Review of Algorithms for Detection the QRS-complex Based on Machine Learning

G.V. Van*, K.V. Podmasteryev**
State University – Education-Science-Production Complex
Oryol, Russia
gennady.van@bmecenter.ru*, asms-orel@mail.ru**

Abstract—A review of methods for detection the QRS-complex on the electrocardiographic signal is done. The results can be used for choosing the prototype QRS detection algorithm, also as for its further development and improvement.

Keywords—ECG signal, QRS-complex, , ANN, K-means, kNN, SVM, R-wave.

I. INTRODUCTION

QRS complex is a dominant of electrocardiographic signal, Fig. 1. Its amplitude and time analysis, shape and appearance time of adjacent rhythms estimation can be used to diagnose a wide range of heart diseases. Thus, the obvious problem is the precise definition of the occurrence time and other various parameters of QRS-complex. Moreover, correct solution of this problem is complicated by the individual variability, wide range of values in health and disease, as well as the noise presence and artifacts in the signal, Fig. 2.

Fig. 1. Electrocardiogram of a single heartbeat. Modified from source picture (url:http://en.wikipedia.org/wiki/Electrocardiography#/media/File:EKG_Complex_en.svg).

Digital signal processing development led to significant improvement of QRS-complexes detection algorithms. These algorithms can be classified by the mathematical apparatus taken as their bases: algorithms based on time and frequency domain transformations; syntactic algorithms; combined algorithms; algorithms based on machine learning methods [1].

However, despite the consistent development of new algorithms, problem of fast and accurate QRS detection is still urgent. One of the most promising approaches is the use of artificial intelligence methods and machine learning. These methods are relatively easy to use, allows real-time implementation and have high sensitivity and specificity.

In general there are several common steps for machine learning based algorithms: obtaining a noisy signal; filtering; forming a set of informative features for classification and adaptive thresholding; calculation of classification function with the help of machine learning methods; use of the resulting classification function. Illustration of certain listed before steps presented on Fig.3.

II. METHODS

There are four basic algorithms for QRS detection based on machine learning methods:

1. Algorithms based on artificial neural network models. These algorithms are usually used for ECG signal morphology analysis and its waves classification [2]. Neural network models provide high adaptability to signal changes, and therefore they are usually used to distinguish the differences between the informative and noise components. Algorithms of this class can be used in real time, but their effectiveness depends essentially on the learning algorithm and training sample quality.
2. Algorithms based on metric classification methods, such as k-nearest neighbors algorithm are also used for morphology analysis and ECG signal elements classification [3]. K-nearest neighbors algorithm based on comparing the unknown ECG signal elements with reference data. The distinctive feature of these algorithms is that training is performed only at the first stage. Thus the efficiency is more dependent on the training set than in neural networks method [3].

3. The specificity of the algorithms based on clustering, in particular k-means method, is that there is no need in forming training set for learning. To use this method it is necessary to determine the number of expected clusters in advance, as well as distance function. The algorithm based on k-means method can be used to distinguish pacemaker QRS-complexes, within normal rhythm. This feature greatly improves the reliability and efficiency [4].

4. Algorithms based on support vector machines (SVM) are based on separating set of analytical features by calculated hyperplane (since the dimension of the selected features set may be arbitrary).

III. SOFTWARE IMPLEMENTATION

This section describes some features of the QRS detection algorithms steps.

A. Algorithms based on artificial neural networks models.

The method proposed by K. Arbateni and A. Bennia [2] consists of two steps. First step suggests using an optimized whitening filter in combination with matching mask. The second step is about the decision of unknown signal fragment membership to the QRS-complexes group.

B. Algorithms based on metric classification methods.

I. Saini, D. Singh, A. Khosla proposed an algorithm described in [3]. Main algorithm stages are: signal filtering, derivative calculation, and QRS-complex detection by using the k-nearest neighbors method.

Low-pass filter can be described with following equation:

\[ y(n) = 2y(n-1) - y(n-2) + \frac{1}{32} [x(n) - 2x(n-6) + x(n-1)] \]  

Next filter is high-pass, described by the following difference equation:

\[ p(n) = p(n-1) - \frac{1}{32} x(n) + x(n-16) - x(n-17) + \frac{1}{32} x(n-32) \]  

QRS-complex has the largest amplitude, so it is reasonable to calculate signal gradient for R-peak better recognition:

\[ \nabla f = \left( \frac{\partial f}{\partial x_1}, ..., \frac{\partial f}{\partial x_k} \right) \]

After signal filtering and gradient calculation k-nearest neighbors method used for QRS-complex detection. In proposed algorithm Euclidian metric is used for distance estimation, and k-value, as number of neighbors, equals 3.

C. Algorithms based on clustering methods.

S. S. Mehta, D. A. Shete, N. S. Lingayat, V. S. Chouhan proposed an algorithm presented in [4]. This algorithm is based on following stages: digital signal filtering, derivative calculation and adaptive thresholding, k-means based QRS-complex detection with subsequent false positive removing.

The finite impulse response notch filter proposed by Van Alste and Schilder [6] is used to remove baseline wander. The filter proposed by Furno and Tompkins [7] is used to remove 50 Hz power line interference.

As an adaptive threshold function the difference between two adjacent samples is considered, as the QRS-complex has the highest slope signal.

Next, the determination of QRS-complexes using the k-means is carried out. This method consists of determining the relation of an unknown element to one of two clusters (QRS-complex, non-QRS-complex). Clusters centers are determined by using annotated signals during iterative learning according to the formula:

\[ Z_j(k+1) = \frac{1}{N_j} \sum_{x \in S_j(k)} X \]  

when \( j=1,2,\ldots,K \); \( Z_j(k+1) \) – override the cluster center; \( N_j \) – the number of elements in the cluster; \( X \) – cluster member. If \( Z_j(k) = Z_j(k+1) \), the iteration terminated.

False-positive removing is based on similar k-means based procedure.

D. Algorithms based on support vector machine.

Filtering procedure, derivative calculation and adaptive thresholding are identical to that described in the third paragraph.

After thresholding and filtering support vector machines based QRS-detection is performed. Hyperplane calculation performed in 10-samples length sliding window.

IV. TESTING AND RESULTS.

In all publications MIT-BIH Arrhythmia database is used for testing. Algorithm efficiency can be estimated by calculating specificity $Sp$ and sensitivity $Se$:

$$Sp = \frac{TN}{TN + FP}$$

$$Se = \frac{TP}{TP + FN}$$

where $TN$ – the number of correct about the absence of QRS complex in the group that does not have the QRS complex; $TP$ – the number of correct conclusions about the QRS complex in the group with the QRS complex; $FP$ – the number of false definitions of QRS complexes in the group that does not have this QRS complex; $FN$ – the number of permits detection of the QRS complex in the group that does not have the QRS complex.

Efficiency values of discussed in present paper QRS-detection algorithms can be found in Table 1.

<table>
<thead>
<tr>
<th>Methods</th>
<th>$Sp$</th>
<th>$Se$</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-means [4]</td>
<td>98.66</td>
<td>98.86</td>
</tr>
</tbody>
</table>

V. CONCLUSION

Discussed machine learning based algorithms show high QRS-complexes detection efficiency (about 98 – 99 %) on test databases. However, one of the crucial drawbacks of presented in review algorithms, is that their effectiveness depends on the quality of the training set. Thus, the actual problem is selection of effective informative features for the classification also as training set dependency reduction.

REFERENCES