
Normal Versus Abnormal ECG Classification by the Aid of Deep Learning

Linpeng Jin and Jun Dong

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.75546>

Abstract

With the development of telemedicine systems, collected ECG records are accumulated on a large scale. Aiming to lessen domain experts' workload, we propose a new method based on lead convolutional neural network (LCNN) and rule inference for classification of normal and abnormal ECG records with short duration. First, two different LCNN models are obtained through different filtering methods and different training methods, and then the multipoint-prediction technology and the Bayesian fusion method are successively applied to them. As beneficial complements, four newly developed disease rules are also involved. Finally, we utilize the bias-average method to output the predictive value. On the Chinese Cardiovascular Disease Database with more than 150,000 ECG records, our proposed method yields an accuracy of 86.22% and 0.9322 AUC (Area under ROC curve), comparable to the state-of-the-art results for this subject.

Keywords: telemedicine, cardiovascular disease, ECG, deep learning, convolutional neural network, rule inference, classification

1. Introduction

As an inexpensive, noninvasive and well-established diagnostic tool for cardiovascular disease, the electrocardiogram (ECG) has been widely applied since the 1980s. In clinics, its main applications are as follows: (1) long-time ECG monitoring, such as 24-hour ambulatory ECG and (2) short-term ECG recording, such as standard 10-second, 12-lead ECG.

With the improvement of living standards, more and more people pay attention to their own health problems and performing an ECG test is the preferred method for preventing cardiovascular disease. Although it is easy to collect ECG now, diagnostic conclusions cannot be

accurately provided since domain experts are in short supply, especially in the basic community medical insurance system (BCMIS) of China. A feasible solution is that ECG records are sent to be interpreted by domain experts in telemedicine centers, who send back their diagnostic conclusions via the Internet. In fact, such institutions are common in China now, for instance, Shanghai Aerial Hospital Network, Henan Telemedicine Center, and so on. But in telemedicine centers, a large number of ECG records need to be interpreted and the workload of domain experts will be very heavy considering the huge possible audience. Since ECG records are mainly collected from people attending physical examinations, their diagnostic conclusions are likely to be “normal.” If computer-assisted ECG analysis algorithms filter out most normal records and then domain experts only focus on interpreting the remaining abnormal ones, that is, man-machine integration [1], the diagnostic efficiency will be greatly increased and the social benefits will be significant. The key technical indicator is that the detection precision of normal records must be at least 90% for long-time ECG monitoring and 95% for short-term ECG recording [2].

As heartbeat classification is an important step in computer-based ECG analysis regardless of application scenarios, a considerable amount of research has been dedicated to this subject. In general, the processing flow of these works is such that feature vectors, including physiology characteristics with diagnostic value (such as RR interval and PR interval) [3] and statistical characteristics (such as wavelet transform [4] and independent component analysis [5]), are extracted from heartbeat segments first, and feature selection [6] is conducted when necessary. Afterwards, a number of machine learning algorithms are employed for classification, such as support vector machine (SVM) [7] and Neural Network [8]. The relevant literature can be categorized into two different types based on the adopted evaluation scheme, namely, “intra-patient” and “inter-patient” [9].

The intra-patient evaluation has been adopted by extensive literature and its main characteristic is that the training and testing sets contain heartbeat segments from the same patients. By this scheme that is not in conformity with clinical practice, classification results tend to be overly optimistic, since the human heartbeat can be used for identity recognition [10]. Using the MIT-BIH arrhythmias (MIT-BIH-AR) database [11] and the Advancement of Medical Instrumentation (AAMI) recommendation [12], de Chazal et al. [9] proposed the inter-patient evaluation where the training and testing sets are constructed by heartbeat segments from different ECG records so that inter-individual variation can be taken into account. This scheme adopted by recent works [5, 6, 8, 13, 14] can evaluate the clinical performance of heartbeat classification algorithms in a relatively effective manner.

Note that heartbeat classification is just an intermediate step and what domain experts are concerned about is the classification result of an ECG record provided by computer-assisted ECG analysis algorithms. However, due to limited standard ECG databases, there are less research works relevant to this subject compared with heartbeat classification [15–17]. For the MIT-BIH-AR database, there are a total of 48 two-lead ECG records with approximately 30 min in duration, each containing heartbeat annotations for both R-peak position and disease type. The lead configurations (sensor locations) are not the same at all and only 40 records have both II and VI. Based on this database, we cannot use the obtained classifier for telemedicine

centers. For the Common Standards for Quantitative ECG (CSE) database [18], there are approximately 1000 standard 10-second, 12-lead ECG records, each containing annotations for P, QRS and T positions. However, this database is not free, which makes many scholars unable to carry out research on it. For other databases (such as American Heart Association (AHA) database [19] and ST-T database [20]), there are also some advantages and disadvantages.

Aiming to carry out relevant research for telemedicine centers, we constructed the Chinese Cardiovascular Disease Database (CCDD) [21] containing 193,690 standard 12-lead ECG records with about 10–20 s in duration. As shown in **Figure 1**, the ECG record consists of 6 limb leads of I, II, III, aVR, aVL, aVF and six chest leads of V1, V2, V3, V4, V5, V6 that describe cardiac electrical activity jointly. If it is collected for 10 s at the sampling frequency of 500 Hz, there are 60,000 ($=500 \times 10 \times 12$) sampling points.

In this study, we develop a method for classification of normal and abnormal ECG records with short duration (record classification), but one that can also be easily extended to other cases since a long-term ECG record can be divided every T s into segments of length T [12]. Based on the CCDD, our research group has tried and proposed some methods for this subject. Using existing methods for heartbeat classification [22, 23] and the average-prediction technology, Zhu [24] obtained an accuracy of 72% with a specificity of 94% and a sensitivity of 25% when testing 11,760 ECG records. Wang [25] proposed a Multi-Way Tree Based Hybrid Classifier (MTHC), including analysis of RR interval, similarity analysis of QRS complex, analysis of physiology characteristics (such as P wave, T wave and PR interval) and statistical learning based on morphological and numerical characteristics. The specificity, sensitivity and accuracy are 54.62, 95.13 and 72.49%, respectively, on the 140,098 ECG records. Zhang [13] proposed a heartbeat classification method which has certain advantages over the relevant literature [5, 6, 8, 9, 14], but the accuracy is about 50–60% when he used it for record classification. For this, our prior work [26] analyzed traditional feature-based methods from the ability to construct nonlinear functions and proposed lead convolutional neural networks (LCNNs) for multi-lead ECGs. Using the explicit training method and the single-point-prediction technology, we

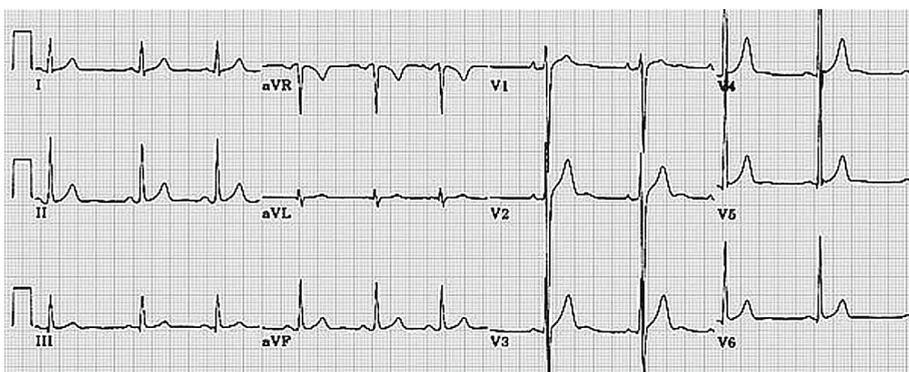


Figure 1. Standard 12-lead synchronous ECG.

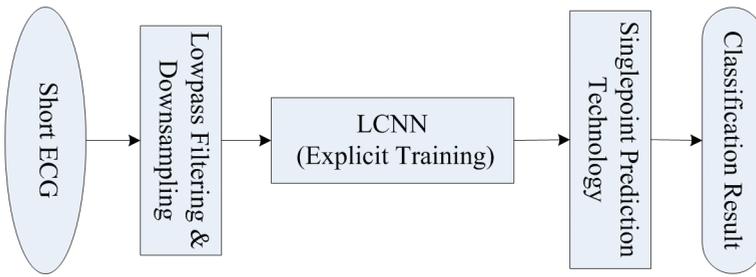


Figure 2. LCNN-based method for record classification.

achieved an accuracy of 83.66% with a specificity of 83.84% and a sensitivity of 83.43% when testing 151,274 records. **Figure 2** depicts the whole process flow.

To improve the classification performance further, this study extends our prior work in the following aspects. (1) Two different LCNN models are obtained through different filtering methods (a low-pass filter and a band-pass filter) and different training methods (the explicit method and the implicit method), and then the multipoint-prediction technology and the Bayesian fusion method are successively applied to them. (2) Four effective disease rules based on R peaks are developed for further analysis. (3) The final classification result is determined by utilizing the bias-average method.

The rest of the chapter is organized as follows: in Section 2, the CCDD used in this study is first introduced; in Section 3, the proposed method is described in detail; then in Section 4, the experimental results, as well as comparison with the published results are presented; discussions and conclusions are provided in Section 5 and 6, respectively.

2. Dataset

There are 193,690 standard 12-lead ECG records with about 10–20s in duration and sampling frequency of 500 Hz in the CCDD. These data were successively obtained from hospitals (i.e., real clinical environment) located at different districts in Shanghai, Suzhou and Changsha. For data1–251, all the records have detailed heartbeat annotations including P-onset, P-peaks, P-offset, QRS-onset, R-peaks, QRS-offset, T-onset, T-peaks, T-offset and disease type. For data252–943, all the records only have heartbeat annotations for R-peak position and disease type. Just like the MIT-BIH-AR database, all the records in data1–943 were chosen purposefully and two or more cardiologists provided the annotations. As heartbeat annotations are available, we can carry out research on heartbeat classification. For data944–193,690, all the records only have diagnostic conclusions provided by one cardiologist. Note that a diagnostic conclusion may contain more than one disease type.

As shown in **Figure 3**, we use hexadecimal codes, that is, form of “0×dddddd”, to encode disease types divided into three grades. There are 12 first-grade types (i.e., Invalid ECG,

0x00	Invalid ECG	
0x01	Normal ECG	
0x02	Sinus Rhythm	//The first-grade type
0x0201	Sinus Rhythm	//The second-grade type
0x020101	Sinus Rhythm	//The third-grade type
0x020102	Normal Sinus Rhythm	
0x020103	Heart Rate is Uncertain, May be Sinus Rhythm	
0x020104	Heart Rate is Uncertain, May be Sinus Irregularity	
	...	
0x0202	Sinus Irregularity	
	...	
0x03	Atrial Arrhythmia	
0x04	Junctional Rhythm	
0x05	Ventricular Arrhythmia	
0x06	Conduction Block	
0x07	Atrial Hypertrophy	
0x08	Ventricular Hypertrophy	
0x09	Myocardial Infarction	
0x0a	ST-T Change	
0x0b	Other Abnormalities	

Figure 3. Disease types in the CCDD.

Normal ECG, Sinus Rhythm, Atrial Arrhythmia, Junctional Rhythm, Ventricular Arrhythmia, Conduction Block, Atrial Hypertrophy, Ventricular Hypertrophy, Myocardial Infarction, ST-T Change and Other Abnormalities), 72 second-grade types and 335 third-grade types, including all the possible diagnostic conclusions provided by cardiologists in clinic. More details can be seen on our website (<http://58.210.56.164:88/ccdd/>).

In telemedicine centers, ECG records that are not explicitly diagnosed as “normal” all need to be further interpreted by domain experts, so that the potential cardiovascular disease of a patient can be detected as early as possible. Therefore, in this study we only regard ECG records whose diagnostic conclusions are “0x01” or “0x020101” or “0x020102” as “normal” (denoted as 0-class) and all the others as “abnormal” (denoted as 1-class). Moreover, we throw away some exception data, that is, the diagnostic conclusion is “0x00” or the duration is less than 9.625 s.

The training and testing sets are organized as follows [26]: “data944–25,693” is partitioned into three parts where the numbers of training samples, validation samples and testing samples (i.e., the small-scale testing set) are 12,320, 560 and 11,789, respectively, and the large-scale testing set is composed of 151,274 ECG records from data25694–179,130. Note that we will combine training samples and validation samples together for implicit training. **Table 1**

#	Dataset	Normal	Abnormal	Irregular	Total
Training samples	data 944–25,693	8800	3520	763	12,320
Validation samples	data 944–25,693	280	280	56	560
Small-scale testing set	data 944–25,693	8387	3402	822	11,789
Large-scale testing set	data 25694–179,130	85,141	66,133	9172	151,274

Table 1. Data distribution.

summarizes the detailed information, where “irregular” denotes the abnormal ECG records whose heart rhythms are irregular (their diagnostic conclusions are “0×0202”).

3. Methodologies

Figure 4 shows the overall framework of the proposed method. As we can see, it consists of three parts, namely, statistical learning, rule inference as well as summarizing. In the part of statistical learning, the ECG record is first preprocessed by two different methods respectively, and then two probability values are outputted by utilizing LCNNs and the multipoint-prediction technology (LCNN[A] and LCNN[B] have the same architecture and are obtained by “explicit training” and “implicit training”, respectively). After that, we use the Bayesian fusion method to incorporate the two outputs. In the part of rule inference, all R-peak positions in the ECG record are detected first [27], and then four disease rules based on them are used for further analysis. Finally, in the part of summarizing, we utilize the bias-average method to determine the classification result, that is, “normal” or “abnormal.”

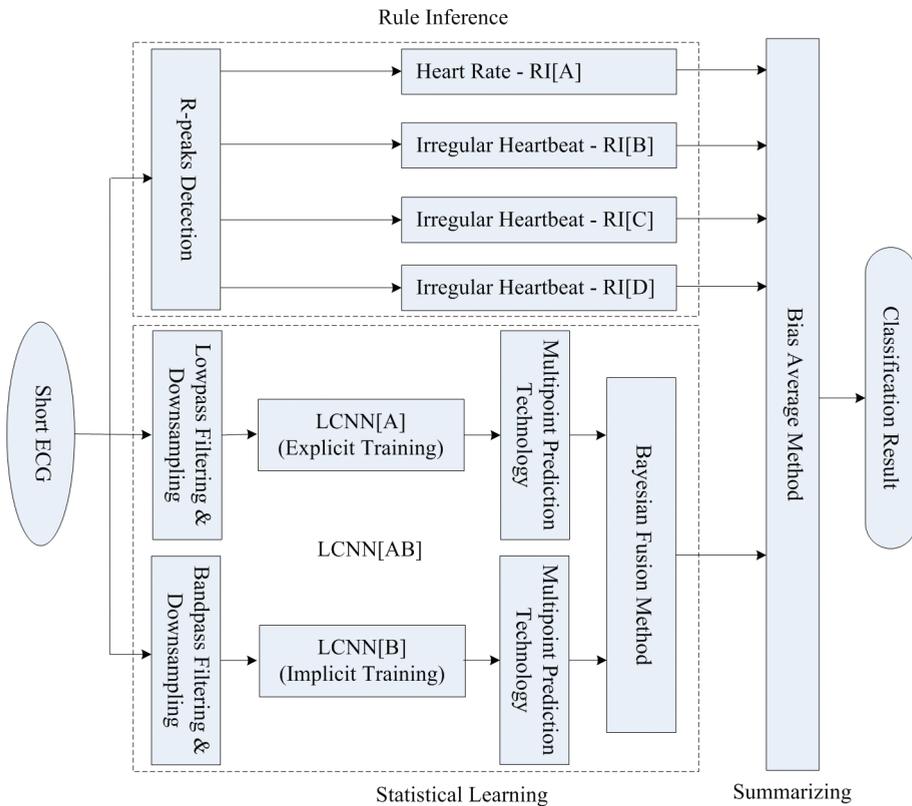


Figure 4. Overall framework of the proposed method.

Essentially, **Figure 4** describes an ensemble method including two homogeneous classifiers (i.e., LCNN[A] and LCNN[B]) and five heterogeneous classifiers (i.e., RI[A], RI[B], RI[C], RI[D], LCNN[AB] consisting of LCNN[A] and LCNN[B]). For the homogeneous ensemble, different base classifiers must be obtained if we want to be able to enhance classification performance. In fact, our generation strategy has certain advantages over some well-known ensemble methods such as Bagging and AdaBoost [28]. For the heterogeneous ensemble, different classifiers should be complementary to each other. Since LCNNs are not good at detecting abnormal heart rate and irregular heart rhythm, we use simple disease rules. Next, we will introduce each of the steps in detail.

3.1. Statistical learning

3.1.1. Preprocessing

Generally speaking, ECG records collected in clinics are often contaminated by several types of interfering noise, such as low-frequency waves (less than 0.5 Hz) caused by breathing movements, high-frequency waves (50 or 60 Hz) caused by electricity, and biological waves (about 33 Hz) caused by physical activities. Although we can use many methods to de-noise, some useful information may be lost after doing that. Since LCNNs which belong to deep learning [29, 30] have the ability to capture useful information but ignore interfering noise after learning from a certain number of training samples, we do not perform special de-noising.

An effective strategy for homogeneous ensemble learning is to make input data different, so we apply a low-pass filter and a band-pass filter of 0.5–40 Hz [31] to the ECG record, respectively. Of course, there is no problem if we exchange the low-pass filter for the band-pass filter; here we just make one path consistent with our prior work. After that, we conduct the down-sampling operation (from original 500 to 200 Hz) and extract a data segment of 9.5 s from the incoming ECG record after ignoring the first 0.125 s. Only eight basic leads, namely II, III, V1, V2, V3, V4, V5 and V6 are reserved since the remaining four leads can be linearly derived from them. As a result, each ECG record consists of 8×1900 sampling points.

3.1.2. Lead convolutional neural network

Convolutional neural networks (CNNs) [32] have been successful in the field of ECG signal classification. Zhu used the CNN-based method for both heartbeat classification and record classification, and obtained an accuracy of 99.20% on the MIT-BIH-AR database with 47,190 heartbeats in the “intra-patient” evaluation, and an accuracy of 83.49% and 0.8819 AUC on the CCDD with 11,760 ECG records, respectively [24]. Kiranyaz applied CNNs on 1-lead ECGs for patient-specific heartbeat classification and achieved good results [33, 34]. However, multi-lead ECGs are different from two-dimensional images: data in the horizontal direction (intra-lead) are relevant while data in the vertical direction (inter-lead) are independent. For this, our prior work proposed lead convolutional neural networks (LCNNs) for multi-lead ECGs, which have better classification performance. **Figure 5** shows an example of three-stage LCNN.

In **Figure 5**, CU denotes a convolution unit consisting of a convolutional layer and a sub-sampling (max-pooling) layer; 1DCov and SubSamp denote the one-dimensional convolution

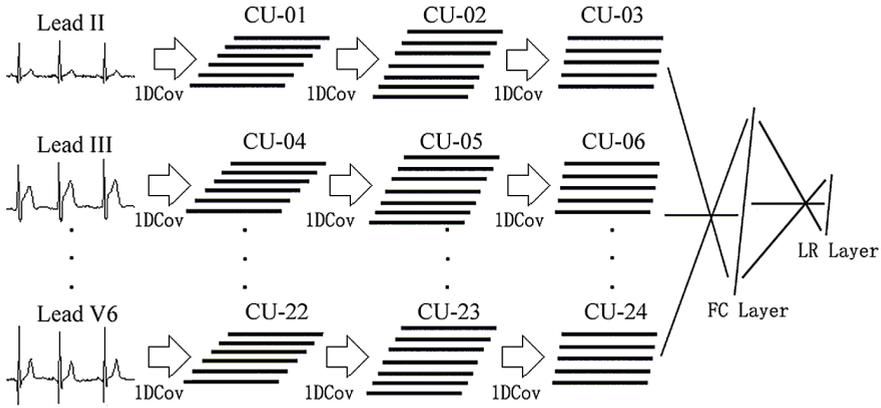


Figure 5. Architecture of a three-stage LCNN.

operation and the sub-sampling operation, respectively. The computational process consists of data for each lead going through three different CUs, and then inputting information from all leads into a fully connected (FC) layer. Finally, the logistic-regression (LR) layer output the predictive value. In fact, we can regard each CU as a feature extractor, and the subsequent multilayer perceptron consisting of the FC layer and the LR layer as a classifier.

To describe the LCNN clearly, we present its computational formula as follows: let $[x_1, x_2, \dots, x_8]$ be the incoming ECG record where x_i (a vector, $1 \leq i \leq 8$) is the data of the i -th lead, then

$$\begin{cases} f(x) = g_E \left(g_D \left(\bigcup_{i=1}^8 g_{C_i} \left(g_{B_i} \left(g_{A_i} (x_i) \right) \right) \right) \right) \\ g_D(x) = \varphi(W^T x + b) \\ g_E(x) = \frac{1}{1 + e^{-(W^T x + b)}} \end{cases} \quad (1)$$

Here, g_E and g_D are the computational formulas for the LR layer and the FC layer, respectively, W and b are the weights and the biases of the corresponding layer, $\varphi(x)$ is the sigmoid function. g_{A_i} , g_{B_i} and g_{C_i} ($1 \leq i \leq 8$) are the computational formulas for the CUs whose expression are all $f_{sub}(f_{cov}(x))$, the only differences between them exist in weights and biases. For the i -th lead, ECG data is inputted into g_{A_i} first, and then the output of g_{A_i} continues to be inputted into g_{B_i} and g_{C_i} . After that, the outputs of g_{C1} , g_{C2} , ..., g_{C8} (namely 8 vectors) are concatenated together (namely 1 vector) using the union operation \cup . Finally, the resulting vector is inputted into g_D and g_E successively and a value ranging from 0 to 1 is outputted. Note that both the input and the output of each function are vectors and the expressions of $f_{sub}(x)$ and $f_{cov}(x)$ are given by

$$\left\{ \begin{array}{l} f_{\text{cov}}(v) = \bigcup_{j,k} v_{ij}^k = \bigcup_{j,k} \left(b_{ij} + \sum_m \sum_{p=0}^{P_i-1} w_{ij,(i-1)m}^p v_{(i-1)m}^{(k+p)} \right) \\ f_{\text{sub}}(v) = \bigcup_{j,k} v_{ij}^k = \bigcup_{j,k} \varphi \left(\max \left(\bigcup_{\substack{q=(k-1)Q_i+1, \\ (k-1)Q_i+2, \\ \dots \\ kQ_i}} v_{(i-1)j}^q \right) \right) \end{array} \right. \quad (2)$$

Here, P_i and Q_i denote the size of the convolutional kernel and the sub-sampling step in the i -th layer, respectively, \bigcup denotes the union operation. v_{ij}^k denotes the output value of the k -th neuron (start from 1) in the j -th feature map (unit) of the i -th layer. $w_{ij,(i-1)m}^p$ is the weight that connects the j -th unit of the i -th layer with the m -th unit of the $(i-1)$ -th layer, and b_{ij} is the bias in the j -th unit of the i -th layer.

The setting of each parameter (such as P_i , Q_i and the number of CUs for each lead) will greatly influence the classification performance, and too large or too small parameters all result in unfavorable outcome. In practical applications, they can be determined by a trial-and-error method. To keep things simple and ensure a fair comparison, we use two 3-stages LCNNs with the same architecture, namely LCNN[A] and LCNN[B] in **Figure 4**, and the corresponding parameters are the same as those in our prior work. The number of the neurons in the input layer is 8×1700 ; the sizes of three convolutional kernels are 21, 13 and 9, respectively; the sizes of three sub-sampling steps are 7, 6, and 6, respectively; the numbers of three feature maps are 6, 7 and 5, respectively; the number of the neurons in the FC layer is 50.

3.1.3. Training method

To obtain different LCNN models, we develop two training methods for LCNN based on mini-batch gradient descent [35] in supervised learning, that is, the explicit method and the implicit method. The main difference between them exists in validation mechanisms. Since gradient descent is a typical local search algorithm, we incorporate “translating starting point” and “adding noise” into it to increase the number of training samples presented to the network, so that classifiers with good generalization performance can be obtained.

The explicit method is commonly used for training neural networks, which utilizes independent validation samples to evaluate the obtained classifier during the training phase. As shown in **Figure 6(a)**, the training process can be described as follows: a random 8×1700 local segment is extracted from the 8×1900 training sample first (translating starting point), and then it is added with random signals whose maximal amplitude is less than 0.15 millivolts (adding noise) with high probability, finally back propagation is invoked. When Batchsize training samples have been presented to the network, the weights will be adjusted and the current LCNN model will be tested by the particular 8×1700 local segments (start from the first position) extracted from 8×1900 validation samples. If the accuracy is the best up to the

```

[Algorithm]:   Explicit Training Method
[Input]:      Training Samples X={xt[i],yt[i]}, i=1,2,...,N1}, Validation Samples V={xv[i],yv[i]}, i=1,2,...,N2}
[Output]:     An LCNN Model
BestAcc = 0
while(!MaxEpoch)
  dW = {0,0,...,0} //Initialize the weight changes
  for i = 1 to N1 //Make sure that N1 is divisible by BatchSize
    xt1[i] = PreprocessA(xt[i]) //Lowpass filtering & downsampling
    xt2[i] = LocalSegment(xt1[i], rand) //Extract a 8 × 1700 local segment starting from a random position
    xt3[i] = AddRandSignal(xt2[i], rand) //Add random signals with high probability
    dW = dW + BackPropagation(xt3[i], yt[i]) //Invoke Backpropagation and accumulate dW
  if (i % BatchSize == 0) //BatchSize training samples have been presented to the network
    UpdateNN(dW) //Adjust the weights
    dW = {0,0,...,0}
    CMatrix = [0,0;0,0] //Initialize the confusion matrix
    for j = 1 to N2
      xv1[j] = PreProcessA(xv[j]) //Preprocessing
      xv2[j] = LocalSegment(xv1[j], 1) //Extract a 8 × 1700 local segment starting from the 1st position
      CMatrix += Test(xv2[j], yv[j]) //Test the current LCNN model
    end
    if(Accuracy(CMatrix) > BestAcc)
      BestAcc = Accuracy(CMatrix)
      SaveNN(Model)
    end
  end
end
end
end

```

a

```

[Algorithm]:   Implicit Training Method
[Input]:      Training Samples X={xt[i],yt[i]}, i=1,2,...,N}
[Output]:     An LCNN Model
BestAcc = 0
while(!MaxEpoch)
  dW = {0,0,...,0} //Initialize the weight changes
  CMatrix = [0,0;0,0] //Initialize the confusion matrix
  for i = 1 to N //Make sure that N is divisible by BatchSize
    xt1[i] = PreprocessB(xt[i]) //Bandpass filtering & downsampling
    xt2[i] = LocalSegment(xt1[i], rand) //Extract a 8 × 1700 local segment starting from a random position
    xt3[i] = AddRandSignal(xt2[i], rand) //Add random signals with high probability
    dW = dW + BackPropagation(xt3[i],yt[i]) //Invoke Backpropagation and accumulate dW
  if (i % BatchSize == 0) //BatchSize training samples have been presented to the network
    UpdateNN(dW) //Adjust the weights
    dW = {0,0,...,0}
    for j = 1 to BatchSize //Samples used between two adjacent weight-updating processes
      xt1[j] = PreprocessB(xt[i])
      xt2[i-j+1] = LocalSegment(xt1[i-j+1],1) //Extract a 8 × 1700 local segment starting from the 1st position
      CMatrix += Test({xt2[i-j+1],yt[i-j+1]}) //Test the current LCNN model
    end
  end
  end
  if(Accuracy(CMatrix) > BestAcc)
    BestAcc = Accuracy(CMatrix)
    SaveNN(Model)
  end
end
end
end

```

b

Figure 6. (a). Pseudo-code of explicit training method. (b). Pseudo-code of implicit training method.

present, the current LCNN model will be saved. Of course, the training will stop when reaching the maximum number of epochs.

Besides the explicit method, we develop an implicit method for training LCNN, whose main characteristic is that a small number of local segments extracted from training samples are

used for evaluating the obtained model. As shown in **Figure 6(b)**, most of the steps are the same as those shown in **Figure 6(a)**. When Batchsize training samples have been presented to the network, the weights will be adjusted and the current LCNN model will be tested by the particular 8×1700 local segments (start from the first position) extracted from the 8×1900 training samples, which are used between two adjacent weight-updating processes. At the end of each epoch, the current LCNN model will be saved if the total accuracy is the best up to the present.

While some may think this method would result in overfitting, this is a false assumption. We extract random local segments and add random signals to them with high probability for training. Differently, we extract particular local segments and do not add random signals to them for validation. Therefore, the probability of overlap is very small.

To be consistent with our prior work, we employ the back propagation algorithm of inertia moment and variable step [36] and do not use any unsupervised pre-training. The relevant parameters are set as follows: the initial step length is 0.02, the step decay is 0.01 except for the second and the third epoch (set as 0.0505), Batchsize is 560, and the maximum number of training epochs is 500.

3.1.4. Multipoint prediction technology

We can immediately get the classification result after inputting a 8×1700 local segment extracted from the incoming ECG record into the obtained LCNN model. This is the method our prior work adopted, namely the single-point-prediction technology. However, the output of the LCNN is a probability value ranging from 0 to 1 and the classification confidence is low if the value is around 0.5. For this, we develop a new testing method, namely the multipoint-prediction technology.

As shown in **Figure 7**, nine 8×1700 local segments which start from the 1st, 26th, 51st, 76th, 101st, 126th, 151st, 176th and 201st positions are extracted from the ECG record, respectively, then their predictive values outputted by the LCNN are aggregated by the average rule (for simplicity's sake), so that the classification confidence can be enhanced.

3.1.5. Bayesian fusion method

In this study, we employ a Bayesian fusion method [37] to incorporate the outputs of the two LCNNs: given M classifiers and K classes, the predicted class k can be determined based on the final probability estimates $P(y = i | c_1, c_2, \dots, c_m)$ and is given by

$$\begin{cases} P(y = i | c_1, c_2, \dots, c_m) = \frac{1}{M} \sum_{m=1}^M P(y = i | c_m) \\ k = \arg \max_{1 \leq i \leq K} \{P(y = i | c_1, c_2, \dots, c_m)\} \end{cases} \quad (3)$$

Here, $P(y = i | c_m)$ is the probability value predicted on the i -th class by the classifier c_m . What we focus on is a binary classification problem (i.e., "normal" vs. "abnormal"), thus $K = 2$, and

```

[Input]:   Testing Samples X={x[i],1<=i<=N}, LCNN[A] and LCNN[B]
[Output]:  Classification Result
for i = 1 to N
  //Multipoint Rrediction Technology
  AvgPb1 = 0
  AvgPb2 = 0
  x1[i] = PreprocessA(x[i])           //Lowpass Filtering & Downsampling
  x2[i] = PreprocessB(x[i])           //Bandpass Filtering & Downsampling
  for j = 1 to 9
    start = (j-1) * 25 + 1           //9 starting position
    x1s[i] = LocalSegment(x1[i], start) //Extract a 8×1700 local segment starting from the given position
    x2s[i] = LocalSegment(x2[i], start)
    AvgPb1 = AvgPb1 + TestA(x1s[i])   //Use LCNN[A] as a classifier
    AvgPb2 = AvgPb2 + TestB(x2s[i])   //Use LCNN[B] as a classifier
  end
  AvgPb1 = AvgPb1/9
  AvgPb2 = AvgPb2/9

  //Bayesian Fusion Method
  nnc = (AvgPb1 + AvgPb2)/2

  //Bias Average Method
  if (max(ro1,ro2,ro3,ro4) == 1) //ro1,ro2,ro3,ro4 are the outputs of four rule-based classifiers
    PbEm = (max(ro1,ro2,ro3,ro4) + nnc)/2
  else
    PbEm = nnc
  End

  //Determine the classification result
  if (PbEm < 0.5)
    //x[i] is identified as "normal"
  else
    //x[i] is identified as "abnormal"
  end
end

```

Figure 7. Pseudo-code of testing process.

obviously, $M = 2$. Of course, other fusion methods can be employed for this purpose. For example, we can train one logistic regression using the concatenation of the outputs of both LCNNs, and it may be able to give better results. The part of statistical learning is denoted as the LCNN[AB] classifier.

3.2. Rule inference

R-peaks detection algorithms have reached a very high level at present [38]. On the other hand, there is no doubt that an ECG record is abnormal if it is identified as a specific disease. Hence, we develop four disease rules based on R peaks to detect abnormal heart rate and irregular heart rhythm.

Let fs be the sampling frequency, R_i ($1 < i < n$) be the i -th R-peak position in an ECG record, and $std(.)$ be the function of calculating the standard deviation. The formula for calculating the average RR interval is given by

$$AvgRR = \frac{1}{n-1} \sum_{i=2}^n (R_i - R_{i-1}) \quad (4)$$

The four disease rules are defined as follows:

1. Heart rate

The formula for calculating heart rate is given by

$$HR = \frac{60 \times fs \times (n - 1)}{R_n - R_1} \tag{5}$$

Normal heart rate is defined as 60–100 CPM (i.e., counts per minute) clinically; here we regard 59–101 CPM as normal, otherwise as abnormal. This rule is denoted as the RI[A] classifier.

2. Irregular heart rhythm based on local characteristics.

The first rule for detecting irregular heart rhythm is defined as follows: three successive RR intervals exceed 15% of the average RR interval, that is, $\exists k \in [1, n - 3], \forall j \in [1, 3]$ satisfies

$$\left| \frac{(R_{k+j} - R_{k+j-1}) - AvgRR}{AvgRR} \right| > 0.15 \tag{6}$$

We regard an ECG record as abnormal if Eq. (6) is satisfied, otherwise as normal. This rule is denoted as the RI[B] classifier.

3. Irregular heart rhythm based on local and global characteristics.

The second rule for detecting irregular heart rhythm is defined as follows: one RR interval exceeds 15% of the average RR interval and the standard deviation of the rates between two neighboring RR intervals is greater than 0.05, that is, $\exists k \in [1, n - 1]$ satisfies

$$\left\{ \begin{array}{l} \left| \frac{(R_{k+1} - R_k) - AvgRR}{AvgRR} \right| > 0.15 \\ std \left(\left\{ \frac{R_{i+2} - R_{i+1}}{R_{i+1} - R_i} \mid 1 \leq i \leq n - 2 \right\} \right) > 0.05 \end{array} \right. \tag{7}$$

Likewise, we regard an ECG record as abnormal if Eq. (7) is satisfied, otherwise as normal. This rule is denoted as the RI[C] classifier.

4. Irregular heart rhythm based on global characteristics.

The last rule for detecting irregular heart rhythm is defined as follows: the standard deviation of RR intervals is greater than $(0.05 \times AvgRR)$, that is, satisfies

$$std(\{R_{i+1} - R_i \mid 1 \leq i \leq n - 1\}) > 0.05 \times AvgRR \tag{8}$$

In the same way, we regard an ECG record as abnormal if Eq. (8) is satisfied, otherwise as normal. This rule is denoted as the RI[D] classifier.

Of course, false R-peak detections will influence the classification performance of the four disease rules and result in an unfavorable outcome. In this study, we utilize the Zhu method

[27] to detect R peaks since its accuracy is 99.83% in the MIT-BIH-AR database with 48 ECG records and 99.78% in the CCDD with 251 ECG records, respectively. Note that 0 is outputted if the classification result is “normal”, otherwise 1 is outputted.

3.3. Summarizing

As we can see, there are five classifiers, namely RI[A], RI[B], RI[C], RI[D] and LCNN[AB], involved in **Figure 4**. We can regard an ECG record as “abnormal” if any one of the classifiers gives such a result. However, the outputs of the four rule-based classifiers are 0 or 1, while the output of the LCNN[AB] classifier is a probability value ranging from 0 to 1. We hope that the predictive value can be outputted in the form of a probability, so a new fusion method, namely the bias-average method, is developed, given by

$$\begin{cases} ro = \max(ro_1, ro_2, ro_3, ro_4) \\ output = \begin{cases} \frac{ro + nno}{2}, & ro == 1 \\ nno, & otherwise \end{cases} \end{cases} \quad (9)$$

Where, ro_1, ro_2, ro_3, ro_4 and nno are the outputs of the RI[A], RI[B], RI[C], RI[D] and LCNN [AB] classifiers, $\max(\cdot)$ is to choose the maximum value. If we use the first three disease rules, just replace “ $\max(ro_1, ro_2, ro_3, ro_4)$ ” with “ $\max(ro_1, ro_2, ro_3)$.” The final output value ranges from 0 to 1 and the classification result is “normal” if it is less than 0.5, otherwise it is “abnormal.” **Figure 7** shows the whole testing process including the multipoint-prediction technology, the Bayesian fusion method and the bias-average method.

4. Result

4.1. Performance metrics

In this study, we use several metrics to investigate the performance of different algorithms: Sp (specificity), NPV (negative predictive value), Se (sensitivity) and Acc (accuracy), given by

$$\begin{cases} Sp = \frac{TN}{TN + FP} \\ NPV = \frac{TN}{TN + FN} \\ Se = \frac{TP}{TP + FN} \\ Acc = \frac{TP + TN}{TN + FP + FN + TP} \end{cases} \quad (10)$$

Here, TP is the number of abnormal ECG records that are correctly classified as abnormal, TN is the number of normal ECG records that are correctly classified as normal, FP is the number of normal ECG records that are incorrectly classified as abnormal, and FN is the number of

abnormal ECG records that are incorrectly classified as normal. We also use receiver operating characteristic (ROC) curve-related metrics, including area under ROC curve (AUC), true positive value (TPR) (e.g., “Sp” in this study) and false positive value (FPR)(e.g., “1-Se” in this study). A rough suggestion for measuring classification performance using “AUC” is as follows [39]:

1. 0.90–1.00 = excellent classification;
2. 0.80–0.90 = good classification;
3. 0.70–0.80 = fair classification;
4. 0.60–0.70 = poor classification;
5. 0.50–0.60 = failure.

Note that “NPV” in this study is also called the detection precision of normal ECG records.

4.2. Numerical experiment

Our prior work [26] has achieved the best results for record classification on the CCDD up to now, so it is meaningful to compare the proposed method with it. As described in Section 2, the same training, validation and testing samples are used. To show the contribution of each strategy intuitively, we present the corresponding results in **Tables 2** and **3** in turn. The classification results on the small-scale testing set is only used for reference; what we mostly focus on is the results on the large-scale testing set, which can be deemed as a more realistic estimate of potential performance in real applications.

In **Tables 2** and **3**, “0”, “1”, “2”, “3” and “4” denote the LCNN[AB], RI[A], RI[B], RI[C] and RI[D] classifiers, respectively. From the results, we can see that although NPV and Se are slightly lower, “0” outperformed our prior work in all other metrics regardless of the small-scale testing set and the large-scale testing set. On this basis, most metrics continue to increase when “1”, “2” and “3” are added, while many metrics start to decrease when “4” is added. The role of “4” is to increase NPV and Se, so that TPR under the condition of NPV being equal to

Model	Sp	NPV	Se	Acc	AUC	FPR = 1%		NPV = 95%		NPV = 90%	
						TPR	NPV	FPR	TPR	FPR	TPR
LCNN [26]	88.84	90.48	76.95	85.41	0.9034	17.5	97.8	8.22	63.3	24.7	90.1
0	91.81	90.04	74.96	86.95	0.9172	23.0	98.3	9.61	74.1	25.2	91.9
0 + 1	91.77	90.11	75.16	86.98	0.9176	23.3	98.3	9.63	74.2	25.2	91.9
0 + 1 + 2	91.62	90.22	75.51	86.97	0.9180	23.3	98.3	9.63	74.2	25.2	91.9
0 + 1 + 2 + 3	90.28	91.11	78.28	86.82	0.9220	24.6	98.4	9.90	76.3	25.3	92.3
0 + 1 + 2 + 3 + 4	80.60	93.09	85.24	81.94	0.9196	25.8	98.5	9.71	74.8	25.1	91.7

Table 2. The results of different classification models on the small-scale testing set.

Model	Sp	NPV	Se	Acc	AUC	FPR = 1%		NPV = 95%		NPV = 90%	
						TPR	NPV	FPR	TPR	FPR	TPR
LCNN [26]	83.84	86.69	83.43	83.66	0.9086	16.0	95.4	1.81	26.7	10.6	73.9
0	87.49	86.30	82.12	85.15	0.9199	22.1	96.6	3.42	50.4	11.3	79.3
0 + 1	87.43	86.75	82.81	85.41	0.9225	24.4	96.9	3.60	53.1	11.5	80.4
0 + 1 + 2	87.20	87.30	83.67	85.66	0.9251	26.2	97.1	3.83	56.5	11.7	81.6
0 + 1 + 2 + 3	86.03	89.11	86.46	86.22	0.9322	34.3	97.8	4.43	65.4	12.1	84.4
0 + 1 + 2 + 3 + 4	80.37	91.04	89.82	84.50	0.9320	36.7	97.9	4.60	67.9	11.8	82.4

Table 3. The results of different classification models on the large-scale testing set.

95% (TPR95) can be increased. Compared with our prior work, the “0 + 1 + 2 + 3” model increased all the metrics and significantly improved the classification performance.

As mentioned previously, the aim of this study is to develop a computer-assisted ECG analysis algorithm for telemedicine centers. Specially, the normal ECG records are filtered out first, and then the remaining abnormal ones are delivered to domain experts for further interpretation, so that their workload can be lessened and diagnostic efficiency can be improved [1]. The key technical indicator is to make TPR95 as high as possible [2]. From **Tables 1** and **3** we can see that, 38.21% ($=56.28\% \times 67.9\%$) of domain experts’ workload can be lessened. The more normal ECG records there are, the less work domain experts do. Although the proposed method increases the computational complexity, we can use the classification system in practical applications. The total computing time for an ECG record is about 125 ms on an Intel Core2 CPU@2.93GHz, 2GB RAM, 32bit Window 7 OS.

Some may think that the self-compared results are not supportive enough to show the effectiveness. In fact, Zhu [24], Wang [25] and Zhang [13] have proposed methods for this subject, but their results are significantly inferior to those of our prior work. Since the heartbeat classification method [5] has achieved the highest accuracy of 99.3% in the “intra-patient” evaluation by now and a good accuracy of 86.4% in the “inter-patient” evaluation, we re-implement it using the CCDD and classify each ECG record with any abnormal heartbeats as “abnormal”, but the performance is poor. The open-source software ECG-KIT [11] has recently been made available on Physionet (<http://physionet.org/physiotools/ecg-kit/>). However, different from the classification standard in this study, ECG-KIT uses the AAMI recommendation to output the result, so it would not necessarily be meaningful to compare our proposed method with ECG-KIT.

Finally, let us consider the industrial applications. The statistical analysis on the classification results given by General Electric Medical System shows that the total accuracy on 2112 ECG records is 88.0%, where the accuracies of interpreting sinus rhythms and non-sinus rhythms are 95 and 53.5%, respectively [40]. Hence, the total accuracy will only be 74.25% if the ratio of the number of sinus rhythms to the number of non-sinus rhythms is 1:1. The similar statistical analysis also shows that on 576 ECG records, the accuracies of Philips Medical Systems and

Draeger Medical Systems are 80 and 75%, respectively, and the average accuracy of non-experts is 85% [41]. Our proposed method achieved an accuracy of 86.22% and 0.9322 AUC on 151,274 ECG records, indicating that it has competitive classification performance.

5. Discussion

Due to inter-individual variation in ECG characteristics and the complexity of clinical data, record classification is a highly difficult problem. As a new research direction in recent years, deep learning has already achieved great success in hard Artificial Intelligence tasks, such as speech recognition [42] and image classification [43]. As an improved deep learning architecture, the LCNN show good performance for record classification. However, one LCNN with limited layers and neurons is not strong enough, so we use an ensemble method based on LCNN in this study and the experimental results show its effectiveness. In fact, using the same training and testing samples (heartbeat segments) from the MIT-BIH-AR database, the LCNN-based ensemble method has achieved an accuracy of 99.46% with a specificity of 99.69% and a sensitivity of 98.73% for detecting ectopic heartbeat, comparable to the state-of-the art results for heartbeat classification in the “intra-patient” evaluation [5].

One may doubt the viability of LCNNs since no feature-extraction operations are involved. In fact, if we let $f(x)$ and $g(x)$ be the corresponding function of feature extraction and classification, respectively, the decision-making function of traditional machine-learning methods can be written as $g(f(x))$. Therefore, we can directly construct the function $g(f(x))$ by deep-learning techniques [26, 44]. Nevertheless, there is not a general-purpose method for all problems, and what one needs to do is to choose or develop the most appropriate method for a special problem. For instance, abnormal ECG records whose heart rhythms are irregular should be collected first if we use the LCNN to detect irregular heart rhythm. It works well; however, the LCNN will capture not only characteristics of heart rhythm, but also other characteristics such as morphology and intervals during the training phase. That will result in more types of ECG records to be collected if we want to obtain a good classifier. In fact, on the small-scale and large-scale testing sets, the sensitivities of the LCNN[AB] classifier in the detection of irregular heart rhythm are 58.15 and 58.54%, respectively, while the corresponding results for detecting abnormal ECG records whose heart rhythm are regular are 80.31 and 85.92%. From **Table 1**, we know that the two testing sets only have a small proportion of irregular heart rhythm. If there are more abnormal ECG records whose heart rhythm is irregular, the LCNN[AB] classifier will not achieve such good results. In a word, accurate detection of irregular heart rhythm plays an important role in enhancing classification performance. Fortunately, we can achieve this aim by running the calculation based on R-peak positions. It is also effective for detecting abnormal heart rate, thus four disease rules are developed in this study.

From the perspective of simulating cardiologists’ diagnostic thinking, LCNNs and disease rules express experiential knowledge and intuitive knowledge, respectively, and the combination of them is a complete simulation [1]. This is another reason why the proposed method can achieve good classification results.

In fact, our proposed method can be divided into two parts: the classification of “normal” versus “abnormal” serves as a global classifier and the classification of each specific disease serve as a local classifier. The advantage of the global classifier is that the error accumulation due to introducing each local classifier can be avoided, since the detection performance for many cardiovascular diseases is not very good regardless of whether LCNNs or disease rules are used. There are not enough samples used for training LCNNs in many cases (especially the disease with low detection rate in the general population); while for disease rules, the inaccurate detection of fiducial points can result in subsequent misclassification. Nevertheless, to enhance classification performance further, we can integrate classifiers of easily detectable diseases into **Figure 4**. For instance, using LCNNs for atrial fibrillation detection, the sensitivity, specificity and accuracy are 98.93, 98.76 and 98.76%, respectively, on 142,167 ECG records in the CCDD, and the detection performance is still excellent on ECG records collected by Shanghai MicroPort Co. Ltd. (almost 100% accuracy). Likewise, using disease rules, we can effectively identify QRS complexes with abnormal amplitude.

6. Conclusion

In order to lessen the workload of domain experts in telemedicine centers, we present a systematic approach for record classification in this chapter. Based on LCNNs and disease rules, an effective ensemble method including two homogeneous classifiers and five heterogeneous classifiers is developed. On the CCDD, our method yields an accuracy of 86.22% and 0.9322 AUC (excellent), which is a significant improvement on previously reported results [13, 24–26]. Specifically, TPR under the condition of NPV being equal to 95% can reach 67.9%, which means that the workload of domain experts can be decreased by $(N\% \times 67.9\%)$ in a clinically acceptable scope (especially in the BCMIS of China) if the percentage of normal ECG records is $N\%$. In general, at least $70 \times 67.9 = 47.53\%$ of domain experts’ workload can be reduced since $N\%$ can be more than 70%. We have deployed the classification system on the real-time cloud platform of Shanghai Aerial Hospital Network.

Regarding future work, our aim is to develop effective detection algorithms for other common cardiovascular diseases such as premature atrial contraction and premature ventricular contraction, and to integrate them into **Figure 4**. Since both LCNNs and disease rules have advantages and disadvantages, the combination of the two is a preferable research direction.

Author details

Linpeng Jin^{1,2} and Jun Dong^{1*}

*Address all correspondence to: jdong2010@sinano.ac.cn

1 Suzhou Institute of Nano-tech and Nano-bionics, Chinese Academy of Sciences, Suzhou, China

2 University of Chinese Academy of Sciences, Beijing, China

References

- [1] Dong J, Zhang JW, Zhu HH, et al. Wearable ECG monitors and its remote diagnosis service platform. *IEEE Intelligent Systems*. 2012;**6**(27):36-43. DOI: 10.1109/MIS.2012.4
- [2] Liu X. *Atlas of Classical Electrocardiograms*. 1st ed. Shanghai: Shanghai Science and Technology Press; 2011
- [3] de Chazal P, Reilly RB. A patient-adapting heartbeat classifier using ECG morphology and heartbeat interval features. *IEEE Transactions on Biomedical Engineering*. 2006;**53**(12): 2535-2543. DOI: 10.1109/TBME.2006.883802
- [4] Sumathi S, Beaulah HL, Vanithamani R. A wavelet transform based feature extraction and classification of cardiac disorder. *Journal of Medical Systems*. 2014;**38**(9):1-11. DOI: 10.1007/s10916-014-0098-x
- [5] Ye C, Kumar BV, Coimbra MT. Heartbeat classification using morphological and dynamic features of ECG signals. *IEEE Transactions on Biomedical Engineering*. 2012;**59**(10):2930-2941. DOI: 10.1109/TBME.2012.2213253
- [6] Llamedo M, Martinez JP. Heartbeat classification using feature selection driven by database generalization criteria. *IEEE Transactions on Biomedical Engineering*. 2011;**58**(3):616-625. DOI: 10.1109/TBME.2010.2068048
- [7] Osowski S, Hoa LT, Markiewicz T. Support vector machine-based expert system for reliable heartbeat recognition. *IEEE Transactions on Biomedical Engineering*. 2004;**51**(4):582-589. DOI: 10.1109/TBME.2004.824138
- [8] Mar T, Zaunseder S, Martinez JP, et al. Optimization of ECG classification by means of feature selection. *IEEE Transactions on Biomedical Engineering*. 2011;**58**(8):2168-2177. DOI: 10.1109/TBME.2011.2113395
- [9] de Chazal P, O'Dwyer M, Reilly RB. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE Transactions on Biomedical Engineering*. 2004;**51**(7):1196-1206. DOI: 10.1109/TBME.2004.827359
- [10] Biel L, Pettersson O, Philipson L, et al. ECG analysis: A new approach in human identification. *IEEE Transactions on Instrumentation and Measurement*. 2001;**50**(3):808-812. DOI: 10.1109/19.930458
- [11] Goldberger AL, Amaral L, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*. 2000;**101**(23):215-220. DOI: 10.1161/01.CIR.101.23.e215
- [12] *Testing and Reporting Performance Results of Cardiac Rhythm and ST Segment Measurement Algorithms*. ANSI/AAMI EC57:1998, Rev; 2012
- [13] Zhang ZC, Dong J, Luo XQ, et al. Heartbeat classification using disease-specific feature selection. *Computers in Biology and Medicine*. 2014;**46**:79-89. DOI: 10.1016/j. compbiomed. 2013.11.019

- [14] Escalona-Moran MA, Soriano MC, Fischer I, et al. Electrocardiogram classification using reservoir computing with logistic regression. *IEEE Journal of Biomedical and Health Informatics*. 2015;**19**(3):892-898. DOI: 10.1109/JBHI.2014.2332001
- [15] Bortolan G, Degani R, Pedrycz W. A fuzzy pattern matching technique for diagnostic ECG classification. In: *Proceedings of Computers in Cardiology; 25–28 September 1988; Washington, DC*. New York: IEEE; 1988. pp. 551-554
- [16] Willems JL, Abreu-Lima C, Arnaud P, et al. The diagnostic performance of computer programs for the interpretation of electrocardiograms. *The New England Journal of Medicine*. 1991;**325**(25):1767-1773. DOI: 10.1056/NEJM199112193252503
- [17] Rahman QA, Tereshchenko LG, Kongkatong M, et al. Utilizing ECG-based heartbeat classification for hypertrophic cardiomyopathy identification. *IEEE Transactions on Nanobioscience*. 2015;**14**(5):505-512. DOI: 10.1109/TNB.2015.2426213
- [18] Willems JL, Arnaud P, Bommel JHV, et al. A reference database for multilead electrocardiographic computer measurement programs. *Journal of the American College of Cardiology*. 1987;**10**(6):1313-1321. DOI: 10.1016/S0735-1097(87)80136-5
- [19] American Heart Association Database [Internet]. 1985. Available from: <http://www.physionet.org/physiobank/other.shtml> [Accessed: January 06, 2018]
- [20] Taddei A, Distanti G, Emdin E, et al. The European ST-T database: Standard for evaluating Systems for the Analysis of ST-T changes in ambulatory electrocardiography. *European Heart Journal*. 1992;**13**(9):1164-1172. DOI: 10.1093/oxfordjournals.eurheartj.a060332
- [21] Zhang JW, Liu X, Dong J. CCDD: An enhanced standard ECG database with its Management & Annotation Tools. *International Journal on Artificial Intelligence Tools*. 2012;**21**(5): 1-26. DOI: 10.1142/S0218213012400209
- [22] Minhas FU, Arif M. Robust electrocardiogram beat classification using discrete wavelet transform. *Physiological Measurement*. 2008;**29**(5):555-570. DOI: 10.1109/ICBBE.2008.796
- [23] Martis RJ, Chakraborty C, Ray AK. A two-stage mechanism for registration and classification of ECG using Gaussian mixture model. *Pattern Recognition*. 2009;**42**(11):2979-2988. DOI: 10.1016/j.patcog.2009.02.008
- [24] Zhu HH. Research on ECG recognition critical methods and development on remote multi body characteristic signal monitoring system [Thesis]. Beijing: University of Chinese Academy of Sciences; 2013
- [25] Wang LP. Study on Approach of ECG Classification with Domain Knowledge [Thesis]. Shanghai: East China Normal University; 2013
- [26] Jin LP, Dong J. Deep learning research on clinical electrocardiogram analysis. *Science China Information Sciences*. 2015;**45**(3):398-415. DOI: 10.1360/N112014-00060
- [27] Zhu HH, Dong J. An R-peak detection method based on peaks of Shannon energy envelope. *Biomedical Signal Processing and Control*. 2013;**8**(5):466-474. DOI: 10.1016/j.bspc.2013.01.001

- [28] Jin LP, Dong J. Ensemble deep learning for biomedical time series classification. *Computational Intelligence and Neuroscience*, 2016. Article ID: 6212684. DOI: 10.1155/2016/6212684
- [29] Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. *Science*. 2006;**313**(5786):504-507. DOI: 10.1126/science.1127647
- [30] Bengio Y. Learning deep architectures for AI. *Foundations and Trends in Machine Learning*. 2009;**2**(1):1-127. DOI: 10.1561/2200000006
- [31] Thakor NV, Webster JG, Tompkins WJ. Estimation of QRS complex power spectra for design of a QRS filter. *IEEE Transactions on Biomedical Engineering*. 1984;**31**(11):702-706. DOI: 10.1109/TBME.1984.325393
- [32] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*. 1998;**86**(11):2278-2324. DOI: 10.1109/5.726791
- [33] Kiranyaz S, Ince T, Hamila R, et al. Convolutional neural networks for patient-specific ECG classification. In: *The 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*; 25-29 August 2015, Milan, Italy. New York: IEEE; 2015. pp. 2608-2611
- [34] Kiranyaz S, Ince T, Hamila R, et al. Real-time patient-specific ECG classification by 1D convolutional neural networks. *IEEE Transactions on Biomedical Engineering*. 2016;**63**(3):664-675. DOI: 10.1109/TBME.2015.2468589
- [35] Li M, Zhang T, Chen YQ, et al. Efficient mini-batch training for stochastic optimization. In: *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; 24-27 August, 2014; 2014. New York: ACM; 2014. pp. 661-670
- [36] Vogl TP, Mangis JK, Rigler AK, et al. Accelerating the convergence of the back-propagation method. *Biological Cybernetics*. 1988;**59**(4):257-263. DOI: 10.1007/BF00332914
- [37] Kittler J, Hatef M, Duin RPW, et al. On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1998;**20**(3):226-239. DOI: 10.1109/34.667881
- [38] Kohler BU, Hennig C, Orglmeister R. The principles of software QRS detection. *IEEE Engineering in Medicine and Biology Magazine*. 2002;**21**(1):42-57. DOI: 10.1109/51.993193
- [39] Gorunescu F. *Data Mining Concepts Techniques and Models*. 1st ed. Berlin: Springer; 2011. pp. 185-317. DOI: 10.1007/978-3-642-19721-5
- [40] Shah AP, Rubin SA. Errors in the computerized electrocardiogram interpretation of Cardiacrhythm. *Journal of Electrocardiology*. 2007;**40**(5):385-390. DOI: 10.1016/j.jelectrocard.2007.03.008
- [41] Hakacova N, Tragardh-Johansson E, Wagner GS, et al. Computer-based rhythm diagnosis and its possible influence on nonexpert electrocardiogram readers. *Journal of Electrocardiology*. 2012;**45**(1):18-22. DOI: 10.1016/j.jelectrocard.2011.05.007

- [42] Dahl GE, Dong Y, Li D, et al. Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition. *IEEE Transactions on Audio, Speech and Language Processing*. 2012;**20**(1):30-42. DOI: 10.1109/TASL.2011.2134090
- [43] Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems*; 3–8 December. Vol. 2012. Vancouver: NIPS; 2012. pp. 153-160
- [44] Sutskever I, Hinton GE. Deep narrow sigmoid belief networks are universal Approximators. *Neural Computation*. 2008;**20**:2629-2636. DOI: 10.1162/neco.2008.12-07-661