Received August 15, 2021, accepted August 22, 2021, date of publication August 24, 2021, date of current version September 2, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3107687

Screening Cardiovascular Autonomic Neuropathy in Diabetic Patients With Microvascular Complications Using Machine Learning: A 24-Hour Heart Rate Variability Study

MOHANAD ALKHODARI[1], MAMUNUR RASHID[2], MOHAMMAD ABDUL MUKIT[2,3], KHAWZA I. AHMED[2], (Member, IEEE), RAQIBUL MOSTAFA[2], (Senior Member, IEEE), SHARMIN PARVEEN[4], (Member, IEEE), AND AHSAN H. KHANDOKER[1], (Senior Member, IEEE)

[1]Healthcare Engineering Innovation Center (HEIC), Department of Biomedical Engineering, Khalifa University, Abu Dhabi, United Arab Emirates
[2]Department of Electrical and Electronic Engineering, United International University, Dhaka 1212, Bangladesh
[3]Department of Electrical and Computer Engineering, The University of Oklahoma, Tulsa, OK 73019, USA
[4]Department of Health Informatics, Bangladesh University of Health Sciences, Dhaka 1216, Bangladesh

Corresponding author: Mohanad Alkhodari (mohanad.alkhodari@ku.ac.ae)

This work was supported in part by the Institute of Advance Research (IAR), United International University, Dhaka, Bangladesh, under Grant UIU-RG-162013, and in part by the Healthcare Engineering Innovation Center (HEIC), Khalifa Univeristy, Abu Dhabi, United Arab Emirates, under Grant 8474000132.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethical Review Committee of Bangladesh University of Health Sciences under Application No. BUHS/BIO/EA/17/01, and performed in line with the Declaration of Helsinki and the Ministry of Health and Family Welfare of Bangladesh.

ABSTRACT Cardiovascular autonomic neuropathy (CAN) is one of the most overlooked complications associated with diabetes. It is characterized by damage in the autonomic nerves regulating heart rate and vascular compliance. Ewing battery is currently the diagnostic tool of choice but is unable to detect sub-clinical CAN and requires patient cooperation. In addition, appropriate timing (day/night) of CAN diagnostic test was not explored in the past. Therefore, a novel approach is proposed herein to investigate the feasibility of using heart rate variability (HRV) features over 24 hours embedded within machine learning algorithms to provide a complete screening for patients suffering from CAN. 24-hour Holter ECG data were acquired from a Bangladeshi cohort (n = 95 patients [75 Diabetic and 25 healthy]). HRV features were extracted from every 5-minute segment of the HRV signal and used as input to four machine learning algorithms for hourly training and testing. A complete hierarchical step by step diagnosis procedure (4 tests) was developed; namely test 1 to check for being healthy or diabetic; test 2 to check for any microvascular complications (including neuropathy such as CAN, peripheral neuropathy (DPN), nephropathy (NEP), and retinopathy (RET)) or not; test 3 to check for presence of only CAN; test 4 to check for combined or multiple complications along with CAN. The highest levels of performance were achieved with accuracy measures of 85.5% (test 1 - convolutional neural network (CNN)), 98.5% (test 2 - CNN), 98.3% (test 3 - one-class support vector machines (SVM)), and 90.9% (test 4 - random forest). Hours 7:00 AM and 7:00 PM were found to be most significant in the diagnosis of CAN in diabetic patients (test 1, 3, and 4). Early screening of CAN by our proposed models could help primary healthcare centers stratify the risk leading to early treatment in preventing sudden cardiac death due to silent myocardial infarction. The approach is considered to be simple and effective, especially for under-resourced clinical settings.

INDEX TERMS Diabetes, cardiovascular autonomic neuropathy (CAN), 24-hour electrocardiography (ECG), heart rate variability (HRV), machine learning, one-class support vector machine (SVM).

The associate editor coordinating the review of this manuscript and approving it for publication was Mohammad Zia Ur Rahman.

I. INTRODUCTION

Diabetes is a chronic disease that occurs when high levels of glucose are accumulated in the blood. It consists of two

types; type 1 and type 2. In type 1, the pancreas does not efficiently produce enough insulin hormone to the body to regulate blood sugar levels. On the other hand, in type 2, the body cannot effectively use the produced insulin to move sugar from the blood into cells and subsequently, use it for energy [1], [2]. This interruption in the secretion and transport of this hormone in the body raises glucose levels in the blood causing serious damage to nerves and blood vessels. According to the world health organization (WHO) [3], diabetes is the major cause of heart attacks, kidney failure, stroke, lower limb amputation, and blindness in people worldwide. It was estimated that more than 422 million people are suffering from diabetes in 2014, reaching a total of 1.5 million deaths directly caused by this disease in 2019.

Most diabetic patients experience additional metabolic events and comorbidities such as hyperlipidemia, hypertension, metabolic syndrome, and obesity. In addition, the existence of microvascular complications of all types is considered to be the major cause of morbidity and mortality in these patients [4]. A neuropathy including cardiovascular autonomic neuropathy (CAN), diabetic peripheral neuropathy (DPN), nephropathy (NEP), and retinopathy (RET) affects all major organs in the body. In particular, autonomic neuropathy (AN), when it involves the cardiac system as in CAN, results in substantial cardiovascular dysfunction, cardiac death, arrhythmias, and myocardial infarctions (MI) [5], [6], In CAN, neuropathy is characterized by changes in heart rate variability (HRV) as a result of damage in the nerves regulating the heart rate. It is prevalent in around 20%-60% of patients suffering from diabetes worldwide with five times higher mortality rate than other AN types [7], [8].

In many cases, the occurrence of neuropathy in diabetic patients goes under-diagnosed, especially for CAN cases, where the cardiovascular dysfunction is accompanied by a progression of myocardial ischemia that is usually painless and silent [9], [10]. The early diagnosis usually reduces the risks associated with CAN including myocardial infarcts and sudden cardiac death [11], thus, routine CAN screening is recommended for patients with diabetes. The current gold standard approach to diagnose and assess CAN patients is to perform five Ewing cardiac reflex tests [12]. However, these tests are considered cumbersome especially for patients with cardiorespiratory dysfunction, frailty, and severe obesity [13]. In addition, since CAN onset is usually occult, screening it through these tests at the time of the onset is less efficient [14]. Therefore, recent research works have suggested electrocardiography (ECG) attributes, including heart rate and HRV information, to address the drawbacks of the Ewing tests, as CAN is mostly characterized by differences in the time and frequency domains of HRV extracted from ECG signals [7], [8]. Furthermore, the recent development in artificial intelligence (AI) algorithms has achieved higher levels of accuracy when diagnosing patients with such cases.

In the literature, a number of previous studies focused on the use of HRV for the analysis of cardiovascular pathologies [15]–[18]. Furthermore, several studies highlighted the association between HRV and diabetes using data analysis tools [19]–[21] and machine or deep learning algorithms [22]–[25]. In [26]–[29], researchers utilized Ewing tests to predict and diagnose patients with mild and moderate stages of CAN. On the other hand, very few studies investigated HRV as a diagnostic tool to evaluate diabetic patients suffering from CAN [30]–[32]. Despite the promising performance in these studies, there still exists a gap in the knowledge about how HRV impacts the screening of the presence of cardiac neuropathy in 24-hour duration of a day, especially when combined with AI algorithms to provide a complete approach for diagnosing CAN cases [14] in the probable presence of other diabetes-induced microvascular complications such as DPN, NEP, and RET.

A. MAJOR CONTRIBUTION

In this paper, a study is conducted to investigate the ability of 5-minute HRV features, extracted over the span of 24 hours of a day, in screening diabetic patients suffering from CAN with and without other co-morbid microvascular complications such as DPN, NEP, and RET. The importance of this work lies in developing a tool capable of identifying those diabetic patients with only CAN, which is usually underdiagnosed due to the silent nature of this complication, and discriminate them from other patients suffering from additional complications along with CAN that makes it even harder to diagnose CAN properly. In addition, the study provides a complete hierarchical machine learning approach to enhance the step by step clinical evaluation of patients starting from screening their diabetic status down to microvascular complications associated with it, with the focus on presence of CAN. (Fig. 1).

To the best of the authors' knowledge, the decision on an optimal time for clinical diagnosis of CAN patients was not estimated in the past. Only a single research work [8] tried to use machine learning algorithms to detect severe CAN cases using HRV features spanned over only 20 minutes. However, it did not reflect the detection ability of HRV throughout the day/night circadian cardiac cycle. Thus, the proposed study presented herein targets analyzing the dynamics of the cardiovascular system in diabetic patients, particularly CAN subjects, on an hour-by-hour basis throughout the 24-hour duration of a day. In addition, it highlights the 24-hour circadian functionality of the heart (correlated within HRV features) to suggest the most suitable hours for a better diagnosis of CAN cases, thus, providing a suitable screening strategy with an optimal time window to prevent further development of diabetic neuropathies.

Furthermore, this study provides a complete hierarchical diagnostic procedure through four sequential clinical testing schemes to successfully diagnose neuropathy complications for diabetic patients. This can serve as a guide for clinicians to enhance the ability and significance of HRV features in diabetes analysis and complications detection. In addition, it paves the way towards transforming the proposed procedure from a testing scenario to direct clinical implementation.

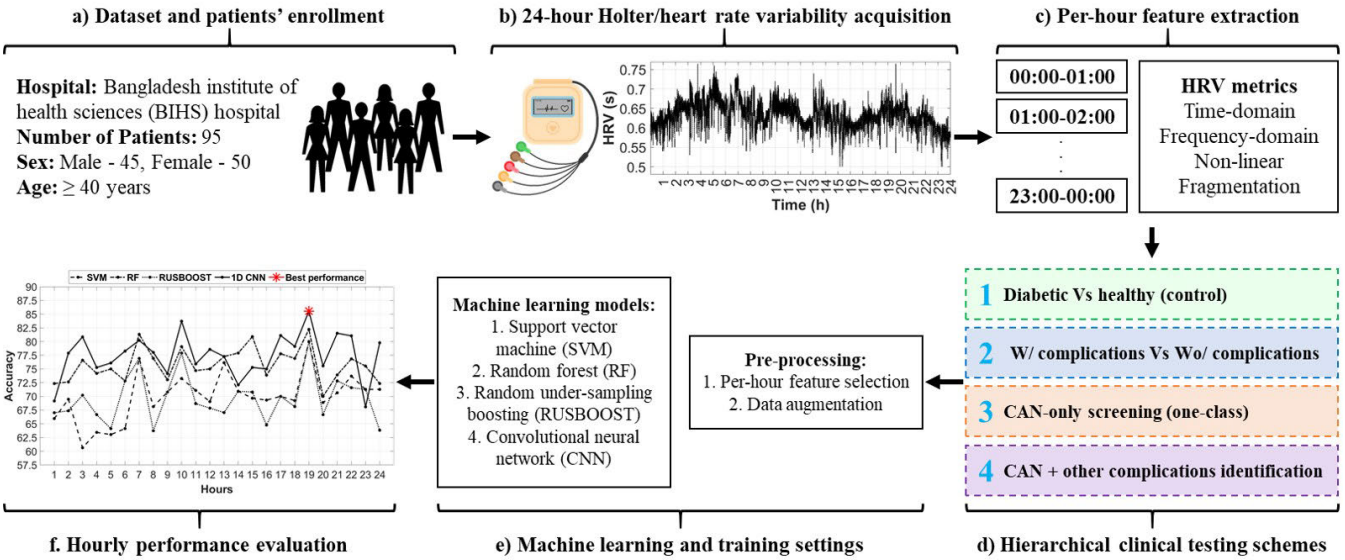


FIGURE 1. A graphical abstract of the complete procedure followed in the study. (a) 95 Patient enrollment at Bangladesh Institute of Health Sciences (BIHS) Hospital. (b) heart rate variability (HRV) acquisition from 24-hour Holter electrocardiography (ECG) recordings. (c) 5-minute HRV features extraction from time-domain, frequency-domain, non-linear, and fragmentation metrics for each hour of a day. (d) Hierarchical clinical testing schemes including four tests to deeply diagnose diabetic neuropathies cases, with the focus on cardiovascular autonomic neuropathy (CAN). (e) The proposed machine learning algorithms including support vector machine (SVM), random forest (RF), random under-sampling boosting (RUSBOOST), and convolutional neural network (CNN) with initial pre-processing steps. (f) Per-hour performance evaluation of each clinical test scheme using each machine learning model.

Therefore, in clinical practice, the ability to follow a sequential testing approach at certain hours of a day is considered of high importance in successfully diagnosing diabetic patients and thus, providing them with timely medication. The proposed approach reduces the demand on clinics, especially for under-resourced clinical settings, due to the simplicity of ECG based approach presented herein and due to targeting only specific hours in the circadian cardiac cycle. Furthermore, the developed screening tool presented herein allows clinicians to perform a complete check-up on the overall health condition of the diabetic patient enrolled, with a focus on discriminating between the existence of only CAN or CAN with additional microvascular complications.

Additionally, a thorough investigation on the use of multiple AI algorithms is provided in this study to evaluate the impact of using several machine learning techniques in enhancing the diagnostic ability of HRV in neuropathy screening for diabetes patients. The reason behind utilizing multiple machine learning algorithms is to provide a wider artificial inspection of HRV features by training models that use higher-order feature space, random sets of decision trees, or neural networks. Thus, it ensures the recommendation of which model to be followed for predicting diabetic patients with different autonomic neuropathy conditions for every clinical testing scheme.

II. MATERIALS AND METHODS

A. DATA SET AND PATIENT ENROLLMENT

The study included a total of 95 participants (Male: 45, Female: 50) enrolled during routine visits to Bangladesh Institute of Health Sciences (BIHS) Hospital between December, 18th 2017, and April, 26th 2018. Out of these participants, 70 were suffering from type 2 diabetes for more than 10 years, and 25 were healthy and considered as the control patients' group. Among the 70 diabetic patients, 66 had complications, i.e., neuropathies, while 4 only did not suffer from any diabetic complications. All participants were Bangladeshi citizens above 40 years of age. The study was approved by the ethical review committee of Bangladesh university of health sciences (BUHS/BIO/EA/17/01) and conforms to the ethical principles outlined in the declaration of Helsinki and the Ministry of Health and Family Welfare of Bangladesh. In addition, a consent form was taken from every participant to be eligible for enrollment in the study. No patients were allowed to participate in the study (exclusion criteria) if they had one of the following conditions; stroke history, heart diseases, diabetes duration of less than 10 years, and the presence of any other pathophysiology that may lead to one or more complications associated with diabetes such as the ones caused by cancer.

All participants undertook a 24-hour Holter ECG recording using Shimmer3 ECG Unit. The device was attached to every patient with 4 ECG leads (5 wires/electrodes). Out of the four leads of the device, namely LARA, LLLA, LLRA, VxRL, the third channel (LLRA) was selected as the reference lead for R-peak calculations as it is the lead connected closely to the chest. First, Consensys Pro application was used to convert the recorded data into a usable computer format (.dat). All recorded ECG signals were sampled at 200 Hz and filtered using low-pass and high-pass filters

TABLE 1. Brief definitions of all heart rate variability (HRV) features [35].

Feature	Definition
Time-domain	
AVNN (ms)	Average N-N interval
SDNN (ms)	Standard deviation of N-N intervals
RMSSD (ms)	Square root of the mean of the sum of squares of differences between adjacent N-N intervals
pNN50 (%)	Percentage of N-N intervals $>$ 50ms
SEM (ms)	Standard error of the average N-N interval
Frequency-domain	
BETA	Slope of the linear interpolation of the spectrum for frequencies less than VLF band upper bound
HF Norm (%)	Normalized high frequency (HF) power
HF Peak (Hz)	Peak frequency in the HF band
HF Power (ms^2)	Power in the HF band
LF Norm (%)	Normalized low frequency (LF) power
LF Peak (Hz)	Peak frequency in the LF band
LF Power (ms^2)	Power in the LF band
LF/HF	Ratio between LF power and the HF power
Total Power (ms^2)	Total power in both frequency bands
VLF Norm (%)	Normalized very-low frequency (VLF) power
VLF Power (ms^2)	Power in the VLF band
Non-linear	
SD1 (ms)	Standard deviation of N-N intervals along the perpendicular to the line-of-identity
SD2 (ms)	Standard deviation of N-N intervals along the line-of-identity
Alpha1	Detrended fluctuation analysis with low-scale slope
Alpha2	Detrended fluctuation analysis with high-scale slope
Sample Entropy	Complexity of physiological time-series signals
Fragmentation	
PIP (%)	Percentage of inflection points in the N-N interval
IALS	Acceleration/deceleration segments inverse average length
PSS (%)	Percentage of short segments
PAS (%)	Percentage of alternation segments

of 45 Hz and 0.5 Hz cutoff frequencies, respectively. Then, the corresponding R-peak locations were annotated carefully on the ECG signal using LabChart 7 Pro application through cyclic measurements. In this application, the ECG signal was manually cleaned and marked for time information including sleeping, resting, exercising, and hour-indexing segments. Then, it was exported into MATLAB with all the marking information (R-peak and time frames) for further analysis.

Initially, the exported signal was converted into an inter-beat interval (IBI) signal from the R-peak information to form the HRV data for every hour starting with hour 00:00-01:00 (1:00 PM). Then, the IBI-HRV signal was filtered twice to remove outliers using MATLAB function (`filloutliers()`). Furthermore, it was further denoised from any artifacts, noise, or wrong labeling using the signal-dependent rank order mean (SD-ROM) [33] and adaptive filtering [34] techniques to ensure no abnormalities in the HRV data. For 7 patients, the Holter recording was not fully completed (full 24-hour), thus, they had missing ECG and HRV data. This has caused a reduction of hours-count for these patients out of 24 hours of a day. Therefore, at certain hours, the total number of patients included in the study was reduced.

B. HRV FEATURES EXTRACTION

HRV is a series of R-peaks of normal beats (N-N intervals) representing heart rate values. Each ECG signal was annotated to obtained a complete HRV data. Each patient's data was initially arranged to start from hour 00:00-01:00 (1:00 AM) and complete the full 24-hour circadian rhythm recording by hour 23:00-00:00 (12:00 AM). For each hourly HRV data, 25 HRV features as shown in Table 1 were extracted for every 5-minute segment from time-domain, frequency-domain, non-linear, and fragmentation metrics using PhysioNet toolbox [35] and MATLAB R2021a. In time and frequency domains, feature were extracted according to the task force of the European society of cardiology [36]. In addition, non-linear indices were extracted based on the Poincare plot, de-trended fluctuation analysis (DFA), and multi-scale entropy (MSE) [37], [38]. Furthermore, the newly introduced HRV fragmentation metrics were included in the study and were extracted with regards to Costa *et al.* [39]. A brief description of each HRV feature is provided in Table 1. A total of 300 values per HRV feature were collected for every hourly segment, that is 25 HRV features in each of the 12 5-minute segments in an hour, yielding a total of 7,200 feature values in 24-hour data for every patient.

C. HIERARCHICAL CLINICAL TESTING SCHEMES

To provide a complete diagnosis for patients with CAN, four tests were applied sequentially on patients' HRV data on an hourly basis (Fig. 2). These tests allow for better screening of diabetic patients, especially those who are suffering from cardiovascular neuropathy, thus, providing them the suitable medication at the optimal time. Each test is described briefly in the following subsections.

1) FIRST TEST: DIAGNOSING DIABETIC PATIENTS

This test targets discriminating between diabetic patients and those participants at the healthy (control) group. Under clinical settings, it is essential to initially identify if a patient is suffering from diabetes before proceeding with additional testing phases. In this test, the whole data set was used as input to the proposed machine learning models for training and testing (more details in Section II-D). The data set used herein consisted of 70 diabetic patients and 25 participants from the control patients group.

2) SECOND TEST: IDENTIFYING PATIENTS WITH COMPLICATIONS

After identifying patients who are diabetic according to their extracted HRV features, this test ensures the detection of patients whether being suffering from further complications, i.e., AN, or not. To train the models, a total of 66 diabetic patients with diabetic complications and 4 patients with no complications were used in this test. Although this test included less number of diabetic patients with no complications, it was considered a challenge to carefully

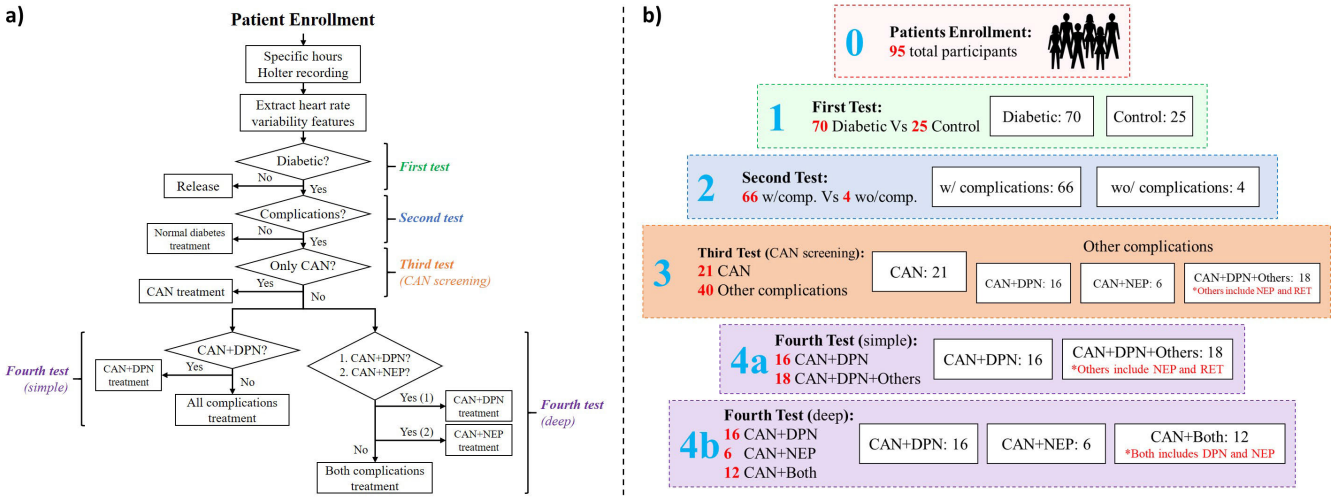


FIGURE 2. **The proposed procedure for diagnosing patients based on their diabetic status. (a) The clinical diagnosis flowchart where each patient goes initially through data acquisition (acquiring 24-Hour Holter electrocardiography (ECG) recording at specific time-stamps and extracting the corresponding heart rate variability (HRV) features). Four tests are then applied sequentially to check for: being healthy or diabetic, having complications or not, suffering from cardiovascular autonomic neuropathy (CAN) only, or having multiple complications along with CAN including diabetic peripheral neuropathy (DPN), nephropathy (NEP), and retinopathy (RET). The correct treatment after each test can be suggested to the patient based on his diabetic status. (b) Detailed representation of each test for the proposed machine learning approach. Each test shows the number of patients' 24-hour data included to train the models. For the third test (CAN screening), the model was trained using CAN patients' data only and used to give a probability (score) for a patient as suffering from CAN. More details on the dataset and the machine learning approach are given in Section II.**

train and test the machine learning models. Furthermore, we hypothesize that patients suffering from different AN types should be easier for the models to discriminate from other normal diabetic patients. It is to be noted that all the subjects considered in this study are above 40 years of age with 10 years of history of diabetes. Thus, it is hard to get someone who has a record of diabetes for 10 years but does not have any AN complications. To compensate for this class bias, i.e., the presence of a less number of train and test samples, feature augmentation is used in the machine learning model for the second test. The details are explained in Section II-D.

3) THIRD TEST: CAN-ONLY SCREENING TOOL

Patients who are detected as suffering from additional diabetic neuropathies were subjected to this test. The third test is considered as a CAN-only screening tool, where it is focusing on the evaluation of patients based on their degree (score) of being suffering only from cardiovascular complications associated with diabetes. In this test, a one-class machine learning model based on SVM was used for training using CAN patients' data only (more details in Section II-D2). The outcomes of the trained model are scores defining a probability for every input data as being diagnosed as CAN-only. The higher the probability (more than 0.70), the more it resembles HRV data coming from a CAN-only diabetic patient. In this test, a total of 21 patients with CAN-only were included in the trained model, and the trained models were further used to predict 40 CAN patients' data with additional neuropathies consisting of 16 CAN + DPN, 6 CAN + NEP, and 18 CAN + DPN + Others, where 'Others' includes NEP and RET.

4) FOURTH TEST: DETECTING ADDITIONAL COMPLICATIONS

The fourth test provides a thorough view of the patients who are suffering from additional diabetic neuropathies alongside CAN. It is split into two parts; simple and deep scenarios. In the simple scenario, a total of 34 patients were used as inputs to machine learning models for training and testing, out of which 16 were suffering from CAN + DPN and 18 were suffering from CAN + DPN + Others, where 'Others' includes NEP and RET. On the other hand, the deep scenario included the same number of patients, however, they were split deeply as 16 CAN + DPN, 6 CAN + NEP, and 12 CAN + Both, where 'Both' includes DPN and NEP complications at the same time.

D. MACHINE LEARNING AND TRAINING SETTINGS

The ability to use machine learning algorithms for the training and prediction of diabetic patients is essential to provide an efficient diagnosis process. Here, 4 models were used, including support vector machine (SVM) and one-class SVM, Random forest (RF), random under-sampling boosting (RUSBOOST), and convolutional neural network (CNN), to evaluate the performance of several machine learning algorithms. Before starting the training and testing phases for each of the aforementioned tests, HRV data pre-processing and preparation steps were performed to maximize the performance. These steps include best HRV features selection as well as HRV data augmentation. A complete description

of the whole machine learning approach is provided in the following subsections.

1) PRE-PROCESSING AND DATA PREPARATION

The pre-processing and data preparation steps were applied for each and every test to ensure a maximized performance within each machine learning model.

First, a feature selection approach was followed based on a univariate chi-square (χ^2) test to select the most significant HRV features on an hourly basis. In this test, a statistical hypothesis analysis is performed for every HRV feature to test whether the observed calculations match with the expected ones, i.e., patient group label. Furthermore, it provides a significant difference p-value measure (p-value $<$ 0.05) between categories based on the statistical calculations and expectation [40]. A low p-value for a feature denotes that this feature is most likely dependent on the group label, therefore, it is important for predicting the group. Thus, a score of importance is returned for every HRV feature used in the test as $score = -log(p)$. This test was performed in MATLAB R2021a using function fscchi2(). It is worth to be mentioned that for the third test, the feature selection approach was based on Laplacian scores [41] that relies on the nearest neighbor similarity graph. This approach was selected due to have only a single class (CAN) used for training, thus, feature selection was considered as an unsupervised clustering problem to identify most significant features and eliminate outliers [42].

Second, HRV data augmentation was essential in preventing each machine learning model from over-fitting. If the model tends to over-fit, it does not allow for optimal training parameters as the model trains to learn exactly the input data. In other words, the model does not generalize well the trained parameters for the training data to the new unseen data [43]. To prevent such phenomena, we provide each model with 500 additional variations of HRV data per category. These variations include only slight changes in the amplitude of every feature, as it was preferred not to over-augment the original data and thus lose important information in the HRV data. Additionally, data augmentation allows the model to face various changes in the data during training. These slight changes enhance the ability of the model to generalize the parameters and maximize the performance on predicting unseen data, as the model becomes exposed to extra information that was not available in the original HRV data.

It is worth noting that data imbalance, which is a common issue occurring in most medical data, was handled carefully in each machine learning model prior to training and testing to prevent biased predictions. First, HRV data augmentation, as mentioned previously, was used to create 500 additional variations for each class, therefore, the training was allowed to run on a more evenly distributed data and classes. Second, each machine learning model allows for the utilization of prior probabilities, i.e., initial weights, for every class. Therefore, the training parameters in each model were optimized according to these prior probabilities for each class to prevent it from learning on biased parameters in favor of one class over the other. More information on the optimization of each model's weights are provided in details in the following subsections. In addition, extra information about data imbalance handling in machine learning can be found in [44]–[46].

2) SUPPORT VECTOR MACHINE (SVM)

Support vector machine (SVM) is a commonly used machine learning and data mining algorithm for its robustness in classification and regression problems. It has the ability to find an optimal hyper-plane (kernel) in a high N^{th} dimension, where N is the number of features, that separates data from two categories with a maximum gap of margin [47].

In terms of mathematical formulation, given a set of features vectors $X = \{x_1, x_2, \ldots, x_N\}$ and their associated labels $Y = \{y_1, y_2, \ldots, y_N\}$, where N is the total number of features, SVM goal is to minimize a regularized risk function R given by,

$$\min_{w,b,\zeta,\zeta^*} R(w, b, \zeta, \zeta^*) = \frac{1}{2}w^2 + C\sum_{i=1}^{N} \zeta_i + \zeta_i^* \qquad (1)$$

where w and b describe the selected hyperplane parameters, C is used to control the trade-off between risk minimization and potential over-fitting, and ζ_i and ζ_i^* are slack variables used to measure the degree of misclassification of the i^{th} data point x and to cope with infeasible constraints.

In addition to multi-class classification problems, SVM allows for one-class classification, which is considered as part of unsupervised learning. It is commonly used as an outlier detector, where the model is trained to distinguish training data from any other unrelated data [48]. The selected class for training is recognized as the positive class, and therefore, the model tends to detect which new objects are closely representing this positive class or are distinguished as an outlier with no close relation [49]. The model returns scores (probabilities) to show the degree of matching between samples and the positive class. Similar to multi-class SVM, a hyperplane is used to maximize the margin between samples and the origin.

In this work, a multi-class SVM was used for training models in the first, second, and fourth tests. On the other hand, a one-class SVM was used for training the model in the third test by considering CAN class as the positive class (CAN screening). The kernel function was selected to be a non-linear radial basis function (RBF) with fine-tuned hyperparameters to ensure a maximized performance. To prevent the model from training on biased parameters due to data imbalance in the clinical testing schemes, a prior probability was utilized to initialize the initial training weights per class. In this model, the prior probability (P_i) was calculated empirically based on the number of samples within every class (N_i) relative to the total number of samples (N) as follows,

$$P_i = 1 - \frac{N_i}{N} \qquad (2)$$

where i refers to the index of the selected class.

It is worth noting that for one-class SVM, no prior probabilities were incorporated, as the model tends to train only using a single class with no possibility of data imbalance.

3) RANDOM FOREST (RF)

Random forest (RF) is a form of decision tree, also known as classification and regression tree (CART), where a set of tree-like attribute nodes is connected by a set of sub-trees of decision nodes [50], [51]. This algorithm is considered as a bootstrap aggregation technique that uses the concept of bagging. Given a set of feature values $X = \{x_1, x_2, \ldots, x_N\}$ and a set of responses (labels) $Y = \{y_1, y_2, \ldots, y_N\}$, where N is the total number of samples, the model trains by repeatedly (B times) selecting random features to fit trees.

To provide a prediction, each decision node is calculated based on the corresponding prediction consequences including the resource cost, outcomes chances, and utility. At each tree, the prediction process starts by assigning an instance to its root node. Then, for each of the following sub-nodes, the outcomes are calculated sequentially. Once a leaf is encountered, the tree-like nodes stop and an instance is assigned with a prediction. The sum of all instances and predictions (voting mechanism) forms the final decision made by the tree model [52]. Using the trained model (G), the prediction of an unseen data (X') is given by,

$$Y' = \frac{1}{B}\sum_{b=1}^{B} G(X') \tag{3}$$

In this work, a range of 10-100 decision trees was used to build the model. The decision on the number of trees for every test mentioned in Section II-C was fine-tuned to ensure the maximum possible performance. In RF, handling the imbalance in the data at every clinical testing scheme was handled similarly to SVM (mentioned in the previous subsection). The prior probability was calculated empirically for every class as shown in Equation 2.

4) RANDOM UNDER-SAMPLING BOOSTING (RUSBOOST)

A hybrid ensemble algorithm that uses decision trees but with data sampling and boosting techniques is known as Random under-sampling boosting (RUSBOOST) [53]. This algorithm is most suitable whenever there is a huge unbalance in the data used for training. In this algorithm, the majority class is under-sampled by randomly discarding samples during each iteration. In addition, weak learners are built through linear combinations in a process called boosting [54].

Initially, RUSBOOST defines the weights of each class as $D_t = 1/m$, where t is the iteration number and m is the total number of samples for every class. After identifying the class with the majority number of samples, random under-sampling is performed to generate a set of temporary training data (S'_t) that includes an equal number of samples per class after random removal of samples from the majority class. Then, a weak hypothesis (h_t) is created based on the learning process in each iteration and used to update the weights D_t as,

$$\alpha_t = \frac{\epsilon_t}{1-\epsilon_t} \tag{4}$$

$$D_{t+1} = \frac{D_t \alpha_t^{\frac{1}{2}(1+h_t(x_i,y_i)-h_t(x_i,y:y\neq y_i))}}{Z_t} \tag{5}$$

where α_t is the weight update parameter, x_i is a point in the feature space, y_i is a class label, and Z_t is the summation of all weights.

After updating all weights, the final hypothesis (H_t) is created as,

$$H_t(x) = \underset{y\in Y}{\operatorname{argmax}} \sum_{t=1}^{T} h_t(x, y) \log \frac{1}{\alpha} \tag{6}$$

where Y is the set of class labels, and T is the maximum number of iterations.

In this work, the learning rate (shrinkage) was selected to be 0.001 as it was shown that the lower the learning rate, the better the convergence and performance of boosting models [55], [56]. In RUSBOOST, data imbalance was handled through incorporating prior probabilities for every class. The prior probability was calculated as an under-sampling proportion ratio relative to the lowest-represented class. The algorithm uses this ratio to randomly under-sample the majority classes to match with the minority class on every iteration, thus, the training can be performed without any bias within the learned parameters. The prior probability (under-sampling proportion ratio) was calculated as,

$$P_i = \frac{Ni}{N_{minority}} \tag{7}$$

where P_i is the ratio, N_i is the number of samples in the selected i^th class, and $N_{minority}$ is the number of samples in the minority class.

5) CONVOLUTIONAL NEURAL NETWORK (CNN)

Convolutional neural network (CNN) is considered as part of deep learning algorithms that utilizes a feed-forward network that is capable of applying various transnational and rotational invariance analysis on input data [57]–[60]. In CNN, a set of convolutions (dot products) are applied to input data (signals or images) to obtain corresponding deep features contaminated within its nodes or pixels. For an input data $X = \{x_1, x_2, \ldots, x_N\}$, where N is the total number of points, the convolutions are done as follows,

$$c_n^{uj} = h_a(b_j + \sum_{m=1}^{M} w_m^j x_{n+m-1}^j) \tag{8}$$

where u is the layer index, h_a is the activation function, b_j is the bias of the j^{th} feature map, M is the kernel size, w_m^j is the weight of the feature map and filter index m^{th}.

In this work, a one-dimensional (1D) CNN was used to train the model using 1D signal inputs (HRV data). The model was selected to be simple with one convolutional layer of kernel size equalling half of the input data vector size, which

varies from hour to hour based on the number of features selected. Furthermore, a total of 96 filters were used for feature extraction with a stride of 1 and padding of 0. The network was followed by a softmax layer prior to the final classification layer. For CNN, the imbalance in the data at every clinical testing scheme was handled through adjusting the initial weights and corresponding losses for every class. The initial weights were calculated empirically as shown in Equation 2. In addition, these weights were utilized within the final classification layer to calculate a weighted cross-entropy loss (L) as follows,

$$L = -\frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{K} w_i ln(y_{ni}) \tag{9}$$

where N is the total number of samples, K is the total number of classes, w_i is the initial weight for the i^{th} class, and y_{ni} is the output of the softmax function during training on extracted features.

6) TRAINING AND TESTING CONFIGURATION

For training and testing, a leave-one-out scheme was followed in all tests to ensure the inclusion of the maximum possible number of samples within the trained models. Furthermore, it was essential to provide a prediction for each and every patient, thus, it allows for mimicking clinical testing situations, where an already developed/trained model can be in hand, while a new testing sample (patient) is completely hidden to the model and used to provide a prediction. In this scheme, an iterative process is applied by selecting 1 subject as the testing subject, while the remaining subjects are used for training. The process repeats on every iteration until a prediction is given for every patient. During training, data imbalance was handled by the incorporation of prior probabilities as discussed in the aforementioned subsections.

The performance of machine learning models was evaluated for every hour using accurate measurements. In addition, the analysis of the area under the receiver operating characteristic (AUROC) was provided for each model in every test to show the true positive rate (TPR) versus the false positive rate (FPR). In the third test, where a one-class SVM was used, the evaluation was based on the observed scores (positive: ≥ 0.7).

III. RESULTS

The performance of each model was evaluated separately for each clinical test. The training of each model required several minutes (around 1-5 minutes) based on the optimized parameters on an Intel processor (i7-9700) with 32 GBs of RAM and NVIDIA GeForce GTX 1070 graphics processing unit (GPU) of 8 GBs display memory (VRAM). The whole training and testing phases took 5-10 minutes per model due to following a leave-one-out scheme as well as testing the performance on an hourly basis. More information on the results of every clinical test is provided in the following subsections.

A. FIRST TEST

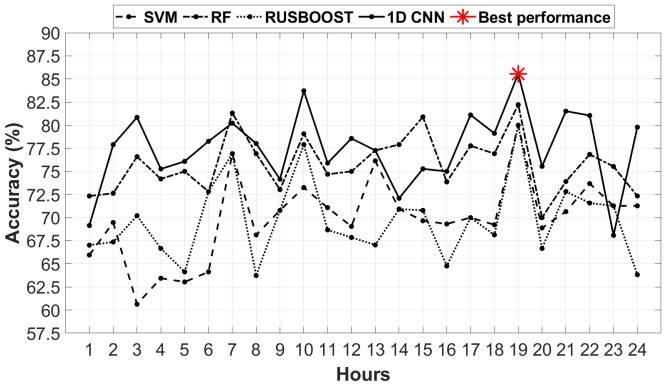
The performance (accuracy) of each machine learning model in the first test is illustrated in Fig. 3. From the figures, the highest performance was observed at hour 18:0-19:00 (7:00 PM) using the CNN model. The discrimination accuracy between diabetic and control patients for the CNN model reached 85.5% compared to 80.0%, 82.2%, and 80.0% for SVM, RF, and RUSBOOST, respectively. It is worth mentioning that hour 09:00-10:00 (10:00 AM) had a good level of accuracy (83.0%) for the CNN model. To analyze deeply the performance of the CNN model at the best hour (7:00 PM), the confusion matrix of predictions is shown in Fig. 3(c). Out of the 68 diabetic patients, 60 were correctly identified as diabetic while the remaining 8 were miss-classified as normal. On the other hand, the majority of control patients (22) were identified correctly with only 5 wrongly classified as diabetic. It is to be noted that the reduced number of the overall patients in each category was due to having missing data for some patients at this specific hour (more information in Section. II-A). Relatively, the analysis of the ROC curves for each category using the CNN model is depicted in Fig. 3(e). The overall area under ROC (AUROC) curves at the best performing hour (7:00 PM) was 0.80 for diabetic and control patients.

B. SECOND TEST

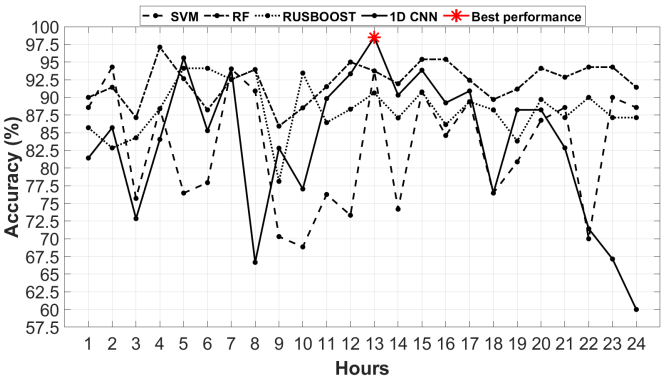
The overall accuracy for the second test throughout the 24-hour HRV segments is shown in Fig. 3(b). From the figure, the best performance was achieved using the CNN model at hour 12:00-13:00 (1:00 PM) with a discrimination accuracy of 98.5% between diabetic patients with complications (W/ Comp.) and without complications (Wo Comp.). Relatively, the SVM, RF, and RUSBOOST models yielded slightly lower accuracy levels of 93.8%, 93.8%, and 90.6%, respectively. The corresponding confusion matrix (Fig. 3(d)) at this hour and using the CNN model (best performance) shows that almost all patients were correctly identified in their corresponding category. Out of the 66 patients with complications, only 2 were miss-classified as patients without complications. On the other hand, all 4 non-complications patients were correctly identified by the model. In addition, the ROC curves for both categories are shown in Fig. 3(f). The overall AUROC curves reached 0.99, suggesting it as a strong machine learning model for testing diabetic patients with or without complications during the afternoon (1:00 PM) time period.

C. THIRD TEST

Using one-class SVM, the third test performance is illustrated in Fig. 4. The 24-hour accuracy measurements are shown in Fig. 4(a) for the CAN-only set, CAN + Others set, as well as the average value when both sets are considered. The CAN + Others set includes additional neuropathies described in Section II-C3. The highest average accuracy level (98.3%) was achieved at hour 06:00-07:00 (7:00 AM) with a 95.2% accuracy for the CAN-only set and 100.0%

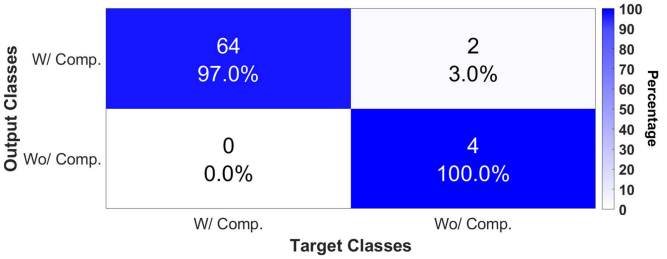


(a) First test: per-hour accuracy

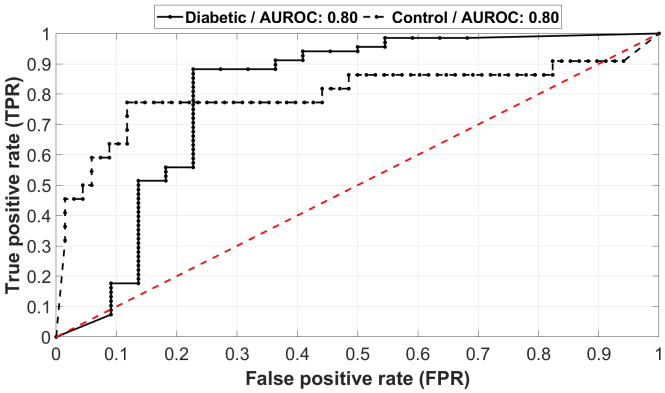


(b) Second test: per-hour accuracy

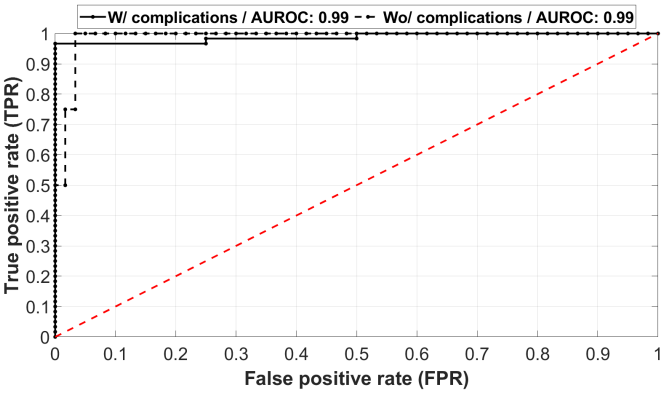
(c) First test (CNN): confusion matrix at 18:00-19:00 (7:00 PM)



(d) Second test (CNN): confusion matrix at 12:00-13:00 (1:00 PM)



(e) First test (CNN): ROC curves at 18:00-19:00 (7:00 PM)



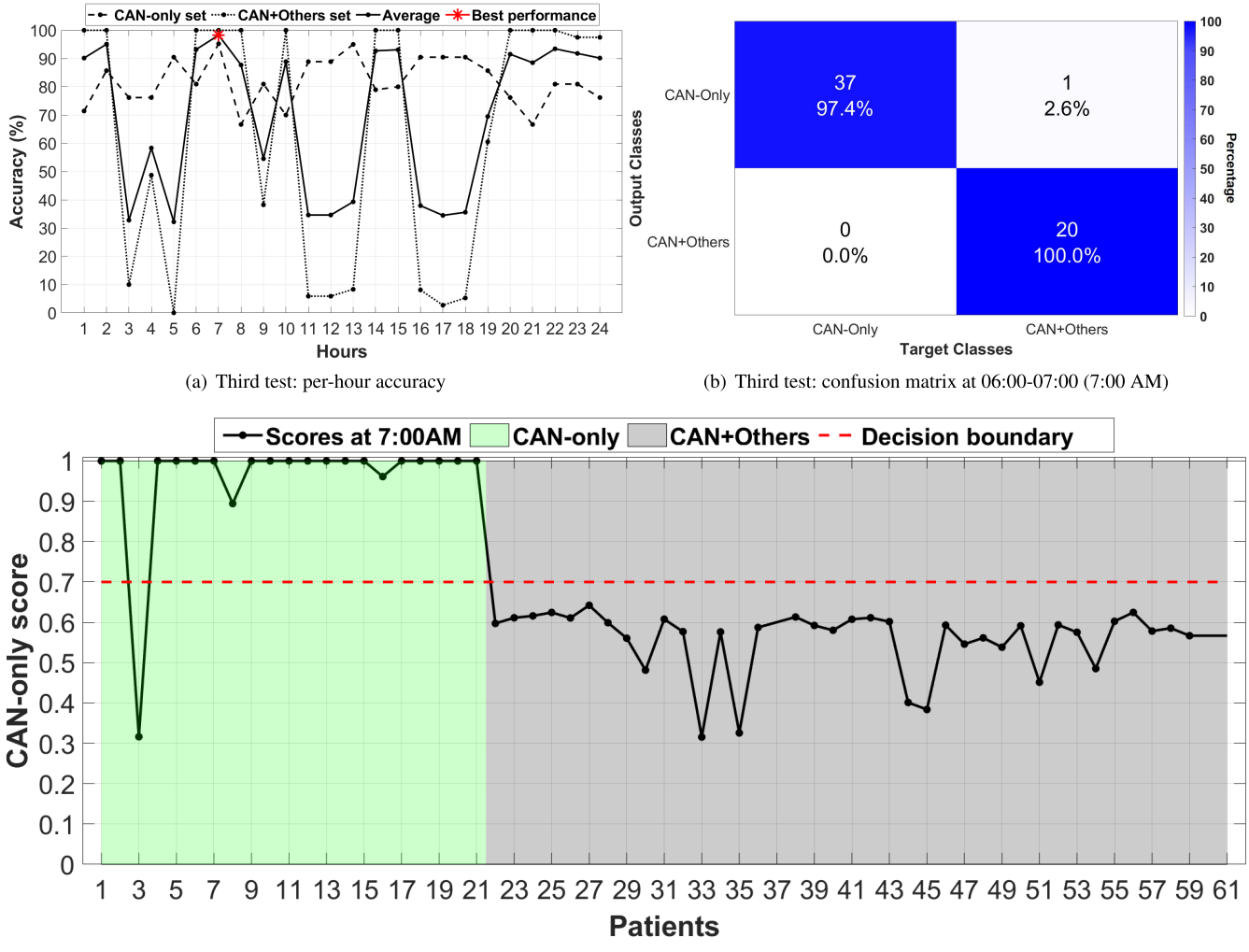
(f) Second test (CNN): ROC curves at 12:00-13:00 (1:00 PM)

FIGURE 3. The performance of machine learning models in the first and second clinical testing schemes. (a,c,e) First test: the 24-hour accuracy measurements using support vector machine (SVM), random forest (RF), random under-sampling boosting (RUSBOOST), and convolutional neural network (CNN) machine learning models, the confusion matrix at the best performing hour (18:00-19:00 (7:00 PM)) and using the best performing model (CNN), and the receiver operating characteristic (ROC) curves at the best performing hour (18:00-19:00 (7:00 PM)), respectively. (b,d,f) Second test: the 24-hour accuracy measurements using SVM, RF, RUSBOOST, and CNN machine learning models, the confusion matrix at the best performing hour (12:00-13:00 (1:00 PM)) and using the best performing model (CNN), and the ROC curves at the best performing hour (12:00-13:00 (1:00 PM)), respectively.

accuracy for the CAN + Others set. Fig. 4(b) shows the prediction performance at this hour within a confusion matrix. Almost all patients were correctly identified according to their original category. However, only a single CAN-only patient was miss-classified as suffering from additional complications. To investigate the scores returned by the one-class SVM model, Fig. 4(c) shows per-patient scores for a positive CAN-only categorization. Most CAN-only patients were correctly identified with a score of 1.00. Only a single patient had a CAN-only score of 0.32, which was wrongly classified as suffering from additional neuropathies attached with CAN. It is worth mentioning that two patients had scores of 0.89 and 0.96, however, they were above the decision boundary set for correctly predicting positive CAN-only patients (0.70). On the other hand, during the best performing hour (06:00-07:00 (7:00 AM)), all CAN + Others patients were correctly scored below the decision boundary (0.70). It can be noted that due to having the characteristics of CAN within their HRV data, many patients had scores ranging between 0.6-0.7.

D. FOURTH TEST

To investigate further the complications attached with CAN diagnosed patients, the fourth test has included simple and deep classifications scenarios for these additional complications (more details in Section II-C4). The performance of both scenarios is shown in Fig. 5 over the 24-hour time interval. The simple classification scenario (Fig. 5(a))



(a) Third test: per-hour accuracy

(b) Third test: confusion matrix at 06:00-07:00 (7:00 AM)

(c) Third test: CAN-only scores at 06:00-07:00 (7:00 AM)

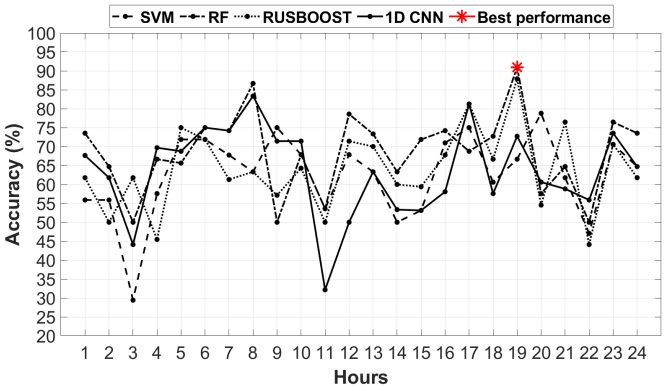
FIGURE 4. The performance of the one-class support vector machine (SVM) model trained using CAN-only diabetic patients. (a) 24-hour accuracy measurements for testing on CAN-only set and CAN + Others set. (b) The confusion matrix at the best performing hour (06:00-07:00 (7:00 AM)). (c) CAN-only scores for every patient in the overall testing set at the best performing hour (06:00-07:00 (7:00 AM)). Positive CAN-only is indicted if the score is ≥0.7.

had a maximum accuracy of 90.9% at hour 18:00-19:00 (7:00 PM) using the RF machine learning model. Furthermore, the RUSBOOST model has returned a very close accuracy level of 87.9%. Similarly, the deep classification scenario (Fig. 5(b)) gave the maximum performance (accuracy = 84.4%) at the same 7:00 PM hour. However, the highest performance was achieved through the CNN trained model. Nevertheless, RF and RUSBOOST models had very close accuracy levels of 79.4% and 81.3%, respectively. It is worth noting that models based on decision trees, i.e., RF and RUSBOOST, have performed efficiently in these tests. Both of the tests' confusion matrices show that the best performing models (RF and CNN) predicted almost every patient in his correct category (Figs. 5(c),(d)). Furthermore, the ROC curves for both tests shown in Figs. 5(e),(f) had an area of more than 0.90 for all groups. In the simple scenario, the AUROC reach up to 0.91 for predicting CANDPN and CANDPN + Others. On the other hand, in the deep scenario, the model predicted CANDPN, CANNEP, and CANBoth with AUROC levels of 0.93, 0.93, and 0.95, respectively.

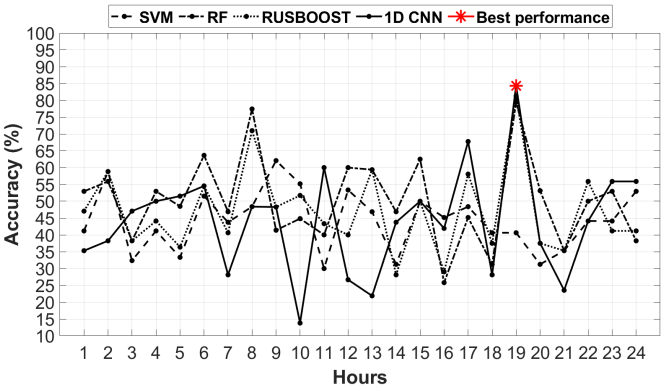
E. MOST SIGNIFICANT HRV FEATURES

As previously mentioned in Section II-D1, the high performance observed in the four clinical testing schemes was based on the selection of most significant 5-minute HRV features at every hour using chi-square and Laplacian score tests. Table 2 shows these features for each testing scheme during the best performing hour. Each feature included originally 12 5-minute segments in every hour, and the feature selection approach selected only best segments (most significant).

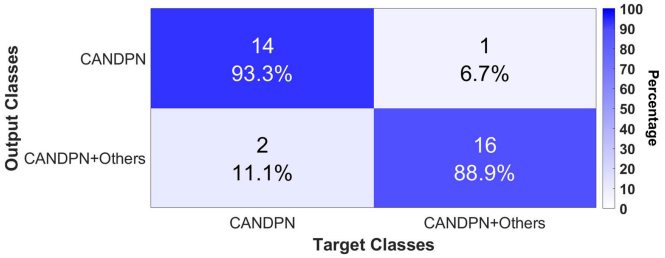
From the table, it can be noticed that the first test had no significant fragmentation features. Furthermore, besides frequency-domain and non-linear features, only two time-domain features were included, namely SDNN and SEM, mostly during 20 to 50 minutes of Hour 7 pm. In the second test, most of the features at different 5-minute segments were included. However, it is worth noting that only SD1 and SD2 were included from the non-linear metrics. The third test included only two features from the frequency-domain, namely Total Power and VLF Power, with no features from the time-domain, non-linear, and fragmentation



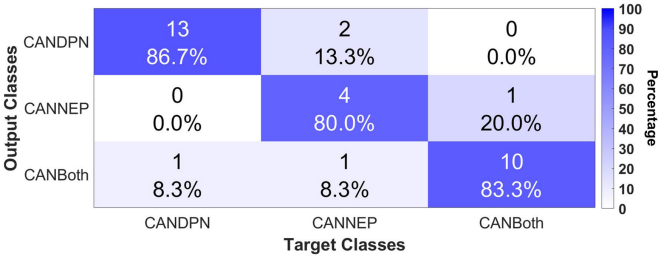
(a) Fourth test (simple): per-hour accuracy



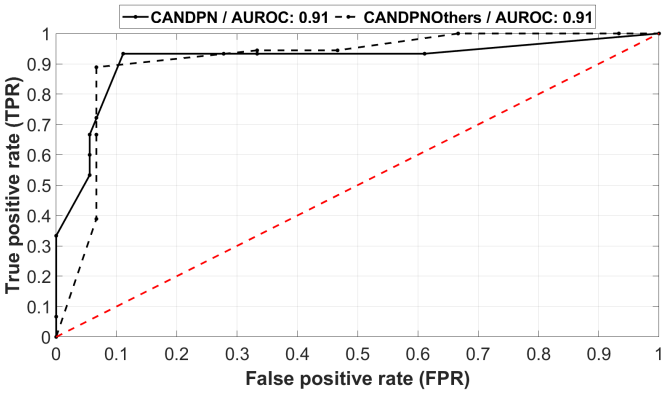
(b) Fourth test (deep): per-hour accuracy



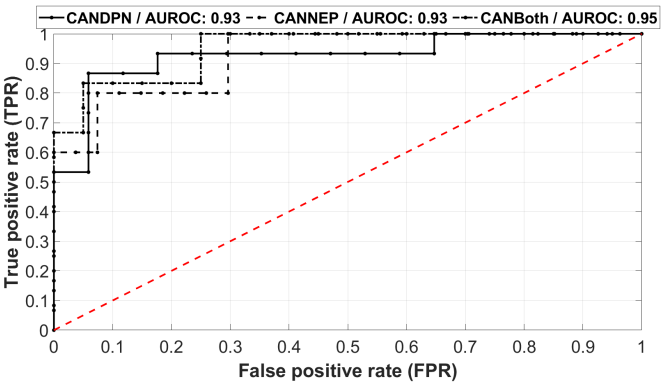
(c) Fourth test (simple - RF): confusion matrix at 18:00-19:00 (7:00 PM)



(d) Fourth test (deep - CNN): confusion matrix at 18:00-19:00 (7:00 PM)



(e) Fourth test (simple - RF): ROC curves at 18:00-19:00 (7:00 PM)



(f) Fourth test (deep - CNN): ROC curves at 18:00-19:00 (7:00 PM)

FIGURE 5. **The performance of machine learning models in the fourth clinical testing scheme. (a,c,e) Fourth test (simple): the 24-hour accuracy measurements using support vector machine (SVM), random forest (RF), random under-sampling boosting (RUSBOOST), and convolutional neural network (CNN) machine learning models for the simple classification scenario, the confusion matrix at the best performing hour (18:00-19:00 (7:00 PM)) and using the best performing model (RF), and the receiver operating characteristic (ROC) curves at the best performing hour (18:00-19:00 (7:00 PM)), respectively. (b,d,f) Fourth test (deep): the 24-hour accuracy measurements using SVM, RF, RUSBOOST, and CNN machine learning models for the deep classification scenario, the confusion matrix at the best performing hour (18:00-19:00 (7:00 PM)) and using the best performing model (CNN), and the ROC curves at the best performing hour (18:00-19:00 (7:00 PM)), respectively.**

metrics. These features were during the first and last 10 minutes for the Total Power and VLF Power and during the 7th and 8th 5-minute segments of Total Power. Finally, the fourth test (simple and deep scenarios) had features from all HRV metrics at different 5-minute segments. In addition, only SD1 and SD2 were included from the non-linear metrics.

IV. DISCUSSION

This study demonstrated the significance of using 24-hour HRV features in diagnosing diabetic patients suffering from cardiovascular complications. Furthermore, it provided a complete machine learning-based clinical approach to screen diabetic patients suffering from CAN and other complications associated with it. The high performance achieved in each of the four clinical testing schemes at specific hours in the heart's circadian rhythm strongly suggests HRV embedded within machine learning algorithms as an assistive screening tool in diabetes diagnosis and treatment.

A. 24-HOUR HRV DATA

The ability to use HRV and its corresponding features overcomes many of the limitations faced when using the current gold standard way to diagnose diabetic patients with CAN, i.e., Ewing tests. Although these tests provide high levels of accuracy [27], they are not considered suitable for patients with respiratory dysfunctions related to the heart or

TABLE 2. Most significant ($p < 0.05$) 5-minute heart rate variability (HRV) features during the best performing hours for each clinical testing scheme. Each HRV feature had 12 5-minute segments during each hour.

Clinical testing schemes	5-minute HRV features during the best performing hours			
	Time-domain	*Frequency-domain*	*Non-linear*	*Fragmentation*
First test Diagnosing diabetic patients (18:00-19:00)	SDNN (18:20-40) SEM (18:20-50)	HF Peak (18:05) HF Power (18:10, 18:30, 18:50) LF Peak (18:20) LF Norm (19:00) LF Power (18:00-19:00) LF/HF (18:25-55) Total Power (18:20-40) VLF Norm (18:00-19:00) VLF Power (18:10-50)	SD1 (18:05, 18:15, 18:30, 18:50) SD2 (18:20, 18:35) Alpha1 (18:30, 18:50) Alpha2 (18:40)	N/A
Second test Identifying patients with complications (12:00-13:00)	SDNN (12:30-35) RMSSD (12:15-55) pNN50 (12:10, 12:30, 12:40-45) SEM (12:10, 12:30-45)	BETA (12:35) HF Peak (12:45) HF Power (12:10, 12:30-40) LF Norm (12:25-35) LF Peak (12:45) LF Power (12:00-13:00) Total Power (12:10) VLF Norm (12:45) VLF Power (12:10)	SD1 (12:15, 12:25-40, 12:55) SD2 (12:30, 12:50-13:00)	PIP (12:20) IALS (12:20) PAS (12:40)
Third test CAN-only screening tool (06:00-07:00)	N/A	Total Power (06:00-10, 06:35-45) VLF Power (6:00-10, 06:50-07:00)	N/A	N/A
Fourth test - simple Detecting additional complications (18:00-19:00)	SDNN (18:25) RMSSD (18:20, 18:50) pNN50 (18:20, 18:45-19:00) SEM (18:25)	HF Peak (18:10) HF Power (19:00) LF Peak (18:25) LF Power (18:10) LF/HF (19:00) Total Power (18:45) VLF Norm (18:40)	SD1 (18:20, 18:50) SD2 (18:25-30)	PSS (19:00) PAS (18:30)
Fourth test - deep Detecting additional complications (18:00-19:00)	AVNN (18:05,18:20) SDNN (18:20-25) RMSSD (18:10, 18:20, 18:55) pNN50 (19:00) SEM (18:20)	HF Power (19:00) LF Peak (18:25-30) Total Power (18:45)	SD1 (18:10, 18:20, 18:55)	PIP (19:00) IALS (19:00) PSS (18:25, 19:00)

- N/A: Not applicable

for patients with obesity or arthritis. Thus, HRV fills this gap by being a simple diagnostic approach that requires only a good quality Holter ECG recording taken from the subject. Furthermore, for 24-hour clinical evaluation, ECG and its corresponding HRV features are considered as an easier and faster method. In addition, it allows for understanding the heart changes, as a result of diabetic neuropathies, throughout the circadian cardiac rhythm. This overcomes the currently followed approach that analyzes only 20-minute portion of HRV [8], [61] by covering the whole 24-hour time interval.

B. SIGNIFICANCE OF 5-MINUTE HRV FEATURES

We would like to underscore the fact that in our study, as HRV features are collected from each 5-minute interval of ECG data during 24 hours period, only 5 minutes of ECG data of the patient within an effective time window will suffice for his or her screening. S/he does not need to undergo 24-hour intensive measurement protocol. Observing all the four tests, 6:00 PM - 7:00 PM and 6:00 AM - 7:00 AM time window can be suggested for the collection of 5-minute data so that the maximum to close to maximum accuracies are achieved in those four tests.

HRV features such as those extracted from time and frequency domains usually elaborate on the sympathetic and parasympathetic activities of the heart. In diabetes, early CAN cases could be identified through a decrease in the average normal-to-normal (AVNN) and standard deviation of normal-to-normal (SDNN) variability in HRV [62], [63].

Furthermore, a parasympathetic dysfunction in patients with type 2 diabetes was observed associated with a reduction in the high-frequency power (HF Power) values of HRV, which could be related directly to measures of vagal nerve integrity [64], [65]. In addition, an association with carotid atherosclerosis in diabetic patients was found with a reduced value in the low-frequency power (LF) [66], [67]. Such variations in these metrics may be useful when trying to discriminate between diabetic and healthy subjects, which have been observed in this study on an hourly basis. It was interesting to observe that HRV features in time and frequency domains were most likely significant during the last 40 minutes of an hour (7 PM, Table 2). This illustrate more cardiac activity toward the evening hours over the afternoon hours. Additionally, we have observed only two features, namely Total Power and VLF Power, in CAN-only detection. Both features are biomarkers that imply cardiovascular autonomic dysregulation and impaired parasympathetic reactions [68], [69], which supports the findings in this study as features used for discriminating CAN-only from other microvascular complications. Additionally, the relation between frequency domain features and CAN in diabetic patients was evaluated previously in literature observing an overall reduction in the power of all spectral bands [70], [71]. In [71], researchers have found a significant drop ($p < 0.001$) in the Total Power HRV feature, which have reached its minimum value during the morning time period (matching with the findings in the current study at 7 AM). Furthermore, features such as the Total Power and LF Power had their lowest values during the morning and evening time periods compared to their values in the afternoon and night, which elaborates on the high performance achieved during 7 AM and 7 PM in the current study when diagnosing CAN-only patients. In [72], a significant difference was observed in the LF/HF HRV feature between the day and night time periods, with an overall decrease in values during the night time. Furthermore, in [73], an overall decrease in HRV parameters, including time and frequency domains, with a significant differences was observed relative to control subjects or between patients with or without complications. In addition, Total Power and VLF Power were from a set of the most preferred parameters to measure improvements in CAN patients during type 1 diabetes treatment [74].

Furthermore, non-linear metrics provide a great indication for diabetic patients, especially SD1 and SD2 features that are correlated with parasympathetic and sympathetic modulation, respectively [75]. Due to diabetes, SD1 usually decreases in patients due to weaknesses in parasympathetic regulation (peripheral neuropathy), whereas SD2 usually increases in patients because of compensatory sympathetic input [76]. These two features were the most significant ones across the four clinical testing schemes with 5-minute segments after the first 20 minutes of the recording (Table 2). It is worth noting that the sample entropy feature was not important in any of the four tests, which is usually associated with an overall reduction in the complexity of glucose dynamics in diabetic patients [77]. It is also interesting to utilize HRV fragmentation features in diabetes diagnostics, as these features are still very recent and usually give an indication about the importance of the changes in the heart rate acceleration and deceleration in diabetic patients [78].

C. MACHINE LEARNING AS A SCREENING TOOL

It is considered essential to be able to gain the benefits of the recent advances in machine learning in the area of diabetes diagnostics. Although the involvement of a clinician in the diagnosis and treatment of diabetic patients is a must, machine learning is expected to provide an early-stage screening that can prevent many complications from further developments. Compared with the current gold standard in CAN screening, the Ewing reflex tests, the proposed approach can overcome issues that these tests may cause such as being cumbersome especially for patients with cardiorespiratory dysfunction, frailty, and severe obesity. Thus, for early and continuous diagnosis, the proposed machine learning-based tool can ease the screening of diabetic patients as well as providing timely and efficiently medication procedure. Furthermore, it can reduce the demand for medical doctors, especially when the data available are infinite, while at the same time provide a complete diagnosis for every patient efficiently. Therefore, a pre-trained machine learning model can make the process faster and less stressful for healthcare providers and practitioners.

Additionally, many diabetic complications exhibit a silent and non-symptomatic nature, i.e. CAN subject with silent myocardial ischemia. Thus, continuous monitoring of diabetic patients who are suspected to develop additional diabetic neuropathies can be easily performed through machine learning models. These models can automatically analyze, diagnose, and predict patients at a very early stage of the disease without the need for repeated visits by patients to clinics and hospitals. This becomes very effective especially for diabetes diagnostics, where most of the patients are from the elderly category. The high performance achieved by the RF and 1D CNN models strongly suggests them as suitable machine learning algorithms for diabetes analysis. 1D CNN was found effective in discriminating between diabetic and control patients (accuracy: 85.5%), between diabetic patients with and without complications (accuracy: 98.5%), and in deeply analyzing complications associated with CAN (accuracy: 90.9%). By analyzing this performance, it can be observed that although HRV feature extraction is essential in diabetes diagnosis, it is also important to obtain hidden and internal characteristics between these features. Utilizing a CNN model allowed for a better interpretation of those features within convolutions layers. Thus, the performance significantly increases in most of the clinical testing schemes. This was supported by usually having the RF model as the best performing model in the simple scenario of CAN and associated complications classification (accuracy: 90.9%) and as the second-best performing model in most of the other schemes. These findings indicate the need for

a deeper model, such as decision trees, to be able to better discriminate between diabetic patients in accordance to their associated complications. Lastly, it was quite promising to observe a high performance using a one-class SVM model for CAN-only screening (accuracy: 98.3%), as the model was trained over one patients' group only (focused training on CAN). This strongly suggests this machine learning model as an assistive tool for clinicians to provide a quicker diagnosis for patients who are only suffering from cardiovascular complications such as CAN, and thus, providing them with the needed medications at optimal timing.

D. CLINICAL RELEVANCE

The testing schemes followed in this work provide an important clinical diagnosis procedure for medical doctors when evaluating and diagnosing diabetic patients. As the scenarios proceed sequentially, models become deeper in analyzing diabetic patients with simple and complicated combinations of autonomic neuropathies. The first test was essential, as it is important to be able of discriminating between healthy and diabetic patients before proceeding with any further clinical tests. This could save a lot of time, expenses, and efforts that may result from performing extra unnecessary testing on non-diabetic patients. Furthermore, the second test identifies those who are having additional autonomic neuropathies, which have shown high levels of performance using the proposed machine learning algorithms. The third test sets the core of this work by providing a screening tool for CAN-only patients. Such clinical test ensures providing a better diagnosis to the cardiovascular and heart functionality, as it discriminates between the causes of only CAN and other complications associated with it. Furthermore, given the silent nature of this complication, it makes it even harder to be diagnosed properly, especially when combined with other microvascular complications. The high performance achieved in this test through machine learning can prevent many under-diagnosed CAN-only cases. The schemes end by additional analysis for these extra complications associated with CAN, which under clinical circumstances, is considered very effective when medical treatment is required. As a medical procedure for only CAN may not return efficient treatment if the extra complications were not identified correctly.

The tools presented herein not only provide high levels of accuracy in diagnosing diabetic patients but also sets a base for medical testing of diabetes using HRV embedded within a machine learning framework. For example, the observations found in this work suggests specific hours for a better diagnosis of the diabetic status. It was repeatedly observed that the hour of 7:00 PM is crucial for screening CAN efficiently. Furthermore, the hour 7:00 AM was also found critical for screening CAN-only patients. These observations could be related to the changes in the functionality of the heart during the early morning and evening hours, which have been previously discussed in many research works on cardiovascular chronopharmacology [79], [80]. In addition, these hours are usually known as times where higher mortality rates are usually observed due to cardiac infarctions and heart disturbances [81].

E. STUDY LIMITATIONS

Although this study shows that machine learning-based models have performed efficiently in diabetes and associated microvascular complications diagnosis, it has a number of shortcomings. First, despite of providing a complete and successful clinical testing approach, further research on a wider cohort of patients across different countries is essential to support the observations found in this study. Second, although machine learning-based models are automated, fast, and can be used as an assistive tool in clinical situations, more investigations are required on the actual design and implementation of such system after careful consideration of all clinical resources. Third and last, more validation needs to be made under a longitudinal follow-up study to investigate further the efficiency of the proposed approach on diagnosing diabetic patients, especially those suffering from CAN. Further, future works may consider forecasting the existence of CAN in non-CAN diabetic patients on a long-term.

V. CONCLUSION

This study suggests HRV features constructed from 5-minute ECG data collected over 24 hours duration of a day as a good screening tool for cardiovascular complications caused by diabetes. Furthermore, it elaborates on the use of such features in a machine learning framework for faster and automated diagnosis of diabetic patients. The high levels of performance achieved in four clinical testing schemes obtained at a specific hourly time window of a day paves the way towards utilizing these tests in clinical practice for diabetes diagnosis. In addition, it provides a deeper understanding of the microvascular complications associated with diabetes, particularly CAN, and the feasibility of efficiently screening them in a continuous and fast manner.

REFERENCES

[1] *Definition and Diagnosis of Diabetes Mellitus and Intermediate Hyperglycaemia: Report of a WHO/IDF Consultation*, World Health Org., Geneva, Switzerland, 2006.
[2] F. Aguiree, A. Brown, N. H. Cho, G. Dahlquist, S. Dodd, T. Dunning, M. Hirst, C. Hwang, D. Magliano, C. Patterson, and C. Scott, "IDF diabetes atlas," Int. Diabetes Fed., Basel, Switzerland, Tech. Rep., 2013. [Online]. Available: https://www.idf.org/component/attachments/attachments.html?id=813&task=download
[3] World Health Organization (WHO). *Diabetes*. Accessed: Apr. 13, 2021. [Online]. Available: https://www.https://www.who.int/news-room/fact-sheets/detail/diabetes
[4] A. I. Vinik, C. Casellini, H. K. Parson, S. R. Colberg, and M.-L. Nevoret, "Cardiac autonomic neuropathy in diabetes: A predictor of cardiometabolic events," *Frontiers Neurosci.*, vol. 12, p. 591, Aug. 2018.
[5] R. E. Maser, B. D. Mitchell, A. I. Vinik, and R. Freeman, "The association between cardiovascular autonomic neuropathy and mortality in individuals with diabetes: A meta-analysis," *Diabetes Care*, vol. 26, no. 6, pp. 1895–1901, Jun. 2003.
[6] R. Pop-Busui, A. J. M. Boulton, E. L. Feldman, V. Bril, R. Freeman, R. A. Malik, J. M. Sosenko, and D. Ziegler, "Diabetic neuropathy: A position statement by the American diabetes association," *Diabetes Care*, vol. 40, no. 1, pp. 136–154, Jan. 2017.

[7] A. I. Vinik, T. Erbas, and C. M. Casellini, "Diabetic cardiac autonomic neuropathy, inflammation and cardiovascular disease," *J. Diabetes Invest.*, vol. 4, no. 1, pp. 4–18, Jan. 2013.
[8] H. F. Jelinek, D. J. Cornforth, and A. V. Kelarev, "Machine learning methods for automated detection of severe diabetic neuropathy," *J. Diabetic Complications Med.*, vol. 1, no. 2, pp. 1–7, 2016.
[9] A. I. Vinik and D. Ziegler, "Diabetic cardiovascular autonomic neuropathy," *Circulation*, vol. 115, no. 3, pp. 387–397, Jan. 2007.
[10] R. Pop-Busui, G. W. Evans, H. C. Gerstein, V. Fonseca, J. L. Fleg, B. J. Hoogwerf, S. Genuth, R. H. Grimm, M. A. Corson, R. Prineas, and t. ACCORD Study Group, "Effects of cardiac autonomic dysfunction on mortality risk in the action to control cardiovascular risk in diabetes (ACCORD) trial," *Diabetes Care*, vol. 33, no. 7, pp. 1578–1584, Jul. 2010.
[11] A. S. Balcioglu and H. Müderrisoğlu, "Diabetes and cardiac autonomic neuropathy: Clinical manifestations, cardiovascular consequences, diagnosis and treatment," *World J diabetes*, vol. 6, no. 1, pp. 80–91, 2015.
[12] D. J. Ewing, C. N. Martyn, R. J. Young, and B. F. Clarke, "The value of cardiovascular autonomic function tests: 10 years experience in diabetes," *Diabetes Care*, vol. 8, no. 5, pp. 491–498, 1985.
[13] R. Baron and D. J. Ewing, "Heart rate variability. The international federation of clinical neurophysiology," *Electroencephalogr. Clin. Neurophysiol. Suppl.*, vol. 52, pp. 283–286, Jan. 1999.
[14] K. Lin, L. Wei, Z. Huang, and Q. Zeng, "Combination of ewing test, heart rate variability, and heart rate turbulence analysis for early diagnosis of diabetic cardiac autonomic neuropathy," *Medicine*, vol. 96, no. 45, p. e8296, 2017.
[15] H. V. Huikuri, M. J. P. Raatikainen, R. Moerch-Joergensen, J. Hartikainen, V. Virtanen, J. Boland, O. Anttonen, N. Hoest, L. V. A. Boersma, E. S. Platou, M. D. Messier, and P.-E. Bloch-Thomsen, "Prediction of fatal or near-fatal cardiac arrhythmia events in patients with depressed left ventricular function after an acute myocardial infarction," *Eur. Heart J.*, vol. 30, no. 6, pp. 689–698, Aug. 2008.
[16] J. Hayano, K. Kiyono, Z. R. Struzik, Y. Yamamoto, E. Watanabe, P. K. Stein, L. L. Watkins, J. A. Blumenthal, and R. M. Carney, "Increased non-Gaussianity of heart rate variability predicts cardiac mortality after an acute myocardial infarction," *Frontiers Physiol.*, vol. 2, p. 65, Sep. 2011. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fphys.2011.00065/full
[17] M. Alkhodari, D. K. Islayem, F. A. Alskafi, and A. H. Khandoker, "Predicting hypertensive patients with higher risk of developing vascular events using heart rate variability and machine learning," *IEEE Access*, vol. 8, pp. 192727–192739, 2020.
[18] J. Hayano, N. Ueda, M. Kisohara, E. Yuda, R. M. Carney, and J. A. Blumenthal, "Survival predictors of heart rate variability after myocardial infarction with and without low left ventricular ejection fraction," *Frontiers Neurosci.*, vol. 15, p. 32, Jan. 2021.
[19] H. F. Jelinek, H. M. Imam, H. Al-Aubaidy, and A. H. Khandoker, "Association of cardiovascular risk using non-linear heart rate variability measures with the Framingham risk score in a rural population," *Frontiers Physiol.*, vol. 4, p. 186, Jul. 2013. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fphys.2013.00186/full
[20] S. L. Cichosz, J. Frystyk, O. K. Hejlesen, L. Tarnow, and J. Fleischer, "A novel algorithm for prediction and detection of hypoglycemia based on continuous glucose monitoring and heart rate variability in patients with type 1 diabetes," *J. Diabetes Sci. Technol.*, vol. 8, no. 4, pp. 731–737, Jul. 2014.
[21] U. R. Acharya, K. S. Vidya, D. N. Ghista, W. J. E. Lim, F. Molinari, and M. Sankaranarayanan, "Computer-aided diagnosis of diabetic subjects by heart rate variability signals using discrete wavelet transform method," *Knowl.-Based Syst.*, vol. 81, pp. 56–64, Jun. 2015.
[22] G. Swapna, R. Vinayakumar, and K. Soman, "Diabetes detection using deep learning algorithms," *ICT Exp.*, vol. 4, no. 4, pp. 243–246, 2018.
[23] G. Swapna, S. Kp, and R. Vinayakumar, "Automated detection of diabetes using CNN and CNN-LSTM network and heart rate signals," *Procedia Comput. Sci.*, vol. 132, pp. 1253–1262, Mar. 2018.
[24] Y. Aggarwal, J. Das, P. M. Mazumder, R. Kumar, and R. K. Sinha, "Heart rate variability features from nonlinear cardiac dynamics in identification of diabetes using artificial neural network and support vector machine," *Biocybern. Biomed. Eng.*, vol. 40, no. 3, pp. 1002–1009, Jul. 2020.
[25] L. Wang, Y. Mu, J. Zhao, X. Wang, and H. Che, "IGRNet: A deep learning model for non-invasive, real-time diagnosis of prediabetes through electrocardiograms," *Sensors*, vol. 20, no. 9, p. 2556, Apr. 2020.
[26] A. V. Kelarev, A. Stranieri, J. Yearwood, J. Abawajy, and H. F. Jelinek, "Improving classifications for cardiac autonomic neuropathy using multi-level ensemble classifiers and feature selection based on random forest," in *Proc. 10th Australas. Data Mining Conf. (AusDM)*, 2012, pp. 93–101.
[27] A. Stranieri, J. Abawajy, A. Kelarev, S. Huda, M. Chowdhury, and H. F. Jelinek, "An approach for ewing test selection to support the clinical assessment of cardiac autonomic neuropathy," *Artif. Intell. Med.*, vol. 58, no. 3, pp. 185–193, Jul. 2013.
[28] H. Jelinek, J. Abawajy, A. Kelarev, M. Chowdhury, and A. Stranieri, "Decision trees and multi-level ensemble classifiers for neurological diagnostics," *AIMS Med. Sci.*, vol. 1, no. 1, pp. 1–12, 2014.
[29] J. Abawajy, A. Kelarev, M. U. Chowdhury, and H. F. Jelinek, "Enhancing predictive accuracy of cardiac autonomic neuropathy using blood biochemistry features and iterative multitier ensembles," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 1, pp. 408–415, Oct. 2014.
[30] A. H. Khandoker, H. F. Jelinek, and M. Palaniswami, "Identifying diabetic patients with cardiac autonomic neuropathy by heart rate complexity analysis," *Biomed. Eng.*, vol. 8, no. 1, pp. 1–12, 2009.
[31] M.-Y. Chun, H.-K. Park, H.-S. Hwang, J.-I. Han, Y.-J. Chee, and J.-S. Lee, "The association between symptoms of autonomic neuropathy and the heart rate variability in diabetics," *Korean J. Family Med.*, vol. 32, no. 5, p. 292, 2011.
[32] H. Fakhrzadeh, A. Yamini-Sharif, F. Sharifi, Y. Tajalizadekhoob, M. Mirarefin, M. Mohammadzadeh, S. Sadeghian, Z. Badamchizadeh, and B. Larijani, "Cardiac autonomic neuropathy measured by heart rate variability and markers of subclinical atherosclerosis in early type 2 diabetes," *ISRN Endocrinology*, vol. 2012, pp. 1–7, Dec. 2012.
[33] E. Abreu, "Signal-dependent rank-ordered-mean (SD-ROM) filter," in *Nonlinear Image Processing*. Amsterdam, The Netherlands: Elsevier, 2001, pp. 111–133.
[34] P. S. R. Diniz, *Adaptive Filtering*, vol. 4. New York, NY, USA: Springer, 1997. [Online]. Available: https://www.springer.com/gp/book/9780792399124
[35] J. A. Behar, A. A. Rosenberg, I. Weiser-Bitoun, O. Shemla, A. Alexandrovich, E. Konyukhov, and Y. Yaniv, "PhysioZoo: A novel open access platform for heart rate variability analysis of mammalian electrocardiographic data," *Frontiers Physiol.*, vol. 9, p. 1390, Oct. 2018.
[36] M. Malik, "Heart rate variability.: Standards of measurement, physiological interpretation, and clinical use: Task force of the European society of cardiology and the North American society for pacing and electrophysiology," *Ann. Noninvasive Electrocardiol.*, vol. 1, no. 2, pp. 151–181, Apr. 1996.
[37] C. K. Peng, J. M. Hausdorff, and A. L. Goldberger, "Fractal mechanisms in neuronal control: human heartbeat and gait dynamics in health and disease," *Nonlinear Dyn., Self-Org. Biomed.*, pp. 66–96, 2000. [Online]. Available: https://www.cambridge.org/core/books/self-organized-biological-dynamics-and-nonlinear-control/fractal-mechanisms-in-neuronal-control-human-heartbeat-and-gait-dynamics-in-health-and-disease/ 9FA0558BC32B768AB94CE944421824F2
[38] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of biological signals," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 71, no. 2, Feb. 2005, Art. no. 021906.
[39] M. D. Costa, R. B. Davis, and A. L. Goldberger, "Heart rate fragmentation: A new approach to the analysis of cardiac interbeat interval dynamics," *Frontiers Physiol.*, vol. 8, p. 255, May 2017.
[40] P. E. Greenwood and M. S. Nikulin, *A Guide to Chi-Squared Testing*, vol. 280. Hoboken, NJ, USA: Wiley, 1996.
[41] X. He, D. Cai, and P. Niyogi, "Laplacian score for feature selection," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 18, 2005, pp. 507–514.
[42] U. von Luxburg, "A tutorial on spectral clustering," *Statist. Comput.*, vol. 17, no. 4, pp. 395–416, 2007.
[43] B. Christian and T. Griffiths, *Algorithms to Live by: The Computer Science of Human Decisions*. New York, NY, USA: Macmillan, 2016.
[44] J. C. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," *Adv. Large Margin Classifiers*, vol. 10, no. 3, pp. 61–74, 1999.
[45] M. Galar, A. Fernandez, E. Barrenechea, H. Bustince, and F. Herrera, "A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches," *IEEE Trans. Syst., Man, C, Appl. Rev.*, vol. 42, no. 4, pp. 463–484, Jul. 2011.
[46] M. Buda, A. Maki, and M. A. Mazurowski, "A systematic study of the class imbalance problem in convolutional neural networks," *Neural Netw.*, vol. 106, pp. 249–259, Oct. 2018.
[47] V. N. Vapnik, *Statistical Learning Theory*. Hoboken, NJ, USA: Wiley, 1998.
[48] K.-R. Müller, S. Mika, G. Rätsch, K. Tsuda, and B. Schölkopf, "An introduction to kernel-based learning algorithms," *IEEE Trans. Neural Netw.*, vol. 12, no. 2, pp. 181–201, Mar. 2001.

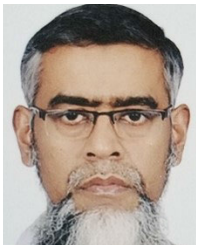
[49] M. M. Moya, M. W. Koch, and L. D. Hostetler, "One-class classifier networks for target recognition applications," *NASA STI/Recon Technical Report N*, vol. 93, p. 24043, Jan. 1993. [Online]. Available: https://www.osti.gov/biblio/6755553
[50] T. K. Ho, "Random decision forests," in *Proc. 3rd Int. Conf. Document Anal. Recognit.*, vol. 1, Aug. 1995, pp. 278–282.
[51] X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, S. Y. Philip, and Z. H. Zhou, "Top 10 algorithms in data mining," *Knowl. Inf. Syst.*, vol. 14, no. 1, pp. 1–37, 2008.
[52] J. Quinlan, *C4.5: Programs for Machine Learning*. San Francisco, CA, USA: Morgan Kaufmann, Oct. 1992, pp. 155–164. [Online]. Available: https://www.elsevier.com/books/c45/quinlan/978-0-08-050058-4 and https://www.bibsonomy.org/bibtex/1a265267f55efc59cd96ecb93a69b520
[53] C. Seiffert, T. M. Khoshgoftaar, J. Van Hulse, and A. Napolitano, "RUSBoost: A hybrid approach to alleviating class imbalance," *IEEE Trans. Syst., Man, Cybern., A, Syst., Humans*, vol. 40, no. 1, pp. 185–197, Jan. 2009.
[54] S. Mounce, K. Ellis, J. Edwards, V. Speight, N. Jakomis, and J. Boxall, "Ensemble decision tree models using RUSBoost for estimating risk of iron failure in drinking water distribution systems," *Water Resour. Manage.*, vol. 31, no. 5, pp. 1575–1589, 2017.
[55] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, Oct. 2001.
[56] J. H. Friedman, "Stochastic gradient boosting," *Comput. Statist. Data Anal.*, vol. 38, no. 4, pp. 367–378, 2002.
[57] S. A. Radzi and M. Khalil-Hani, "Character recognition of license plate number using convolutional neural network," in *Proc. 2nd Int. Vis. Inform. Conf. (IVIC)*. Selangor, Malaysia: Springer, Nov. 2011, pp. 45–55. [Online]. Available: https://link.springer.com/chapter/10.1007%2F978-3-642-25191-7_6
[58] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.
[59] M. Alkhodari and L. Fraiwan, "Convolutional and recurrent neural networks for the detection of valvular heart diseases in phonocardiogram recordings," *Comput. Methods Programs Biomed.*, vol. 200, Mar. 2021, Art. no. 105940.
[60] M. Alkhodari, L. J. Hadjileontiadis, and A. H. Khandoker, "Identification of cardiac arrhythmias from 12-lead ECG using beat-wise analysis and a combination of CNN and LSTM," in *Proc. Comput. Cardiol.*, Sep. 2020, pp. 1–4.
[61] A. H. Khandoker, H. F. Jelinek, T. Moritani, and M. Palaniswami, "Association of cardiac autonomic neuropathy with alteration of sympatho-vagal balance through heart rate variability analysis," *Med. Eng. Phys.*, vol. 32, no. 2, pp. 161–167, Mar. 2010.
[62] *Standards of Medical Care in Diabetes-2013*, vol. 36, Diabetes Care, Amer. Diabetes Assoc., Arlington County, VA, USA, Jan. 2013, pp. S11–S66.
[63] S.-A. Cha, Y.-M. Park, J.-S. Yun, S.-H. Lee, Y.-B. Ahn, S.-R. Kim, and S.-H. Ko, "Time- and frequency-domain measures of heart rate variability predict cardiovascular outcome in patients with type 2 diabetes," *Diabetes Res. Clin. Pract.*, vol. 143, pp. 159–169, Sep. 2018.
[64] B. Takase, H. Kitamura, M. Noritake, T. Nagase, A. Kurita, F. Ohsuzu, and T. Mathuoka, "Assessment of diabetic autonomic neuropathy using twenty-four-hour spectral analysis of heart rate variability. A comparison with the findings of the ewing battery: A comparison with the findings of the ewing battery," *Jpn. Heart J.*, vol. 43, no. 2, pp. 127–135, 2002.
[65] M. Alkhodari, H. F. Jelinek, N. Werghi, L. J. Hadjileontiadis, and A. H. Khandoker, "Investigating circadian heart rate variability in coronary artery disease patients with various degrees of left ventricle ejection fraction," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 714–717.
[66] A. Gottsäter, Å. R. Ahlgren, S. Taimour, and G. Sundkvist, "Decreased heart rate variability may predict the progression of carotid atherosclerosis in type 2 diabetes," *Clin. Autonomic Res.*, vol. 16, no. 3, pp. 228–234, Jun. 2006.
[67] M. Alkhodari, H. Jelinek, N. Werghi, L. Hadjileontiadis, and A. Khandoker, "Discrimination amongst various degrees of left ventricular ejection fraction in CAD patients using circadian heart rate variability features," in *Proc. 11th Conf. Eur. Study Group Cardiovascular Oscillations (ESGCO)*, Jul. 2020, pp. 1–2.
[68] A. H. Khandoker, H. M. Al-Angari, K. Khalaf, S. Lee, W. Almahmeed, H. S. Al Safar, and H. F. Jelinek, "Association of diabetes related complications with heart rate variability among a diabetic population in the UAE," *PLoS ONE*, vol. 12, no. 1, Jan. 2017, Art. no. e0168584.
[69] S.-W. Niu, J.-C. Huang, S.-C. Chen, H. Y.-H. Lin, I.-C. Kuo, P.-Y. Wu, Y.-W. Chiu, and J.-M. Chang, "Association between age and changes in heart rate variability after hemodialysis in patients with diabetes," *Frontiers Aging Neurosci.*, vol. 10, p. 43, Feb. 2018.
[70] D. Ewing, I. Campbell, and B. Clarke, "The natural history of diabetic autonomic neuropathy," *QJM, An Int. J. Med.*, vol. 49, no. 1, pp. 95–108, 1980.
[71] A. Perciaccante, A. Fiorentini, A. Paris, P. Serra, and L. Tubani, "Circadian rhythm of the autonomic nervous system in insulin resistant subjects with normoglycemia, impaired fasting glycemia, impaired glucose tolerance, type 2 diabetes mellitus," *BMC Cardiovascular Disorders*, vol. 6, no. 1, pp. 1–10, Dec. 2006.
[72] D. Aronson, L. A. Weinrauch, J. A. D'Elia, G. H. Tofler, and A. J. Burger, "Circadian patterns of heart rate variability, fibrinolytic activity, and hemostatic factors in type I diabetes mellitus with cardiac autonomic neuropathy," *Amer. J. Cardiol.*, vol. 84, no. 4, pp. 449–453, Aug. 1999.
[73] H. Kudat, V. Akkaya, A. Sozen, S. Salman, S. Demirel, M. Ozcan, D. Atilgan, M. Yilmaz, and O. Guven, "Heart rate variability in diabetes patients," *J. Int. Med. Res.*, vol. 34, no. 3, pp. 291–296, 2006.
[74] L. D. S. D. Silva, N. N. M. de Queiroz, F. T. C. de Melo de Melo, J. F. A. Neto, L. C. Janaú, N. J. K. de Souza Neto, M. N. de Lemos, M. C. N. I. de Oliveira, A. L. de Alcântara, L. V. de Moraes, and W. M. da Silva, "Improvement in cardiovascular autonomic neuropathy after high-dose vitamin d supplementation in patients with type 1 diabetes," *Frontiers Endocrinol.*, vol. 11, p. 890, Nov. 2020.
[75] M. Brennan, M. Palaniswami, and P. Kamen, "Poincaré plot interpretation using a physiological model of HRV based on a network of oscillators," *Amer. J. Physiol.-Heart Circulatory Physiol.*, vol. 283, no. 5, pp. H1873–H1886, Nov. 2002.
[76] B. Roy and S. Ghatak, "Nonlinear methods to assess changes in heart rate variability in type 2 diabetic patients," *Arquivos Brasileiros Cardiologia*, vol. 101, pp. 317–327, Oct. 2013. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/24008652/
[77] J.-L. Chen, P.-F. Chen, and H.-M. Wang, "Decreased complexity of glucose dynamics in diabetes: Evidence from multiscale entropy analysis of continuous glucose monitoring system data," *Amer. J. Physiol.-Regulatory, Integrative Comparative Physiol.*, vol. 307, no. 2, pp. R179–R183, 2014.
[78] M. D. Costa, R. B. Davis, and A. L. Goldberger, "Heart rate fragmentation: A symbolic dynamical approach," *Frontiers Physiol.*, vol. 8, p. 827, Nov. 2017.
[79] H. V. Huikuri, M. K. Linnaluoto, T. Seppänen, K. E. J. Airaksinen, K. M. Kessler, J. T. Takkunen, and R. J. Myerburg, "Circadian rhythm of heart rate variability in survivors of cardiac arrest," *Amer. J. Cardiol.*, vol. 70, no. 6, pp. 610–615, Sep. 1992.
[80] B. G. Schwartz, G. S. Mayeda, S. Burstein, C. Economides, and R. A. Kloner, "When and why do heart attacks occur? Cardiovascular triggers and their potential role," *Hospital Pract.*, vol. 38, no. 3, pp. 144–152, Jun. 2010.
[81] M. Alkhodari, H. F. Jelinek, N. Werghi, L. J. Hadjileontiadis, and A. H. Khandoker, "Estimating left ventricle ejection fraction levels using circadian heart rate variability features and support vector regression models," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 3, pp. 746–754, Mar. 2021.

MOHANAD ALKHODARI received the B.S. degree (Hons.) in electrical engineering from Abu Dhabi University (ADU), United Arab Emirates, in 2017, and the M.S. degree in biomedical engineering from American University of Sharjah (AUS), United Arab Emirates, in 2019. He is currently a Research Associate at the Healthcare Engineering Innovation Center (HEIC), Department of Biomedical Engineering, Khalifa University (KU), United Arab Emirates. Prior to joining KU, he held research appointments at the Department of Computer Science and Engineering, AUS, and the Department of Electrical and Computer Engineering, ADU. While pursuing his master's degree at AUS, he worked as a Teaching and Research Assistant with the College of Engineering, under a full scholarship granted by the university to distinguished students. He has published one book chapter and more than 20 scientific papers in international journals and conferences. His current research interests include biological signal analysis, medical imaging development, and deep/machine learning. He won the first place at UAE Undergraduate Research Competition (URC) and UAE Ministry of Health and Prevention (MOHAP) Innovations in Health Hackathon, in 2017 and 2019, respectively.

MAMUNUR RASHID received the B.Sc. degree (Hons.) in electrical and electronic engineering from United International University (UIU), Bangladesh, in 2019. He is currently working as a Research Assistant at the Biomedical Image and Signals (BIMS) Research Group, Institute of Advanced Research, UIU. Before joining BIMS Research Group, he was a Teaching Assistant with the Department of Electrical and Electronic Engineering, UIU. He was awarded a full-tuition scholarship at UIU for his academic excellence. His current research interests include biomedical signal processing, bio-chip design, and machine learning.

MOHAMMAD ABDUL MUKIT received the bachelor's degree in electrical and electronics engineering from United International University, Bangladesh. He is currently pursuing the M.S. degree in electrical and computer engineering from The University of Oklahoma, USA. His M.S. work focuses on applications of extended reality, machine learning, and deep learning in medical surgeries. His research has resulted in a startup focused on AI and data visualization in medical surgeries. He was awarded Wethington Graduate Fellowship for his academic excellence.

KHAWZA I. AHMED (Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical and electronic engineering from Bangladesh University of Engineering Technology (BUET), in 1996 and 2000, respectively, and the Ph.D. degree in electrical engineering from Arizona State University (ASU), Tempe, USA, in 2005. He served as a Staff Member at Olympus Communication Technology, USA. He was one of the Founding Directors of Bangladesh Submarine Cable Company Ltd. Currently, he is a Professor at United International University (UIU), Dhaka, Bangladesh, and is involved with the activities of the Biomedical, IMage and Signals (BIMS) Research Group. His current research interests include bio-signal processing, smart devices for e/m-health, the development of a sustainable model for e/m-health in Bangladesh, and big data in education.



RAQIBUL MOSTAFA (Senior Member, IEEE) received the undergraduate degree in EEE from BUET, in 1991, and the M.S. and Ph.D. degrees from the Electrical Engineering Department, Virginia Tech. His research focus during his graduate study was electromagnetic time domain measurement, antenna array and digital signal processing algorithms for wireless communication, prototype development for smart antenna systems, and software defined radio (SDR). He joined as a Lecturer with the EEE Department, BUET, after serving there for one and a half years, he went to the USA to complete his M.S. and Ph.D. degrees. He worked with research sponsors that include Texas Instruments (TI), U.S. Navy, DARPA of USA, and LG. After completing his Ph.D. Program, he joined Virginia Tech as a Postdoctoral Research Associate. He then joined Qualcomm Inc., San Diego, USA, and worked at the Research and Development Branch for about five years. He came back to Bangladesh to join United International University (UIU), in 2009. He is currently a Professor with the Electrical and Electronic Engineering (EEE) Department and the Dean of the School of Science and Engineering (SoSE) at United International University (UIU). He is also currently the Director of the Biomedical, IMage and Signal (BIMS) Research Group, EEE Department, UIU.

SHARMIN PARVEEN (Member, IEEE) received the B.Sc. and M.Sc. degrees (Hons.) in applied physics electronics, the M.Phil. degree in applied physics and electronics communication engineering from the University of Dhaka, and the Ph.D. degree in communication engineering, under a Sandwich Program, between the University of Dhaka, Bangladesh, and Jadavpur University, India, in 2008. She has 25 years of teaching and research experience with the universities at Bangladesh, Malaysia, India, and Indonesia. She is an Ex-Faculty Member of the Department of Computer System Technology, Faculty of Computer Science and Information Technology, University of Malaya, and the Faculty of Computer Science, University of Indonesia. She is currently an Associate Professor and the Head of the Department of Health Informatics and the Dean of Faculty of Allied Health Sciences at Bangladesh University of Health Sciences, Dhaka. She has multidisciplinary research experience in the field of high-speed computer networks, vehicular networks, green computing, medical image processing, telemedicine, mobile health, and healthcare management information systems. She has published a good number of journal articles and conference papers. She has supervised 19 master's students and two Ph.D. students. Her current research and scientific activity involves various areas of health informatics. She has been a member of IEEE COMSOC, since 2003. She has also contributed as a Key Member of the Steering Committee towards the development of the National Digital Health Strategy of Bangladesh. She has received various professional training from Italy, India, Malaysia, and Indonesia.

AHSAN H. KHANDOKER (Senior Member, IEEE) received the Ph.D. degree in the area of electronics and biomedical engineering from Muroran Institute of Technology, Japan, in 2004, followed by an Australian Research Council Fellowship at the Department of Electrical and Electronic Engineering, The University of Melbourne, Australia. He is currently an Associate Professor of biomedical engineering at Khalifa University, Abu Dhabi, United Arab Emirates. He has published over 85 journal articles and more than 130 conference papers. He has multidisciplinary research accomplishments in the area of sleep, diabetes, fetal medicine, psychiatry, biomechanics, bioinstrumentation, bio-signal processing and circuits, and nonlinear modeling. His research projects are funded by Abu Dhabi Department of Education and Knowledge, Bill and Melinda Gates Foundation, Australian Research Council, as well as Khalifa University internal funds in cardiac and mental health monitoring research area in collaboration with Cleveland Clinic Abu Dhabi and several key international medical research facilities in Australia, Germany, and Japan.