AN X-RAY ON METHODS AIMING AT ARRHYTHMIA CLASSIFICATION IN ECG SIGNALS

E. Luz and D. Menotti

Department of Computing - Universidade Federal de Ouro Preto
Campus Universitário, Ouro Preto, Brazil
{eduluz,menottid}@gmail.com

ABSTRACT

Arrhythmias (*i.e.*, irregular cardiac beat) classification in electrocardiogram (ECG) signals consists in an important issue for heart disease diagnosis due to the non-invasive nature of the ECG exam. In this paper, we present an X-ray, a generic view, on methods aiming at arrhythmia classification in ECG signals, which starts with signal preprocessing, and then segmentation of each heartbeat and so before classification, the feature extraction step. We also analyze and criticize the results of some arrhythmia classification methods present in the literature in terms of how the samples are chosen for train/test the classifier and the impact of this choice in their accuracies/sensitivities.

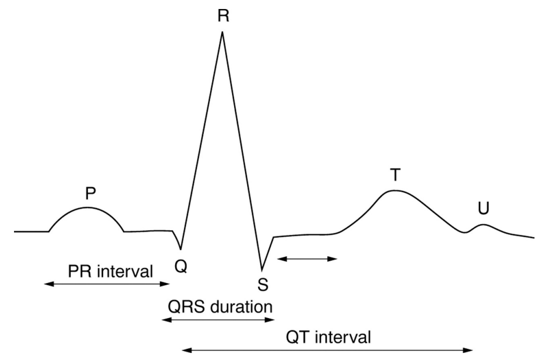

Fig. 1. A normal heartbeat ECG signal

1. INTRODUCTION

The electrocardiogram (ECG) is the most widely used non-invasive technique in heart disease diagnoses. It can be described as a record of the electrical phenomena originated from cardiac activity. Fig. 1 shows a schematic record of a normal heartbeat. The ECG is frequently used to detect cardiac rhythm abnormalities, otherwise known as, arrhythmias. Arrhythmias can be defined in two ways: as a unique irregular cardiac beat or as a set of irregular beats. Arrhythmias can be rare and harmless, but may also result in serious cardiac issues.

There are several methods proposed in the literature for the purpose of automatic arrhythmia classification in ECG signals and a complete system for such an aim can be divided into four subsequent categories (as shown in Fig. 2): preprocessing, segmentation, feature extraction, and classification.

The most widely used database for evaluation of the accuracy/sensitivity/specificity (from now on performance) of arrhythmia classification systems is the MIT-BIH Arrhythmia Database [1]. This database was the first available for such a purpose and it has gone through several improvements over the years to encompass the broadest possible range of waveforms [2]. The Association for the Advancement of Medical Instrumentation (AAMI) also recommends the use of the MIT-BIH Arrhythmia Database for performance evaluation of arrhythmia systems. The AAMI has developed a standard for testing and reporting performance results of algorithms aiming at arrhythmia classification (ANSI/AAMI EC57:1998/(R)2008). According to [3, 4] few researchers have used the AAMI recommendations and standards, leading to clinically unreliable results since several methods in the literature are favored by a biased dataset (*i.e.*, where heartbeats from the same patient are used for both training and testing the classifiers, which makes a fair comparison among methods difficult).

The aiming of this work is twofold. First to summarize recent techniques aiming at arrhythmia classification. And, second, analyzing the results obtained by different designs of automatic classification system using two ways for choosing samples for training/testing the performance of these systems - one following the AAMI recommendations and another one which disregard such recommendations.

The remainder of this work is organized as follows. In Section 2, we briefly describe each category of an arrhythmia classification system presenting the most relevant works, in our point of view, proposed so far. The methods used in our analysis are cited and grouped in Section 3. Finally, discussion of the results reported in those works and conclusions are pointed out in Section 4 and 5, respectively.

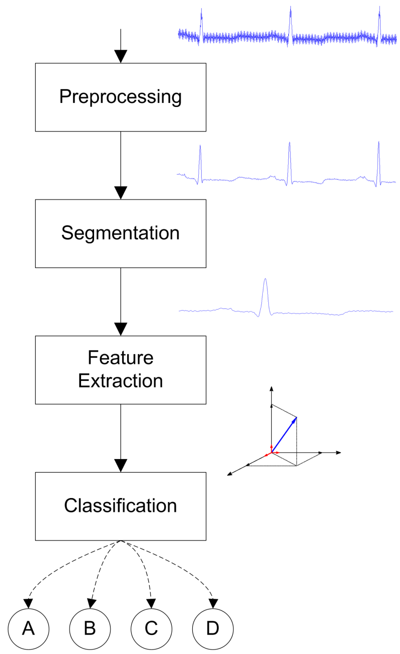

Fig. 2. A diagram of a classification system of arrhythmia

2. ARRHYTHMIA CLASSIFICATION SYSTEM

In this section, we present methods proposed for building a complete system for arrhythmia classification. The system can be divided into four subsequent categories, which starts with signal preprocessing, and then segmentation of each heartbeat and so before classification, the feature extraction step.

2.1. Preprocessing

The preprocessing consists mainly in detecting and attenuating frequencies of the ECG signal related to artifacts. Those artifacts can be from a biological source, like muscular activity, or can be originated from an external source, such as 50/60Hz from electrical network. It is also desired, in the preprocessing, to perform a signal normalization and complex QRS enhancement, in order to help the segmentation process.

Many methods have been proposed to reduce noise in the ECG signal. The most simple and fairly used is the implementation of digital filters [5]. Other architectures, like adaptive filters [6], have also been used to attenuate noise in ECG signal. Most sophisticated methods like adaptive filters based in neural network [7] have brought a significative improvement in noise attenuation process and then raised the effectiveness of segmentation and classification methods.

Statistics techniques, such as principal component analysis (PCA) [8] and independent component analysis (ICA) [9] are also powerful tools for noise attenuation in ECG signal, due to the fact that they allow one extraction of noises represented by frequencies very related or near to the ones of the ECG signal.

Nowadays, methods based in the wavelet transform are widely used. Due to a more accurate filtering process, they preserve the ECG signal, avoiding the loss of important physiological details [10, 11].

Other methods have also presented interesting results. In [12], a non-linear Bayesian filter is proposed to reduce the noise in ECG signal. In [13], a new algorithm based on extended Kalman filter structure incorporates the ECG dynamical model to attenuate noise and data compression of ECG signal. This approach has brought a significative improvement on noise suppression and overcome the most effective methods so far.

2.2. Segmentation

Regarding ECG signals analysis, segmentation consists in delimitating the part of the signal of more interest, the QRS complex, since it reflects the electrical activity of the heart (see Fig. 1). Once the segmentation of QRS complex is done one can obtain many physiological information, such as cardiac frequency, and so the techniques to extract features from the signal can be applied.

Several algorithms have been proposed in the literature for ECG beat segmentation. The problem faced by researches are many, since the ECG beat morphology can be vary for both inter- and intra-patient. A common approach for ECG signal segmentation, *i.e.*, the heartbeat detection, is based on digital filters for preprocessing, linear transformation for R peak enhancement, and adaptive thresholds for heartbeat recognition [14].

QRS detection methods have been proposed over three decades [14, 15, 16] and the evolution of those algorithms reflects the evolution of processing power of computers. Nowadays, more advanced methods are used and the most popular methods are based in neural network [17], genetic algorithm [15], wavelet transform [18], filter banks [19], and support vector machines [16].

2.3. Feature extraction

Feature extraction is the key point for the final classification performance. Features can be extracted directly from ECG wave morphology in time or frequency domain. More sophisticated methods have been used in order to find features less sensitive to noise, such as the autoregressive model coefficients, higher-order cumulant (higher order statistics) [20] and variations of wavelet transform. Researchers claim that wavelet transform is the most promising technique to extract features from the ECG signal [20, 21, 22]. However, in [23], the author argues that methods based on wavelet transform may have some limitations and its use should depend on the application.

Table 1. Mapping the MIT-BIH Arrhythmia types to the AAMI Classes

The AAMI heartbeat class	N	SVEB	VEB	F	Q
Description	Any heartbeat not in the S, V, F, or Q class	Supraventricular ectopic beat	Ventricular ectopic beat	Fusion beat	Unknown beat
MIT-BIH heartbeat types (code)	normal beat (N)	atrial premature beat (A)	premature ventricular contraction (V)	fusion of ventricular and normal beat (F)	paced beat (P)
	left bundle branch block beat (L)	aberrated atrial premature beat (a)	ventricular escape beat (E)		fusion of paced and normal beat (f)
	right bundle branch block beat (R)	nodal (junctional) premature beat (J)			unclassified beat (U)
	atrial escape beat (e)	supraventricular premature beat (S)			
	nodal (junctional) escape beat (j)				

The authors [24] claim that using techniques to reduce the dimension of feature space, such as PCA or linear discriminant analysis (LDA), can offer advantages such as reducing of time and amount of data required for training the classifier. According to them, the usage of techniques for reducing the feature space can worth the loss on accuracy. In [16], for a SVM classifier, the usage of LDA for reducing the feature dimension has shown greater accuracy than the usage of PCA. Moreover, those authors point out that the accuracy of the SVM classifier with reduced feature space using LDA is greater even than the accuracy with the original feature set.

2.4. Classifiers

In order to accurately detect cardiac frequency, it is necessary to consider sporadic arrhythmias occur. An accurate arrhythmia classification is also desirable to correctly diagnose cardiac issues and in some cases, the early detection can save lives. With that motivation in mind, researches keep the efforts to develop better and better methods.

Artificial neural networks (ANN) are widely used to arrhythmias classification in ECG signals [25], and the multilayer perception (MLP), the most popular ANN, is often used for that purpose [20].

The conventional MLP has shown high accuracy in classification of arrhythmias. Nevertheless it suffers from slow local convergence, global minimum localization and random initial weights. These drawbacks could make it inappropriate to clinical usage [24]. To overcome this issues, hybrid systems, combining MLP with another ANN are normally indispensable [26]. In those kind of systems, the first level of networks are responsible to initially classify the heartbeats and also build models generating new feature inputs. The MLP completes the second task of multi-classification [27]. With that approach, many weakness of MLP are surmounted.

In [27] and [20], a method based on higher order statistics to extract features, and a hybrid neuro-fuzzy method for classification [28], which uses type-2 fuzzy c-means algorithm to improve the accuracy of the neural network, have reached higher accuracies than conventional MLP methods.

SVM has also been widely used to classify arrhythmias. In [29], a comparison of different methods using SVM and ANN has shown that SVM methods should be choose when training time matters. Otherwise ANN methods have demonstrated better results. In [16], the authors have used linear discriminant analysis (LDA) in order to reduce the size of the feature space, and despite that fact a high accuracy has been shown.

In [30], it is proposed a method with fast learning rates and high accuracy (% 98.72), using morphology filtering, principal component analysis (PCA) and extreme learning machine (ELM). The algorithm is used to detect six types of arrhythmias and the results have shown that the method is faster than others like MLP and SVM.

3. METHODS

We chose eight methods to analyze their performances. Three of them, in our consideration, are state-of-the-art methods, since its authors have followed the AAMI recommendations [3, 4, 31]. In the remaining five methods, the authors did not follow the AAMI recommendations [32, 33, 25, 34, 35]. However they report performance in average near to 100% as shown in Table 2.

The MIT-BIH arrhythmia database contain 48 half-hour records, sampled at 360Hz, and eighteen types of heartbeats were classified and labeled. To comply with the AAMI recommendations, only 44 records of MIT-BIH arrhythmia

Table 2. Classification performance of methods using random selection of samples (heartbeats) - biased selection

Method	Accuracy	Sensitivities (%)														
		N	L	R	A	V	P	a	!	F	x	j	f	E	J	e
Ye *et al.* [32]	99.91	99.95	100	99.99	99.65	99.26	100	92.86	100	99.73	100	100	100	100	97.06	100
Yu & Chen [25]	99.65	99.97	99.33	99.54	99.76	99.04	100	-	-	-	-	-	-	-	-	-
Yu & Chou [33]	98.71	99.65	96.25	99.15	98.40	98.45	99.37	-	90.12	-	-	-	-	91.53	-	-
Korürek & Nizam [34]	-	95.49	-	97.56	86.78	93.33	-	-	-	74.51	-	-	84.06	-	-	-
Tsipouras *et al.* [35]	96.43	93.89	-	98.65	-	91.35	-	-	97.74	-	-	-	-	-	-	-

database should be used for evaluation of arrhythmia classification methods, excluding the 4 records that contain paced beats. The ANSI/AAMI EC57:1998/(R)2008 standards recommends to group those heartbeats into five classes: 1) normal beat; 2) ventricular ectopic beat (VEB); 3) supraventricular ectopic beat (SVEB); 4) fusion of a VEB and a normal beat; and 5) unknown beat type (see Table 1). Moreover, the AAMI standards also recommends to divide the recordings into two datasets, one for training and another for testing, such that heartbeats from one recording (patient) are not used simultaneously for both training and testing the classifier.

The methods which do not follow the AAMI standards for building the arrhythmia classifiers create randomly their datasets for training and testing, in such a manner that unavoidably heartbeats from one recording are present in both sets. This practice, *i.e.*, to put data from the same patient in both sets, should be avoided as already stated in [3].

There is also a lack of standard regarding classes of heartbeats to be analyzed. In some cases, the classifiers are design to classify a specific number of classes, e.g., 2, 3, 10. In other cases, the authors present the performance of methods for non standard classes (*i.e.*, non beat annotation codes), such as Ventricular Flutter Wave (!) and Non-Conducted P-wave (x) [32, 33].

4. DISCUSSIONS

In order to analyze the classification performance, two measures are used, *i.e.* accuracy and sensitivity. Accuracy is defined as the ratio of total beats correctly classified and the number of total beats, *i.e.*,

$$Accuracy = \frac{beats\ correctly\ classified}{number\ of\ total\ beats}. \quad (1)$$

Sensitivity stands for the ratio of correctly classified beats of one class and the total beats classified as that class, including the miss classification beats, *i.e.*,

$$Sensitivity = \frac{true\ negatives}{true\ negatives\ +\ false\ positives}. \quad (2)$$

Sensitivity[1] is the most important measure for our analysis, since the number of heartbeats for each class in the MIT-BIH arrhythmia database is very imbalanced and a single class (*e.g.*, the normal beats) could represent most of the total accuracy, while the sensitivity and specificity directly depend on the number of samples for each class.

[1]together with specificity, which is not used in this study due to the lack of these data in the studied works.

Table 3. Classification performance of methods following the AAMI recommendations

Method	Accuracy	Sensitivities (%)				
		N	SVEB	VEB	F	Q
Chazal *et al.* [3]	85.9	86.86	75.93	77.73	89.43	0
Ince *et al.* [4]	93.6	97.04	62.11	88.39	61.36	0
Jiang & Kong [31]	94.5	98.73	50.58	86.61	35.78	0

Table 4. Classification performance of methods which do not follow the AAMI recommendations. The classes and method presented in Table 2 are grouped according to Table 1, to comply with the AAMI recommended classes

Method	Accuracy	Sensitivities (%)				
		N	SVEB	VEB	F	Q
Ye *et al.* [32]	99.91	99.96	98.48	99.83	99.21	99.96
Yu & Chen [25]	99.65	99.67	99.53	99.22	-	100
Yu & Chou [33]	98.71	99.81	98.50	97.74	-	100
Korürek & Nizam [34]	-	95.51	86.78	-	74.51	84.06
Tsipouras *et al.* [35]	96.43	93.90	-	91.35	-	-

Comparing the results achieved by methods using the AAMI recommendations for designing the arrhythmia classification systems and the ones which do not follow them (Tables 3 and 4), respectively, we can observe a significant difference in terms of the sensitivities reported. This remark can be extended to the accuracy figures.

For both measures, the methods which do not follow the AAMI recommendations present higher values. It is noticeable that all methods analyzed in this work are consistent and use advanced techniques to solve the arrhythmia classification problem. Thus, we suggest that this significant difference in the performances are mostly related to datasets used for training and testing the classifiers. The use of a dataset for training a classifier and then testing it with samples (heartbeats) from the same patients helps the classifier to yield better classification results, since it is specialized in those data.

Besides the fact that heartbeats from same recording, used both for training and testing, can favor the classifier, there is another practice that can lead to biased conclusions as well. Several methods do not use the complete data from the MIT-BIH arrhythmia database as done in [25] and [33], where only 23200 and 9800 heartbeats are used, respectively. In those approaches, the heartbeats were randomly chosen and the classifiers can be favored by eventually easily heartbeat patterns.

Moreover, according to [4], only a few of the methods presented in the literature have, in fact, used the AAMI standards. This statement suggest that the results of several methods in literature are unreliable and should not be taken into account clinically before a robust performance test can be performed.

5. CONCLUSIONS

In this work, we presented an X-ray, a generic view, on methods aiming at arrhythmia classification in ECG signals. Moreover, we showed that the challenges to properly classify arrhythmias in ECG signal are many.

Researchers have been working on improvements, and many of them have shown remarkable results. Nonetheless, few authors have considered the impact on the performance of the classifiers caused by the way the samples (heartbeats) were selected for building the dataset used for training and testing the classifiers. This work have cited methods that may use heartbeats from same patients for training and testing a classifier which could favor their results in terms of performance. However, those reported performances are not realistic, since those methods will classify "never seen" heartbeats (*e.g.*, a new patient), and in these situations, the performance obtained by the method can be quite small.

Thus, the choice of unbiased dataset, such as recommended by the AAMI, should be used for arrhythmia classification methods in order to obtain more reliable results. Having this fact in mind, several methods in the literature can be re-run using unbiased datasets. These results should be used for report new prediction values for these methods, establishing a new state-of-the-art method in terms of performance.

6. REFERENCES

[1] Massachusetts Institute of Technology, "MIT-BIH ECG database," available at http://ecg.mit.edu/.

[2] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.

[3] P. Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 7, pp. 1196–1206, 2004.

[4] T. Ince, S. Kiranyaz, and M. Gabbouj, "A generic and robust system for automated patient-specific classification of ECG signals," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 5, pp. 1415–1427, 2009.

[5] P. Lynn, "Recursive digital filters for biological signals," *Medical and Biological Engineering and Computing*, vol. 9, no. 1, pp. 37–43, 1979.

[6] M. Yelderman, B. Widrow, J. M. Cioffi, E. Hesler, and J. A. Leddy, "ECG enhancement by adaptive cancellation of electrosurgical interference," *IEEE Transactions on Biomedical Engineering*, vol. 30, no. 7, pp. 392 –398, 1983.

[7] Q. Xue, Y. H. Hu, and W. J. Tompkins, "Neural-network-based adaptive matched filtering for QRS detection," *IEEE Transactions on Biomedical Engineering*, vol. 39, no. 4, pp. 317–329, 1992.

[8] F. Castells, P. Laguna, L. Sörnmo, A. Bollmann, and J. M. Roig, "Principal component analysis in ECG signal processing," *'EURASIP Journal on Applied Signal Processing*, vol. 2007, no. 1, pp. 98–98, 2007.

[9] T. He, G. Clifford, and L. Tarassenko, "Application of independent component analysis in removing artefacts from the electrocardiogram," *Neural Computing and Applications*, vol. 15, no. 2, pp. 105–116, 2006.

[10] B. N. Singh and A. K. Tiwari, "Optimal selection of wavelet basis function applied to ECG signal denoising," *Digital Signal Processing*, vol. 16, no. 3, pp. 275–287, 2006.

[11] O. Sayadi and M. B. Shamsollahi, "Multiadaptive bionic wavelet transform: Application to ECG denoising and baseline wandering reduction," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, no. 14, pp. 1–11, 2007.

[12] R. Sameni, M. B. Shamsollahi, C. Jutten, and G. D. Clifford, "A nonlinear Bayesian filtering framework for ECG denoising," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 12, pp. 2172–2185, 2007.

[13] O. Sayadi and M. B. Shamsollahi, "ECG denoising and compression using a modified extended kalman filter structure," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 9, pp. 2240–2248, 2008.

[14] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Transactions on Biomedical Engineering*, vol. 32, no. 3, pp. 230–236, 1985.

[15] R. Poli, S. Cagnoni, and G. Valli, "Genetic design of optimum linear and nonlinear QRS detectors," *IEEE Transactions on Biomedical Engineering*, vol. 42, no. 11, pp. 1137–1141, 1995.

[16] M. H. Song, J. Lee, S. P. Cho, K. J. Lee, and S. K. Yoo, “Support vector machine based arrhythmia classification using reduced features,” *International Journal of Control, Automation, and Systems*, vol. 3, no. 4, pp. 509–654, 2005.

[17] I. Güler and E. D. Übeyli, “ECG beat classifier designed by combined neural network model,” *Pattern Recognition*, vol. 38, no. 2, pp. 199–208, 2005.

[18] Y. Jung and W. J. Tompkins, “Detecting and classifying life-threatening ECG ventricular arrythmias using wavelet decomposition,” in *IEEE International Conference on Engineering in Medicine and Biology Society*, 2003, vol. 3, pp. 2390–2393.

[19] V. X. Afonso, W. J. Tompkins, T. Q. Nguyen, and S. Luo, “ECG beat detection using filter banks,” *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 2, pp. 192–202, 1999.

[20] E. Mehmet, “ECG beat classification using neuro-fuzzy network,” *Pattern Recognition Letters*, vol. 25, no. 15, pp. 1715–1722, 2004.

[21] I. Güler and E. D. Übeyli, “ECG beat classifier designed by combined neural network model,” *Pattern Recognition*, vol. 38, no. 2, pp. 199–208, 2005.

[22] C. Lin, Y. Du, and T. Chen, “Adaptive wavelet network for multiple cardiac arrhythmias recognition,” *Expert Systems with Applications*, vol. 34, no. 4, pp. 2601–2611, 2008.

[23] Y. Özbay, “A new approach to detection of ECG arrhythmias: Complex discrete wavelet transform based complex valued artificial neural network,” *Journal of Medical Systems*, vol. 33, no. 6, pp. 435–445, 2009.

[24] R. Ceylanand and Y. Özbay, “Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network,” *Expert Systems with Applications*, vol. 33, no. 2, pp. 286–295, 2007.

[25] S. Yu and Y. Chen, “Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network,” *Pattern Recognition Letters*, vol. 28, no. 10, pp. 1142–1150, 2007.

[26] M. H. Fredric and H. Soowhan, “Classification of cardiac arrhythmias using fuzzy ARTMAP,” *IEEE Transactions on Biomedical Engineering*, vol. 43, no. 4, pp. 425–429, 2002.

[27] S. Osowski and T. H. Linh, “ECG beat recognition using fuzzy hybrid neural network,” *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 11, pp. 1265–1271, 2001.

[28] R. Ceylan, Y. Özbay, and B. Karlik, “A novel approach for classification of ECG arrhythmias: Type-2 fuzzy clustering neural network,” *Expert Systems with Applications*, vol. 36, no. 3, pp. 6721–6726, 2009.

[29] M. Moavenian and H. Khorrami, “A qualitative comparison of artificial neural networks and support vector machines in ECG arrhythmias classification,” *Expert Systems with Applications*, vol. 37, no. 4, pp. 3088–3093, 2010.

[30] J. Kim, H. S. Shin, J. Shin, and M. Lee, “Robust algorithm for arrhythmia classification in ECG using extreme learning machine,” *BioMedical Engineering OnLine*, vol. 8, no. 1, pp. 1–12, 2009.

[31] W. Jiang and G. S. Kong, “Block-based neural networks for personalized ECG signal classification,” *IEEE Transactions on Neural Networks*, vol. 18, no. 6, pp. 750–1761, 2007.

[32] C. Ye, M. T. Coimbra, and B. V. K. V. Kumar, “Arrhythmia detection and classification using morphological and dynamic features of ECG signals,” in *IEEE International Conference on Engineering in Medicine and Biology Society (EMBC)*, 2010, pp. 1918–1921.

[33] S. Yu and K. Chou, “Integration of independent component analysis and neural networks for ECG beat classification,” *Expert Systems with Applications*, vol. 34, no. 4, pp. 2841–2846, 2008.

[34] M. Korürek and A. Nizam, “A new arrhythmia clustering technique based on ant colony optimization,” *Journal of Biomedical Informatics*, vol. 41, pp. 874–881, 2008.

[35] M. G. Tsipouras, C. Voglis, and D. I. Fotiadis, “A framework for fuzzy expert system creation-application to cardiovascular diseases,” *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 11, pp. 2089–2105, 2007.