Premature Ventricular Contraction (PVC) Detection System Based on Tunable Q-Factor Wavelet Transform

Mohamad Hadi Mazidi[1]*[ID], Mohammad Eshghi[2], Mohammad Reza Raoufy[3]

[1]PhD, Department of Electric Engineering, Qeshm branch, Islamic Azad University, Qeshm, Iran
[2]PhD, Department of Electronics, Faculty of Electrical Engineering, Shahid Beheshti University, Tehran, Iran
[3]PhD, Department of Physiology, Faculty of Medical Sciences, Tarbiat Modares University, Tehran, Iran

ABSTRACT

Background: The Electrocardiogram (ECG) is an important measure for diagnosing the presence or absence of heart arrhythmias. Premature ventricular contractions (PVC) is a relatively large arrhythmia occurring outside the normal tract and being triggered outside the Sino atrial (SA) node of heart.

Objective: This study has focused on tunable Q-factor wavelet transform (TQWT) algorithm and statistical methods to detect PVC.

Material and Methods: In this analytical and statistical study, 22 ECGs records were selected from the MIT/BIH arrhythmia database. In the first stage the noise of signal remove and then five sub-bands create by TQWT. In the second stage nine features (minimum, maximum, root mean square, mean, interquartile range, standard deviation (SD), skewness, and variance) extracted of ECG and then the best features selected by using analysis of variance (ANOVA) test. Finally, the system is evaluated by using the learning machines of support vector machine (SVM), the K-Nearest Neighbor (KNN), and artificial neural network (ANN).

Results: The best results were verified with KNN learning machine: the sensitivity Se= 98.23% and accuracy Ac= 97.81%.

Conclusion: A comparative analysis with the related existing methods shows the method proposed in this study is higher than the other method for classification PVC and can help physicians to classify normal and PVC heart signals in the screening of the patients with coronary artery diseases (CADs).

Introduction

The heart is one of the most important parts of the body that is responsible for pumping blood into the coronary arteries. If the heart function is out of its normal state (rhythm), blood flow is not done well and this can lead to a serious risk for the individual [1]. Therefore, the accurate and timely diagnosis of heart arrhythmias is very important. The Electrocardiogram (ECG) is an important measure for diagnosing the presence or absence of heart arrhythmias [2].

ECG is a simple test and can be used to check the heart's rhythm and electrical activity [3]. The electrical activity of the heart begins in the Sinoatrial (SA) node, the primary pacemaker of the heart. A damaged

*Corresponding author:
Mohamad Hadi Mazidi
Department of Electric Engineering, Qeshm branch, Islamic Azad University, Qeshm, Iran
E-mail: mazidi_hadi@yahoo.com

or malfunctioning SA node results in the heart relying on the secondary pacemaker, the AV node, to control the rate of contraction of the heart muscle. If the AV node fails, the Purkinje fibers will control pacemaking in the heart [3].

The components of ECG are P wave, QRS complex, T- wave and U- waves. P-wave represents atrial depolarization, and QRS complex represents ventricular depolarization, the T- wave and eventually the U-wave, which reflect ventricular repolarization [3].

Premature ventricular contraction is one of the most common abnormalities that seriously affects older people (about 17% of people over 70 years of age) [4]. In addition, it includes about 33% of the total heart disease around the world [4]. PVCs may be a warning signal, alerting people to pay attention to this symptom and cure it early on. Otherwise, PVC can progress to serious heart disease that requires further medical treatment [5]. In the event of PVC arrhythmias, QRS complexes are wider, RR intervals increased, and P wave may not exist.

Processing of the ECG signals has attracted increasing attention in recent years [5]. Researchers have developed many PVC detection algorithms. Jung and Heeyoung used an automated diagnostic method for the PVC disease. In this paper, the statistical parameters of ECG using wavelet function were extracted and we tried to diagnose the PVC arrhythmias [6]. Zarei et al. extracted morphological features of the heart, using the PAT algorithm. They put the ECG samples in a row matrix using a linear analyzer for diagnosing PVC arrhythmias [7]. Zhou FY tried to diagnose the PVC arrhythmias using deep neural networks, such as lead convolutional neural network (LCNN) or long short-term memory (LSTM) network, and rules inference [8]. Rameshwari et al. suggested a simple and effective algorithm for extraction of the ECG signal characteristics and calculated the sensitivity and specificity parameters for this arrhythmia [9].

Lek-uthai et al. used several features of the ECG for diagnosing PVC. They used a state vector machine (SVM) for classifying the ECG signal [10]. Sabarimalai et al. detected PVC using the sparse signal decomposition, temporal features, and decision rules [11]. Cuesta et al. used only two QRS features for detection of PVC, and their proposed algorithm was tested on the MIT-BIH Arrhythmia database [12]. Shyu et al. used wavelet transform for feature extraction and fuzzy neural network (FNN) for classification. The proposed algorithm has a gross sensitivity of 97.04% for 7 patients from the MIT Arrhythmia Database [13].

Alajlan has proposed the capabilities of the hybrid support vector machine (SVM) and Gaussian process methods in the diagnosis and classification of premature ventricular contraction arrhythmias. In this study, various specifications of ECG signals, including morphology, discrete wavelet transform, and S transform have been calculated [14]. Gutiérrez-Gnecchi et al. have segmented the ECG signal into 6-second pieces. They used discrete-wavelet transform for extracting the properties of ECG signal and also used the probabilistic neural network (PNN) for classification. The results achieved by their method demonstrated an accuracy of 71.04% [15]. Other techniques are also used for the classification of PVC arrhythmias, which include symbolic dynamic analysis [16], morphological transformation and cross-correlation [17], and using linear and nonlinear techniques [18].

As shown, various methods have been presented for detecting the PVC arrhythmia based on the ECG signals, each using different techniques and each algorithm with different behavior in the diagnosis of this arrhythmia. However, no method has been presented to provide a large amount of information with high accuracy and sensitivity to identify PVC arrhythmias. For instance, in the study conducted by Robert Chen et al. only 15 min of data from the MIT-BIH database were utilized [5]. Besides, some of the methods proposed demonstrate high diagnostic results, while us-

ing a small amount of data. For example, in reference [13], PVCs have been categorized with an accuracy of 97.4% using wavelet method and fuzzy classification. Meanwhile, the above method has been performed on 7 data of the MIT-BIH database.

To address the aforementioned problems, in the present study, a new method has been presented for the diagnosis of PVC arrhythmias based on feature extraction using tunable Q-factor WT (TQWT). This method is capable of processing the 30-min ECG in the MIT-BIH database and achieving better results. In recent years, TQWT has been used for analyzing and processing oscillating signals in various problems [19]. TQWT can determine the dynamic properties of signals with abnormal beats according to normal signals. Most wavelet transforms are incapable of tuning their Q-factors (the Q parameter is the main factor in the signal oscillation), however, the TQWT is capable of adjusting this parameter and has thus been used as a powerful tool for analyzing oscillatory signals. In TQWT, the signal can be better examined by changing the Q parameter. Moreover, in TQWT technique, the signal is developed using the ideal structure of the rebuilding of filter banks, which will be further described in later sections.

This study presents a nonlinear algorithm for classification of PVC arrhythmias. After applying the TQWT algorithm to the signal, the statistical properties were extracted in each sub-band and then three popular classifiers: k-nearest neighbor (KNN), and support vector machine (SVM) and artificial neural network (ANN) were employed to evaluate the performance of the proposed scheme. This method will provide effective assistance to doctors when screening for patients with cardiovascular diseases (CVDs).

The MIT-BIH arrhythmia database was utilized as a database to test the effectiveness of the proposed method [20]. This database includes 48 signals recorded from different individuals from 32 to 89 years-old. Each record-

ed signal contains 30 min, and two leads have been used for ECG recording: a modified body lead and one of the modified leads of v1, v2, v4 and v5. In this database, the sampling frequency is adjusted at 360 Hz and the data has a pass band of 0.1-100 Hz. The signals have a resolution of 11 bits in the range of 10 mV. In this study, lead 2 has been used. Furthermore, in accordance with the AAMI recommended practice, records #102, #104, #107 and #217 were discarded because these beats do not have sufficient signal quality for reliable processing [21], then 44 records remained. These records were divided into two datasets (DS1 and DS2) with each dataset containing ECG data from 22 recordings. We used DS1 dataset in this work.

The remainder of this paper is organized as follows: The ECG database is discussed in Section 2. Section 3 presents our proposed PVC classification method, including filtering, TQWT based decomposition, feature extraction, classification method, and evaluation parameters. Section 4 presents the results, including experiment description, validation of the model's efficacy, and comparisons with previously published work. Finally, Section 5 concludes this paper.

Material and Methods

In this analytical and statistical study, the block diagram of proposed PVC detection is shown in Figure 1, consisting of four stages, namely data preprocessing, feature extraction, classification, and PVC arrhythmia detection.

Filtering

The ECG signal is generally composed of noises, including interface noise, power line, motion artifact due to electrode, skin interface, patient breathing, etc. [22]. On the other hand, in order to increase the accuracy of detection of the PVC arrhythmias, the noise in the ECG signal should be eliminated, thus a filtering operation is required. From the power spectral analysis of different signal components
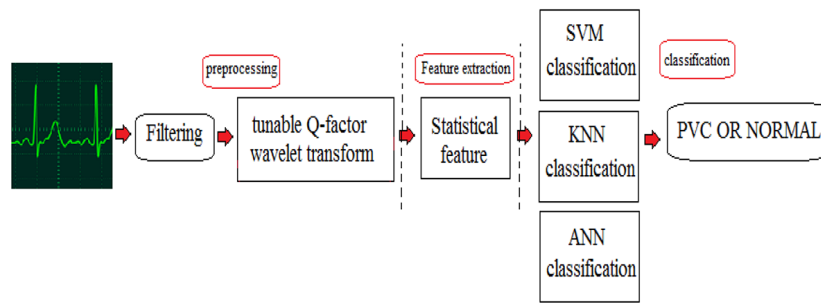
Figure 1: Block diagram of the proposed premature ventricular contraction (PVC) detection

in the ECG signal, we notice that the maximum power density of the QRS complex lies between 5-20 Hz. Therefore, a finite impulse response (FIR) band-pass filter with a band of 5 to 16 Hz is used for removing noise [23].

As the sampling frequency of the ECG signal is 360 Hz for the MIT database, stop-band is 0 to 4 Hz and 17 to 180 Hz. The stop-band of band pass filter between 0 and 4 Hz removes the baseline wander noise (between 0 and 1 Hz) and low frequency noises. Moreover, in a frequency range between 16 and 180 Hz, the stop-band filter eliminates power line interface noise, electromagnet noise, and high frequency noises. In this way, a noise-free signal will be achieved applicable to the remainder of algorithm process.

TQWT-based decomposition

The TQWT method has been described in [24]. This transform depends on three main adjustable parameters, including Q-factor (Q), redundancy (R), and decomposition level (J). In addition, parameters such as Center Frequency (FS), bandwidth and coefficients α and β are important in TQWT. Q-factor is the oscillation of the ECG signal, R is the redundancy of the wavelet transform, and J is the number of steps in which the signal is divided into the sub-bands. Moreover, FS and BW indicate the sampling rate of input and the band width, respectively [25].

The input signal S [M] with the sampling frequency FS passes through a filter bank.

This filter bank consists of a low-pass filter (LP) and a high-pass filter (HP) with αfs and βfs sampling frequencies, respectively, where α and β are the low pass scaling and high pass scaling factors, respectively.

In order to get the low-pass sub-band, low-pass filter H_0 (W) and low-pass scaling LPS-α are required. The high-pass sub-band H is obtained using H_1 (W) and high-pass scaling HPS- β [26]. Figure 2 illustrates the signal decomposition stages.

The same frequency responses for low-pass and high-pass sub-bands signal, which were obtained from J-stages and shown with H_0^J (*W*) and H_1^J (*W*), respectively, are expressed as Eqs. 1 and 2 [24, 26].

$$H_0^{(j)}(\omega) = \begin{cases} \Pi_{m=0}^{j-1} H_0\left(\dfrac{\omega}{\alpha^m}\right), \; |\omega| \le \alpha^j \pi \\ 0, \; \alpha^j \pi < |\omega| \le \pi \end{cases} \quad (1)$$

$$H_1^{(j)}(\omega) = \begin{cases} H_1(\dfrac{\omega}{\alpha^{j-1}}) \Pi_{m=0}^{j-2} H_0\left(\dfrac{\omega}{\alpha^m}\right) \\ (1-\beta)\alpha^{J-1} \; \pi \le |\omega| \le \alpha^{J-1}\pi \\ 0, \; for\, other\, \omega \in [-\pi, \pi] \end{cases} \quad (2)$$

Where,

$$H_0(\omega) = \theta(\frac{\omega + (\beta - 1)\pi}{\alpha + \beta - 1}) \quad (3)$$

$$H_1(\omega) = \theta(\frac{\alpha \pi - \omega}{\alpha + \beta - 1}) \quad (4)$$
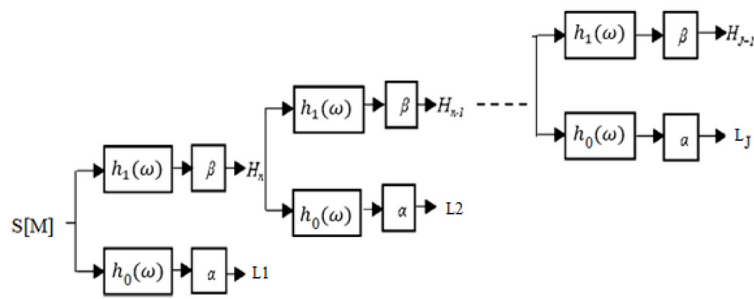
Figure 2: Decomposition of the input signal into low pass sub-band and high pass sub-band at Jth levels by tunable Q-factor wavelet transform (TQWT)

In which $(1-\beta)\pi<\omega<\alpha\pi$. It should be taken into account that in this relation, $\theta(\omega)$ is the frequency response of the Daubechies filter with two vanishing moments. $\theta(\omega)$ can be expressed as follows [24]:

$$\theta(\omega) = 0.5(1+\cos(\omega))\sqrt{2-\cos(\omega)}, |\omega| \leq \pi \quad (5)$$

Moreover, the values of α and β in relation 3 and 4 with respect to the values of R and Q as Q >= 1 and R > 3, are calculated from the following equation:

$$\beta = \frac{2}{Q+1}, \alpha = 1 - \frac{\beta}{r} \quad (6)$$

Based on TQWT decomposition in J stages, J + 1 sub-bands are obtained. These sub-bands signal can be given by cell array D as follows [24]:

$$D = \{W_1, W_2, \ldots W_J, W_{J+1}\} \quad (7)$$

Where, $W_{(J+1)}$ is a sub-band signal having the lowest frequency, and W_1 to W_J are sub-band signals that contain high signal frequencies. The number of samples obtained at each stage based on the sampling frequency of the signal and the coefficients α and β with respect to J >= 1, is expressed as follows [24]:

$$D_1 = \beta\alpha^{J-1}F_S \quad (8)$$

In this study, as explained in Section 2, the MIT-BIH database was used. Based on the investigations and application of different Q values, the best value for analyzing ECG data was obtained as Q = 5. Besides, the J decomposition level was determined for 4 sub-bands. Figures 3a and b demonstrate the TQWT analysis of the ECG signal from a healthy person and a person with the PVC disease. From Figure 3, the difference between normal and PVC is seen in each sub-band of the proposed method. It means the statistical power is different at each of the sub-bands.

Feature extraction

Feature extraction is one of the most important sections in heart signal processing, as the correct feature extraction can enhance the accuracy of diagnosis of arrhythmias. In this study, the ECG statistical characteristics were extracted for each sub-band obtained by TQWT. The statistical method was divided into two stages. First, the division technique was considered. For this purpose, we used PAT algorithm in this paper [27]. The PAT algorithm is divided into four steps, differentiation of the filtered signal is used to distinguish the QRS complex from other ECG waves. Then, a nonlinear transformation is performed through point-to-point squaring of the filtered ECG. After that, integration is carried out by a moving time window to extract additional features, such as the QRS width. Finally, adaptive amplitude thresholds are applied to the averaged signal to detect R peaks [27]. By extracting the R peak, 1800 samples of any data that were equivalent to 5 s from the ECG were
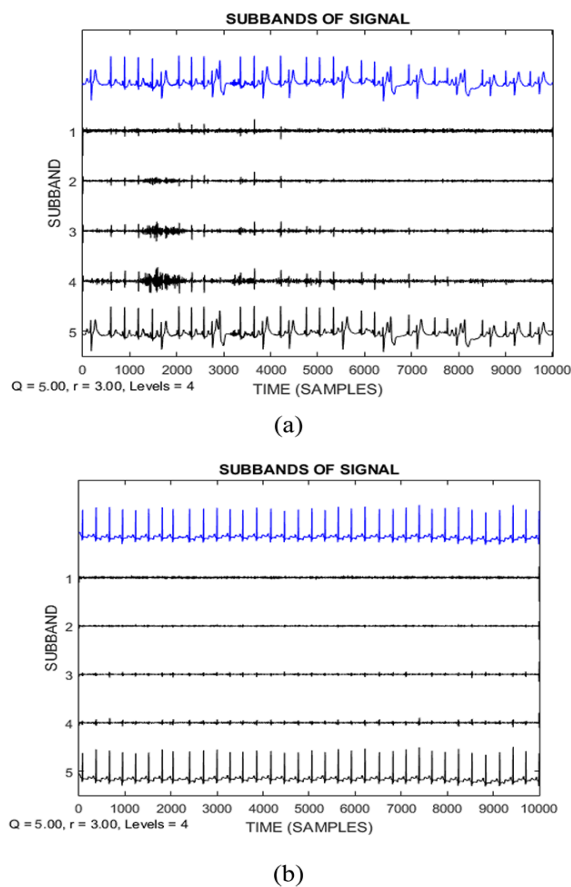
(a)

(b)

Figure 3: Examples of normal and premature ventricular contraction (PVC) and their five sub-brands obtained using the tunable Q-factor wavelet transform (TQWT). a) Example of PVC (data106 of Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH) database) decomposition, b) Example of a normal Electrocardiogram (ECG) (# 100 of MIT-BIH) decomposition

considered, and the TQWT was implemented. The values of SB1, SB2,…, SB5 sub-bands were calculated for these 5 s, and repeated for the whole signal that is 30 min for each record.

In the second step, 9 statistical characteristics were extracted in each step, including the minimum, maximum, root mean square, mean, interquartile range, standard deviation (SD), skewness, and variance [28]. Statistical features can provide useful information regarding the ECG in order to diagnose the PVC arrhythmia. Some of the related mathematical

relations have been presented below.

$$SD = \sqrt{\frac{\sum_{i=1}^{N}(x_i - SM)^2}{N-1}} \qquad (9)$$

$$Min = \min(x_i) \qquad (10)$$

$$Max \quad \max(x) \qquad (11)$$

$$Mean = \frac{1}{N}(\sum_{i=1}^{N} x_i) \qquad (12)$$

$$Med = (\frac{N+1}{2})^{th} \qquad (13)$$

$$Var = \sum_{i=1}^{N} \frac{(x_i - SM)^2}{N} \qquad (14)$$

$$RMS = \sqrt{\frac{1}{N}(\sum_{i-1}^{N} x_i^2)} \qquad (15)$$

$$Ran = Max - Min \qquad (16)$$

$$SK = \sum_{i=1}^{N} \frac{(x_i - SM)^3}{(N-1)SD^3} \qquad (17)$$

With the value of x(n) = 1, 2,…, n, is a time series, N and SM are the number of data points and the mean value of the samples, respectively.

Certainly, all of the extracted features are not suitable for this research to achieve the maximum classification performance in which the boundaries of each class are optimally distinctively defined from others. It is also very problematic and time-consuming to examine all subsets associated with the extracted features because if there are n features extracted in the previous step, there will be 2n-1 of non-empty subsets entirely. One of the most essential phases during a learning machine problem is introducing the best subset of the extracted features by a feature selection methodology. There are variations of feature selection meth-

od for selecting the best feature, and in this paper, analysis of variance (ANOVA) test is used [29]. The ANONA is tested using the one way ANOVA, taking one random sample from each population under consideration, with significance level (α) of 0.05. Using this test, the parameters rms, mean, median, std and var were selected. In the next step, the desired data will be classified.

Classification method

In the current study, three machine learning classifiers of SVM, KNN, and ANN were utilized to evaluate the features extracted technique and finally, the best classifier was considered for diagnosis. In the following section, these machine learning classifiers will be further described in detail.

K nearest neighbors (k-NN)

Searching for the nearest neighbor is an optimization problem to find the closest points in metric spaces. The problem is: the set S containing a number of points in a metric space such as M and also a query point $q \in M$ are given, and the objective is to find the closest point to q in S. In most cases, the space M is measured as a d-dimensional Euclidean space and the distance between points is specified by Euclidean distance, Manhattan distance, or other metric distances [30].

In the KNN method, K returns the neighbor closer to the query point. This method is usually exploited in prediction analysis to estimate or categorize a point based on the consensus of its neighbors. Therefore, choosing the k parameter in the KNN algorithm is important and should be performed with accuracy. In this study, the k = 1 value was considered. The elimination of the noise generated in the data is one of the advantages of the KNN algorithm, increasing the accuracy of the PVC detection.

Support Vector Machine (SVM)

In this research, a support vector machine (SVM) has been used to classify the data. In today's machine learning applications, SVM is considered as one of the most powerful and precise methods among other popular algorithms. SVM is used in various applications such as biomedical signal classification, image classification, text categorization, and bioinformatics.

The SVMs can be considered as a linear or nonlinear classifier as they have different kinds of kernel function. With a linear kernel function, the SVM becomes a linear classifier, and with a nonlinear kernel function, such as a polynomial, and a sigmoid function, it becomes a non-linear classifier [31].

In this paper, the non-linear Radial Basis Function (RBF) kernel has been used. One of the reasons for the use of SVM in this study was that the SVM grants a better generalization solution if its parameters are well chosen.

Artificial Neural Networks (ANN)

Artificial Neural Networks (ANNs), or more simply the neural networks, are modern systems and computational methods for machine learning, presentation of knowledge, and, finally, the application of knowledge obtained to predict the output responses from complex systems. The main idea behind such networks is some extent inspired by the function of the biological nervous system to process data and information in order to learn and create knowledge [32].

This system consists of a large number of super-integrated processing elements called neurons working together in coordination to solve a problem and transmit information through synapses (electromagnetic communications). In the current study, a back-propagation neural network was applied to diagnose PVC arrhythmia with the PURLINE output function.

Evaluation parameters

After presenting the desired algorithm, the evaluation of effectiveness of this algorithm in detecting the PVC arrhythmia should be examined. The data in this study have been selected from the MIT-BIH database. Table 1 represents the number of training and testing

data for each data set from database used in this paper.

To diagnose PVC arrhythmias, four unique parameters called True Positive (TP), False Negative (FN), False Positive (FP), and True Negative (TN) have been used in this paper as shown in Table 2 [6]. For example, TN, the algorithm presented in this study, detects the healthy ECG as normal, and FP detects healthy ECG as PVC.

Based on these parameters, to assess the practical results obtained in this study and to compare them with other studies, two parameters have been defined as [6]:

1. Accuracy (AC): the result of dividing of the number of correctly detected peaks by the total number of peaks;

$$AC = \frac{TP+TN}{TP+TN+FN+FP} \times 100\% \qquad (18)$$

2. Sensitivity (SE): The result of dividing the number of peaks that correctly detected PVC arrhythmias by the total number of PVC peaks;

$$SE = \frac{TP}{TP+FN} \times 100\% \qquad (19).$$

Results

Experiment description

In this paper, the experimental results were obtained through the MIT-BIH database. In addition, the TQWT method was used to extract features. In order to extract the TQWT sub-bands, first, the peak R was obtained from PAT algorithm, and in the second step, the signal was converted to small 5-second segments, and 5 sub-bands were calculated for each of them. This was repeated for the whole 30-min signal. Then, 9 statistical characteristics were extracted for each of the sub-bands. The important features were investigated separately through three SVM, KNN, and ANN machine learning classifiers. Finally, a classifier was selected as the preferred classifier for the detection of PVC arrhythmias. Experimental results were obtained using the MATLAB software (version 2015) on a computer system with the configuration of Intel (R) Core™ i7 and 4500U CPU and 6 GB of RAM.

Validation of the model's efficacy

The proposed method was evaluated using

Table 1: Numbers of the training and testing sets used in this study

SIG	Total	Training	Testing
101	1865	1305	560
106	2027	1216	811
108	1774	1241	533
109	2532	1772	760
112	2539	1778	761
114	1879	1315	564
115	1953	1367	586
116	2412	1689	723
118	2288	1601	687
119	1987	1390	597
122	2476	1734	742
124	1619	1133	486
201	2000	1400	600
203	2980	2086	894
205	2656	1859	797
208	2955	2068	887
209	3005	2103	902
215	3363	2345	1018
220	2048	1434	614
223	2605	1823	782
230	2256	1597	659

Table 2: Definition of True Positive (TP), False Negative (FN), False Positive (FP), and True Negative (TN)

		Classified labels	
		PVC	
			Normal
True	Normal	TN	FP
labels	PVC	FN	TP

PVC: Premature ventricular contraction, TN: True Negative, FP: False Positive, FN: True Negative, TP: True Positive

the evaluation parameters. Table 3 demonstrates the measured parameters of each sub-band using the SVM, KNN, and ANN classification algorithms for diagnosis of the PVC arrhythmias based on the MIT-BIH database. Regarding Table 3, the ECG signal has been divided into 5 sub-bands using the TQWT algorithm and different values of signal accuracy and sensitivity for MIT-BIH database have been obtained with three different methods.

As it is clear based on the classification methods, various values are obtained according to the TQWT method for each record of the MIT-BIH database in each sub-band. For instance, the accuracy obtained from record 124 associated with a 77-yearold man with PVC arrhythmia, has been calculated 97.75, 98.85, and 97.95, respectively, from the SVM, KNN, and ANN classification methods in sub-band 1. Meanwhile, the accuracy obtained in sub-band 5 for this record was 99.38, 100, and 98.59, respectively.

Table 4 indicates the mean values obtained from the MIT-BIH data. According to the results of the Table 4, there were significant differences in the accuracy and sensitivity of the SVM, KNN, and ANN classification methods based on the proposed method. Accordingly, the lowest accuracy and sensitivity were associated with the results of the ANN and SVM classification methods with values of 94.52 and 94.80, respectively.

Moreover, it can be observed from Table 4 that the KNN algorithm in comparison to the SVM and ANN methods has a higher accuracy and sensitivity of the proposed method in all of its sub-bands as shown in Figure 4.

Figures 4a and b demonstrate a comparison between the average values of parameters obtained from the feature extraction by the TQWT method with the SVM, ANN, and KNN classifiers. It is clear that the highest accuracy and sensitivity in the KNN method for the SB5 sub-band were obtained as 98.20 and 100%, respectively. Thus, the combination of the TQWT with the KNN algorithm is more ef-

fective in comparison with the ANN and SVM methods for diagnosis of the PVC arrhythmia (as illustrated in Table 5). Table 5 represents the mean values of each sub-band using the results obtained from the proposed algorithm and based on Table 4. It can be seen that after KNN, the highest accuracy and sensitivity were related to the SVM and ANN algorithms, respectively.

Discussion

The present study aims to investigate PVC detection arrhythmia using TQWT and statistical features, regarding the accuracy of ECG beat classification. In the beat-based training scheme, three learning machine (KNN, SVM and ANN) are used to classify ECG signal. Moreover, ANOVA technique is used to reduce the number of features used in the classification process.

The results of our proposed method were compared with 5 published studies which have used the MIT-BIH database. Their performance results from our approach were compared with the corresponding reported results. Table 6 shows the results of the current study with other ones in detail. The results show that performance of our proposed algorithm compared to other studies is more appropriate to detect the PVC arrhythmias using KNN classifiers.

Robert Chen et al. proposed a PVC classification method based on wavelet transform (WT) and a combination of the sum of trough and sum of R peak [5]. They used a test-set of 5 recordings 5 data from the MIT database to determine the performance of their classifier. The accuracy parameter was reported 94.73 in this study, however, our proposed method was evaluated on 22 records of the MIT database and the results of ours are higher than this reference.

As stated in the introduction, Jung Y and Kim H. proposed an ECG monitoring procedure for the diagnosis of PVC beats. After ECG, signals were decomposed by discrete wavelet

Table 3: Calculation of evaluation parameters of Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH) database base on tunable Q-factor wavelet transform (TQWT)

SIG	SB1						SB2						SB3					
	SVM		KNN		ANN		SVM		KNN		ANN		SVM		KNN		ANN	
	AC	SE	AC	SE	AC	SE	AC	SE	AC	SE	AC	SE	AC	SE	AC	SE	AC	SE
101	98.3	98	99.2	98.8	96.07	93.03	98.7	99.1	100	100	95.2	93.2	98.7	99.3	100	100	93.59	94.48
106	97	97.2	98.5	98.7	93.58	91.4	99.4	99.3	100	100	93.47	94.74	86.35	90.88	100	98.6	87.3	90.14
108	99.53	99.63	99.02	98.35	98.5	99.21	98.73	98.9	98.7	99.6	99.26	100	99.38	99.44	99.07	99.21	99.21	100
109	99.60	99.06	100	98.43	98.55	100	99.43	99.19	100	100	98.41	100	98.28	98.66	98.41	100	98.41	100
112	99.8	99.6	100	100	92.77	91.93	100	100	100	100	95.42	95.48	100	100	100	100	97.04	98.09
114	98.6	98.9	100	100	97.52	99.81	98.05	98.19	98.05	99.45	97.87	99.81	97.87	99.63	100	98.05	100	100
115	97.7	98.2	99.97	99.95	95.47	98.98	96.11	97.23	98.85	99.85	95.36	96.02	96.02	96.58	98.2	99.25	95.23	92.58
116	98.7	97.74	97.28	96.47	94.27	95.85	92.27	94.58	96.19	98.98	95.27	97.58	96.27	97.58	97.65	98.78	95.19	94.28
118	99.12	99.85	98.54	99.12	99.12	100	99.12	99.85	99.41	99.85	98.44	100	99.83	99.56	99.27	99.7	99.27	99.85
119	95.85	96.09	96.25	96.35	94.89	98.41	93.62	94.87	94.25	96.78	93.46	97.29	98.22	98.09	95.53	97.57	94.57	98.13
122	96.02	97.12	96.23	97.98	95.14	97.55	96.25	97.23	98.87	96.35	95.27	97.23	97.85	95.68	98.23	99	95.28	97.12
124	97.75	100	98.85	100	97.95	99.78	97.13	98.73	97.75	100	97.75	99.36	97.13	98.94	98.15	99.36	98.36	99.37
201	92.62	93.1	94.09	96.55	93.93	97.25	91.47	95.01	95.62	95.09	93.28	97.78	91.63	92.01	93.93	96.37	93.27	98.9
203	87.93	88.6	93.45	92.35	85.96	97.23	87.25	88.98	93.2	91.78	87.23	99.23	87.15	88.01	93.24	92.73	85.49	97.23
205	98.25	99.87	99.87	99.22	98.37	99.48	98.37	99.61	99	100	98.75	99.61	98	99.84	100	100	98.12	99.01
207	92.5	94.46	94.25	94.12	94.59	94.84	92.85	94.25	95.98	94.52	92.28	94.71	93.25	94.12	97.5	97.89	95.89	98.12
208	82.87	83.25	94.23	93.47	76.97	91.58	82.58	81.27	90.88	86.85	72.95	89.43	83.85	87.25	91.2	90.25	70.25	88.22
209	100	100	100	100	99.95	100	99.85	100	100	100	99.98	100	100	99.99	100	100	99.95	100
215	96.95	97.93	99.34	100	97.35	99.47	97.44	98.86	100	100	97.44	99.89	96.85	98.55	98.03	99.25	96.36	98.85
220	97.8	97.65	94.29	97.25	95.76	96.89	95.43	94.24	95.45	95.72	95.97	94.27	94.98	97.47	97.43	97.65	97.85	95.89
223	90.89	91.98	94.47	94.14	92.03	97.98	92.12	95.18	96.38	94.60	87.86	95.30	90.53	94.39	91.26	95.84	91.15	94.66
230	100	100	100	100	99.94	100	100	100	100	100	98.98	99.87	99.87	100	100	100	99.78	100
total	96.26	96.73	97.62	97.78	94.99	97.30	95.60	96.57	97.63	97.70	94.54	97.18	95.54	96.63	97.49	98.28	94.52	97.04

SIG	SB4 SVM AC	SB4 SVM SE	SB4 KNN AC	SB4 KNN SE	SB4 ANN AC	SB4 ANN SE	SB5 SVM AC	SB5 SVM SE	SB5 KNN AC	SB5 KNN SE	SB5 ANN AC	SB5 ANN SE
101	98.5	98.2	98.2	98.89	89.57	90.88	99.6	98.8	99.6	100	96.12	98.12
106	92.45	91.43	100	100	88.8	91	94.8	92.48	99.8	100	91.54	90.4
108	98.35	99.37	99.08	99.63	99.21	99.08	97.8	96.1	100	100	99.63	99.8
109	98.55	99.06	98.55	99.33	98.28	99.86	94.5	93.7	98.9	100	98.28	99.86
112	98.8	99.6	100	99.8	97.73	97.44	100	100	100	100	95.16	95.66
114	96.99	98.91	97.7	99.09	97.7	100	95.58	96.38	98.81	100	97.52	99.63
115	98.23	97.02	96.87	98.23	96.23	98.23	94.98	95.23	98.89	100	98.5	98.5
116	98.32	97.23	96.27	97.14	97.14	97.89	95.89	97.25	98.85	100	98.26	99.52
118	99.12	99.85	99.12	99.56	98.98	99.85	97.09	97.66	98.98	100	99.41	100
119	96.16	95.48	100	98.7	94.93	98.13	88.78	87.93	98.25	100	90.35	96.54
122	97.23	97.12	98.87	96.89	97.25	97.78	95.14	95.21	98.81	100	96.87	99.25
124	99.47	99.15	97.34	98.73	98.33	100	99.38	98.52	100	100	98.59	100
201	93.27	94.37	96.11	95.64	94.59	99.46	87.04	89.47	95.68	100	91.96	99.82
203	87.43	83.33	89.55	87.84	86.77	99.49	86.38	80.8	96.23	100	96.23	99.75
205	98.25	99.74	98.52	100	97.50	98.58	97.87	99.22	98.58	100	97.5	98.97
207	94.89	95.12	95.23	98.21	95.78	97.59	94.89	95.28	97.23	100	94.23	95.28
208	89.16	88.68	89.98	90.25	85.25	96.25	83.28	89.25	94.3	100	96.23	99.5
209	99.98	100	100	100	99.97	99.89	100	100	100	100	99.89	100
215	96.26	98.65	97.54	100	98.33	99.89	97.23	98.14	98.66	100	97.25	99.69
220	97.27	97.28	98.38	98.47	97.12	97.25	95.12	95.14	99.23	100	100	98.28
223	91.02	91.52	95.56	95.37	90.89	96.24	90.64	91.87	96.91	100	89.88	95.09
230	100	100	100	100	99.89	100	99.78	99.52	99.79	100	99.69	100
total	96.35	96.41	97.40	97.80	95.46	97.94	94.80	94.90	98.52	100	96.46	98.34

SB: Sub-band, SVM: Support vector machine, KNN: K-nearest neighbors, ANN: Artificial neural network, AC: Accuracy, SE: Sensitivity

Table 4: Mean values of Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH) data evaluation using tunable Q-factor wavelet transform (TQWT) method

SB	PRF	SVM	KNN	ANN
SB1	AC	96.26	97.62	94.99
SB1	SE	96.73	97.78	97.30
SB2	AC	95.60	97.63	94.54
SB2	SE	96.57	97.70	97.18
SB3	AC	95.54	97.49	94.52
SB3	SE	96.63	98.28	97.04
SB4	AC	96.35	97.80	95.46
SB4	SE	96.41	97.40	97.94
SB5	AC	94.80	98.52	96.46
SB5	SE	94.90	100	98.34

SB: Sub-band, PRF: Performance, SVM: Support vector machine, KNN: K-nearest neighbors, ANN: Artificial neural network, AC: Accuracy, SE: Sensitivity

transforms, significant wavelet coefficients were extracted for constructing a monitoring statistic based on Hotelling's T2 statistics [6]. The accuracy of this study was approximately the same as the proposed method, however, the sensitivity obtained in the proposed method of the present paper is significantly different from this reference.

Roozbeh Zarei et al. [7] performed extraction of the ECG features and diagnosis of the PVC arrhythmia based on the replacement method and the PAT algorithm [27]. The sensitivity of the method proposed in the current study, which is an important parameter in the evaluation of studies, is better than the method offered in [7]. Furthermore, the method suggested in [7] can only detect PVC arrhythmias, whereas the scheme proposed in the present

paper is capable of detecting other cardiac arrhythmias.

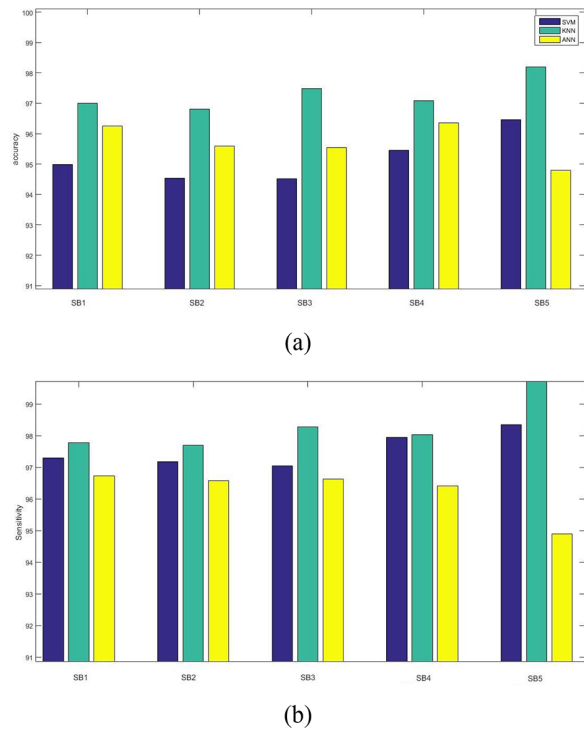F Zhou et al. [8] proposed the algorithm that combined deep neural networks and rules inference for PVC detection. The accuracy and the sensitivity of this study on the MIT-BIH database were 99.41%, 97.59%, respectively. LCNN and LSTM are both deep learning models, that they used in their work and their architectures are more complex than our architectures KNN. Additionally, the experimental results of our research showed that our method had obtained high sensitivity [8].

A premature ventricular contraction (PVC) detection method based on the sparse signal decomposition, temporal features, and decision rules is proposed by Sabarimalai et al, [11]. This method classifies PVC beats by combining ten generalized temporal features with decision-rule- detection algorithm. Comparing this study with the current investigation, the sensitivity of the proposed method is considerably higher, besides, the accuracy has not been declared in the former study. Based on the results of Table 6, we can conclude that our work, using TQWT and the KNN learning machine, has offered some improvement com-

(a)

(b)

Figure 4: Comparison of performance of the three classifiers of support vector machine (SVM), K-nearest neighbors (KNN), and artificial neural network (ANN) based on the proposed feature extraction method in terms of: a) accuracy, b) sensitivity

Table 5: Comparison of mean parameters of the proposed method measured with support vector machine (SVM), K-nearest neighbors (KNN), and artificial neural network (ANN)

PRF	SVM	KNN	ANN
AC	95.71	97.81	95.19
SE	96.30	98.23	97.56

PRF: Performance, SVM: Support vector machine, KNN: K-nearest neighbors, ANN: Artificial neural network, AC: Accuracy, SE: Sensitivity

Table 6: Comparison of the results with previously-studied articles

Reference	Rec.No.	Technique	AC	SE
Robert Chen-Hao Chang [5]	5	wavelet transform+combines the sum of trough and sum of R_peak	94.73	-
Jung Y and Kim [6]	9	Wavelet+SPC	99	94.3
Roozbeh Zarei [7]	22	"Replacing" strategy+PCA	98.77	96.12
Zhou FY [8]	22	deep neural networks+ rules inference	99.41	97.59
Manikandan MS [11]	47	signal decomposition+temporal features combined with decision-rule	-	89.69
This paper	22	TQWT+ statistical features +KNN	97.81	98.23

AC: Accuracy, SE: Sensitivity, SPC: Statistical process control, PCA: Principal component analysis, TQWT: Tunable Q-factor wavelet transform, KNN: K-nearest neighbors

pared to other works.

A limitation of the current study is that the proposed method has limited performance with the processing tools in real time due to the complexity of mathematical computations in the TQWT algorithm. Nevertheless, detecting other CVDs in addition to PVC diagnosis is one of the strengths of the method proposed in this study, hence it is an appropriate method to be used in cardiac clinics to diagnose CVDs. Moreover, the main advantage of the TQWT-based feature extraction is that this technique divides the large volume of ECG data into smaller sets, thus it is capable of processing a large amount of data at a lower cost relative to other methods.

Conclusion

In this study, detection of the PVC arrhythmia was proposed according to the TQWT method based on MIT-BIH database. The ECG signal passed through the filter and after finding the R peak divided into 5 sub-bands using the TQWT method based on its parameters, Q, R, and J level. After signal decomposition, nine statistical features were used to segment the sub-bands and the most important features were extracted. These features were put in one set for each sub-band and then the signal was analyzed using three machine learning classifiers SVM, KNN, and ANN. The experimental results reveal that the combination of the proposed TQWT method with the KNN classifier has better performance compared to the SVM and ANN methods. For detection of PVC, our algorithm achieves an accuracy of 97.81% and sensitivity of 98.23% for 30 min of each MIT-BIH recording. In the future, we will implement the proposed method for wearable tools using FPGA.

Authors' Contribution

Mazidi MH, M Eshghi M, Raoufy MR, designed and directed the project; Mazidi MH, M Eshghi M, performed the experiments; Mazidi MH, M Eshghi M, Raoufy MR, made the simulations. All the authors read, modified, and approved the final ver-

sion of the manuscript.

Ethical Approval

This article does not contain any studies with human participants or animals.

Conflict of Interest

None

References

1. Asl BM, Setarehdan SK, Mohebbi M. Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal. *Artif Intell Med.* 2008;**44**(1):51-64. doi: 10.1016/j.artmed.2008.04.007. PubMed PMID: 18585905.

2. Chan AD, Hamdy MM, Badre A, Badee V. Wavelet distance measure for person identification using electrocardiograms. *IEEE Trans Instrum Meas.* 2008;**57**(2):248-53. doi: 10.1109/TIM.2007.909996.

3. Kadish AH, Buxton AE, Kennedy HL, et al. ACC/AHA clinical competence statement on electrocardiography and ambulatory electrocardiography: A report of the ACC/AHA/ACP–ASIM Task Force on Clinical Competence (ACC/AHA Committee to Develop a Clinical Competence Statement on Electrocardiography and Ambulatory Electrocardiography) Endorsed by the International Society for Holter and Noninvasive Electrocardiology. *J Am Coll Cardiol.* 2001;**38**(7):2091-100. doi: 10.1161/circ.104.25.3169. PubMed PMID: 11748119.

4. Acharya UR, Joseph KP, Kannathal N, et al. Heart rate variability: a review. *Med Biol Eng Comput.* 2006;**44**(12):1031-51. doi: 10.1007/s11517-006-0119-0. PubMed PMID: 17111118.

5. Chang RC, Lin CH, Wei MF, Lin KH, Chen SR. High-precision real-time premature ventricular contraction (PVC) detection system based on wavelet transform. *J Signal Process Syst.* 2014;**77**(3):289-96. doi: 10.1007/s11265-013-0823-6.

6. Jung Y, Kim H. Detection of PVC by using a wavelet-based statistical ECG monitoring procedure. *Biomed Signal Process Control.* 2017;**36**:176-82. doi:10.1016/j.bspc.2017.03.023.

7. Zarei R, He J, Huang G, Zhang Y. Effective and efficient detection of premature ventricular contractions based on variation of principal directions. *Digit Signal Process.* 2016;**50**:93-102. doi: 10.1016/j.dsp.2015.12.002.

8. Zhou FY, Jin LP, Dong J. Premature ventricular contraction detection combining deep neural networks and rules inference. *Artif Intell Med.* 2017;**79**:42-51. doi: 10.1016/j.artmed.2017.06.004. PubMed PMID: 28662816.

9. Mane RS, Cheeran AN, Awandekar VD, Rani P. Cardiac

arrhythmia detection by ecg feature extraction. *Int J Eng Res Appl.* 2013;**3**(2):327-32.

10. Lek-uthai A, Ittatirut S, Teeramongkonrasmee A. Algorithm development for real-time detection of premature ventricular contraction. TENCON 2014 - IEEE Region 10 Conference; Bangkok, Thailand: IEEE; 2014. p. 1-5. doi: 10.1109/TENCON.2014.7022418.

11. Manikandan MS, Ramkumar B, Deshpande PS, Choudhary T. Robust detection of premature ventricular contractions using sparse signal decomposition and temporal features. *Healthcare Technology Letters.* 2015;**2**(6):141-8. doi: 10.1049/htl.2015.0006. PubMed PMID: 26713158. PubMed PMCID: PMC4678438.

12. Cuesta P, Lado MJ, Vila XA, Alonso R. Detection of premature ventricular contractions using the RR-interval signal: a simple algorithm for mobile devices. *Technol Health Care.* 2014;**22**(4):651-6. doi: 10.3233/THC-140818. PubMed PMID: 24898863.

13. Shyu LY, Wu YH, Hu W. Using wavelet transform and fuzzy neural network for VPC detection from the Holter ECG. *IEEE Trans Biomed Eng.* 2004;**51**(7):1269-73. doi: 10.1109/TBME.2004.824131. PubMed PMID: 15248543.

14. Alajlan N, Bazi Y, Melgani F, Malek S, Bencherif MA. Detection of premature ventricular contraction arrhythmias in electrocardiogram signals with kernel methods. *Signal Image Video Process.* 2014;**8**(5):931-42. doi: 10.1007/s11760-012-0339-8.

15. Gutiérrez-Gnecchi JA, Morfin-Magaña R, Lorias-Espinoza D, et al. DSP-based arrhythmia classification using wavelet transform and probabilistic neural network. *Biomed Signal Process Control.* 2017;**32**:44-56. doi: 10.1016/j.bspc.2016.10.005.

16. Zhao L, Wiggins M, Vachtsevanos G. Premature ventricular contraction beat detection based on symbolic dynamics analysis. International Conference CIRCUITS, SIGNALS AND SYSTEMS; Cancun, Mexico: IASTED; 2003. p. 48-50.

17. Nahar S, Bin Munir MS. Automatic detection of premature ventricular contraction beat using morphological transformation and cross-correlation. 3rd International Conference on Signal Processing and Communication Systems; Omaha, NE, USA: IEEE; 2009. p. 1-4. doi: 10.1109/ICSPCS.2009.5306426.

18. Mazidi MH, Eshghi M, Raoufy MR. Detection of premature ventricular contraction (PVC) using linear and nonlinear techniques: an experimental study. *Cluster Comput.* 2019:1-6. doi: 10.1007/s10586-019-02953-x.

19. Patidar S, Pachori RB, Upadhyay A, Acharya UR. An integrated alcoholic index using tunable-Q wavelet transform based features extracted from EEG signals for diagnosis of alcoholism. *Appl Soft Comput.* 2017;**50**:71-8. doi: 10.1016/j.asoc.2016.11.002.

20. Suppappola S, Sun Y, Chiaramida SA. Gaussian pulse decomposition: an intuitive model of electrocardiogram

waveforms. *Ann Biomed Eng.* 1997;**25**(2):252-60. doi: 10.1007/BF02648039. PubMed PMID: 9084830.

21. AAMI. Testing and reporting performance results of ventricular arrhythmia detection algorithms. ANSI/AAMI EC57:1998/(R)2003, United States: Association for the Advancement of Medical Instrumentation; 1986.

22. Clifford GD, Azuaje F, Mcsharry P. Advanced methods and tools for ECG data analysis, Chapter 3: ECG Statistics, Noise, Artifacts, and Missing Data. Clifford Lab; 2006. p. 18.

23. Thakor NV, Webster JG, Tompkins WJ. Estimation of QRS complex power spectra for design of a QRS filter. *IEEE Trans Biomed Eng.* 1984;**31**(11):702-6. doi: 10.1109/TBME.1984.325393. PubMed PMID: 6500590.

24. Selesnick IW. Wavelet transform with tunable Q-factor. *IEEE Trans Signal Process.* 2011;**59**(8):3560-75. doi: 10.1109/TSP.2011.2143711.

25. Patidar S, Pachori RB. Segmentation of cardiac sound signals by removing murmurs using constrained tunable-Q wavelet transform. *Biomed Signal Process Control.* 2013;**8**(6):559-67. doi: 10.1016/j.bspc.2013.05.004.

26. Patidar S, Pachori RB, Acharya UR. Automated diagnosis of coronary artery disease using tunable-Q wavelet transform applied on heart rate signals. *Knowl-Based Syst.* 2015;**82**:1-10. doi: 10.1016/j.knosys.2015.02.011.

27. Pan J, Tompkins WJ. A real-time QRS detection algorithm. *IEEE Trans Biomed Eng.* 1985;**32**(3):230-6. doi: 10.1109/TBME.1985.325532. PubMed PMID: 3997178.

28. Sharma P, Ray KC. Efficient methodology for electrocardiogram beat classification. *IET Signal Processing.* 2016;**10**(7):825-32. doi: 10.1049/iet-spr.2015.0274.

29. Faul F, Erdfelder E, Lang AG, Buchner A. G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods.* 2007;**39**(2):175-91. doi: 10.3758/BF03193146.

30. Homaeinezhad MR, Atyabi SA, Tavakkoli E, Toosi HN, Ghaffari A, Ebrahimpour R. ECG arrhythmia recognition via a neuro-SVM–KNN hybrid classifier with virtual QRS image-based geometrical features. *Expert Syst Appl.* 2012;**39**(2):2047-58. doi: 10.1016/j.eswa.2011.08.025.

31. Mazidi MH, Eshghi M. Detection of Heart Attack using Cross Wavelet Transformation and Support Vector Machine. *Appl Med Inform.* 2019;**41**(3):77-92.

32. Özbay Y, Ceylan R, Karlik B. Integration of type-2 fuzzy clustering and wavelet transform in a neural network based ECG classifier. *Expert Syst Appl.* 2011;**38**(1):1004-10. doi: 10.1016/j.eswa.2010.07.118.