

Classification of Premature Ventricular Contraction in ECG

Yasin Kaya and Hüseyin Pehlivan
Department of Computer Engineering
Karadeniz Technical University
Trabzon, TURKEY

Abstract—Cardiac arrhythmia is one of the most important indicators of heart disease. Premature ventricular contractions (PVCs) are a common form of cardiac arrhythmia caused by ectopic heartbeats. The detection of PVCs by means of ECG (electrocardiogram) signals is important for the prediction of possible heart failure. This study focuses on the classification of PVC heartbeats from ECG signals and, in particular, on the performance evaluation of time series approaches to the classification of PVC abnormality. Moreover, the performance effects of several dimension reduction approaches were also tested. Experiments were carried out using well-known machine learning methods, including neural networks, k-nearest neighbour, decision trees, and support vector machines. Findings were expressed in terms of accuracy, sensitivity, specificity, and running time for the MIT-BIH Arrhythmia Database. Among the different classification algorithms, the k-NN algorithm achieved the best classification rate. The results demonstrated that the proposed model exhibited higher accuracy rates than those of other works on this topic. According to the experimental results, the proposed approach achieved classification accuracy, sensitivity, and specificity rates of 99.63%, 99.29% and 99.89%, respectively.

Keywords—ECG; arrhythmia; classification; k-NN; PVC

I. INTRODUCTION

According to recent reports, cardiovascular disease (CVD) is listed as a major underlying cause of death, accounting for 54.5% and 47.73% of all deaths in the United States[1] and in Turkey[2], respectively. In order to reduce the mortality rate caused by CVD, monitoring heart cycles for the recognition of early complications is a vital concern for cardiologists and related medical personnel.

An arrhythmia is an abnormal cardiac rhythm. Heart arrhythmias are caused by any disruption in the regularity, rate, or transmission of the cardiac electrical impulse [3]. Among the various abnormalities, premature ventricular contraction (PVC) is one of the most significant arrhythmias [4]. PVC results from the early depolarisation of the myocardium originating in the ventricular area and is a widespread form of arrhythmia in adults. PVC is common, with an estimated occurrence of 1 to 4% in the general population. It is often seen along with structural heart disease and increases the risk of sudden death. Moreover, its assessment and treatment are complex [4], [5]. This paper focuses on the classification of PVC arrhythmias.

In recent years, numerous studies have been conducted on automatic recognition of cardiovascular system problems. Researchers attempting to classify PVC arrhythmias have mostly used time-frequency analysis techniques, statistical measurements, and hybrid methods. The most recently published works are those presented in [6–15]. In [6], the authors applied a dynamic Bayesian network for PVC classification. In [7], Ittatirut et al. attempted to detect PVCs for real-time applications. Their work employed a real-time algorithm for PVC detection based on a low computational method. Simple decision rules were used in the classifier process, which was suitable for embedded applications. Another study [8] compared the learning capability and classification skill for normal heartbeats with PVC clustering using four classification techniques: neural networks (NN), the k-nearest neighbour method (k-NN), discriminant analysis (DA) and fuzzy logic (FL). In [9], the authors used the k-NN method to classify PVC beats and normal beats, while the authors in [10] tried to detect PVC using a neural network-based classification scheme and extracted 10 ECG (electrocardiogram) structural features and one timing interval feature. In [11], a low-complexity data-adaptive method for PVC recognition was designed which achieved an accuracy of 98.2% in the tests. In [12], the authors focused on manifold learning for PVC detection and proposed a method for PVC recognition using manifold learning and support vector machines (SVM). A neural network-based ECG pattern recognition method was presented in [13]. In that study, NN correctly distinguished normal heartbeats and PVCs in 92% of the proposed cases. In [14], the authors tried to classify PVC via an NN classifier and used a wavelet transform to extract morphological features from ECG data. In [15], Independent Component Analysis (ICA) was used for feature extraction and k-means and Fuzzy C-Means (FCM) classifiers were employed to recognize the PVC beat. All of these studies [6-15] used ECG records from the MIT-BIH Arrhythmia Database.

In this paper, an effective and comparative approach was developed for the classification of PVC arrhythmias. The main objective was to improve the accuracy of cardiac arrhythmia classification and examine the performance of time series and their equivalent reduced-size features of ECG signals. The time series of the signal was used to evaluate performance metrics for classification. In addition, principal component analysis (PCA), independent component analysis (ICA), and self-organising maps (SOM) were used to reduce the size of input

feature vectors. To obtain the experimental results, NN, k-NN, SVM and decision tree (DT) classification algorithms were applied using different schemes. In order to provide a better representation, the test data used in the analysis were selected from the MIT-BIH Arrhythmia Database. The results showed that the proposed approach obtained the considerably high classification accuracy rate of 99.63% and provided better classification performance than other approaches studied previously.

II. MATERIAL AND METHODS

All ECG signals comprising Lead II (containing normal or PVC beats) from the MIT-BIH Arrhythmia Database were used in this work. The signal was passed through pre-processing for de-noising. Beat parsing was performed on the noise-free signal, and 200 samples were selected as the cycle of the ECG beat. Because the sampling frequency of the signal was 360 Hz., the 200 points around the QRS complex as a signal window were the approximate equivalent of one cardiac cycle. In total, 7000 windowed ECG beats were used for the analysis.

In this study, the focus was on the improvement of PVC classification performance. Memory requirements and the complexity of the model were reduced by optimising the input vectors. Thus, the model required less operational time. Fig. 1 illustrates a block diagram of the proposed approach for classifying the PVC beat in the ECG of an arrhythmia. The functioning of each step is described in detail in the following sections.

A. ECG Database

The MIT-BIH Arrhythmia Database [16], [17] was used as the data source for this study. The database contains 48 signals of 30 min duration each, and two leads – Lead II and one of the modified leads (V1, V2, V4, or V5). The signals of the database were sampled at 360 Hz. Twenty-three files were randomly selected to serve as a representative sample of routine clinical recordings and 25 files were selected to include uncommon complex ventricular, junctional, and supraventricular arrhythmias. The database was annotated both in timing information and beat label. In this work, the annotation labels were used to locate the beats in the signal files. A total of 43 data files were used, marked as: 100, 101, 103, 105, 106, 107, 108, 109, 111, 112, 113, 115, 116, 117, 118, 119, 121, 122, 123, 124, 200, 201, 202, 203, 205, 207, 208, 209, 210, 212, 213, 214, 215, 217, 219, 220, 221, 222, 223, 228, 230, 231, and 234. The remaining files were not used because they did not contain Lead II or related beats. Eight files of the selected records did not contain normal beats and ten did not contain PVC beats. Approximately 100 normal beats were selected for the test from each file. The data used consisted of 3500 (from 35 files) normal (N) beats and 3500 (from 33 files) PVC beats. The PVC beats were intermittently selected from the files because these beats were unevenly distributed in the files. Table I gives details of the distribution of the selected beats from the MIT-BIH Arrhythmia Database.

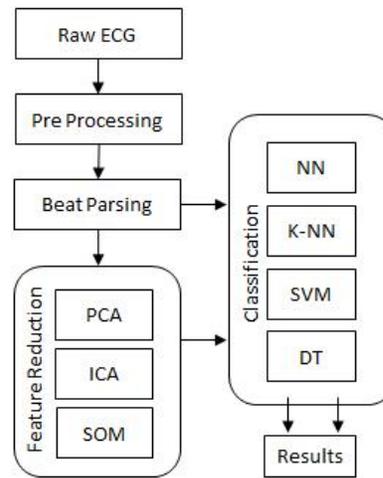


Fig. 1. Overall system architecture

TABLE I. TOTAL NUMBER OF SELECTED BEATS FROM MIT-BIH ARRHYTHMIA DATABASE

Files	Beats		
	N	PVC	TOTAL
100,101,103,205	401	42	443
106,107,108,109	99	622	721
111,112,113,115	300	1	301
116,117,118,119	300	568	868
121,122,123,124	300	51	351
200,201,202,203	400	456	856
205,207,208,209	300	296	596
210,212,213,214	300	669	969
215,217,219,220	400	389	789
221,222,223,228	400	400	800
230,231,234	300	6	306
TOTAL	3500	3500	7000

B. Preprocessing

Noise reduction in ECG signals is a significant problem. There are several noise factors in the ECG: EMG noise, power line noise, baseline wander, and composite noise [18]. Fluctuations in the amplitude of ECG signals have a negative effect on the calculated feature vectors. The same type of ECG signals taken from different patients can show remarkable variances. The differences in ECG signals are minimised by performing normalisation and pre-processing operations.

In this study, the mean of the signal was set to zero. The zero mean signal $\{y(t) | 1 \leq t \leq L\}$ was calculated using Equation (1):

$$y(t) = x(t) - \bar{x} \quad (1)$$

where $y(t)$ is the calculated signal, $\{x(t) | 1 \leq t \leq L\}$ is the raw ECG, \bar{x} is the arithmetic mean of $x(t)$, and L is the length of the signal.

Thereafter, a median filter was used to reduce noise. The median filter is a simple nonlinear smoother that can suppress noise while holding sharp edges in signal values [19].

The filtered signal $\{Y(t) | 1 \leq t \leq L\}$ was calculated using Equation (2):

$$Y(t) = \text{med}\{y(t-1), y(t), y(t+1)\} \quad (2)$$

where $Y(t)$ is the filtered signal and $y(t)$ is the input signal.

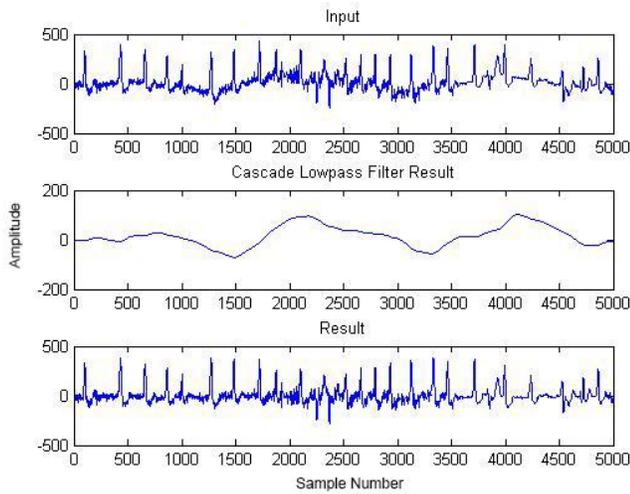


Fig. 2. Input signal, cascade low-pass filter result, and final result of the filter from data file 203

A cascade low-pass filter to remove frequency components below 2 and 0.5 Hz from the signal was applied in the final signal $Y(t)$ to remove the baseline wander and powerline noise. Frequency components of the baseline wander are generally below 0.5 Hz; however, in the event of a stress test, this value can be higher. Consequently, the frequency limit was adjusted to 2 Hz [20]. The required change of filter type from low- to high-pass filters can be achieved by subtracting the output of the low-pass filter from the suitably delayed input signal. Fig. 2 demonstrates the input signal, cascade-filtered signal, and filter results of the first 5000 samples of data file 203 on the MIT-BIH Database.

C. Beat Parsing

Each beat's window length of 200 points was established from the filtered ECG signal according to the location of the R point in the QRS complex (99 points on the left side of the R point, 100 points on the right side of the R point, and the R point itself). The associated location of the R points composed the annotation files of the MIT-BIH Database. No QRS detection algorithm was used. The selected beats constituted a 7000×200 data matrix.

D. Feature Reduction

In this study, the time series of the one beat was used for classification. In addition, feature reduction methods were used for dimension reduction. Consequently, the performance of the classification algorithms using the time series and their reduced features were compared. PCA, ICA, and SOM were used to reduce the size of the input vectors, and the computation time of classification was diminished. Both single-beat time series and reduced dimension data were used as input vectors of the classifiers for comparison and a notable acceleration was obtained.

PCA is a numerical technique that uses perpendicular conversion to transform a set of observations of possibly correlated features into a set of values of uncorrelated features, called principal components [21]. ICA is a very versatile statistical method in which observed random data are linearly transformed into elements that are maximally independent from each other [22]. SOM is an unsupervised neural network method, improved by Kohonen, which proposes an effective and easily interpretable mapping from a higher dimensional input space into a lower dimensional (especially, two-dimensional) space [23], [24]. The feature reduction parameters are described in the result section.

E. Classification

In this work, NN, k-NN, SVM, and DT classification algorithms were used for classification, and are briefly discussed below.

A three-layered feed-forward neural network was applied for pattern classification in this study [25]. The input layer was composed of 200 nodes corresponding to the 200 points of one beat. Moreover, results of the PCA, ICA and SOM feature reduction approaches were also tested using this method. In that case, the sizes of the input layer were 2, 17, and 10 for the PCA, ICA and SOM, respectively. The output layer consisted of two nodes.

The k-NN algorithm is one of the most conventional methods in pattern recognition because of its effective non-parametric nature. The nearest neighbour decision rule assigns the classification of the closest training samples in the feature space to an unclassified sample point [26]. This algorithm does not depend on the statistical distribution of training samples. The classification process of the samples is realised according to the nearest neighbourhood of training examples. The algorithm uses numerous distance measures. An instance is classified by a majority vote of its k-nearest neighbours. In this work, k was established as 1 after the parameter optimisation step. Euclidean distance was used as the measure function.

SVM is popular in machine learning for pattern recognition, especially for binary classification [27], [28]. The input data are transformed into a high-dimensional feature space. In this space, the data points are linearly separated by a hyper-plane. Because the patterns are not linearly separable in most cases, the patterns are mapped into a high-dimensional space using an appropriate kernel, and then, the optimisation step is fulfilled. Various kernel transformations are used for mapping the data into high-dimensional space, some of which include linear, sigmoid, polynomial, and radial basis function (RBF). In this study, parameter optimisation was used to find the optimum SVM parameters. After this stage, the C parameter was set as 100, the Gamma parameter was set as 4, and the polynomial was selected as the kernel-type parameter.

DT is a predictive model which can be used to characterise both classifiers and regression models. DT refers to a hierarchical model of decisions and their results and is used to classify a sample into a predefined set of classes based on their feature values. DT consists of nodes that form a rooted tree meaning. It is a directed tree with a node called a root that has

no entering edges. All other nodes have only one entering edge. A node with outgoing edges is referred to as a test node. All other nodes are known as leaves, or decision nodes [29]. Each leaf is allocated to one class demonstrating the most accurate target value. The leaf holds a probability vector specifying the probability of the target feature with a definite value.

Thus, from the last leaf to the root, the most likely path to the destination can be calculated by multiplying all other probability values of the leaves. The efficiency of the calculation can be improved by cutting specific branches of the tree or changing the defining characteristics. There are many common decision tree algorithms, some of which are ID3, C4.5, CART, CHAID, and MARS. At the generating stage of the DT, the gain ratio was used as the criterion parameter, 4 as the minimal size for the split, 2 as the minimal leaf size, and 20 as the maximal depth.

III. EXPERIMENTS AND RESULTS

The approach was tested on 200 time series samples of one beat. These samples were applied to the classification methods discussed in Section II as the input vectors. A parameter optimisation step was performed to obtain optimum parameter values.

In the NN classifier, a hidden layer consisting of 10 neurons was used. The output layer consisted of two neurons. The size of the hidden layer was selected by empirical observation. Even numbers between 2 and 20 were tested for hidden layer size. In the hidden layer, maximum accuracy was obtained at around 10 neurons. Table II shows classification accuracies versus neuron size of the hidden layer of the neural network classifier using time series, ICA, PCA, and SOM features as input vectors. The NN was trained by a back propagation algorithm. At the training and testing stage, training cycle and learning rate parameters were set as 500 and 0.3, respectively. The error threshold parameter was set as 0.00001 to terminate the iterations when mean square error (MSE) was attained.

As a result of a grid search, the present experiments showed that the best k value of the k-NN algorithm was found at one; however, all k values in the test range achieved high results. Euclidean distance was used as the distance measure in this study. Since the k-NN classifier obtained the highest results, it was used in the parameter optimisation stages of the feature reduction algorithms.

For the SVM classification experiments, parameters were determined using a grid search like that done with the k-NN experiments. Four kernel functions (polynomial, RBF, sigmoid and linear), a complex SVM fixed parameter (C) having 12 different kernels with values in the exponential range of 0-1000, and 18 different Gamma parameters having values in the exponential range of 0-100 were tested by the grid search. After the optimisation stage, the polynomial kernel function was selected, C was set as 100, and Gamma set as 4.

TABLE II. NEURAL NETWORK CLASSIFIER CLASSIFICATION ACCURACIES (%) FOR DIFFERENT INPUT VECTORS AND HIDDEN LAYER SIZE

Hidden layer size	Time series	ICA	PCA	SOM
2	0.9729	0.961	0.9677	0.7879
4	0.9783	0.9704	0.9791	0.8744
6	0.9873	0.9766	0.9829	0.924
8	0.983	0.9771	0.9827	0.9343
10	0.9846	0.9779	0.9814	0.9466
12	0.98	0.9773	0.982	0.9574
14	0.981	0.977	0.9864	0.9703
16	0.982	0.9776	0.9831	0.9756
18	0.9801	0.9783	0.9817	0.9723
20	0.9809	0.9779	0.9851	0.977

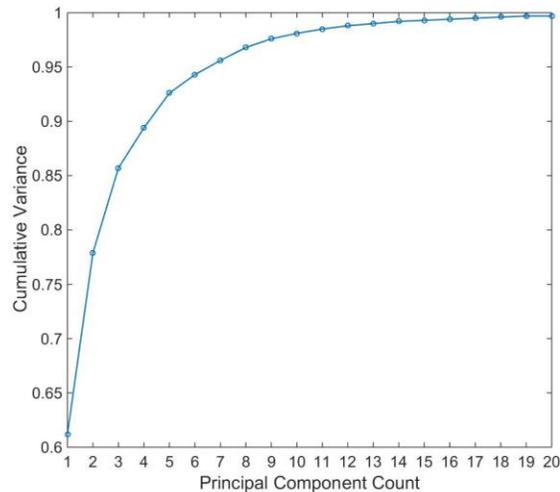


Fig. 3. Cumulative variance versus number of PCs for first 20 principal components

In addition, PCA, ICA and SOM were applied to reduce the size of the feature vectors. Processed data were used in the same classifiers. Remarkable achievements were obtained and classification time was reduced. Before implementing the classification test, the grid search was used to find the counts of the best principal components (PCs) and independent components (ICs) resulting in the highest accuracy rate for the classifiers. It was found experimentally that k-NN classifiers feeding PCA features achieved the highest accuracy. The calculation of the SOM features took more time than other dimension reduction methods. The computation times of the PCA, ICA, and SOM feature reduction methods were 2.5, 1.2, and 67.3 s, respectively.

As shown in Fig. 3, when calculating principal components, cumulative variance started with a small number and increased rapidly. Cumulative variance reached 0.926 and 0.997 at principal component counts 5 and 20, respectively. All principal components between 0 and 30 were tried out and the number of principal components that provided the best result for the classification algorithms was calculated.

When the principal component count was 17, the cumulative variance was 0.995. This value obtained the best result; therefore, the PCs = 17 value was used as the principal component for the tests in this study.

The FastICA algorithm was used for calculating the independent components [22] in the ICA experiments. All independent components between 0 and 30 were tried out with parameter optimisation and the number of independent components providing the highest result for the classification algorithms was calculated. The ICs = 10 value obtained the highest results according to the experiments.

SOM was used to reduce the size of the input vector to 2. The network size was taken as 30 and the training rounds were specified as 30. The two-dimensional output vector was calculated by the SOM network to be used as the input vectors of the classification algorithms.

Classification models have a common strategy of dividing the dataset into two parts, one for training and the other for testing. The classification accuracy obtained from the test part more precisely projects the performance. An upgraded version of this technique is known as cross-validation. A 10-fold cross-validation method was used in this study for training and testing of the classification algorithms. In the 10-fold cross-validation, first, the dataset was split into 10 subsets of the same size. Sequentially, one subset was evaluated using the classification algorithm trained on the other 9 subsets. Thus, each subset of the whole dataset was predicted once. The average accuracy of these 10 trials was calculated as a classification result. The cross-validation accuracy is the percentage of data which are properly classified. The cross-validation technique can prevent the problem of over-fitting [28].

The classification performance of the classifiers can be measured by calculating the accuracy, sensitivity, and specificity. These performance parameters are defined as shown in Equations (3)-(5).

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (3)$$

$$Specificity = \frac{TN}{TN+FP} \quad (4)$$

$$Sensitivity = \frac{TP}{TP+FN} \quad (5)$$

where TP and TN symbolise the total number of correctly classified PVC beat (true positive) samples and N beat (true negative) samples. The FP and FN symbolise the total number of misclassified PVC beat (false positive) samples and N beat (false negative) samples.

Table III shows the classification performance parameters (accuracy, specificity, and sensitivity) of classifiers using the time series of the signal as an input feature vector. When the time series of the signal was fed to the classifiers, the k-NN classifier achieved the highest accuracy of 99.56%.

Tables IV-VI present a comparison of classification results for classifiers fed to the PCA, ICA and SOM features, respectively. Classification results showed that the k-NN

classifier achieved the highest accuracy for the reduced data by ICA, PCA and SOM. The SOM features achieved less success than the other features.

TABLE III. CLASSIFICATION RESULTS (%) FOR TIME SERIES AS INPUT

	Accuracy	Sensitivity	Specificity
<i>k-NN</i>	99.56	99.29	99.83
<i>NN</i>	98.46	99.06	97.86
<i>SVM</i>	98.09	96.86	99.31
<i>DT</i>	97.96	97.40	98.51

TABLE IV. CLASSIFICATION RESULTS (%) FOR PCS AS INPUT

	Accuracy	Sensitivity	Specificity
<i>k-NN</i>	99.63	99.29	99.89
<i>NN</i>	98.14	98.97	97.31
<i>SVM</i>	98.56	97.43	99.69
<i>DT</i>	95.90	93.23	98.57

TABLE V. CLASSIFICATION RESULTS (%) FOR ICS AS INPUT

	Accuracy	Sensitivity	Specificity
<i>k-NN</i>	99.26	98.80	99.71
<i>NN</i>	97.79	99.06	96.51
<i>SVM</i>	98.04	97.00	99.09
<i>DT</i>	94.73	92.11	97.34

TABLE VI. CLASSIFICATION RESULTS (%) FOR SOM FEATURES AS INPUT

	Accuracy	Sensitivity	Specificity
<i>k-NN</i>	98.36	97.14	99.57
<i>NN</i>	94.60	93.89	95.31
<i>SVM</i>	81.39	69.60	93.17
<i>DT</i>	77.54	55.29	99.80

Fig. 4 shows the average accuracy achieved by the k-NN classifier versus the number of PCs and ICs. The number of PCs varied from 1 to 25 and their effects on classification accuracy were determined. The count of PCs for the k-NN classifier was found as 17. The cumulative variance of the first 17 principal components was 0.995. After beginning with small numbers of PCs, the average accuracy increased rapidly and then levelled off at around 7 PCs. The average accuracy stayed at around 99% at higher PC numbers. Additional increase in PC numbers did not significantly increase the accuracy of the classifier.

The number of ICs from 1 to 25 and their effects on classification accuracy were also examined. The count of ICs for the k-NN classifier was calculated as 10. As is seen in Fig. 4, the average accuracy began with small IC numbers and then increased sharply. On the other hand, there were slight fluctuations in the classification performance of the k-NN classifier at ICs higher than 15. The ICs from 8 to 15 achieved high classification accuracy results.

Parameter optimisation was applied in order to find the optimum k value that gave the best result for the k-NN classifier for the input vector time series, PCs, ICs, and SOM features. The odd numbers from 1 to 15 were tried as a k value. The highest average accuracy of 99.63% was reached at k = 1, but all k values in the range achieved high results (> 98.8%) using the time series and PCs features. Fig. 5 shows the average accuracy versus the k number of the k-NN classifier for input vector time series, ICA, PCA, and SOM features.

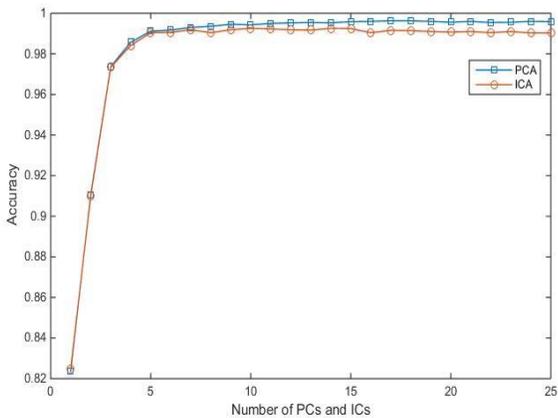


Fig. 4. Average accuracy versus number of PCs and ICs for k-NN classifier

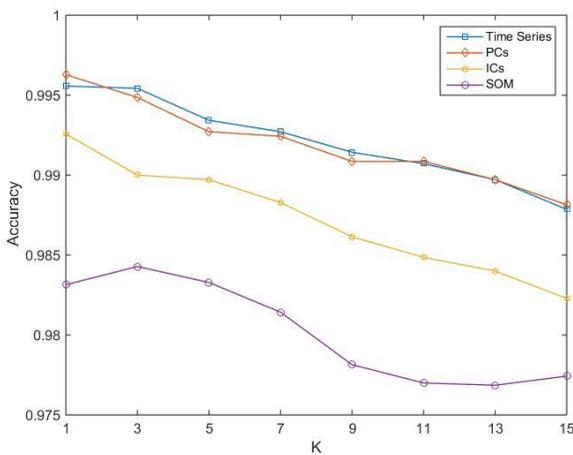


Fig. 5. Average accuracy versus number of k for input vector time series, PCs, ICs, and SOM

As seen in Fig. 5, when the k value increased, the classification performance of the algorithm decreased slightly. The time series and PCs features achieved approximately the same classification accuracy; however, the ICs and SOM features remained slightly lower for all k values. When the SOM features were used as an input vector of the k-NN classifier, the highest classification accuracy was achieved by the k = 3 value.

Table VII shows the classification times (in seconds) for the 10-fold cross-validation of the classifiers. As seen in the results, classifiers that were fed with time series took more time for calculation because the size of the input vector was 200. The k-NN classifier presented an acceptable calculation time for all types of input vectors. The NN classifier took more time than the other classifiers because of its complex computation mechanism. Table VIII shows the results of the proposed method in this work in comparison with results of other methods available in the literature dealing with the classification of PVC. In the proposed method, a k-NN classifier which was fed PCs features was used and an average accuracy of 99.63% for the 10-fold cross-validation was achieved. The proposed approach obtained a higher performance than the existing methods.

TABLE VII. CLASSIFICATION TIMES (S) FOR 10-FOLD CROSS-VALIDATION

	Time Series	PCA	ICA	SOM
<i>k</i> -NN	45	4	3	1
NN	2008	225	16	75
SVM	56	13	17	636
DT	181	21	19	7

TABLE VIII. PERFORMANCE METRICS (SPECIFICITY, SENSITIVITY, ACCURACY), CLASSIFIERS AND DATABASE FILE COUNT USED IN TEST OF PROPOSED METHOD AND PUBLISHED PVC CLASSIFIERS AS REPORTED BY THE AUTHORS

Researchers	Classifiers	Spe.	Sen.	Acc.	File Count
Jie Zhou [5]	Quantum NN	-	-	97.74	11
De Oliveira et al.[6]	Bayesian Network	99.86	95.09	-	-
Ittaturut et al.[7]	Simple decision rule	99.55	91.05	-	26
Bortolan et al.[8]	NN, k-NN, DA, FL	98.7	91.3	-	48
Christov et al. [9]	k-NN	96.7	96.9	-	-
Ebrahimzadeh and Khazaei [10]	MLP NN	-	-	95.4	7
Li et al.[11]	Template Matching	-	93.1	98.2	22
Ribeiro et al. [12]	SVM	98.28	89.39	-	-
Foo et al.[13]	NN	-	-	92.2	4
Inan et al. [14]	NN	-	98.33	95.16	40
Jenny et al. [15]	k-Means, Fuzzy c-Means	80.10	81.10	80.94	-
Proposed method	k-NN, NN, SVM, DT	99.89	99.29	99.63	43

IV. CONCLUSION

In this paper, an approach was proposed to correctly classify PVC beats. At the classification stage, 10-fold cross-validation was used to ensure the reliability of the classification process. Most of the tested classifiers obtained high accuracy rates. In particular, the k-NN classifier achieved the highest accuracy results of 99.63% using PCA features as input vectors. The DT classifier produced the least satisfactory results of all the feature sets. The SVM and DT classifiers using SOM features attained the lowest accuracy rates of 81.39% and 77.54%, respectively. Considering the computation time, the k-NN classifier attained the best results using reduced feature vectors. All of the tested classifiers achieved remarkable acceleration by reducing the size of the feature vectors. However, the computational time of the NN was higher than the others, even when using reduced input feature vectors.

The accuracy, sensitivity, and specificity were calculated in order to compare the training algorithms. In terms of recognition accuracy, it can be seen that the k-NN classification algorithm achieved the best performance according to the experiments.

In comparison with other works, the PVC classification approach presented in this paper showed a higher performance of classification accuracy. Most of the current studies ([5] [7] [10] [11] [13]) have used a specific subset of data in the database. In this study, rather than using a specific subset,

almost all PVC beats existing in the database were used. De Oliveira et al. used 947 PVC beats for classification, with 80% of the data used for training and 20% for testing. However, they did not give sufficient details of their experimental implementation, such as the number of cross validations. Furthermore, the number of records used in the study was not specified [6].

In another work, Ittairut et al. tested their method with 26 records. They excluded some records such as those using pacemakers and those containing heart blockage and atrial fibrillation from their experiments [7].

On the other hand, Bortolan et al. used all the ECG recordings from the MIT-BIH Database. However, the size of the learning set was very small (260 beats for the global set, 76 beats for the local set). The best accuracy achieved was 88.5% from the global set with a DA classifier and 98.7% from the local set with a k-NN classifier [8]. Similarly, Christov et al. used a k-NN algorithm to classify PVC beats in all files in the Database and achieved sensitivity and specificity rates of 96.9% and 96.7%, respectively [9]. Inan et al. used most of the signal files, tested the data with an NN classification algorithm and achieved an accuracy of 95.16% [14]. Jenny et al. used an unsupervised learning algorithm and, therefore, achieved lower accuracy rates than those of the other works [15].

This study showed that high classification accuracy can be obtained without implementing any feature extraction method and by using time series of the signal for input. PCA can be used to reduce the size of the input vectors representing the data. Because of its high computational speed, the proposed method in this work may advance the capability of any system performing real-time PVC analysis. The classification approach presented in this paper can be implemented as part of a computer-aided diagnosis system and can speed up the diagnosis process. The proposed method can be further developed for future use in detecting more ECG arrhythmias.

REFERENCES

- [1] A. S. Go, D. Mozaffarian, V. L. Roger, E. J. Benjamin, J. D. Berry, M. J. Blaha, S. Dai, E. S. Ford, C. S. Fox, S. Franco, H. J. Fullerton, C. Gillespie, S. M. Hailpern, J. A. Heit, V. J. Howard, M. D. Huffman, S. E. Judd, B. M. Kissela, S. J. Kittner, D. T. Lackland, J. H. Lichtman, L. D. Lisabeth, R. H. Mackey, D. J. Magid, G. M. Marcus, A. Marelli, D. B. Matchar, D. K. McGuire, E. R. Mohler, C. S. Moy, M. E. Mussolino, R. W. Neumar, G. Nichol, D. K. Pandey, N. P. Paynter, M. J. Reeves, P. D. Sorlie, J. Stein, A. Towfighi, T. N. Turan, S. S. Virani, N. D. Wong, D. Woo, and M. B. Turner, "Heart disease and stroke statistics-2014 update: a report from the American Heart Association," *Circulation*, vol. 129, no. 3, pp. e28–e292, Jan. 2014.
- [2] N. Tosun, Y. Erkoç, T. Buzgan, B. Keskinliç, D. Aras, N. Yardım, S. Gögen, G. Sarioğlu, and M. Soyulu, *Türkiye Kalp ve Damar Hastalıklarının Önleme ve Kontrol Programı (2010-2014)*, Ankara: Anıl Matbaası, 2014.
- [3] G.D. Clifford, F. Azuaje, and P.E. McSharry (Eds.), *Advanced Methods and Tools for ECG Data Analysis*, Boston: Artech House, 2006.
- [4] G. K. Lee, K. W. Klarich, M. Grogan, and Y.-M. Cha, "Premature ventricular contraction-induced cardiomyopathy: a treatable condition," *Circ. Arrhythm. Electrophysiol.*, vol. 5, no. 1, pp. 229–36, Feb. 2012.
- [5] Jie Zhou, "Automatic detection of premature ventricular contraction using quantum neural networks," in *Third IEEE Symposium on Bioinformatics and Bioengineering*, 2003. Proceedings, pp. 169–173, 2003.
- [6] L. S. C. De Oliveira, R. V. Andreão, and M. Sarcinelli-Filho, "Premature Ventricular beat classification using a dynamic Bayesian Network," *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, vol. 2011, pp. 4984–7, 2011.
- [7] S. Ittairut, A. Lek-Uthai, and A. Teeramongkonrasmee, "Detection of Premature Ventricular Contraction for real-time applications," in *2013 10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, ECTI-CON 2013*, 2013.
- [8] G. Bortolan, I. Jekova, and I. Christov, "Comparison of four methods for premature ventricular contraction and normal beat clustering," *Computers in Cardiology*, vol. 32, pp. 921–924, 2005.
- [9] I. Christov, I. Jekova, and G. Bortolan, "Premature ventricular contraction classification by the K th nearest-neighbours rule," *Physiol. Meas.*, vol. 26, no. 1, pp. 123–130, Feb. 2005.
- [10] A. Ebrahimzadeh and A. Khazaei, "Detection of premature ventricular contractions using MLP neural networks: A comparative study," *Meas. J. Int. Meas. Confed.*, vol. 43, pp. 103–112, 2010.
- [11] P. Li, C. Liu, X. Wang, D. Zheng, Y. Li, and C. Liu, "A low-complexity data-adaptive approach for premature ventricular contraction recognition," *Signal, Image Video Process*, vol. 8, no. 1, pp. 111–120, Apr. 2013.
- [12] B. R. Ribeiro, J. H. Henriques, A. M. Marques, and M. A. Antunes, "Manifold learning for premature ventricular contraction detection," *Computers in Cardiology*, vol. 35, pp. 917–920, 2008.
- [13] S. Y. Foo, G. Stuart, B. Harvey, and A. Meyer-Baese, "Neural network-based EKG pattern recognition," *Eng. Appl. Artif. Intell.*, vol. 15, no. 3–4, pp. 253–260, June 2002.
- [14] O. T. Inan, L. Giovangrandi, and G. T. A. Kovacs, "Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pt. 1, pp. 2507–15, Dec. 2006.
- [15] N. Z. N. Jenny, O. Faust, and W. Yu, "Automated Classification of Normal and Premature Ventricular Contractions in Electrocardiogram Signals," *J. Med. Imaging Heal. Informatics*, vol. 4, no. 6, pp. 886–892, Dec. 2014.
- [16] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, 2001.
- [17] G. Moody and R. Mark, "The MIT-BIH Arrhythmia Database on CD-ROM and software for use with it," in [1990] *Proceedings Computers in Cardiology*, pp. 185–188, 1990.
- [18] M. Kania, M. Fereniec, and R. Maniewski, "Wavelet denoising for multi-lead high resolution ECG signals," *Meas. Sci. Rev.*, vol. 7, section 2, no. 3, pp. 30–33, 2007.
- [19] T. Nodes and N. Gallagher, "Median filters: Some modifications and their properties," *IEEE Trans. Acoust.*, vol. 30, no. 5, pp. 739–746, Oct. 1982.
- [20] D. Kicmerova, "Methods for detection and classification in ECG analysis," *Brno University of Technology*, 2009.
- [21] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 4th Ed. Burlington: Academic Press, 2009.
- [22] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural Netw.*, vol. 13, pp. 411–30, 2000.
- [23] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biol. Cybern.*, vol. 43, no. 1, pp. 59–69, 1982.
- [24] T. Kohonen, "The self-organizing map," *Proc. IEEE*, vol. 78, no. 9, pp. 1464–1480, 1990.
- [25] S. Haykin, *Neural Networks: a comprehensive foundation*, 2nd Ed. Ontario: Pearson, 1999.
- [26] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, Jan. 1967.
- [27] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sept. 1995.
- [28] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd Ed. New York: John Wiley & Sons, Inc., 2000.
- [29] L. Rokach and O. Maimon, *Data Mining with Decision Trees: Theory and Applications*. Singapore: World Scientific Publishing Co., 2008.