IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 26 October 2022, accepted 23 November 2022, date of publication 2 December 2022,
date of current version 13 December 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3225899

RESEARCH ARTICLE

DeepRTSNet: Deep Robust Two-Stage Networks for ECG Denoising in Practical Use Case

POORYA AGHAOMIDI[1], AMIR MOHAMMADISARAB[1], (Member, IEEE), JALIL MAZLOUM[2], MOHAMMAD ALI AKBARZADEH[3], MAHDI OROOJI[4], NADER MOKARI[1], (Senior Member, IEEE), AND HALIM YANIKOMEROGLU[5], (Fellow, IEEE)

[1]Department of Electrical and Computer Engineering, Tarbiat Modares University, Tehran 1411713116, Iran
[2]Department of Electrical Engineering, Amirkabir University of Technology, Tehran 158754413, Iran
[3]Cardiovascular Research Center, Shahid Beheshti University of Medical Sciences, Tehran 1516745811, Iran
[4]Department of Electrical and Computer Engineering, University of California Davis, Davis, CA 95616, USA
[5]Department of Systems and Computer Engineering, Carleton University, Ottawa, ON K1S 5B6, Canada

Corresponding author: Nader Mokari (nader.mokari@modares.ac.ir)

ABSTRACT In this paper, we develop a low-cost cellular internet of medical things (IoMT)-based electrocardiogram (ECG) recorder for monitoring heart conditions and used in practical cases. In order to remove noise from signals recorded by these non-clinical devices, we propose a cloud-based denoising approach that focuses on utilizing deep neural network techniques in the time-frequency domain through the two stages. Accordingly, we exploit the fractional Stockwell transform (FrST) to transfer the ECG signal into the time-frequency domain and apply the deep robust two-stage network (DeepRTSNet) for noise cancellation. Due to the practical use case, the various heart physiologies and noise levels in different amplitudes and frequencies are needed to be robust against wide-range noises in actual conditions. We utilize the MIT-BIH Apnea-ECG database (APNEA-ECG) with several different heart physiologies. Next, the different noises consisting of muscle artifacts (MA), baseline wander (BW), and electrode motion (EM) from the MIT-BIH Noise Stress Test Database (NSTDB) and random noise, are added to the signals. The main focus of the noise generation part is the fast Fourier transformation (FFT) of the simulated noisy signal and the practical noisy signal has a maximum cross-correlation to gain a better morphological resemblance between realistic signals and the prepared datasets. Based on the results, DeepRTSNet outperforms prior learning-based methods and conventional non-learning approaches in terms of signal-to-noise ratio (SNR), root mean square error (RMSE), and percent root mean square difference (PRD). Moreover, outcomes reveal that DeepRTSNet has an extraordinary performance with a certain amount of further complexity than others.

INDEX TERMS Internet of Medical Things, electrocardiogram signal, denoising, time-frequency domain, deep learning.

I. INTRODUCTION

A. MOTIVATIONS AND STATE OF THE ART

Cardiovascular diseases (CVDs) are causing a significant increase in deaths, based on data from the World Health Organization (WHO) and the American Heart Association (AHA) [1], [2]. CVDs, especially heart rhythm problems or arrhythmias, are the most significant cause of unexpected death [3]. Heart arrhythmias are irregular heartbeats, and arrhythmias occur when natural rhythms do not coordinate correctly with electrical impulses in the heart. Electrocardiogram (ECG) is a simple, minimally invasive, and inexpensive way of diagnosing and assessing arrhythmia. Introducing the Internet of Things (IoT) leads to a new generation of mobile monitoring healthcare devices which consist of vital sign sensors, processors, and cellular communication modules. Utilizing these wearable gadgets allows healthcare staff to monitor the condition of patients from anywhere and anytime [4], [5]. Besides the regular clinical tests, a cellular wearable IoT-based ECG recorder can be a suitable solution for real-time

The associate editor coordinating the review of this manuscript and approving it for publication was Qichun Zhang.

monitoring and early heart disease diagnosis. Furthermore, the patient's signs can be recorded during various activities and then sent to a cloud for analysis due to their portability. Eventually, experts can find anomalies and alarm the critical event of the patient. In terms of price and size, these devices have fewer quality components than clinical recorders [6], [7]. As a result, the obtained signals from these gadgets have an additional amount of noise. Nevertheless, the main noise components are baseline wander (BW), muscle artifact (MA), and electrode motion (EM) [8], [9], [10]. Known as BW, it is a low-frequency noise generated by breathing, electrically charged electrodes or the movement of subjects [11]. Muscles cause MA during activity [12], and EM is produced by changing electrode impedance during recording signal [13]. All these noises cover the significant features of ECG signals utilized for disease diagnosis. Therefore, the use of denoising methods for ECG signals is pivotal.

Consequently, before any other analysis, approximately removing noise from ECG is compulsory. Bandpass filters are used in the 0.1-100 Hz range, and their output proceeds via a moving average filter for further smoothing [14]. Furthermore, we can utilize an analog low-pass filter to eliminate high-frequency artifacts and keep the essential part of data for processing [15]. For the target of canceling BW and DC offset, high-pass can be a good suggestion [15]. Moreover, other filters like notch, median, Savitsky-Golay, wavelets, and adaptive filters have good performance and accuracy for ECG denoising during these years, [16]. However, all of these filters can successfully remove noises, but facing different heart physiologies and types of noises, especially when the morphology of the signal is crashed, their accuracy decreases. Developing artificial intelligence (AI) and neural networks (NN) leads to various advancements in signal processing. Learning-based ECG denoising reveals a good ability and performance to face noisy ECG signals. In this case, convolutional neural networks (CNNs) and denoising encoder-decoders have a primary role in learning the sparsity of data [17], [18]. As the correlation between noise and signal increases, the one-dimensional approaches cannot satisfy our goal of signal improvement. Accordingly, two-dimensional (2D) methods must be employed to achieve a better separation of noise from the clean signal, such as time-frequency transformation that can be used to convert the ECG signal to the time-frequency domain. Following that, utilizing 2D CNN for denoising and, finally, hiring the inverse of them to return the time-frequency representation to the standard form of ECG signal for diagnosis [19], [20].

B. RELATED WORKS

1) IoMT-BASED ECG EVENT RECORDER

A recent study [21] presents an IoT-based ECG monitoring system that employed an AD8382 ECG sensor to read patients' data, Arduino Uno, and ESP8266 Wi-Fi module for remote healthcare. This study created an IoT Blynk application that can be installed on smartphones to process and visualize ECG signals everywhere and at any time without clinical infrastructure. Latency, connectivity, and battery lifetime are three issues that wearable sensors face regarding vulnerability and limitations. To cope with these issues, [22] suggested a method for empowering the gateway; therefore, all tasks that require high power consumption are accomplished locally on multi-core processors. Furthermore, the authors explore the performance of real-time Compressive sensing (CS)-based recovery of ECG signals on an IoT gateway embedded with ARM's big.littleTM multicore. ECG signal monitoring with an abnormal beat detection algorithm is introduced in [23]. An average single format-based identification technique requires much memory to create a layout, and strange beats make it hard to produce a common beat format. Hence authors propose a dependable strategy for producing a typical beat layout utilizing a format cluster with Pearson likeness. The proposed technique utilizes the weighted mean to limit memory use in the format cluster generation step. The authors in [24] show the possibility of ECG observation with sewn material anodes rather than customary gel terminals in a 3-lead, chest-mounted setup. The material terminals are sewn with silver-covered string covering crisscross examples into an inextensible texture. Sensor approval included ECG observing and solace overviews with human subjects, stretch testing, and wash cycling. The cathodes are tried with the BIOPAC MP160 ECG information-obtaining module. In [25], a portable, practical framework,' NeuroSpy,' utilizing off-the-rack jellybean segments, is planned, which is useful and insightful like industrially accessible gadgets from various merchants. The created framework has lower creation and execution costs by 2.22-90 than conventional business gadgets. This framework can process various biological inputs, including EEG and ECG, and track the human internal heat level over time. In order to identify certain cardiac diseases, a processing algorithm in IoT-based devices is introduced in [26]. Even though this device can detect heart diseases alone, it can simultaneously transmit data to a remote cardiologist. Furthermore, the proposed computing algorithm is developed in the LabVIEW environment, and this paper focuses on removing and updating the device's memory.

2) TIME-FREQUENCY DOMAIN AND ECG DENOISING

The QRS complex is one of the significant features of ECG signals that can help us analyze heart conditions. In order to increase the number of correct R-peaks detection, [27] introduces a new technique based on Fractional Stockwell transform (FrST). The authors ' main focuses are artifact cancelation, enhancing accuracy, and detecting the QRS complex for unconventional heartbeats that do not belong to the five heart standard beat classes. Different noise sources influence the ECG signal, leading to its morphological damage. That is the motivation of [28] to develop a Stockwell transform (ST)-based noise removing method, in which ST is used for estimating the noise level and above the specific level removed with better performance compared to wavelet-based techniques. Another FrST-based denoising technique with greater

robustness is investigated in [29] that the MIT-BIH Arrhythmia and the European ST-T databases are used as a dataset. For simulating BW and MA noises, the authors add Gaussian noise at 5 dB, 10 dB, and 15 dB classes to datasets in which their proposed method can successfully cancel noise evaluated with Root-Mean-Square Error (RMSE), Percent Root Mean Square Difference (PRD), and Signal-to-Noise Ratio (SNR) parameters. The ECG test of patients with implanted pacemakers is burned with EMG noises that disrupt the pace of pulse artifact detection. Accordingly, [30] introduces a methodology based on ST, Shannon energy computing, and threshold rule for identifying pace artifacts on noisy ECG caused by EMG noise. The time-frequency domain is potentially one of the promising techniques that can accurately recognize integrated noises with ECG. The mixture of high-order synchrosqueezing transform (FSSTH) and non-local means (NLM) for the aim of ECG noise cancellation is used in [31]. First, the signal is converted into intrinsic modes functions (IMFs), and noisy IMF components are canceled by detrended fluctuation analysis (DFA) for the next step. Next, the remained components are filtered by NLM. Ultimately, a clean ECG can be brought from filtered IMFs.

3) AI AND CONVENTIONAL METHODS FOR ECG DENOISING

It is generally held that standard methods for denoising ECG such as Savitsky-Golay, wavelet, notch, median, and band-pass filters, can be accurate with less complexity. In [32], a q-lag unbiased finite impulse response (UFIR) filter is used in discrete-time state-space for denoising ECG signal and enhancing feature extraction, which has a better performance in comparison with conventional methods. In [33], sparse optimization and a low-pass filter are utilized for ECG noise cancellation and BW estimation regarding the diverse signal features. An iterative optimization model for ECG signal denoising is discussed in [34], where the BW signal is considered as a low-pass signal, and the ECG signal is referred to as a sequence of sparse signals. Moreover, the authors use the banded matrix to characterize the optimization problem and the majorization-minimization (MM) algorithm for convergence. However, model-based Bayesian filtering is one of the promising techniques for denoising, relying on predefined causes of heavy pre-processing, and cannot guarantee performance when morphologies are changed. In this regard, [35] proposes a Bayesian technique based on the Kalman filter, one for removing QRS complex noise and the other for R and T wave denoising. Emerging Artificial Intelligence (AI), especially Deep Learning (DL), set the stage for automatic noise cancellation in medical signals. The modified frequency slice wavelet transform (MFSWT) is mixed with a convolutional neural network (CNN) in [36] to tackle the issues of the influence of noise on ECGs that are recorded by wearable devices. In [37], a CNN-based two-step ECG noise reduction is presented. First, the U-net model is employed to remove noises that lead to distorting the ECG waveform. To cope with this issue, the DR-net model is used in the next step to correct the waveform. A different approach for ECG signal denoising is proposed in [20], where the FSSTH transform is used for transforming the signal into a time-frequency domain image. Then, the image's imaginary and real parts are considered as inputs to the deep convolutional encoder-decoder network, evaluated with PRD, SNR, and RMSE. Reference [38] studies the redundant convolutional encoder-decoder for canceling noise from the QRS complex for the ECG signals obtained from armband gadgets. The authors assess the method with SNR, cross-correlation, and the percent of correct R-peaks detection.

C. NOVELTY AND CONTRIBUTIONS

The conventional non-learning denoising approaches, mainly designing filters, are considered a primary solution due to efficiency and less computational complexity. As the variety of noise resources and data sparsity increases, these methods cannot provide reliable performance. DL-based approaches are developed to cope with this problem, which are more trustworthy. These models need an appropriate training dataset in practical use cases (nonclinical ECG recording environments such as cellular IoMT-based ECG recorders). It is required to familiarize the DL algorithm with the characteristics of actual conditions in terms of the wide range of clients with different physiologies, morphological signal suppressing, and features masking caused by high noise levels. We are encouraged to propose a new solution to tackle the previous model's problem based on these bottlenecks. The summary of the main related works and comparison with our contributions can be found in Table. 1. Our contributions can be summarized as follows:

- The typical noise generation methods use BW, MA, EM, and random noise with different intensities, frequencies, and a little morphological change leading to denoise by existing methods. In contrast, these approaches do not close to the practical use cases where noises severely suppress the ECG signal, and its vital features are masked. To solve this problem, we generate the noise using two different methods. The first is a combination of BW, MA, EM, and random noise, significantly changing the clean signal's morphology. The second is related to adding each of these noise components lonely. In addition, the fast Fourier transform (FFT) of these simulated noisy signals has a high cross-correlation with the noisy practical ones.
- Due to practical use cases, our denoising approach faces many clients with a wide range of heart physiologies and morphologies that can differ from each other caused of natural factors or diseases. To be adaptive in dealing with distinct morphologies and physiologies, we utilize a dataset of several heart physiologies and morphologies that are not considered in prior works.
- We propose a deep robust two-stage network (DeepRTSNet) based on an encoder-decoder in the time-frequency domain with more remarkable performance than similar DL-based denoisers and conventional filters. The significant superiority of DeepRTSNet is that

it cancels the noisy ECG signals with minimum possible loss of valuable data. Moreover, the results illustrate that for the cost of more computational complexity, DeepRTSNet achieves an outstanding denoising performance compared to other benchmarks.

- To the best of our knowledge, there is no work to exploit the fractional Stockwell transform (FrST) for transforming ECG signals from time to the time-frequency domain along with a convolutional encoder-decoder that has a better resolution leading to that DeepRTSNet powerfully removing noise from valuable data.
- Based on the authors' knowledge, no paper exploits perceptual loss functions or networks for DL in the field of ECG denoising that causes more impressive results. In this paper, the mse + HFENN (where HFENN stands for High-Frequency Error Norm Normalized) function as a loss analyzer for first-stage denoising parts.

D. PAPER ORGANIZATION

The rest of this paper is organized as follows: In Section II, we explain the methodology. The methodology is divided into four parts: The information about our ECG recorder can be found in Subsection II-A, we present the used dataset and the procedure of preparing it in Subsection II-B, the details about FrST are illustrated in Subsection II-C, and our proposed deep learning method is introduced in Subsection II-D. The numerical results are described in Section III. In Section IV, the discussion is provided. Eventually, the conclusion and future works are expressed in Section V.

II. METHODOLOGY

A. CIRCUITS AND HARDWARE

We designed an ECG event recorder called I-Heart, as is seen in Fig. 1. It is based on the STM32L151C8T6 microcontroller (MCU), an ultra-low-power Arm Cortex-M3 MCU with 64 Kbytes of flash memory, 10 KB SRAM memory, 37 digital and analog pins, and a 32 MHz CPU. It can be connected to the mobile operator with a free subscription for one year by using SIM800L GSM/GPRS IC, which supports quad-band 850 / 950 / 1800 /1900 MHz frequencies (bandwidths). The SIM800L offers data rates between 1200 and 115200 bits per second (bps). Moreover, AD8232 is used for recording ECG with three dry electrodes. One is a common electrode, and the others are for right and left thumbs, as observed from Fig. 1. The battery used is a Lithium Polymer Battery (LP422139) of 3.7 V and 300 mAh with dimensions of 4.2 x 21 x 39 millimeter (mm). Moreover, the capacity of external micro SD is 4 GB.

On the other hand, to connect to the cloud, we need pub-key, sub-key, and channel names to be obtained by registering on our server. The low-powered, cellular-based IoT heart monitoring system architecture is simple: the board collects the ECG data recorded by the probes. It sends data directly to the I-Heart via the UART protocol. The amount of data is 6000 samples, i.e., 6000 bytes. Therefore, we store the received data and transmit it via a cellular internet connection

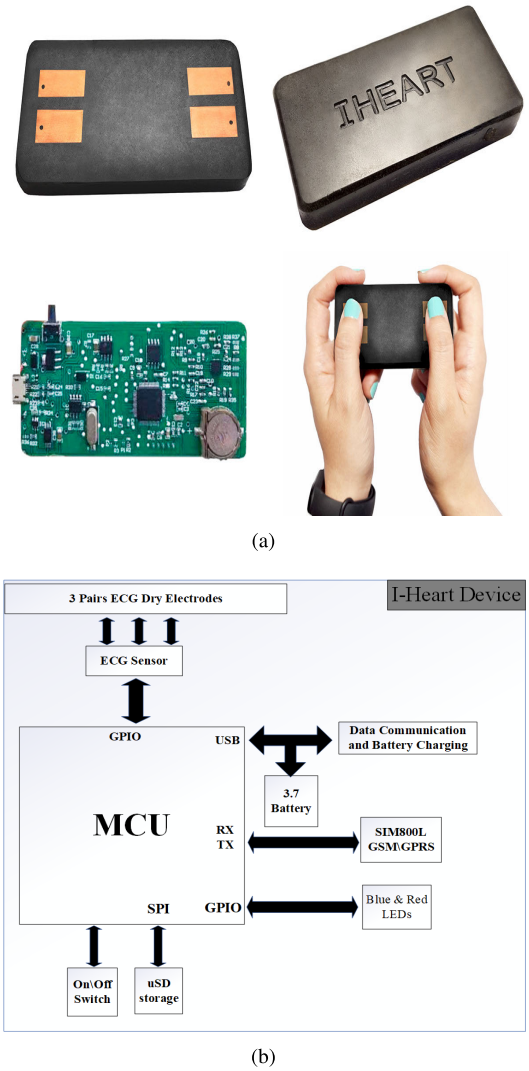


FIGURE 1. Cellular IoMT-based ECG recorder. (a) I-Heart device. (b) Schematic diagram of I-Heart device.

to the cloud. The cloud that we are using is an optimal server along with Node-RED. As can be seen from Fig. 2, the data are analyzed (denoised) and then sent to be visualized by internet-connected devices (computers, tablets, or mobile phones). Message Queuing Telemetry Transport (MQTT) is used for data transmission between patient and cloud or cloud and doctor. The doctor and patient (I-Heart) are considered MQTT-client, and the MQTT broker manages the data transmission. This architecture makes the system modular. The data acquisition board works independently, so installing many boards do not affect the system's behavior. Another advantage is the cellular connectivity of our architecture, as it can simply connect to the internet by just powering on the device. This type of architecture has four levels:

- Data collection level (DCL): This level contains the probes that transform heartbeat signals into electrical signals (digital signals). Then, these signals are sent to the I-Heart.

TABLE 1. Main Related Works Summary.

Refrences	Robustness against various noises that are morphologically similar to the ECG.	Using time-frequency domain	Utilizing various heart physiologies	2D Convolutional encoder-decoder or other models	Two stage denoising	Exploiting perceptual loss function	Practical use cases
[17]	×	×	×	✓	×	×	×
[19]	✓	×	×	✓	×	×	×
[20]	✓	✓	×	✓	×	×	×
[28]	×	✓	×	×	×	×	×
[29]	✓	✓	×	×	×	×	×
[31]	×	✓	×	×	×	×	×
[32]	×	×	×	×	×	×	×
[33]	×	×	×	×	×	×	×
[34]	×	×	×	×	×	×	×
[35]	×	×	×	×	×	×	×
[36]	✓	✓	×	✓	×	×	✓
[37]	×	×	×	×	✓	×	×
[39]	×	×	×	✓	×	×	×
Our work	✓	✓	✓	✓	✓	✓	✓

- Communication level (CL): The communication system must be suitable for applying the information recorded by the ECG probe. Our proposed system is oriented to communicate distant elements using a cellular internet. The SIM800L GSM module is used to connect to the internet.
- Cloud database and analysis level (CDAL). The elements of storage and the interpretation of data coincide at this level. The cloud data storage allows remote access to the collected data. In addition, the cloud can process the data for different visualizations. This work utilizes cloud-based processing for ECG signal denoising obtained from cellular-based recorders.
- Visualization level (VL): As mentioned above, the information is accessible for cardiologists via the registered devices such as tablets or computers.

B. DATESET

In this paper, the MIT-BIH Apnea-ECG database (APNEA-ECG) is used due to the diversity of its database in terms of the number of records. Indeed, the primary problem of other ECG databases, such as Arrhythmia and Long-term, is the number of fewer physiologies and records than apnea. Moreover, utilizing this database with long-term and diverse records can help our networks become familiar with many heart physiologies and denoise different noisy signals. The signal source is accessible at [40]. What is more, approximately 70 records in the length of slightly less than 7 hours to nearly 10 hours. Each recording includes a continuous digitized ECG signal. Using up-sampling, we changed the sampling frequency from 100 Hz to 200 Hz to make it similar to our device's sampling frequency (if it is not similar, our denoising DL-based method cannot work in dealing with practical signals obtained from I-Heart). Additionally, as can be observed from Fig. 3(a), the APNEA-ECG signal is clean (it is noteworthy that APNEA-ECG signals are not clean, and before adding our simulated noises, we cancel the negligible noises), and to simulate noisy signals, we need to add noise. Consequently, we utilize noises from the MIT-BIH Noise Stress Test Database (NSTDB), containing 12 half-hour ECG recordings and three half-hour noise recordings [39], [40]. In this study, we employ BW, EM, and MA noises of the NSTDB along with random noises as follows:

$$\text{Noise} = \eta \times \text{BW} + \mu \times \text{EM} + \gamma \times \text{MA} + \rho \times \text{random noise}, \quad (1)$$

where η, μ, γ, and ρ are coefficients and determine the effect of corresponding noise in the simulated signal. Moreover, we add this noise to the clean APNEA-ECG signals that an instance of a simulated noisy signal is presented in Fig. 3(b) where it is so obvious that the simulated signal is so close to the real noisy signal obtained from our device that is shown in Fig. 3(c). Our noise generation aims to simulate the noisy signal with maximum cross-correlation with the practical obtained signal from our recorder. In fact, we generate the desired noise by changing the coefficient and frequency of each noise component. In the final step, we split each record into short signals of 256 samples after adding noise to each record. In addition to the signals with 256 samples, we tested signals with 128 and 512 samples, but the denoising result was unsatisfactory. Eventually, we split the whole dataset into train, test, and validation with 80, 10, and 10 percent ratios, respectively.

C. FRACTIONAL STOCKWELL TRANSFORM

A time-frequency analysis method, fractional Stockwell transform (FrST), can be developed from the mixture of Stockwell transform (ST) and fractional Fourier transform (FrFT). The FrST follows ST but in the fractional domain. Firstly, for a signal $x(t)$, the ST transform is calculated as follows [41]:

$$s(\tau, f) = \int_{-\infty}^{\infty} x(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(\tau - t)^2 f^2}{2}} e^{-j2\pi f t} dt, \quad (2)$$

where t, f, and τ are represented time, frequency, and the window function's position control parameter on the t axis, respectively. Also, $|.|$ defines the absolute value. Without loss of generality, the ath order FrFT of signal $x(t)$, $X_a(u)$, is calculated using the following formula [42], [43], [44]:

$$X_a(u) = F^a(x(t)) = \int_{-\infty}^{\infty} x(t) K_a(t, u_\phi) dt, \quad (3)$$

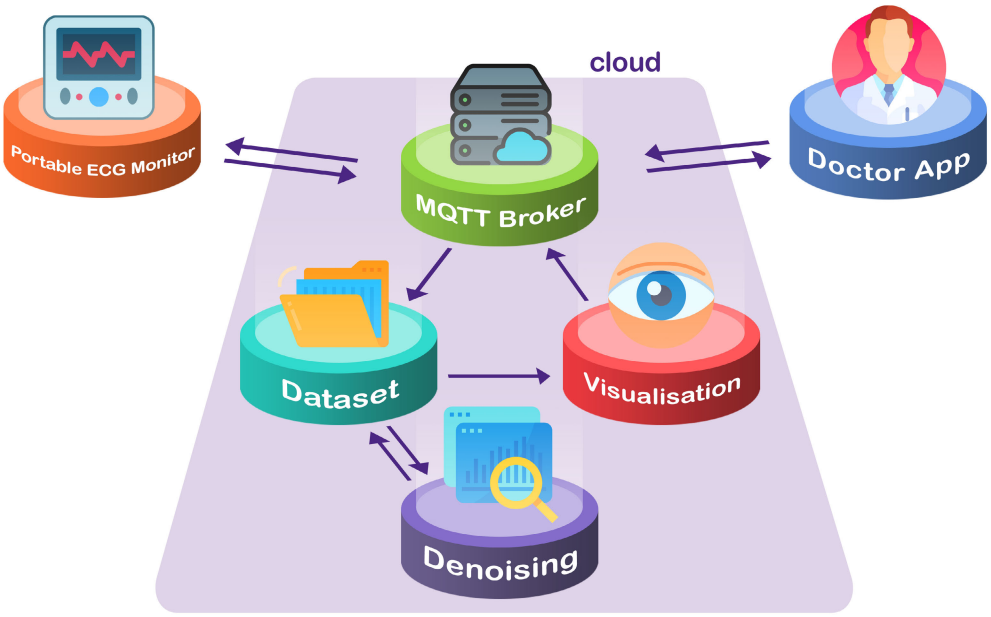


FIGURE 2. Our proposed system architecture.

where u_ϕ is the fractional Fourier frequency and $K_a(t, u_\phi)$ denotes the kernel function of the FrFT that can be formulated as follows:

$$K_a(t, u_\phi) = \begin{cases} A_\phi \exp[\chi], & \phi \neq n\pi \\ \delta(t - u_\phi), & \phi = 2n\pi \\ \delta(t + u_\phi), & \phi = (2n \pm 1)\pi \end{cases} \quad (4)$$

where $A_\phi = \sqrt{1 - j\cot\phi}$, $\phi = a\pi/2$, $\chi = j\pi\left(u_\phi^2 \cot\phi - 2u_\phi t \csc\phi + t^2 \cot\phi\right)$, and j represents the complex unit. According to [45], for a signal $x(t)$, the ath order FrST is defined as:

$$\mathrm{FrST}_x^a\left(\tau, u_\phi\right) = \int_{-\infty}^{+\infty} x(t) g\left(\tau - t, u_\phi\right) K_a\left(t, u_\phi\right) dt, \quad (5)$$

where on time t and fractional Fourier frequency u_ϕ, $g\left(\tau - t, u_\phi\right)$ is a adaptable Gaussian window function that is described as follows:

$$g\left(t, u_\phi\right) = \frac{\left|u_\phi \csc\phi\right|^p}{\sqrt{2\pi} q} \exp\left(\frac{-t^2\left(u_\phi \csc\phi\right)^{2p}}{2q^2}\right), \quad (6)$$

here p and q are two parameters that, along with fractional order a, are utilized for the window shape control. Moreover, *csc* represents the cosecant function. Together, these parameters improve the time-frequency resolution by mapping the signal's energy in the transformed domain [29]. FrST parameters (a, p, q) are determined in this study with $a = 0.7$, $p = 0.5$, and $q = 0.5$, which are derived from [29] where it is found that FrST with this values has a better resolution.

D. DEEP LEARNING

1) DENOISING AUTOENCODER

As part of deep learning, denoising autoencoders (DAE) attempt to separate the noise from the signal by recreating the input vector precisely by minimizing the loss function [46]. A simple representation of an encoder and decoder can be formulated as follows,

$$\begin{aligned} \alpha &= \varphi\left(Ax(t) + b\right) \\ \beta &= \hat{\varphi}\left(\hat{A}\alpha + \hat{b}\right), \end{aligned} \quad (7)$$

where $x(t)$ represents the noisy input data that is deterministically transformed to the hidden form, α, by using the nonlinear activation function φ. A and b denote the encoder's weight matrix and bias vector. Next, α is converted back to the reconstituted vector β. Likewise, $\hat{\varphi}$, $\hat{A}$, and $\hat{b}$ are nonlinear activation function, weight matrix, and bias vector of decoder, respectively. Optimizing the model's parameters can be calculated with the minimization of the following perceptual loss,

$$\mathrm{Loss} = \arg\min_{A,b,\hat{A},\hat{b}} \frac{1}{M} \sum_{m=1}^{M} \|x_m(t) - \beta_m\|_2^2 + \lambda \times \mathrm{HFENN}, \quad (8)$$

where the number of noisy data instances is M and the data index is m. λ is the weight of high-frequency error norm normalized (HFENN) is the particular type of loss function that can be found in [47]. In addition, $||.||$ indicates the L2 norm function. Our motivation for using the perceptual loss is to compare the time-frequency image with each other instead of the output time-domain signal. For this purpose, mse + HFENN has a better performance in this application [47].

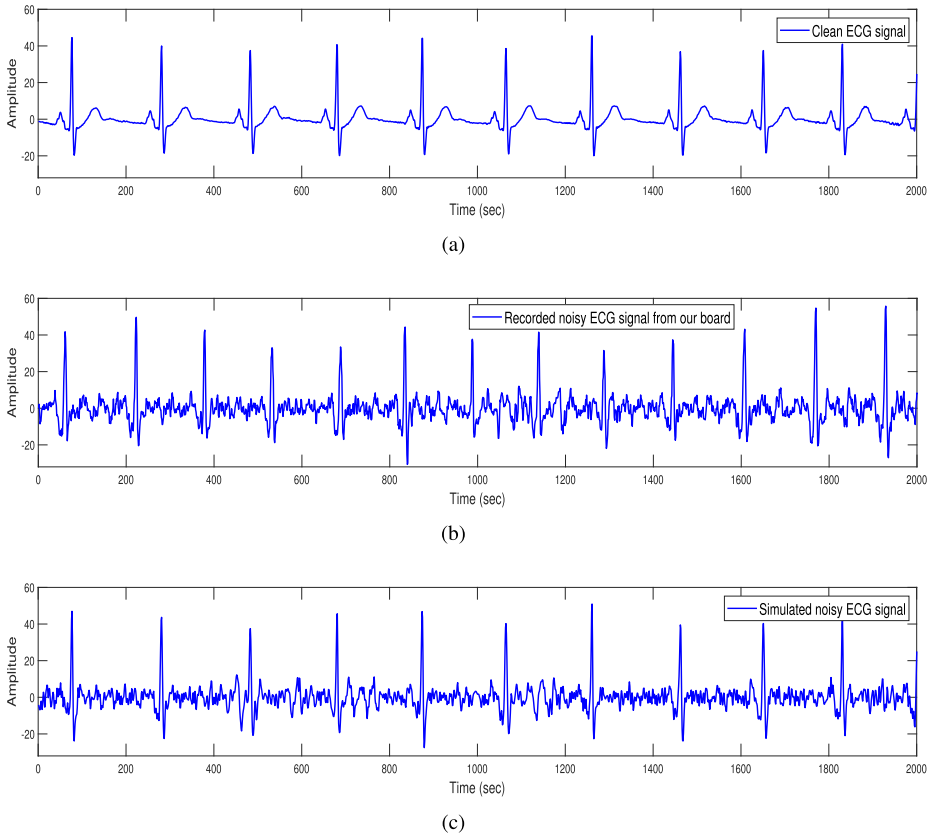


FIGURE 3. Noise generating for labeling dataset. (a) Clean ECG signal. (b) Recorded noisy ECG signal from our board. (c) Simulated noisy ECG signal.

2) OUR PROPOSED APPROACH

Firstly, a noisy ECG signal is converted to the time-frequency domain by applying FrST [29], which gives a higher resolution in comparison with ST transform. The noisy time-frequency ECG signal can be formulated as follows:

$$X(t,f) = X_c(t,f) + X_n(t,f), \tag{9}$$

where $X_c(t,f)$ is useful part of signal and $X_n(t,f)$ represents the noise of ECG signal. The denoising goal is to estimate a clean signal $\tilde{X}_c(t,f)$ from its noisy input signal $X_c(t,f)$ in the time-domain regarding the minimization of perceptual error in Eq. (8).

As well as the advantage of nonlinear mapping, artificial neural networks can learn time-frequency data sparsity. Learning data sparsity allows them to recognize useful signals from noise accurately. Auto-encoders are a promising technique for denoising applications with outstanding performance in learning data sparsity among various deep learning methods. As shown in Fig. 4, we propose a deep robust two-stage network (DeepRTSNet) architecture in the time-frequency domain. Due to perceptual loss and RGB images being employed, it is necessary to duplicate the obtained denoised time-frequency into three redcopies, as it is shown in Fig. 4. In addition, skip connections between two parts are exploited so training convergence can occur more rapidly and a better local optimum can be found more easily [48]. Apart from that, for signals that are highly corrupted by noise, we suggest another deep network that is entirely similar to the first one, as can be seen in Fig. 4. The specifications of our two-step deep encoder-decoder networks are as follows:

- **Encoder:** As discussed above, first, the input ECG signal with 256 samples is transformed into a time-frequency domain by using FrST, and then we consider the magnitude of this signal as an input of encoder that can be calculated by,

$$X_{\text{mag}}(t,f) = \sqrt{|X_{\text{real}}(t,f)|^2 + |X_{\text{img}}(t,f)|^2}, \tag{10}$$

where $X_{\text{real}}(t,f)$, $X_{\text{img}}(t,f)$ are represents the real and imaginary parts of $X_{(t,f)}$. The encoder aims to obtain α by mapping input to multidimensional feature space with nonlinear function ϕ as mentioned in Eq. (7). ϕ is implemented by a series of $2D$ convolutional layers, where the first one is a resizing layer with the $256 \times 512 \times 1$ dimension. Next, three layers are starting from 3×3 convolutional layer with exponential linear units (elu) as an activation function, continuing with batch normalization (BN) layer, and the input is entered to 3×3 convolutional with elu, stride1, stride2, and BN where stride1 is exploited to filter the feature space. However, stride 2 is utilized to decrease the feature

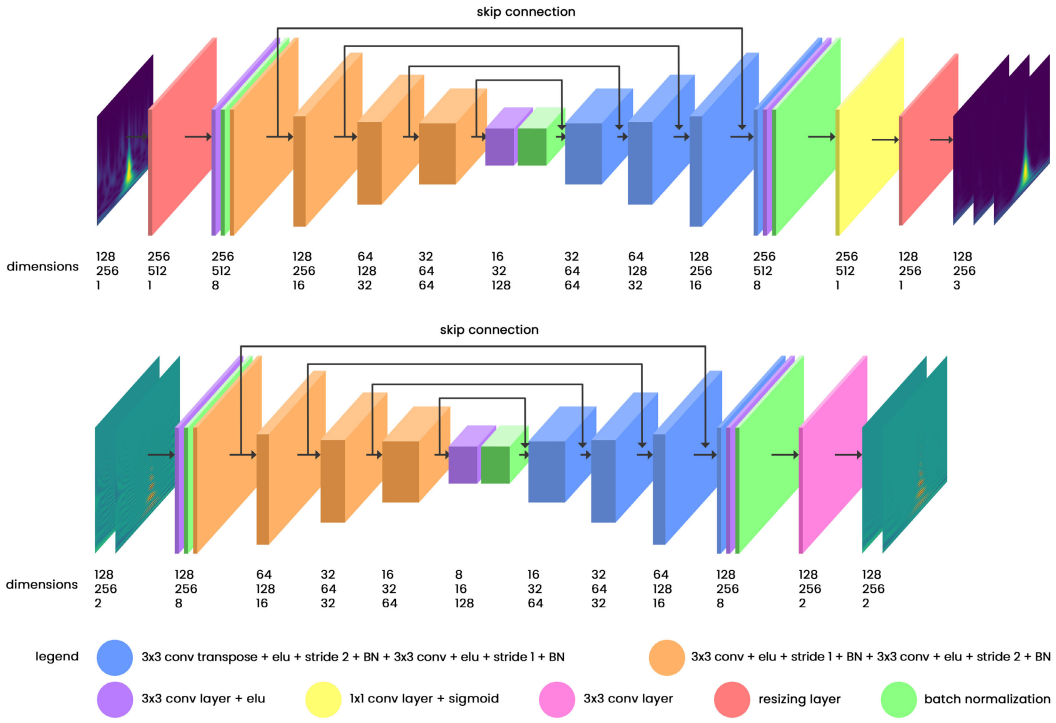


FIGURE 4. Architecture of DeepRTSNet.

space size. It is worth noting that the filter size is 3×3 and remained unchanged throughout the whole network. Moreover, this procedure is followed convolutional with specifications like a last layer for the next three layers. Furthermore, convolutional with elu is implemented in the last layer of the encoder. Consequently, each successive layer accurately learns the sparse representation of time-frequency data and separates noise from the signal.

- **Decoder:** Although the primary purpose of the decoding module is to regain time-frequency data from high-dimensional feature space generated in the encoder module, it is ultimately an inverse version of the encoder part that leads to exploring β in Eq. (7). Notably, it starts with convolutional transpose along with elu, stride 2, stride 1, and BN for three sequential layers. Furthermore, a joined layer consists of 3 parts that start one's structure entirely similar to the last layer, and the second one is a convolutional layer with elu, and this block is finished with a BN layer. After that, the data flow is perpetuated to convolutional with elu, and the encoder module is completed with the resizing layer to tune the size of the denoised magnitude.

While the first stage of our proposed deep learning approach can successfully cancel the noise from the signal, the level of noise cancellation is not acceptable for the ECG signals with high-level noise corruption. On the other hand, the first stage network can profoundly deal with noisy signals and remove them from them, while it can not operate for very noisy signals. To cope with this problem, we suggest the second stage denoising network, which is the same as the first stage's network without resizing layers at the start and end of the encoder and the decoder, respectively. Furthermore, the dimension of layers are changed in the second stage, which can be seen in Fig. 4. In order to detect which signal should enter the second network, SNR improvement of the output of the first stage denoising that in the magnitude format must be evaluated. Remark that the input of the first stage is the magnitude of real and imaginary parts of the time-frequency version of input data that is obtained by applying FrST. It is noticeable that we only use the magnitude of data and do not consider the phase due to the less complexity of one input's network. Further, Keeping the noisy phase till the end of the first stage can help us conserve the morphology of the signal. Suppose the SNR improvement of the signal is suitable. In that case, the denoised magnitude and noisy phase directly reform the complex form by using $X_c(t,f) = X_{\text{mag}}(t,f) \times \exp i \times X_{\text{phase}}(t,f)$, and then apply the inverse FrST to rebuild the denoised ECG signal that is shown in Fig. 5(a). If it is not valid, the real and imaginary parts of the denoised signal in the first stage, by applying again, the FrST obtained and enter the second network and are denoised. After this stage, the inverse FrST on denoised real and imaginary can reconstruct the denoised signal that can be observed in Fig. 5(B). In addition, it can impose extra run-time on the procedure of real-time cloud base ECG

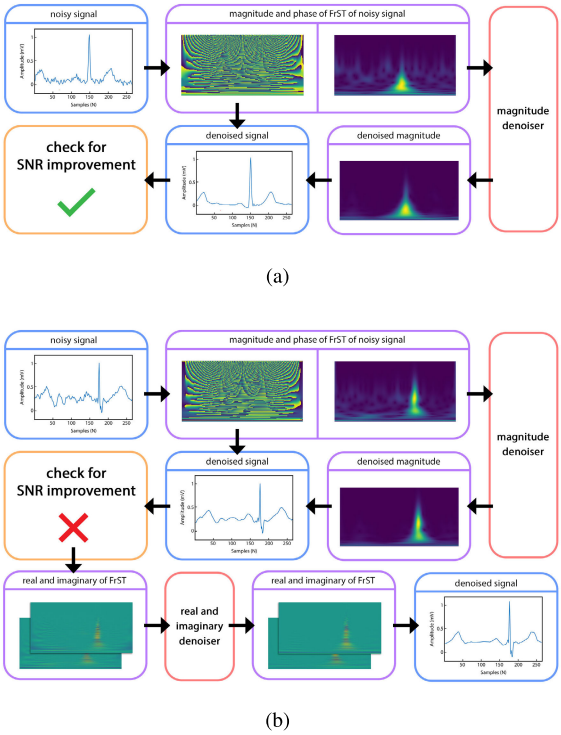


FIGURE 5. Procedure of the signals denoising for the ones that needed one or two step denoising. (a) One step denoising. (b) Two step denoising.

noise cancellation. To tackle this issue, we use the SNR estimation using a periodogram, kaiser window with the beta parameter equal to 38 [49], [50], [51].

The training dataset contains many noisy signals fed into a data generator that produces the noisy signals and the labels associated with them (clean form). Afterward, they are converted into the time-frequency representation, and required pre-processing is completed, such as normalization. This time-frequency representation is used in the first stage network, whereas in the second stage, we split the real and imaginary parts of FrST. In order to avoid memory-filling problems, the data generator feeds the network one by one once the input and mask are ready. Following the first batch of inputs, the network begins training, and the encoder attempts to create a feature map for the remaining parts. We employ a network of encoders that extract features in noisy signals from their time-frequency representations and decoders that reconstruct denoised signals from their time-frequency presentations. Using this architecture, we used a skip connection to minimize the loss function of the decoder part in the reconstruction process. We train two-stage networks separately in our approach. The first step is to design and train a first-stage network to denoise the magnitude part of FrST. Exploiting this trained network, we denoise our dataset and created a dataset with higher SNRs, which we used to train our second network that removed noise from real and imaginary parts of FrST. As a result, our second-stage network focuses on the noises left over. Remarkably, the primary concern of DeepRTSNet is to deal with noisy ECG signals in practical use cases. For this reason, the various noise resources and physiologies are regarded. In addition, DeepRTSNet can deal with any ECG signals with its standard input format in terms of sampling frequency and sample number.

In order to clarify why we use this architecture and layers in Fig. 4, the following justifications are added. Our first step was to double the input size of our network by resizing layers since this allowed us to go deeper into the network and made the result more practical. Afterward, new convolution layers, ELU, and batch normalization are applied to the reshaped inputs. Various activation functions are tested, but ELU produces the best results. As ELU becomes smooth, it tends to create more accurate results faster and converge to zero cost more quickly. We use two convolutional layers of each size (sometimes three), the first one comes with stride 1, which focuses on the convolutional filtering operation and producing the feature space, and the second comes with stride 2 to make the feature space smaller and also improves the computational efficiency. As a result of a larger kernel, the convolution calculation obtains more features from neighboring signals. Nevertheless, the large convolution kernel increases computational time, limiting the depth of the neural network. Thus, all convolution/deconvolution layers have three kernels [52]. The batch normalization method is used for training deep neural networks by normalizing the contributions made by each mini-batch. As a result, it settles the learning process and drastically decreases the number of training epochs needed to train deep neural networks. In order to prepare the input distribution of the next layer, we used a batch normalization layer after every convolutional layer. Encoding and decoding process feature maps are concatenated with skip connections, which improve the convergence of training and reconstruction information. Encoding and decoding process feature maps are concatenated with skip connections, which improve the convergence of training and reconstruction information [53]. Masks are produced in the penultimate layer of the denoising network using sigmoid activation functions. Further, transposed convolutional layers are needed to extend the feature map in order to produce an accurate mask. Alternatively, we could use up-sampling layers instead of transposed convolutional ones, but up-sampling layers have no trainable parameters and increase the size.

III. EXPERIMENTS

A. EVALUATION CRITERIA

The performance of our denoising method (DeepRTSNet) is assessed quantitatively using three criteria, namely, signal-to-noise ratio (SNR), root mean square error (RMSE), and percent root mean square difference (PRD) that can be formulated as follows [20], [37]:

$$\mathrm{SNR} = 10\log_{10}\frac{\sum_{t=1}^{M}[x_c(t)]^2}{\sum_{t=1}^{M}[x_c(t)-\tilde{x_c}(t)]^2}, \qquad (11)$$

$$\text{RMSE} = \sqrt{\frac{1}{M}\sum_{t=1}^{M}[x_c(t) - \tilde{x_c}(t)]^2}, \tag{12}$$

$$\text{PRD} = \sqrt{\frac{\sum_{t=1}^{M}[x_c(t) - \tilde{x_c}(t)]^2}{\sum_{t=1}^{M}[x_c(t)]^2}} \times 100, \tag{13}$$

where M represents the ECG length. $x_c(t)$, $\tilde{x_c}(t)$ denote the clean and the denoised ECG signal, consecutively. Our ability to measure noise corruption can be determined by analyzing SNR. Higher noise levels cause lower SNR, and higher SNR means the signal has a lower noise level. Measurement of the difference between expected and actual results is possible through RMSE, which higher RMSE signifies that the acceptable result is not achieved. Moreover, PRD can be affirmative in determining how the noisy signal is reconstructed accurately, which higher PRD means this reconstructed signal is not appropriate.

B. EXPERIMENTAL RESULTS

As we mentioned, the processes involved in our system are divided into two parts. Firstly, the signal recording's computation is done by I-Heart, where hardware specifications were mentioned in Section II part II-A. Secondly, the denoising process is thoroughly performed on the server. Indeed, training, validation, and testing are computed on our server. The hardware specifications of the server are listed in the following. It has 64 gigabytes (GB) of random-access memory (RAM), Intel XEON E5 2680 V2 central processing unit (CPU) with 10 cores (20 threads), two terabytes (TB) hard disk drive (HDD), two 500 GB solid state drives, and 2 GB graphic Nvidia Quadro K2000. Additionally, the capacity of its network port is around one gig of bits per second (Gbps), and the operating is Windows Server 2019 Version 1809 Build 17763.2114 Retail. Apart from that, both two-stage networks are implemented in Python 3.7, Tensorflow 2.6 library, and Keras 1.7 library.

1) EXPERIMENTAL RESULTS BASED ON MIT-BIH TEST DATASET

In this section, the performance of DeepRTSNet is compared with two other deep learning-based methods closest to our approach and the non-learning techniques such as wavelet, FIR, and IIR filters to prove our superiority. Furthermore, DeepRTSNet's source code can be found in [54]. As can be observed in Fig. 6(a) and Fig. 6(b), the losses of training and validation datasets of the first and second proposed networks are quickly lessened and converged to a small amount. After a particular number of epochs, the gap between training and validation losses is reduced and reaches negligible quantities. It means that our models are not overfitted along with the substantial learning potential of the proposed networks. Fully convolutional network (FCN)-based DAE [19] and deep convolutional encoder-decoder network (DeepCEDNet) [20] are more closest methods to our work. It encourages us to implement them and apply them to our dataset to compare the results. FCN-based DAE consists of 13 convolutional and deconvolution layers. Additionally, DeepCEDNet is designed with 24 convolutional and deconvolution consecutive layers. Three different physiologies from APNEA-ECG are considered to examine the DeepRTSNet, DeepCEDNet, and FCN and give a good insight into our proposed noise cancellation effectiveness.

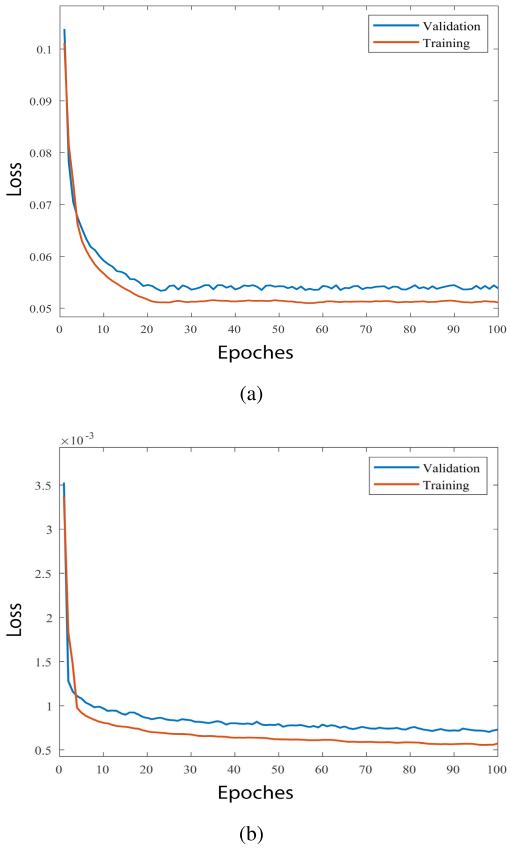


FIGURE 6. Training and validation process losses for one and two step denoising. (a) One step denoising loss. (b) Two step denoising loss.

TABLE 2. Main Learning parameters of DeepRTSNet.

	First Stage	Second Stage
Batch Size	64	16
Epoches	100	100
Optimizer	RMSprop	Adam
Loss	Perceptual Loss (mse+HFENN)	mse
Network Input Format	Magnitude	Real & Imaginary
Learning Rate	10^{-3}	5×10^{-4}
Total Number of Parameters	346034	346034

Further, the learning parameters of DeepRTSNet can be found in Table. 2. Fig. 7, 8, 9 illustrate the result of denoising on three different physiologies of our test dataset where (a) is a clean and noisy ECG signal, (b) is the denoised ECG signal by FCN-DAE along with the clean signal, (c) shows the denoised ECG signal by CEDN along with the clean signal, and (d) depicts the denoised ECG signal by our approach along with the clean signal. As can be seen in Fig. 7(a), the

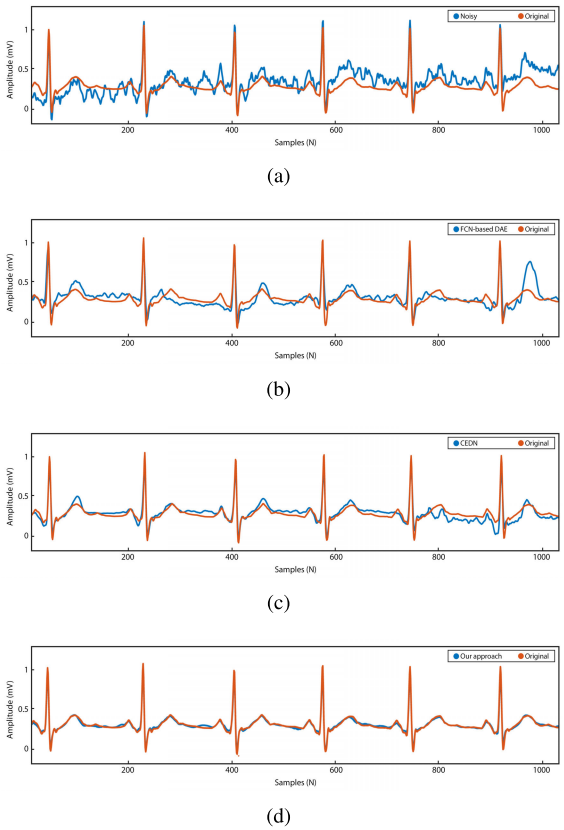


FIGURE 7. First denoised signal. (a) Original ECG signal and noisy version of it. (b) Original ECG signal and denoised version of it by FCN-based DAE. (c) Original ECG signal and denoised version of it by CEDN. (d) Original ECG signal and denoised version of it by our approach.

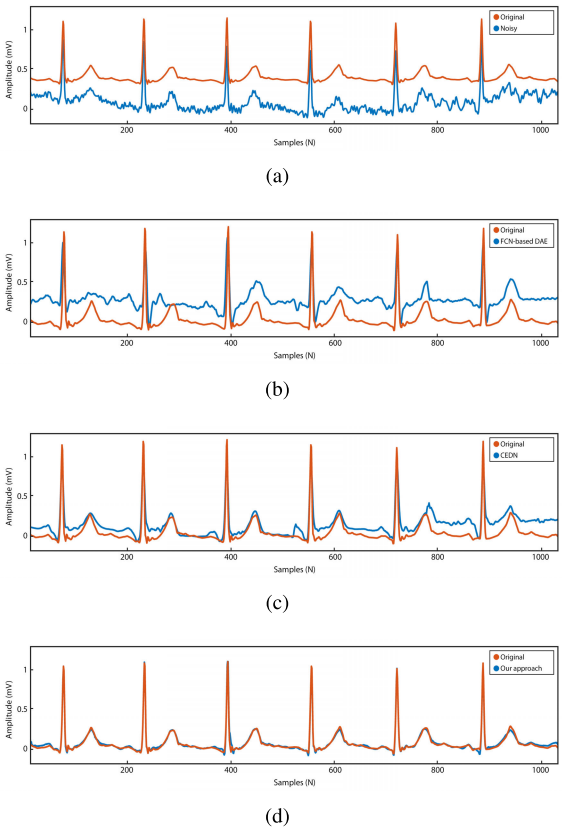


FIGURE 8. Second denoised signal. (a) Original ECG signal and noisy version of it. (b) Original ECG signal and denoised version of it by FCN-based DAE. (c) Original ECG signal and denoised version of it by CEDN. (d) Original ECG signal and denoised version of it by our approach.

noise has severely deformed the P, T, and QRS waves in the way that the significant ECG features are masked. The denoised result of FCN-DAE can be observed in Fig. 7(b) that can not smoothly remove noise from the signal with input SNR 0 dB. Although the CEDN has a better performance than FCN-based DAE, our denoising result with just one stage denoising more effectively and smoothly removes the noise than two other networks shown in Fig. 7(c), 7(d). It is worth noting that the more promising outcome of our methodology could be that the first stage network is one input (only magnitude), and our training datasets contain various high and low-frequency noises due to the practical condition. The second noisy signal can be found in Fig. 8(a) with input SNR -1 dB and a higher frequency overlap (means the various high and low-frequency noises like ECG waves) between noise and signal than Fig. 7. The FCN-based and CEDN in Fig. 8(b), 8(c) can not effectively follow the valuable signal and cancel the noise because of the misunderstanding caused by huge noise sparsity and overlap between ECG waves and noise. Alternatively, our approach can successfully remove noise and keep the morphology with a negligible amount of suppressing with two-stage denoising that can be viewed in Fig. 8(d). The first stage keeps the signal morphology and smoothes the high-frequency noises that may be achieved by denoising the only magnitude and keeping the noisy phase version of the time-frequency version of the signal. Next, in the second stage, the denoiser tries to remove the remaining noises using the smooth real and imaginary version of the denoised signal in the first stage.

Sometimes due to cardiovascular diseases and genetic backgrounds, the heart physiology is dissimilar to the regular ones. For exploiting a denoising solution in practical use cases, the denoiser must be able to deal with any physiologies because various clients are involved in the practice. As a result, the less common physiology with input SNR -3 dB in Fig. 9(a) is used to evaluate the considered denoiser to deal with different conditions. FCN-based DAE and CEDN almost failed to remove noise adequately even though CEDN has more acceptable performance, as observed in Fig. 9(b), 9(c). Meanwhile, our approach achieves satisfactory results in Fig. 9(d) while the cross-correlation between noise and signal in frequency resemblances is 72.2 percent. The reason for our more acceptable results may be that several noises in various frequencies and a massive number of physiologies (like the ECG waves) are involved in our training dataset. Recall that our motivation for designing this dataset is to develop a cellular IoMT-based ECG event recorder and an online denoising solution that should be able to deal with any physiologies.

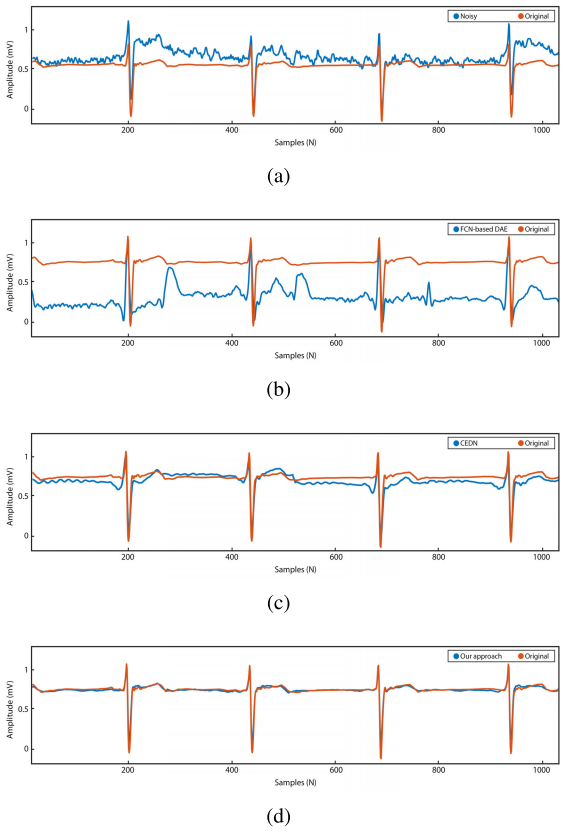


FIGURE 9. Third denoised signal. (a) Original ECG signal and noisy version of it. (b) Original ECG signal and denoised version of it by FCN-based DAE. (c) Original ECG signal and denoised version of it by CEDN. (d) Original ECG signal and denoised version of it by our approach.

To provide more justification for our above claims, apart from the signals in Fig. 7, 8, 9, the six physiologies with three input SNR -1,-2, and -3 dB are visualized in Fig. 10. Our approach removes the noises accurately with a small level of suppressing the physiological signs (P, T, and QRS waves) and other ECG features. This enhancement is because we make an effort to generate noises similar to ECG waves due to our practical application.

In order to reveal more justifications for the superiority of our approach and prove the better performance quantitatively, we calculate the average output SNR, RMSE, and PRD regarding the nine different input SNR levels that can be seen in Fig. 11. Our test dataset consists of signals with different Input SNR from -12 dB to 18 dB, and we categorized these signals into nine levels of $-4, -3, -2, -1, 0, 1, 2, 4$, and 8 dB. In addition to CEDN and FCN-based DAE, the Kalman bank filter [35], a finite impulse response (FIR) filter [32], and an infinite impulse response (IIR) filter are implemented to compare non-learning and learning ECG denoising approaches with each other. Undoubtedly, DeepRTSNet achieves a higher output SNR than the other approaches, which is increased as the input SNR is raised, which can be found in Fig. 11(a). In other words, the greater the output SNR means the more suitable noise cancellation. As shown in Fig. 11(a), the DeepRTSNet SNR improvements (especially the second stage) are dramatically superior to others. Considering our noise generation method, which causes misunderstandings between noise and the inherent wave of the signal, non-deep learning methods such as the Kalman bank filter, unbiased FIR filter, and IIR filter are unable to improve SNR significantly even though they suppress significant waves and damage the vital information of ECG. Apart from that, two major factors are influential. Firstly, exploiting FrST lead to a higher resolution which sets the stages for denoising networks to better separate noise and signal [29]. Secondly, in the second denoising stage, the real and imaginary of the denoised signal are used. It leads to more SNR improvement due to the impact of utilizing these parts instead of the magnitude and impression of the first stage denoising. Due to the high correlation between noise and signal, it is necessary to reduce noise efficiently by combining time and frequency information. In comparison with other deep learning-based methods that include the frequency domain, the FCN-based DAE only considers the time domain. This reason can justify the lower SNR improvement of FCN-based DAE compared to other learning methods. Furthermore, DeepRTSNet gains a lower RMSE and PRD than other conventional methods, which can be interpreted as using our approach leads to a better signal reconstruction. Moreover, the denoised signal with one or two stages is closer to the original signal. In contrast, other methods exploiting our dataset cannot perform well, especially in the very noisy signals, as shown in Fig. 11(b), 11(c).

2) EXPERIMENTAL RESULTS BASED ON COLLECTED DATA FROM I-HEART DEVICE

All the results in the last part, III-B1, are based on the test dataset made up of clean ECG signals from the APNEA-ECG dataset that are additive simulated noises from NSTDB. As seen in III-B1, DeepRTSNet successfully canceled the signals from the test dataset. Nevertheless, in order to make sense of the capability of DeepRTSNet to deal with signals obtained from the I-Heart device, we recorded empirical data from four separate subjects. Next, we evaluate the denoising performance with these signals. Moreover, the first subject is denoised utilizing DeepRTSNet, which can be seen in Fig. 12. The noisy version is depicted in Fig. 12(a), and the denoised ones by FCN-based DAE, DeepCEDNet, and DeepRTSNet can be observed in Fig. 12(b), 12(c), and 12(d), respectively. Undoubtedly, DeepRTSNet achieved a better performance than other benchmarks. Furthermore, the results of denoising other subjects are illustrated in Fig. 13, 14, 15. Obviously, DeepRTSNet successfully cancels these four noisy subjects' signals. In short, our approach has a good denoising performance in dealing with practical and simulated noisy signals with one training process. Although the results in III-B1 were evaluated with SNR improvement, RMSE, and PRD, it cannot be achievable for the recorded signals by the I-Heart device because clean versions of them (reference signal) are not accessible. Thus, calculating these evaluation measures

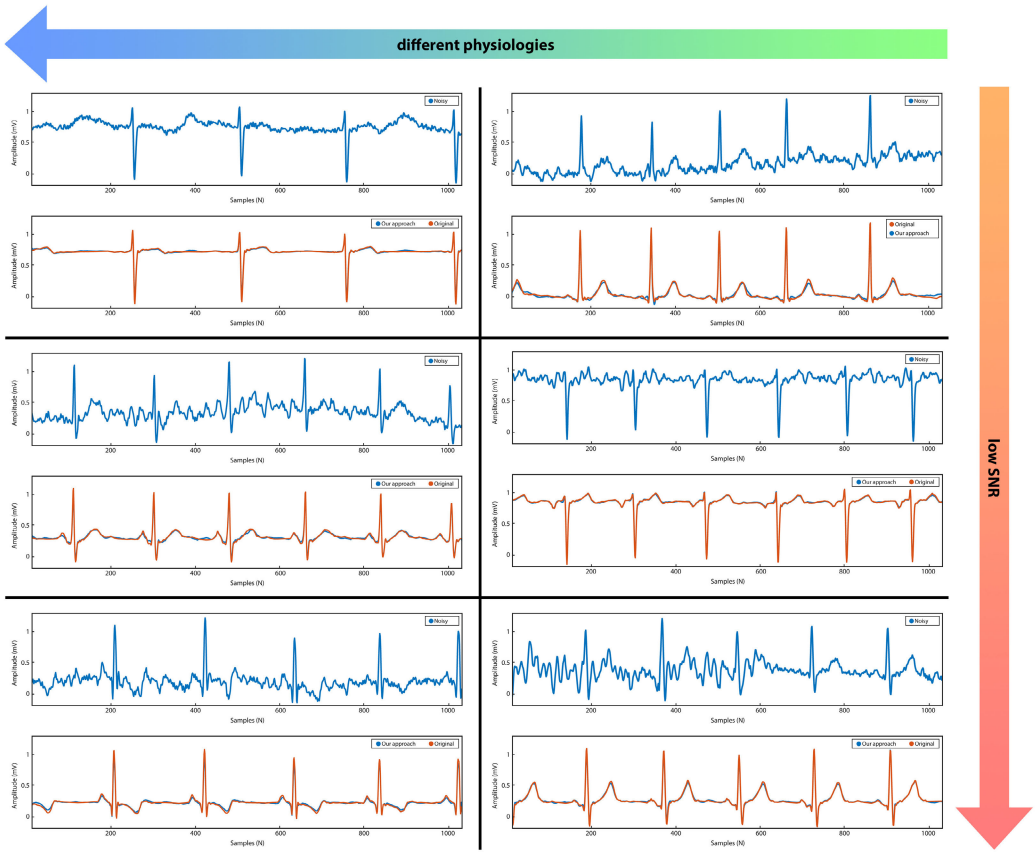


FIGURE 10. Six different heart physilogies with three level SNR input −1 dB, −2 dB, and −3 dB from our dataset that accurately denoised.

in equations (11), (12), and (13) cannot be possible without a reference signal.

3) COMPUTATIONAL COMPLEXITY AND LATENCY ANALYSIS

Computational complexity (CC) is one of the performance metrics to evaluate to what extent one DL-based method is complicated. Moreover, it measures how many computational resources are needed to run this approach [55]. In this regard, we assess the complexity with three criteria: mathematical modeling, number of parameters, and run time. Regarding our approach being based on 2D CNN, the simple way of thinking of this is that we have a $C \times V$ (grayscale) image and an $c \times v$ filter. Therefore, $O(c \cdot v)$ computations are needed for each pixel. Hence, the 2D convolution would have an approximate complexity of $O(C \cdot V \cdot c \cdot v)$ [56]. For calculating the complexity, it is important that images are padded or not. However, assuming that c and v are small compared to C and V. Thus, the differences should be negligible. On the other hand, one-dimensional (1D) CNN is the sum of the row-wise dot products of a filter $W \in \mathbb{R}^{k \times d}$ with a region matrix $Q \in \mathbb{R}^{k \times d}$, where k is the length of the filter and d is the depth dimension. If we have H layer 1D CNN, the CC of 1D CNN can be calculated by $O\left(H \cdot k \cdot M \cdot d^2\right)$ [56], where M is the length of the input. More details can be found in Table 3. According to Table 3, each stage of our approach has 346034 parameters, whereas DeepCEDNet has 2854914 parameters. Remarkably, DeepCEDNet is more complex than DeepRTSNet, with two stages in terms of the number of parameters. Following that, FCN-based DAE is less complex than the two others because it is based on 1D CNN, which evidently can be implemented with less complexity than 2D CNN. Regarding run times, the first stage of DeepRTSNet, 278.25 milliseconds (ms), is less complex than FCN-based DAE. However, for the ECG signals with high noise corruption that needed two-stage denoising, DeepRTSNet is more complex. It is noteworthy that this complexity is justifiable because one more stage of denoising is the only way to cancel highly noisy signals. Mathematical modeling indicates that calculating CC for DeepRTSNet and DeepCEDNet is based on the same formula. The only difference between them is the number of layers in each. Indeed, for comparing the 2D CNN-based approaches, more layers mean more complexity. Although 1D CNN has a less CC, its denoising performance is not comparable with 2D CNN-based methods, especially DeeRTSNet, as shown in Fig. 7, 8, 9, 10, 11, 12, 13, 14, and 15. Generally, more costs in terms of CC (more computational resources) are required for better and more accurate ECG denoising. The trade-off here is needed between the complexity and accuracy

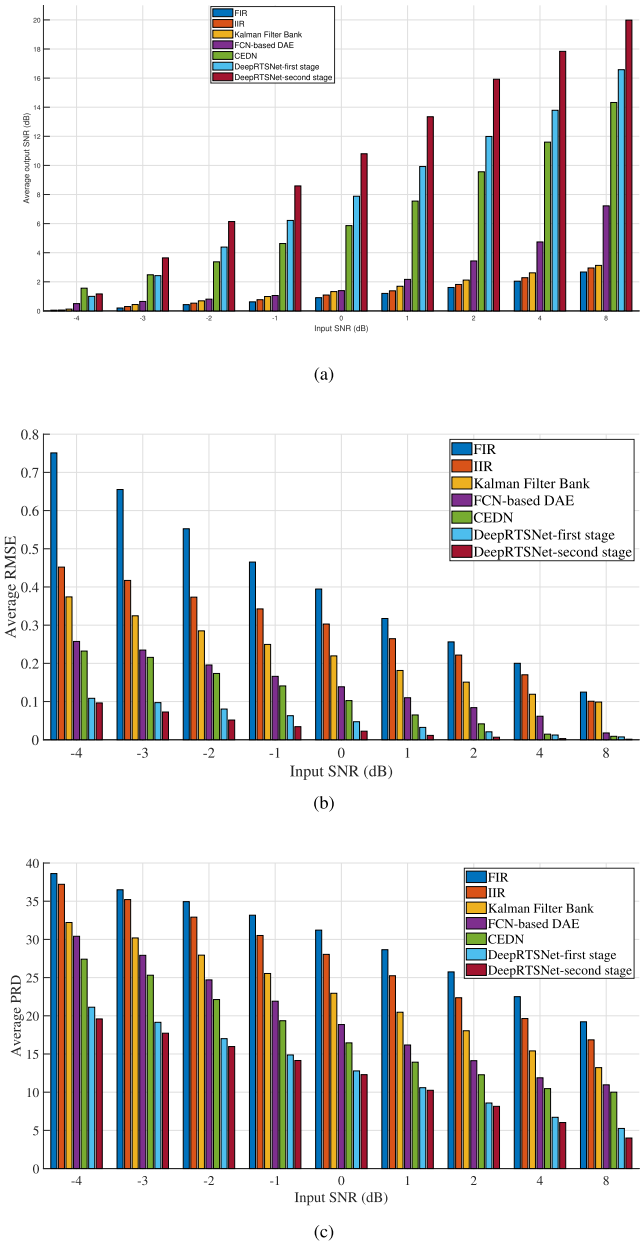


FIGURE 11. Quantitative comparison of denoising performance of DeepRTSNet, CEDN, FCN-based DAE, Kalman bank filter, unbiased FIR filter, and IIR filter in terms of average output SNR, RMSE, and PRD regarding nine different input SNR of -4, -3, -2, -1, 0, 1, 2, 4, and 8 dB. (a) Average output SNR. (b) Average output RMSE. (c) Average output PRD.

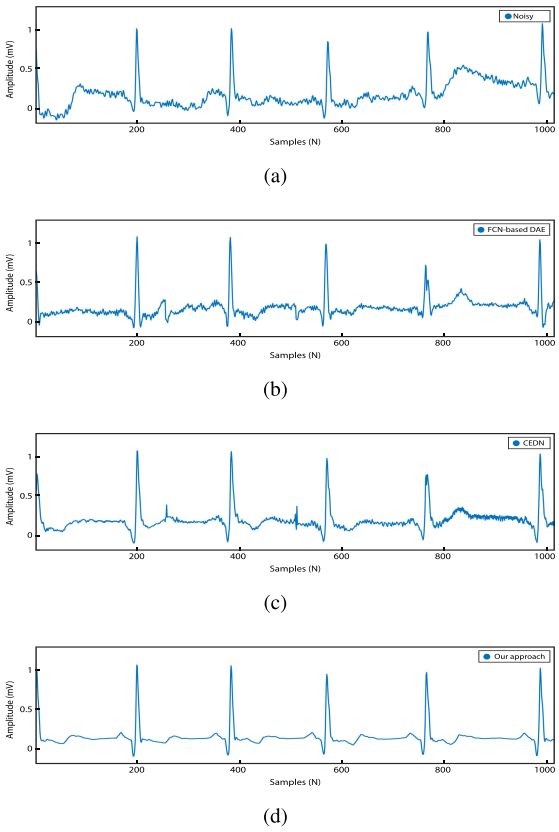


FIGURE 12. Qualitative denoising performance of DeepRTSNet for first subject signal obtained from I-Heart. (a) Noisy signal of first subject. (b) Denoised version of first subject by FCN-based DAE. (c) Denoised version of first subject by DeepCEDNet. (d) Denoised version of first subject by DeepRTSNet.

of the denoiser, which is well considered in implementing DeepRTSNet.

Calculating the latency of our process from recording to denoising (from I-Heart to server) can be affirmative. Firstly, it takes 30 seconds to record the data by I-Heart. As we mentioned, SIM800l is used as a communication module. Accordingly, SIM800l has a bit rate of 1200 bps - 115200 bps. From I-Heart to server, we have an average data rate of 58200 bps. I-Heart records 6000 samples, and their total data size is approximately 96000 bits, including additive headers and control bits. At the average rate of 58200 bits per second, 96000 bits of data are transferred from I-Heart to a server in approximately 1.64 seconds. We begin our denoising process by splitting the 6000 sample signal into 46 shorter signals with 256 samples, which takes 0.99 ms. It is noteworthy that the signals overlap by 128. Pre-processing involves transferring each short signal to a time-frequency presentation with the FrST formula, calculating the magnitude of FrST, and normalizing the magnitude. The server needs to process all of these parts in 0.99 ms. As soon as our magnitude denoiser receives the input, it begins producing the denoised magnitude part of FrST within 278.25 ms. The SNR improvement can be evaluated using the time domain form of the time-frequency presentation. It takes 7.97 ms to compute the inverse transform. To determine whether the signal needs further denoising, we calculate SNR improvement, which takes 52.64 ms. As a worst-case scenario, we would need to repeat the pre-processing part for the second network, but rather than calculating the magnitude, we would split the complex form of FrST into real and imaginary portions, which would take our server 1.99 ms. After that, the second network denoises the real and imaginary parts, and then the inverse transform will be used to calculate the clean signal

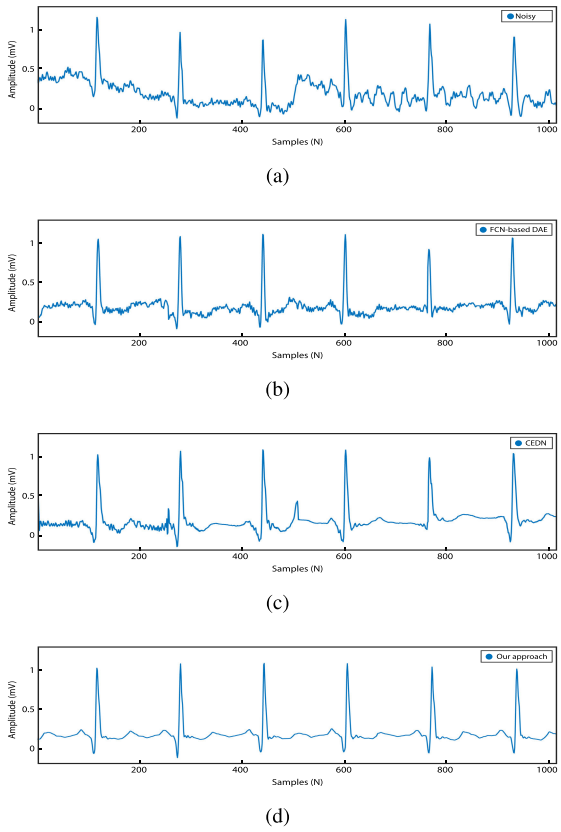


FIGURE 13. Qualitative denoising performance of DeepRTSNet for second subject signal obtained from I-Heart. (a) Noisy signal of second subject. (b) Denoised version of second subject by FCN-based DAE. (c) Denoised version of second subject by DeepCEDNet. (d) Denoised version of second subject by DeepRTSNet.

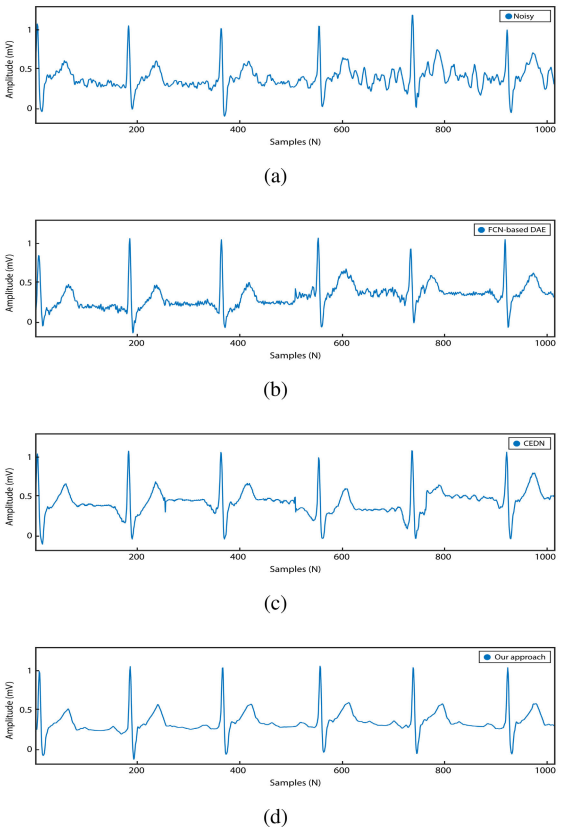


FIGURE 14. Qualitative denoising performance of DeepRTSNet for third subject signal obtained from I-Heart. (a) Noisy signal of third subject. (b) Denoised version of third subject by FCN-based DAE. (c) Denoised version of third subject by DeepCEDNet. (d) Denoised version of third subject by DeepRTSNet.

from the complex form of FrST, which takes 286.23 ms and 5.98 ms. Finally, transferring denoised ECG to a doctor's application takes 15 ms on average. Taking into account just one denoising stage, I-Heart recording to server denoising would take 31.98084 seconds. From recording to denoising, it would take 32.27504 seconds if they required two-stage noise cancellation. There is no significant difference between these two scenarios. It is deserved to be mentioned that we do not have any diseases or irregularities diagnoses. The main focus of DeepRTSNet is signal denoising, and we assume that disease diagnosing is done by the doctor. Finally, if the diseases are diagnosed, the doctor can take appropriate action to inform the patient to visit or refer to clinics.

IV. DISCUSSION

Interpreting the ECG signal is always a primary indicator of diagnosing cardiovascular diseases. Early diagnosing is compulsory for most diseases that can be so affirmative to prevent transforming into high-risk conditions and attempt to fast recovery. Early diagnosing needs continuous patient heart monitoring in the outer clinic. With emerging cellular IoMT technologies, this persistent observation is possible. By exploiting these portable gadgets, the patient's ECG signal can be monitored everywhere and at any time. Apart from the portability of these devices, the recorded signal carry several different noises due to their small size and low quality compared to clinical ones. Cloud-based denoising allows us to present a powerful denoising solution. In this paper, deep robust two-stage network, DeepRTSNet, is proposed based on a convolutional encoder-decoder in the time-frequency domain that can strongly realize data sparsity. Although existing learning-based models such as CEDN and FCN-based DAE have a good performance in dealing with clinical recorded signals consisting of standard noises, their efficiency and functionality are remarkably decreased in practical IoMT use cases. The superiority of our system can be addressed in four factors: (1) the FrST provides a higher resolution means that discrepancies between noise and valuable signal are apparent; (2) DeepRTSNet takes advantage of a two-stage convolutional encoder-decoder and perceptual approaches loss calculation; (3) the noise generation technique involving noises that are severely suppressed and changed the signal's morphology that it is gained more robustness in DeepRTSNet; (4) utilizing wide-range physiologies in our dataset that enables promising adaptiveness in DeepRTSNet. Generally, these elements made our

TABLE 3. Computational complexity comparison between DeepRTSNet, DeepCEDNet, and FCN-based DAE in terms of mathematical modeling, run times, and number of parameters.

	The first Stage of DeepRTSNet	The second Stage of DeepRTSNet	DeepCEDNet	FCN-based DAE
Number of Parameters	346034	346034	2854914	80223
Run Time	278.25 ms	286.23 ms	335.10 ms	321.14 ms
Mathematical Modeling	$O(H \cdot C \cdot V \cdot c \cdot v)$	$O(H \cdot C \cdot V \cdot c \cdot v)$	$O(H \cdot C \cdot V \cdot c \cdot v)$	$O\left(H \cdot k \cdot M \cdot d^2\right)$

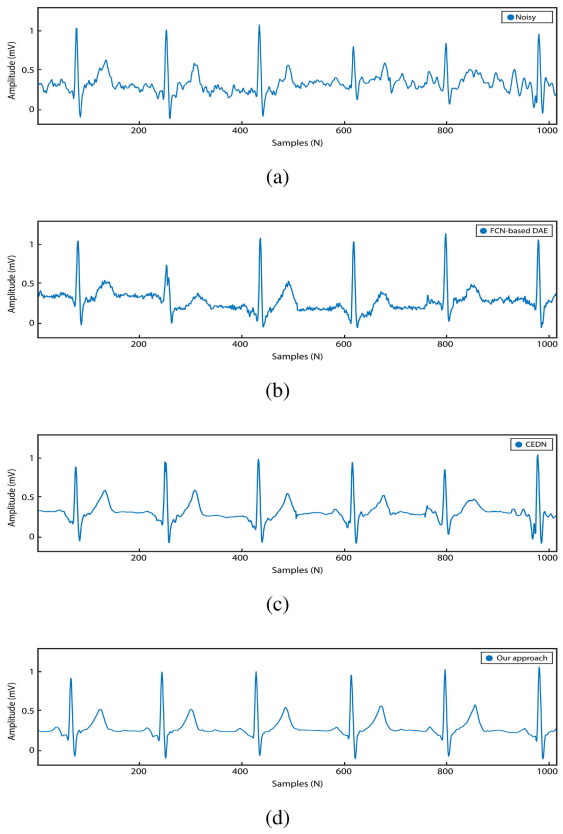


FIGURE 15. Qualitative denoising performance of DeepRTSNet for fourth subject signal obtained from I-Heart. (a) Noisy signal of fourth subject. (b) Denoised version of fourth subject by FCN-based DAE. (c) Denoised version of fourth subject by DeepCEDNet. (d) Denoised version of fourth subject by DeepRTSNet.

approach superior, with more SNR improvement and lower RMSE and PRD compared to learning and non-learning-based approaches.

Although DeepRTSNet adaptively and robustly cancels the wide-range noises in various physiologies, our approach still has some disadvantages. These weaknesses can be expressed as: (1) despite the fact that noise cancellation in two-stage is needed regarding very corrupted signal, it may lead to extra denoising time in real-time practical use cases; (2) performance in very noisy signal need more enhancement and lower suppressing that can be caused by insufficient resolution in the time-frequency domain, and other transformers should be analyzed; (3) mixing denoising and feature extraction can be more affirmative even though combining disease diagnosis; (4) change the DeepRTSNet architecture in order to introduce a new one suitable for edge computing instead of cloud computing in terms of computational complexity regarding our practical use cases. Considering these mentioned drawbacks will be helpful to enhance the DeepRTSNet and lead to exploiting it in different practical applications.

V. CONCLUSION AND FUTURE WORKS

The DeepRTSNet architecture presented in this paper uses deep learning to denoise ECG signals. It can identify the signal from the noise by learning a sparse distribution of data in the time-frequency domain. Empirical results reveal that DeepRTSNet was outstandingly robust against various noises similar to ECG waves (P, QRS, and T) and adaptively dealt with different heart physiologies. Furthermore, evaluation proved that the DeepRTSNet achieved a higher output SNR and gained lower RMSE and PRD compared to DeepCEDNet, FCN-based DAE, Kalman filter bank, FIR, and IIR filters. For Future works, a more efficient DL-based approach in terms of computational complexity and network input data format, a higher-resolution time-frequency transform for better noise separation, new DL models such as self-supervised learning, and one-stage iterative denoising architecture instead of two-stage will be taken into account in order to enhance the DeepRTSNet performance and efficiency.

REFERENCES

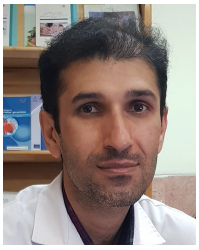
[1] A. Alwan, *Global Status Report on Noncommunicable Diseases 2010*. Geneva, Switzerland: World Health Organization, 2011.
[2] E. J. Benjamin et al., "Heart disease and stroke statistics—2017 update: A report from the American heart association," *Circulation*, vol. 135, no. 10, pp. e146–e603, 2017.
[3] R. M. John, U. B. Tedrow, B. A. Koplan, C. M. Albert, L. M. Epstein, M. O. Sweeney, A. L. Miller, G. F. Michaud, and W. G. Stevenson, "Ventricular arrhythmias and sudden cardiac death," *Lancet*, vol. 380, no. 9852, pp. 1520–1529, Oct. 2012.
[4] S. Chen, H. Xu, D. Liu, B. Hu, and H. Wang, "A vision of IoT: Applications, challenges, and opportunities with China perspective," *IEEE Internet Things J.*, vol. 1, no. 4, pp. 349–359, Aug. 2014.
[5] S. U. Amin and M. S. Hossain, "Edge intelligence and Internet of Things in healthcare: A survey," *IEEE Access*, vol. 9, pp. 45–59, 2021.
[6] M. A. Khan, "An IoT framework for heart disease prediction based on MDCNN classifier," *IEEE Access*, vol. 8, pp. 34717–34727, 2020.
[7] W.-L. Chin, C.-C. Chang, C.-L. Tseng, Y.-Z. Huang, and T. Jiang, "Bayesian real-time QRS complex detector for healthcare system," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5540–5549, Jun. 2019.
[8] P. Xiong, H. Wang, M. Liu, S. Zhou, Z. Hou, and X. Liu, "ECG signal enhancement based on improved denoising auto-encoder," *Eng. Appl. Artif. Intell.*, vol. 52, pp. 194–202, Jun. 2016.
[9] M. Merone, P. Soda, M. Sansone, and C. Sansone, "ECG databases for biometric systems: A systematic review," *Exp. Syst. Appl.*, vol. 67, pp. 189–202, Jan. 2017.
[10] M. Wasimuddin, K. Elleithy, A. Abuzneid, M. Faezipour, and O. Abuzaghleh, "Stages-based ecg signal analysis from traditional signal processing to machine learning approaches: A survey," *IEEE Access*, vol. 8, pp. 177782–177803, 2020.
[11] Y. Luo, R. H. Hargraves, A. Belle, O. Bai, X. Qi, K. R. Ward, M. P. Pfaffenberger, and K. Najarian, "A hierarchical method for removal of baseline drift from biomedical signals: Application in ECG analysis," *Sci. World J.*, vol. 2013, pp. 1–10, May 2013.

[12] V. de Pinto, "Filters for the reduction of baseline wander and muscle artifact in the ECG," *J. Electrocardiol.*, vol. 25, pp. 40–48, Jan. 1992.
[13] Z. Wang, C. M. Wong, J. N. da Cruz, F. Wan, P.-I. Mak, P. U. Mak, and M. I. Vai, "Muscle and electrode motion artifacts reduction in ECG using adaptive Fourier decomposition," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2014, pp. 1456–1461.
[14] R. G. Afkhami, G. Azarnia, and M. A. Tinati, "Cardiac arrhythmia classification using statistical and mixture modeling features of ECG signals," *Pattern Recognit. Lett.*, vol. 70, pp. 45–51, Jan. 2016.
[15] C. Watford, "Understanding ECG filtering," in *Proc. EMS 12-Lead*, Hilton Head Island, SC, USA, Mar. 2014, pp. 1–6. [Online]. Available: https://www.rigacci.org/wiki/lib/exe/fetch.php/tecnica/misc/ecg90a/understanding-ecg-filtering.pdf
[16] S. K. Berkaya, A. K. Uysal, E. S. Gunal, S. Ergin, S. Gunal, and M. B. Gulmezoglu, "A survey on ECG analysis," *Biomed. Signal Process. Control*, vol. 43, pp. 216–235, May 2018.
[17] A. Rasti-Meymandi and A. Ghaffari, "A deep learning-based framework for ECG signal denoising based on stacked cardiac cycle tensor," *Biomed. Signal Process. Control*, vol. 71, Jan. 2022, Art. no. 103275.
[18] D. Yoon, H. S. Lim, K. Jung, T. Y. Kim, and S. Lee, "Deep learning-based electrocardiogram signal noise detection and screening model," *Healthcare Inform. Res.*, vol. 25, no. 3, pp. 201–211, 2019.
[19] H.-T. Chiang, Y.-Y. Hsieh, S.-W. Fu, K.-H. Hung, Y. Tsao, and S.-Y. Chien, "Noise reduction in ECG signals using fully convolutional denoising autoencoders," *IEEE Access*, vol. 7, pp. 60806–60813, 2019.
[20] P. Bing, W. Liu, and Z. Zhang, "DeepCEDNet: An efficient deep convolutional encoder–decoder networks for ECG signal enhancement," *IEEE Access*, vol. 9, pp. 56699–56708, 2021.
[21] D. Hasan and A. Ismaeel, "Designing ECG monitoring healthcare system based on Internet of Things blynk application," *J. Appl. Sci. Technol. Trends*, vol. 1, no. 3, pp. 106–111, Jul. 2020.
[22] H. Djelouat, M. Al Disi, I. Boukhenoufa, A. Amira, F. Bensaali, C. Kotronis, E. Politi, M. Nikolaidou, and G. Dimitrakopoulos, "Real-time ECG monitoring using compressive sensing on a heterogeneous multicore edge-device," *Microprocessors Microsystems*, vol. 72, Feb. 2020, Art. no. 102839.
[23] S. Lee and D. Park, "A real-time abnormal beat detection method using a template cluster for the ECG diagnosis of IoT devices," *Hum.-Centric Comput. Inf. Sci.*, vol. 11, pp. 1–17, Jan. 2021.
[24] K. Arquilla, A. Webb, and A. Anderson, "Textile electrocardiogram (ECG) electrodes for wearable health monitoring," *Sensors*, vol. 20, no. 4, p. 1013, Feb. 2020.
[25] M. E. Hassin and R. Khan, "NeuroSpy: A low-cost portable IoT enabled EEG and ECG data processor," in *Proc. 2nd Int. Conf. Robot., Electr. Signal Process. Techn. (ICREST)*, Jan. 2021, pp. 225–229.
[26] R. Morello, F. Ruffa, I. Jablonski, L. Fabbiano, and C. De Capua, "An IoT based ECG system to diagnose cardiac pathologies for healthcare applications in smart cities," *Measurement*, vol. 190, Feb. 2022, Art. no. 110685.
[27] A. Bajaj and S. Kumar, "QRS complex detection using fractional stockwell transform and fractional stockwell Shannon energy," *Biomed. Signal Process. Control*, vol. 54, Sep. 2019, Art. no. 101628.
[28] A. Mishra, S. S. Sahu, R. Sharma, and S. K. Mishra, "Denoising of electrocardiogram signal using S-transform based time–frequency filtering approach," *Arabian J. Sci. Eng.*, vol. 46, no. 10, pp. 9515–9525, Oct. 2021.
[29] A. Bajaj and S. Kumar, "A robust approach to denoise ECG signals based on fractional stockwell transform," *Biomed. Signal Process. Control*, vol. 62, Sep. 2020, Art. no. 102090.
[30] I. Jekova, I. Iliev, and S. Tabakov, "Application of stockwell transform and Shannon energy for pace pulses detection in a single-lead ECG corrupted by EMG artifacts," *Appl. Sci.*, vol. 10, no. 21, p. 7505, Oct. 2020.
[31] P. Bing, W. Liu, Z. Wang, and Z. Zhang, "Noise reduction in ECG signal using an effective hybrid scheme," *IEEE Access*, vol. 8, pp. 160790–160801, 2020.
[32] C. Lastre-Dominguez, Y. S. Shmaliy, O. Ibarra-Manzano, and M. Vazquez-Olguin, "Denoising and features extraction of ECG signals in state space using unbiased FIR smoothing," *IEEE Access*, vol. 7, pp. 152166–152178, 2019.
[33] H. Shi, R. Liu, C. Chen, M. Shu, and Y. Wang, "ECG baseline estimation and denoising with group sparse regularization," *IEEE Access*, vol. 9, pp. 23595–23607, 2021.
[34] X. Wang, Y. Zhou, M. Shu, Y. Wang, and A. Dong, "ECG baseline wander correction and denoising based on sparsity," *IEEE Access*, vol. 7, pp. 31573–31585, 2019.
[35] H. D. Hesar and M. Mohebbi, "An adaptive Kalman filter bank for ECG denoising," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 1, pp. 13–21, Jan. 2021.
[36] Z. Zhao, C. Liu, Y. Li, Y. Li, J. Wang, B.-S. Lin, and J. Li, "Noise rejection for wearable ECGs using modified frequency slice wavelet transform and convolutional neural networks," *IEEE Access*, vol. 7, pp. 34060–34067, 2019.
[37] L. Qiu, W. Cai, M. Zhang, W. Zhu, and L. Wang, "Two-stage ECG signal denoising based on deep convolutional network," *Physiological Meas.*, vol. 42, no. 11, Nov. 2021, Art. no. 115002.
[38] N. Reljin, J. Lazaro, M. B. Hossain, Y. S. Noh, C. H. Cho, and K. H. Chon, "Using the redundant convolutional encoder–decoder to denoise QRS complexes in ECG signals recorded with an armband wearable device," *Sensors*, vol. 20, no. 16, p. 4611, Aug. 2020.
[39] S. Nurmaini, A. Darmawahyuni, A. N. Sakti Mukti, M. N. Rachmatullah, F. Firdaus, and B. Tutuko, "Deep learning-based stacked denoising and autoencoder for ECG heartbeat classification," *Electronics*, vol. 9, no. 1, p. 135, Jan. 2020.
[40] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, Jun. 2000, doi: 10.1161/01.CIR.101.23.E215.
[41] R. G. Stockwell, L. Mansinha, and R. P. Lowe, "Localization of the complex spectrum: The S transform," *IEEE Trans. Signal Process.*, vol. 44, no. 4, pp. 998–1001, Apr. 1996.
[42] H. M. Ozaktas, O. Arikan, M. A. Kutay, and G. Bozdagt, "Digital computation of the fractional Fourier transform," *IEEE Trans. Signal Process.*, vol. 44, no. 9, pp. 2141–2150, Sep. 1996.
[43] C. Candan, M. A. Kutay, and H. M. Ozaktas, "The discrete fractional Fourier transform," *IEEE Trans. Signal Process.*, vol. 48, no. 5, pp. 1329–1337, May 2000.
[44] Y. Yu, T. Xu, and P. Yang, "Analysis of the fractional S transform," in *Advances in Acoustic Emission Technology*. Cham, Switzerland: Springer, 2017, pp. 75–84.
[45] D.-P. Xu and K. Guo, "Fractional S transform—Part 1: Theory," *Appl. Geophys.*, vol. 9, no. 1, pp. 73–79, 2012.
[46] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with dnoising autoencoders," in *Proc. 25th Int. Conf. Mach. Learn. (ICML)*, 2008, pp. 1096–1103.
[47] Ilopezfr. (2022). *Github*. [Online]. Available: https://github.com/ilopezfr/image-superres.git
[48] X. Mao, C. Shen, and Y.-B. Yang, "Image restoration using very deep convolutional encoder–decoder networks with symmetric skip connections," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 29, 2016, pp. 1–9.
[49] S. ElRamly, F. Newagy, H. Yousry, and A. Elezabi, "Novel modified energy detection spectrum sensing technique for FM wireless microphone signals," in *Proc. IEEE 3rd Int. Conf. Commun. Softw. Netw.*, May 2011, pp. 59–63.
[50] M. Kaur and B. Singh, "Comparison of different approaches for removal of baseline wander from ECG signal," in *Proc. Int. Conf. Workshop Emerg. Trends Technol. (ICWET)*, 2011, pp. 1290–1294.
[51] K. S. Kumar, B. Yazdanpanah, and P. R. Kumar, "Removal of noise from electrocardiogram using digital FIR and IIR filters with various methods," in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, Apr. 2015, pp. 0157–0162.
[52] H. Zhang, C. Ma, V. Pazzi, Y. Zou, and N. Casagli, "Microseismic signal denoising and separation based on fully convolutional encoder–decoder network," *Appl. Sci.*, vol. 10, no. 18, p. 6621, Sep. 2020.
[53] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1492–1500.
[54] P. Aghaomidi, A. Mohammadisarab, J. Mazloum, M. A. Akbarzadeh, M. Orooji, N. Mokari, and H. Yanikomeroglu, "DeepRTSNet: Deep robust two-stage network for practical wearable ECG recorders denoising under severe noise corruption and various physiologies," *IEEE Dataport*, Jun. 23, 2022. [Online]. Available: https://ieee-dataport.org/documents/deeprtsnet-deep-robust-two-stage-network-practical-wearable-ecg-recorders-denoising-under, doi: 10.21227/tzgg-fw33.
[55] A. U. Haq, J. P. Li, B. L. Y. Agbley, C. B. Mawuli, Z. Ali, S. Nazir, and S. U. Din, "A survey of deep learning techniques based Parkinson's disease recognition methods employing clinical data," *Exp. Syst. Appl.*, vol. 208, Dec. 2022, Art. no. 118045.
[56] J. Park, J. Lee, and D. Sim, "Low-complexity CNN with 1D and 2D filters for super-resolution," *J. Real-Time Image Process.*, vol. 17, no. 6, pp. 2065–2076, Dec. 2020.

POORYA AGHAOMIDI received the B.Sc. degree in biomedical engineering from the Sahand University of Technology (SUT), Tabriz, Iran, in 2020. He is currently pursuing the M.Sc. degree in biomedical engineering—bioelectric from Tarbiat Modares University (TMU), Tehran, Iran. He is the last year's student, and his thesis is about AI-based recognition of driver's distractions. His research interests include neuroscience and artificial intelligence in different biomedical engineering applications, including computer-aided diagnostics, and signal and image denoising.

AMIR MOHAMMADISARAB (Member, IEEE) received the B.Sc. degree (Hons.) in electrical and electronic engineering from the University of Kurdistan (UOK), Kurdistan, Iran, in 2020. He is currently pursuing the M.Sc. degree (Hons.) in telecommunications system with Tarbiat Modares University, Tehran, Iran. He is a Senior Master's Student, and his thesis is about AI-based Resource Management for Critical IoT Use Cases. His research interests include wireless communication networks (B5G/6G), intelligent reflecting surface (IRS), non-terrestrial networks (NTNs) communications, the Internet of Things (IoT), deep reinforcement learning (DRL), resource allocation, and optimization theory.

JALIL MAZLOUM was born in Tehran, Iran, 1974. He received the M.Sc. degree in biomedical engineering from the Amirkabir University of Technology, Tehran, and the Ph.D. degree in electrical engineering from Shahid Beheshti University, Tehran. Since March 2000, he has been a Research Fellow and a Teaching Assistant. His research interests include high frequency circuit design and intelligent signal processing.



MOHAMMAD ALI AKBARZADEH received the M.D. degree from the Shiraz University of Medical Sciences (SUMS), in 2004, and was the high degree (4th rank) in Iranian National Cardiology Board Exam, in 2010. He is currently an Associate Professor and a Fellowship of invasive electrophysiology with the Cardiology Department, Shahid Beheshti University of Medical Sciences (SBMU), Tehran, where he is also a Cardiologist. He is a Research Assistant with the Cardiovascular Research Center, SBMU, and a Consultant to Iranian Legal Medicine Organizations, since 2015.

MAHDI OROOJI was born in Tehran, Iran, in 1980. He received the Bachelor of Science degree in electrical and computer engineering from the University of Tehran, Iran, in 2003, and the Master of Science and Ph.D. degrees from Louisiana State University, Baton Rouge, LA, USA, in 2011 and 2013, respectively. In 2013, he joined the Center for Computational Imaging and Personalized Diagnostics (CCIPD), Case Western Reserve University, Cleveland, OH, USA, as a Research Associate. He joined the Department of Electrical and Computer Engineering, Tarbiat Modares University (TMU), Tehran, as an Assistant Professor of biomedical engineering, in 2016. Since 2021, he has been with the Department of Electrical and Computer Engineering, University of California, Davis, as a Visiting Professor. His current research interests include artificial intelligence and machine learning in different applications of biomedical engineering, including computer-aided diagnostics and prognostics.

NADER MOKARI (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from Tarbiat Modares University, Tehran, Iran, in 2014. He joined the Department of Electrical and Computer Engineering, Tarbiat Modares University, as an Assistant Professor, in October 2015. He is currently an Associated Professor with the Department of Electrical and Computer Engineering, Tarbiat Modares University. In recent years, his research has been funded by Iranian Mobile Telecommunication Companies, Iranian National Science Foundation. His research interests include many aspects of wireless technologies with a special emphasis on wireless networks. His thesis received the IEEE Outstanding Ph.D. Thesis Award. He received the Best Paper Award at ITU K-2020. He has been elected as an IEEE Exemplary Reviewer in 2016 by IEEE Communications Society. He is on the Editorial Board of the IEEE TRANSACTIONS ON COMMUNICATIONS. He was also involved in a number of large-scale network design and consulting projects in the telecom industry.

HALIM YANIKOMEROGLU (Fellow, IEEE) is currently a Professor with the Department of Systems and Computer Engineering, Carleton University, Ottawa, Canada. His research group has made substantial contributions to 4G and 5G wireless technologies. His group's current focus is the aerial and satellite networks for the 6G and beyond-6G era. His extensive collaboration with industry resulted in 39 granted patents. He is a Fellow of the Engineering Institute of Canada (EIC) and the Canadian Academy of Engineering (CAE). He has received several awards for his research, teaching, and service. He is a Distinguished Speaker for both IEEE Communications Society and IEEE Vehicular Technology Society.