Bio-Medical Materials and Engineering 26 (2015) S1045–S1052
DOI 10.3233/BME-151400
IOS Press

A novel approach for arrhythmia diagnosis: Self-adaptive and distribution-free mode

Fenghuan Li, Dequan Zheng * and Tiejun Zhao
MOE-MS Key Laboratory of Natural Language Processing and Speech, Harbin Institute of Technology, 150001, Harbin, PR China

Abstract. Arrhythmia diagnosis is very significant to ensure human health. In this paper, a new model is developed for arrhythmia diagnosis. A salient feature of the algorithm is a synergistic combination of statistical and fuzzy set-based techniques. It is distribution-free and is realized in an unsupervised mode. Arrhythmia diagnosis is viewed as a certain statistical hypothesis testing. 'Abnormal' is typically a much complex concept, so it can be described with the technology of fuzzy sets which bring a facet of robustness to the overall scheme and play an important role in the successive step of hypothesis testing. Intensive fuzzification is engaged in parameters determination which is self-adaptive and no parameter needs to be specified by the user. The algorithm is validated with a number of experiments, which prove its effectiveness for arrhythmia diagnosis.

Keywords: Arrhythmia diagnosis, fuzzy sets, statistical testing, self-adaptive, distribution-free

1. Introduction

The electrocardiogram (ECG) provides great information on normal and pathological physiology. Main problems in arrhythmias detection are that one disease may have dissimilar signs for different patients and two diseases may have similar effects on ECG signals. Various methods were developed, including classification[1], nearest neighbor-based techniques[2], clustering[3], statistical techniques[4], et al. However, classification[5] cannot detect unknown ECG signals. Although many clustering methods[6] show good performance in the lab, there are few techniques gaining popularity in practical applications. There are three main reasons, including sensitivity to initialization, trapping into local minima and lack of prior knowledge. For statistical methods[7], statistical features are taken into account. They fit a statistical model to given data, then apply a statistical inference test to determine whether an instance belongs to this model. The main drawbacks are statistical features are unknown and various, what's more, they may not coincide with the specified value of the process characteristic and several parameters need to be determined[8]. However, they can detect unseen patterns. The nearest neighbor-based technique is a nonparametric method[9]. Despite its simplicity, its critical shortcomings are the low tolerance to noise

*Address for correspondence: Dequan Zheng, Machine Intelligence & Translation Lab, Harbin Institute of Technology, 150001, Harbin, China. Tel.: +86-0451-86412449-606; Fax: +86-0451-86416225; E-mail: dqzheng2007@gmail.com.

and the fact that it relies on existing data, assuming that training set defines decision boundaries among the classes perfectly, which is not always the case[10]. The fact that 'normal' and 'abnormal' are fuzzy makes fuzzy logic theory used successfully[11,12]. Fuzzy logic theory can avoid hard threshold, thereby increasing the tolerance towards contradictions in the data. However, fuzzy rules and membership functions are difficult to determine. If information is insufficient, the performance will become worse.

According to characteristics of ECG signals and above-mentioned methods, a combination of statistical and fuzzy set-based technique is proposed. Arrhythmia diagnosis is viewed as a statistical test. Because ECG signals have non-stationary statistical attributes, mean of samples' characteristics isn't adopted, but a threshold. 'Abnormal' includes fuzziness, so fuzzy set theory is taken into account. What's more, predicates in statistical test are treated as fuzzy predicates (the degree of inclusion). Intensive fuzzification is adopted to realize parameters determination which is self-adaptive. So parameters are not required to be predefined. Due to fuzzy set theory, statistical test is distribution-free and totally unsupervised.

2. Material and preprocessing

In this work, MIT/BIH arrhythmia database[13] is used for arrhythmia diagnosis, which includes 48 records. The length of each record is around 650,000 samples. Each record is an annotation file, in which each beat has been identified by several expert cardiologist annotators. These labels are referred to as annotations of ground truth and can be used to do the training and evaluate the performance.

All ECG records are first segmented into beat cycles. More than 100,000 cycles are segmented with all samples. Each cycle and the position of QRS complex can be located. QRS complex is aligned at the central point of each cycle. The cycle time and duration from P to T wave for a normal cycle is around 0.5 and 0.8s, respectively. Sampling rate of an ECG signal is 360 samples/s[14]. Since most features are located in PQRST waves, 200 samples are taken, which correspond to 0.56s, for each cycle. This length can cover the range for PQRST waves. After detection of QRS complex, 99 samples were chosen from the left side of QRS mid-point and 100 samples after QRS mid-point and the QRS mid-point itself as a beat of 200 samples[15]. After preprocessing, an ECG beat complex is represented by a sequence of 200 discrete samples. 75023 Normal and 25002 abnormal (7256 Right Bundle Branch Block (RBBB), 8071 Left Bundle Branch Block (LBBB), 2546 Atrial Premature Contraction (APC) and 7129 Ventricular Premature Contraction (VPC)) beats are extracted. A brief description of used beats is explained in [16].

3. Arrhythmia diagnosis as a statistical test

Arrhythmia is abnormal cardiac activity. So, arrhythmia diagnosis is a kind of anomaly detection. A hypothesis testing is to discover whether the data is significantly different from the present in the population. The testing is evaluated by fundamental tradeoff between false positive and false negative rates. Hypothesis testing explains how to pick thresholds decision when one problem is faced with balancing this tradeoff. In fact, any anomaly detection method requires a threshold. So hypothesis testing is available to discover whether current data is significantly different from the whole information. Any anomaly detection method at some point uses a statistical test to verify whether a hypothesis is true.

3.1. Anomaly score

Considering a dataset, P is an instance(beat) with q samples. According to section 2, q is 200 in this study. A key component of the algorithm is the distance used to determine how closely two instances

are matched. For each instance, the degree to which the instance differs from others is defined, termed anomaly score. It is defined as the average of distances between an instance and its K nearest neighbors. Formally speaking, anomaly score of an instance P is computed by Eq.(1), where $P_t \in KNNSet\,(P)$, which denotes the set of K nearest neighbors of P. $Distance\,(P, P_t)$ is measured by Euclidean distance.

$$Score\,(P) = \frac{1}{K}\sum_{t=1}^{K} Distance\,(P, P_t) \tag{1}$$

3.2. Statistical hypothesis testing

Traditional hypothesis testing relies heavily on prior knowledge of data distribution and is not suitable for unknown data. What's more, hypothesis testing is based on statistical estimation of distribution parameters. Because the distribution is unknown, it is impossible to estimate its parameters. In practice, the process mean may not coincide with the specified value of the feature. So threshold θ is considered. Following hypothesis $H0$ for the case when there is an anomaly, and an alternate hypothesis $H1$ for the case when there is no anomaly are formed.

$$\begin{aligned} H_0 &: anomaly\ score > \theta \\ H_1 &: anomaly\ score \leq \theta \end{aligned} \tag{2}$$

Neither the distribution nor the values of its parameters is known, such as mean and standard deviation. But according to the definition of anomaly score, the average of distances between an instance and its K nearest neighbors can be determined. Meanwhile, standard deviation of this instance can be calculated. *T-test* is a hypothesis test which is commonly applied when the test statistic follows the normal distribution. When the scaling term is unknown and is replaced by an estimate, the test statistic follows a Student's *t-test*, which doesn't need known distribution parameters. The instance mean and standard deviation are used to replace distribution parameters. When considered in the problem, the test statistic $t(P)$ and rejection region come in Eq.(3). Where $Score\,(P)$ is the anomaly score of instance P, s is the standard deviation between an instance and its K nearest neighbors, and α is significance level.

$$t\,(P) = \frac{Score\,(P) - \theta}{s/\sqrt{K}} \leq t_{1-\alpha}(K-1) \tag{3}$$

There are two parameters to be determined, which are θ and K. 'Abnormal' includes fuzziness. For example, if anomaly score is equal to $\theta - \varepsilon$ (ε is an infinitesimal positive number), this instance is normal. But if anomaly score is equal to $\theta + \varepsilon$, it is difficult to determine whether it is normal. In this sense, fuzzy sets arise naturally. Correspondingly, inequalities in statistical test are treated as fuzzy predicates. θ determination and the reason why there is no need to care data distribution will be illustrated in next section. The algorithm is unsupervised, but training can be used as decision support. Optimal results can be derived by optimizing K whose value is estimated based on training. In order to determine optimum K, a simple way is to alter K from 1 to a large enough value. Based on diagnosis results with each K, performance measures are computed and the K for which best performance is obtained is chosen. In this way, optimum K is determined. Furthermore, only K optimization needs training. In this paper, a predefined K which assumes any value can be considered. In this case, the algorithm is unsupervised.

4. Self-adaptive and distribution-free mode

As mentioned earlier, threshold determination can be realized by engaging fuzzy sets. In this section, threshold determination and a certain intensive fuzzification process that leads to inequalities as fuzzy predicates are shown. The aim of fuzzification is to achieve optimized parameters in statistical test. Because of fuzzification, the algorithm is a self-adaptive and distribution-free statistical test.

4.1. Fuzzification

As the aim is to realize arrhythmia diagnosis, two fuzzy sets are selected, namely 'normal' and 'abnormal'. Because of their smoothness, infinite support, Gaussian membership functions are adopted, which are described in Eq.(4). Each membership function is characterized by two important parameters. a (or b) represents function centre, whereas σ denotes the spread. Higher values of σ correspond to larger spreads of fuzzy sets. Same σ is selected for two functions. In this case, σ can be viewed as a constant. So it is a or b, not σ, which determines an overlap between these two functions. When $abnormal\,(x) > normal\,(x)$, there is an anomaly. The threshold is given by the condition $abnormal\,(x) = normal\,(x)$.

$$normal\,(x) = \begin{cases} 1, & x \le a \\ exp(-\left(\frac{x-a}{\sigma}\right)^2), & x > a \end{cases}, \qquad abnormal\,(x) = \begin{cases} exp(-\left(\frac{x-b}{\sigma}\right)^2), & x \le b \\ 1, & x > b \end{cases} \tag{4}$$

Now, how to determine parameters a and b is illustrated. In principle, a sound membership function not only can reflect fuzzy characteristics of a fuzzy concept, but also can describe the objective content as clear as possible. The clarity can be measured by fuzziness degree of one fuzzy set. When fuzziness degree is smaller, fuzzy set expresses the problem more clearly. So the principle of minimum fuzziness degree is used to determine related parameters. Fuzzy entropy is selected to express the fuzziness degree, leading to the optimization model in Eqs.(5)and (6), where N is the number of samples in the dataset.

$$\min H\,(a,b) = \frac{1}{N \ln 2} \sum_{i=1}^{N} \{s\,(normal\,(x_i)) + s\,(abnormal\,(x_i))\} \tag{5}$$

$$s(x) = \begin{cases} -x \ln x - (1-x)\ln(1-x), & x \in (0,1) \\ 0, & x = 0, 1 \end{cases} \tag{6}$$

4.2. Intensive fuzzification

The fuzzy set-based scheme which is an intensive fuzzification process is proposed in Figure 1. Solid line represents the first fuzzification in which anomaly score is the data object. The aim of the first fuzzification is to obtain an optimized threshold. Dashed line shows the second fuzzification, in which $t(P)$ is the data object. Because optimized threshold is a numeric value and inequality ($\le$) is a fuzzy predicate, a second fuzzification is conducted to determine the degree of inclusion. The scheme is developed in two stages: training and testing, which are almost the same. Both of them are intensive fuzzification scheme. Their only difference is K optimization. Training is needed only when an optimized K is required. Then, optimized K is used at testing stage. Otherwise, training and K optimization are not necessary. K will be predefined by users when testing.The scheme consists of five major components as follows:

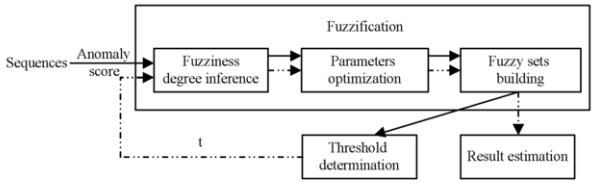

Fig. 1. Intensive fuzzification detection scheme.

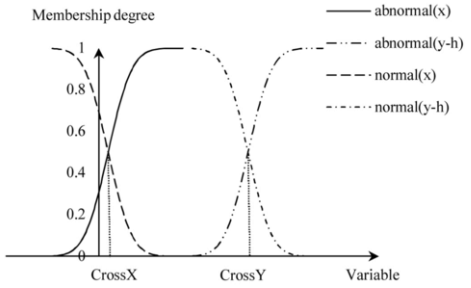

Fig. 2. Membership functions with different variables.

1. Fuzziness degree inference: Fuzziness degree is expressed by membership functions in Eq.(4).
2. Parameters optimization: a and b are optimized by the minimum of fuzziness degree in Eq.(5).
3. Fuzzy sets building: Membership functions are constructed by optimized a, b and predefined σ.
4. Threshold determination: θ in Eq.(3) exists when two membership degrees are the same.
5. Result estimation: P is an anomaly and $H0$ is accepted when $abnormal(t(P)) > normal(t(P))$, and K optimization is carried out.

4.3. Distribution-free statistical test

Now, the reason why there is no need to care the data distribution is illustrated. Only in the second fuzzification, data distribution need be considered. As shown in Eq.(3), $t(P)$ and $t_{1-\alpha}(K-1)$ have to be calculated. If K and α are known, $t_{1-\alpha}(K-1)$ will be a constant. Predicted result of $t(P)$ will be the same as that of $t(P) - t_{1-\alpha}(K-1)$. For example, x and y are two variables, and $y = x + h$, h is a constant. If membership functions of x are $normal(x)$ and $abnormal(x)$, membership functions of y are $normal(y-h)$ and $abnormal(y-h)$. $abnormal(y-h)$ and $normal(y-h)$ are on the right of $abnormal(x)$ and $normal(x)$ as shown in Figure 2. Their distance is h and shapes are the same. $CrossX$ and $CrossY$ are their crossing points and the distance is h. It means $CrossY = CrossX + h$. So the hypothesis $x < CrossX$ is equivalent to the hypothesis $y < CrossY$. Additionally, if their membership functions can be determined respectively, their crossing points can be determined without caring the value of h. Therefore, $t_{1-\alpha}(K-1)$ in Eq.(3) is not considered, and only $t(P)$ is regarded as the test object. So, there is no need to care whether the data follows a normal distribution or not.

5. Experiments and discussions

5.1. Performance measures

Four measures are used to assess results, which are sensitivity(Sen), specificity(Spe), positive predictive value(PPV) and accuracy(Acc). They are defined as follows: $Sen = TP/(TP+FN)$, $Spe = TN/(FP+TN)$, $PPV = TP/(TP+FP)$, $Acc = (TP+TN)/(TP+FN+FP+TN)$. True positive (TP) is the number of true arrhythmias successfully detected; false positive(FP) is the number of true arrhythmias that are missed; true negative (TN) is the number of non-target arrhythmias correctly detected; false negative(FN) is the count of non-target arrhythmias detected wrongly.

To determine optimum K corresponding to the best accuracy, K is altered from 1 to a large enough value(in this study, $Kmax$ = 7000). In this process, ten-fold cross validation method is used. The entire

Table 1
Comparison for several techniques from two points of view

Measures	Algorithm	APC	LBBB	VPC	RBBB
PPV	SP(1630)	0.10	0.31	0.45	0.46
	SO	**0.12**	**0.32**	**0.51**	**0.60**
	SFP(1630)	0.38	0.40	0.88	0.89
	SFO	**0.55**	**0.77**	**0.88**	**0.92**
Sen	SP(1630)	0.83	1.00	1.00	1.00
	SO	**0.86**	**1.00**	**1.00**	**1.00**
	SFP(1630)	0.55	1.00	0.84	1.00
	SFO	0.50	0.95	0.88	0.98
Spe	SP(1630)	0.74	0.76	0.89	0.89
	SO	**0.78**	**0.77**	**0.91**	**0.93**
	SFP(1630)	0.97	0.84	0.99	0.99
	SFO	**0.99**	**0.97**	**0.99**	**0.99**
Acc	SP(1630)	0.77	0.80	0.90	0.90
	SO	**0.80**	**0.82**	**0.92**	**0.94**
	SFP(1630)	0.96	0.86	0.98	0.99
	SFO	**0.97**	**0.97**	**0.98**	**0.99**

Table 2
Comparison with four competing algorithms

Measures	Algorithm	APC	LBBB	VPC	RBBB
PPV	SFO	**0.55**	**0.77**	**0.88**	**0.92**
	K-distance	0.43	0.40	**0.88**	0.89
	LOF	0.42	0.40	0.86	0.89
	Random Walk	0.31	0.40	**0.88**	0.53
	Brute Force	0.41	0.18	0.75	0.57
Sen	SFO	0.50	0.95	**0.88**	0.98
	K-distance	**0.64**	**1.00**	0.84	**1.00**
	LOF	0.62	**1.00**	0.86	**1.00**
	Random Walk	0.46	**1.00**	0.84	0.57
	Brute Force	0.60	0.44	0.71	0.63
Spe	SFO	**0.99**	**0.97**	**0.99**	**0.99**
	K-distance	0.97	0.84	**0.99**	**0.99**
	LOF	0.97	0.84	**0.99**	**0.99**
	Random Walk	0.97	0.84	**0.99**	0.95
	Brute Force	0.97	0.78	0.98	0.95
Acc	SFO	**0.97**	**0.97**	**0.98**	**0.99**
	K-distance	0.96	0.86	**0.98**	**0.99**
	LOF	0.96	0.86	**0.98**	**0.99**
	Random Walk	0.95	0.86	**0.98**	0.92
	Brute Force	0.96	0.75	0.96	0.93

training set is divided into ten disjoint subsets of almost equal size. In every fold, one subset is used for the training and the rest nine subsets are used for the test. The process is repeated ten times so that every subset will be tested in different folds. Based on correct diagnosis results with each K, performance measures are computed in each fold and K for which the best accuracy is obtained in all folds is chosen.

5.2. Performance analysis

Some techniques in this algorithm are analyzed. Because every arrhythmia has its own characteristics, a predefined and same K isn't used. Of course, if there is limited training data, a predefined and same K can be used for all arrhythmias. In this case, the algorithm will be a totally unsupervised model. According to characteristics of the algorithm, four experiments are performed as follows:

1. The algorithm based on statistical testing and a predefined K is called SP.
2. The algorithm based on statistical testing and an optimized K is called SO.
3. The algorithm based on statistical testing, fuzzy set theory and a predefined K is called SFP.
4. The algorithm based on statistical testing, fuzzy set theory and an optimized K is called SFO.

Because the aim is arrhythmia diagnosis, results of four arrhythmias are just listed, which are RBBB, LBBB, APC and VPC. The performance is analyzed from two points of view: K selection and fuzzification. In the first (SP) and third (SFP) experiments, a fixed $K(1630)$ is selected for the experiments. In other two experiments, optimized K is used, which may be different for different arrhythmias. All results are listed in Table 1. If K optimization is needed, results when the best accuracy is obtained are listed, otherwise, experiments are performed in every subset and average results are listed for every arrhythmia.

K optimization is to obtain best accuracy. When K is different, the performance may be different. This conclusion can be come up with from performance comparison of SO with SP(1630) and comparison of SFO with SFP(1630) in Table 1. Taking comparison of SO with SP(1630) for example, better

results are marked in bold. All results of SO are better than those of SP(1630). All positive predictive value, specificity and accuracy are improved significantly. Values of sensitivity are unchanged or slightly better. Similar conclusions occur in comparison of SFO with SFP(1630). SFO achieves better positive predictive value, specificity and accuracy. Values of sensitivity are unchanged or slightly worse. It means K optimization is necessary. If there is no training data, K can be any value. In this case, sensitivity is not significantly affected by different K. That means that the performance is improved greatly.

The analysis of fuzzification is shown by comparing SFP(1630) with SP(1630) and comparing SFO with SO in Table 1. Taking comparison of SFO with SO for example, better results with fuzzification are marked in bold. All positive predictive value, specificity and accuracy are better than those of SO. All sensitivity values are smaller or stay the same. Similar conclusions occur in the comparison of SFP(1630) with SP(1630). According to the analysis, SP and SO just use statistical testing to do arrhythmia diagnosis. Data distribution is unknown, so it is unreasonable. In this case, the cut-off is too small, making too many instances are detected as abnormal beats, so sensitivity is best and other measures are worst in SP and SO. According to Table 1, SFO algorithm outperforms other algorithms on positive predictive value, specificity and accuracy. Fuzzification is workable and effective for performance improvement.

5.3. Performance comparison

Because this algorithm is unsupervised, it is compared against four unsupervised methods, which are Brute force[17], LOF[18], Random Walk[19] and K-distance[20]. LOF and K-distance use a minimum point(K) to define the neighborhood. All competing methods need parameter N to determine the number of predicted anomalies. Selection of K and N for these approaches must be predefined.

Different Sizes of datasets lead to the difficulty in N selection. N has significant influence on four measures which can't reach their maximum values at same N. In general, when N is larger, sensitivity is better and other three measures are worse. So there is no meaning to compare the performance when N is different. Same N is selected, which is automatically generated by SFO. For SFO, LOF and K-distance, K optimization is conducted by the method in subsection 5.1. For Random Walk and Brute Force, results when accuracy is maximized are shown. It means that results for all algorithms when accuracy is the best are shown in Table 2. Best results are marked in bold. More best results means this algorithm is better. Results indicate SFO can obtain best performances in measures except sensitivity. Even so, its sensitivity is better than that of Random Walk and Brute Force. Random Walk performs a random walk on Markov chain. There are some correlations for instances, but Markov property may not exist. SFO, K-distance, Brute Force and LOF are distance-based methods, which are simple and intelligible, but the performance is not stable, especially for sparse data. The advantage of a local instead of a global view is detailed in LOF. However, if densities of two different instances are similar, LOF is not effective. K-distance and Brute Force just consider one distance that is the Kth largest distance to the nearest neighbor, therefore they are not thoughtful. Overall, the proposed algorithm is better than competing algorithms.

6. Conclusion

In this work, a novel algorithm for arrhythmia diagnosis is first developed. The problem is viewed as a statistical hypothesis testing. Exploiting fuzzy set theory, hypothesis testing is distribution-free and self-adaptive, and is an unsupervised model. K optimization and fuzzification are necessary for performance improvement. Positive predictive value, specificity and accuracy have been greatly improved by K optimization and intensive fuzzification process, while sensitivity is not significantly affected by different K. A careful comparison confirms the effectiveness of the proposed algorithm.

Acknowledgment

This work is supported by the National Natural Science Foundation of China(61402134) and the National High Technology Research and Development Program of China(863 Program) (2015AA015405).

References

[1] A. Özcift, Random forests ensemble classifier trained with data resampling strategy to improve cardiac arrhythmia diagnosis, Computers in Biology and Medicine **41** (2011), 265–271.
[2] M.R. Homaeinezhad, S.A. Atyabi, E. Tavakkoli, H.N. Toosi, A. Ghaffari and R. Ebrahimpour, ECG arrhythmia recognition via a neuro-SVM-KNN hybrid classifier with virtual QRS image-based geometrical features, Expert Systems with Applications **39** (2012), 2047–2058.
[3] B.H. Zhu, Y.S. Ding and K.R. Hao, Multiclass maximum margin clustering via immune evolutionary algorithm for automatic diagnosis of electrocardiogram arrhythmias, Applied Mathematics and Computation **227** (2014), 428–436.
[4] A.V.D. Bruaene, P. Moons, A. Belmans, M.C. Post, J.G. Luermans, M. Delcroix, A. Pasquet, J.D. Backer, B. Paelinck, M. Morissens and W. Budts, Predictive model for late atrial arrhythmia after closure of an atrial septal defect, International Journal of Cardiology **164** (2013), 318–322.
[5] J.J. Zhu, L.S. He and Z.Q. Gao, Feature extraction from a novel ECG model for arrhythmia diagnosis, Bio-Medical Materials and Engineering **24** (2014), 2883–2891.
[6] B. Doğan and M. Korürek, A new ECG beat clustering method based on kernelized fuzzy c-means and hybrid ant colony optimization for continuous domains, Applied Soft Computing **12** (2012), 3442–3451.
[7] G.F. Lewis, S.A. Furman, M.F. McCool and S.W. Porges, Statistical strategies to quantify respiratory sinus arrhythmia are commonly used metrics equivalent?, Biological Psychology **89** (2012), 349–364.
[8] G. Koulaouzidis, S. Das, G. Cappiello, E.B. Mazomenos, K. Maharatna, P.E. Puddu and J.M. Morgan, Prompt and accurate diagnosis of ventricular arrhythmias with a novel index based on phase space reconstruction of ECG, International Journal of Cardiology **182** (2015), 38–43.
[9] I. Saini, D. Singh and A. Khosla, QRS detection using K-nearest neighbor algorithm (KNN) and evaluation on standard ECG databases, Journal of Advanced Research **4** (2013), 331–344.
[10] J. Derrac, S. García and F. Herrera, Fuzzy nearest neighbor algorithms: Taxonomy, experimental analysis and prospects, Information Sciences **260** (2014), 98–119.
[11] Y.P. Huang, C.Y. Huang and S.I. Liu, Hybrid intelligent methods for arrhythmia detection and geriatric depression diagnosis, Applied Soft Computing **14** (2014), 38–46.
[12] W.K. Lei, B.N. Li, M.C. Dong and M.I. Vai, AFC-ECG: An adaptive fuzzy ECG classifier, Soft Computing in Industrial Applications **39** (2007), 189–199.
[13] G.B. Moody and R.G. Mark, MIT-BIH arrhythmia database directory, Cambridge: MA: MIT. Available: http://www.physionet.org/physiobank/database/mitdb/, Last accessed February, 2013.
[14] C.P. Shen, W.C. Kao, Y.Y. Yang, M.C. Hsu, Y.T. Wu and F.P. Lai, Detection of cardiac arrhythmia in electrocardiograms using adaptive feature extraction and modified support vector machines, Expert Systems with Applications **39** (2012), 7845–7852.
[15] R.J. Martis, U.R. Acharya, K.M. Mandana, A.K. Ray and C. Chakraborty, Application of principal component analysis to ECG signals for automated diagnosis of cardiac health, Expert Systems with Applications **39** (2012), 11792–11800.
[16] A.L. Goldberger, Clinical Electrocardiography: A Simplified Approach 7th Ed., Mosby Elsevier, Missouri, 2006.
[17] E. Keogh, J. Lin, A.W. Fu and H.V. Herle, Finding unusual medical time-series subsequences: Algorithms and applications, IEEE Transactions on Information Technology in Biomedicine **10** (2006), 429–439.
[18] J. Lee, B. Kang and S. Kang, Integrating independent component analysis and local outlier factor for plant-wide process monitoring, Journal of Process Control **21** (2011), 1011–1021.
[19] H.D.K. Moonesinghe and P.N. Tan, Outlier detection using random walks, Proceedings of the 18th IEEE International Conference on Tools with Artificial Intelligence(ICTAI'06), Washington, DC, 2006, pp. 532–539.
[20] A. Nurunnabi, G. West and D. Belton, Outlier detection and robust normal-curvature estimation in mobile laser scanning 3D point cloud data, Pattern Recognition **48** (2015), 1404–1419.