www.igi-global.com

Comparative Study of Neural Network-based Approaches for QRS Segmentation

Anna Borde
Bauman Moscow State Technical University, Medicom LLC, Russia

George Kolokolnikov
Bauman Moscow State Technical University, Medicom LLC, Russia

Victor Skuratov
All-Russian Research Institute of Radio Engineering, Russia

Roman Gaponov
Medicom LLC, Russia

Anastasiya Rumyantseva
Bauman Moscow State Technical University, Russia

ABSTRACT

The paper is devoted to the development of QRS segmentation system based on deep learning approach. The considered segmentation problem plays an important role in the automatic analysis of heart rhythms, which makes it possible to identify life-threatening pathologies. The main goal of the research is to choose the best segmentation pipeline in terms of accuracy and time-efficiency. Process of ECG-signal analysis is descripted and the problem of QRS segmentation is discussed. State-of-the-art algorithms are analyzed in literature review section and the most prominent are chosen for further research. In the course of the research four hypotheses about appropriate deep learning model are checked: LSTM-based model, 2-input 1-dimensional CNN model, "signal-to-picture" approach based on 2-dimensional CNN and the simplest 1-dimensional CNN model. All the architectures are tested, and their advantages and disadvantages are discussed. The proposed ECG segmentation pipeline is developed for Holter monitor software.

Keywords: Deep Learning, Electrocardiography (ECG), Holter Monitoring, Machine Learning, Convolution Neural Network (CNN), Long Short-term Memory (LSTM), ECG Analysis Software, ECG Segmentation,

INTRODUCTION

Diagnosis of patients with cardiovascular diseases (CVD) has now become easier thanks to advances in Holter monitoring (HM) – the monitoring of electrical activity of the cardiovascular system for at least 24 hours. However, death rate from CVD globally increased over the last 10 years by 14.5% and amounted nearly 17.6 million deaths in 2016 (Benjamin, Muntner, & Bittencourt, 2019). Solution to this problem requires improvement in automatic ECG interpretation algorithms for HM analysis software.

BACKGROUND

One of the most important tasks in automatic ECG signal analysis is detection of specific points: P, QRS and T waves, i.e. ECG segmentation. Specific points include onset, peak and offset of wave (Figure 1).

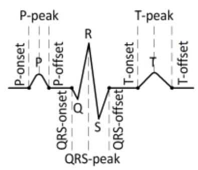


Figure 1. Specific points of ECG signal wave

ECG segmentation can be performed by classical mathematical methods, including such algorithms as heuristic rules and probabilistic models (Beraza, & Romero, 2017). However, these approaches require large computing resources in cases of very noisy signals and various morphologies of ECG signal waves. Deep learning models, especially artificial neural networks (ANN), are currently the most promising way to overcome these limitations (Xiang, Lin, & Meng, 2018). Clinical conditions of HM ECG analysis software use are strictly time limited due to high patients' flow. In this connection, the main criterion for choosing the ANN architecture along with accuracy maximization is inference time minimization. The research discusses one software module for ECG analysis that solves the problem of ECG waves segmentation. The entire Holter ECG signal analysis pipeline is shown in Figure 2.

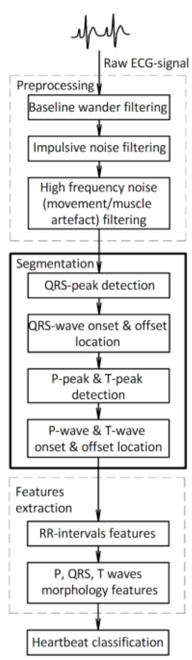


Figure 2. Holter ECG signal analysis pipeline (module, developed in the paper, is highlighted with solid bold lines, other modules – with dashed lines)

MAIN FOCUS OF THE ARTICLE

Goal Settings and Development Pipeline

The development of QRS segmentation system goes through the steps shown in Figure 3. The research is aimed to define the most precise signal segmentation algorithm with the least computing power consumption.

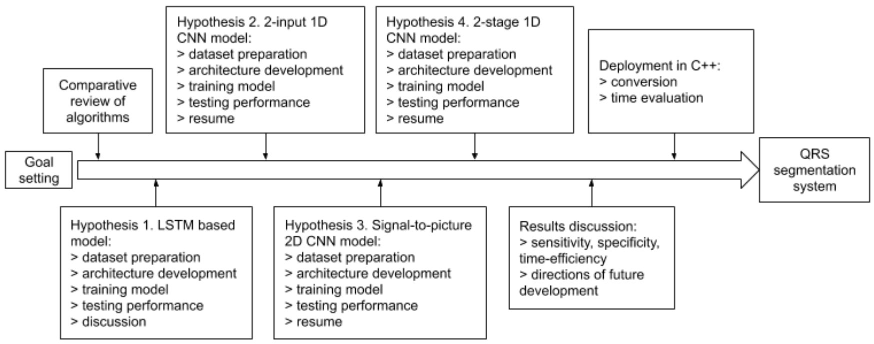


Figure 3. Development pipeline

At the first step, comparative study of modern algorithms is conducted. Then the first hypothesis is checked: LSTM-based model with pre- and post-processing operations is developed and dataset for training is formed. After training the model, testing takes place. According to the acquired results, the hypothesis about the applicability of the approach is accepted or rejected. At the next stage, the same sequence of operations is applied to the 2nd hypothesis about 2-input 1-dimensional convolutional neural network model. Hypothesis 3 about “signal-to-picture” approach based on 2-dimensional convolutional model is checked. Finally, the drawbacks of previous approach are considered, and the fourth method based on 2-stage 1-D CNN approach is developed. After aggregating the performance characteristics of each method, the best approach is defined, deployment in C++ HM ECG analysis software is carried out, and directions of model development are discussed.

Literature Review

Standard Evaluation Procedure of QRS Detection Algorithms

According to IEC 60601-2-33 Ed. 3.0:2010 (Medical electrical equipment - Part 2-33: Particular requirements for the basic safety and essential performance of ambulatory electrocardiographic systems) for a reliable evaluation of the algorithm, the sensitivity and specificity of heartbeat detection should be presented. Evaluation indexes are the sensitivity (Sen) and positive predictivity (PPR) (International Electrotechnical Commission, 2010), which are defined as follows:

$$\begin{aligned} Sens &= \frac{TP}{TP+FN} \times 100\% \\ PPR &= \frac{TP}{TP+FP} \times 100\% \end{aligned} \qquad (1)$$

where TP is the number of truly detected QRS complexes, FN is the number of false negative (missed detected) QRS complexes, and FP is the number of false positive (extra falsely detected) QRS complexes.

A large number of ECG databases (DB) with annotated ECG records is available on the PhysioNet Resource (Goldberger et al., 2000). For evaluation procedure AHA, MIT-BIH, NST, CU and ESC DB are recommended (International Electrotechnical Commission, 2010). For segmentation algorithms evaluation QTDB is also commonly used (Laguna, Mark, Goldberg, & Moody, 1997). QTDB provides onset, peak, and offset markers for P, QRS and T waves.

Comparison of Existing ANN-based Approaches

According to the literature review of ECG specific points detection algorithms in the last 10 years, there is a prevalence of ANN-based approaches over classical mathematical methods (Borde, 2018). Due to the huge number of such publications the literature review is limited by the results of working groups published since 2018. Evaluation results of ANN-based approaches are presented in Table 1. Terms and definitions used in Table 1: CNN – convolution neural network; LSTM – long short-term memory; MLP – multilayer perceptron; RNN – recurrent neural network.

Table 1. Comparison of ANN-based approaches published since 2018

Authors	ANN architecture	ECG DB	Sen	PPR	Explanations
Xiang Y., Lin Z., Meng J.	2-input 1-D CNN	MIT-BIH	99.86%	99.89%	QRS-peak detection of raw ECG and with noise added
Jun T. J., Nguyen H. M., Kang D., Kim D., Kim D., Kim Y. H	2-D CNN	MIT-BIH	97.85%	98.55%	QRS-peak detection and heartbeat classification
Abrishami H., Han C., Zhou X., Campbell M., Czosek R.	Bidirectional RNN with LSTM layers	QTDB	95.00%* 98.00%** 97.00%***	–	Segmentation of P, QRS and T waves (* - P segmentation, ** - QRS segmentation, *** - T segmentation; accuracy for all waves is given instead of Sen)
Yildirim Ö.	Bidirectional RNN with LSTM layers and additional wavelet-based layer	MIT-BIH	99.39%	–	QRS-peak detection and heartbeat classification
Saadatnejad S., Oveisi M., Hashemi M.	Multiple RNN with LSTM layers and wavelet transform features	MIT-BIH	93.00%* 66.90%**	98.20%* 95.70%**	QRS-peak detection and heartbeat classification (* - results for ventricular ectopic beats, ** - results for supraventricular ectopic beats)

Full segmentation pipeline by LSTM neural network is presented only in work (Abrishami, Han, Zhou, Campbell, & Czosek, 2018). The accuracy characteristics are lower than in 1-D CNN approach (Xiang, Lin, & Meng, 2018), however this approach focuses only on the peak detection. A detailed review on classical mathematical methods for ECG segmentation is presented in (Beraza, & Romero, 2017) and reports that methods based on the wavelet transform are the most promising. Thereby research (Yildirim, 2018) and (Saadatnejad, Oveisi, & Hashemi, 2019) based on discrete wavelet transform (DWT) features are interesting in terms of increasing the accuracy of ANN algorithms.

Materials and Methods

Technological Stack and Datasets

Technological stack used in the research is based on Python 3.6 programming language with utilization of the following libraries:

- Pandas – for data management.
- Scikit-learn – for dataset preparation.
- WFDB – for reading, writing, and processing PhysioNet signals and annotations.
- SciPy – for ECG-signal processing.
- Keras with TensorFlow backend – for deep learning model training and testing.
- Auxiliary packages like Numpy, Matplotlib, etc.

The development of segmentation systems is carried out with the following hardware:

- Nvidia GeForce GT 1030 GPU for prototyping and testing models.
- Google Colaboratory with Tesla K80 GPU for training.

The following databases are used in the research:

- QTDB (Laguna, Mark, Goldberg, & Moody, 1997) – a database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG.
- MIT-BIH Arrhythmia database – a database with 48 half-hour excerpts of two-channel ambulatory ECG recordings.

The QTDB contains in total 105 fifteen-minute two-lead ECG records (many excerpted from other databases), with onset, peak, and offset markers for P, QRS, T, and (where present) U waves of from 30 to 50 selected beats in each record. Description of the MIT-BIH Arrhythmia DB is presented in Table 2.

Table 2. Description of the MIT-BIH Arrhythmia DB

<table>
<tr><th>ECG DB</th><th>Record ID</th><th>Description</th><th>Records number</th></tr>
<tr><td rowspan="4">MIT-BIH</td><td>100, 101, 103, 105, 106, 108, 109, 111-119, 121-124</td><td>Records without arrhythmia</td><td>20</td></tr>
<tr><td>200-203, 205, 207-210, 212-215, 219-223, 228, 230-234</td><td>Records with arrhythmia</td><td>24</td></tr>
<tr><td colspan="2">Total records number</td><td>44</td></tr>
<tr><td colspan="2">102, 104, 107, 217
Records contain impulses of an artificial pacemaker, excluded</td><td>4</td></tr>
</table>

Each record of QTDB is preprocessed according to the algorithm shown in Figure 4 and saved as CSV file. The similar preprocessing pipeline is applied to records of MIT-BIH database (Figure 5). In each studied method datasets are preprocessed additionally in a specific way, depending on the used neural network architecture, to form training and testing datasets.

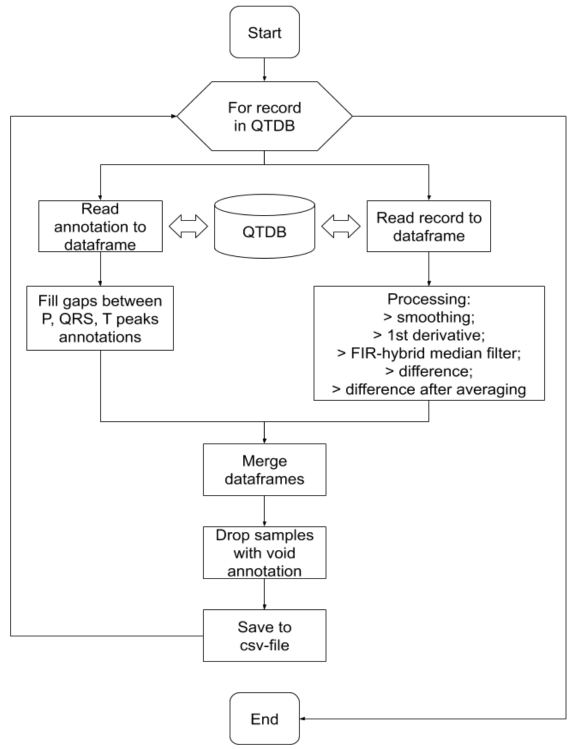


Figure 4. Algorithm of QTDB preprocessing

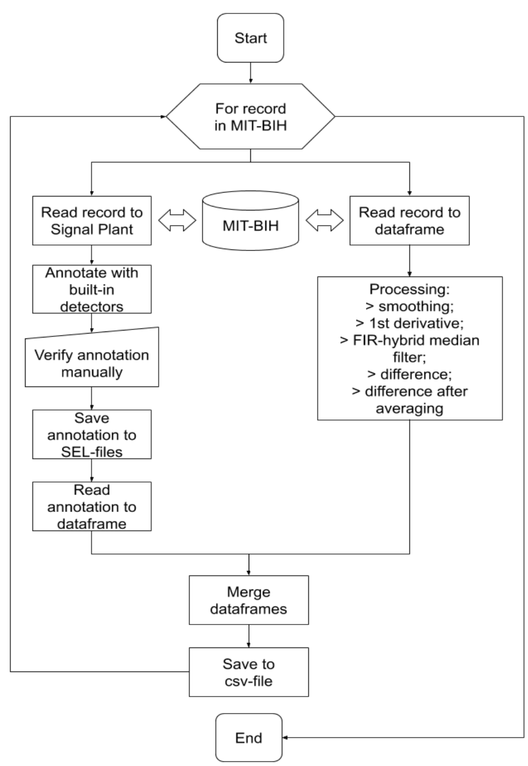


Figure 5. Algorithm of MIT-BIH preprocessing

Hypothesis 1: LSTM-based Model

The first method is based on a bidirectional long short-term memory (BLSTM) neural network (Schuster & Paliwal, 1997). This type of recurrent neural network is well-known for high performance while working with time-domain signals and sequential data. It allows to reach previous and future input information from the current state, so the context of the input helps to distinguish different classes. It is useful in ECG analysis as the signal has periodical trend: QRS complex goes after P wave and is followed by T wave. The proposed implementation of network is working in many-to-many sequence classification mode. The model is trained to detect P, QRS, and T waves.

The signal segmentation pipeline based on the BLSTM model is shown in Figure 6. The initial signal is preprocessed (smoothed, filtered, or differentiated). Then 250-samples window is moving along the signal. This window is given to LSTM network which generates 250 labels of annotation. After the signal analysis is done, generated annotation is subjected to median filtering that removes wrong spikes (dramatic changes in annotation). Additional logic filter controls the correct sequence of P, QRS, and T waves. Median filter window size is chosen according to the information about ECG waveforms possible duration: both normal and abnormal (Table 3 – information provided by Department of Therapeutic faculty, Pediatric faculty, Russian National Research Medical University named after N.I. Pirogov RNRMU). For avoiding QRS false positive detection according to minimum QRS width in MIT-BIH DB window size is chosen equal to 23.

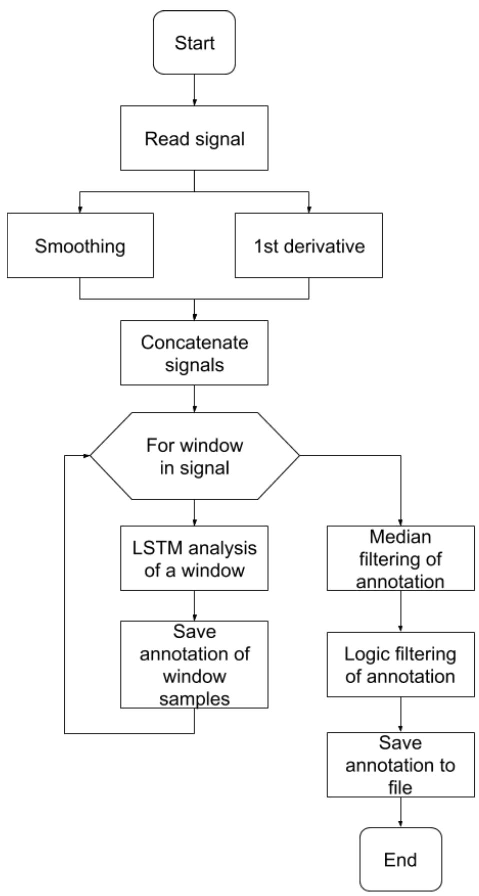


Figure 6. Signal segmentation pipeline based on the BLSTM model

Table 3. ECG waveforms duration

ECG feature	Duration, ms	Duration (fs = 360 Hz)
P width	30-200	10-72
PQ interval	80-400	28-144
QRS width	50-240	18-87
QT interval	200-700	72-252
T width	30-240	10-87

The architecture of bidirectional long short-term memory network (Figure 7) is inspired by paper (Abrishami, Han, Zhou, Campbell, & Czosek, 2018) and consists of two bidirectional LSTM layers and one time-distributed dense layer.

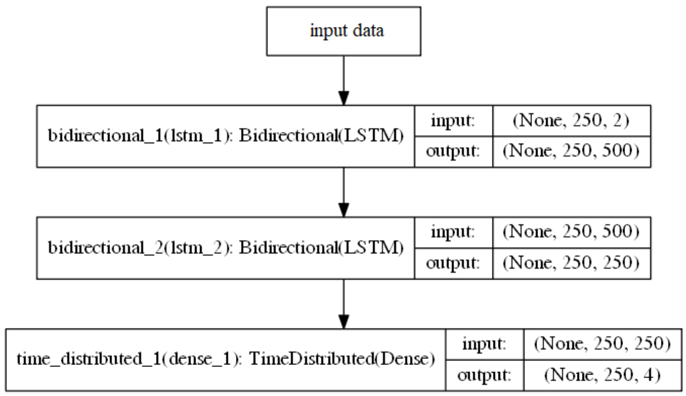
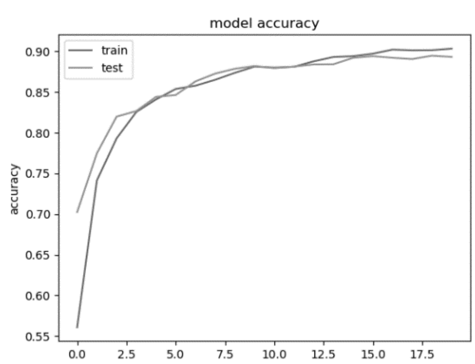


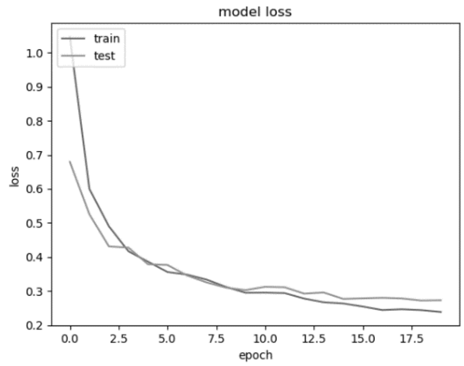
Figure 7. Architecture of BLSTM neural network

The BLSTM network requires specific input data form: it can process at one time number of samples that is equal to the number of input neurons. It can also get time-dependent signal with more than one feature in one sample. This means that the BLSTM can get filtered, differentiated or smoothed signal along with the initial one as input. Consequently, the input data size is (windows size x number of features per sample). To form dataset for BLSTM, each record of QTDB in csv-format is split into segments of 250-samples length. The total number of signal segments in training set is 2900 and 350 in testing set.

The BLSTM model is trained for 200 epochs with batch size of 64 signal segments. Categorical cross entropy is chosen as loss function and Adam method – for optimization. After 15 epochs learning curves go to a plateau. The learning curves for the BLSTM model with smoothed and differentiated signals as input are shown in Figure 8.



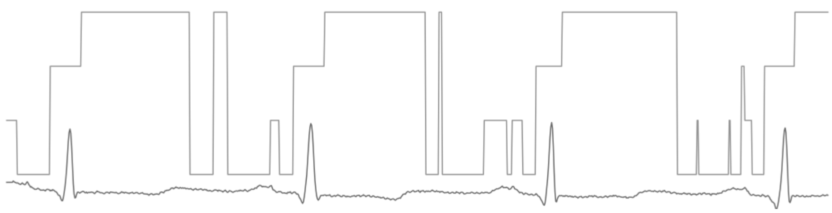
a)



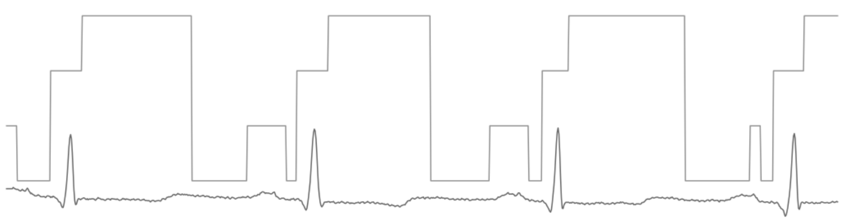
b)

Figure 8. Learning curves for BLSTM network: a) accuracy learning curve, b) loss learning curve

The proposed pipeline requires approximately 0.40 second to process 1 second of signal with 250 Hz sampling frequency. Example of signal segmented with trained BLSTM is shown in Figure 9. The lower curve is ECG signal and the upper – annotation (0 stands for "neutral", 1 – "P", 2 – "QRS", 3 – "T"). The segmented signal has many wrong detections, which can be removed by median and logic filtering (Figure 9). Although the pipeline manages to locate complexes, the accuracy of edge detection is not very high. Moreover, the method is not time-efficient, thus, it is declined.



a)



b)

Figure 9. BLSTM-segmented signal: a) before filtering, b) after median and logic filtering

Hypothesis 2: 2-input 1-D CNN-based Model

As ECG signal is a time sequence, using 1-D variation of CNN models seems to be the most suitable processing method. This method is based on the idea of current signal count classification into 4 classes: P-wave, QRS-wave, T-wave or neutral (baseline). The central count of a window is classified according to the surrounding context.

The signal segmentation pipeline based on the 2-input 1-D CNN model is shown in Figure 10. The initial signal is preprocessed (differentiated and differentiated after averaging). Then 250-samples window is moving along the signal and the 125th count of a current window is classified according to neighbor counts. After the signal analysis is done, generated annotation is subjected to median filtering that removes wrong spikes, as in hypothesis 1.

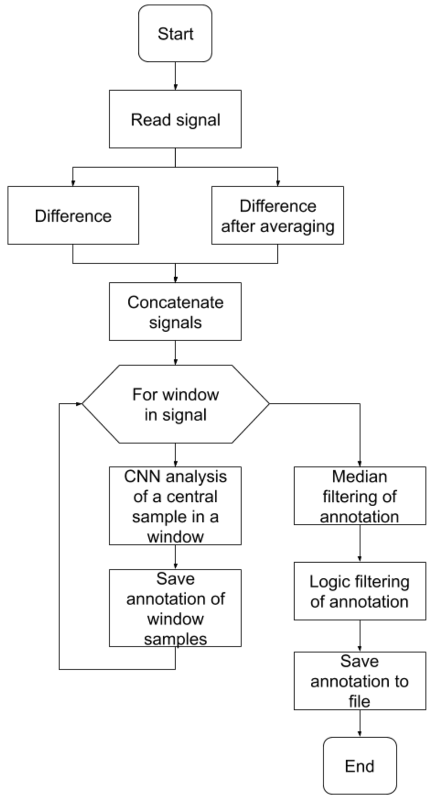


Figure 10. Signal segmentation pipeline based on the 2-input 1-D CNN model

The architecture of 2-input convolution neural network (Figure 11) is inspired by paper (Xiang, Lin, & Meng, 2018). Differentiated signal is given as the first input and differentiated after averaging signal as the second input of the neural network. The first branch of the network is represented by two convolutional layers, whereas the second one contains one convolutional layer. Features extracted by two convolutional branches are concatenated and passed to fully connected classification layers. Figure 11 shows the variation of the architecture developed only for QRS segmentation: it annotates signal in 1 for QRS and 0 for non-QRS segment. That is why the last layer consists of 2 neurons.

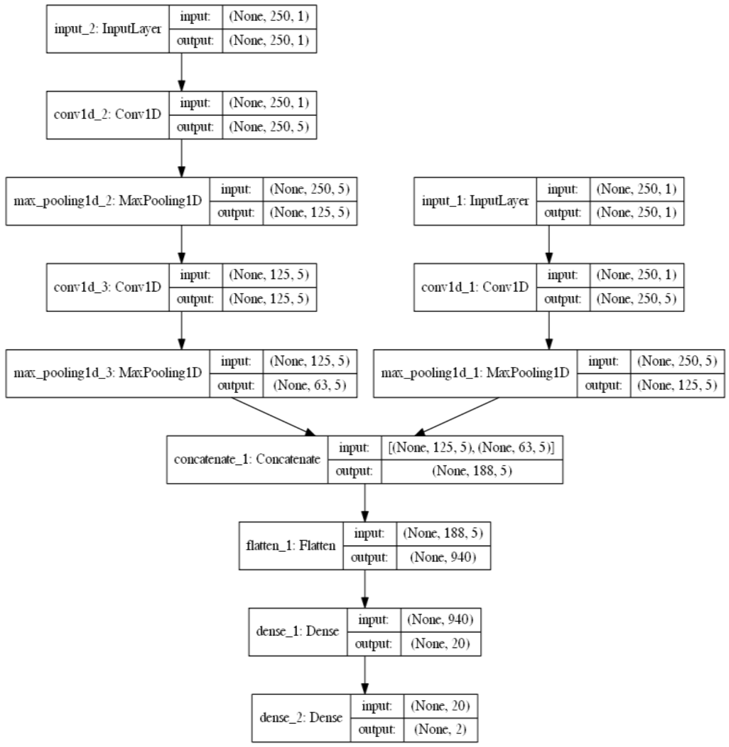


Figure 11. Architecture of 2-input 1-D CNN

2-input 1-D CNN model gets 250-sample frames as input, so the dataset should be processed to fit the input size: number of frames x window size x feature per sample. For instance, record containing 650000 samples is split into 649750 frames (record length minus frame length).

The method under consideration turned out to be demanding on a large amount of data. To generate additional training set, QRS onset and offset annotation of downsampled MIT-BIH DB with 250-Hz frequency is carried out using SignalPlant, an open signal processing software platform (Plesinger, Jurco, Halamek, & Jurak, 2016). QRS complexes edges are obtained with SignalPlant built-in detector and then manually verified. Thus, training set consists of QTDB and MIT-BIH DB records.

The model is trained for 100 epochs with batch size of 128 frames. Categorical cross entropy is chosen as loss function and Adam method – for optimization. Validation of the signal segmentation pipeline is conducted on MIT-BIH DB (360 Hz). The results of performance evaluation are shown in Table 4. Figure 12 shows QRS segmentation after median and logic filtering. As follows from the table and figure, method based on the 2-input 1-D CNN model performs high accuracy of both peaks and edges detection. Nevertheless, the method requires approximately 0.45 second to process 1 second of signal with 360 Hz sampling frequency, so it is not time-efficient enough for real-world applications.

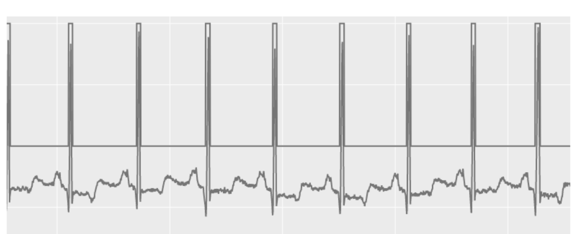


Figure 12. 2-input 1-D CNN QRS segmentation after median and logic filtering

Table 4. Performance evaluation of the hypothesis 2 using the MIT-BIH DB

Record	**Total beats**	**TP**	**FP**	**FN**	**Sen, %**	**PPR, %**
100	2273	2270	0	3	99.87	100.00
101	1856	1850	4	6	99.68	99.78
103	2084	2082	2	2	99.90	99.90
105	2572	2524	84	48	98.13	96.78
106	2027	2017	20	10	99.51	99.02
108	1763	1682	36	81	95.41	97.90
109	2532	2519	10	13	99.49	99.60
111	2124	2093	3	31	98.54	99.86
112	2539	2517	4	22	99.13	99.84
113	1795	1784	114	11	99.39	93.99
114	1879	1874	1	5	99.73	99.95
115	1953	1951	1	2	99.90	99.95
116	2412	2391	3	21	99.13	99.87
117	1535	1533	1	2	99.87	99.93
118	2278	2255	28	23	98.99	98.77
119	1987	1986	11	1	99.95	99.45
121	1863	1852	14	11	99.41	99.25
122	2476	2472	0	4	99.84	100.00
123	1518	1516	4	2	99.87	99.74
124	1619	1617	14	2	99.88	99.14
200	2601	2573	37	28	98.92	98.58
201	1963	1961	1	2	99.90	99.95
202	2136	2134	20	2	99.91	99.07
203	2980	2893	34	87	97.08	98.84
205	2656	2653	0	3	99.89	100.00
207	1860	1844	232	16	99.14	88.82
208	2955	2935	3	20	99.32	99.90
209	3005	2982	1	23	99.23	99.90
210	2650	2631	9	19	99.28	99.66
212	2748	2717	0	31	98.87	100.00
213	3251	3248	0	3	99.91	100.00
214	2262	2242	196	20	99.12	91.96
215	3363	3358	2	5	99.85	99.94
219	2154	2153	18	1	99.95	99.17
220	2048	2045	1	3	99.85	99.95
221	2427	2422	4	5	99.79	99.84
222	2483	2469	8	14	99.44	99.68
223	2605	2597	28	8	99.69	98.93
228	2053	2012	44	41	98.00	97.86
230	2256	2253	6	3	99.87	99.73
231	1571	1570	3	1	99.94	99.81
232	1780	1776	12	4	99.78	99.33
233	3079	3075	4	4	99.87	99.87
234	2753	2752	0	1	99.96	100.00
Overall	100724	100080	1017	644	99.36	98.99

Hypothesis 3: 2-D CNN Model

So called "signal-to-picture" approach is based on converting 1-D signal into 2-D image and its classification with simple 2-D CNN. The main intuitions of this method are in the following:

- Cardiologist's visual analysis analogy: segmentation system is "looking through" the signal, finds QRS peaks, and detects edges.
- CNN architecture is originally intended for image analysis.
- The proposed method uses the simplest CNN models to increase time-efficiency.
- System uses two CNN models: one for QRS peak detection, another for QRS edges detection.

The pipeline of "signal-to-picture" approach is shown in Figure 13. The ECG signal is scaled and transformed into a set of ECG images with 75x150 size where the number of samples is plotted on one axis, and the normalized signal amplitude on the other. Each image is the result of 75-samples window moving along the initial signal with a sequential 1-sample shift. After transformation, image data is sent to CNN for analysis to determine the position of the QRS complex, after which its maximum is determined. Further, in a certain area around the maximum obtained, a "window" of variable duration is run, and the received images are sent for analysis to the second CNN, which acts as an edge detector of QRS complexes. According to the frame selected by CNN, the QRS complex onset and offset are located. The annotation of the signal containing QRS onset, peak, and offset is saved to a file. The developed system is trained to detect QRS peak, onset and offset.

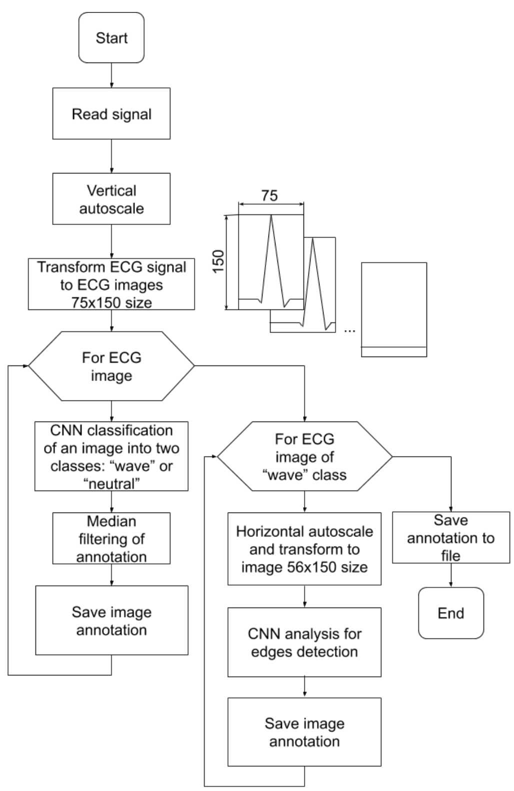


Figure 13. Signal segmentation pipeline based on 2-D CNN model

Two CNN of the same architecture with difference in the last layer are used. Figure 14 shows CNN model for peak detection.

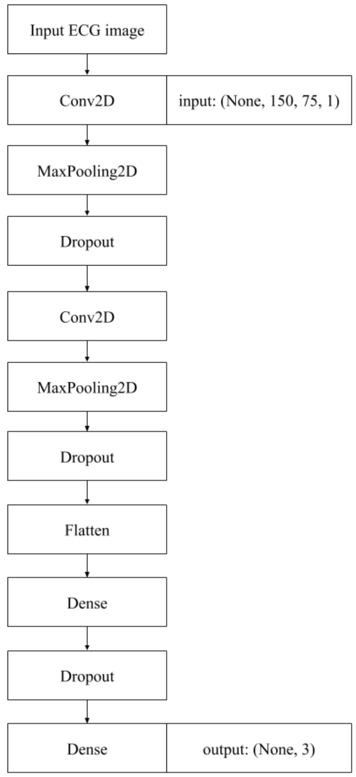


Figure 14. Architecture of CNN for QRS peak detection

Two datasets are required to train CNN models used in “signal-to-picture” approach. The datasets are formed from selected MIT-BIH records. The first set contains “wave” and “non-wave” classes. “Wave” class has instances in which QRS peak is located in the middle of the image. All the other images with non-QRS waves or with QRS peaks shifted from the center position belong to “Non-wave” class. The second dataset consists of images with QRS peaks located in the middle but with different onset and offset location. There are two classes in this set: “Correct” and “Incorrect” edge location.

Training process of each neural network takes 20 epochs. Accuracy on test set is more than 0.99 in the of the training process. The proposed pipeline can process 1 second of MIT-BIH record with 360 Hz in 0.03 second. Results of ECG signal segmentation are demonstrated in Figure 15.

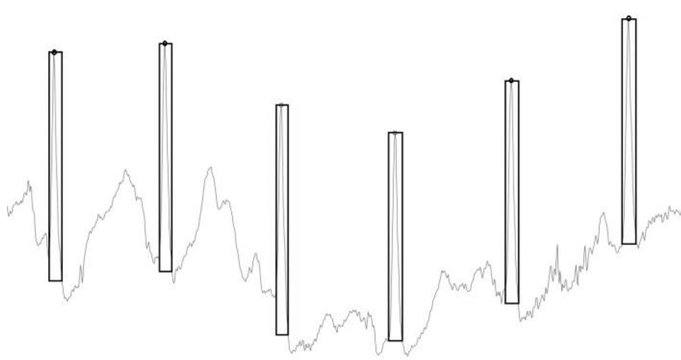


Figure 15. ECG segmentation by “signal-to-picture” approach

Although it seems that two-dimensional "signal-to-picture" approach is the most appropriate for QRS segmentation among others (Borde, Kolokolnikov, & Skuratov, 2019), it has one major disadvantage: the data preprocessing step. This is the main motivation for conversion this two-dimensional approach into one-dimensional. It makes possible to achieve utilization of simpler models with faster performance and get rid of the specific data preprocessing step. Detailed information of this method is represented in the next section.

Hypothesis 4: 2-stage 1-D CNN approach

2-stage 1-D CNN approach is a rethinking of a previous method. It utilizes the simplest possible one-dimensional convolutional neural networks in two steps: at first for peak location, and then for edge detection. It allows to reduce computing power requirements, hence increase time-efficiency. The algorithm based on 2-stage 1-D CNN approach is shown in Figure 16.

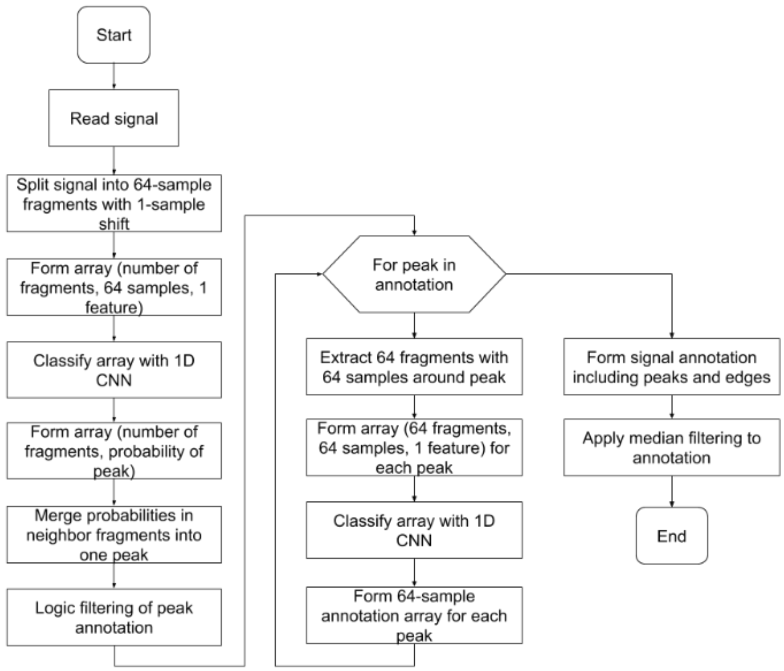


Figure 16. 2-stage signal segmentation pipeline based on 1-D CNN model

The initially read signal is divided into fragments of 64 samples with 1-sample shift (overlap of two fragments is 63 counts). The first one-dimensional convolutional neural network receives formed two-dimensional array with the following dimensions: number of fragments, 64 samples, 1 feature. At the same time, the neural network classifies the central 32nd sample of the fragment according to the first 31 and last 32 samples. As a result, the model skips 32 samples at the beginning and the end of the record. As a result, in accordance with the input array, an output array of dimensions (number of fragments, 1 value) is formed containing the probability of a peak in this fragment. Next, the output array is analyzed to determine the coordinate of the peak. Variation of the decision-making threshold leads to changes of sensitivity and specificity. With an increase in the threshold of acceptance, sensitivity increases, and specificity decreases. Changing the threshold of adoption from 0.5 to 0.9 leads to the difference in sensitivity and specificity by ~ 1%. As a result of peak detector operation, a list of signal peaks coordinates is formed. The list of peak coordinates is additionally filtered, so that the distance between adjacent peaks is at least 72 counts. Example of peak annotation is shown in Figure 17.

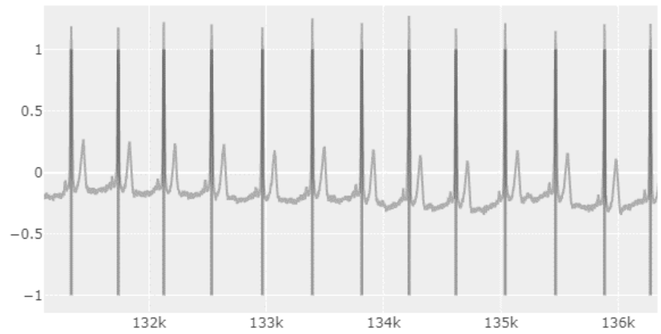


Figure 17. Peak annotation with 1-D CNN peak detector

The obtained coordinates are used to extract the areas around these peaks. An array of fragments is formed containing 64 fragments with a size of 64 samples for each peak of the signal. The formed array is fed to the input of a one-dimensional convolutional neural network. At the same time, the network receives (64 fragments, 64 samples, 1 feature) for each individual signal peak in the for-loop. The probability of belonging to QRS complex of the central sample in the fragment is generated. Next, the decision on the threshold is applied. Depending on the threshold value, sensitivity and specificity may change as in the previous case. As a result, an annotation array is formed for an area of 64 samples around each peak. Further, annotation for the entire record is formed from the markings in the region of each peak. At the end of edge detector operation, a median filter can be applied to the annotation to remove double triggering on a single QRS complex. Example of edge detection is shown in Figure 18.

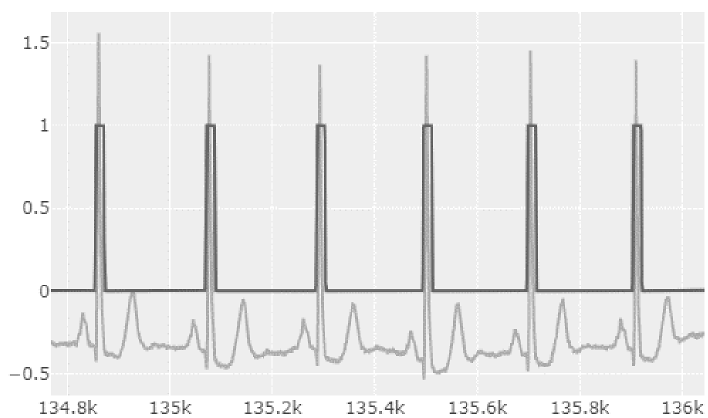


Figure 18. QRS segmentation with 1-D CNN edge detector

Architecture of both convolutional neural networks is the same (Figure 19). Due to this fact, it is possible to reduce storage requirements by using the same computational graph of the network for peak and edge detectors. The network contains one convolutional layer, followed by dropout for regularization. Features extracted by the first layer are flattened and passed to classification fully connected layers with dropout. At the end of the network the probability is generated for the central sample of the analyzed fragment. In both cases the problem is reduced to binary classification:

- Is the central sample of the fragment represented by QRS peak (true/false)?
- Does the central sample of the fragment belong to QRS complex (true/false)?

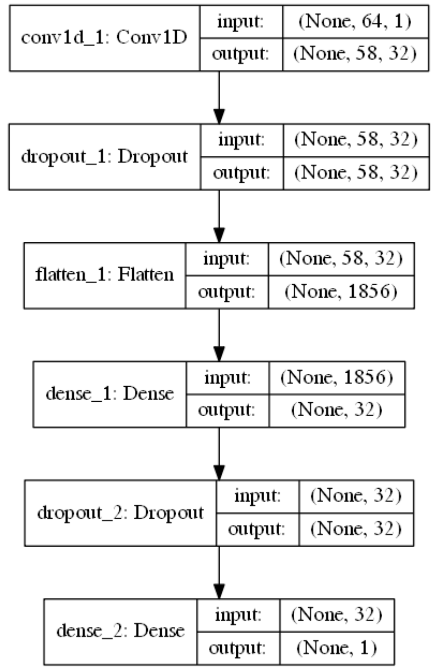


Figure 19. Architecture of CNN used in peak and edge detectors

Both models (for peak and edge detection) were trained for 50 epochs. After 20 epochs the learning curve went to a plateau of 97%. Validation of the models was carried out on MIT-BIH dataset. The performance evaluation of peak detector with decision threshold 0.9 is represented in Table 5. Sensitivity and PPR averaged 96.1% and 97.8% respectively. The total time required for analysis of 43 30-minute record (in total 21.5 hours) is approximately 19 minutes in Python environment without speedup.

Table 5. Performance evaluation of 1-D CNN peak detector using the MIT-BIH DB

Record	Total beats	TP	FP	FN	Sen, %	PPR, %	Time, sec
100	2273	2272	0	1	99.96	100.00	26.03
101	1865	1864	4	1	99.95	99.79	26.26
103	2084	2084	0	0	100	100.00	26.25
105	2572	2497	105	75	97.08	95.96	25.17
106	2027	1759	175	268	86.78	90.95	24.63
108	1763	1364	314	399	77.37	81.29	25.92
109	2532	2473	31	59	97.67	98.76	26.19
111	2124	2056	56	68	96.8	97.35	26.13
112	2539	2539	0	0	100	100.00	26.1
113	1795	1794	0	1	99.94	100.00	26.14
114	1879	1369	478	510	72.86	74.12	26.26
115	1953	1953	0	0	100	100.00	26.12
116	2412	2384	11	28	98.84	99.54	26.12
117	1535	1534	4	1	99.93	99.74	26.01
118	2278	2222	56	56	97.54	97.54	25.64
119	1987	1627	0	360	81.88	100.00	26.04
121	1863	1848	4	15	99.19	99.78	27.34
122	2476	2476	1	0	100	99.96	26.14
123	1518	1517	1	1	99.93	99.93	26.06
124	1619	1596	11	23	98.58	99.32	26.24

200	2601	2549	36	52	98	98.61	26.17
201	1963	1814	72	149	92.41	96.18	26.11
202	2136	2120	13	16	99.25	99.39	26.23
203	2980	2830	101	150	94.97	96.55	25.7
205	2656	2650	0	6	99.77	100.00	25.4
207	1860	1676	32	184	90.11	98.13	25.38
208	2955	2453	41	502	83.01	98.36	25.96
209	3005	3002	8	3	99.9	99.73	27.39
210	2650	2606	23	44	98.34	99.13	26.3
212	2748	2747	3	1	99.96	99.89	26.17
213	3251	3129	17	122	96.25	99.46	25.15
214	2262	2122	126	140	93.81	94.40	26.21
215	3363	3332	12	31	99.08	99.64	26.22
219	2154	2106	1	48	97.77	99.95	26.23
220	2048	2047	0	1	99.95	100.00	26.21
221	2427	2418	0	9	99.63	100.00	26.39
222	2483	2439	3	44	98.23	99.88	26.17
223	2605	2500	92	105	95.97	96.45	26.77
228	2053	1911	56	142	93.08	97.15	26.29
230	2256	2195	68	61	97.3	97.00	25.98
231	1571	1571	0	0	100	100.00	25.42
232	1780	1739	4	41	97.7	99.77	25.71
233	3079	2866	176	213	93.08	94.21	26.17
234	2753	2750	0	3	99.89	100.00	26.14
Total	100733	96800	2135	3933	96.1	97.84	1146.66 ≈ 19 min

The performance evaluation of edge detector with decision threshold 0.5 is represented in Table 6. Sensitivity and PPR of onset and offset detection averaged 96.9%. Processing of 38 30-minutes records (in total 19 hours) took approximately 4 minutes in the same conditions as peak detector evaluation. It should be noted that due to the lack of open access for marking the boundaries of QRS complexes from the MIT database, this database was manually marked up. A training sample was formed from 80% of the data, and testing was also carried out by comparison with manual annotation. Since the presence of errors in the manual annotation process can be admitted, the main criterion for a successful result was the separation of the width of normal complexes (up to 100 ms) and pathological wide ones, such as ventricular complexes and blockade (over 100 ms) according to the original annotation of the MIT database.

Table 6. Performance evaluation of 1-D CNN edge detector using the MIT-BIH DB

Record	Sen_onset, %	PPR_onset, %	Sen_offset, %	PPR_offset, %	Time, sec
100	99.82	99.82	99.96	99.96	9.5
101	99.84	99.89	99.89	99.95	5.49
103	100	100	100	100	6.14
105	98.91	97.32	97.58	96.01	7.32
106	97.34	97.24	93.44	93.35	5.75
108	91.15	95.88	77.99	82.04	4.89
109	99.84	99.84	99.49	99.49	7.24
111	98.59	98.63	96.89	96.94	5.99
112	99.8	99.65	99.76	99.61	7.28
113	100	100	100	100	4.96
114	89.89	90.18	94.47	94.77	5.34
115	99.95	99.95	99.95	99.95	5.55
116	99.09	99.75	99.13	99.79	6.94

117	99.93	99.8	99.93	99.8	4.37
118	98.42	98.38	98.81	98.77	6.5
119	99.9	99.9	99.6	99.6	5.56
122	100	100	100	100	7.07
124	98.64	97.98	99.07	98.4	4.44
200	88.89	88.75	97.65	97.5	7.31
201	97.5	98.51	95.62	96.6	5.66
202	99.81	99.81	99.49	99.49	6.05
203	93.12	93.06	95.07	95	8.37
205	98.04	98.15	99.62	99.74	7.47
207	82.1	84.55	77.58	79.9	5.28
208	94.01	94.2	84.06	84.23	8.45
209	99.93	99.93	99.97	99.97	8.48
210	96.79	97.2	97.47	97.88	7.4
212	99.96	99.6	97.56	97.21	7.83
213	96.71	96.68	99.35	99.32	9.21
214	98.63	98.5	92.31	92.19	6.29
215	99.23	99.17	98.81	98.75	9.59
219	99.58	99.58	99.95	99.95	6.07
220	99.95	100	99.95	100	5.82
221	99.59	99.51	99.84	99.75	6.9
222	97.66	98.7	96.09	97.11	7
223	98.93	98.85	98.08	98.01	7.42
228	94.4	93.9	96.05	95.54	6.13
232	62.96	59.18	95.23	89.5	5.06
233	99.09	99.06	99.61	99.58	8.78
Total	96.88	96.9	96.97	96.99	260.87 ≈ 4 min

Results and Discussion

For further discussion three time-efficient approaches with the highest accuracy are selected, leaving aside signal segmentation pipeline based on the BLSTM model. According to the models' evaluation results, dramatic increase in time consumption is observed in case of "signal-to-picture" 2-D CNN approach. However, this method requires bulky data preprocessing step that significantly loads the system. To get around this computationally intensive stage, 2-stage 1-D CNN approach is developed. The main intuitions of the hypothesis 3 are implemented in hypothesis 4 taking into account the transition from two-dimensional to one-dimensional architecture. Along with much simpler data preprocessing this approach requires only 0.01 second for analysis of 1 second of 360 Hz ECG record. Moreover, it gives sufficient accuracy which is comparable to other approaches. The results of models' evaluation are shown in Table 7.

Table 7. Comparison of three most suitable approaches

Model	**Average accuracy, %**	**Time required for analysis of 1 second of 360 Hz ECG record, seconds**
2-input 1-D CNN model	99.18	0.45
2-D CNN model	99.00	0.03
2-stage 1-D CNN model	96.97	0.01

The further directions of QRS segmentation system development include:

- Development of both P and T waves detection modules with the selected approach.
- Improving time-efficiency with code base conversion to C++ and parallel programming.

- Deployment of the algorithm into the Holter monitoring ECG analysis software.

Deployment of Neural Network-based QRS detector in C++

Reasons for deployment in C++

The aim of this step is to integrate the developed ECG segmentation model into the HM ECG analysis software. The main problem is that usually there is no personal computer with either integrated or discrete graphical processing unit (GPU) at the typical medical institution. In this regard, it is necessary to implement an approach that allows to perform all the operations only with central processing unit (CPU). 2-stage 1-D CNN approach (hypothesis 4) is chosen as the most efficient method for implementation in C++.

Method of conversion

The authors have applied several different CPU-built extensions in order to provide time-efficiency of the algorithm. Universal approach consisting in automatic selection of the fastest extension supported by the particular CPU have been developed. The following extensions are used for time-efficiency increasing:

- Floating-point unit (FPU) – a part of a computer system specially designed to carry out operations on floating-point numbers.
- Streaming SIMD extensions (SSE).
- Advanced vector extensions (AVX).

SSE (streaming SIMD extensions) and AVX (advanced vector extensions) are SIMD (single instruction multiple data streams) instruction sets supported by CPUs manufactured by Intel and AMD (Jeong, Kim, Lee, & Myung, 2012). These extensions allow parallel processing by multiple cores in a single CPU. Basic arithmetic and data transfer operations such as sum, multiplication and square root can be processed simultaneously. In particular, deep learning methods working with vectors and matrices can be easily optimized by the SIMD programming.

Time evaluation

Example of QRS segmentation by 2-stage 1-D CNN approach implemented in C++ is shown in Figure 20. Comparison of processing time using different CPU-built extensions is shown in Table 8.

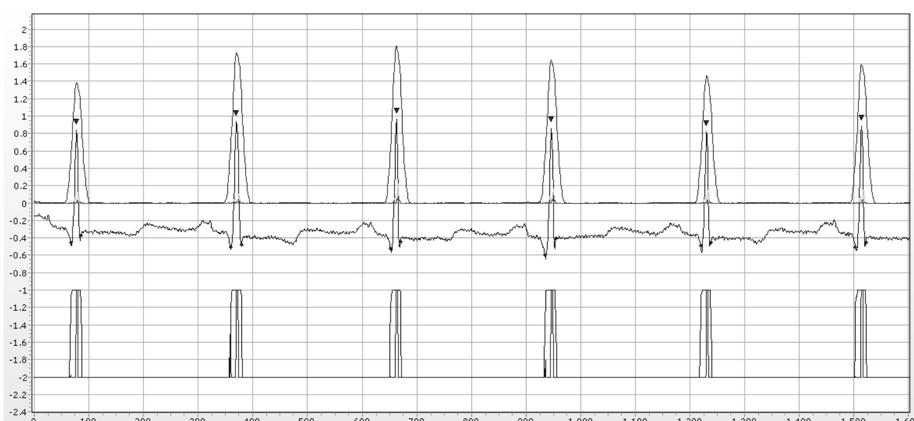


Figure 20. QRS segmentation with 2-stage 1-D CNN approach converted to C++ (first 5 seconds of MIT-DB 100 record with performance of Pan-Tompkins real time QRS detection algorithm — on the top, detected peaks and edges using 2-stage 1-D CNN approach — on the bottom)

Table 8. Comparison of processing time using different CPU-built extensions

Extension	Average time required for segmentation of 30 min MIT-BIH ECG record, seconds	Average time required for segmentation of 1-channel 24-hours ECG record, seconds	Average time required for segmentation of 3-channel 24-hours ECG record using 4 cores CPU, seconds
FPU	25.6	1225.9	817.2
SSE	4.3	204.9	136.6
AVX	2.8	133.8	89.2

According to the results shown in Table 7 with using only FPU extension it's possible to reach equal time efficiency with the model tested on GPU (Table 4-5). SSE extension gives 6 times gain in time efficiency, and AVX extension gives 9 times gain. Moreover, average time for 3-channel 24-hours ECG record (running on 4 core CPU supporting AVX) segmentation is less than two minutes. This value is acceptable for HM ECG analysis software. The results obtained, in turn, make it possible to put forward recommendations for personal computers for equipping medical institutions.

CONCLUSION

In the course of the research devoted to the development of automated QRS segmentation system, the authors have carried out literature review to define approaches with the highest performance characteristics. Four hypotheses about ECG segmentation methods are checked and the best one is accepted. Detailed description of dataset preparation, algorithm architecture and training the models is provided for each approach. BLSTM-based model is built and tested, but evaluation shows the impossibility of applying the model in real-life applications. 2-input 1-D CNN model has demonstrated high accuracy of peak and edges detection – average accuracy is 99.18%, but long processing time – it takes 0.45 seconds to process 1 second of 360 Hz MIT DB ECG record. "Signal-to-picture" approach has proved its efficiency in terms of accuracy – 99% which is close to the results of the previous approach, and time-efficiency – the proposed algorithm is 15 times faster than the previous one. Although it seems to be the most appropriate approach for QRS segmentation, it has one major disadvantage which is data preprocessing step. ECG signal registered in time domain form needs to be transformed into 2-D data, in turn it leads to time and computing power consumption at the data preparation step. Finally, 2-stage 1-D CNN model has shown a bit lower accuracy characteristic, although significantly less time and computing power consumption. This method is chosen for further C++ implementation and deployment into HM ECG analysis software. The following technical solutions are applied for better performance of the algorithm operation on CPU: CPU-built extensions including FPU, SSE, AVX. As a result, the best averaged processing time of 3-channel 24-hours ECG record performed by using AVX extension is 1.5 minutes. The further directions of work include development of P and T waves detection modules with the selected approach.

ACKNOWLEDGMENT

Authors would like to thank Medicom LLC for providing computing power and material support to the research.

REFERENCES

Abrishami, H., Han, C., Zhou, X., Campbell, M., & Czosek, R. (2018). Supervised ecg interval segmentation using lstm neural network. In Proceedings of the International Conference on Bioinformatics & Computational Biology (BIOCOMP) (pp. 71-77). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).

Benjamin, E. J., Muntner, P., & Bittencourt, M. S. (2019). Heart disease and stroke statistics-2019 update: a report from the American Heart Association. Circulation, 139(10), e56-e528.

Beraza, I., & Romero, I. (2017). Comparative study of algorithms for ECG segmentation. Biomedical Signal Processing and Control, 34, 166-173.

Borde A. (2018, October). Intelligent decision support system in cardiology. In Russian Journal of Cardiology 2018 (pp. 163a-163b).

Borde, A., Kolokolnikov, G., & Skuratov, V. (2019). Development of neural network-based approach for QRS segmentation. In Proceedings of the Proceedings of the 25th Conference of Open Innovations Association FRUCT, Helsinki, Finland (FRUCT 25) (pp. 77-84).

Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. circulation, 101(23), e215-e220.

International Electrotechnical Commission. (2010). Medical electrical equipment-Part 2-33: Particular requirements for the basic safety and essential performance of magnetic resonance equipment for medical diagnosis. IEC 60601-2-33 Ed. 3.0.

Jeong, H., Kim, S., Lee, W., & Myung, S. H. (2012). Performance of SSE and AVX instruction sets. arXiv preprint arXiv:1211.0820.

Jun, T. J., Nguyen, H. M., Kang, D., Kim, D., Kim, D., & Kim, Y. H. (2018). ECG arrhythmia classification using a 2-D convolutional neural network. arXiv preprint arXiv:1804.06812.

Laguna, P., Mark, R. G., Goldberg, A., & Moody, G. B. (1997, September). A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG. In Computers in cardiology 1997 (pp. 673-676). IEEE.

Saadatnejad, S., Oveisi, M., & Hashemi, M. (2019). LSTM-based ECG classification for continuous monitoring on personal wearable devices. IEEE journal of biomedical and health informatics.

Schuster, M., & Paliwal, K. K. (1997). Bidirectional recurrent neural networks. IEEE transactions on Signal Processing, 45(11), 2673-2681.

Plesinger, F., Jurco, J., Halamek, J., & Jurak, P. (2016). SignalPlant: an open signal processing software platform. Physiological measurement, 37(7), N38.

Xiang, Y., Lin, Z., & Meng, J. (2018). Automatic QRS complex detection using two-level convolutional neural network. Biomedical engineering online, 17(1), 13.

Yildirim, Ö. (2018). A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification. Computers in biology and medicine, 96, 189-202.