COMPUTER AIDED DIAGNOSIS OF VENTRICULAR ARRHYTHMIAS FROM ELECTROCARDIOGRAM LEAD II SIGNALS

Saleh A. Alzahrani[1], and Umar S. Alqasemi[2]

[1]Biomedical Engineer, National Guards Hospital, Jeddah; and Master Student at King Abdulaziz University, Jeddah, Saudi Arabia
[2]Assistant Professor, Dept. of Electrical and Computer Engineering, Biomedical Engineering Program, King Abdulaziz University, P. O. Box. 80204, Jeddah 21589, Saudi Arabia

ABSTRACT

In this work, we use computer aided diagnosis (CADx) to extract features from ECG signals and detect different types of cardiac ventricular arrhythmias including Ventricular Tachycardia (VT),Ventricular Fibrillation (VF), Ventricular Couplet (VC), and Ventricular Bigeminy (VB).Our methodology is unique in computing features of lower and higher order statistical parameters from six different data domains: time domain, Fourier domain, and four Wavelet domains (Daubechies, Coiflet, Symlet, and Meyer). These features proved to give superior classification performance, in general, regardless of the type of classifier used as compared with previous studies. However, Support Vector Machine (SVM) and Artificial Neural Network (ANN) classifiers got better performance than other classifiers tried including KNN and Naïve Bayes classifiers. Our unique features enabled classifiers to perform better in comparison with previous studies: for VT, 100% accuracy while best previous work got 95.8%, for VF, 100% accuracy while best previous work got 97.5%, for VC, 100% sensitivity while best previous work got 71.8%, and for VB, 100% sensitivity while best previous work got 84.6%.

KEYWORDS

Electrocardiogram, Computer-Aided Detection (CAD), Computer-Aided Diagnosis (CADx), Wavelet Transform, Feature Extraction, Digital Signal Processing, Artificial Neural Networks (ANN), Support Vector Machine (SVM), Machine Learning.

1. INTRODUCTION

The heart is a muscular organ and it is one of the most crucial organs in the entire human body. It is responsible for pumping the blood throughout the body. The purpose of the circulation blood is to supply oxygen and essential nutrients to the tissues of the body and eliminate carbon dioxide and waste products. Thus, the development of technology for monitoring the status of the heart is of particular importance in medical science [1].

Cardiology is the medical science concerned with abnormalities and diseases of the heart. Cardiac arrhythmia is a group of irregular heartbeat or abnormal heart rhythm. There are various types of arrhythmia, some of them are harmless and others are life-threatening, and could cause death due to ventricular arrhythmias, coronary artery disease, and valve disease. Therefore there is a need for a method to study and monitor arrhythmias, this can be done using an Electrocardiogram (ECG)[2]. The establishment of ECG technology began at the start of 20th century by Dutch

DOI : 10.5121/sipij.2018.9501

physiologist Willem Einthoven in 1903 by using a string galvanometer [3]. ECG is a diagnostic tool that can be used to measures and records the electrical activity of the heartbeat. So, it is crucial to extract the minute information from the ECG signal to obtain an accurate analysis of the heart to allow physician diagnose various forms of heart disease [4].

Arrhythmias are becoming a significant reason for sudden death around the world. The aim of this research is to detect and classify ECG ventricular arrhythmias. Classical techniques have been used to address this problem such as the analysis of ECG signals for arrhythmia detection using the Fourier Coefficients, statistical features and wavelet domain features, etc.

ECG arrhythmias classes used in this thesis are ventricular arrhythmias are more serious life-threatening than aterialarrhythmias. These ECG arrhythmias data were obtained from MIT-BIH database, which consist of four abnormal of the types including Ventricular Couplet (VC), Ventricular Tachycardia (VT), Ventricular Bigeminy (VB) and Ventricular Fibrillation (VF), and one normal control class.

Physicians in healthcare facilities can diagnose different types of arrhythmias after doing 12-lead ECG analysis. However, for emergency admissions, first aid, and ambulances, it is usually not accessible for physicians to detect such abnormalities from vital sign monitors. However, CAD algorithms can be incorporated in such monitors for online detection of such cases to help identify people of such abnormalities for preventive and follow-up purposes that aid health care practitioners to provide necessary measures.

2. PREVIOUS RESEARCH

Issac et at [5], proposed a method for the classification of the heartbeat of ECG, based on the use of the Artificial Neural Network (ANN). The feature sets considered include RR intervals, Heartbeat intervals, and Spectral entropy. The ECG signals were also obtained from MIT-BIH database, which were used to classify the normal beat and nine different arrhythmias namely, Right Bundle Branch Block (RBBB), Left Bundle Branch Block (LBBB), Supraventricular ectopic Premature beat (SP), Atrial Premature beat (AP), Premature Ventricular Contraction (PVC), Atrial Fibrillation (AF), Sick Sinus Syndrome (SSS), Ventricular Fibrillation (VF), and Fusion of ventricular and normal beat (FVN). The total accuracy of classification of the proposed method is 0.990.

Asl et al. [6], proposed an algorithm for ECG arrhythmias classification by reduced features. The data set consist of four different classes (Normal sinus rhythm (NSR), arterial premature contraction and supraventricular tachycardia (APC/SVT), PVC/ Ventricular Tachycardia (VT), and VF) obtained from MIT-BIH database, too. Seventeen features were extracted by wavelet transform; two features related to rhythm and fifteen wavelet coefficient features. They used three classifiers with accuracy of 0.986 for Multilayer Perceptrons (MLP), 0.989 for Fuzzy Inference System (FIS), and 0.993 for support vector machine (SVM). The best performances have been obtained by SVM.

Hasan, Kadah[7], detected and classify ECG arrhythmias using ANN, K-nearest neighbor (KNN), Multi-class support vector machine (MC-SVM), and linear discriminant analysis (LDA) classifier. The data set consist of five classes (Normal rhythm (NR), ventricular couplet (VC), VT, ventricular bigeminy (VB), and VF) obtained from MIT-BIH database as well. Features extraction of ECG signal based on the accompanied poles using Prony’s method and complex resonance frequencies. They reported accuracies as 1 for ANN, 0.933 for KNN, 0.924 for MC-SVM, and 0.857 for LDA. ANN classifier has proved its high accuracy compared to other classifiers for the type of features used in this study.

Kim et al. [8], proposed a method of ECG feature to detect ECG arrhythmias through combination of wavelet transform. The data set consist of three types (Normal ECG, VT, and PVC) obtained from developed patch type electrode by researcher. Features set was based on Daubechies, Symlets and Coiflets wavelet transforms on ECG signals. The maximum detection accuracy achieved 0.962 using ANN classifier.

Orozco-Duque et al. [9], implement SVM and ANN for real-time detection of ventricular arrhythmia. The database sets consist of three classes NR, VT and VF obtained from same MIT-BIH arrhythmia database. The features set considered include a fast wavelet transform (FWT) and sub-bands wavelet energy. The overall accuracy of classification of the proposed method was 0.995 for both classifier ANN and SVM.

Othman et al.[10], used semantic mining (SM) based algorithm for detecting VT and VF. The database set considered three types of ECG signal normal, VT and VF obtained from same MIT-BIH arrhythmia database. They opted to use classifier called semantic mining to characterize VF and VT by using three syntax parameters (Natural frequency, damping coefficient, and input signal). They obtained high accuracy 0.967 because they used well targeted features.

Li and Rajagopalan [11], classified VF and VT by using SVM. The ECG signals were derived from (the Creighton University Ventricular Tachyarrhythmia Database (CUDB), the MIT-BIH Malignant Ventricular Arrhythmia Database (VFDB), and the American Heart Association Database (AHADB)) which include VF and VT. The features sets considered include time domain features, frequency domain features, complexity features, and statistical features extracted from specific window length of ECG signal. The SVM classifier achieves 0.982 accuracy.

Kavitha and Christopher [12], proposed a method of ECG feature to detect ECG arrhythmias using Particle Swarm Optimization (PSO), and Fuzzy C-Means (FCM) Clustering. The data set consisted of five classes (PVC, NSR, AF, VF, and 2° heart block (BII)) obtained from the most commonly used MIT-BIH database. Features extraction were based on linear analysis (time domain features, andfrequency domain features) and nonlinear analysis (largest Lyapunov exponent, poincare plot, correlation dimension, and spectral entropy). With implemented SVM, they obtained high accuracy 0.984.

Pooyan and Akhoondi[13], applied morphological features for classification of ventricular arrhythmias. The data set consist of five classes VT, ventricular flutter (VFL), VF, ventricular escape beat (VEB), and PVC obtained from MIT-BIH database. Features were extracted using morphological features include (the amplitude of R peak, QS interval, the rising and falling slopes of QRS complex, and positive and negative areas of the complex QRS. Accuracy of ventricular abnormalities and normal sinus rhythm obtained 0.959 by using SVM with Gaussian kernel.

Weixin[14], Classified ECG ventricular arrhythmias using fuzzy logic classifier. The data set consist of three types disorganized VF (DVF), VT and organized VF (OVF) obtained from the MIT-BIH Malignant Ventricular Arrhythmia Database (VFDB), and Creighton University Ventricular Tachyarrhythmia Database (CUDB). Three major feature extraction methods: frequency domain features, time domain features, and complexity features. The total accuracies obtained with two levels were: first detect VT with an accuracy of 0.926 and then the discrimination between DVF and OVF was detected with an accuracy of 0.845.

Tripathy et al. [15], introduces new method for detection and classification of ventricular arrhythmia using the least square support vector machine (LSSVM). The data set consisted of three types (NR, VT, VF) obtained from the MIT-BIH Malignant Ventricular arrhythmia Database (VFDB), and Creighton University Ventricular Tachyarrhythmia Database (CUDB).

Moreover the features extracted based on digital Taylor-Fourier transform (DTFT). They achieved an accuracy of 0.898.

Sreedevi and Anuradha [16], evaluated for detection of heart arrhythmias by using Discrete Wavelet Transform (DWT) method. The data set consist of five types (bradycardia, VT, PVC, supraventricular tachycardia, and myocardial infarction (MI) obtained from the MIT-BIH arrhythmia database. The proposed method for extracting was a daubechies Wavelet Transform. Overall accuracy of 0.971 was achieved.

Mohanty et al.[17], detected and classify VT and VF arrhythmias using cubic SVM and C4.5 classifier. The data set consist of three types (NR, VT, andVF) obtained from MIT-BIH arrhythmia database. The features sets considered include temporal, spectral, and statistical features. The experiments showed accuracy of 0.970 for C4.5 classifier which was better than cubic SVM 0.922.

Mohammadalipour et al. [18], proposed a method for discrimination ECG arrhythmias using nonlinear features, time and frequency domain. The ECG data consisted of ten types (VF, VT, AF, NSR, bigeminy (BG), trigeminy (TG), quadrigeminy (QG), couplet, triplet, and PVC) from MIT-BIH arrhythmia database. The features sets considered include Image-Based Phase Plot for Morphological Analysis, frequency domain feature, nonlinear feature, and Shannon Entropy (SE). The accuracy of binary decision tree BDT, and SVM are 0.962, and 0.929, respectively.The BDT provided slightly higher accuracy than SVM classifier.

Table 1. Summarizes the previous research

Author and date	**Main features**	**Classifiers**	**ACC**
Issac et at [5]	RR intervals Heartbeat intervals and Spectral entropy	ANN	0.990
Asl et al. [6]	17 features were extracted by wavelet transform; two features related to rhythm and 15 wavelet coefficient features	MLP	0.986
		FIS	0.989
		SVM	0.993
Hasan, Kadah[7],	-Prony's method -complex resonance frequencies	ANN	1
		MC-SVM	0.924
		LDA	0.857
		KNN	0.933
Kim, M.S., et al (2011) [8]	Daubechies, Coiflets and Symlets order 5 wavelet transform	CWTANN	0.962
Orozco-Duque et al. (2013)[9]	fast wavelet transform (FWT) and sub-bands wavelet energy	ANN SVM	0.995
Othman et al. (2013) [10]	-Natural frequency -Dynamic ECG features for atrial fibrillation recognition.	Semantic mining	0.967

Li and Rajagopalan (2014) [11]	-Time domain features, -Frequency domain features, -Complexity features and -Statistical feature	SVM	0.963
Kavitha and Christopher [12]	Nonlinear analysis and -Nonlinear analysis	SVM	0.984
Pooyan and Akhoondi[13]	Morphological features	SVM	0.959
Weixin[14]	-Frequency domain features, -Time domain features and -Complexity features	Fuzzy logic classifier	0.885
Tripathy et al. [15]	Digital Taylor-Fourier transforms (DTFT).	LS-SVM	0.898
Sreedevi and Anuradha (2017) [16]	Daubechies Wavelet Transform	ANN	0.971
Mohanty et al. (2018) [17]	Time-frequency and statistical features	C4.5	0.970
		Cubic SVM	0.922
Mohammadalipour et al (2018) [18]	-Image-Based Phase Plot for Morphological Analysis -Frequency Domain Feature -Nonlinear Feature SE	SVM (For 4 different stages)	0.929
		BDT (For 4 different stages)	0.962

The significant points of the above literature review are as follows:

- Almost all of the researchers have used MIT-BIH arrhythmia database.
- Researchers have worked on different ECG datasets that of different cardiac disorders like: VT, VF, VB, VC, PVC, AF, BG, MI and so on, but mostly on ventricular arrythmias.
- An advancement has been observed in computer-aided ECG signal analysis and diagnosis during the last decade.
- The combined fast wavelet transform (FWT) and sub-bands wavelet energy features gave the best accuracy all of studies involving (0.995) ANN classifier. It has been found that is it better than using Daubechies Wavelet Transform alone (accuracy 0.971), or morphological features likeRR intervals, Heartbeat intervals, and Spectral entropy (accuracy 0.990).

- Wavelet transform analysis boosts the accuracy of detecting of ECG arrhythmias, if used with SVM and ANN classifiers compared to nonlinear analysis, linear analysis, statistical features, and morphological features.

- Most studies showed good accuracy using ANN and SVM classifiers compared to other classifiers.

3. METHODOLOGY

All the methods used in this research will be discussed. All computations were implemented using MATLAB 2014a on a personal computer. The rest of the chapter will be divided into six subsequent categories data collection, feature extraction, T-test feature selection, and finally the classification as shown in Figure 1.

Our methodology in this study can be summarized as follows:

- For each case of the five studied arrhythmias we consider the case as abnormal and compare it to normal case as a control set.

- Features are extracted by computing higher and lower order statistics from six different domains time domain, Fourier transform, and four Wavelet transforms (Daubechies, Coiflets, Symlets, and Meyer).

- We used t-test to screen out statistically insignificant features, while maintaining only useful features.

- We used different classifiers to detect abnormal ECG signals including K-Nearest Neighbour (KNN) with different neighbour number, Artificial Neural Networks (ANN), Naive Bayes, and Support Vector Machine (SVM) with different kernels include: Polynomial, Linear, RBF, quadratic and MLP.

- We used quantitative evaluation criteria to assess the best classifier for each arrhythmia type.

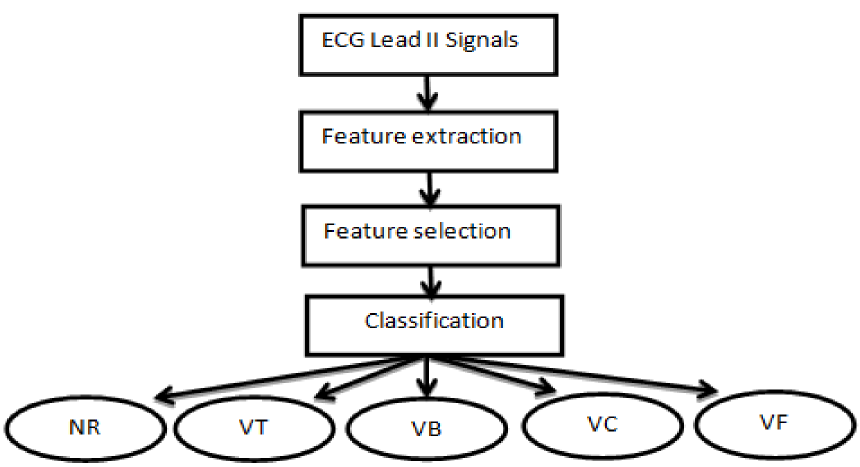


Figure 1. A diagram of our arrhythmia CADx algorithm design

4. DATABASE

ECG lead-II signals are used throughout this study acquired from the MIT-BIH arrhythmia [19]. The first available set of standard test dataset for evaluation of arrhythmia database was this MIT-BIH Arrhythmia Database. Furthermore, it has been also used for research purposes in more than five hundred sites around the world since 1980[20].This online database is formed using a set of large number of independent 3s intervals of ECG signals.

The dataset signals used in this study have set of time series sampled at 360 samples/sec for duration of 3secs. This is to comply with the ANSI/AAMI EC13-1992 standard, which requires alarms for abnormal ECG signals to be started within 10secs of their onset. The use of two different sampling was not found to be critical as long as the ECG signal is adequately sampled[21].

The data used is divided into two sets, one is used to train the classifier, and the other one is used to test its performance on non-training data. The two sets are as follow:

1. 320 ECG signal for training set: 64 for each arrhythmia class, and 64 for normal class.

2. 160 ECG signal for testing sample: 32 for each arrhythmia class, and 32 for normal control class.

5. FEATURE EXTRACTION

ECG Feature Extraction plays a significant role in diagnosing most of the cardiac diseases. For physicians, one cardiac cycle in an ECG signal consists of the P-QRS-T waves. Physicians are trained to detect any abnormality in the amplitudes and intervals in the ECG signal. The amplitudes and intervals values of P-QRS-T segment determines the functionality of the heart. For CAD, recently numerous researchers have developed different numerical techniques for analysing the features from ECG signal pattern.

In our study we computed different higher and lower order statistics from six different domains, as follows:

Mean

To measure the average of the values[22].

$$\mu = \frac{1}{m}\sum_{i=1}^{m} x_i \qquad \mathbf{1}$$

Standard deviation

To measure how the values are spread out around the mean [22].

$$\sigma = \sqrt{\frac{1}{m-1}\sum_{i=1}^{m}(x_i-\mu)^2} \qquad \mathbf{2}$$

Kurtosis

To measure the peakedness of the probability distribution of the data[22].

$$ku = \frac{\mu_4}{\rho^4} \quad \mathbf{3}$$

Where $\boldsymbol{\mu}_4$

$$\mu_4 = \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu)^4 \quad \mathbf{4}$$

Skewness

To measure the asymmetry of the data pattern[22].

$$Sk = \frac{\mu_3}{\rho^3} \quad \mathbf{5}$$

Where $\boldsymbol{\mu_3}$

$$\mu_3 = \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu)^3 \quad \mathbf{6}$$

Percentiles

Percentiles were used to measure the position. In this study, we computed percentiles at 10, 30, 40, 70 and 90%.

Median

Medianwas used to measure most frequent value of the data pattern.

Mode

Mode was used to measure most probable value in the signal pattern.

In addition to the above mentioned measures, we also computed variance, mean of derivatives, standard deviation of derivatives, and third moment. All these features were computed using MATLAB built-in functions. Such statistical features are computed in time domain, Fourier domain, and four different wavelet domains [Daubechies (db1), Coiflets (coif1), Symlets (sym4), and Dmey] to extract different information from the signal. In wavelet domain, these statistical features are computed from the approximation coefficients as well as from the detailed coefficients. So, we have 30 features from each wavelet domain. We repeated this for each different wavelet domain as mentioned earlier, so, we have 120 features total extracted from all wavelet domains.

Note that we do not know which of the features will be significant for detection of each disease, that's why we used t-test to screen out statistically insignificant features for each of the four diseases studied here. Eventually, four different algorithms are set to best detect each of the four ventricular abnormalities.

T-Test

Student t-test is most commonly used in the context of hypothesis testing. Student t-testuse t-distribution to identify the statistical significance of each feature. The method can be described as follows:

- Consider a particular feature of interest.
- Divide the values into two sets for normal and abnormal cases.
- Compute the mean and standard deviation for both sets.
- Use the t-test to compute the p-value of the null hypothesis that both sets do not have a statistically significant difference.
- The feature is suitable if the P-value is 0.05 or less.
- Eliminate any feature if the P-value is greater than 0.05 because there is no relation with the type of signal and it will over burden the classifier and waste computational power for nothing[23].

6. CLASSIFICATION

Machine Learning is the technology used for mining knowledge from data. It plays a central role in pattern/image recognition by classifying two or more classes of data patterns. The learning techniques that are used in training the classes depend on the patterns that are extracted from the raw data (features). In this study we use four main types of classifiers with existing functions in MATLAB, which are Support Vector Machine (SVM), Naïve Bayes, K-Nearest Neighbor (KNN), and Artificial Neural Network(ANN).

In SVM classifier, we tried different kernels, including linear, polynomial, and quadratic. In Bayesian classifier, we used the built-in function in MATLAB "Classify" with "diaglinear" type. In KNN classifier, we tried K= 1, 2, 3, 4, 5. In ANN classifier, we used feed-forward network with size [250 250 55 35 15], 'tansig' thresholding function for all network levels,'trainrp' training algorithm, and 'learngdm' for weight updating algorithm. Note that, in our results we showed; for each arrhythmia type, only the results of best five of all classifiers tried.

7. RESULTS AND DISCUSSION

In this chapter, we compare between the Accuracy, Sensitivity, Specificity, and Area Under the Curve (AUC) of the Receiver Operator Characteristic (ROC) curve to declare which of the classifiers are feasible for each type of arrhythmias studied. After that, we compared our results with previous work.

7.1 Ventricular Fibrillation (VF)

Figure 1 illustrates an example of 3secsVF lead II ECG time series sampled at 360 samples/sec.

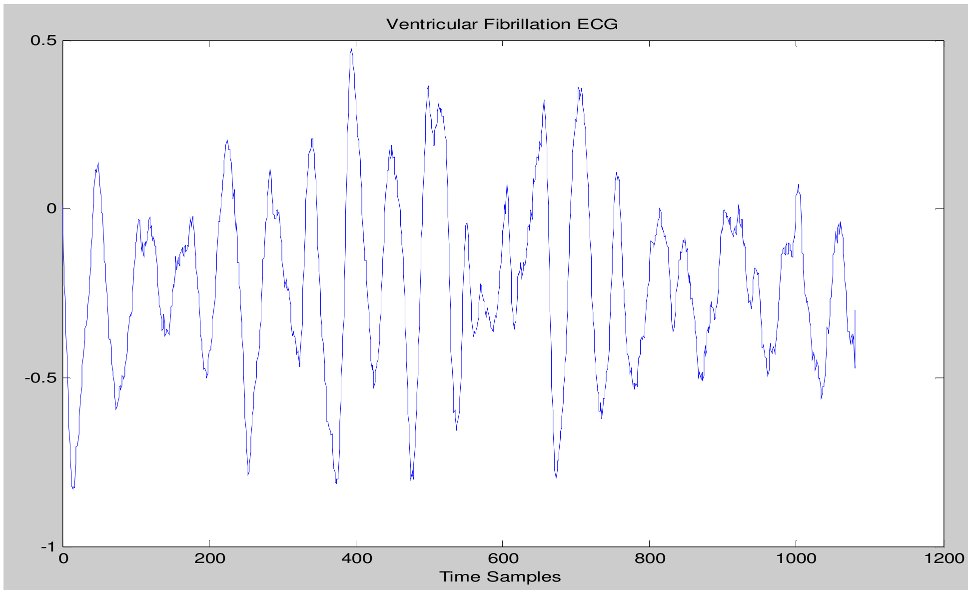


Figure 1. Ventricular Fibrillation (VF)

Table 2 summarizes the performance parameters of the best five classifiers among the ten we tried.

Table 2. The Performance evaluation of the best five classifiers (VF)

	*K*NN, K=1	SVM Linear	SVM Polynomial	Naïve Bayes	ANN
ERROR RATE	0.094	**0**	**0**	0.094	0.0192
Accuracy	0.954	**1**	**1**	0.954	0.991
Sensitivity	1	**1**	**1**	1	0.982
Specificity	0.915	**1**	**1**	0.915	1
PPV	0.906	**1**	**1**	1	1
NPV	1	**1**	**1**	0.968	0.968
AUC	0.958	**1**	**1**	0.958	0.991

The performance parameters of ANN shown in the table above are the average of five different runs.SVM classifier (with both linear and polynomial kernels) shows the best results among all classifiers tried as shown in Table 2 while the next best classifier is ANN.

Comparison of the results for VF types with published results:

The following table compares our best classifiers results with previous studies on the Ventricular Fibrillation (VF):

Table 3. Comparison of our best results for VF types with previous studies

Author and date	Main features	Classifiers	ACC
Pooyan and Akhoondi (2016) [13]	Morphological features	SVM	0.945
Lee et al. (2013) [24]	RR interval, QRS slope, and QRS shape similarity	SVM	0.883
Kavitha and Christopher [12]	Nonlinear analysis and -Nonlinear analysis	SVM	0.934
Issac et at (2005) [5]	RR intervals Heartbeat intervals and Spectral entropy	ANN	0.975
Bai et al. (2011) [25]	Frequency Spectrum Entropy (SpEn) and Energy Rate ERIMF	Naïve Bayes	0.9737
This study	FFT, wavelet transforms and statistical features (141 features)	KNN, K=1	0.954
		SVM Linear	1
		SVM Polynomial	1
		Naïve Bayes	0.954
		ANN	0.991

From Table 3, our SVM and ANN classifiers are better than previous studies on Ventricular Fibrillation (VF) with more than 1.5% in accuracy improvement.

7.2 Ventricular Tachycardia (VT)

Figure 2 an example of 3secs VT lead II ECG time series sampled at 360 samples/ sec.

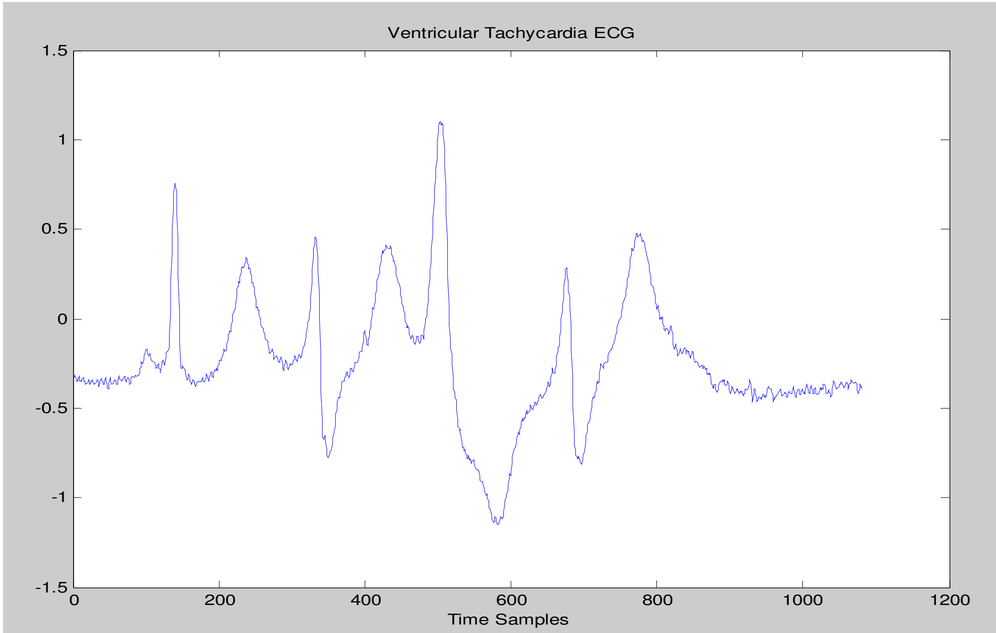


Figure 2. Ventricular Tachycardia (VT)

Table 4 summarizes the performance parameters of the best five classifiers among the ten we tried:

Table 4. The Performance evaluation of the best five classifiers (VT)

	*K*NN, K=1	SVM Linear	SVM Polynomial	Naïve Bayes	ANN
ERROR RATE	0.188	0.094	0.032	0.188	0
Accuracy	0.907	0.954	0.985	0.907	1
Sensitivity	0.965	1	1	0.934	1
Specificity	0.862	0.915	0.97	0.883	1
PPV	0.844	0.907	0.969	0.875	1
NPV	0.969	1	1	0.938	1
AUC	0.913	0.958	0.985	0.908	1

Note that, the performance parameters of ANN shown in the table above are the average of five different runs. ANN classifier shows the best result among all classifiers tried as shown in Table 4, while the next best classifier is SVM polynomial kernels.

Comparison of the results for VT types with published results:

The following table compares our best classifiers results with previous studies on the Ventricular Tachycardia (VT):

Table 5: Comparison of our best results for VT types with previous studies

Author and date	Main features	Classifiers	ACC
Pooyan and Akhoondi (2016) [13]	Morphological features	SVM	0.958
Lee et al. (2013) [24]	RR interval, QRS slope, and QRS shape similarity	SVM	0.926
Aparna et al. (2017) [26]	morphological features	SVM	0.954
Bai et al. (2011) [25]	Frequency Spectrum Entropy (SpEn) and Energy Rate ERIMF	Naïve Bayes	0.907
This study	FFT, wavelet transforms and statistical features (141 features)	KNN, K=1	0.907
		SVM Linear	0.954
		SVM Polynomial	0.985
		Naïve Bayes	0.954
		ANN	1

Table 5 our SVM (polynomial kernel) and Naïve Bayes classifiers accuracy is better than previous studies on Ventricular Tachycardia (VT).

7.3 Ventricular Couplet (VC)

Figure 3 illustrates an example of 3secs VC lead II ECG time series sampled at 360 samples/ sec.

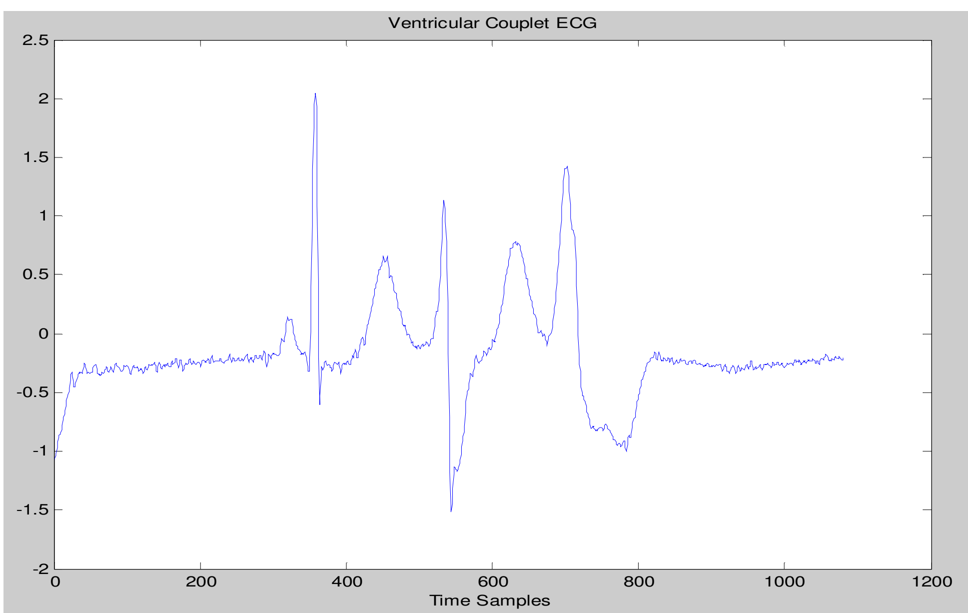


Figure 3. Ventricular Couplet (VC)

Table 6 summarizes the performance parameters of the best five classifiers among the ten we tried:

Table 6. The Performance evaluation of the best five classifiers (VC)

	*K*NN, K=1	*K*NN, K=2	SVM	SVM	ANN
			Linear	Polynomial	
ERROR RATE	0.25	0.25	0.125	0.25	0.063
Accuracy	0.875	0.875	0.938	0.875	0.969
Sensitivity	0.929	0.929	0.938	1	0.942
Specificity	0.834	0.834	0.938	0.8	1
PPV	0.813	0.813	0.938	0.75	1
NPV	0.938	0.938	0.938	1	0.938
AUC	0.881	0.881	0.938	0.9	0.971

The performance parameters of ANN shown in the table above are also the average of five different runs. ANN classifier shows the best result among all classifiers tried as shown in the Table 6, while the next best classifier is SVM (with linear kernel).

Comparison of the results for VC types with published results:

The following table compares our best classifiers results with previous studies on the VC:

Table 7: Comparison of our best results for VC types with previous studies

Author and date	Main features	Classifiers	Sensitivity
Owis et al. (2001) [21]	The correlation dimension and largest Lyapunov exponent	KNN, K=1	0.594
		KNN, K=2	0.656
		KNN, K=3	0.687
		KNN, K=4	0.687
		KNN, K=5	0.718
This study	FFT, wavelet transforms and statistical features (141 features)	KNN, K=1	0.929
		KNN, K=2	0.929
		SVM Linear	0.938
		SVM Polynomial	1
		ANN	0.942

Table 7 shows our KNN K=1 and K=2 classifiers Sensitivity is better than previous studies on Ventricular couplet. Furthermore, though not classifier comparable, our SVM and ANN results exceeded their best KNN results.

7.4 Ventricular Bigeminy (VB)

Figure 4 illustrates an example of 3secs VB lead II ECG time series sampled at 360 samples/ sec.

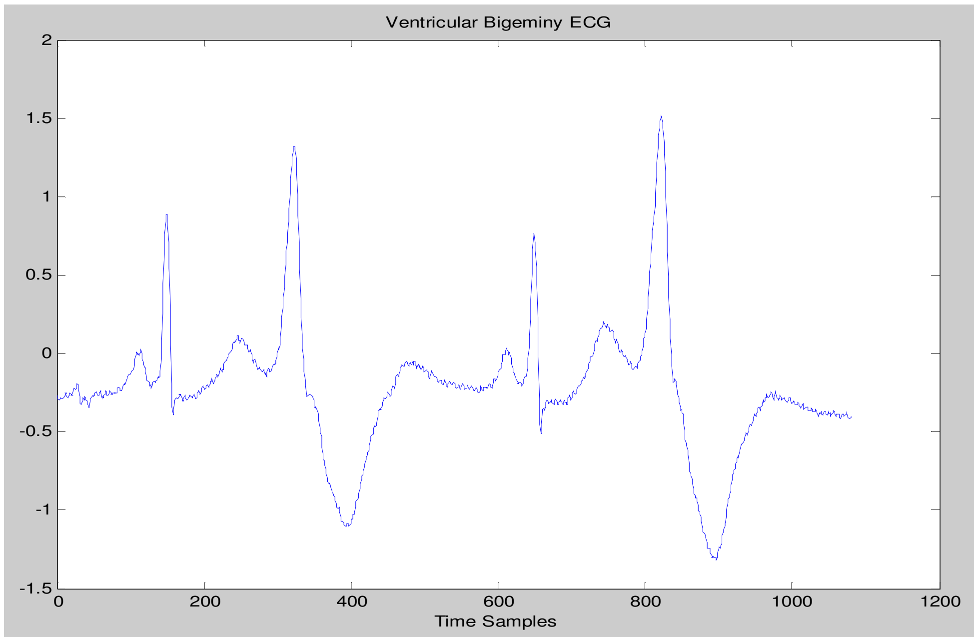


Figure 4. Ventricular Bigeminy (VB)

Table 8summarizes the performance parameters of the best five classifiers:

Table 8. The Performance evaluation of the best five classifiers (VB)

	*K*NN, K=3	*K*NN, K=4	SVM Linear	SVM Polynomial	ANN
ERROR RATE	0.219	0.25	0.094	0.219	0.063
Accuracy	0.891	0.875	0.954	0.891	0.969
Sensitivity	0.858	0.853	0.940	0.963	1
Specificity	0.932	0.9	0.968	0.838	0.942
PPV	0.938	0.907	0.969	0.813	0.938
NPV	0.844	0.844	0.938	0.969	1
AUC	0.895	0.877	0.954	0.901	0.971

Note that, the performance parameters of ANN shown in the table above are the average of five different runs. ANN classifier shows the best result among all classifiers tried as shown in the table above, while the next best classifier is SVM linear kernels.

Comparison of the results for VB types with published results:

The following table compares our best classifiers results with previous studies on the VB:

Table 9. Comparison of our best results for VB types with previous studies

Author and date	Main features	Classifiers	Sensitivity
Al-Atabany et al. (2004)[27]	CON, ASM, ENT, COR, MAX, and Inverse difference moment)	Naïve Bayes	0.846
Owis et al. (2001) [21]	The correlation dimension and largest Lyapunov exponent	KNN, K=1	0.593
		KNN, K=2	0.718
		KNN, K=3	0.718
		KNN, K=4	0.843
		KNN, K=5	0.812
This study	FFT, wavelet transforms and statistical features (141 features)	KNN, K=3	0.858
		KNN, K=4	0.853
		SVM Linear	0.940
		SVM Polynomial	0.963
		ANN	1

CHAPTER 1: From

Table 9 we notice that, our KNN (K=3 and K=4) classifiers Sensitivity is better than previous studies on Ventricular Bigeminy. Furthermore, SVM and ANN got even better results.

8. CONCLUSION

It is obvious that our study is better than all previous studies by combining lower and higher order statistical features computed on data in six different domains: time domain, Fourier domain, and four Wavelet domains. This made our CAD algorithm perform generally better as compared to previous studies regardless of studied abnormality or classifier used. However, we can see that SVM and ANN classifiers perform generally better than KNN and Naïve Bayes classifiers.

REFERENCES

[1] Wilhelm, K., (2008)“Classification Techniques for Undersampled Electromyography and Electrocardiography”, ProQuest Dissertations Publishing, pp 10.

[2] Chin, F., (2012)“A real-time data mining technique applied for critical ECG rhythm on handheld device”, pp. 19.

[3] AlGhatrif, M. & Lindsay J., (2012)“A brief review: history to understand fundamentals of electrocardiography”, Journal of community hospital internal medicine perspectives, Vol. 2, No. 1.pp14383.

[4] Islam M. K., Haque A. N. M. M., Tangim G., Ahammad T., and Khondokar M. R. H., (2012)“Study and analysis of ECG signal using MATLAB & LABVIEW as effective tools”, International journal of Computer and Electrical engineering,Vol. 4, No. 3, pp 404.

[5] Niwas S. I., Kumari R. S., and Sadasivam V., (2005)"Artificial neural network based automatic cardiac abnormalities classification", in Computational Intelligence and Multimedia Applications, Sixth International Conference on IEEE, pp 41- 46.

[6] Asl, B.M., Setarehdan S.K., and Mohebbi M., (2008)"Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal",Artificial intelligence in medicine, Vol.44, No.1, pp 51-64.

[7] Bani-Hasan, M., Kadah Y. M., and El-Hefnawi F., (2010)"Identification of cardiac arrhythmias using natural resonance complex frequencies",International Journal of Biological and Life Sciences, Vol. 6, No.3, pp 143-149.

[8] Kim, M.S., cho, Y. C., Seo, S., Son, C., and Kim, Y., (2011)"A new method of ECG feature detection based on combined wavelet transform for u-health service",Biomedical Engineering Letters,Vol. 1, No. 2, pp 108-115.

[9] Duque, A. O., SantiagoR´ua, SantiagoZuluaga, AlfredoRedondo, JoseV.Restrepo and JohnBustamante, (2013) "Support Vector Machine and Artificial Neural Network Implementation in Embedded Systems for Real Time Arrhythmias Detection", CentrodeBioingenieria, pp 310-313.

[10] Othman, M.A., Safri, N. M., Ghani, I. A., Harun, F. K. C. H., Ariffin, I, (2013) "A new semantic mining approach for detecting ventricular tachycardia and ventricular fibrillation",Biomedical Signal Processing and Control, Vol. 8, No. 2,pp. 222-227.

[11] Li, Q., Rajagopalan C., and Clifford G.D., (2014)"Ventricular fibrillation and tachycardia classification using a machine learning approach",IEEE Transactions on Biomedical Engineering, Vol.61, No. 6 pp 1607-1613.

[12] Kavitha, R. and Christopher T., (2015)"An effective classification of heart rate data using PSO-FCM clustering and enhanced support vector machine", Indian Journal of Science and Technology,Vol.8, No.30.

[13] Pooyan, M. and Akhoondi, F., (2016)"Providing an efficient algorithm for finding R peaks in ECG signals and detecting ventricular abnormalities with morphological features", Journal of medical signals and sensors, Vol. 6, No. 4,pp 218.

[14] Weixin, N., (2016) "A novel algorithm for ventricular arrhythmia classification using a fuzzy logic approach", Australasian physical & engineering sciences in medicine, Vol. 39, No. 4,pp 903-912.

[15] Tripathy, R., Sharma L., and Dandapat S., (2016)"Detection of shockable ventricular arrhythmia using variational mode decomposition", Journal of medical systems, Vol. 40, No.4, pp. 79.

[16] Sreedevi, G. & Anuradha, B., (2017)"ECG Feature Extraction and Parameter Evaluation for Detection of Heart Arrhythmias", i-Manager's Journal on Digital Signal Processing, Vol. 5, No. 1, pp 29.

[17] Mohanty, M., Sahooa, S., Biswalb, P., Sabutc, S. (2018) "Efficient classification of ventricular arrhythmias using feature selection and C4.5 classifier", Biomedical Signal Processing and Control, Vol. 44, pp 200-208.

[18] Mohammadalipour, S. H., Ahmadi, M., Shahghadami, R. and Chon, K. H.,(2018) "Automated Method for Discrimination of Arrhythmias Using Time, Frequency, and Nonlinear Features of Electrocardiogram Signals", Sensors (Basel, Switzerland), Vol. 18, No. 7.

[19] Moody, G.B., Mark, R.G. and Goldberger A.L., "PhysioNet: a web-based resource for the study of physiologic signals", IEEE Engineering in Medicine and Biology Magazine, Vol. 20, No. 3, pp 70-75.

[20] Goldberger, A.L., Amaral L.A., Glass L., Hausdorff J.M., Ivanov P.C., Mark R.G., Mietus J.E., Moody G.B., Peng C.K., Stanley H.E., (2000)"PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals", Circulation, Vol. 101, No.23,pp E215-20.

[21] Owis, E.M.I.I., Youssef, A., and Kadal, Y., (2001)"Novel techniques for cardiac arrhythmia detection", FACULTY OF ENGINEERING, CAIRO UNIVERSITY GIZA.

[22] Esmael, B., Arnaout, A., Fruhwirth, R. K., and Thonhauser, G., (2015) "A statistical feature-based approach for operations recognition in drilling time series", International Journal of Computer Information Systems and Industrial Management Applications, Vol. 3 pp 454-461.

[23] Kadah, P.Y.M. (2015) BIOSIGNAL PROCESSING. [Online] Available at: https://www.k-space.org/Class_Info/EE675/EE675_lec2.pdf [Accessed 14 Oct. 2018].

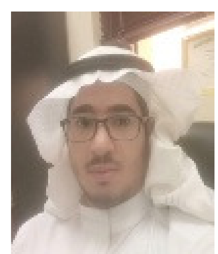
[24] Lee, S.H., H.-C. Ko, & Yoon, Y.-R., (2013)"Classification of ventricular arrhythmia using a support vector machine based on morphological features", in Engineering in Medicine and Biology Society (EMBC), 35th Annual International Conference of the IEEE.

[25] Bai, B. & Wang, Y., (2011)"Ventricular fibrillation detection based on empirical mode decomposition", in Bioinformatics and Biomedical Engineering, 5th International ConferenceIEEE.

[26] Aparna, P., Mirajkar, P., and Prabhu, R., (2017) "Detection and Classification of Ventricular Tachycardia Using SVM" International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering, National Conference on Advances in Electrical Engineering, Vol. 5, pp 116-120.

[27] El-Atabany, W., Youssef, A.-B.M. and Kadah, Y.M. (2004)"Nonlinear dynamical modeling of ECG signals based on phase density matrix representation", in Proc. 2nd Cairo International Biomedical Eng. Conference.

AUTHORS



Umar S. Alqasemi is currently Assistant Professor of Biomedical Engineering at the Dept. of Electrical and Computer Engineering, King Abdulaziz University (KAU), Jeddah, Saudi Arabia. In 2007, he received B.Sc. degree in Electrical and Computer Engineering (Biomedical Eng. Minor) from King Abdulaziz University (KAU). In 2011, he received M.Sc. degree in Biomedical Engineering from the University of Connecticut, Storrs, USA. In 2013, he received Ph.D. degree in Biomedical Engineering fromthe University of Connecticut, Storrs, USA. He published 14 peer-reviewed ISI journal papers, and 14 conference papers in the past 10 years. He is currently an active member of the Institute of Electrical and Electronics Engineers (IEEE), and the Society of Photo-optical Instrumentation Engineers (SPIE).

Saleh A. Alzahrani is currently Biomedical Engineer at the National Guards Hospital, Jeddah, Saudi Arabia. In 2007, he received B.Sc. degree in Electrical and Computer Engineering (Biomedical Eng. Minor) from King Abdulaziz University (KAU), Jeddah, Saudi Arabia. In 2011, he Received M.Sc. in Engineering Management from Huddersfield University, UK.In 2016 He started M.Sc. Study at KAU in Biomedical Engineering.