IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received September 24, 2020, accepted October 19, 2020, date of publication October 22, 2020, date of current version November 2, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3033004

Predicting Hypertensive Patients With Higher Risk of Developing Vascular Events Using Heart Rate Variability and Machine Learning

MOHANAD ALKHODARI, DEEMA K. ISLAYEM, FERYAL A. ALSKAFI, AND AHSAN H. KHANDOKER, (Senior Member, IEEE)

Healthcare Engineering Innovation Center (HEIC), Department of Biomedical Engineering, Khalifa University, Abu Dhabi, United Arab Emirates

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

This work was supported by a grant (award number: 8474000132) from the Healthcare Engineering Innovation Center (HEIC), Khalifa University, Abu Dhabi, United Arab Emirates.

ABSTRACT The prognosis of cardiovascular and cerebrovascular events for patients suffering from hypertension is considered of a high importance in preventing any further development of cardiac diseases. Despite of the ability of current gold standard techniques in predicting vascular events risks, they still lack the required clinical efficiency. In this vein, the study proposed herein provides an investigation on the feasibility of using heart rate variability (HRV) utilized through a machine learning approach to predict hypertensive patients at higher risk of developing vascular events. Initially, HRV features were extracted from all patient's data using time-domain, frequency-domain, non-linear, and fragmentation metrics. The extraction of features was based on a 24-hour cycle analysis segmented into four time periods; namely late-night, early-morning, afternoon, and evening. Analysis of all features was performed using a one-way analysis of variance (ANOVA) test on period by period basis. Furthermore, the selection of best features was performed following a Chi-squared test for demographic and HRV features. Then, a model based on decision trees and random under-sampling boosting (RUSBOOST) was trained using demographic features, HRV features, and a combination of both features. The performance of the trained model achieved a maximum accuracy of 97.08% using the combined set of features during the afternoon time period. In addition, the precision and F1-score in predicting high risk patients reached 81.25% and 86.67%, respectively. The overall area under the curve for the model was at 0.98, suggesting a high performance in the sensitivity and specificity measures. This study paves the way towards utilizing machine learning models and heart rate variability for the prognosis of vascular events in hypertensive patient. Furthermore, it assists clinicians in decision making by providing a simple, yet effective, and continuous prediction approach when compared to other available techniques.

INDEX TERMS Hypertension, cardiovascular events, cerebrovascular events, heart rate variability, ANOVA, machine learning, feature selection, RUSBOOST.

I. INTRODUCTION

Hypertension is a major risk factor for many cardiac diseases such as heart failure (HF), coronary artery disease (CAD), and stroke. It is defined as an increase in the blood pressure to above 140 mmHg systolic and 90 mmHg diastolic, causing arteries to become thicker and narrower. When combined with other arteries defects including high cholesterol deposition, the demanded blood pumping force required by the heart is increased, thus, not enough oxygen-rich blood is pumped to the rest of the body [1], [2]. There are many reasons for hypertension including smoking, less physical activity, stress, and family history [3]. According to the recent world health organization (WHO) report [4], it is estimated that hypertension in 2015 was affecting more than 1.3 billion people worldwide. In addition, 1 in 4 men and 1 in 5 women suffer from increased blood pressure.

Despite of receiving proper treatment, only 25% of patients maintain their blood pressure under control [5]. Many cardiovascular and cerebrovascular events may arise,

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

i.e., myocardial infarction (MI) and stroke, even after receiving adequate medications. Therefore, the diagnosis and prognosis of cardiac events is of a high importance to reduce the risk of premature death and disability. Prognosis of vascular events is usually done by arterial intima media thickness (IMT) assessment using carotid ultrasound [6], left ventricular mass calculations using echocardiography [7], and blood pressure measurements, which is the gold standards in hypertension analysis, using mercury sphygmomanometer [8]. Although the efficiency of these techniques is high, they require further improvements in their positive predictive value to ensure accurate levels of quality in clinical settings [9].

A promising alternative to the aforementioned techniques in vascular events risk prediction is electrocardiography (ECG) and its corresponding heart rate variability (HRV). HRV is defined as the change in the distances between R-peaks in an ECG signal with respect to time [10]. It has been a standard method to study heart functionality through statistical, geometrical, spectral, and non-linear analysis. Previous studies have shown a strong clinical relation between HRV and the control mechanism of the autonomic nervous system (ANS) [11], [12]. For example, a reduced variability in the R-R distances suggests ANS dysfunction [13]. Furthermore, low HRV levels have been found to increase death rates caused by coronary artery disease [11]. In addition to heart functionality analysis, cardiac events risk prediction and hypertension detection using HRV have been of a high interest in literature. Recently, the emergence of computerized and artificial intelligence (AI) algorithms allowed for a better prognostic techniques to assist clinical practitioners in decision making.

Several studies suggested support vector machine (SVM) models for the quantification of cardiac death risk in patients after myocardial infarction [14] as well as the detection of patients more prone to sudden death [15]. In addition, the detection of hypertensive patients aside from other cardiac diseases was performed in [9] and [16] with machine/deep learning models (highest performance of 96.67%) including k-nearest neighbor (k-NN), decision trees (DT), and convolutional neural networks (CNs). A study done by Rajput *et al.* [17] developed a hypertension diagnosis index using a wavelet filter bank to discriminate patients with high risk hypertension (HRHT) from low risk hypertension (LRHT). Furthermore, a classification tree model was designed to predict 90 days mortality in non-ST elevation acute coronary syndrome patients [18]. Despite achieving high levels of performance in vascular event prognostics caused by hypertension in these studies, there have been no 24-hour studies investigating the direct relation between HRV features and hypertension for cardiovascular risk analysis on hourly basis throughout the day/night periods. Furthermore, the analysis of the relationship between HRV features and vascular events still requires more investigations in literature.

A. OUR CONTRIBUTION

In this work, a study is conducted to investigate the ability of HRV in predicting patients at higher risk of developing cardiovascular and cerebrovascular events using a machine learning approach. It is considered of a high importance to be able of providing continuous analysis of the cardiac system to prevent the development of further diseases or even death in severe cases. Unlike the current risk prediction techniques, the novelty of the proposed work lies in designing a general computerized model that is simple, cost effective, and can be used frequently. In addition, it is able of handling data imbalance often found when analyzing medical data.

Furthermore, the current study provides deeper analysis on the 24-hour circadian cardiac functionality as reflected by the HRV. This allows to overcome limitations often found in previous studies where only short segments (5 minutes long) from the HRV were taken into consideration. Instead, HRV analysis is performed herein over the four time periods in the day/night cycle (length of 6 hours) with the knowledge that the heart exhibits different characteristics during the late-night, early-morning, afternoon and evening times [19], [20]. Therefore, certain features at a certain time period throughout the 24-hour time interval would be able of highly predicting the possible development of vascular events in future. Thus, being able to prevent it from happening by taking the required medication at the right time. In addition, a 5-minutes segment taken at random would be a very small portion of the HRV to represent any cardiac abnormalities. This has been tackled in the proposed approach by taking a wider range of HRV at known time periods throughout the 24-hour cardiac cycle.

It is worth noting that aside from conventional HRV metrics commonly used in previous studies, the proposed work analyzes HRV using a new fragmentation metric, which to the best of our knowledge has not been investigated in literature. In addition, the proposed study further elucidates on the impact of demographic information on the overall performance of the model. It suggests taking into consideration certain patient-specific information, close to those normally obtained by gold standard techniques, along with HRV features to maximize the diagnosis efficiency.

Finally, this study provides a comparison between the proposed model and several machine learning models commonly found in literature.

The main contributions of this study can be summed up into several points as follows,

- Implementing a simple, yet effective, machine learning model that handles data imbalance and performs efficiently in predicting vascular events in hypertensive patients.
- Providing deeper analysis on HRV features in a 24-hour circadian rhythm manner by covering the late-night, early-morning, afternoon, and evening time periods.
- Investigating the impact of the new HRV fragmentation features in cardiac functionality analysis besides the conventional metrics commonly used in literature.

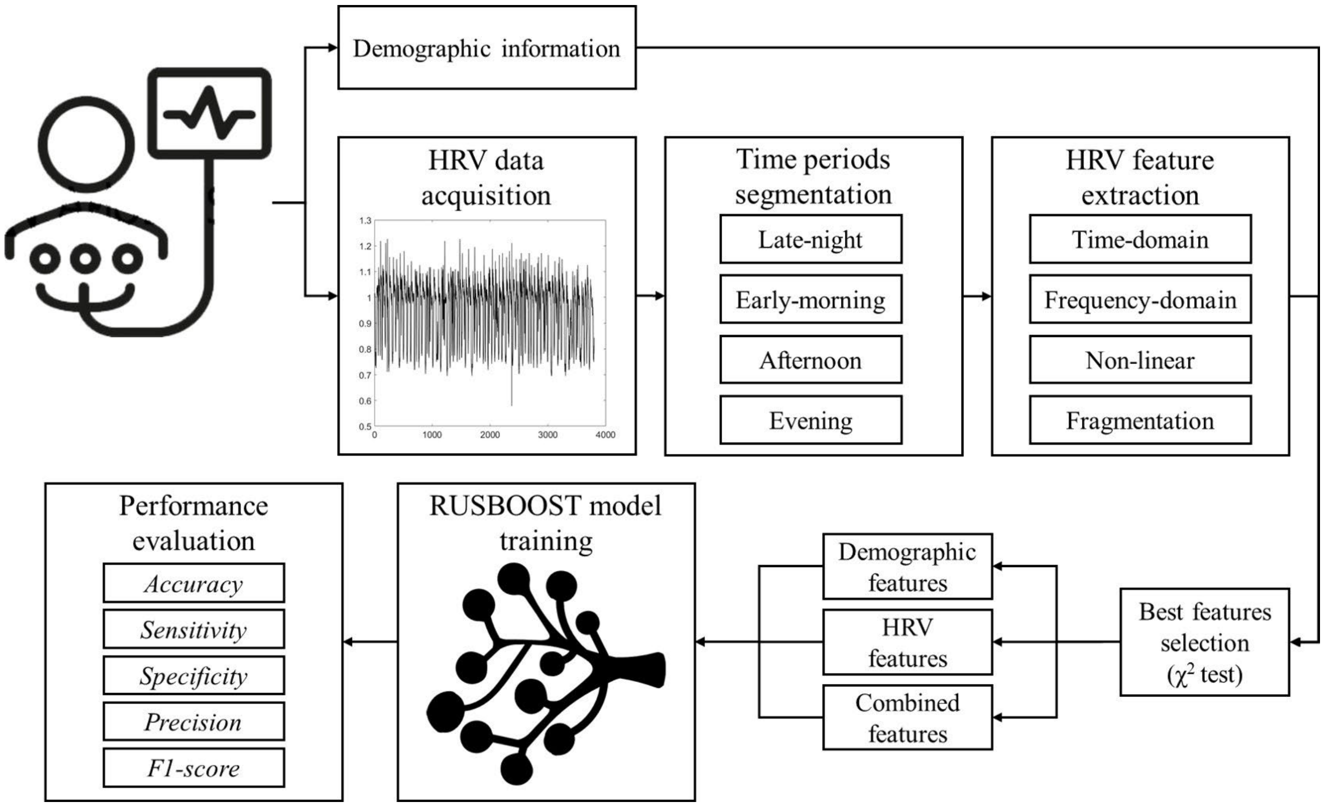


FIGURE 1. A graphical abstract of the complete procedure followed in the proposed study.

- Elucidating further on the impact of utilizing demographic information, close to those normally used in gold standard approaches, within the trained model on the overall performance.
- Providing a comparison between the proposed approach and several machine learning models commonly used in literature.

The complete procedure followed in this study is illustrated in Fig. 1.

B. PAPER ORGANIZATION

The paper is organized as follows. Section II provides a brief description of the procedure followed in this study. It starts with information about the data-set used followed by the extraction methodology of HRV features. Furthermore, the analysis of these feature on four time periods throughout the 24-hour day/night cycle is discussed along with the whole machine learning training and classification process. The section ends by a detailed information about the metrics used to evaluate the performance. Section III represents the observations from the HRV analysis as well as the results of the machine learning model. Section IV provides a detailed discussion over the observations found in this study from machine learning and clinical perspectives. Finally, the paper is concluded and summarized briefly with additional future works in Section V.

II. MATERIALS AND METHODS

A. DATA-SET

The selected data-set was obtained from the PhysioNet SHAREE database [21] that includes a total of 139 hypertensive patients' data of 24-hour ECG Holter recordings. All patients were recruited at the center of hypertension of the University Hospital of Naples Federico II, Naples, Italy. Patients included in the study were 90 males and 49 females in the age range of >55 years (average 72±7 years). The collection of ECG recordings was done after a one-month anti-hypertensive therapy wash-out. In addition, all patients were followed up for the first 12 months after getting their ECG recordings to observe any cardiovascular or cerebrovascular events such as myocardial infarction, syncope, coronary revascularization, fatal or non-fatal stroke, and transient ischemic attack. All events were adjudicated based on patient history, event/arrhythmia record, and contact with the general practitioner by the committee for event adjudication in the center. In addition, all patients had cardiac and carotid ultrasonography to determine their left ventricular mass as per the American society of echocardiography (ASE) recommendations. Among the patients, only 17 of them had events including 11 with myocardial infarctions, 3 with strokes, and 3 with syncopal events.

Each ECG recording had a 24-hour length sampled with a sampling frequency of 128 Hz. QRS complexes and

TABLE 1. **The complete demographic information (Mean ± STD) of patients covered in the database.**

Gender	Vascular Events	Number (n)	Smoker (Yes/No)	Age	BSA	BMI	SBP	DBP	IMT	LVMI	EF
Male	*Normal*	81	28/53	71 ± 7	1.95 ± 0.14	27.59 ± 3.48	135.90 ± 17.95	77.78 ± 8.90	2.38 ± 0.69	133.88 ± 23.83	57.53 ± 11.29
	MI	6	3/3	77 ± 7	2.00 ± 0.15	28.46 ± 3.40	143.17 ± 20.93	74.00 ± 10.10	2.98 ± 1.40	151.33 ± 23.37	56.00 ± 7.77
	Stroke	2	1/1	70 ± 4	1.71 ± 0.05	22.72 ± 0.44	127.50 ± 10.61	75.00 ± 7.07	2.65 ± 0.21	122.00 ± 5.66	43.50 ± 26.16
	Syncope	1	1/0	76	2.02	28.73	120.00	70.00	3.80	184.00	33.00
	Overall	**90**	**32/57**	$\mathbf{72 \pm 7}$	$\mathbf{1.95 \pm 0.14}$	$\mathbf{27.55 \pm 3.48}$	$\mathbf{136.02 \pm 17.98}$	$\mathbf{77.37 \pm 8.89}$	$\mathbf{2.44 \pm 0.76}$	$\mathbf{135.49 \pm 24.36}$	$\mathbf{56.77 \pm 11.72}$
Female	*Normal*	41	8/33	71 ± 6	1.81 ± 0.17	27.81 ± 4.71	138.40 ± 22.50	73.40 ± 8.93	2.24 ± 0.75	122.95 ± 29.21	62.82 ± 9.47
	MI	5	0/5	71 ± 8	1.81 ± 0.20	29.47 ± 6.63	151.00 ± 29.24	77.00 ± 4.47	1.55 ± 0.49	133.00 ± 23.09	64.00 ± 5.24
	Stroke	1	0/1	74	2.02	34.67	140.00	80.00	2.30	101.00	72.00
	Syncope	2	0/2	78 ± 1	1.62 ± 0.10	23.66 ± 1.90	140.00 ± 42.43	60.00	1.90 ± 0.57	141.00 ± 7.07	67.00 ± 4.24
	Overall	**49**	**8/41**	$\mathbf{71 \pm 6}$	$\mathbf{1.80 \pm 0.18}$	$\mathbf{27.95 \pm 4.92}$	$\mathbf{139.81 \pm 23.37}$	$\mathbf{73.35 \pm 8.82}$	$\mathbf{2.17 \pm 0.74}$	$\mathbf{124.35 \pm 28.00}$	$\mathbf{63.33 \pm 8.90}$

BSA: body surface area (m^2), BMI: body mass index (kg/m^2), SBP: systolic blood pressure ($mmHg$), DBP: dyastolic blood pressure ($mmHg$), IMT: intima media thickness (mm), LVMI: left ventricular mass index (g/m^2), EF: ejection fraction (%).

R-peaks were automatically annotated using length transform and onset detection algorithms [22]. It is worth noting that no manual corrections were performed on the acquired annotations. All demographic and clinical information were recorded and provided for every patient. More details on the database can be found on [9] and the complete demographic information of patients is provided in Table 1.

B. HRV FEATURES EXTRACTION

HRV is a series of heart rate values denoting the intervals between heart beats (R-R intervals). The database provides annotation files that contain the location of each R-peak. Therefore, using the location of each R-peak, the distance between peaks (HRV) is calculated and stored for every patient. It is worth noting that not all patients completed a full 24-hour Holter recording, therefore, missing hours were padded with empty values.

The annotation file comes with the knowledge about the initial starting time of each recording. Thus, all recordings were segmented into per-hour segments and re-arranged to start from hour 00:00 (12AM). Then, four periods of 6-hours, corresponding to late-night (00:00-06:00), early-morning (06:00-12:00), afternoon (12:00-18:00), and evening (18:00-00:00), were segmented from the re-arranged data. Any further HRV analysis was applied on each of these 6-hours time periods. In case of missing ECG recordings, few patients may have empty time periods with missing HRV data. These patients at these time periods were not included in the training and testing of the proposed model.

Then, HRV features were extracted from time-domain, frequency-domain, non-linear, and fragmentation analysis using PhysioNet toolbox [23] and MATLAB R2020a. Features were extracted per patient for each one of the four 6-hours time periods mentioned earlier.

In time-domain, features were extracted based on the task force of the European society of cardiology [24] and included: average of all normal-to-normal (NN) intervals (AVNN (*ms*)), standard deviation of all NN intervals (SDNN (*ms*)), square root of the mean of the sum of squares of differences between adjacent N-to-N intervals (RMSSD (*ms*)), percentage of NN intervals greater than 50 *ms* (pNN50 (%)), and standard error of the average NN intervals (SEM (*ms*))

In frequency domain, features were extracted based on the power spectral density (PSD) analysis and included: slope of the linear interpolation of the spectrum for frequencies less than very-low frequency (VLF) band upper bound (BETA), normalized high frequency (HF) power (HF Norm (%)), peak frequency in the HF band (HF Peak (*Hz*)), power in the HF band (HF Power (ms^2)), normalized low frequency (LF) power (LF Norm (%)), peak frequency in the LF band (LF Peak (*Hz*)), power in the LF band (LF Power (ms^2)), ratio of the LF power to the HF power (LF/HF), total power in both frequency bands (Total Power (ms^2)), normalized VLF power (VLF Norm (%)), and power in the VLF band (VLF Power (ms^2)).

Furthermore, non-linear HRV features were extracted from the Poincare plot, de-trended fluctuation analysis (DFA), and multi-scale entropy (MSE) [25], [26] including SD1, SD2, alpha2, alpha2, and sample entropy (SampEn). In addition. fragmentation features were extracted using according to the newly introduced methods of [27] including percentage of inflection points in the N-to-N interval (PIP (%)), acceleration/deceleration segments inverse average length (IALS), percentage of short segments (PSS (%)), and percentage of alternation segments (PAS (%)).

C. STATISTICAL ANALYSIS OF FEATURES

It is of a high importance to analyze HRV features across the four time periods using statistical measures. Each feature exhibits information about the cardiac activity of each patients' heart lying in the two groups; patients that did not develop vascular events and patients with vascular events. To statistically analyze these features, a one-way analysis of variance (ANOVA) test was performed. In this test, a comparison between the average value of multiple groups can be performed to obtain statistical evidences that these groups are of a significant difference ($p < 0.05$) [28].

A significant difference represents the strong ability for the selected feature in discriminating between the two groups, therefore, suggesting it as a significantly important feature. In addition, it suggests which period throughout the 24-hour cycle is critical to identify such differences.

D. TRAINING AND CLASSIFICATION

Utilizing machine learning for the purpose of identifying patients with higher risk of developing vascular events from patients with normal cardiac conditions was implemented in this study. Initially, a feature selecting approach based on chi-square (χ^2) test was performed to pick features with high importance prior to any training/classification procedures. These features were based on demographic information mentioned in Table 1, HRV features, and a combination of both. Then, the selected model was based on the the random under-sampling boosting (RUSBOOST) method. The complete procedure followed is briefly discussed in the following subsections.

1) FEATURE SELECTION

The selection of best features was performed based on a chi-square (χ^2) test. In this test, a statistical hypothesis test is performed to compare the observed data to expectations. In other words, it determines the significant differences between the observed and expected frequencies of samples within the data [29]. The decision on using this test besides the aforementioned ANOVA test was to analyze the features not only with respect to their average and variance measures, but also with with respect to each individual feature per-patient as a Chi-squared distribution.

In the problem presented in this study, the test ranks each demographic and HRV feature from being highly important to low based on the observed p-value after applying the test on each one. In general, the lower the p-value, the higher the importance of the selected feature. The selection of best features ($\geq$1 scoring value) set was performed for every time period using MATLAB R2020a and function *fscchi2()*.

2) MACHINE LEARNING MODEL

Decision trees are a set of tree-like attribute nodes connected to sub-trees of decision nodes. Each decision node is labeled with a class referring to the predicted class. The predictions are based on the model decisions and the corresponding consequences including the resource cost, outcomes chances, and utility. To give a prediction, the process starts by giving an instance to the root node of the tree. Then, the outcomes for this instance are measured for the following sub-nodes. Whenever a leaf is encountered, the process ends and the label is given as the prediction of this instance. More information about this technique is provided in depth in [30]. One form of decision trees are random forest (RF) or CART models [31], [32]. In CART, when the variables are discrete values, the model is training based on a classification bag of trees. On the other hand, a regression model is trained if the variables are continues.

Despite achieving high levels of performance in decision trees models, they may lack the required accuracy when classifying un-balanced data-sets. Therefore, data sampling and boosting algorithms based on the original decision trees model can handle skewness in the data. The combination of both approaches forms a hybrid ensemble algorithm named as random under-sampling boosting (RUSBOOST) [33], [34]. In this algorithm, the data is balanced by randomly removing samples from the majority class. In addition, the boosting allows of iteratively build weak learners using different linear combinations.

RUSBOOST algorithm functions in three major steps; namely setting initial weights, iterating over the weak learners and removing majority class samples, and returning the final weights of the weak learners. In step 1, the weights of each class, D_t, are defined as $1/m$, where t is the iteration number and m is the total number of samples per-class. This steps allows the algorithm to identify which learners are weaker than the others as well as identifying the class with the majority number of samples. Step 2 starts by applying RUS to decrease the number of the majority class samples. For example, if the desired ratio between two classes is 50:50, then samples are randomly removed until the minor classes are equal to the majority classes in the temporary training data-set. S_t'. This steps will change the initial weights distribution (D_t') identified in step 1. Furthermore, S_t' and D_t' are then utilized within the base learner to create the weak hypothesis, h_t. Using this learned hypothesis, the weights (D_t) are updated as follows,

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

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

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 several iterations, the final hypothesis ($H(x)$) is calculated as follows,

$$H(x) = \underset{y\in Y}{\operatorname{argmax}} \sum_{t=1}^{T} h_t(x,y) \log \frac{1}{\alpha} \tag{3}$$

where Y is the set of class labels, and T is the maximum number of iterations. More details on this algorithm are briefly provided in [33].

In this study, a RUSBOOST algorithm was implemented as the training model to handle data imbalance. The total number of splits (weak learners) was 139 corresponding to the maximum number of possible labels. The minimum leaf size and the minimum parent size were set to 1 and 2, respectively. In addition, the learning rate (shrinkage) was set to 0.1. This rate was selected based on [35], [36], where a rate $\leq$0.1 is usually suggested. In literature, it was shown that the lower the learning rate, the better the convergence and performance of boosting models. Thus, after several fine tuning tests, 0.1 was selected as the optimum rate. Furthermore, the weights were adjusted based on the amount of each class relative to the smallest class number. It is worth noting that the training/classification process was performed following three scenarios; using demographic features alone, HRV features alone, and combination of both features. The training was selected to follow a leave-one-out scheme due to the low number of samples in the data-set. In addition, it was selected to ensure the maximum possible sample within the trained model and to provide a prediction for every single patient.

Furthermore, the performance of several classifiers including RF, SVM and 1D convolutional neural networks (CNNs) was also tested, evaluated, and compared with the proposed RUSBOOST model. The settings of the RF algorithm included the use of 300 bagged trees with 5 randomly chosen feature at each split. In SVM, the model followed a radial basis function (RBF) kernel function with a 1.4 kernel scaling (gamma). For the 1D CNN network, the architecture included 1 single convolutional layer of kernel size [1,1] and a total number of filter of 16. The layer was followed by additional batch normalization (BN) and rectified linear unit (ReLU) layers. The optimizer was based on the adaptive moment estimation (ADAM) solver. The decision of picking optimum parameters for the models was based on previous findings from literature [9], as well as several tuning tests to maximize the performance.

3) PERFORMANCE EVALUATION

The performance of the algorithm in classifying patients was observed based on the confusion matrix of the predictions. Several evaluation metrics were measured including accuracy, sensitivity, specificity, precision, and F1-score. Each metric is given as follows,

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{5}$$

$$Specificity = \frac{TN}{TN + FP} \tag{6}$$

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

$$F1 - score = \frac{2TP}{2TP + FP + FN} \tag{8}$$

where TP, TN, FP, and FN corresponds to the true positives, true negative, false positives, and false negatives, respectively. In addition, the receiver operating characteristic (ROC) was obtained along with the area under the curve (AUC).

III. RESULTS

A. STATISTICALLY SIGNIFICANT FEATURES

The results of the one-way ANOVA test are shown in Table 2. Each feature was selected for the comparison at each time period across the two groups. Significant differences were obtained for LF Peak at the afternoon (p = 0.041), LF Power at the evening (p = 0.015), Total Power at the evening (p = 0.031), and VLF Power at the evening (p = 0.026). This test did not identify any time-domain, non-linear, or fragmentation features as significant. However, few features were close to significant results such as SEM, HF Power, and LF Norm with p-values of 0.073, 0.063, and 0.07, respectively.

TABLE 2. Statistical one-way ANOVA results showing HRV features p-value for the four time periods between patients with no vascular events and patients who developed vascular events.

	Time Periods			
	Late-night (00:00-06:00)	*Early-morning (06:00-12:00)*	*Afternoon (12:00-18:00)*	*Evening (18:00-00:00)*
AVNN	0.647	0.219	0.645	0.874
SDNN	0.965	0.296	0.582	0.171
RMSSD	0.839	0.926	0.965	0.433
pNN50	0.719	0.591	0.741	0.937
SEM	0.891	0.229	0.769	0.073
BETA	0.956	0.133	0.366	0.922
HF Norm	0.381	0.550	0.417	0.779
HF Peak	0.789	0.750	0.377	0.368
HF Power	0.467	0.628	0.926	0.063
LF Norm	0.895	0.070	0.418	0.163
LF Peak	0.894	0.500	**0.041***	0.875
LF Power	0.487	0.229	0.567	**0.015***
LF/HF	0.348	0.951	0.532	0.780
Total Power	0.478	0.417	0.745	**0.031***
VLF Norm	0.352	0.230	0.193	0.334
VLF Power	0.691	0.411	0.688	**0.026***
SD1	0.839	0.926	0.965	0.433
SD2	0.780	0.173	0.452	0.123
Alpha1	0.991	0.312	0.870	0.623
Alpha2	0.408	0.647	0.555	0.759
SampEn	0.495	0.460	0.920	0.756
PIP	0.647	0.895	0.194	0.452
IALS	0.647	0.895	0.194	0.452
PSS	0.528	0.952	0.271	0.382
PAS	0.837	0.696	0.250	0.735

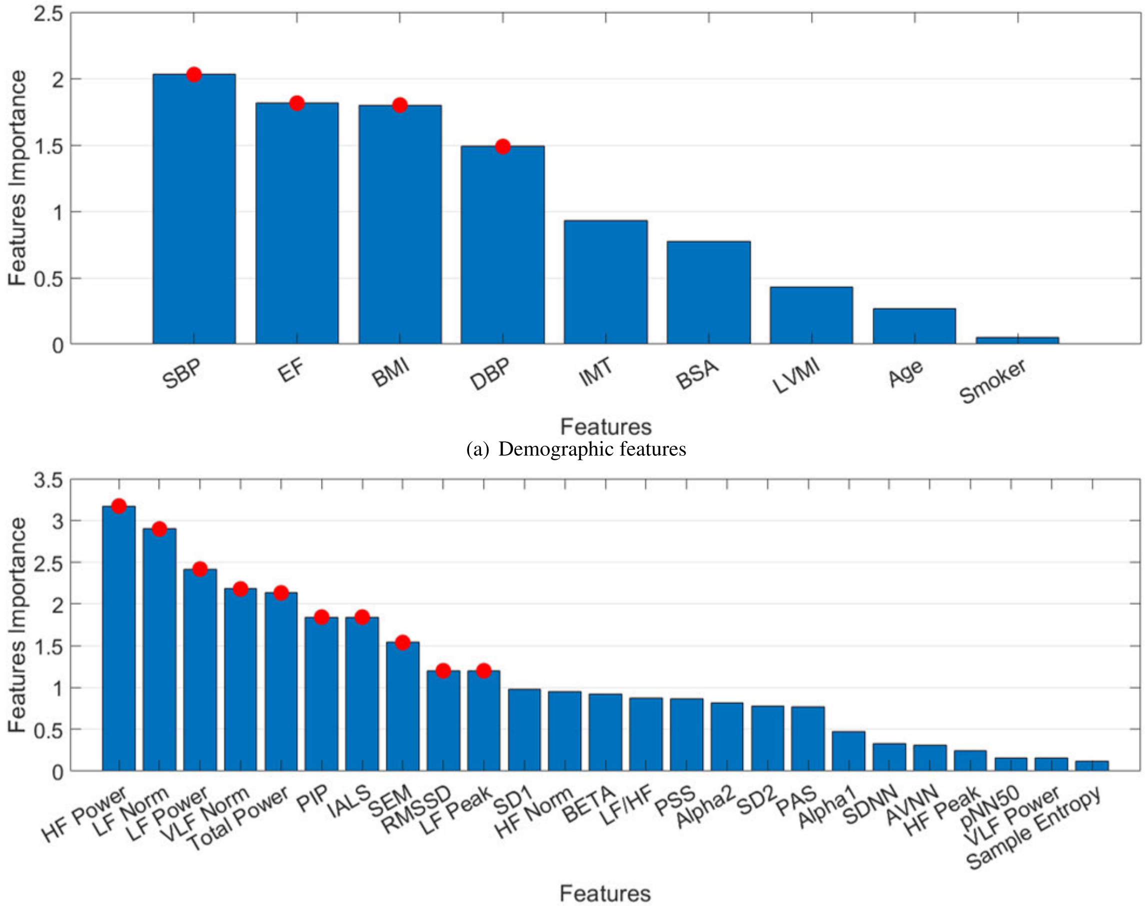
* Significant difference ($p < 0.05$).

B. TRAINING/CLASSIFICATION PERFORMANCE

Initially, the Chi-squared feature selection approach (Fig. 2) suggested the following features as significant features prior to the training of the RUSBOOST model:

- **Demographic features:** BMI, SPB, DPB, and EF.
- **HRV features:** RMSSD, SEM, HF Power, LF Norm, LF Peak, LF Power, Total Power, VLF Norm, PIP, and IALS.

For demographic features, SBP had the highest importance score (>2). In addition, EF, BMI, and DBP values were significant with scores higher than 1. The least important features were found to be the LVMI, age, and smoking with scores less than 0.5. On the other hand, HRV features (Fig. 2 (b)) are shown for the time period that gave the highest levels of performance (afternoon period). The most important features were based on the frequency-domain analysis of the HRV with scores higher than 2 except for the LF Peak (>1). Furthermore, important features from the fragmentation metrics (PIP and IALS) were obtained after the frequency-based



(a) Demographic features

(b) HRV features

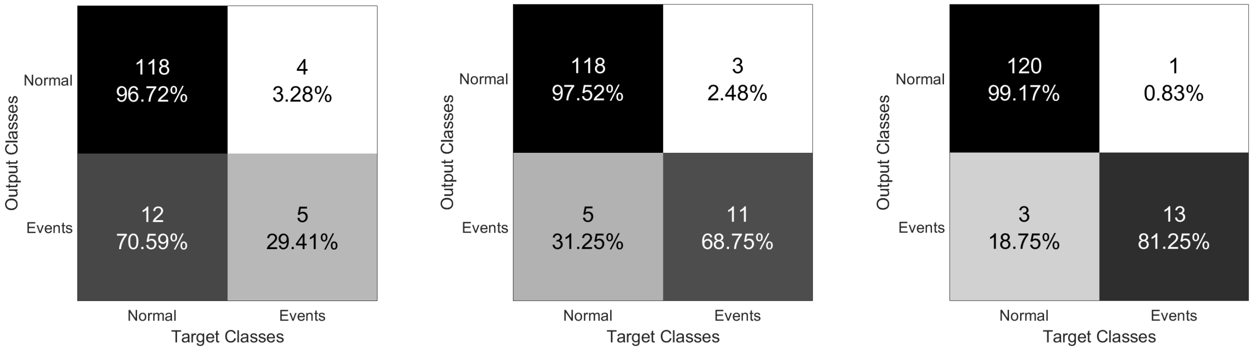
FIGURE 2. Features importance as observed from the Chi-square (χ^2) test at the best performing time period (afternoon) for: (a) demographic features, (b) HRV features. The red circles denote the significant features selected for model training (≥ 1 scoring value).

features. Two additional time-domain features (SEM and RMSSD) were considered important with importance scores higher than 1. No non-linear features were found significant during this time period.

When compared to ANOVA significant features, LF Peak feature was observed in both tests. However, Chi-squared test identified several additional features that were able of maximizing the performance even further. Two of these additional features (LF Power and Total Power) were already identified as significant features in the evening time period using the ANOVA test. In addition, the VLF Power was replaced by the normalized VLF power (VLF Norm) feature. It is worth noting that ANOVA test identified features only from the frequency domain metrics. On the other hand, Chi-squared suggested RMSSD and SEM from the time-domain, HF Power, LF Norm, LF Peak, LF Power, Total Power and VLF Norm from the frequency domain, and PIP and IALS from the fragmentation metrics. Both tests did not identify non-linear features as significant features during this time period.

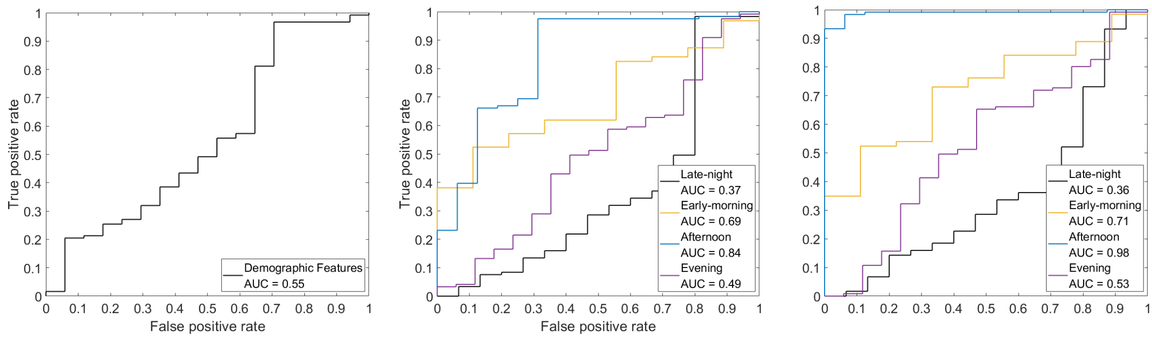
The confusion matrices of the prediction process are shown in Fig. 3 for (a) using demographic features, (b) HRV features, and (c) combined demographic and HRV features. The HRV confusion matrices were observed during the highest performing time period (afternoon). Using demographic features, the accuracy reached 88.49% on average. However, the precision of predicting vascular events was the lowest with 29.41%. Furthermore, HRV features increased the identification performance of vascular events. The precision reached 68.75%, with 11 correctly predicted patients (out of 16). The combination of both features gave the maximum performance with an average accuracy of 97.08%. The total number of correctly predicted high-risk patients was 13 out of 16 patients (precision of 81.25%). As previously discussed in Section II-B, the total number of samples during the afternoon was 137 patients out of 139 due to having 2 patients with empty ECG recordings at this time period, thus, having missing HRV values.

Tables 3, 4, and 5 show the performance metrics for the classification process using demographic, HRV, and combined features, respectively. Using demographic features, the classification sensitivity and specificity were at 90.77%



(a) Demographic features (b) HRV features (c) Combined features

FIGURE 3. confusion matrix and the corresponding true positive (TP) percentage of each class using: (a) demographic features, (b) HRV features (Afternoon), (c) combined features (Afternoon).



(a) Demographic features (b) HRV features (c) Combined features

FIGURE 4. Receiver operating characteristic (ROC) and area under curve (AUC) for the four time periods using: (a) demographic features, (b) HRV features, (c) combined features.

TABLE 3. The performance of the RUSBOOST algorithm in classifying normal and high-risk vascular events patients using demographic features.

Patients	Metrics	Percentage
Normal	*Accuracy*	88.49%
	Sensitivity	90.77%
	Specificity	55.56%
	Precision	96.72%
	F1-score	93.65%
Vascular Events	*Accuracy*	88.49%
	Sensitivity	55.56%
	Specificity	90.77%
	Precision	29.41%
	F1-score	38.46%

and 55.56%. The precision and F1-score were higher on the normal patients predictions with 96.72% and 93.65%, respectively. On the other hand, the prediction performance of vascular events had a precision of 29.41% and F1-score of 38.46%. Furthermore, HRV features improved the overall performance in predicting vascular events during the afternoon (12:00-16:00) time period with sensitivity, specificity, precision, and F1-score of 78.57%, 95.93%, 68.75%, and 73.33% respectively. Combining both features had the best precision and F1-score of 81.25% and 86.67%, as well as the best sensitivity and specificity of 92.86% and 97.56%. Table 7 shows a summary for all three methods in performance at the best performing time period (afternoon).

Furthermore, the ROC curves as well as the AUC values following each scenario are shown in Fig. 4. Using demographic features, the AUC was 0.55. However, this was improved by using HRV features during early-morning, afternoon, and evening time periods with 0.69, 0.84, and 0.49, respectively. The combination of both features maximized the performance with a 0.98 AUC during the afternoon time period.

To compare the performance of the proposed model (RUSBOOST) during the afternoon as the best performing time period, Table 7 shows the AUC, accuracy, sensitivity, specificity, precision, and F1-score values obtained by other machine learning classifiers. Using SVM, the model had an overall accuracy of 86.13% with an AUC of 0.80. The model was high in sensitivity for normal patients (97.22%) and low for vascular events patients (44.84%). However, the use of RF enhanced the performance of predicting vascular events slightly with an AUC, accuracy, sensitivity, and

TABLE 4. The performance of the RUSBOOST algorithm in classifying normal and high-risk vascular events patients throughout the four time periods using HRV features.

		Periods			
		Late-night (00:00-06:00)	*Early-morning (06:00-12:00)*	***Afternoon (12:00-16:00)***	*Evening (16:00-00:00)*
Patients	**Metrics**	**Percentage**			
Normal	*Accuracy*	88.06%	69.44%	**94.16**%	86.96%
	Sensitivity	90.55%	90.2%	**95.93**%	88.72%
	Specificity	42.86%	19.05%	**78.57**%	40.00%
	Precision	96.64%	73.02%	**97.52**%	97.52%
	F1-score	93.5%	80.7%	**96.72**%	92.91%
Vascular Events	*Accuracy*	88.06%	69.44%	**94.16**%	86.96%
	Sensitivity	42.86%	19.05%	**78.57**%	40.00%
	Specificity	90.55%	90.2%	**95.93**%	88.72%
	Precision	20.00%	44.44%	**68.75**%	11.76%
	F1-score	27.27%	26.67%	**73.33**%	18.18%

TABLE 5. The performance of the RUSBOOST algorithm in classifying normal and high-risk vascular events patients throughout the four time periods using combined features.

		Periods			
		Late-night (00:00-06:00)	*Early-morning (06:00-12:00)*	***Afternoon (12:00-16:00)***	*Evening (16:00-00:00)*
Patients	**Metrics**	**Percentage**			
Normal	*Accuracy*	84.33%	73.61%	**97.08**%	88.41%
	Sensitivity	89.52%	92.31%	**97.56**%	88.89%
	Specificity	20.00%	25.00%	**92.86**%	66.67%
	Precision	93.28%	76.19%	**99.17**%	99.17%
	F1-score	91.36%	83.48%	**98.36**%	93.75%
Vascular Events	*Accuracy*	84.33%	73.61%	**97.08**%	88.41%
	Sensitivity	20.00%	25.00%	**92.86**%	66.67%
	Specificity	89.52%	92.31%	**97.56**%	88.89%
	Precision	13.33%	55.56%	**81.25**%	11.76%
	F1-score	16.00%	34.48%	**86.67**%	20.00%

TABLE 6. Performance summary at the most important time period (afternoon) using demographic, HRV, and combined sets of features.

Patients	**Features used**	Accuracy	Sensitivity	Specificity	Precision	F1-score
Normal	*Demographic*	88.49%	90.77%	55.56%	96.72%	93.65%
	HRV	94.16%	95.93%	78.57%	97.52%	96.72%
	Combined	**97.08**%	**97.56**%	**92.86**%	**99.17**%	**98.36**%
Vascular Events	*Demographic*	88.49%	55.56%	90.77%	29.41%	38.46%
	HRV	94.16%	78.57%	95.93%	68.75%	73.33%
	Combined	**97.08**%	**92.86**%	**97.56**%	**81.25**%	**86.67**%

specificity of 0.81, 93.43%, 73.33%, and 95.90%. The 1D CNN had an overall AUC of 0.88 with an accuracy of 94.89%. In addition, the prediction of vascular events had a sensitivity of 76.47%. a specificity of 97.50%, precision of 81.25%, and F1-score of 78.79%. Despite being high in performance, the RUSBOOST algorithm performed the highest with an AUC of 0.98 and performance metrics high than other classifiers.

TABLE 7. Performance comparison between the proposed RUSBOOST model and various models such as SVM, RF, and 1D CNN at the most important time period (afternoon).

Patients	Classifier	Performance Metrics					
		AUC	*Accuracy*	*Sensitivty*	*Specificity*	*Precision*	*F1-score*
Normal	SVM	0.80	86.13%	97.22%	44.83%	86.78%	91.70%
	RF	0.81	93.43%	95.90%	73.33%	96.69%	96.30%
	1D CNN	0.88	94.89%	97.50%	76.47%	96.69%	97.10%
	RUSBOOST	**0.98**	**97.08**%	**97.56**%	**92.86**%	**99.17**%	**98.36**%
Vascular Events	SVM	0.80	86.13%	44.83%	97.22%	81.25%	57.78%
	RF	0.81	93.43%	73.33%	95.90%	68.75%	70.97%
	1D CNN	0.88	94.89%	76.47%	97.50%	81.25%	78.79%
	RUSBOOST	**0.98**	**97.08**%	**92.86**%	**97.56**%	**81.25**%	**86.67**%

IV. DISCUSSION

This study investigated the ability of HRV features in discriminating between patients that are with high or low risk of developing vascular events. With the utilization of a machine learning approach, the proposed model reached higher levels of performance (97.08%) when predicting each patients expected cardiac condition.

A. ANOVA TEST OBSERVATIONS

The statistical analysis have shown significant differences during the afternoon and evening time periods, which could be correlated to the high activation of the autonomic nervous system unlike the late-night or early morning time periods. It has been also shown by Melillo *et al.* [9] that LF and VLF features are highly important in risk prediction analysis using a random forest algorithm, which is similar to the findings of this work. The ANOVA test have identified LF Peak as significant feature during the afternoon, where this feature is usually considered as an accurate reflection of the activity of the sympathetic nervous system [37], [38]. Furthermore, the afternoon time-period is known in literature for a slight increase in the overall blood pressure [39]. On the other hand, during the evening, features such as LF Power and VLF Power were identified significant. These features play a major role in the cardiac autonomic outflows controlled by baroreflexes [40]. In addition, Total Power HRV feature increases when there is a sympathetic activation and decreases during vagal activation [41]. The observations obtained during the evening time period is matching the literature findings found in [42], where the evening is usually considered as a suitable time period for observing information about the risk of vascular events. Non-linear and fragmentation features had no significant effects on the analysis throughout the four time periods.

It is worth noting that the ANOVA test is usually performed on the average value of the feature across the groups in comparison. In other words, it represents the feature by its average values across the whole samples in the group. Therefore, a deeper look into patient-by-patient HRV features was followed.

B. MACHINE LEARNING PERFORMANCE

The model was trained based on a leave-one-out scheme to allow for the maximum number of samples to be included within the training model as well as to provide a prediction for each patient separately. The prediction of high risk patients was performed initially based on demographic features of patients alone. The feature selection approach resulted in showing the blood pressure measurements as significant features along with the BMI and EF, which matches the findings of studies performed on a wider range of population [43]–[45]. However, the risk prediction of vascular events was not of a high efficiency despite the high performance in predicting normal patients. Only 5 patients were correctly predicted to be in higher risk out of the overall 17 patients, which could be due to having close to normal demographic information in some patients and severely different information on others. Therefore, an extended set of features was required to observe the variations between such patients that did not arise using demographic features alone.

The use of HRV features allowed for an improved performance in identifying patients at high risk of developing vascular events. LF Norm HRV feature was identified as significant feature in both the ANOVA and chi squared tests during the afternoon. However, Chi-squared test allowed for the inclusion of more features within the model to maximize its performance. The training and classification process was performed using both features sets; namely ANOVA and Chi-square features, however, Chi-squared features helped in achieving higher levels in the overall performance which could be due the inclusion of more features that better represent the cardiac functionality. It is interesting to see two fragmentation features (PIP and IALS) to be of high importance in discriminating the two groups, while no features from the non-linear metrics were important. Fragmentation metrics were suggested in this study and have not been investigated previously in literature for hypertension risk prognosis.

A total of 11 patients were correctly identified as at high risk with a precision of 68.75% during the afternoon time period. There was no significant performance in the classification process in any other time period, which suggests the afternoon to be a critical hour to notice disturbances in the cardiac condition. Combining both feature sets from the demographic information and HRV allowed for the maximum possible performance in the algorithm. Only 3 patients were not classified correctly as high risk patients leading to an overall AUC of 0.98. This high performance was achieved for the afternoon time period. On the other hand, normal patients were also identified even better with only a single patient falling into the at-risk category. The ROC curves clearly show the high sensitivity and specificity of the trained model during the afternoon (97.56% and 82.86%).

The usage of other classifiers provided acceptable performance metrics, however, it was outperformed by the proposed RUSBOOST model. This could be due to the internal ability of the RUSBOOST model in balancing the data-set. The huge unbalance between the samples requires prior data im-balance handling techniques. Instead, the RUSBOOST model does not require any prior interference with the data. It is interesting to note that the 1D CNN performed better than other conventional classifiers (SVM and RF). Further investigations on the building of an advanced network may enhance its performance even further.

C. CLINICAL RELEVANCE

The investigations were performed over the 24-hour cycle of the cardiac system as seen by the ECG recordings. It is well known that HRV is a result of heart rate changes caused by the ANS fluctuations (sympathetic and parasympathetic outflow). Therefore, hypertensive patients that are more prone to develop vascular events are usually suffering from a less adaptive ANS [9]. Thus, having minor changes in the hemodynamics of the ANS can be reflected by HRV features such as the LF and HF frequency band features [6], [46]. Any decrease in the HRV through the LF and HF Power features are usually a reflection of an increased risk for cardiovascular morbidity and mortality. On the other hand, patients with high levels of HRV are at a lower risk of having cardiac abnormalities following a hypertensive event [47].

In the study presented herein, most HRV features were from the frequency-domain. In addition, the afternoon hour was more significant in showing such differences between patients, as usually the heart functionality is at its maximum during the middle of the day. A decrease in HRV as seen from the feature would have been observed at this time period which allowed the trained model to discriminate between them with higher levels of performance. It is worth noting that the fragmentation features (PIP and IALS) were useful in this study. This could be correlated to the fact that has been found in [48] that suggests that any decrease in the fluctuations of these features are hardly to be detected in a recorded ECG data. In addition, a paradoxical increase in these fluctuations comes as a results of a reduced vagal tone that would suggest the occurrence of heart diseases [49].

In terms of demographic information, it was expected to have the systolic and diastolic blood pressure measurements in the most significant features set. The current gold standard method to diagnose and predict hypertension diseases development are the mercury sphygmomanometer that provides these measurements. Furthermore, the BMI was found in literature to be strongly associated with increasing the risk of hypertension [50]. Finally, the ejection fraction ratios was also included as it represent the pumping efficiency of the heart. Any diastolic dysfunctionality could lead to hypertension, which may develop even further for a heart failure or death [51].

V. CONCLUSION

This study suggests HRV as a strong indicator of hypertensive patients that more prone to develop cardiovascular and cerebrovascular events. Unlike the gold standard techniques used for the same purpose, the proposed machine learning model is simple, efficient, cost-effective, and can be used for continuous cardiac analysis. Future works include using a bigger data-set as well as utilizing deep learning to compare the performance relative to the proposed model.

REFERENCES

[1] N. Tabassum and F. Ahmad, "Role of natural herbs in the treatment of hypertension," *Pharmacognosy Rev.*, vol. 5, no. 9, p. 30, 2011.

[2] D. Drozdz and K. Kawecka-Jaszcz, "Cardiovascular changes during chronic hypertensive states," *Pediatric Nephrol.*, vol. 29, no. 9, pp. 1507–1516, Sep. 2014.

[3] *Global Brief on Hypertension. Silent Killer, Global Public Health Crisis* World Health Org., Geneva, Switzerland, 2013.

[4] WHO. *Hypertension*. Accessed: Sep. 13, 2019. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/hypertension

[5] P. M. Kearney, M. Whelton, K. Reynolds, P. Muntner, P. K. Whelton, and J. He, "Global burden of hypertension: Analysis of worldwide data," *Lancet*, vol. 365, no. 9455, pp. 217–223, Jan. 2005.

[6] M. W. Lorenz, M. Sitzer, H. S. Markus, M. L. Bots, and M. Rosvall, "Prediction of clinical cardiovascular events with carotid intima-media thickness: A systematic review and meta-analysis-response," *Circulation*, vol. 116, no. 9, p. 318, 2007.

[7] G. Schillaci, P. Verdecchia, C. Porcellati, O. Cuccurullo, C. Cosco, and F. Perticone, "Continuous relation between left ventricular mass and cardiovascular risk in essential hypertension," *Hypertension*, vol. 35, no. 2, pp. 580–586, Feb. 2000.

[8] T. G. Pickering, G. D. James, C. Boddie, G. A. Harshfield, S. Blank, and J. H. Laragh, "How common is white coat hypertension?" *Jama*, vol. 259, no. 2, pp. 225–228, 1988.

[9] P. Melillo, R. Izzo, A. Orrico, P. Scala, M. Attanasio, M. Mirra, N. De Luca, and L. Pecchia, "Automatic prediction of cardiovascular and cerebrovascular events using heart rate variability analysis," *PLoS ONE*, vol. 10, no. 3, Mar. 2015, Art. no. e0118504.

[10] A. J. Camm, M. Malik, J. T. Bigger, G. Breithardt, S. Cerutti, R. J. Cohen, P. Coumel, E. L. Fallen, H. L. Kennedy, R. E. Kleiger, and F. Lombardi, "Heart rate variability. Standards of measurement, physiological interpretation, and clinical use," *Eur. Heart J.*, vol. 17, no. 3, 1996.

[11] J. M. Dekker, R. S. Crow, A. R. Folsom, P. J. Hannan, D. Liao, C. A. Swenne, and E. G. Schouten, "Low heart rate variability in a 2-minute rhythm strip predicts risk of coronary heart disease and mortality from several causes: The ARIC study," *Circulation*, vol. 102, no. 11, pp. 1239–1244, Sep. 2000.

[12] Z. Binici, M. R. Mouridsen, L. Køber, and A. Sajadieh, "Decreased nighttime heart rate variability is associated with increased stroke risk," *Stroke*, vol. 42, no. 11, pp. 3196–3201, Nov. 2011.

[13] H. V. Huikuri and P. K. Stein, "Heart rate variability in risk stratification of cardiac patients," *Prog. Cardiovascular Diseases*, vol. 56, no. 2, pp. 153–159, Sep. 2013.

[14] T. Song, X. F. Qu, Y. T. Zhang, W. Cao, B. H. Han, Y. Li, J. Y. Piao, L. L. Yin, and H. Da Cheng, "Usefulness of the heart-rate variability complex for predicting cardiac mortality after acute myocardial infarction," *BMC Cardiovascular Disorders*, vol. 14, no. 1, pp. 1–8, Dec. 2014.

[15] E. Ebrahimzadeh, M. Pooyan, and A. Bijar, "A novel approach to predict sudden cardiac death (SCD) using nonlinear and time-frequency analyses from HRV signals," *PLoS ONE*, vol. 9, no. 2, Feb. 2014, Art. no. e81896.
[16] M. G. Poddar, A. C. Birajdar, and J. Kriti, "Automated classification of hypertension and coronary artery disease patients by PNN, KNN, and SVM classifiers using HRV analysis," in *Machine Learning in Bio-Signal Analysis and Diagnostic Imaging*. Amsterdam, The Netherlands: Elsevier, 2019, pp. 99–125.
[17] J. S. Rajput, M. Sharma, and U. R. Acharya, "Hypertension diagnosis index for discrimination of high-risk hypertension ECG signals using optimal orthogonal wavelet filter bank," *Int. J. Environ. Res. Public Health*, vol. 16, no. 21, p. 4068, Oct. 2019.
[18] A. Singh and J. V. Guttag, "A comparison of non-symmetric entropy-based classification trees and support vector machine for cardiovascular risk stratification," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2011, pp. 79–82.
[19] 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.
[20] 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.
[21] PhysioNet. *Smart Health for Assessing the Risk of Events via ECG Database*. Accessed: May 19, 2015. [Online]. Available: https://physionet.org/content/shareedb/1.0.0/
[22] W. Zong, G. Moody, and D. Jiang, "A robust open-source algorithm to detect onset and duration of QRS complexes," in *Proc. Comput. Cardiol.*, Sep. 2003, pp. 737–740.
[23] 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.
[24] 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.
[25] C. K. Peng, J. M. Hausdorff, and A. L. Goldberger, "Fractal mechanisms in neuronal control: human heartbeat and gait dynamics in health and disease," in *Nonlinear Dynamics, Self-Organization and Biomedicine*. Cambridge, U.K.: Cambridge Univ. Press, 2000, pp. 66–96.
[26] 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.
[27] 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.
[28] M. Hamada and J. Wu, *Experiments: Planning, Analysis, and Parameter Design Optimization*. New York, NY, USA: Wiley, 2000.
[29] P. E. Greenwood and M. S. Nikulin, *A Guide to Chi-Squared Testing*, vol. 280. Hoboken, NJ, USA: Wiley, 1996.
[30] J. Quinlan, *C4. 5: Programs for Machine Learning Morgan Kaufmann Publishers San Francisco*. Burlington, MA, USA: Morgan Kaufmann, 1993, pp. 155–164.
[31] T. K. Ho, "Random decision forests," in *Proc. 3rd Int. Conf. Document Anal. Recognit.*, vol. 1, Aug. 1995, pp. 278–282.
[32] X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yu, Z.-H. Zhou, M. Steinbach, D. J. Hand, and D. Steinberg, "Top 10 algorithms in data mining," *Knowl. Inf. Syst.*, vol. 14, no. 1, pp. 1–37, 2008.
[33] 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. 2010.
[34] S. R. Mounce, K. Ellis, J. M. Edwards, V. L. Speight, N. Jakomis, and J. B. 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, Mar. 2017.
[35] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, Oct. 2001.
[36] J. H. Friedman, "Stochastic gradient boosting," *Comput. Statist. Data Anal.*, vol. 38, no. 4, pp. 367–378, Feb. 2002.
[37] M. S. Houle and G. E. Billman, "Low-frequency component of the heart rate variability spectrum: A poor marker of sympathetic activity," *Amer. J. Physiol.-Heart Circulatory Physiol.*, vol. 276, no. 1, pp. H215–H223, Jan. 1999.
[38] K. Trimmel, J. Sacha, and H. V. Huikuri, *Heart Rate Variability: Clinical Applications and Interaction Between HRV and Heart Rate*. Lausanne, Switzerland: Frontiers Media SA, 2015.
[39] C. K. Kramer, C. B. Leitao, L. H. Canani, and J. L. Gross, "Afternoon blood pressure increase: A blood pressure pattern associated with microvascular complications in type 2 diabetes mellitus," *Amer. J. Hypertension*, vol. 24, no. 1, pp. 64–69, Jan. 2011.
[40] D. S. Goldstein, O. Bentho, M.-Y. Park, and Y. Sharabi, "Low-frequency power of heart rate variability is not a measure of cardiac sympathetic tone but may be a measure of modulation of cardiac autonomic outflows by baroreflexes," *Exp. Physiol.*, vol. 96, no. 12, pp. 1255–1261, Dec. 2011.
[41] J. CAMM, "Task force of the European society of cardiology and the North American society of pacing and electrophysiology. Heart rate variability: Standarts of measurement, physiological interpretation and clinical use," *Circulation*, vol. 93, no. 5, pp. 1043–1065, Mar. 1996.
[42] K. Asayama, T. Ohkubo, M. Kikuya, T. Obara, H. Metoki, R. Inoue, A. Hara, T. Hirose, H. Hoshi, J. Hashimoto, and K. Totsune, "Prediction of stroke by home 'morning' versus 'evening' blood pressure values: The Ohasama study," *Hypertension*, vol. 48, no. 4, pp. 737–743, 2006.
[43] J. A. A. G. Damen, L. Hooft, E. Schuit, T. P. A. Debray, G. S. Collins, I. Tzoulaki, C. M. Lassale, G. C. M. Siontis, V. Chiocchia, C. Roberts, M. M. Schlüssel, S. Gerry, J. A. Black, P. Heus, Y. T. van der Schouw, L. M. Peelen, and K. G. M. Moons, "Prediction models for cardiovascular disease risk in the general population: Systematic review," *BMJ*, vol. 353, p. i2416, May 2016.
[44] S. Teramukai, Y. Okuda, S. Miyazaki, R. Kawamori, M. Shirayama, and T. Teramoto, "Dynamic prediction model and risk assessment chart for cardiovascular disease based on on-treatment blood pressure and baseline risk factors," *Hypertension Res.*, vol. 39, no. 2, pp. 113–118, Feb. 2016.
[45] M. Rosu, S. Popa, E. Mota, A. Popa, M. Manolache, C. Guja, C. Bala, C. Mota, and M. Mota, "Cardiovascular risk assessment in the adult (aged 40-79 years) Romanian population," (Bucharest), *Acta Endocrinologica*, vol. 14, no. 2, p. 227, 2018.
[46] K. Nagai, S. Shibata, M. Akishita, N. Sudoh, T. Obara, K. Toba, and K. Kozaki, "Efficacy of combined use of three non-invasive atherosclerosis tests to predict vascular events in the elderly; carotid intima-media thickness, flow-mediated dilation of brachial artery and pulse wave velocity," *Atherosclerosis*, vol. 231, no. 2, pp. 365–370, Dec. 2013.
[47] S. Hillebrand, K. B. Gast, R. de Mutsert, C. A. Swenne, J. W. Jukema, S. Middeldorp, F. R. Rosendaal, and O. M. Dekkers, "Heart rate variability and first cardiovascular event in populations without known cardiovascular disease: Meta-analysis and dose–response meta-regression," *EP Europace*, vol. 15, no. 5, pp. 742–749, May 2013.
[48] M. D. Costa, S. Redline, R. B. Davis, S. R. Heckbert, E. Z. Soliman, and A. L. Goldberger, "Heart rate fragmentation as a novel biomarker of adverse cardiovascular events: The multi-ethnic study of atherosclerosis," *Frontiers Physiol.*, vol. 9, p. 1117, Sep. 2018.
[49] J. Wdowczyk, D. Makowiec, M. Gruchała, D. Wejer, and Z. R. Struzik, "Dynamical landscape of heart rhythm in long-term heart transplant recipients: A way to discern erratic rhythms," *Frontiers Physiol.*, vol. 9, p. 274, Apr. 2018.
[50] F. Landi, R. Calvani, A. Picca, M. Tosato, A. M. Martone, E. Ortolani, A. Sisto, E. D'Angelo, E. Serafini, G. Desideri, M. T. Fuga, and E. Marzetti, "Body mass index is strongly associated with hypertension: Results from the longevity check-up 7+ study," *Nutrients*, vol. 10, no. 12, p. 1976, Dec. 2018.
[51] F. H. Messerli, S. F. Rimoldi, and S. Bangalore, "The transition from hypertension to heart failure: Contemporary update," *JACC, Heart Failure*, vol. 5, no. 8, pp. 543–551, 2017.

MOHANAD ALKHODARI received the B.S. degree (Hons.) in electrical engineering from Abu Dhabi University (ADU), United Arab Emirates (UAE), in 2017, and the M.S. degree in biomedical engineering from the American University of Sharjah (AUS), UAE, in 2019. He is currently a Research Associate with the Healthcare Engineering Innovation Center (HEIC), Department of Biomedical Engineering, Khalifa University (KU), UAE. Prior to joining KU, he held research appointments at the Department of Computer Science and Engineering of AUS and at the Department of Electrical and Computer Engineering of ADU in summer 2019. 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. So far, he has published one book chapter and more than 12 scientific articles in international journals and conferences. He is also an Active Reviewer of the *Journal of Medical Imaging and Health Informatics* (JMIHI). His current research interests include biological signal analysis, medical imaging development, and deep/machine learning. He won the first place at the UAE Undergraduate Research Competition (URC) and at the UAE Ministry of Health and Prevention (MOHAP) Innovations in Health Hackathon in 2017 and 2019, respectively.

DEEMA K. ISLAYEM received the B.Sc. degree in biomedical engineering from Khalifa University (KU), Abu Dhabi, UAE, in 2019, where she is currently pursuing the M.Sc. degree in biomedical engineering. She is also working as a Graduate Teaching and Research Assistant, where her research interests include the wet part of biomedical engineering focusing in RNA post-transcriptional modification and RNA sorting.

FERYAL A. ALSKAFI received the bachelor's degree in biomedical engineering from the Jordan University of Science and Technology (JUST), in 2019. She is currently pursuing the master's degree in biomedical engineering with Khalifa University, Abu Dhabi, United Arab Emirates. She is also working as a Teaching/Research Assistant with the College of Engineering under a full-ride academic scholarship granted by the university. Her current research interests include biomedical signal processing and machine learning.

AHSAN H. KHANDOKER (Senior Member, IEEE) received the Ph.D. degree in electronics and biomedical engineering from the Muroran Institute of Technology, Japan, in 2004, followed by an Australian Research Council Fellowship in the Department of Electrical and Electronic Engineering, University of Melbourne, Australia. He is currently an Associate Professor of biomedical engineering with Khalifa University, Abu Dhabi, UAE. 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, and bioinstrumentation; bio-signal processing & circuits and nonlinear modeling. His research projects are funded by the 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.