Cyberheart-Diagnostics Software Package for Automated Electrocardiogram Analysis Based on Machine Learning Techniques

Ⓒ

V.A. Moskalenko, Graduate Student[1];
A.V. Nikolskiy, MD, PhD, Cardiovascular Surgeon[2];
N.Yu. Zolotykh, DSc, Associate Professor, Professor, Department of Algebra, Geometry and Discrete Mathematics[1];
A.A. Kozlov, Anesthesiologist, Resuscitator[3];
K.A. Kosonogov, MD, PhD, Cardiovascular Surgeon[2];
A.I. Kalyakulina, PhD Student[1];
I.I. Yusipov, PhD Student[1];
V.M. Levanov, MD, DSc, Associate Professor, Professor, Department of Social Medicine and Health Care[4]

[1]National Research Lobachevsky State University of Nizhny Novgorod, 23 Prospekt Gagarina, Nizhny Novgorod, 603950, Russia;

[2]City Clinical Hospital No.5, 34 Nesterova St., Nizhny Novgorod, 603005, Russia;

[3]Nizhny Novgorod Regional Clinical Hospital named after N.A. Semashko, 190 Rodionova St., Nizhny Novgorod, 603126, Russia;

[4]Privolzhsky Research Medical University, 10/1 Minin and Pozharsky Square, Nizhny Novgorod, 603005, Russia

The aim of the study was to develop the Cyberheart-Diagnostics software module, an automated electrocardiogram analysis system being part of the Cyberheart software and hardware complex, and to select machine learning techniques for testing the system based on the comparative analysis of their capabilities.

Materials and Methods. The software package was developed using various machine learning techniques working on a large sample of labeled data, i.e. ECG database with known diagnostic conclusions: support-vector machines, decision tree, artificial neural networks, linear and quadratic discriminant analysis, the random subspace method, AdaBoost, random forest, logistic regression (McCulloch–Pitts neuron model). For comparative analysis and evaluation of the obtained results, the Cyberheart-Diagnostics software was tested using open international ECG databases: Arrhythmia Data Set, PhysioNet PTBDB, PhysioNet Competition 2017 as well as our own database comprising 1652 records of a standard 12-lead resting ECG. The ECG records were interpreted by expert physicians who then formed structured medical conclusions considered as reference.

Results. In different classes of attributes, the diagnostic accuracy of the Cyberheart-Diagnostics software appeared to be 83.8 to 94.5% as compared to the conclusions of expert doctors — 66.3 to 95.1%. Thus, the developed software is comparable with the world analogues in quality of electrocardiogram analysis.

Key words: electrocardiogram; automated analysis; ECG database; machine learning techniques.

Introduction

Development of automated data processing and advances in healthcare information technology along with the use of portable medical equipment open up new opportunities for improving the methods for early diagnosis of circulatory diseases and remote patient monitoring [1, 2]. Electrocardiogram (ECG) still remains the most common method of instrumental diagnostics [3]. One of the challenging directions in this field is development of an intellectual electrocardiograph, i.e. an automated system for decoding electrocardiogram signals followed by issuing a medical report most similar to that of a doctor [2, 4]. Machine learning techniques based on representative ECG sampling are used to achieve this goal.

Nowadays, mobile software products based on methods of automated ECG analysis appear in different countries. At the same time, the degree of confidence in the automated ECG diagnosis without human intervention has remained a controversial issue for several decades. Therefore, the algorithm providing the necessary accuracy of diagnostic conclusion is likely to be the crucial factor in choosing a particular system.

The scientists of the National Research Lobachevsky State University of Nizhny Novgorod created Cyberheart

Corresponding author: Vladimir M. Levanov, e-mail: levanov53@rambler.ru

software and hardware complex for collection, storage, and automated analysis of inhomogeneous medical data. Cyberheart consists of medical probes monitoring the performance of the cardiovascular system and a software module, Cyberheart-Diagnostics. The purpose of this module is automated analysis of ECG of various duration and making pre-hospital diagnostic conclusion.

To improve diagnostic accuracy, the authors have developed and tested an algorithm that includes training and testing of the module using open international databases and structured medical conclusions.

The aim of the study was to develop the Cyberheart-Diagnostics software module, an automated electrocardiogram analysis system being part of the Cyberheart software and hardware complex, and to select machine learning techniques for testing this system based on the comparative analysis of their capabilities.

Materials and Methods

The authors have developed a basic version of ECG analysis software, Cyberheart-Diagnostics, which allows subsequent training. The software interface is written in C# language, mathematical methods are in Python using the matrix computing library NumPy (www.numpy.org), libraries of wavelet analysis PyWavelets (github.com/PyWavelets), and libraries of machine learning techniques Scikit-learn (scikit-learn.org/stable).

The authors have created their own database (Cardiobase) comprising 1652 records of a standard 12-lead resting ECG in digital EDF format. ECG was obtained in adult patients aged 17–80 years, including 743 males (45%) and 909 females (55%). ECG was recorded from patients seeking medical care of cardiologists, arrhythmologists, cardiac surgeons in out-patient departments of Nizhny Novgorod and from those who were hospitalized in cardiology departments and cardiovascular centers.

All 1652 ECG records were interpreted by expert physicians (cardiologists and physicians of functional diagnostics), who then formed structured medical conclusions. Subsequently, the ECG records were analyzed automatically using key point detection (KPD), segmented and automatically described in the form of a pre-hospital report with the software developed by the authors according to the classical criteria for ECG analysis [5].

KPD application algorithm included the following stages:

ECG signal pre-processing: filtering (noise suppression), detecting the baseline (isoline) [1, 3];

KPD signal: detection of the beginning, peak and end of the QRS-complex, P- and T-waves and determining their morphology (Figure 1);

automated report generation.

These data served as the basis for calculating standard signal characteristics (see, for example, [3]).
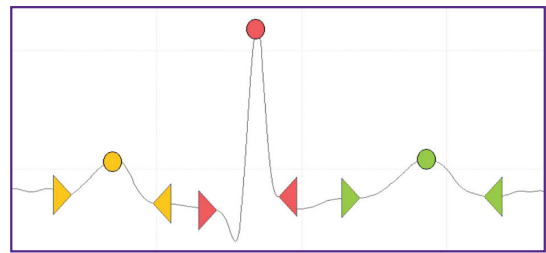

Figure 1. PQRST complex segmentation

Standard functions were used to calculate numerical characteristics (average duration and height of ECG signal complexes, their standard deviations, etc.). As a result, 38 attributes were obtained to describe each ECG lead.

The Cyberheart-Diagnostics software was trained using 1242 out of 1652 ECGs available in the Cardiobase and formalized medical conclusions. The most common machine learning techniques were used for this purpose: the support-vector machine, decision tree, linear and quadratic discriminant analysis [6], the random subspace method [7], AdaBoost [8], random forest [9], logistic regression (McCulloch–Pitts neuron model) with attribute preprocessing through the Batch Normalization layer [10].

In these techniques, the algorithms are automatically built using a large sample of labeled data, i.e. ECG database with known diagnostic conclusions, while the algorithm of establishing a diagnosis (decisive function) is not programmed explicitly. The model is "set up" according to the training sample data, a set of characteristics and attributes of the patient's ECG with a known diagnostic conclusion.

The automated ECG analysis software was trained using machine learning techniques on attributes obtained with KPD (Figure 2). More information about the KPD algorithm is provided in work [5].

To test the Cyberheart-Diagnostics software with the above machine learning techniques, experimental ECG analysis was carried out using the largest open ECG databases processed by medical experts: Arrhythmia Data Set (http://archive.ics.uci.edu/ml/datasets/arrhythmia), PhysioNet PTBDB (https://www.physionet.org/physiobank/database/ptbdb), PhysioNet Competition 2017 (https://physionet.org/challenge/2017). The diagnostic accuracy with respect to the main ECG attributes and classes was taken into account when analyzing the results.

Arrhythmia DataSet database is represented by 410 records of a 12-lead electrocardiogram. Each ECG is described by 279 attributes and belongs to one of the 16 classes of ECG attributes.

PhysioNet PTBDB database consists of 549 ECG records from 290 patients. Each patient belongs to one of nine classes.

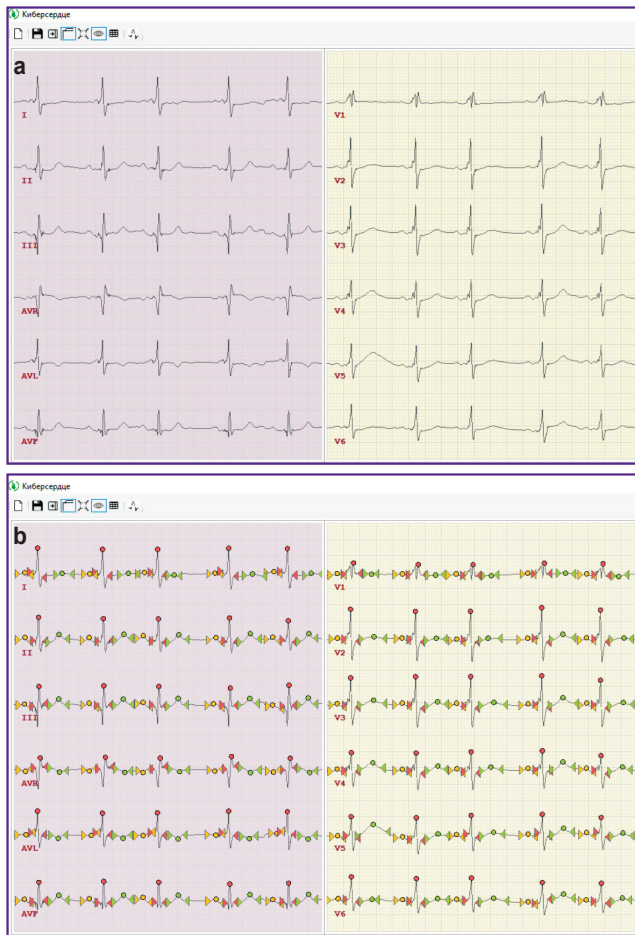PhysioNet Competition 2017 database consists of

Figure 2. ECG (12 leads) of patient U., 69 years:
(a) initial; (b) with segmentation

Table 1

Cardiobase structure according to classes of diagnostic conclusions

Class number	Diagnosis	Number of records	
		Absolute number	%
1	Absolute norm	340	21.2
2	Sinus arrhythmia	287	17.9
3	Right bundle branch block	89	5.5
4	Bifascicular block	48	3.0
5	Atrial fibrillation	170	10.6
6	Ischemic changes	272	16.9
7	Hypertrophy of the left heart	315	19.6
8	Implanted electrocardiostimulator	65	4.0
9	Other	21	1.3
Total		1607*	100.0

* Some ECGs from 1242 patients were included in two or more classes.
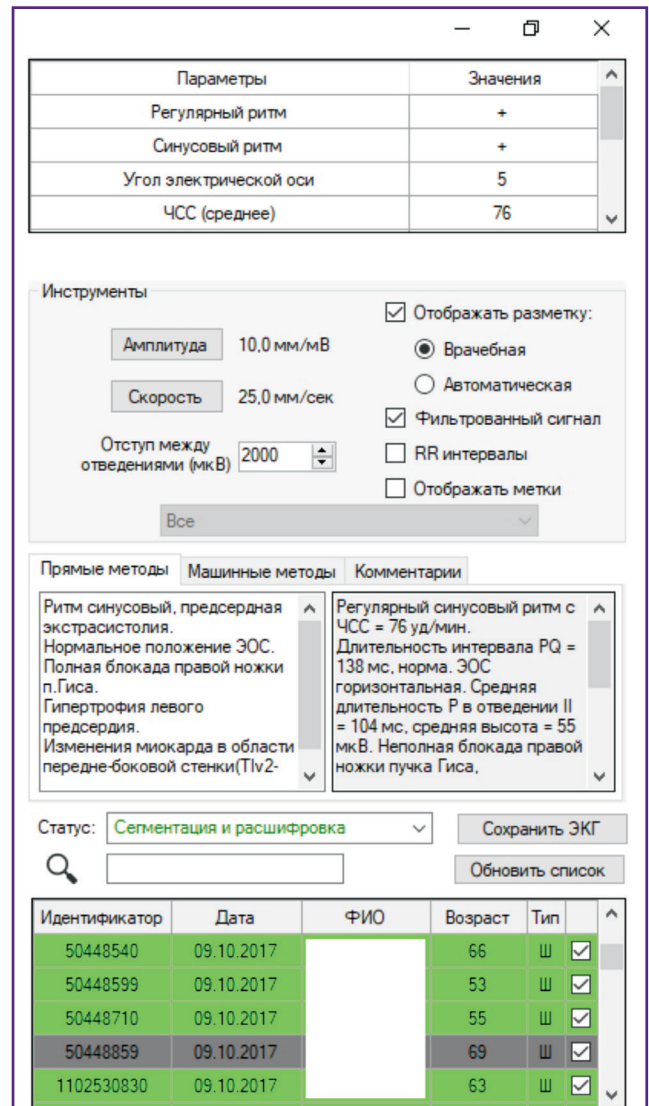
Figure 3. Example of ECG report interface

8528 single-channel five-minute ECG records assigned to one of four classes.

Developing the Cyberheart-Diagnostics software, the authors created their own database (Cardiobase) represented by 1652 ECG records divided into 9 classes and used ECG derived from 1242 patients for training (Table 1).

The Cardiobase classes were developed on the basis of classic clinical and electrophysiological approaches to ECG findings (see Table 1). The following groups of findings were the most frequently presented: normal ECG, sinus arrhythmia, hypertrophy of the left heart, atrial fibrillation, ischemic changes.

The findings obtained by automated analysis of 410 ECGs from the Cardiobase were compared with the data in medical expert reports, which were nominally considered as reference. Figure 3 shows the ECG report interface.

The findings were comprehensively evaluated according to attribute classes. Fivefold cross-checking was carried out using ROC-AUC metric for quality assessment of diagnosis [11, 12].

V.A. Moskalenko, A.V. Nikolskiy, N.Yu. Zolotykh, A.A. Kozlov, K.A. Kosonogov, …, V.M. Levanov

Results and Discussion

In accordance with the objectives of the investigation, sequential testing of the developed software was carried out using machine learning techniques (algorithms) and available international databases.

Arrhythmia database. Testing experiments were carried out with a sample consisting of objects (observations) belonging to the most representative classes: class 1 ("Norm"), class 2 ("Ischemic changes"), and class 10 ("Right bundle branch block") (a total of 339 records). The software was tested using such methods as support-vector machine, random subspace, the nearest neighbors, decision tree, neural networks, linear and quadratic discriminant analysis. Fivefold cross-checking was carried out using ROC-AUC metric for quality assessment. Table 2 shows the parameters of the algorithms that yielded the best results.

An experiment on three-class classification (classes 1, 2, and 10) without preliminary selection of attributes was also carried out. The following machine learning techniques were used: support-vector machine, logistic regression, AdaBoost, random forest. The data were randomly divided into training (70%) and test (30%)

samples. This division was carried out 100 times. The achieved accuracy of the algorithms is presented in Table 3.

Testing on PhysioNet PTBDB database. The task of binary classification between classes 1 ("Myocardial infarction") and 9 ("Healthy") was considered. Other classes were under-represented, therefore it was difficult to identify any patterns useful for proper diagnostics. Logistic regression yielded the best result (McCulloch–

Table 2

Results of the experiment with preliminary selection of attributes (%)

Method	ROC-AUC (for "Ischemic changes" class)	ROC-AUC (for "Right bundle branch block" class)
Support-vector machine	94.1	89.0
Random subspace method	94.5	87.2
Discriminant analysis	93.6	83.8

Table 3

Three-class classification accuracy obtained on the training sample

Method	Accuracy (%)
Support-vector machine	86.1
Gradient boosting	85.3
Random forest	85.2
Logistic regression	82.7

Table 4

Results of the experiments using PhysioNet Competition 2017 database

Database classes		Accuracy (%)
First	Second	
N, A, O	~	94.7
N	A	87.2
N	A, O	79.5
O	A	78.0
N	O	77.4

N o t e: N — normal rhythm; A — atrial fibrillation; O — other rhythm; ~ — noise.

Table 5

Results of testing the Cyberheart-Diagnostics software on the created Cardiobase (n=410)

ECG class	Class name	The number of ECGs according to expert report data	The number of truly positive conclusions	Sensitivity	The number of truly negative conclusions	Specificity	Accuracy
		P	TP	TPR=TP/Px100%	TN	TNR=TN/(n–P)x100%	A=(TP+TN)/n (%)
1	Absolute norm	156	120	76.9	234	92.1	86.3
2	Sinus arrhythmia	105	74	70.5	302	99.0	91.7
3	Right bundle branch block	22	20	90.9	252	64.9	66.3
4	Bifascicular block	6	4	66.7	307	76.0	75.9
5	Atrial fibrillation	33	31	93.9	271	71.9	73.7
6	Ischemic changes	70	63	90.0	248	72.9	75.9
7	Hypertrophy of the left heart	11	10	90.9	323	80.9	81.2
8	Implanted electrocardiostimulator	5	5	100.0	385	95.1	95.1
9	Other	3	2	66.7	256	62.9	62.9

Pitts neuron model) with attribute preprocessing through the Batch Normalization layer. The accuracy of 86.1% was achieved with the balanced sample.

Testing on PhysioNet Competition 2017 database. A series of binary classification tasks were set, the results of the experiments are described below (Table 4).

The value of F1-score (the harmonic average between specificity and sensitivity) was found to be 0.81. It should be noted, the best F1-score indices in the international study equaled 0.83, according to PhysioNet Competition 2017. Thus, the results of algorithms implemented by us are comparable in diagnostic accuracy with the world analogues.

Testing on the created Cardiobase. The results of testing the software on the created Cardiobase are presented in Table 5.

The sensitivity of diagnostic conclusions made by the Cyberheart-Diagnostics software equaled 66.7–100.0% for different classes as compared to the conclusions of doctors-diagnosticians. Specificity of the method amounted to 62.9–99.0%, accuracy — 62.9–95.1%.

Conclusion

The results of testing the developed Cyberheart-Diagnostics software with the main machine learning techniques showed rather high accuracy of diagnostic conclusions achieved by this software as compared to conventionally "ideal" conclusions made by doctors-diagnosticians.

The data obtained by the authors as a result of automated ECG decoding with the presented software and the authors' own Cardiobase materials were compared with the data from three large publicly available ECG databases. Despite the fact that these databases had differences in the duration of cardiosignal recording, the number of leads, sets of classes and other criteria, diagnostic accuracy of the algorithm developed by the authors in the corresponding classes of attributes amounted to 62.9–95.1%.

Conflict of interests. The authors have no conflict of interests to disclose.

References

1. Yurovskiy A.Yu., Sukhov S.S. Distant analysis of ECG and computerized electrocardiography — modern alternatives to classic "paper" solutions. *Practicheskaya meditsina* 2017; 2: 14–17.

2. Strutynskiy A.V. *Elektrokardiogramma: analiz i interpretatsiya* [Electrocardiogram: analysis and interpretation]. Moscow: MEDpress-inform; 2017; 224 p.

3. Vorobiov L.V. ECG analysis of cardiac activity of a healthy person. *Mezhdunarodnyy zhurnal prikladnykh i fundamental'nykh issledovaniy* 2016; 10: 549–553.

4. Drozdov D.V., Levanov V.M. Automatic ECG analysis: problems and prospects. *Zdravookhranenie i meditsinskaya tekhnika* 2004; 1: 10.

5. Kalyakulina A.I., Yusipov I.I., Moskalenko V.A., Nikolskiy A.V., Kozlov A.A., Zolotykh N.Yu., Ivanchenko M.V. Finding morphology points of electrocardiographic signal waves using wavelet analysis. *Izvestiya Vuzov. Radiofizika* 2018; 61(8): 773–789.

6. Duda R.O., Hart P.E., Stork D.G. *Pattern classification*. Wiley-Interscience; 2000.

7. Tin Kam Ho. The random subspace method for constructing decision forests. *IEEE Trans Pattern Anal Mach Intell* 1998; 20(8): 832–844, https://doi.org/10.1109/34.709601.

8. Freund Y., Schapire R.E. A decision-theoretic generalization of on-line learning and an application to boosting. *J Comput Syst Sci* 1997; 55(1): 119–139, https://doi.org/10.1006/jcss.1997.1504.

9. Breiman L. Random forests. *Machine Learning* 2001; 45(1): 5–32.

10. Ioffe S., Szegedy C. *Batch normalization: accelerating deep network training by reducing internal covariate shift.* 2015. URL: https://arxiv.org/pdf/1502.03167.pdf.

11. Brown C.D., Davis H.T. Receiver operating characteristics curves and related decision measures: a tutorial. *Chemometr Intell Lab Syst* 2006; 80(1): 24–38, https://doi.org/10.1016/j.chemolab.2005.05.004.

12. Petrov V., Lebedev S., Pirova A., Vasilyev E., Nikolskiy A., Turlapov V., Meyerov I., Osipov G. CardioModel — new software for cardiac electrophysiology simulation. In: Voevodin V., Sobolev S. (editors). *Supercomputing. RuSCDays 2018. Communications in computer and information science.* Vol. 965. Springer, Cham; 2018; p. 195–207, https://doi.org/10.1007/978-3-030-05807-4_17.

V.A. Moskalenko, A.V. Nikolskiy, N.Yu. Zolotykh, A.A. Kozlov, K.A. Kosonogov, ..., V.M. Levanov