signals during acquisition and are not recognized by the naked eye. These noises interfere with the quality of recorded signals and hence lead to the wrong interpretation of data. The five main contaminants discussed in our work are: 1) Power Line Interference; caused by electromagnetic radiations originating from different electrical devices, fluorescent lights and power lines (50 or 60 Hz). 2) Motion Artifact; It is produced by the movement of the sensor due to the voluntary or involuntary movement of the patient. It may occur due to the direct impact on the body or the sensor and with the stretching or shortening of muscles that causes a chemical imbalance of skin electrode. 3) Cross-Talk; This contamination occurs from the signal propagation of neighboring muscles. 4) Clipping; It is a form of distortion caused by the saturation in the amplitude of the EMG signal or due to sensor detachment. 5) Physiological Noise; The unwanted signals from other organs and tissues. In some cases, ECG also adds up to the EMG signal. In this research work, the real EMG signals and real EMG contaminated signals have been taken for analysis.

Secure transmission of data from the IoT enabled Internet of health things (IoHT) devices is a big challenge in implementing the whole framework. The issues involve the tampering of data at the device level and sending false information to the cloud. Another case concerns the eavesdropping in which the data is fetched by the intruder. The aim should be to keep the trust of the user and the remote monitoring healthcare staff's intact. Furthermore, the latency in the IoT further degrades the quality of service, which leads to the requirement of higher bandwidths

and the use of 5 G, specifically the millimeter wave (mmWave). High data rate and large bandwidth are the key offerings of 5 G. The study we present in this paper challenges the support vector machine (SVM) method by using a 1D convolutional neural network (1D-CNN). As pointed out in [3], 1D-CNN has become the state-of-the-art technique for critical signal processing applications such as biomedical signal classification, structural health monitoring, anomaly or motor-fault detection. The CNNs combine feature extraction and classification, and they are capable of learning the optimal features and classifier parameters based solely on raw data. Compact 1D-CNNs can be efficiently trained with a limited dataset of 1D signals, and due to their low computational requirements, they are well-suited for real-time and low-cost applications [4].

In this paper, we use 1D-CNN to classify contaminants without first doing feature engineering. One of this research's goal is to find out if the feature set from the SVM method is sufficient for accurate contaminant identification. We propose to use 1D-CNN that is trained on real and simulated EMG signals. Furthermore, the work in [5] and [2] does not cover real-life situations when multiple contaminants of the ECG signal are present at once. In this work, we train and test our model to distinguish contaminants that are simultaneously present in the signal.

The contributions of this paper can be summarized as follows. First, we have introduced 1D-CNN for identification of contaminants in EMG signals. Second, we have solved the identification problem when multiple contaminants are present simultaneously. And third, we present lightweight security algorithms using which the data could be transferred to the clinician over secured IoHT to keep the trust intact.

## II. Related Work

Recent advances in different rehabilitation applications have increased the importance of EMG signal analysis. High-density surface electromyography (HD-EMG) gives detailed information about muscle activity. A large number of channels can be obtained, providing an insight into a broad muscle region. This information helps in identifying motor tasks [7] and allows us to understand complete muscle anatomy [8]. HD-EMG recordings enable us to extract the motor unit (MU) activities non-invasively [9]. However, HD-EMG recordings are prone to contaminants like motion artifacts (MA) and power line interference (PL). Different methods like filtering, feature extraction etc. have been proposed and developed for the de-noising of EMG signals. An effective de-noising tool has been provided by blind source separation techniques, including independent component analysis (ICA) and canonical component analysis (CCA). In [10], the CCA method was proposed to reduce additive white Gaussian noise (AWGN). Moreover, some studies also used blind source separation techniques by considering noise elements zero [10]. The drawbacks of this approach were that the noisy component may still contain some of the EMG signals and setting the noisy part to zero leads to the formation of a singular covariance matrix which is not useful in EMG decomposition by ICA. MA and PL were detected by the spectrum and removed by the notch filter. This technique seemed to be valid for the signals with a high signal-to-noise ratio (SNR). However, the authors did not address other interferences like physiological cross-talk, instrumental artifacts etc. and signals with low SNR.

Various ways have also been presented, targeting a particular contaminant in EMG recordings. These include: 1) quantization noise [11]; ECG interference [12]; analog-to-digital conversion clipping [11]; PL [13]; and amplifier saturation [11]. In myoelectric applications like pattern recognition, macro features (root mean square (RMS) and median frequency (MDF)) are extracted from EMG recordings [14].

In [5] and [2], the authors presented a method based on feature extraction and SVM classifier referred as classical SVM method. They have addressed the detection of different contaminants present in an EMG signal. Seven features are extracted: SMR, signal-to-power-line ratio, SNR, signal-to-ECG ratio (SER), maximum drop in power density (DPR), power spectrum deformation ($\Omega$-ratio), and Pearson correlation coefficient test for normality. The SVM classifier is trained for low SNR values at which contaminants could be fairly distinguished. The SVM classifier was trained on different features to produce an absolute result and is capable of distinguishing single contaminant only. The SVM method is a classical machine learning (ML) method that encompasses feature extraction and feature classification. However, the computational effort of the feature extraction can prevent its use in real-time applications. Also, it is not clear whether the chosen features completely characterize contaminants and if the classification performance could be improved or efficiently solved if other features are selected. Our previous work used SVM and multi-class SVM for the detection and classification of ECG signal under the presence of multiple contaminants simultaneously. However, it misses the security element for IoHT applications and important contaminants like AS and AWGN [15], [16].

The authors in [17] have proposed a routing algorithm for unmanned aircraft systems (UAS) in which the security is ensured by considering blockchain based model. The proof-of-traffic (PoT) is constructed for the for swarm UAS networking which reduces the the energy consumption. In [18], the authors have used the physical-layer channel information to prevent any adversarial activity in the IoT network. The deep-learning (DL) algorithms are used to identify the IoT devices in scalable IoT. The proposed work in [19] presents incremental learning (IL) for non-cryptographic device identification in which various computational analysis are presented. Furthermore, in [20], the authors have proposed a blockchain based method to overcome the threats while accessing the adversarial base station (BS). The proposed algorithm detects the adversarial BS and eliminates it from the list of all the available BSs which helps the device to not use these BSs while carrying out various activities in the network.

## III. Identification System

Fig. 1 shows the system used for the identification of contaminants in surface EMG signals. Recorded EMG signals are carried to the CNN. The network is trained to classify the

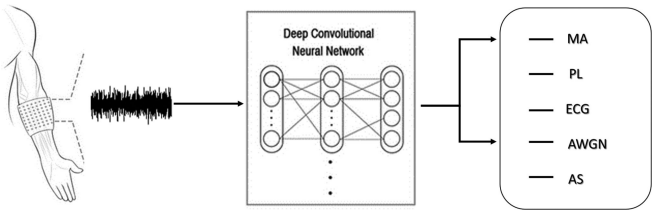


Fig. 1. Classification of EMG contaminants based on deep neural network (DNN) [5], [6].

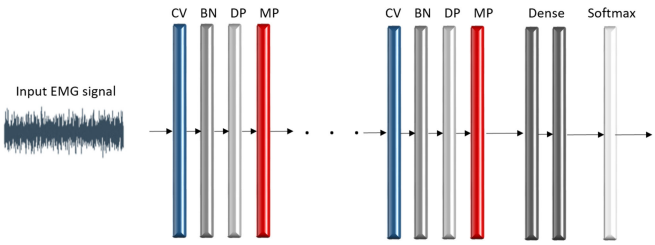


Fig. 2. 1D convolutional neural network model.

contaminants in five different types: MA, PL, ECG signal, AWGN and amplitude saturation (AS).

### A. 1D Convolutional Neural Network

A general 1D-CNN that will operate on raw EMG signal is shown in Fig. 2. The system consists of a series of layers: convolutional (CL), batch normalization (BN), dropout (DP), max pooling (MP), dense layers and a softmax layer. Feature extraction is performed by a series of convolutional layers that execute a linear convolutional operation on the signal followed by a non-linear activation function. Convolution involves a series of multiplications of 1D input and a 1D array of weights called kernel or filter. The filter is applied multiple times to the input, resulting in a 1D array of output values called a feature map. Typical non-linear activation function used is a rectifier unit (ReLU). MP is accompanying convolutional layers to reduce the dimensions of the problem by preserving signal features, invariant to transformations of the input, such as signal translation. To prevent the network from over-fitting to the dataset, networks are usually supplemented by layers for regularization. DP is a computationally inexpensive technique that offers regularization by randomly dropping out (i.e., ignoring) some network nodes during training. BN normalizes activation of the previous layer to achieve zero mean and unit variance of the features. These layers help speed up learning and stability of the network and also have a regularization effect [21]. The classification part is done by dense layers followed by softmax function for each of the five classes/contaminants.

## IV. Lightweight Security for IoHT

The implementation of 5 G has opened many doors in IoHT. Higher data rates and low latency are achieved due to which the seamless communication is made possible [22]. Securing the network of IoHT is a challenge. Traditional cryptographic solutions consume more energy and require devices with extensively large processing capabilities [23]. The medical device is in a static position and sends data to the server. The server is further connected to the cloud using 5 G back-haul network. The device is connected to another static node which is considered as anchor node which computes the received signal strength indicators (RSSI) and sends the information to the server. Any medical device in the IoHT will be connected to the anchor node and sends RSSI to the server. The server is considered as highly secured and has high computational capabilities. RSSI values of any connected devices have a linear relationship in terms of variations.

### A. Threat Model and Risk Assessment

Man in the middle attack: When any adversary changes the path of transmission and comes in between the communication path of connected IoHT device and anchor node devices, the adversary is able to listen to the data being transmitted from one device to another. This is an issue for the breach in privacy. The adversary has now access of all the important information of patient which is generated by medical devices.

Tampering of IoHT devices: The adversary is able to forge the data of IoHT devices which may cause a great impact on the patient's health situation. The adversary can inject wrong data to be transferred to the server. This can have severe implications as the relevant authorities would consider the forged data as correct.

Replay attack: The adversary after knowing the data pattern repeats the same data again and again. The system considers the data to be correct because the data can easily be decoded and the message would make sense. But because of this kind of attack, many important information is neglected/missed which results in wrong data being sent to the server at all the times.

### B. Methodology for Securing IoHT

The RSSI values which are acquired by MICAz motes in decimal are converted to binary which is termed as link fingerprint ($LF$) and is encoded with a private key ($K$) which is unique for the device and a pseudo-random nonce ($N$) which helps in getting different values after encoding each time.

$$LF_{encoded(1\rightarrow n)} = LF_{1\rightarrow n} \oplus K_i \oplus N_i. \tag{1}$$

$LF_{encoded}$ and $N$ are sent to the server by both IoHT device and the anchor node. In order to extract the link fingerprint, the adversary requires to have the information of both private key and Nonce. The nonce keeps on changing at every instant of time which is system generated and is synchronized at both the medical devices and server.

The encoded $LF$ and the $N$ are decoded at the server. The private keys associated with each IoHT device and server is used to perform the decoding operation.

$$LF_{1\rightarrow n} = K_i \oplus LF_{encoded(1\rightarrow n)} \oplus N_i. \tag{2}$$

Next step is to convert the $LF$ to decimal values in dBm. The RSSI variations are compared with the defined threshold.

Another threshold $P_{r(avg)}$ can be defined as:

$$P_{r(avg)} = \left| \frac{P_{r(i)} - P_{r(i+1)}}{2} \right|, \tag{3}$$

In (3), the consecutive RSSI values are represented as $P_{r(i)}$ and $P_{r(i+1)}$ for IoHT devices. The difference between RSSI values of IoHT device and anchor node is computed as $P_{r(IoHT)} - P_{r(server)}$, which defines the synchronized difference $(diff)$ of RSSI values of IoHT device and anchor node. Equation 4 describes the mathematical representation of the average difference formula of consecutive difference values. The result is checked against the predefined threshold.

$$diff_{avg} = \left| \frac{diff_{(i)} - diff_{(i+1)}}{2} \right|. \tag{4}$$

The detection parameters of any adversary in the IoHT are:

1) the RSSI values go higher or get lower for a large period of time,
2) the $P_{r(avg)}$ values are not 0 or 1,
3) non-constant values of $diff$,
4) variations in $diff_{avg}$ are observed.

## V. Results and Discussions

### A. Experimental Setup

Experiments are performed for securing IoHT network using MICAz motes which are Zigbee enabled devices. MICAz motes are used as IoHT device, anchor node and adversarial node to carry out experiments which record RSSI values after every 60 seconds. The RSSI logs are generated on a Linux based system which is further simulated using MATLAB to validate the results.

### B. Simulations

We train 1D-CNN on dedicated DL workstation. The workstation for training uses NVIDIA GeForce $GTX-2070$ super graphics card and 8-core AMD processor Ryzen-7(3800X). As a software tool, we use MATLAB with DL Toolbox. We simulate EMG signals by passing white Gaussian noise through a band-pass filter as in [2], [5]. Length of the signal is set to 15000 samples and a sampling frequency of the filter to 3 kHz. For training, we use only EMG at low SNR with respect to the contaminants: $-20$ dB, $-15$ dB, $-10$ dB and $-5$ dB. The real EMG signals used in this paper are the same data used in [2], [5].

The data is recorded on five subjects from three muscles: 1) biceps brachii; 2) quadriceps femoris (rectus femoris); and 3) tibialis anterior. The signals are sampled at 3000 Hz. After processing and manual selection, there are a total of 266 EMG recordings of 5 s duration. After filtration we end up with 170 signals. The reader is referred to [2], [5] for more details on the recording procedure. Test and training data are split 20% and 80%, respectively. For training, we use Adam optimizer. The mini-batch size is set to 256, and the initial learning rate is set to 0.0001. In Fig. 3, results are shown for the classification of five different contaminants as a function of the SNR. Simulation parameters are mentioned in Table I. As it is trained for lower SNR only, the classification is accurate for lower SNR values. In Fig. 5(a), the results of the classification of single contaminant as a function of the SNR for simulated EMG signals are shown. The network is trained for negative SNR values or high contamination cases for which it correctly classifies contaminants. Moreover, the network was able to generalize its behaviour for small positive SNR values. Ideal accuracy is still present for all contaminants at 5 dB, for AS at 10 dB, and PL for all tested SNR values. At SNR 20 dB, contaminant values are low, and it is not possible to distinguish among them.

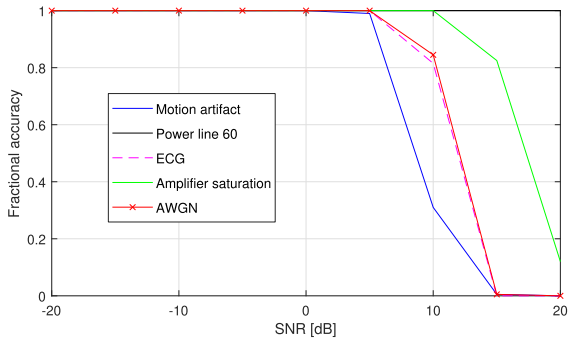


Fig. 3. Accuracy of classification of 1D-CNN identification method when single contaminant (MA, PL, ECG, AS, or AWGN) is present in the EMG signal.

TABLE I
Simulation Parameters

| EMG signal parameters | | 1D-CNN parameters | |
|---|---|---|---|
| length | 15000 [samples] | optimizer | Adam |
| sampling frequenccey | 3 KHz | threshold | 0.65 (simulated) 0.60 (real) |
| number of signals | 1000 (simulated) 170*64 (real) | initial learning rate | 0.0001 |
| - | - | batch size | 256 |

Since the number of real EMG signals is only 170, it gives small dataset for training a neural network. Such small dataset would cause over-fitting of the network and would give lower classification accuracy. In order to increase the size of the dataset for the training, two techniques are used for data augmentation: over-sampling and windowing method. Real EMG signals are first over-sampled by some factor $R$. New signals are created by shifting the over-sampled signal by 1 to $R$ number of samples and then down-sampled. The over-sampling factor $R$ is set to 64. Besides, contaminated real EMG signals are divided into 10 slices and training and classification are done on the slice level. A majority voting is performed to decide a predicted label [24].

The results of the classification for the five different contaminants as a function of the SNR for real EMG signals are shown in Fig. 5(b). The deportment of AS and PL are similar to the ideal case. Good accuracy for AWGN is preserved up to 0 dB SNR for which it is trained. Accuracy of classification of low-frequency interference (ECG and MA) is degraded (0.85 at 0 dB). This could be explained by excessive oversampling and confusion between the two interferences.

To simulate multiple contaminants present at once, we introduce a binary contamination matrix. Each row of the matrix decides which combination of contaminants is being added to the

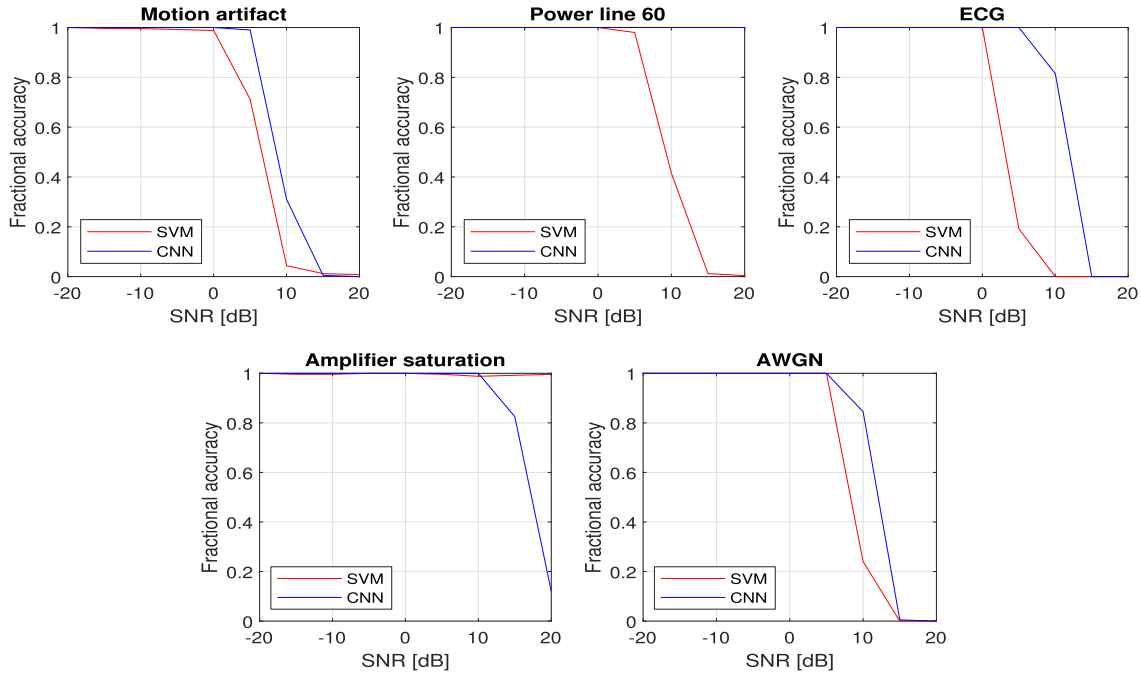


Fig. 4. Comparing the accuracy of SVM and 1D-CNN identification methods when a single contaminant is present. Five different contaminant types analyzed: MA, PL, ECG, AS, and AWGN.

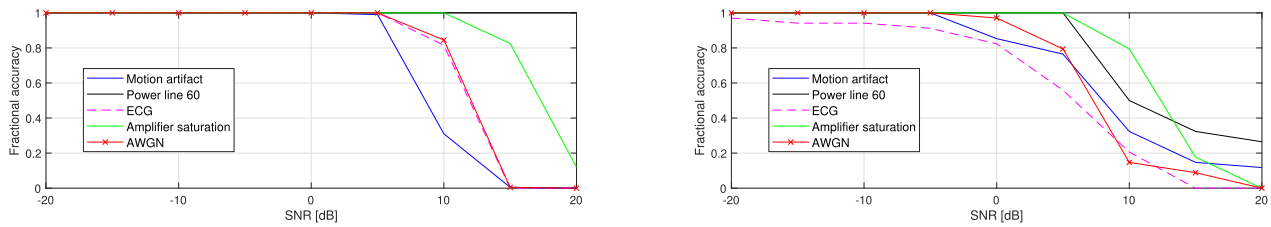


Fig. 5. Classification of five different contaminants as a function of SNR for simulated (a) and real (b) EMG signals.

EMG signal. Here, 1 corresponds to the case where the contaminant is included and 0 when the contaminant is not included in the simulations. We consider combinations of two, three, four, and all five contaminants present simultaneously. This combination gives the matrix of size 31 × 5. For the simulation of multiple contaminants, we have treated this problem as a regression problem. The output is a vector of five elements for each contaminant. The vector elements are compared to a threshold to determine if a contaminant is present (1) or not (0). The threshold value that gives the best separation of contaminants is found empirically. For real signals the threshold has optimal value of 0.6 and for simulated signals it is 0.65. We have also added one more set of layers to the neural network. Table III summarizes the network architecture of multiple contaminants.

## C. Discussion

We compare the 1D-CNN classification of EMG signals with single contaminant with the referent system [2] that is using

TABLE II
1D-CNN ARCHITECTURE USED FOR SINGLE CONTAMINANT IDENTIFICATION

| No. | Layer | No. of filters * kernel size | Layer parameters | Output size |
|---|---|---|---|---|
| 0 | Input | - | - | 15000 |
| 1 | Conv1D | 24*64 | Stride=1, ReLU | 14937*24 |
| 2 | BatchNorm | 24*64 | - | 14937*24 |
| 3 | Dropout | 24*64 | p=1 | 14937*24 |
| 4 | MaxPool | - | stride=1, width=2 | 7468*12 |
| 5 | Conv1D | 12*32 | Stride=1, ReLU | 7437*12 |
| 6 | BatchNorm | - | - | 7437*12 |
| 7 | Dropout | - | - | 7437*12 |
| 8 | MaxPool | - | stride=2, width=2 | 3718*12 |
| 9 | Conv1D | 6*16 | Stride=1, ReLU | 3703*6 |
| 10 | BatchNorm | - | - | 3703*6 |
| 11 | Dropout | - | p=0.1 | 3703*6 |
| 12 | MaxPool | - | stride=2, width=2 | 1851*6 |
| 13 | Dense | - | - | 1*48 |
| 14 | Dropout | - | p=0.5 | 1*48 |
| 15 | Dense | - | - | 1*5 |
| 16 | Softmax | - | - | 1*5 |

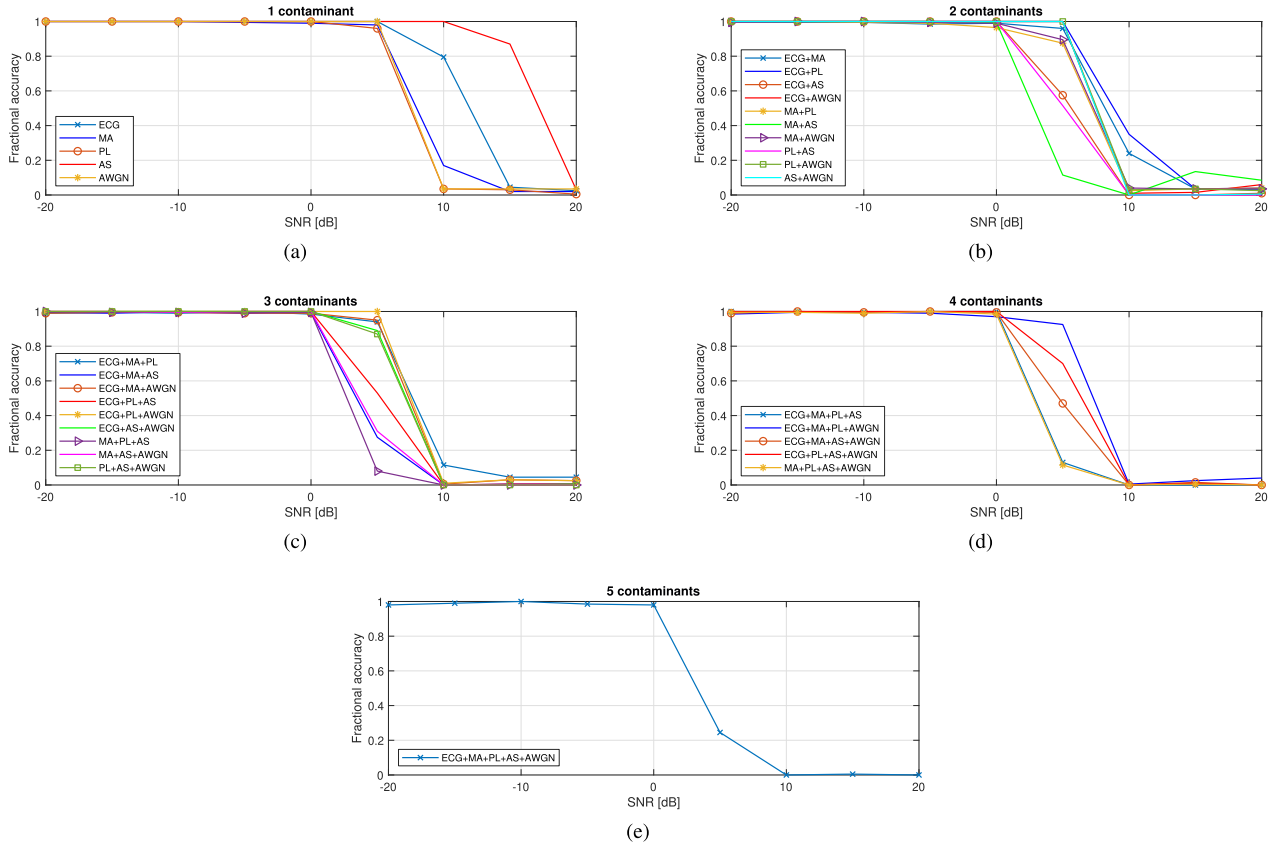


Fig. 6. Simulation results with simulated EMG signal and multiple contaminants.

TABLE III
1D-CNN ARCHITECTURE USED FOR MULTIPLE CONTAMINANTS IDENTIFICATION

| No. | Layer | No. of filters * kernel size | Layer parameters | Output size |
|---|---|---|---|---|
| 0 | Input | - | - | 15000 |
| 1 | Conv1D | 24*64 | Stride=1, ReLU | 14937*20 |
| 2 | BatchNorm | 24*64 | - | 14937*20 |
| 3 | Dropout | 24*64 | p=1 | 14937*20 |
| 4 | MaxPool | - | stride=2, width=2 | 7468*20 |
| 5 | Conv1D | 12*32 | Stride=1, ReLU | 7437*20 |
| 6 | BatchNorm | - | - | 7437*20 |
| 7 | Dropout | - | - | 7437*20 |
| 8 | MaxPool | - | stride=2, width=2 | 3718*12 |
| 9 | Conv1D | 20*8 | Stride=1, ReLU | 3703*20 |
| 10 | BatchNorm | - | - | 3703*20 |
| 11 | Dropout | - | p=0.1 | 3703*20 |
| 12 | MaxPool | - | stride=2, width=2 | 1851*20 |
| 13 | Conv1D | 20*16 | stride=1, ReLU | 1844*20 |
| 14 | BatchNorm | - | - | 1844*20 |
| 15 | Dropout | - | p=0.1 | 1844*20 |
| 16 | MaxPool | - | stride=2, width=2 | 922*20 |
| 17 | Dense | - | - | 1*48 |
| 18 | Dropout | - | p=0.5 | 1*48 |
| 19 | Dense | - | - | 1*5 |

feature extraction and SVM classifier. Both 1D-CNN and SVM are trained and tested on the same simulated EMG signals, and the results for each contaminant are shown in Fig. 4. For low and high SNR, both classifiers exhibit expected behavior of classifying and not classifying contaminants, respectively. The transition region for MA, ECG, and AWGN occurs at slightly higher SNR values for CNN classifier than for SVM. This indicates that the network could fine-tune to additional features beyond the seven features used in SVM classifier that can still help identify contaminants for modest SNR values. Those features could correspond to time-frequency domain features not considered in [2]. In real-life scenarios, it is expected to deal with multiple contaminants present simultaneously in the EMG signal.

Simulation results with simulated EMG contaminated with groups of contaminants from one (single) to five (all) contaminants are shown in Fig. 6(a)–(e). It can be seen that the network can correctly identify contaminants for negative SNR values for which it is trained. Furthermore, depending on the combinations of contaminants, the network was able to generalize its behavior for small positive SNR. At higher SNR values, identification property decays as expected. The transition region shows different behavior depending on the contaminants' combination. As we focus on more number of contaminants but fewer combinations, the transition region shifts toward lower SNR. When two combinations of contaminants are considered, test accuracy is above 0.85 at 5 dB SNR for all combinations of contaminants, apart from when ECG, MA, or PL are combined with AS. When three combinations of contaminants ECG+MA+PL, ECG+AS+AWGN, ECG+MA+AWGN, ECG+PL+AWGN, PL+AS+AWGN, are considered, test accuracy is above 0.85 at 5 dB SNR. The network is not able to further

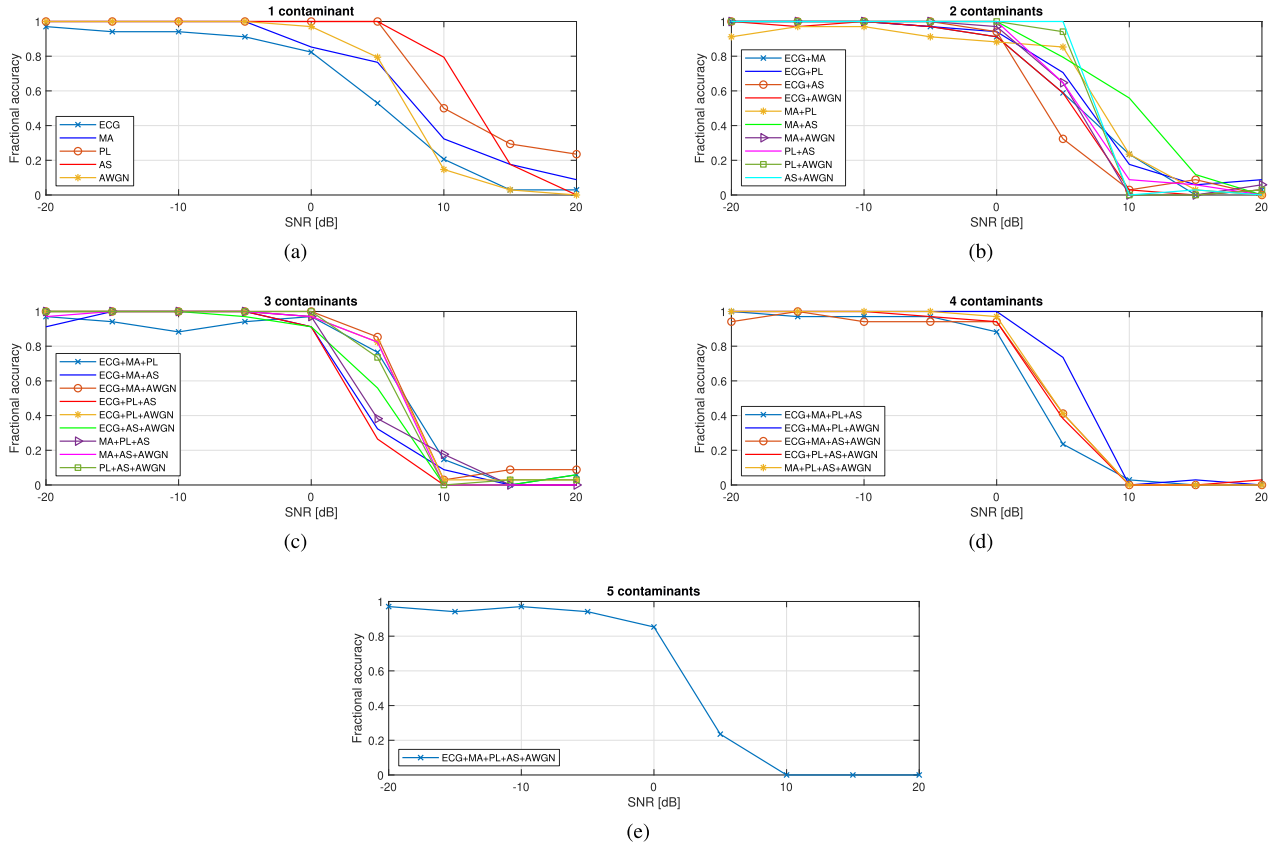


Fig. 7. Simulation results with real EMG signal and multiple contaminants.

generalize on MA and AS combinations. Similar behavior is observed when four (and all five) combinations of contaminants are considered. The generalization property of the network applies to ECG, PL, AWGN, but not on MA and AS combinations. Simulation results when real EMG signals are used are shown in Fig. 7(a)–(e). It can be observed that the classifier experiences higher sensitivity to both contaminants and their combinations then in the case of simulated signals. Also, by using the data augmentation techniques, the pattern of behavior observed in simulated EMG case vanishes. We can say that combinations with AWGN tend to have slightly higher accuracy and with ECG accuracy is lower. Larger dataset or different data augmentation techniques should be used in the future to increase the accuracy of identification.

### D. Security and Privacy Preservation

Various test cases are simulated to validate the results.

*Case 1:* Adversary between IoHT device and anchor node

This case presents the scenario when a static adversary changes the routing path of the communication channel. The path is now from IoHT device to adversarial node and adversarial node to anchor node. A sudden change in the RSSI values is observed because both IoHT device and anchor node receive higher values for RSSI, as shown in Fig. 8. As previously described, this witnesses the eavesdropper in the IoHT network. Though, the difference values do not show any obvious change

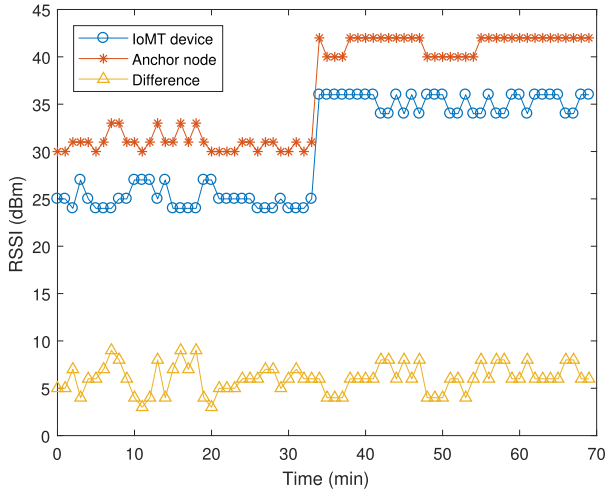


Fig. 8. RSSI variations of IoHT device and anchor node. Adversary is present in the network.

because the RSSI values of both IoHT and anchor node go up synchronously. In Fig. 9, a value other than 0 and 1 is also acquired, which justifies the presence of adversary in the IoHT.

*Case 2:* Data tampering at IoHT

When the data tampers at the IoHT device, the RSSI values which it computes are also changed as they are a part of the packet as header to the payload. The IoHT devices send this

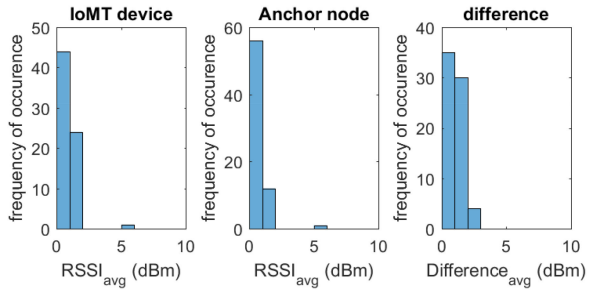


Fig. 9. Average of consecutive RSSI and distance of IoHT device and anchor node when adversary is present in network.

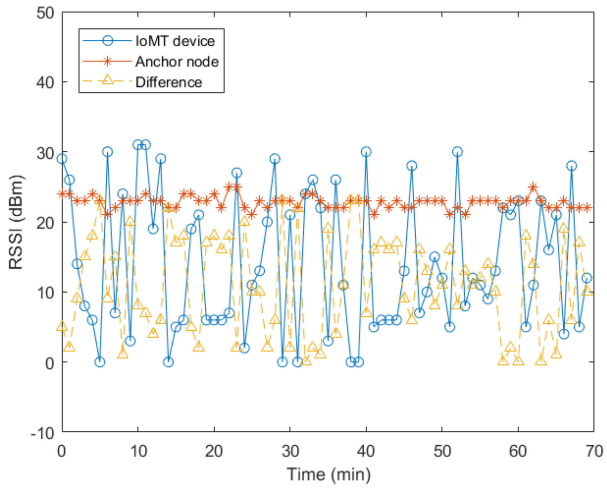


Fig. 10. Variations in RSSI of IoHT device in the presence of reply attack.

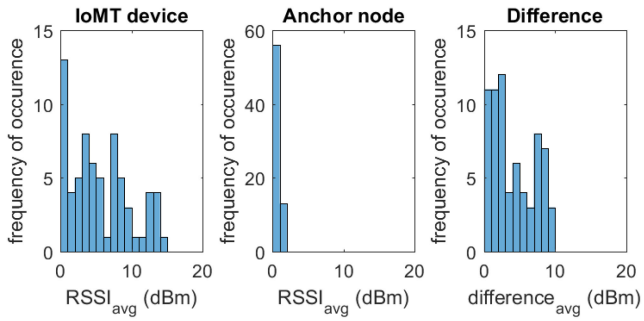


Fig. 11. Average of consecutive RSSI of IoHT device and distance when data is forged at IoHT device level.

tampered data to the server, which applies all the algorithms to compute RSSI values in $dBm$. Fig. 11 presents the spread of RSSI and difference values all over instead of $0's$ and $1's$ which is a metric to justify tampering of data at the IoHT.

*Case 3:* Replay attack by the adversary: When the adversary generates the same data at every instant of time and sends it to the server, the serve considers that data as legitimate and decodes it as per the key and nonce values of concerned IoHT. As described in Section IV-B, the nonce changes at every time instant, which generates different results at each time when decoded using synchronized nonce values. Fig. 10 shows that the LF of anchor node is decoded successfully and shows no or very little variations in RSSI values. On the other hand, the server assumes that the received encoded LF is from medical device having synchronized nonce so it decodes the encoded LF with the key and nonce associated with that. Fig. 10 presents variations in the RSSI values of medical device which according to our algorithms is enough to justify the presence of adversary in the IoHT.

## VI. Conclusion and Future Work

In this paper, we have presented a novel method for classification of contaminants in EMG signals using DNN for low-cost real-time applications. We have demonstrated the classification of multiple contaminants being present at once by using a 1D-CNN network with four convolutional layers. Simulations have been conducted on real and simulated EMG signals. The results have shown that the DNNs are well suited for the identification of contaminants in EMG signals. Moreover, it has been demonstrated that the 1D-CNN approach extracts additional information beyond hand-made features used in the previous classical ML approaches with SVM. Still, the small dataset of real EMG signals for training, has resulted in the degraded performance when multiple contaminants are present. This limitation demands a collection of bigger dataset and/or exploration of data augmentation techniques in the follow-up work. This work can be extended in the direction of modeling and addition of remaining contaminants that are not addressed here. Also, it is worth seeing the capability of SVM and 1D-CNN when all possible contaminants are present simultaneously in an EMG signal. The security for IoHT is provided using RSSI of IoHT and connected anchor nodes, which shows that the proposed solutions are robust against different physical and cyber attacks in IoHT.

## Acknowledgment

The authors would like to thank Dr Graham Fraser [2] for providing the data set of real EMG signals.

## References

[1] P. Funk and N. Xiong, "Extracting knowledge from sensor signals for case-based reasoning with longitudinal time series data," in *Case-Based Reasoning Images Signals*. Berlin, Heidelberg: Springer, 2008, pp. 247–284.

[2] G. D. Fraser, A. D. Chan, J. R. Green, and D. T. MacIsaac, "Automated biosignal quality analysis for electromyography using a one-class support vector machine," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 12, pp. 2919–2930, Dec. 2014.

[3] O. Avci, O. Abdeljaber, S. Kiranyaz, and D. Inman, "Structural damage detection in real time: Implementation of 1D convolutional neural networks for SHM applications," in *Structural Health Monitoring Damage Detection*, Volume 7. Cham, Denmark: Springer, 2017, pp. 49–54.

[4] S. Kiranyaz, T. Ince, O. Abdeljaber, O. Avci, and M. Gabbouj, "1-D convolutional neural networks for signal processing applications," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2019, pp. 8360–8364.

[5] P. McCool, G. D. Fraser, A. D. Chan, L. Petropoulakis, and J. J. Soraghan, "Identification of contaminant type in surface electromyography (EMG) signals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 4, pp. 774–783, Jul. 2014.

[6] W. Geng, Y. Du, W. Jin, W. Wei, Y. Hu, and J. Li, "Gesture recognition by instantaneous surface EMG images," *Sci. Rep.*, vol. 6, 2016, Art. no. 36571.
[7] M. Jordanic, M. Rojas-Martínez, M. A. Mañanas, and J. F. Alonso, "Spatial distribution of HD-EMG improves identification of task and force in patients with incomplete spinal cord injury," *J. Neuroeng. Rehabil.*, vol. 13, no. 1, pp. 1–11, 2016.
[8] Y. Zheng and X. Hu, "Reduced muscle fatigue using kilohertz-frequency subthreshold stimulation of the proximal nerve," *J. Neural Eng.*, vol. 15, no. 6, 2018, Art. no. 066010.
[9] M. Chen and P. Zhou, "A novel framework based on FastICA for high density surface EMG decomposition," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 1, pp. 117–127, Jan. 2016.
[10] M. Al Harrach *et al.*, "Denoising of HD-sEMG signals using canonical correlation analysis," *Med. Biol. Eng. Comput.*, vol. 55, no. 3, pp. 375–388, 2017.
[11] G. D. Fraser, A. D. Chan, J. R. Green, and D. MacIsaac, "Detection of ADC clipping, quantization noise, and amplifier saturation in surface electromyography," in *Proc. IEEE Int. Symp. Med. Meas. Appl. Proc.*, 2012, pp. 1–5.
[12] G. D. Fraser, A. D. Chan, J. R. Green, N. Abser, and D. MacIsaac, "CleanEMG-Power line interference estimation in sEMG using an adaptive least squares algorithm," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2011, pp. 7941–7944.
[13] H. R. Marateb, M. Rojas-Martínez, M. Mansourian, R. Merletti, and M. A. M. Villanueva, "Outlier detection in high-density surface electromyographic signals," *Med. Biol. Eng. Comput.*, vol. 50, no. 1, pp. 79–89, 2012.
[14] N. Nazmi *et al.*, "A review of classification techniques of EMG signals during isotonic and isometric contractions," *Sensors*, vol. 16, no. 8, 2016. [Online]. Available: https://www.mdpi.com/1424-8220/16/8/1304
[15] M. U. Abbasi, A. Rashad, G. Srivastava, and M. Tariq, "Multiple contaminant biosignal quality analysis for electrocardiography," *Biomed. Signal Process. Control*, vol. 71, 2022, Art. no. 103127.
[16] M. U. Abbasi, A. Rashad, A. Basalamah, and M. Tariq, "Detection of epilepsy seizures in neo-natal EEG using LSTM architecture," *IEEE Access*, vol. 7, pp. 179074–179085, 2019.
[17] J. Wang, Y. Liu, S. Niu, and H. Song, "Lightweight blockchain assisted secure routing of swarm UAS networking," *Comput. Commun.*, vol. 165, pp. 131–140, 2021.
[18] Y. Liu *et al.*, "Zero-bias deep learning for accurate identification of Internet-of-Things (IoT) devices," *IEEE Internet Things J.*, vol. 8, no. 4, pp. 2627–2634, Feb. 2021.
[19] Y. Liu, J. Wang, J. Li, S. Niu, and H. Song, "Class-incremental learning for wireless device identification in IoT," *IEEE Internet Things J.*, to be published, doi: 10.1109/JIOT.2021.3078407.
[20] J. Wang, Y. Liu, S. Niu, H. Song, W. Jing, and J. Yuan, "Blockchain enabled verification for cellular-connected unmanned aircraft system networking," *Future Gener. Comput. Syst.*, vol. 123, pp. 233–244, 2021.
[21] Ö. Yıldırım, U. B. Baloglu, and U. R. Acharya, "A deep convolutional neural network model for automated identification of abnormal EEG signals," *Neural Comput. Appl.*, vol. 32, no. 20, pp. 15857–15868, 2020.
[22] M. Kamal and M. Tariq, "Light-weight security and data provenance for multi-hop Internet of Things," *IEEE Access*, vol. 6, pp. 34439–34448, 2018.
[23] M. Kamal, G. Srivastava, and M. Tariq, "Blockchain-based lightweight and secured V2V communication in the Internet of Vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 3997–4004, Jul. 2021.
[24] A. Le Guennec, S. Malinowski, and R. Tavenard, "Data augmentation for time series classification using convolutional neural networks," in *ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data*, 2016.