Deep Learning Models for Magnetic Cardiography Edge Sensors Implementing Noise Processing and Diagnostics

SADMAN SAKIB[1,2], MOSTAFA M. FOUDA[3,4], (Senior Member, IEEE),
MUFTAH AL-MAHDAWI[5,6], (Senior Member, IEEE), ATTAYEB MOHSEN[7],
MIKIHIKO OOGANE[5,6,8], YASUO ANDO[5,6,8], AND
ZUBAIR MD. FADLULLAH[1,2], (Senior Member, IEEE)

[1]Department of Computer Science, Lakehead University, Thunder Bay, ON P7B 5E1, Canada
[2]Thunder Bay Regional Health Research Institute (TBRHRI), Thunder Bay, ON P7B 7A5, Canada
[3]Department of Electrical and Computer Engineering, College of Science and Engineering, Idaho State University, Pocatello, ID 83209, USA
[4]Department of Electrical Engineering, Faculty of Engineering at Shoubra, Benha University, Cairo 11672, Egypt
[5]Center for Science and Innovation in Spintronics (Core Research Cluster), Tohoku University, Sendai 980-8577, Japan
[6]Center for Spintronics Research Network, Tohoku University, Sendai 980-8577, Japan
[7]Artificial Intelligence Center for Health and Biomedical Research (ArCHER), National Institutes of Biomedical Innovation, Health and Nutrition (NIBIOHN), Osaka 567-0085, Japan
[8]Department of Applied Physics, Tohoku University, Sendai 980-8579, Japan

Corresponding author: Mostafa M. Fouda (mfouda@ieee.org)

ABSTRACT Remote health monitoring has become a necessity due to reduced healthcare access resulting from pandemic lockdowns and the increasing aging population. Electrocardiography (ECG) is the standard for cardiac monitoring and arrhythmia identification, but it is inconvenient for long-time remote monitoring. Recently, Magnetocardiography (MCG) sensors that operate at room temperature became available based on spintronic sensors. However, MCG analysis is affected by the low-frequency noise present at the sensors. In this paper, we present an artificial intelligence (AI)-aided multi-model pipeline combining two AI architectures, defined as model-M1 and model-M2, targeted for ultra-edge Internet of Things (IoT) sensors to simulate arrhythmia detection. Model-M1 is a denoising preprocessor based on a sliding-window assisted deep-learning (DL) model. We investigate various methods to achieve high accuracy with lightweight computation. Model-M2 is a lightweight DL model that analyzes denoised ECG output from model-M1 to identify arrhythmia. We use multiple publicly available clinically annotated datasets to evaluate our proposal. We find that denoising by model-M1 retains the features, which assist the model-M2 in achieving high classification accuracy, compared to using a conventional moving average filter. This AI pipeline architecture is promising for privacy-preserving ultra-edge medical sensing devices.

INDEX TERMS Remote health monitoring, arrhythmia, Internet of Things (IoT), electrocardiogram (ECG), magnetocardiography (MCG), deep learning (DL), spintronic sensor, convolutional neural network (CNN), medical analytics.

I. INTRODUCTION

With the recent advancement of the Internet of Things (IoT), the Internet of Medical Things (IoMT), and Artificial Intelligence (AI) the remote health monitoring has significantly

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

evolved [1], [2]. Even with the massive technological push towards smart and connected healthcare systems, health monitoring in remote/non-clinical settings for an extended period still exhibits various challenges. Numerous wearable devices and smartphone-based applications are now available to observe heartbeat variability and even electrocardiography (ECG). Although these technologies are non-invasive,
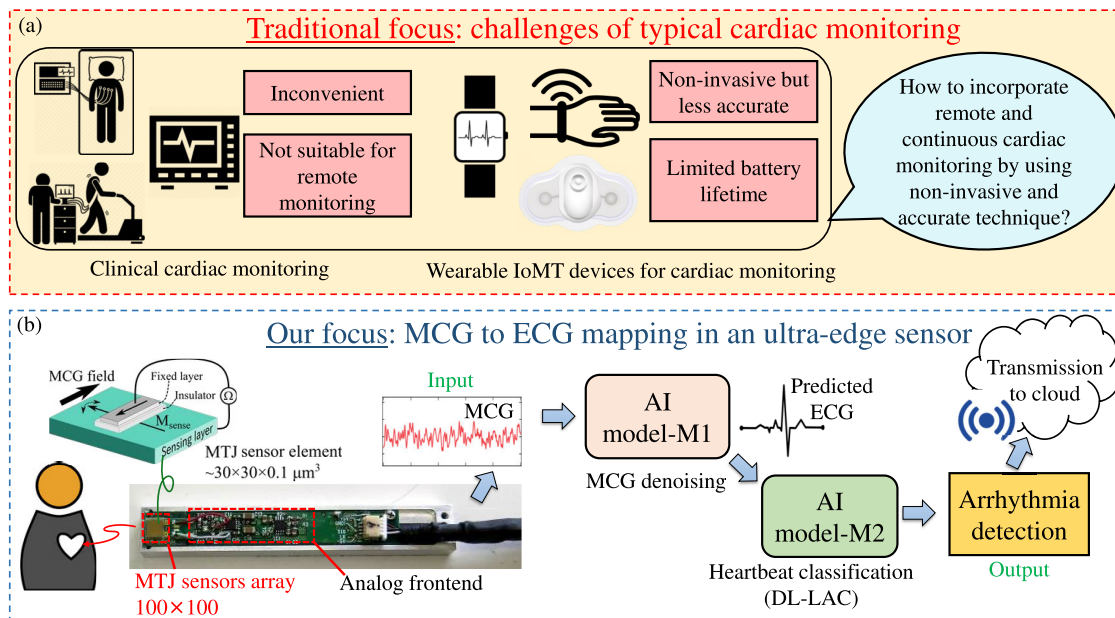
FIGURE 1. (a) Traditional cardiac monitoring, and (b) our proposed AI-aided MCG monitoring system for cardiac monitoring, including a schematic of the MTJ sensor and the outline of a lightweight AI pipeline.

their diagnosis capability is not up to the mark when compared to the standard clinical-grade ECG machines. Cardiac well-being is a vital concern globally, and detecting heart abnormalities such as arrhythmia requires monitoring in a non-intrusive way and for a long time [3]. We need to shift our focus on cardiac activity monitoring in a non-intrusive approach and analyze the cardiac status at the IoT edge level.

Cardiovascular diseases (CVDs) are still the leading cause of death worldwide, accounting for approximately 30% of all deaths, and according to American Heart Association (AHA), CVD expenditures will drastically increase in the coming years [4]. Irregular heartbeat, cardiac arrhythmia (CA), is a dominant cause of CVD [5]. Arrhythmia can happen suddenly and needs urgent medical care [6]. Hence, we picked the use-case of cardiac arrhythmia (CA) classification from a diverse set of use cases. With multiple waves of the novel coronavirus disease (COVID-19) [7], [8], the need for remote cardiac monitoring or arrhythmia detection has become even more critical than before [9], [10]. Access to healthcare got reduced in many regions. For example, 40% of adults in the USA have decreased access to regular care, even with patients at cardiac underlying conditions risks [11]. A similar 20–40% decrease in out-patient visits is also observed in Japan [12]. Hence, continuous cardiac monitoring is vital in remote/non-clinical healthcare settings. For arrhythmia detection, ECG is usually employed [13]. The typical wearable technology-based heart monitoring is not accurate enough, and the standard multi-lead-based clinically-graded ECG is non-invasive, but not suitable for continuous remote monitoring (Fig. 1(a)). Magnetocardiography (MCG) monitoring potentially offers a more direct

mapping to cardiac activity, and it is less intrusive compared to ECG [14]. Conventional MCG devices are based on superconducting sensors that require cryogenic cooling, barring them from wide adoption. Recently, MCG sensors based on spintronic magnetoresistive sensors were demonstrated by various groups, including ours [15]–[18]. The basis of the magnetic field detection is the change of resistance, due to quantum tunneling effect through a thin insulator between two magnetic layers (inset of Fig. 1(b)), thus named magnetic tunnel junction (MTJ) sensors [19], [20]. Ultra-sensitive MTJ sensors were developed to measure both of the cardiac and brain magnetic signals, *i.e.* MCG and magnetoencephalography (MEG), at room temperature [15], [18], [21]. The R-peak in the MCG was observed without using averaging. Also, the significant QRS complex was identified with a high signal-to-noise ratio by averaging for a few seconds. A differential benefit of MCG is the high spatial resolution in real-time, compared to ECG, which can potentially improve cardiac diagnosis [14], [17]. The MTJ sensors offer a low-power operation at room temperature, and they can be incorporated inside integrated circuits on silicon wafers. With the aid of IoT systems, cost-effective spintronic-based point-of-care (POC) systems utilizing wireless technologies to deliver data efficiently are an effective solution with little requirement for maintenance [22], at diverse regions with dense populations or under-served remote regions.

The use of remote IoT medical devices also entails the classification of CVDs from cardiograph data. Sophisticated AI models are proposed in different studies that use ECG data. It is critical to identify CVDs, such as CA, as early as possible, and perpetual ECG monitoring is needed. The traditional Holter machine-based approach of long-term

ECG monitoring is intrusive and costly, and it interferes with subjects' regular lives. Innovative portable sensors and machine learning (ML) algorithms that provide near-real-time diagnosis and developments in cardiovascular monitoring technologies can supply personalized health care services [23]. Cloud-based AI analytics is used to address this problem and incorporate some degree of automation into the ECG monitoring system, in which the ECG signal is typically transferred employing wireless communication techniques (e.g., Bluetooth, 3G/4G, Zigbee, Wi-Fi) to the IoT system for cloud-based data analysis [24], [25]. In the existing literature, various DL-based methods for detecting heart arrhythmia have been employed where most approaches focus on the utilization of ECG signals as data [26], [27]. In terms of the DL-based architecture, variations on deep convolution neural network (CNN) and recurrent neural network (RNN) models were employed to identify and classify cardiac arrhythmia or irregular heartbeats from ECG traces [28]–[34]. However, these techniques are not resource-efficient as they induce high computational complexity, especially when deployed in an edge sensor [35], [36]. Furthermore, these ECG-based monitoring approach is not suitable for long-term remote non-invasive monitoring of subject's cardiac acitivity.

In this paper, we focus on proposing the remote monitoring modality of magnetocardiography (MCG), replacing ECG remote monitoring by using spintronic MTJ sensors, and how to incorporate lightweight AI models for ultra-edge analytics. The contributions are outlined as follows.

1) We propose a system to collect the MCG data in a portable and non-intrusive manner utilizing spintronic magnetic tunnel junction (MTJ) sensors. We refer to this conceptualized system as the ultra-edge sensor/node, which would operate at room temperature in a remote cardiac monitoring environment, as depicted in Fig. 1.

2) We identify the key technical challenges associated with the conceptualized MCG monitoring ultra-edge node for continuous cardiac monitoring. Most of the information of the cardiac activity is skewed to lower frequency side within $f = 0.1$–30 Hz, where f is the spectral frequency. At this very low frequency band, electronic and magnetic noise sources become significant. Most significant low-frequency noise source is of magnetic origin arising in the sensors, which exhibits as time-correlated flicker noise [21], [37]. The power-spectral density (PSD) of this noise is inversely proportional to the spectral frequency (i.e., $1/f$). Therefore, cardiac activity has the same spectral features as the sensor noise, complicating processing and diagnosis. The separation between the noise and relevant signal becomes a formidable challenge, which conventional linear filters cannot solve. Popular noise filtering methods fail to efficiently treat the $1/f$ noise because of the spectral similarity of the noise and heart-generated signal [38]. We address the denoising of MCG data by exploring various deep learning (DL)-based techniques

to reconstruct ECG traces. We analyze three variations of the DL-based methods, namely a customized convolutional neural network (CNN) that includes temporal information, a combination of the custom CNN and a gated recurrent unit (GRU) layer [39], and a combination of the custom CNN and a long short-term memory (LSTM) layer.

3) After minimizing the noise from the MCG, we utilize the AI-aided lightweight Deep Learning-based Lightweight Arrhythmia Classification (DL-LAC) model designed in one of our previous works, to detect arrhythmia heartbeats [40]. Thereby, as manifested in Fig. 1, our focus in this work is to combine the pipeline of DL-based MCG denoising to obtain ECG, and by analyzing ECG, arrhythmia can be detected with high efficiency and accuracy. We employ multiple publicly available datasets in order to validate the performance of our proposed methodology via conducting comprehensive computer-based simulations. The proposed method can be a proof-of-concept for remote cardiac arrhythmia monitoring via analyzing MCG data collected in a non-invasive manner.

The remainder of this paper is organized as follows. Our proposed AI models and AI methodology pipeline are presented in section II. Section III manifests the algorithm for the considered arrhythmia classification task by analyzing MCG. Via extensive experimental simulations, the performance of our proposal is evaluated in section IV. Finally, section V concludes the manuscript.

II. PROPOSED SYSTEM MODEL
In this section, we envision an AI-enabled multi-model pipeline that can identify cardiac arrhythmia by analyzing MCG signals (Fig. 1(b)). First, we develop a preprocessing technique for the MCG temporal data, so that the signals are organized and passed to the convolution layers of a one-dimensional (1-D) CNN model. Then, we investigate how to build, train, and fine-tune a customized model (referred to as model-M1) to mitigate the $1/f$ noise from the MCG signal, so that the trained model can be deployed to the MTJ sensor. The motivation for using DL is to explore the automatic search for the optimal filtering techniques, which can be applied to various other applications. Lastly, we transmit the noise-eliminated ECG obtained from model-M1 to the DL-LAC model (referred to as model-M2). The model-M2 analyzes the data to produce the heartbeat classes for arrhythmia detection.

A. DATA PREPARATION AND PREPROCESSING STAGE
We consider noisy MCG data as the input and the corresponding ECG data as the labels for the training data. The detailed preprocessing and random sequencing of MCG segments for overfitting resilience to generate the training dataset is depicted in Fig. 2. In comparison to the traditional input characterization of a convolution layer whereby the whole MCG training data would be passed to
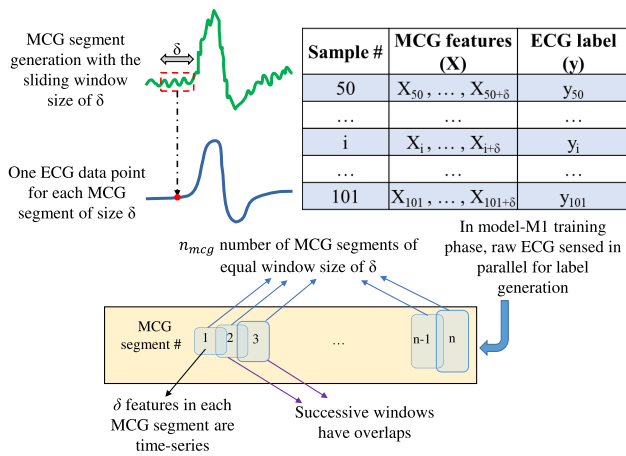
FIGURE 2. Data preprocessing methodology for non-intrusive MCG to ECG noise minimization before passing to model-M1.

FIGURE 3. The proposed architecture of the model-M1 leveraging custom CNN structure for the considered MCG denoising task.

the CNN for training, we divide an MCG cycle into n_{mcg} number of smaller MCG segments, each with a window size of δ. Each MCG segment of size δ is mapped to one ECG sample point during the training phase. This chopping mapping mechanism from input (MCG) to output (ECG) assisted the model-M1 in obtaining better performance. Our systematic investigation reveals that the MCG segments result in data overfitting when sequentially placed in the training dataset. Therefore, the split MCG segments are arbitrarily provided to the CNN layer via shuffling the training segments. However, in the test/inference phase, the MCG segments are given sequentially as input to the trained model-M1 to obtain a real-time mapping of a given MCG segment to its corresponding ECG representation. Each i_{th} MCG segment, denoted by $[X_i, X_{i+1}, \ldots, X_{i+\delta}]$ is passed to the model-M1 for training and corresponding ECG denoised sample, denoted by y_i is obtained as the output.

B. PROPOSED CNN MODEL STRUCTURE FOR MCG DENOISING (MODEL-M1)

A one-dimensional custom CNN is developed and is used as a DL-based solution as model-M1. The proposed CNN model can be outlined as a shallow and simple combination of convolution and fully connected layers. Fig. 3 represents the high-level structure of the proposed model. Here, the model receives the raw MCG segment as input and generates the corresponding ECG sample as output which is later passed to model-M2. The convolution (1D) layer receives the raw MCG segments and extracts relevant features from the MCG, which is then passed through the fully-connected dense layer. The dense layer is followed by the output layer of model-M1 with a linear activation function. The output layer produces one ECG sample prediction for each MCG input segment.

Let the input layer comprise N_{m1} units. Assume that the stride length of the convolution layer is s_{m1}. Using γ_{m1} number of filters on the input MCG segment, the convolution is performed, generating feature maps of size $(N_{m1} - k_{m1} + 1) \times \gamma_{m1}$. The ith output of the convolution layer, denoted by y_i, can
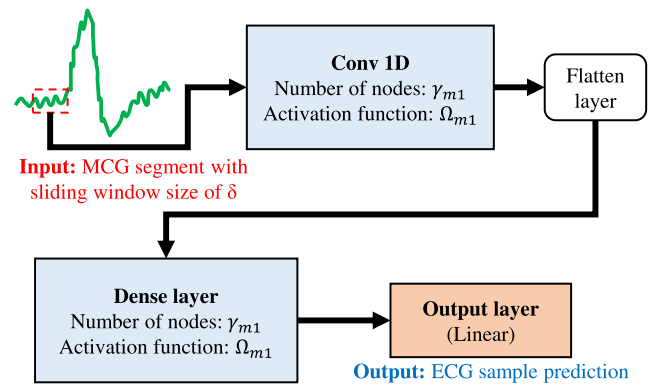
be outlined as follows:

$$y = \Omega_{m1}(\sum_{j \in \gamma_{m1}} (x_j * w_j + b_j)) \qquad (1)$$

Here, x denotes the input MCG segment, and Ω_{m1} refers to the activation function. Also, w and b refer to weight vector and bias. This convolution operation is carried out using the number of filters, γ_{m1} on the input MCG segment.

C. AI MODEL FOR ARRHYTHMIA DETECTION (MODEL-M2)

This sub-section depicts the lightweight heartbeat classification technique for arrhythmia detection using the denoised ECG obtained from model-M1, which can be implemented and integrated with AI-assisted ultra-edge IoT nodes. A lightweight classification model is essential for integrating the AI-aided framework at the ultra-edge IoT nodes for more agile and real-time analysis. One of our previous works established that the deep learning-based lightweight arrhythmia classification (DL-LAC) model outperforms traditional machine learning models such as random forest and K-nearest-neighbour in terms of accuracy and computational expense. Hence, in terms of choosing model-M2, we have employed the DL-LAC framework to detect irregular cardiac status by utilizing single-lead denoised ECG obtained from model-M2 [40]. Fig. 4 illustrates an overview of the architecture of the DL-LAC, which was considered for the model-M2 in the arrhythmia detection task. The model receives denoised ECG from model-M1 as input and generates heartbeat labels as output. The ECG input data size is validated in the data size validation phase, where heartbeats with enough samples are considered to be passed to the model. The model consists of two main components: automated feature extraction (AFE) and automated classification (AC). The feature extraction is conducted by a stack of nl_{AFE} convolution layers. The AC module and the final output layer generate the heartbeat classes. All the details about the model-M2 hyperparameters are listed in our previously published manuscript [40].
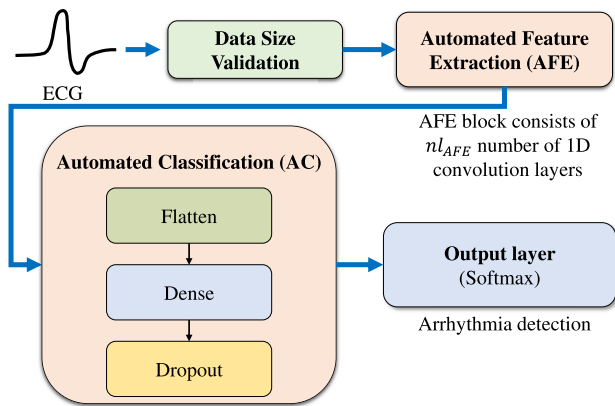
FIGURE 4. The brief architecture of DL-LAC, used as the model-M2. The model-M2 receives denoised ECG as input and produces heartbeat labels as output to detect arrhythmia at the ultra-edge node. The model was initially proposed in [40].

III. PROPOSED DEEP LEARNING-BASED ALGORITHM

In this section, the steps of the training/learning and inference phases of our proposed technique are depicted.

A. LEARNING PHASE USING MODEL-M1

Algorithm 1 demonstrates the steps of selecting the best DL model structure as the model-M1 for noise reduction from MCG data. The algorithm takes training and validation data as input. Algorithm 1 is utilized in the training/learning phase at algorithm 2.

Algorithm 1 returns the best-performing model, proposed as the model-M1, which is later employed as a pre-trained model in the running/inference phase. Next, all the necessary parameters are initialized in steps 1 to 5. Some notable hyperparameters are γ_{m1}, k_{m1}, ε_{m1} that indicates initialize the number of units, filter size, and the number of epochs, respectively. Then, the exploration is conducted for three DL variations (i.e., CNN, CNN-LSTM, and CNN-GRU) and a set of activation functions and optimizers, defined in the nested iterative loops in steps 6 to 8. From step 9, searching for the best DL structure begins by initializing the convolution layer with the essential hyperparameter values. Then, in steps 10 to 15, the algorithm checks for the GRU and LSTM layer to be added with the convolution layer if the corresponding conditions satisfy and update the structure of the current model (denoted by *currModel*). Afterward, the training of the current model is done in step 16 with the training data, and the training loss (*lossCurr*) of the model is computed in step 17. Finally, the current model's performance is compared with the existing best-performing model. If the current model outperforms the previous best one, then the model is stored. Lastly, step 25 returns the best model for future use in the inference phase.

The algorithm of the learning/training phase is described in Algorithm 2. It takes the location of data as input (denoted as *data_path*) and stores the best selected trained model. In the first step, the striding window size (δ) is determined and

---

Algorithm 1: Model Structure Selection of Model-M1

> **Input:** X_{train} (training MCG data), y_{train} (training ECG target), X_{val} (training MCG data), y_{val} (training ECG data)
> **Output:** *bestModel* (best performing model-M1)

1   *allModelTypes* ← [cnn, cnnLstm, cnnGru]
2   γ_{m1}, k_{m1}, ε_{m1} ← initialize the number of units, filter size, and number of epochs
3   *bestScore* ← ∞
4   *bestModel* ← ∅
5   *activations*, *optimizers* ← initialize the list of activation functions and optimizers for the model-M1
6 **foreach** *modelType* ∈ *allModelTypes* **do**
7   **foreach** Ω_{m1} ∈ *activations* **do**
8   **foreach** σ_{m1} ∈ *optimizers* **do**
9   *currModel* ← load convolution layer with γ_{m1}, m, Ω_{m1}, σ_{m1}, ε_{m1}
10   **if** (*modelType* = *cnnLstm*) **then**
11   add LSTM layer to *currModel*
12   **end if**
13   **else if** (*modelType* = *cnnGru*) **then**
14   add GRU layer to *currModel*
15   **end if**
16   train and update model weights of *currModel* using X_{train} and y_{train}
17   *lossCurr* ← evaluate loss function value of model *currModel* using X_{val} and y_{val}
18   **if** (*lossCurr* < *bestScore*) **then**
19   *bestScore* ← *lossCurr*
20   *bestModel* ← *currModel*
21   **end if**
22   **end foreach**
23   **end foreach**
24 **end foreach**
25 return *bestModel*

Algorithm 2: Training/Learning Phase of Model-M1

> **Input:** *data_path* (training data location)

1   δ ← initialize striding window for MCG segment
2   D ← load all training data from *data_path*
3   D ← generate synthetic MCG data by adding low-frequency $1/f$ noise
4   D_{seg} ← []
5 **for** (i = 1 to *length*(D) − δ) **do**
6   D_{seg} ← append $D[i : i + \delta]$ to D_{seg}
7 **end for**
8   X_{train}, y_{train}, X_{val}, y_{val} ← prepare the training and validation data, respectively, from D_{seg} based on the split ratios
9   *model_m1* ← obtain best model-M1 structure by calling Algorithm 1 with arguments X_{train}, y_{train}, X_{val}, y_{val}
10 save model parameters of *model_m1*

---

initialized for MCG segmentation. Then the training data is loaded, and $1/f$ noise is added to generated synthetic

MCG data in steps 2 and 3, respectively. Next, in step 4, the list for storing segmented MCG data is introduced, denoted by D_{seg}. Then, in steps 5 to 7, the initial training data is segmented with δ and saved into D_{seg}. Next, step 8 split the MCG data into training and validation data and stored accordingly. Then the best model-M1 structure is returned in step 9, where Algorithm 1 is utilized with appropriate training and validation data. Finally, the best model-M1 structure is saved along with the model parameters for further utilization as a pre-trained model-M1.

The training algorithm of model-M2 is presented in our previous work [40], and hence we are only exhibiting the inference phase implementation of both models in the following sub-section.

B. INFERENCE PHASE USING BOTH MODEL-M1 AND MODEL-M2

In the running/inference phase, the algorithm proposed with pre-trained model-M1 and model-M2 is exhibited in Algorithm 3. It takes the location of test data for inference and returns the predicted class labels (y_{pred}) for the corresponding data sample. After loading the testing data from step 1, the pre-trained model-M1 and model-M2 (defined as $model_{m1}$ and $model_{m2}$) are loaded in subsequent steps 2 and 3. In step 4, the model-M1 is utilized to predict the ECG of each corresponding MCG record and store the prediction in ecg_{pred}. Next, the DL-LAC model, which was proposed in [40], is used as our proposed model-M2 to generate one of the four heartbeat class labels by using the predicted heartbeat (ecg_{pred}). At the penultimate stage, the class label with the maximum probability is chosen as the final class of each heartbeat sample. Lastly, the predictions list is returned in step 7, which concludes the running algorithm workflow.

Algorithm 3: Inference/Test Phase

Input: $data_{test}$ (test data location)
Output: y_{pred} (heartbeat class predictions)
1 X_{test} ← load all test MCG data $data_{test}$
2 $model_{m1}$ ← load pre-trained model-M1 parameters trained using Algorithm 2
3 $model_{m2}$ ← load pre-trained model-M2 parameters from DL-LAC model designed in [40]
4 ecg_{pred} ← use $model_{m1}$ to predict ECG heartbeat by denoising MCG data
5 y_{prob} ← predict the class probabilities for ECG heartbeat ecg_{pred} employing the model $model_{m2}$
6 y_{pred} ← $argmax(y_{prob})$
7 return y_{pred}

IV. PERFORMANCE EVALUATION

This section assesses the proposed technique's effectiveness using extensive computer-based experiments depending on synthesized noisy MCG data obtained from a publicly available ECG dataset named PTB Diagnostic ECG

Database, referred to as DS1 [41]. The noisy MCG signals are generated from actual ECG data arranged unbiasedly with realistic specifications to establish a substantial training dataset. The proposed AI-aided model can be shifted to the MTJ sensor and then used to minimize noise from MCG, employing ultra-edge logic-in-sensor functionality after the learning/training phase. After obtaining the filtered ECG from the noisy MCG data, we use the DL-LAC model (refereed as model-M2) to get the arrhythmia detection results. Four different real and publicly available arrhythmia datasets are employed to verify the efficiency of the arrhythmia classification task by using the ECG obtained from the MCG. We have utilized a machine with Intel Core i7, 3.00GHz central processing unit (CPU), 16 GB RAM to implement the experimental simulations. We have also employed an Nvidia RTX 2060 graphics processing unit (GPU) to expedite the computing speed. The simulations were conducted with multiple Python libraries (i.e., Scikit-learn, NumPy, Scipy, Pandas, and Matplotlib libraries) for data processing and illustration to assess the efficiency of our strategy. TensorFlow 2.0 and the Keras Python library are used to implement the DL models. We compared our proposed DL-based method to the MA filtering using various performance indicators.

The simulation phases can be summarized into mainly two phases:

1) First, to choose the best structure of the model-M1 and the optimal values of hyperparameters. Among several hyperparameters, we only optimized two vital hyperparameters of model-M1 (i.e., optimizer and activation functions) [42]. Then, we considered stacking multiple configurations with various RNN-based layers (i.e., GRU and LSTM layers) with the convolution layer. Then, we compared the performance of CNN, CNN-GRU, and CNN-LSTM models and selected the most appropriate model architecture for further investigation. PTB Diagnostic ECG Database, referred to as DS1 [41], was employed for this simulation phase.

2) Phase 2: After selecting the appropriate computational technique (i.e., model-M1) for MCG noise minimization, we passed the ECG signal obtained from denoised ECG data to the DL-based model for identifying the presence of arrhythmia heartbeats. We have utilized the previously proposed DL-LAC model [40], referred to as model-M2, to detect irregular arrhythmia heartbeats automatically in a lightweight manner. We have employed multiple publicly available datasets obtained from PhysioNet [43], specified as MIT-BIH Supraventricular Arrhythmia Database (DS2) [44], MIT-BIH Arrhythmia Database (DS3) [45], St Petersburg INCART 12-lead Arrhythmia Database (DS4), and Sudden Cardiac Death Holter Database (DS5) [46].

A. DATASET PREPARATION

For denoising MCG task, we have employed synthesized MCG obtained from ECG cycles of the publicly available

TABLE 1. Frequency of heartbeats of each class in DS2, DS3, DS4, and DS5.

Heartbeat Class	DS2	DS3	DS4	DS5
N (Normal)	87.9%	89.3	87.3%	96.6%
S (Supraventricular ectopic)	6.6%	2.7%	1.1%	0.3%
V (Ventricular ectopic)	5.3%	7.1%	11.3%	3.0%
F (Fusion beat)	0.2%	0.9%	0.3%	0.1%

PTB Diagnostic Database [41]. The dataset arrangement steps are mainly replicated from our previous work in [39] by adopting ECG traces from lead II of all the subjects. The ECG cycles are segmented into individual heartbeats beginning with the R wave to the subsequent QRS complex. Each ECG trace was upsampled to a sampling frequency of 2000 Hz, and then we added low-frequency $1/f$ noise to reflect the spectral characteristics of the MCG signal. The MCG and actual ECG cycles are utilized for training the DL model after the data has been collected and preprocessed.

After obtaining the denoised ECG cycles from the first DL model (model-M1), we then employ the pre-trained DL-LAC model (model-M2) to classify arrhythmia in a lightweight manner. The data preparation steps are primarily comparable to our previous work [40]. We have considered 100 samples in each ECG heartbeat for both the learning/training phase and inference/testing phases. For arrhythmia detection, we have considered ECG signal as the input, denoted by $X = [x_1, x_2, \ldots x_N]$, and the model-M2 outputs heartbeat labels denoted by $y = [y_1, y_2, \ldots y_N]$. In the considered experiments, y_i represents one of four separate ECG heartbeat class labels denoted as N, S, V, and F, representing normal, supraventricular ectopic, ventricular ectopic, and fusion heartbeat, respectively. The details of the four datasets are listed in Table 1.

B. RESULTS AND DISCUSSION
In this sub-section, we illustrate the results of our extensive simulations in two separate phases. In phase 1, we have noted the results while choosing the best DL-based model among three candidate DL models. After selecting the DL model structure of model-M1, in phase 2, we have listed the classification performance when the model-M1 is combined with previously proposed DL-LAC as the model-M2.

1) PHASE 1 (SELECTION OF THE APPROPRIATE MODEL-M1)
In simulation phase 1, to determine the best configuration (i.e., hyperparameters) of the DL-based model-M1, we use the synthetic MCG data from DS1 as input and obtain actual ECG as the model's output. Three variations of DL methods are considered, namely, CNN, CNN-LSTM, and CNN-GRU. As for performance metrics, several quantitive and qualitative indicators are considered. We analyze the power spectral density (PSD) dependence of the remaining noise, the noise

reduction factor, and the Root Mean Square Error (RMSE) to compare how each method performed in the MCG to ECG prediction task. Furthermore, we also compare the output of the three DL-aided models visually to get an intuitive understanding of how the models are performing in terms of the MCG denoising task to obtain ECG signals. Finally, we have also compared the methods in terms of average prediction time to get a quantitive idea about the computational requirement of the model-M1.

Firstly, we have tuned some of the significant hyperparameters of the candidate model-M1 (i.e., CNN, CNN-LSTM, and CNN-GRU). The spectral characteristics of the remaining noise of each technique (i.e., MA and three DL variations) are compared by analyzing the PSD dependence and the noise reduction factor in Fig. 5. Note that the spectral frequency is normalized by the sampling frequency ($f/f_s, f_s = 2$ kHz). Employing the manual tuning technique, the hyperparameters of the DL models are tuned. Our search space consisted of the activation functions rectified linear unit (ReLU) [47], Sigmoid, Tanh, and exponential linear unit (ELU) [48] along with stochastic gradient descent (SGD), root mean square propagation (RMSprop), adaptive moment estimation (Adam), adaptive delta (Adadelta), adaptive gradient algorithm (Adagrad), Adamax, and Nadam as candidate optimizers [49]. We have utilized the set of optimizers and activation functions to determine the best hyperparameters for each of the three DL model variations. Some of the other essential hyperparameters are set without further optimization. The convolution layer in the DL models consists of 300 filters, and the fully-connected layer contains 300 nodes. The learning rate of all three models is set to 0.001. The models are trained for 100 epochs with early stopping criteria of 5 epochs to avoid overfitting while training the models. The CNN-LSTM and CNN-GRU models are designed to have an LSTM and GRU layer after the convolution layer. Both LSTM and GRU layers in the CNN-LSTM and CNN-GRU models consist of 300 nodes and a recurrent dropout of 0.5 to avoid overfitting.
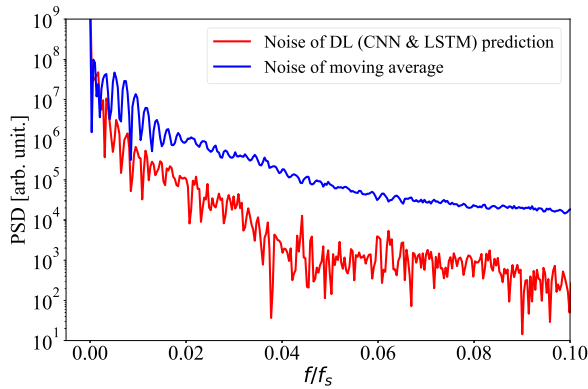
The best optimizer and activation function for the CNN model are Adagrad and Sigmoid; for the CNN-LSTM model, the most suitable optimizer and activation function are Adagrad and ReLU; and for the CNN-GRU model, the best performing optimizer and activation function are Nadam and ReLU. Hence, we have selected these hyperparameters as the most suitable ones for each DL model variation. As manifested in the PSD graphs, all the DL-based variations demonstrate exceptional performance in terms of noise minimization compared to the MA filtering technique. The DL-based methods outperformed the MA technique, especially at the essential low-frequency region ($f/f_s < 0.03$). Similar outcomes are also observed when the noise ratio plots are analyzed for all the methods. Compared to the MA filtering, the noise PSD is decreased by 10–100 times. This improved noise reduction is due to the nonlinear filtering of the CNN layer of the models. A similar effect was found using nonlinear threshold functions applied to dual-tree wavelet
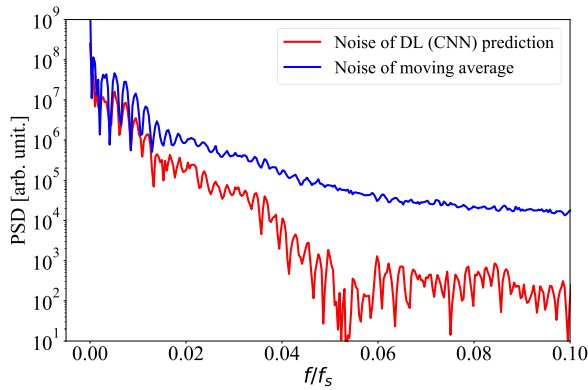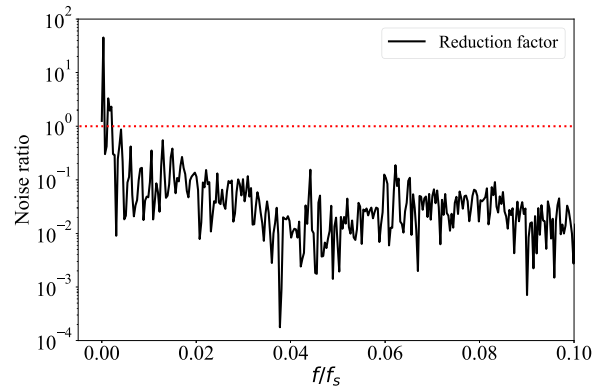
(a) CNN as the deep learning-based prediction (optimizer: Adagrad, activation function: Sigmoid).
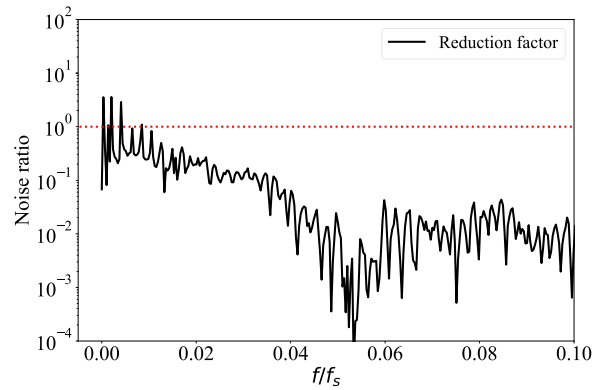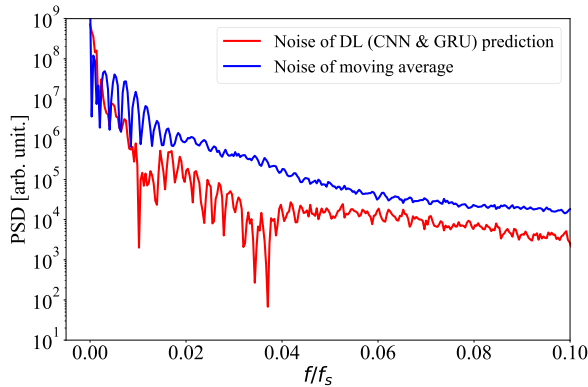
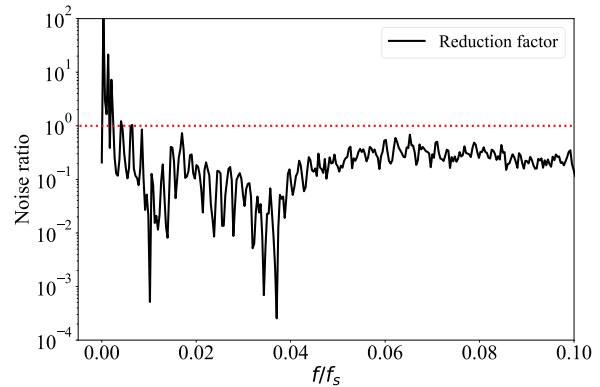(b) CNN as the deep learning-based prediction (optimizer: Adagrad, activation function: Sigmoid).

(c) CNN with LSTM as the deep learning-based prediction (optimizer: Adagrad, activation function: ReLU).

(d) CNN with LSTM as the deep learning-based prediction (optimizer: Adagrad, activation function: ReLU).

(e) CNN with GRU as the deep learning-based prediction (optimizer: Nadam, activation function: ReLU).

(f) CNN with GRU as the deep learning-based prediction (optimizer: Nadam, activation function: ReLU).

FIGURE 5. Comparison of the DL models and the MA filtering in terms of the dependence of noise power on the spectral frequency (in Subfig 5(a), 5(c), 5(e)) and noise PSD ratio (in Subfig 5(b), 5(d), 5(f)). The spectral frequency is normalized by the sampling frequency f_s, and for different optimizers and activation functions, DL model predictions noise is relative to the noise of the MA filtering (using DS1).

transform, with up to 4 times (5.6 dB) reduction in flicker noise power [50]. However, the use of CNN is a more generalized approach to nonlinear filtering.

Fig. 6 provides a visual illustration whereby the actual ECG signal utilized as ground truth is presented at the bottom, which is obtained from the original dataset DS1. The noisy MCG data is also exhibited. Then, the shape of the conventional MA filtering technique is plotted in a blue line. Then we plot the prediction of the DL variations

(i.e., CNN, CNN-LSTM, and CNN-GRU) in the top three line graphs. It is observed that the ECG predictions of the DL variations are notably similar to the actual ECG cycle, and it can robustly distinguish the crucial waves or peaks of the ECG heartbeat such as R-peak, QRS complex, etc. The MA filter with the same window size distorts the QRS complex.

To decide the best-performing model among the DL models, we compare the prediction efficiency in terms of prediction error RMSE and average inference time required.
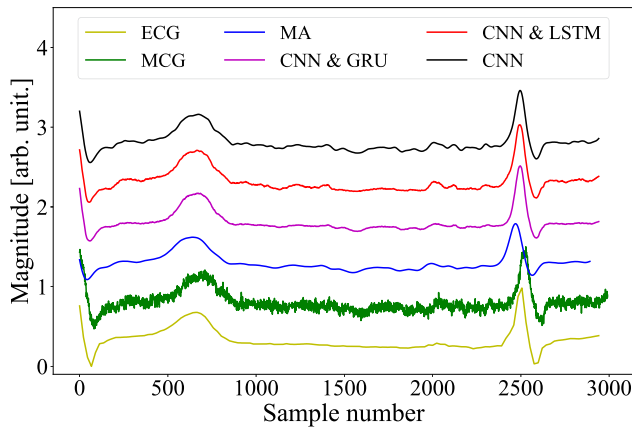
FIGURE 6. Visual performance evaluation demonstrating the original ECG cycle, synthetic noisy MCG cycle used as input, comparison between conventional moving average method, and the DL-based techniques utilized as the model-M1 to process and remove the input MCG signal's noise to obtain denoised ECG. The curves are vertically shifted to increase transparency.
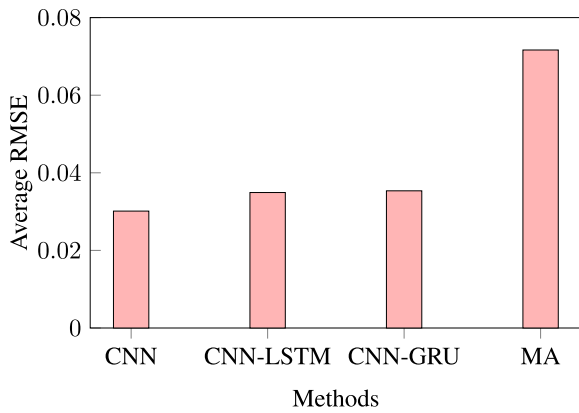

FIGURE 7. Average prediction performance (RMSE) for MA technique and different variations of model-M1.
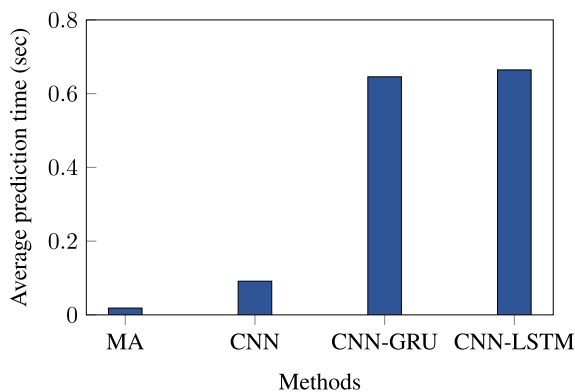

FIGURE 8. Average prediction time per cycle for MA technique and different variations of model-M1.

Fig. 7 manifests the average RMSE, and Fig. 8 represents the mean prediction time induced by each of the methods. The RMSE measure is typically utilized for measuring the prediction performance [51], [52]. It gives an overall view of how each procedure performed in terms of the

ECG prediction task by denoising MCG. The results show that the CNN variation of the DL outperforms all the other DL methods. The MA technique suffers from a high RMSE value due to its smoothing of high-frequency components on the QRS complex,. The other two DL variations (CNN-LSTM and CNN-GRU) performed slightly worse than the CNN model. In terms of the average prediction/inference time, the MA requires the least time, due to the simple computation. However, the CNN model is within an order of magnitude, and also incurs considerably short computation time compared to the CNN-LSTM and CNN-GRU. Thereby, considering the performance efficiency of the CNN, in phase 1 of our experiments, we consider CNN to be the most appropriate DL model, and hence we picked it as the model-M1.

2) PHASE 2 (ARRHYTHMIA DETECTION UTILIZING MULTI-MODEL PIPELINE WITH MODEL-M1 AND MODEL-M2)

In phase-2 of the simulation, we compare three input types to model-M2, such as 1) denoised ECG from the MA filter, 2) denoised ECG from the optimal model for model-M1, and 3) the original ground-truth ECG a reference. Model-M2 is the lightweight arrhythmia classification model (DL-LAC), classifying heartbeats into normal or abnormal heartbeats. Note that the considered architecture of the DL-LAC model had already been optimized substantially in our previous work [40]; therefore, we have employed the optimized model architecture by adopting the appropriate hyperparameters from that work. The datasets DS2–DS5 are new to the model and were not included in the training phase. Furthermore, three different performance indicators (accuracy, precision, and F1-score) are adopted to evaluate the classification efficiency of each method in Table 2. These performance indicators are expected to portray a comprehensive idea of the effectiveness of each technique in terms of classification efficiency.

The classification performance of Model-M2 is high considering that the datasets are new, with an average accuracy of 0.9033. The DL-based pipeline consisting of MCG denoising with model-M1, followed by arrhythmia classification using model-M2, produces similar classification performance. When using the MA-based input to Model-M2, the performance deteriorated. The MA-based pipeline achieved 0.8220 average accuracy, 0.8301 average precision, and 0.8243 average F1-score across the four datasets. Whereas our proposed multi-model pipeline (model-M1 & model-M2) obtained 0.8878, 0.8682, and 0.8626, respectively. Thereby, the analysis reveals that our proposed two-model-based pipeline achieved significantly better performance than using the MA method, with a performance close to using the ground-truth ECG.

Next, we compare the outcomes of the MA filtering method and the proposed multi-model pipeline (model-M1 & model-M2) with the heartbeat prediction of only the base model-M2 (i.e., DL-LAC), which used actual ECG from the dataset as input. In Fig. 9, we compare the cosine similarity, accuracy, precision, and F1-score values to get a more

TABLE 2. Arrhythmia classification performance comparison of the proposed technique with model-M1 and model-M2 combination with MA and model-M2 combination. Both combinations are also compared with baseline (ECG and model-M2). The outcomes noted in the table exhibit the robustness of the methods compared to the original heartbeat labels from the dataset.

Dataset	MA+Model-M2			Model-M1+Model-M2			Original ECG+Model-M2		
	Accuracy	Precision	F1-score	Accuracy	Precision	F1-score	Accuracy	Precision	F1-score
DS2	0.814	0.793	0.803	0.881	0.827	0.839	0.891	0.887	0.863
DS3	0.832	0.819	0.825	0.898	0.876	0.859	0.939	0.907	0.918
DS4	0.822	0.774	0.796	0.867	0.824	0.840	0.869	0.869	0.851
DS5	0.820	0.935	0.873	0.905	0.946	0.913	0.914	0.960	0.930
All DS (average)	0.822	0.830	0.824	0.889	0.868	0.863	0.903	0.906	0.891

(a) Cosine similarity

(b) Accuracy
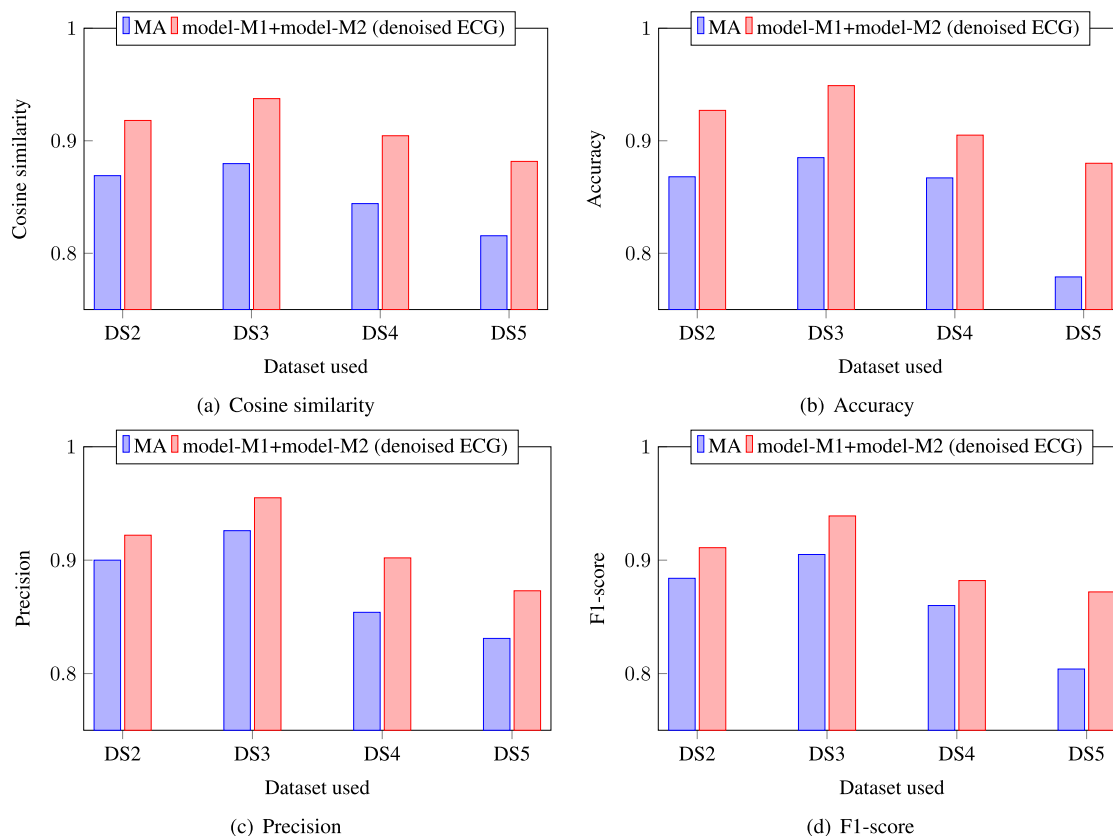
(c) Precision

(d) F1-score

FIGURE 9. Arrhythmia classification performance comparison of the proposed method with MA methods (utilizing noise-minimized/denoised ECG signal for both methods). These results measure the robustness of the methods compared to the outcomes with actual arrhythmia classification model with ECG.

comprehensive idea about the outcomes of both MA and our proposed methodology with respect to the heartbeat prediction of model-M2/DL-LAC model. The cosine similarity, accuracy, precision, and F1-score comparisons are illustrated in sub-figure 9(a), 9(b), 9(c), 9(d), respectively. The mean accuracy, precision, F1-score, and cosine similarity values for the MA filtering technique are 0.8494, 0.8778, 0.8631, and 0.8521. On the other hand, using denoised ECG data, our proposed approach (model-M1 & model-M2) can achieve much higher efficiency (accuracy: 0.9154, precision: 0.9131, F1-score: 0.901, and cosine similarity: 0.9084) than the MA method with respect to the results of the DL-LAC model with actual ECG data. The encouraging outcomes reveal that the proposed multi-model methodology is considerably robust in arrhythmia detection from MCG data. Also, the proposed

technique is quite similar in terms of efficiency when we compare it to the results of the DL-LAC model with ECG data.

V. CONCLUSION

The spread of various worldwide epidemics in recent times has accelerated the demand for remote patient monitoring. Hence, in this paper, we have presented an AI-aided analytics pipeline for next-generation ultra-edge IoT sensors to facilitate remote health monitoring. As a use case, we have elected one of the major health issues worldwide: arrhythmia classification. Although typically automated arrhythmia detection utilizes ECG data, the ECG data collection process using contact leads is inconvenient. Even though ECG can be employed to monitor patients' heart problems, the

discomfort it causes when used for prolonged time warrants that new applications must be established in the IoT industry for cardiovascular monitoring. In this context, spintronic sensors based on Magnetic Tunnel Junction (MTJ) devices show a promising advantage in regard to sensitivity and mobility, as well as the capacity to support ultra-edge logic-in-sensor topology. However, one vital challenge with MCG analysis is the low-frequency $1/f$ noise which needs to be reduced effectively to analyze the data for arrhythmia detection. Firstly, to minimize low-frequency noise from collected MCG data, we propose a custom CNN model, referred to as model-M1, which is used to generate denoised ECG from raw MCG signal. Then, we utilize a lightweight AI framework noted as model-M2 to detect arrhythmia by evaluating model-M1's denoised ECG. We conducted extensive simulations on various publicly available datasets. We found that the classification results from our pipeline are comparable to the ECG-based arrhythmia detection. The proposed AI-aided multi-model MCG analysis using spintronic technology-based ultra-edge IoT sensors can be considered a new path to facilitate remote patient monitoring for a prolonged period, especially in aging urban populations and under-served regions.

Even though we have used multiple publicly available datasets to conduct extensive experimental simulations, one of the limitations of this work is we have not tested the proposed DL-based approach in a real-world setting. Our proposed method demonstrated excellent potential for arrhythmia detection in an AI-aided non-invasive manner, but we have investigated a particular application case study of arrhythmia classification, and this work can be extended in many different directions. For future research directions, the proposed approach can be generalized to other remote subject monitoring tasks (i.e., brain-activity monitoring using MagnetoEncephaloGraphy, MEG) as well, especially for a long time non-invasive monitoring. Furthermore, this MTJ-based logic-in-sensor approach can be extended to other domains, such as the quality assurance of sensors during aerospace and automotive control. Therefore, future research can be conducted to evaluate the potential of this approach in other remote health monitoring use-cases as well as other domains.

REFERENCES

[1] S. Sakib, M. M. Fouda, Z. M. Fadlullah, and N. Nasser, "Migrating intelligence from cloud to ultra-edge smart IoT sensor based on deep learning: An arrhythmia monitoring use-case," in *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, Jun. 2020, pp. 595–600.

[2] Z. Fadlullah, M. M. Fouda, A.-S.-K. Pathan, N. Nasser, A. Benslimane, and Y.-D. Lin, "Smart IoT solutions for combating the COVID-19 pandemic," *IEEE Internet Things Mag.*, vol. 3, no. 3, pp. 10–11, Sep. 2020.

[3] S. Sakib, M. M. Fouda, and Z. M. Fadlullah, "A rigorous analysis of biomedical edge computing: An arrhythmia classification use-case leveraging deep learning," in *Proc. IEEE Int. Conf. Internet Things Intell. Syst. (IoTaIS)*, Jan. 2021, pp. 136–141.

[4] Q. Yao, R. Wang, X. Fan, J. Liu, and Y. Li, "Multi-class arrhythmia detection from 12-lead varied-length ECG using attention-based time-incremental convolutional neural network," *Inf. Fusion*, vol. 53, pp. 174–182, Jan. 2020.

[5] R. Mehra, "Global public health problem of sudden cardiac death," *J. Electrocardiol.*, vol. 40, no. 6, pp. S118–S122, Nov. 2007.

[6] R. S. Barbosa, L. Glass, R. Proietti, B. Burstein, A. Al-Turki, L. Sobolik, Z. Zhang, G. Viart, M. Samuel, A. Shrier, and V. Essebag, "Defining the pattern of initiation of monomorphic ventricular tachycardia using the beat-to-beat intervals recorded on implantable cardioverter defibrillators from the RAFT study: A computer-based algorithm," *J. Electrocardiol.*, vol. 51, no. 3, pp. 470–474, May 2018.

[7] S. Sakib, T. Tazrin, M. M. Fouda, Z. M. Fadlullah, and M. Guizani, "DL-CRC: Deep learning-based chest radiograph classification for COVID-19 detection: A novel approach," *IEEE Access*, vol. 8, pp. 171575–171589, 2020.

[8] S. Sakib, M. M. Fouda, Z. M. Fadlullah, and N. Nasser, "On COVID-19 prediction using asynchronous federated learning-based agile radiograph screening booths," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2021, pp. 1–6.

[9] S. Sakib, M. M. Fouda, M. Al-Mahdawi, A. Mohsen, M. Oogane, Y. Ando, and Z. M. Fadlullah, "Noise-removal from spectrally-similar signals using reservoir computing for MCG monitoring," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2021, pp. 1–6.

[10] B. Shakya, M. M. Fouda, S. C. Chiu, and Z. M. Fadlullah, "A circuit-embedded reservoir computer for smart noise reduction of MCG signals," in *Proc. IEEE Int. Conf. Internet Things Intell. Syst. (IoTaIS)*, Nov. 2021, pp. 56–61.

[11] M. É. Czeisler, K. Marynak, K. E. N. Clarke, Z. Salah, I. Shakya, J. M. Thierry, N. Ali, H. Mcmillan, J. F. Wiley, M. D. Weaver, C. A. Czeisler, S. M. W. Rajaratnam, and M. E. Howard, "Delay or avoidance of medical care because of COVID-19–related concerns—United States, June 2020," *Morbidity Mortality Weekly Rep.*, vol. 69, no. 35, p. 1250, Sep. 2020.

[12] *Health Insurance Claims Review & Reimbursement Services, 2020 Monthly Reports*. Accessed: Aug. 20, 2021. [Online]. Available: https://www.ssk.or.jp/tokeijoho/geppo/geppo_r02.html

[13] S. Sakib, M. M. Fouda, Z. M. Fadlullah, K. Abualsaud, E. Yaacoub, and M. Guizani, "Asynchronous federated learning-based ECG analysis for arrhythmia detection," in *Proc. IEEE Int. Medit. Conf. Commun. Netw. (MeditCom)*, Sep. 2021, pp. 277–282.

[14] S. Yamada and I. Yamaguchi, "Magnetocardiograms in clinical medicine: Unique information on cardiac ischemia, arrhythmias, and fetal diagnosis," *Internal Med.*, vol. 44, no. 1, pp. 1–19, 2005.

[15] K. Fujiwara, M. Oogane, A. Kanno, M. Imada, J. Jono, T. Terauchi, T. Okuno, Y. Aritomi, M. Morikawa, M. Tsuchida, N. Nakasato, and Y. Ando, "Magnetocardiography and magnetoencephalography measurements at room temperature using tunnel magneto-resistance sensors," *Appl. Phys. Exp.*, vol. 11, no. 2, Jan. 2018, Art. no. 023001.

[16] M. Wang, Y. Wang, L. Peng, and C. Ye, "Measurement of triaxial magnetocardiography using high sensitivity tunnel magnetoresistance sensor," *IEEE Sensors J.*, vol. 19, no. 21, pp. 9610–9615, Nov. 2019.

[17] Y. Shirai, K. Hirao, T. Shibuya, S. Okawa, Y. Hasegawa, Y. Adachi, K. Sekihara, and S. Kawabata, "Magnetocardiography using a magnetoresistive sensor array," *Int. Heart J.*, vol. 60, no. 1, pp. 50–54, Jan. 2019.

[18] M. Oogane, K. Fujiwara, A. Kanno, T. Nakano, H. Wagatsuma, T. Arimoto, S. Mizukami, S. Kumagai, H. Matsuzaki, N. Nakasato, and Y. Ando, "Sub-pT magnetic field detection by tunnel magneto-resistive sensors," *Appl. Phys. Exp.*, vol. 14, no. 12, Nov. 2021, Art. no. 123002.

[19] J.-G. Zhu and C. Park, "Magnetic tunnel junctions," *Mater. Today*, vol. 9, no. 11, pp. 36–45, Nov. 2006.

[20] W. Egelhoff, P. Pong, J. Unguris, R. McMichael, E. Nowak, A. Edelstein, J. Burnette, and G. Fischer, "Critical challenges for picoTesla magnetic-tunnel-junction sensors," *Sens. Actuators A, Phys.*, vol. 155, no. 2, pp. 217–225, Oct. 2009.

[21] K. Fujiwara, M. Oogane, T. Nishikawa, H. Naganuma, and Y. Ando, "Detection of sub-nano-Tesla magnetic field by integrated magnetic tunnel junctions with bottom synthetic antiferro-coupled free layer," *Jpn. J. Appl. Phys.*, vol. 52, no. 4, Apr. 2013, Art. no. 04CM07.

[22] X. Liu, K. H. Lam, K. Zhu, C. Zheng, X. Li, Y. Du, C. Liu, and P. W. T. Pong, "Overview of spintronic sensors with Internet of Things for smart living," *IEEE Trans. Magn.*, vol. 55, no. 11, pp. 1–22, Nov. 2019.

[23] C. Krittanawong, A. J. Rogers, K. W. Johnson, Z. Wang, M. P. Turakhia, J. L. Halperin, and S. M. Narayan, "Integration of novel monitoring devices with machine learning technology for scalable cardiovascular management," *Nature Rev. Cardiol.*, vol. 18, no. 2, pp. 75–91, 2021.

[24] M. Bansal and B. Gandhi, "IoT & big data in smart healthcare (ECG monitoring)," in *Proc. Int. Conf. Mach. Learn., Big Data, Cloud Parallel Comput. (COMITCon)*, Feb. 2019, pp. 390–396.

[25] A. Rahman, T. Rahman, N. H. Ghani, S. Hossain, and J. Uddin, "IoT based patient monitoring system using ECG sensor," in *Proc. Int. Conf. Robot., Electr. Signal Process. Techn. (ICREST)*, Jan. 2019, pp. 378–382.

[26] M. Wu, Y. Lu, W. Yang, and S. Y. Wong, "A study on arrhythmia via ECG signal classification using the convolutional neural network," *Frontiers Comput. Neurosci.*, vol. 14, p. 106, Jan. 2021.

[27] Z. Ebrahimi, M. Loni, M. Daneshtalab, and A. Gharehbaghi, "A review on deep learning methods for ECG arrhythmia classification," *Expert Syst. Appl. X*, vol. 7, Sep. 2020, Art. no. 100033.

[28] T.-M. Chen, C.-H. Huang, E. S. C. Shih, Y.-F. Hu, and M.-J. Hwang, "Detection and classification of cardiac arrhythmias by a challenge-best deep learning neural network model," *iScience*, vol. 23, no. 3, Mar. 2020, Art. no. 100886.

[29] E. Izci, M. Degirmenci, M. A. Özdemir, and A. Akan, "ECG arrhythmia detection with deep learning," in *Proc. 28th Signal Process. Commun. Appl. Conf. (SIU)*, Oct. 2020, pp. 1–4.

[30] S. Shraddha, S. K. Pandey, U. Pawar, and R. R. Janghel, "Classification of ECG arrhythmia using recurrent neural networks," *Proc. Comput. Sci.*, vol. 132, pp. 1290–1297, Jan. 2018.

[31] T. Mahmud, S. A. Fattah, and M. Saquib, "DeepArrNet: An efficient deep CNN architecture for automatic arrhythmia detection and classification from denoised ECG beats," *IEEE Access*, vol. 8, pp. 104788–104800, 2020.

[32] A. Isin and S. Ozdalili, "Cardiac arrhythmia detection using deep learning," *Proc. Comput. Sci.*, vol. 120, pp. 268–275, Jan. 2017.

[33] C. Che, P. Zhang, M. Zhu, Y. Qu, and B. Jin, "Constrained transformer network for ECG signal processing and arrhythmia classification," *BMC Med. Informat. Decis. Making*, vol. 21, no. 1, pp. 1–13, Dec. 2021.

[34] K. Weimann and T. O. F. Conrad, "Transfer learning for ECG classification," *Sci. Rep.*, vol. 11, no. 1, pp. 1–12, Dec. 2021.

[35] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[36] S. A. Magid, F. Petrini, and B. Dezfouli, "Image classification on IoT edge devices: Profiling and modeling," *Cluster Comput.*, vol. 23, no. 2, pp. 1025–1043, Jun. 2020.

[37] Z. Q. Lei, G. J. Li, W. F. Egelhoff, P. T. Lai, and P. W. T. Pong, "Review of noise sources in magnetic tunnel junction sensors," *IEEE Trans. Magn.*, vol. 47, no. 3, pp. 602–612, Mar. 2011.

[38] V. Pandey and V. K. Giri, "High frequency noise removal from ECG using moving average filters," in *Proc. Int. Conf. Emerg. Trends Electr. Electron. Sustain. Energy Syst. (ICETEESES)*, Mar. 2016, pp. 191–195.

[39] A. Mohsen, M. Al-Mahdawi, M. M. Fouda, M. Oogane, Y. Ando, and Z. M. Fadlullah, "AI aided noise processing of spintronic based IoT sensor for magnetocardiography application," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.

[40] S. Sakib, M. M. Fouda, Z. M. Fadlullah, N. Nasser, and W. Alasmary, "A proof-of-concept of ultra-edge smart IoT sensor: A continuous and lightweight arrhythmia monitoring approach," *IEEE Access*, vol. 9, pp. 26093–26106, 2021.

[41] R. Bousseljot, D. Kreiseler, and A. Schnabel, "Nutzung der EKG-signaldatenbank CARDIODAT der PTB über das internet," *Biomedizinische Technik/Biomed. Eng.*, vol. 40, no. 1, pp. 317–318, 1995.

[42] L. Yang and A. Shami, "On hyperparameter optimization of machine learning algorithms: Theory and practice," *Neurocomputing*, vol. 415, pp. 295–316, Nov. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231220311693

[43] 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, pp.e215–e220, Jun. 2000.

[44] S. D. Greenwald, R. S. Patil, and R. G. Mark, "Improved detection and classification of arrhythmias in noise-corrupted electrocardiograms using contextual information," in *Proc. Comput. Cardiol.*, Sep. 1990, pp. 461–464.

[45] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, May 2001.

[46] S. D. Greenwald, "The development and analysis of a ventricular fibrillation detector," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, USA, 1986.

[47] A. F. Agarap, "Deep learning using rectified linear units (ReLU)," *CoRR*, vol. abs/1803.08375, pp. 1–7, Mar. 2018.

[48] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation functions: Comparison of trends in practice and research for deep learning," *CoRR*, vol. abs/1811.03378, pp. 1–20, Nov. 2018.

[49] Q. V. Le, J. Ngiam, A. Coates, A. Lahiri, B. Prochnow, and A. Y. Ng, "On optimization methods for deep learning," in *Proc. ICML*. Madison, WI, USA: Omnipress, 2011, pp. 265–272.

[50] O. E. B'charri, R. Latif, K. Elmansouri, A. Abenaou, and W. Jenkal, "ECG signal performance de-noising assessment based on threshold tuning of dual-tree wavelet transform," *Biomed. Eng. OnLine*, vol. 16, no. 1, p. 26, Dec. 2017.

[51] S. Sakib, T. Tazrin, M. M. Fouda, Z. M. Fadlullah, and N. Nasser, "A deep learning method for predictive channel assignment in beyond 5G networks," *IEEE Netw.*, vol. 35, no. 1, pp. 266–272, Jan. 2021.

[52] S. Sakib, T. Tazrin, M. M. Fouda, Z. M. Fadlullah, and N. Nasser, "An efficient and lightweight predictive channel assignment scheme for multiband B5G-enabled massive IoT: A deep learning approach," *IEEE Internet Things J.*, vol. 8, no. 7, pp. 5285–5297, Apr. 2021.

SADMAN SAKIB received the master's degree in computer science (specialization in artificial intelligence) from Lakehead University, Canada, in 2021. He is currently working as a Deep Learning Developer at MERIDIAN, Institute for Big Data Analytics, Dalhousie University. His research interests include intelligent computing/communication systems for mobile health (mHealth) applications for providing better health outcomes in Canada and beyond. He was the winner of the best student's poster award of the Canadian Institutes of Health Research (CIHR) category in the graduate conference held at the Graduate Student Conference—Poster Presentation, Research and Innovation Week, Lakehead University, in 2020. Also, he received an internship opportunity through the Mitacs Accelerate research grant, in 2020.

MOSTAFA M. FOUDA (Senior Member, IEEE) received the Ph.D. degree in information sciences from Tohoku University, Japan, in 2011. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, Idaho State University, ID, USA. He also holds the position of an Associate Professor with Benha University, Egypt. He has worked as an Assistant Professor with Tohoku University. He was a Postdoctoral Research Associate with Tennessee Technological University, TN, USA. He has been engaged in research on cybersecurity, communication networks, wireless mobile communications, smart healthcare, smart grids, AI, blockchain, and the IoT. He has published more than 60 papers in prestigious peer-reviewed journals and conferences. He was a recipient of the prestigious 1st place award during his graduation from the Faculty of Engineering at Shoubra, Benha University, Egypt, in 2002. He has served as the Symposium/Track Chair of IEEE VTC2021-Fall conference. He has also served as the Guest Editor of some special issues of several top-ranked journals, such as IEEE WIRELESS COMMUNICATIONS (WCM) and IEEE INTERNET OF THINGS MAGAZINE (IoTM). He also serves as a Referee of some renowned IEEE journals and magazines, such as IEEE COMMUNICATIONS SURVEYS AND TUTORIALS (COMST), IEEE WIRELESS COMMUNICATIONS (WCM), IEEE TRANSACTIONS ON WIRELESS COMMUNICATION, IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS (TPDS), IEEE TRANSACTIONS ON SMART GRID (TSG), IEEE ACCESS, IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT (TNSM), IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTING (TETC), and IEEE NETWORK. He is an Editor of IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY (TVT) and an Associate Editor of IEEE ACCESS.

MUFTAH AL-MAHDAWI (Senior Member, IEEE) received the B.Sc. degree in electrical and electronic engineering from the University of Benghazi, Libya, in 2007, and the M.Sc. and Ph.D. degrees in electronic engineering from Tohoku University, Japan, in 2013.

Since 2018, he has been an Assistant Professor with the Center for Science and Innovation in Spintronics (CSIS), Tohoku University. He was a Postdoctoral Researcher with the National Institute for Materials Science, from 2017 to 2018, and the School of Engineering, Tohoku University, from 2013 to 2016. He was a Field Engineer with Schlumberger Ltd., from 2007 to 2008. He has a track record of 30 journal articles and 50 conference presentations. His current research interests include within spintronics-based sensors and systems, magnetotransport at the nano-scale, spin dynamics, and multiferroic and magnetoelectric devices.

Dr. Al-Mahdawi is a Senior Member of IEEE Magnetics Society and a member of Japanese Society of Applied Physics.

ATTAYEB MOHSEN received the Ph.D. degree in medical sciences from Tohoku University, Japan, in 2014. He is currently a Postdoctoral Researcher with the Laboratory of Bioinformatics, Artificial Intelligence Center for Health and Biomedical Research (ArCHER), National Institutes of Biomedical Innovation, Health and Nutrition, Osaka Japan. His research interests include computational biology, gene expression, and NGS data analysis.

MIKIHIKO OOGANE received the Ph.D. degree in applied physics from Tohoku University, Japan, in 2003.

From 2004 to 2010, he was a Research Associate with the Department of Applied Physics, Tohoku University, where he has been an Associate Professor with the Department of Applied Physics, since 2010. He is an author of more than 200 articles. His research interests include the development of spintronic materials and devices. In particular, he has made remarkable achievements in the development and application of magnetic tunnel junction devices. He was a recipient of the Magnetics Society of Japan Young Researcher Award for Excellence, in 2007, and JSAP Outstanding Paper Award, in 2012.

YASUO ANDO received the Ph.D. degree in engineering from Tohoku University, Japan, in 1994.

From 1986 to 1992, he was a Researcher with Konica Corporation. Since 1992, he has been with the School of Engineering, Tohoku University, where he has been a Professor, since 2007. He worked as a Visiting Researcher at RIEC, Tohoku University, NHK Research Labs, and TU Kaiserslautern. He has authored eight books or book chapters. He has published more than 400 journal articles and holds 60 patents/patent applications. His research interests include solid-state physics, in particular magnetism and magneto-transport properties, thin film and device preparation, and materials science. He led the demonstration of the first detection of human bio-magnetic fields by magnetic tunnel junction sensors at room temperature. He is a fellow of Japanese Society of Applied Physics and a fellow of the Magnetics Society of Japan.

ZUBAIR MD. FADLULLAH (Senior Member, IEEE) received the Ph.D. degree in information sciences from Tohoku University, Japan, in 2011. He is currently an Associate Professor with the Computer Science Department, Lakehead University, and the Research Chair of the Thunder Bay Regional Health Research Institute (TBRHRI), Thunder Bay, ON, Canada. He was an Associate Professor with the Graduate School of Information Sciences, Tohoku University, from 2017 to 2019. His research interests include cyber physical system layers in sensing, communication, and computing problems and elegant solutions. He is currently an Editor of IEEE Transactions on Vehicular Technology (TVT), IEEE Access, and IEEE Open Journal of the Communications Society (OJ-COMS).

• • •