

# ST Segment Change Classification Based on Multiple Feature Extraction Using ECG

Hongmei Wang, Wei Zhao, Yanwu Xu, Jing Hu, Cong Yan, Dongya Jia, Tianyuan You

Guangzhou Shiyuan Electronics co., Ltd, Guangzhou, China

## Abstract

ST deviation detection using electrocardiogram (ECG) is of great significance for ischemia heart disease diagnosis. In this paper, we proposed an algorithm based on multiple feature extraction to classify the ST deviation beat by beat. First, the ST segment was located. Then, morphological and Poincaré features of ST segment were extracted and combined with global feature. Finally, random forest was adopted to classify the ST segment change into normal, elevated or depressed. The algorithm was evaluated on the European ST-T Database and the average sensitivity of normal, depressed and elevated ST segment was 85.2%, 86.9% and 88.8% respectively. The result shows that the developed algorithm is helpful in automatically detecting the ST segment elevation and depression, showing more details of the ischemic syndrome.

## 1. Introduction

ST segment change is a crucial symptom related with myocardial ischemia and detection of ST deviation plays an important role in myocardial infarction diagnosis. The ST segment elevation most happens in patients with transmural myocardial ischemia or variant angina pectoris while the ST segment depression usually appears in subendocardial ischemia or stable or unstable angina [1]. Electrocardiogram (ECG) is a non-invasion, convenient, cheap and widely used way to detect ST deviation.

A number of algorithms [2-3] based morphological features have been widely used to detect the ST deviation. Stergios et al. [4] proposed a method based on self-organizing map (SOM) for the identification of ischemia in signal with V1-V5 lead. Jinho et al. [5] designed three features and used support vector machine (SVM) and kernel density estimation (KDE) to identify ischemia. However, morphology of the ST segment is various, susceptible to noise and patient-specific, thus it's difficult to detect ST deviation accurately.

In this paper, we proposed an algorithm to classify the ST segment changes into normal, depressed and elevated.

First, preprocessing and delineation of the fiducial points were applied to the ECG signal. Secondly, various local morphological features were extracted and combined with the global features. Finally, random forest was trained to classify the heartbeat. The result shows that the algorithm is helpful to detect various types of ST deviation automatically.

## 2. Methodology

The schema of the proposed methodology is shown in Figure 1, including steps of preprocessing, feature extraction and ST change classification.

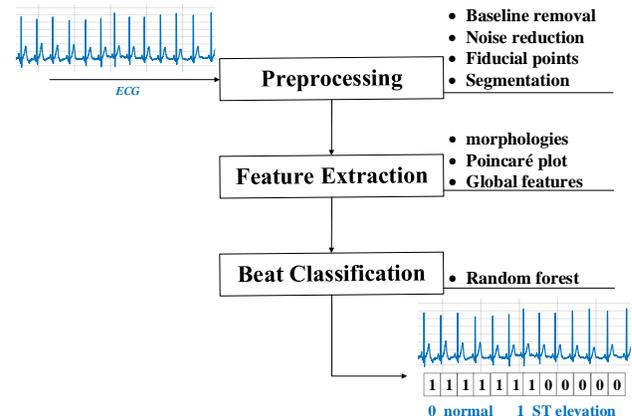


Figure 1. Schema of the proposed methodology

### 2.1. Preprocessing

ECG signal is easily affected by noise such as muscle electricity, power line interference and baseline wander, which often changes the ST segment and the electrical line and further leads to inaccurate detection. The same noise elimination way as Kumar [6] was adopted. Then, the Pan-Tompkins algorithm [7] was used for QRS complex detection. After that, absolute maximum in the window [QRS-0.1s, QRS+0.1s] was searched for R peak. Q, T, P wave and J point were located by the same way as Kumar [6]. Then the ECG signal was segmented, and 5 successive beats were taken as a sample.

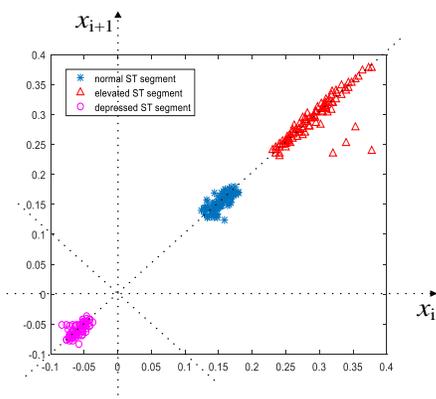
## 2.2. Feature extraction

Features we extracted can be classified into the categories:

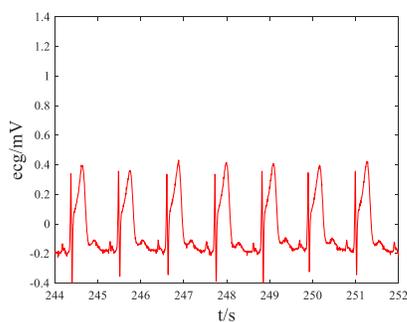
- 1) Poincaré features
- 2) ST segment morphological features
- 3) global features.

For a heartbeat, the duration from J point to J+0.08s was taken as the ST segment, and the change in the ST segment relative to the mean voltage of Q onset was the ST deviation. ST deviation of 5 successive beats were combined into a new series, denoted as  $\{x_i\}, i = 1, 2, \dots, n$ . Then Poincaré plot was constructed by representing the current value of the series  $\{x_i\}$  on the x-axis, and value of the next time  $\{x_{i+1}\}$  on the y-axis. An example of the Poincaré plot is shown in Figure 2(a). For normal and ST elevated heartbeats, points on Poincaré plot mainly distribute in the first quadrant while for ST depressed heartbeats, points mainly distribute on the third quadrant, which is in accordance with the signal shown in Figure 2(b), (c) and (d). On the Poincaré plot, following features were extracted:

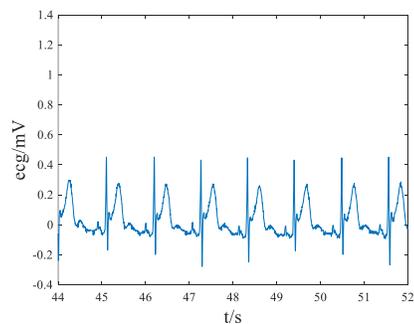
- 1) Average euclidean norm of all points on the Poincaré plot (AEN).
- 2) Average distance of the points to the line  $x_{i+1} = -x_i$  (ADL).
- 3) Intercept of the first-order fitting line of points on the Poincaré plot (PFLI).



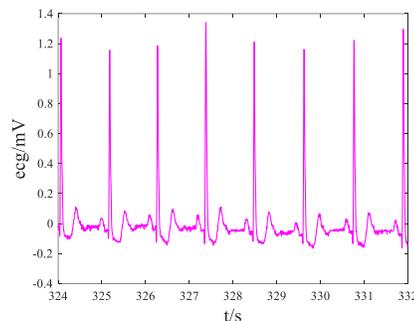
(a)



(b)



(c)



(d)

Figure 2. Poincaré plot of e0105 from European ST-T database and corresponding ECG series. (a) Poincaré plot (b) ST segment elevation series (c) normal series (d) ST segment depression series.

For ST segment morphological feature, following features were extracted.

- 1) Average ST deviation voltage (AST).
- 2) Intercept of the first order fitting line of the ST deviation (FLI).
- 3) Slope (NFLS) and intercept (NFLI) of the Z-score normalized ST deviation's first order fitting line.
- 4) Popular morphological features including the isoelectric energy function (IEEF) [6] and a Rpeak related area feature (SFRA) [5].

Global features including RR intervals (RR), average RR interval (ARR) and lead information (Lead) were extracted to alleviate the patient specificity and lead specificity effect.

## 2.3. Classification

We applied the random forest [8] to classify each beat into normal or with elevated or depressed ST segment. Random forest is helpful to solve the problem of insufficient generalization ability of single classifier. The random forest performance is affected by parameters such as the number of trees (nTree) and the smallest node of each tree (Min\_leaf), which were optimized in the experiment.

### 3. Experiment and Results

The European ST-T Database [9] is used to evaluate the proposed algorithm. The database includes 90 recordings. Each record consists of two-channel 2 hours' ECG signals. The episodes with ST deviations were annotated by experts. We excluded e0107, e0133, e0155, e0163, e0509, e0611. The remaining records were randomly classified into training set and test set, and the same record will never cover both sets. For training set, we randomly select as many normal samples as the ST change samples. The composition of the final dataset is shown in Table 1.

Table 1. Training set and test set.

	Training set	Test set
Record number	58	26
Normal samples	24003	73392
Elevated samples	3033	2008
Depressed samples	22656	7147

Five-fold cross validation was adopted to optimize the classifier parameters. The nTree changed from 100 to 1000 and the Min\_leaf changed in the range of 1 and 100 with step 10. Parameters with the maximum accuracy were selected.

There is a random screening process in the formation of training set. Therefore, under the optimal parameters, we conducted 10 modeling and testing processes. The results are shown in Table 2. As can be seen, mean sensitivity of normal, depressed ST segment (ST-) and elevated ST segment (ST+) was 85.2%, 86.9% and 88.8%. And the sensitivity variance is comparatively very small.

Table 2. Test result.

	RF (mean $\pm$ std)
Normal Se	0.8515 $\pm$ 0.0031
ST+ Se	0.8694 $\pm$ 0.0043
ST- Se	0.8879 $\pm$ 0.0043

Table 3 shows the confusion matrix of the test with maximum overall accuracy. It shows that the wrong prediction of elevated and depressed samples mainly lies in being judged as normal samples rather than misjudging each other.

Table 3. Confusion Matrix.

	Normal	ST-	ST+
Normal	62789	8647	1956
ST-	910	6235	2
ST+	186	40	1782

The proposed algorithm was compared with the

following baseline methods under the same data set. As shown in Table 3, the result of the proposed method is better than the Stergios's [4].

Table 4. Comparison with Stergio's method

	Stergio's method	The proposed method
Accuracy	80.4%	86.9%

Jinho [5] used 10% of each record for training and the rest for testing. As shown in Table 5, under the same dataset, the proposed method is superior and has a better sensitivity ( $S_e$ ) and specificity ( $S_p$ ) on ischemia detection.

Table 5 Comparison with Jinho's method

	The proposed method		
	RF	KDE	SVM
ST deviation $S_e$	0.944	0.939	0.941
ST deviation $S_p$	0.934	0.912	0.923

### 4. Discussion

Importance of each feature was evaluated by permuting out-of-bag observations of the random forest. Figure 3 shows that the PFLI is the most important one for classification. The PFLI has advantage in representing the sloping ST deviation. Assuming the ST segment deviates under following function:

$$x = kt + b, k \in R, t = 1, 2, \dots, n \quad (1)$$

$t$  is time,  $k$  and  $b$  are parameters. when  $k > 0$ ,  $k < 0$  and  $k = 0$ , the ST segment respectively appears upsloping, down sloping and horizontal deviation. Points on Poincaré plot are  $(x_i, x_{i+1})$ , i.e.  $(kt_i + b, kt_{i+1} + b)$ ,  $(x_j, x_{j+1})$ , i.e.  $(kt_j + b, kt_{j+1} + b)$ . Thus first-order fitting line of the points on the Poincaré plot is:

$$z = \frac{x_{j+1} - x_{i+1}}{x_j - x_i} x + \frac{x_{i+1}x_j - x_{j+1}x_i}{x_j - x_i} = x + k \quad (2)$$

PFLI is  $k$ , which can effectively reflect the sloping degree of the ST deviation.

However, as shown in Figure 3, other features also contribute to the classification. The classifier does not work well if only PFLI feature is used. Multiple feature fusion can better characterize the nature of ST segment change and reduce the effect of patient specificity and lead specificity, thus improve accuracy.

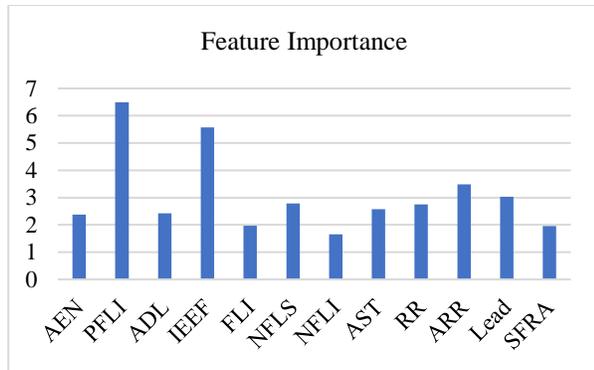


Figure 3. Importance of the extracted features

## 5. Conclusions

We developed an algorithm based on the combination of two-dimension space features, one-dimension space features, global features and random forest classifier was used to classify the ST deviation as normal, depressed or elevated. The performance was tested on the European ST-T database and the results showed  $85.18\pm 0.38\%$ ,  $87.35\pm 0.28\%$ ,  $88.06\pm 0.31\%$  sensitivity respectively for normal, depressed and elevated ST segment. The proposed method outperforms the baseline methods. Multiple features help better explore the essence of ST deviation. The Algorithm can be applied to the ECG monitor and is helpful for automatic diagnosis of ST deviation. In the future, the author will dedicate to methods such as CNN and RNN for automatic feature extraction and ischemia detection.

## References

- [1] Ansari S, Farzaneh N, Duda M et al. A review of automated methods for detection of myocardial ischemia and infarction using electrocardiogram and electronic health records. *IEEE Reviews in Biomedical Engineering* 2017; 10: 264-298.
- [2] Manocha AK, Singh M. An overview of ischemia detection techniques. *International Journal of Scientific and Engineering Research* 2011; 2: 1-6.
- [3] Exarchos TP, Papaloukas C, Fotiadis DI, Michalis LK. An association rule mining-based methodology for automated detection of ischemic ECG beats. *IEEE Transaction on biomedical engineering* 2006; 53: 1531-1540.
- [4] Stergios P, Seferina M, Liviu V et al. Ischemia detection with a self-organizing map supplemented by supervised learning. *IEEE Transactions on Neural Networks and Learning System* 2001; 12: 503-515.
- [5] Jinho P, Witold P, Moongu J. Ischemia episode detection in ECG using kernel density estimation, support vector machine and feature selection. *BioMedical Engineering OnLine* 2012; 1-22.
- [6] Kumar A, Singh M. Ischemia detection using isoelectric energy function. *Computers in Biology and Medicine* 2016; 68: 76-83.
- [7] Pan J, Tompkins WJA. real-time QRS detection algorithm. *IEEE Transaction on Biomedical Engineering* 1985; 32:

230-236.

- [8] Breiman L. Random Forests. *Machine Learning* 2001; 45: 5-32.
- [9] Taddei A, Distanti G, Emdin M, Pisani P, Moody GB, Zeelenberg C, Marchesi C. The European ST-T Database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography. *European Heart Journal* 1992; 13: 1164-1172.

Address for correspondence.

Hongmei Wang  
Guangzhou Shiyuan Electronics co., ltd. No.6, Yunpusi Road,  
Huangpu District, Guangzhou City, Guangdong Province, China  
wanghongmei@cvte.com