sciendo

JAISCR, 2023, Vol. 13, No. 3, pp. 181 – 196
10.2478/jaiscr-2023-0014

A NOVEL METHOD FOR AUTOMATIC DETECTION OF ARRHYTHMIAS USING THE UNSUPERVISED CONVOLUTIONAL NEURAL NETWORK

Junming Zhang[1,2,3,4,*], Ruxian Yao[1,2] , Jinfeng Gao[1,2], Gangqiang Li[1,2], Haitao Wu[1,2]

[1]*College of Information Engineering, Huanghuai University, Henan 463000, China*

[2]*Henan Key Laboratory of Smart Lighting, Henan 463000, China*

[3]*Henan Joint International Research Laboratory of Behavior Optimization Control for Smart Robots, Henan 463000, China*

[4]*Zhumadian Artificial Intelligence & Medical Engineering Technical Research Centre, Henan 463000, China*

[*]*E-mail: 2013zhangjm@tongji.edu.cn*

Submitted: 28th January 2023; Accepted: 17th May 2023

Abstract

In recent years, various models based on convolutional neural networks (CNN) have been proposed to solve the cardiac arrhythmia detection problem and achieved saturated accuracy. However, these models are often viewed as "blackbox" and lack of interpretability, which hinders the understanding of cardiologists, and ultimately hinders the clinical use of intelligent terminals. At the same time, most of these approaches are supervised learning and require label data. It is a time-consuming and expensive process to obtain label data. Furthermore, in human visual cortex, the importance of lateral connection is same as feed-forward connection. Until now, CNN based on lateral connection have not been studied thus far. Consequently, in this paper, we combines CNNs, lateral connection and autoencoder (AE) to propose the building blocks of lateral connection convolutional autoencoder neural networks (LCAN) for cardiac arrhythmia detection, which learn representations in an unsupervised manner. Concretely, the LCAN contains a convolution layer, a lateral connection layer, an AE layer, and a pooling layer. The LCAN detects salient wave features through the lateral connection layer. The AE layer and competitive learning is used to update the filters of the convolution network—an unsupervised process that ensures similar weight distribution for all adjacent filters in each convolution layer and realizes the neurons' semantic arrangement in the LCAN. To evaluate the performances of the proposed model, we have implemented the experiments on the well-known MIT–BIH Arrhythmia Database. The proposed model yields total accuracies and kappa coefficients of 98% and 0.95, respectively. The experiment results show that the LCAN is not only effective, but also a useful tool for arrhythmia detection.

Keywords: convolutional neural network, arrhythmia detection, unsupervised learning, ECG classification.

1 Introduction

The Cardiovascular diseases are known as ”health killers” [1]. According to the report of the World Health Organization, about 17 million people die of cardiovascular diseases every year in the world [2]. Research shows that most cardiovascular diseases are usually accompanied by arrhythmia at the early stage [3]. Cardiac arrhythmia is an important manifestation of cardiovascular disease [4]. Therefore, early monitoring of arrhythmia plays a key role in the prevention of cardiovascular disease. Electrocardiography (ECG) is the most basic and well established method of diagnosing cardiac arrhythmia, as it is a non-invasive and easy to use method that can provide useful information on heart health and pathology [5]. Careful study of the ECG by expert cardiologists is necessary for the diagnosis of life threatening cardiac arrhythmias. However, with the development of wearable devices, massive ECG data is acquired every day. It is a time-consuming and tiresome task for cardiologists to analyze these ECG data [6]. Therefore, several methods are proposed for either fully automatic arrhythmia detection or event selection for further verification by human experts [6].

Most classical machine learning methods based on the extraction of features from single-channel ECG signal have been proposed, such as random forest [7, 8] , support vector machines (SVMs) [9, 10], artificial neural networks [11, 12, 13], KNN [14, 15, 16], and hidden Markov models [17, 18]. All these studies have achieved much better performances. In general, these methods include the following steps: noise removal, beat segmentation, feature extraction and classification. In these processes, feature extraction is the most important step. However, to design effective ECG features are labor-intensive and require expert knowledge, particularly of the restrictions on high-dimensional data [19]. In addition, almost all the models use shallow structures to process biosignals, which may be inadequate to compute complex input [20], such as ECG and electromyography (EMG) signals. Moreover, with the increase of the number of ECGs, the arrhythmia waveform becomes more complex, and traditional arrhythmia feature extraction is difficult to better complete the task of automatic diagnosis of arrhythmia. Therefore, it is of great clinical value and scientific significance to develop an efficient and accurate algorithm for automatic diagnosis of arrhythmia.

In recent decades, the advancement of deep learning makes it possible to learn feature automatically. Different from the traditional methods, the deep learning algorithm can use the original ECG signal as the input for automatic feature learning based on the probability distribution of the data set [21]. Deep learning models consisting of multiple processing layers, with each layer being able to learn increasingly abstract, higher-level representations of the input data relevant to perform specific tasks [21]. Uraab et al. [22] applied a 9-layer deep convolutional neural network to automatically identify 5 different categories of heartbeats in ECG signals.Fiorina L et al. [23] used convolutional neural network to process the original sampling points and realized the automatic classification of ECG signals. Uslu et al. [24] proposed a multi-task network to identify atrial fibrillation. Fan et al. [25] used CNN multi-scale fusion method to screen atrial fibrillation symptoms and obtained advanced performance. Lynn et al. [26] proposed a ECG classification method based on deep bidirectional GRU network. Andersen et al. [27] used CNN and recurrent neural network (RNN) to extract deep level features on RR interval for paroxysmal atrial fibrillation detection. Mousavi et al. [28] used 1D ECG sequences and 2D image features as input, and then used different depth neural networks on each channel for classification. Choi et al. [29] proposed a reverse time attention mechanism in the prediction task of heart failure to determine the impact of hospitalization records and medical codes on the prediction results.

All of these studies have achieved significantly improved performance. However, they are supervised methods, meaning that they require labeled data. With the emergence of wearable medical devices, a vast amount of data has become available. However, annotating this data is both expensive and challenging. Therefore, utilizing a large amount of unlabeled ECG data to diagnosis arrhythmia is very important in the biomedical community. More importantly, people do not know what knowledge the deep models have learned from the data and how to make the final decision. These end-to-end models lead to a very weak interpretability [30]. In many

fields, high requirements are put forward for the interpretability of the models. For example, in terms of detecting arrhythmia, the diagnosis of cardiologists largely depends on the results of machine learning methods, but the lack of interpretability limits the clinicians' trust in such methods.

In accordance with this problem, many researchers proposed different solutions. Koh et al. [31] research the prediction effect of the deep learning model through the influence function. Zhang et al. [32] added a simple yet effective loss for the output feature map of each lter to push the lter towards the representation of an object part. Lipton et al. [33] first analyzed the connotation of interpretability in the deep learning model from four aspects: trustworthiness, causality, transfer learning and information provision. By combining the deep neural network with first-order logic rules, the classification effect is significantly improved, showing good interpretability [34]. Some researchers used offline visualization [35-38], pattern transformation [39-40] and diagnosis [41-42] of pretrained models' representations. Although some have achieved results, interpretability is always the Achilles' heel of CNN[32].

To address the aforementioned challenges, this paper explores the enhancement of CNN using a semantic arrangement mechanism of neurons [43]. As Kandel et al. [44] have noted, neurons in different regions of the human visual cortex are organized in columnar structures, such that adjacent neurons can be activated by semantically similar inputs. Hubel et al. [45] further observed that neurons in adjacent columnar structures respond strongly to input patterns moving in the same direction. Neurons that exhibit a strong response to the same directional input are typically located close to each other in space [45]. It should be noted that this arrangement of neurons in the visual cortex is not innate, but rather, it is developed through training [46]. Therefore, if convolution neurons can learn semantic arrangement expression, the interpretability of CNN will be improved. Through literature analysis, we can see that almost all CNN models are feed-forward neural networks, that is, the information transmission between neurons is unidirectional. In the human visual neural network, in addition to the feed-forward connection, there are also a large number of lateral connections [47-48]. Thus, if convolution neurons can learn to express semantic arrangements, the interpretability of CNNs will be greatly improved. However, enabling convolution neurons to exhibit the semantic arrangement mechanism still poses a signify-cant challenge

According to Kohonen [49], a self-organizing feature mapping (SOM) network can accepts the external input mode. It will be divided into different corresponding regions, each region has different response characteristics to the input mode, and this process is automatically completed [49]. The fact of biological research shows that the organization principle of neurons is orderly arranged in the sensory channels of the human brain [50]. Therefore, when the human brain receives specific spatio-temporal information from the outside world through the senses, specific areas of the cerebral cortex will generate excitement, and similar external information is continuously mapped in the corresponding areas [50]. For a certain pattern or a specific excitation process of a certain frequency, the orderly arrangement of neurons and the continuous mapping of external information are the biological basis of competition mechanism in self-organizing feature map networks [46]. Therefore, SOM can be used to teach a CNN the semantic arrangement expression of a lateral connection.

The main issue with the self-organizing map (SOM) is that when there are many neurons in the network, some may never win [50]. These neurons are commonly referred to as 'hard neurons.' To address this problem, Bebis et al. [50] proposed a Kohonen algorithm with a conscience (SOFM-C), which not only maintains topological invariance mapping, but also effectively avoids the hard neuron problem. From a machine learning perspective, the optimization process of SOM can be viewed as an encoder. However, evaluating the performance of the encoder is a challenging problem. To tackle this, we drew inspiration from the concept of the autoencoder (AE) [51], and incorporated a decoder to reconstruct raw samples. By doing so, the CNN with lateral connections can learn semantic arrangement expression and achieve better feature representation. In this study, we propose a novel model, the lateral connection convolutional autoencoder neural network (LCAN), to automate arrhythmia detection by utilizing the feature representation capabilities

of the LCAN. Most importantly, the proposed training algorithm is unsupervised, which is crucial for portable ECG monitoring equipment.

The article is organized as follows: Section 2 presents the dataset and data description; Section 3 provides the details of SOFM-C, LCAN and the unsupervised feature learning process. The experimentation and discussion are presented in Section 4. Finally, Section 5 concludes the proposed method along with the future scope.

2 Dataset

For this study, the MIT-BIH Arrhythmia database [52], hosted at PhysioNet (http://www.physionet.org), is used to evaluate the performance of the proposed model. The database includes more than 4000 24-hour Holter recordings of 47 patients25 males(age range 32 – 89 years) and 22 females(age range 23 to 89 years). The 23 ECG recording signals beginning with No. 1 represent the common abnormal ECG signals in clinical ECG detection. The 25 ECG recording signals starting with No. 2 represent some uncommon ECG signal waveforms with very important clinical symptoms. 201 and 202 ECG records are from the same patient. Nearly 60% of these recordings were obtained from patients. Each ECG is recorded for about 30 minutes with a sampling frequency of 360Hz. Nearly 110000 heart beats were obtained by dividing all records according to heart beats. Each heart beat is labeled by two or more cardiologists. According to the Association for the Advancement of Medical Instrumentation (AAMI) specifications [53], the MIT-BIH database can be grouped into five arrhythmia groups: normal (N), supraventricular ectopic (S), ventricular ectopic (V), fusion (F), and unknown (Q). Table 1 shows the distribution of the data over the training and testing sets for each of the five types of arrhythmia. To get clean ECG signals, Daubechies wavelet 6 filters [22] are used to remove noises and artifacts.

Table 1. Distribution of five types of arrhythmia.

Class	Training	Testing	All data
N	72471	18118	90589
S	2223	556	2779
V	5788	1448	7236
F	641	162	803
Q	6431	1608	8039
Total	87554	21892	109446

3 Method

3.1 SOFM-C

Assuming that the total number of training samples is N, the training set can be expressed as X=$\{X_1, X_2, \ldots, X_N\}$. Suppose there are K input neurons and N output neurons, and the weight vector from the i-th input neuron to the j-th output neuron is W_{ij}, where $i \in [1,N], j \in [1,K]$. The SOFM-C algorithm steps are as follows.

1. Initialization of the weight matrix W=$\{W_1,W_2,\ldots,W_J\}$, bias vector B=$\{b_1, b_2, \ldots, b_J\}$ is set equal to $1/J$, learning rate $\eta_0(0 < \eta_0 < 1)$, the number of iterations is T, D with small random numbers occurs, where J is the number of neurons.

2. Calculation of the Euclidean distance between the input training vector X_i and neuron j:

$$d_{ji} = \sum_{j=1}^{J} (X_j - W_j(t))^2 - C(t)(\frac{1}{j} - b_j) \quad (1)$$

where t $\leq T$.

3. Select the neuron with the smallest distance d_{j^*} as the winning neuron:

$$d_{j^*} = min_j d_{ij} \quad (2)$$

4. Updating of the weight of the winner neuron j^* and its neighbors:

$$W_j(t+1) = W\,j(t) + \eta(t) f(j)(X_i - W_j(t)) \quad (3)$$

where $f(j) = exp(-||w_j - w_{j^*}||_2^2)/2sigma(t)^2)$, $sigma(t) = \frac{1}{1+e^{-t}}, \eta(t) = \eta_0 * exp(\frac{-t}{10})$.

5. Updating of the bias of the winner neuron:

$$b_{j^*}(t+1) = b_{j^*}(t) + D(1 - b_{j^*}(t)) \qquad (4)$$

$C(t) = C(0)(1 - \frac{t}{T})$. If $t = t_{max}$, then stop; otherwise repeat by going to step 2.

3.2 Lateral connection convolutio-nal neural network

Each region of the human visual cortex is highly connected [54-55]. These connections include feed-forward connections between multiple levels, and lateral connections between regions at the same level [54-55]. The number of lateral connections is far greater than that of feed-forward connections [56]. This is illustrated in Figure 1 (a) where the two consecutive convolutional layers and pooling layers are shown. In essence, a common CNN is a feed-forward neural network (FNN) without lateral connection. In terms of function, lateral connection mechanism can greatly suppress noise by creating a competitive environment for neurons, which is important for physiological signals, such as ECG. Therefore, we propose a novel lateral connection model: LCAN. As shown in Figure 1(b), LCAN consists of consecutive convolution layer, lateral connection layer and pool layer. As seen from Figure 1(b), each neuron is connected to other neurons in the same lateral connection layer (dotted line).

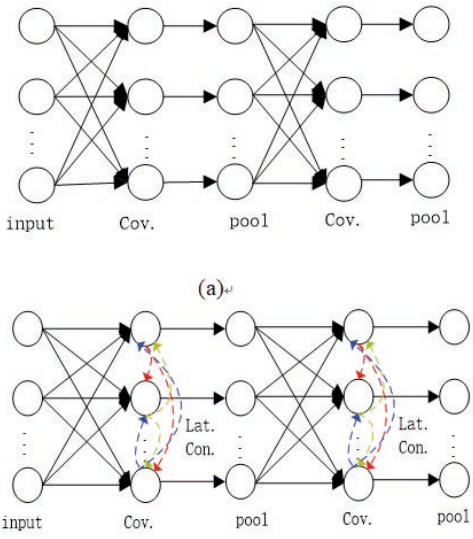


Figure 1. The illustration of common CNN (a) and a lateral connection CNN (b).

3.3 Unsupervised feature learning based on the LCAN

For the LCAN structure, we propose a two-stage unsupervised training algorithm. In the first phase, to study semantic arrangement expression, SOFM-C is used to train the parameters of LCAN. This process can be viewed as encoder. To ensure that the network learns useful features, decoder is used in the second phase. Suppose the number of input map is K. $W = W_{1,1}, W_{1,2}, \ldots, W_{1,K}, W_{2,1}, W_{2,2},$. $W_{j,k}$ denotes the filter between $j-th$ neuron and $k-th$ input map. Set the length of $W_{j,k}$ is L. Suppose the output of $j-th$ convolution neuron is $Z_j Z = Z_1, Z_2, \ldots, Z_j, \ldots, Z_i$. Z_j can be defined as:

$$\sum_{k=1}^{K} W_{jk} * X_k + b_j \qquad (5)$$

where $*$ represents convolution operation.

To utilize SOFM-C algorithm, we investigate the convolution operation in detail. Set $X_k = \{x_{k,1}, x_{k,2}, \ldots, x_{k,M}\}$, $W_{j,k} = \{w_{j,k,1}, w_{j,k,2}, \ldots, w_{j,k,L}\}$, $Z_{j,k} = X_k * W_{j,k}$. According to our previous study [68], convolution is a cross-correlation. Therefore, $Z_{j,k}$ can be written as:

$$Z_{j,k} = \begin{bmatrix} z_{j,k,1} \\ z_{j,k,2} \\ \cdots \\ z_{j,k,M-L+1} \end{bmatrix} = \begin{bmatrix} x_{k,1} & x_{k,2} & \cdots & x_{k,L} \\ x_{k,2} & x_{k,3} & \cdots & x_{k,L+1} \\ \cdots & \cdots & \cdots & \cdots \\ x_{k,M-L+1} & x_{k,M-L+2} & \cdots & x_{k,M} \end{bmatrix} \begin{bmatrix} w_{j,k,1} \\ w_{j,k,2} \\ \cdots \\ w_{j,k,L} \end{bmatrix} \qquad (6)$$

$$\text{Set } X'_k = \begin{bmatrix} x_{k,1} & x_{k,2} & \cdots & x_{k,L} \\ x_{k,2} & x_{k,3} & \cdots & x_{k,L+1} \\ \cdots & \cdots & \cdots & \cdots \\ x_{k,M-L+1} & x_{k,M-L+2} & \cdots & x_{k,M} \end{bmatrix},$$

Eq. (6) can be is defined as:

$$Z_{j,k} = X'_k W_{j,k} \qquad (7)$$

Therefore, Eq. (5) can be defined as:

$$Z_j = \begin{bmatrix} W_{j,1} & W_{j,2} & \cdots & W_{j,K} \end{bmatrix} \begin{bmatrix} X'_1 \\ X'_2 \\ \cdots \\ X'_K \end{bmatrix}, \qquad (8)$$

Hence, we can use matrix multiplication instead of convolution calculation. According to Eq. (1) and Eq. (8), SOFM-C can be used to train the filters of the LCAN.

Set input data X_k, the number of output neurons is J. We first extract subset $X_k^{'}=\{X_{k,1}^{'},X_{k,2}^{'},\cdots,X_{k,d}^{'},\cdots,X_{k,M-L+1}^{'}\}$ from X_k. According to Eq. (7), we can calculate the output feature map of each neuron. For example, for input $X_{k,d}^{'}$, the output of each neuron can be written as $P=\{p_{1,d}, p_{2,d}, p_{3,d}, p_{4,d}\ldots, p_{j,d}, \ldots, p_{J,d}\}$.

$$p_{j,d} = X_{k,d}^{'} W_{j,k} \tag{9}$$

where $X_{k,d}^{'} = [x_{k,d}\ x_{k,d+1}\ \cdots\ x_{k,d+L-1}]$. The result of Eq. (9) is a SOM network (Figure 2). Therefore, we can use SOFM-C to train the filter $W_{j,k}$.

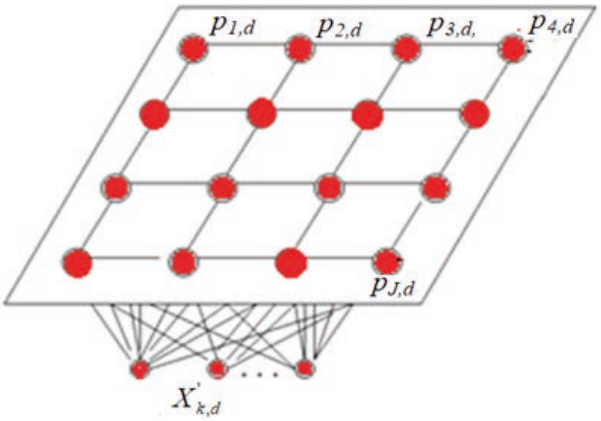


Figure 2. The calculation process of Eq. (9).

Suppose there are K input samples and J convolution neurons. Figure 3 shows the unsupervised training process of LCAN in detail.

In the first phase, the filters of LCAN can be trained by using SOFM-C. The process of Eq. (9) can be regarded as encoder. To ensure that the LCAN learns better feature representation, the decoder is used. According to Eq. (9), we can define the decoder as follows:

$$X_{k,d}^{*} = p_{j,d} W_{j,k}^{T} \tag{10}$$

As in literature [59], the cost function can be defined as :

$$J = J(W,b) + \beta \sum_{j=1} KL(\rho||\hat{\rho}_j) \tag{11}$$

where $KL(\rho||\hat{\rho}_j)=\rho\log\frac{\rho}{\hat{\rho}_j}+(1-\rho)\log\frac{1-\rho}{1-\hat{\rho}_j}$, is a constant, which controls the weight of sparsity penalty factor. Suppose there are K input features and j output neurons. The detailed decoder process is shown in Figure 4

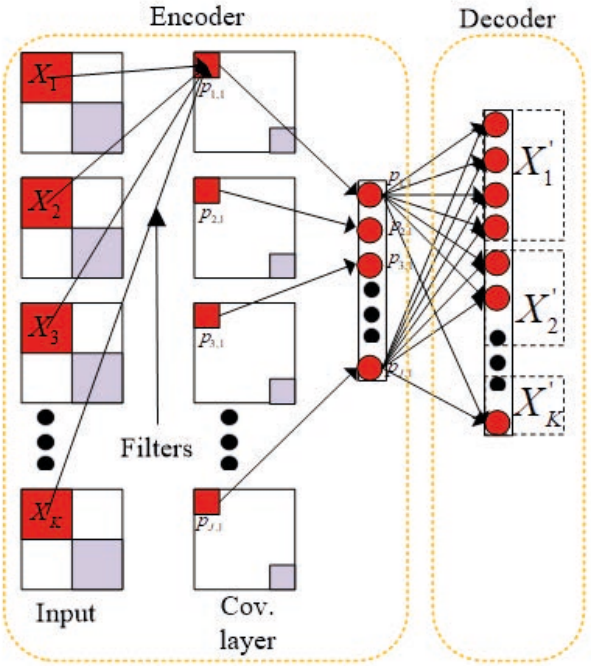


Figure 4. Decoder process of the LCAN

4 Experimentation

In order to measure the classification performance of the proposed model, all experiment results are evaluated with F_1 score, total accuracy (TAC), specificity, sensitivity, and kappa coefficient (KP) [57-59].

TAC = $\frac{TP+TN}{TP+FN+FP+TN}$%,

precision = $\frac{TP}{TP+FP}$%,

recall = sensitivity = $\frac{TP}{TP+FN}$%,

specificity = $\frac{TN}{TN+FP}$%,

F_1 = (2* recall * precision) / (recall + precision), where TP, TN, FP and FN denote true positives, true negatives, false positives and false negatives, respectively. In this study, a ten-fold cross-validation is used to assess the performance of the proposed model. We implemented all the experiments using MATLAB 2019a and Python Google Tensor-

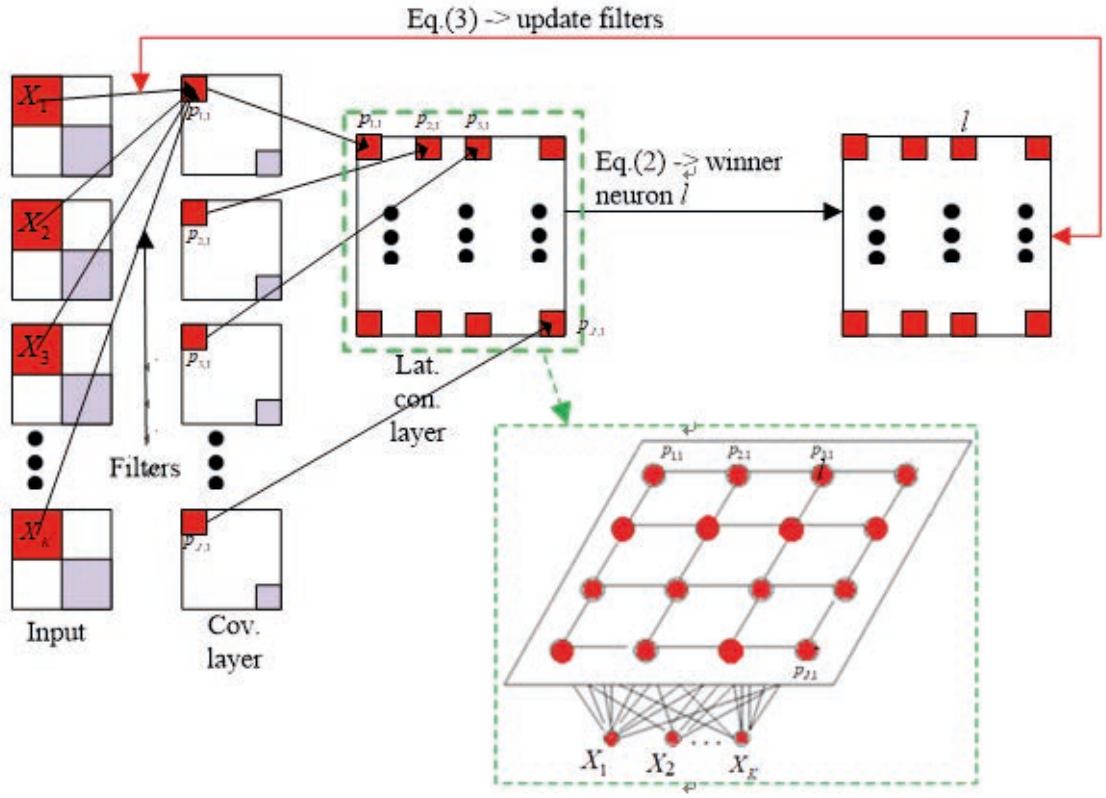


Figure 3. Unsupervised training process of the LCAN.

Table 2. The parameters of the LCAN model

Layers	Units	Unit Type	Size	Stride	Others
Input					
BN					
Con.	100	ReLU	6×1	1	
LC.					C(0)=1, η_0=0.1,
AE					= 0.1.
Pooling			3×1	2	
Con.	100	ReLU	3×1	1	
LC.					C(0)=1, η_0=0.1,
AE					β =0.1.
Pooling			2×1	2	
Con.	100	ReLU	3×1	1	
LC.					C(0)=1, η_0=0.1,
AE					β =0.1.
Pooling			2×1	2	
Flatten					
Dense	100	ReLU			
Dense	100	ReLU			
Dropout					0.46
Dense	5	Softmax			

Con. denotes connection layer; LC. denotes lateral connection layer.

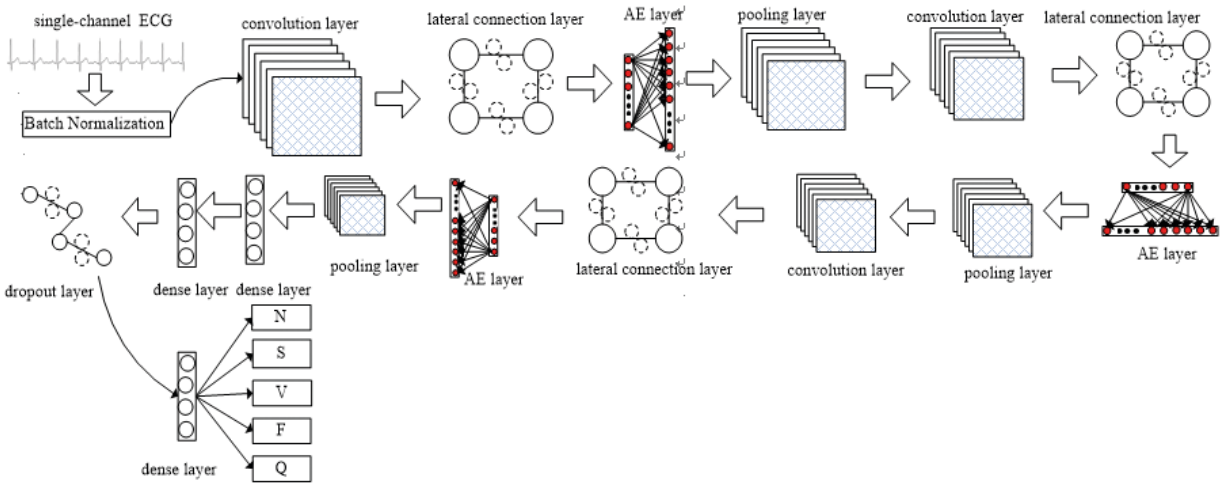


Figure 5. The architecture of the proposed LCAN model.

flow 2.0 on a machine equipped with 64 GB memory and NVidia Geforce GTX 3080 GPU.

4.1 Experimental Setup

In the present study, a 17-layer LCAN (see Figure 5) was designed for the classification of cardiac arrhythmia. The proposed model receives ECG beats as inputs and classifies them into five distinct classes, namely N, S, V, F and Q. More specifically, use the following procedure to divide each ECG signal into separate beats: first, detect all R peaks, and then define each pulse as 250 ms before R peak and 500 ms after R peak. In this way, a one-dimensional ECG signals with 187 sample points are fed as the input of the proposed model. Table 2 shows the detailed parameters, such as the number of filters, the size and stride in each convolution layer, the size and stride of the kernel in each pooling layer.

We trained the proposed model with a maximum number of 300 epochs and a batch size of 32 samples. The Adam optimizer is applied to minimize the loss with a learning rate a = 0.001. To mitigate the effect of the overfitting problem, a dropout technique with a probability of retaining input units of 0.46 is used. The main code of the proposed model is available on the GitHub (https://github.com/hhzjm/automatic-detection-of-arrhythmias).

4.2 Performance Results

To evaluate the proposed model performance, we show the learn curves. Figure 6(a) shows the curves of the loss curve during the training of the model on the training and validation data. Fig.6(b) shows the accuracy curves in the training and validation phases. From Figure 6, we can see that the network converges for both values after 30 epochs. Therefore, we think the training loss is not decreasing, and the accuracy is not increasing with increasing the number of epochs.

Table 3 presents the confusion matrix of the LCAN about heartbeats using the test data. From Table 3, we can see that the highest ACC, PRE, SEN, SPEC and F1 values of the proposed model for N, S, V, F and Q are 99.7%, 99.35%, 99.7%, 99.9% and 0. 993, respectively. The smallest ACC, PRE, SEN, SPEC and F1 values of the LCAN for N, S, V, F and Q are 98.76%, 79%, 81.5%, 94.25% and 0. 802, respectively. Table 4 shows the overall performance of the LCAN. From Table 4, it is observed that the values of the TAC, KP and MF1 are 98%, 0.95 and 0.92, respectively. Based on the above results, it can be deduced that the LCAN can correctly classify most of the heartbeats. Meanwhile, compared with the values of F1 in Table 3, we can observe that the proposed network model has a slight difficulty classification some of the data belonging to the S and F class. The reason is the class imbalance problem existed in the dataset where the group F has only 641 heartbeats and the group S has 2223 heartbeats (as shown in Table 1). An imbalanced

dataset can negatively affect the performance of a machine learning algorithm.

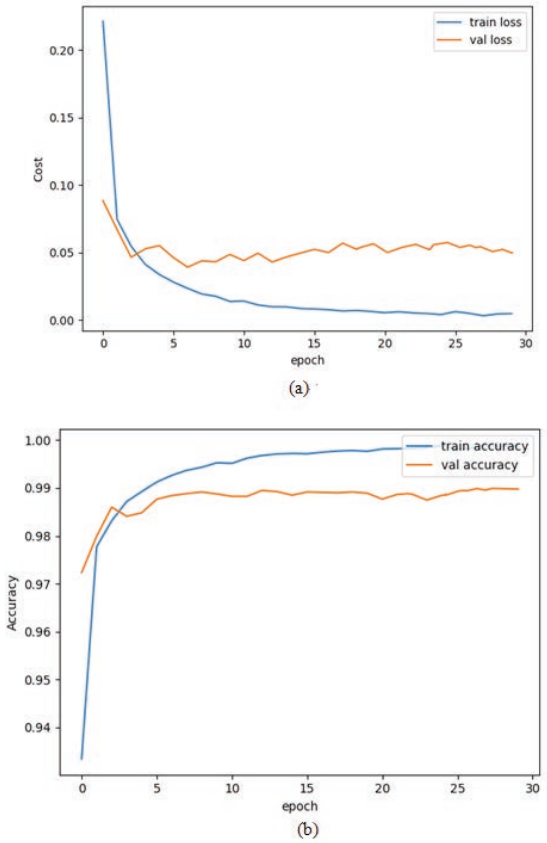


Figure 6. Training of the LCAN model: loss (a) and accuracy (b) curves.

Table 4. Overall Performance of the LCAN (%).

models	TAC	KP	MF_1
LCAN	0.98	0.95	0.92

MF1: Macro-averaging of F1

To evaluate the efficiency of the LCAN, we implemented two experimentations for comparison:

General CNN (CNN-1): CNN-1 is a supervised CNN model. According to Fig.6, we removed the lateral connection layer and AE layer, and the remaining model structure is CNN-1 (as shown in Fig.7). It consists of three convolution layers, three pooling layers, one batch normalization layer, a dropout layer, two fully-connected layers, and one softmax layer. The parameters of CNN-1 resemble those of the LCAN.

Lateral connection CNN (CNN-2): As illustrated in Figure 8, CNN-2 is an unsupervised model. A lateral connection layer is added after each convolution layer. SOFM-C is used to train the filters of CNN-2. From Fig.8, we can see that there are three convolution layers, three lateral connection layers, three pooling layers, one batch normalization layer, a dropout layer, two fully-connected layers, and one softmax layer. The parameters of CNN-2, such as the number of filters, the size and stride in each convolution layer, resemble those of CNN-1. The hyper parameters of the proposed model are chosen as follows: C(0)=1, η_0=0.1, D=0.03, epoch=400.

Table 5 presents the obtained classification results of the CNN-1 and CNN-2 models. From Table 5, it is observed that the LCAN achieved highest 98% TAC, 0.95 KP and 0.92 MF_1 followed by CNN-1 with TAC, KP and MF_1 of 95%, 0.9 and 0.85, respectively. The least recognizable model is CNN-2 with TAC, KP and MF_1 of 89%, 0.82 and 0.81, respectively. Compared with CNN-2, the CNN-1 is a supervised model. This model can better fit the raw ECG data distribution, which is good for classification problems. Therefore, the performance of CNN-1 better than that of CNN-2. Although LCAN is an unsupervised model, it can learn semantic arrangement expression by lateral connection and better feature repression via AE algorithm. Therefore, the performance of LCAN better than that of CNN-1. In order to analyze this reason, we show the learned filters from the three models.

Table 5. Performance evaluation for the three models (%).

models	TAC	KP	MF_1
CNN-1	0.95	0.9	0.85
CNN-2	0.89	0.82	0.81
LCAN	0.98	0.95	0.92

MF1: Macro-averaging of F1

Figure 9 depicts the learned features from CNN-1, Figure 10 presents the learned features from CNN-2, and Figure 11 shows the learned features from LCAN.

Table 3. Presents the confusion matrix of the LCAN about heartbeats using the test data.

		Predicted					Per-class Performance(%)				
		N	S	V	F	Q	ACC	PRE	SEN	SPEC	F1
Actual	N	18064	23	26	3	2	98.76	98.9	99.7	94.25	99.3
	S	83	461	6	4	2	99.39	92.2	82.9	99.8	87.3
	V	70	9	1342	22	5	99.25	95.86	92.68	99.7	94.2
	F	16	3	10	132	1	99.7	79	81.5	99.8	80.2
	Q	48	4	16	6	1534	99.6	99.35	95.4	99.9	97.3

ACC: accuracy; PRE: precision; SEN: sensitivity, SPEC: specificity

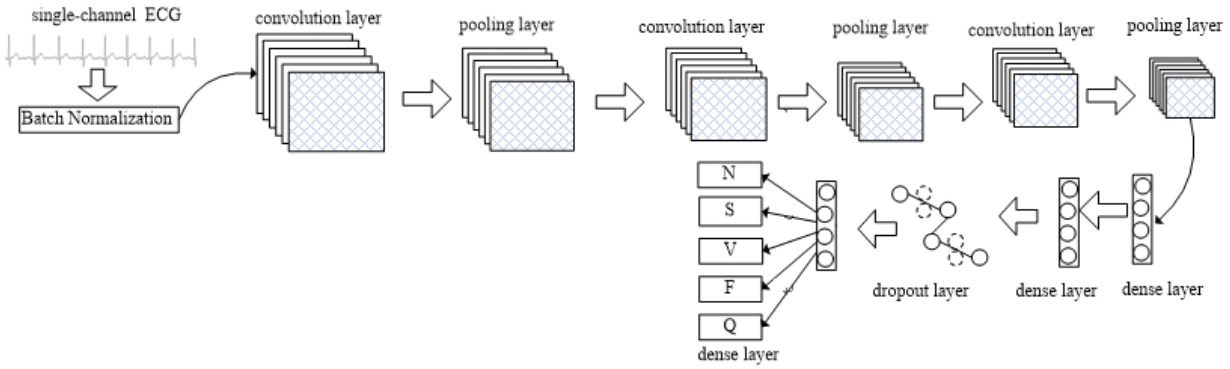


Figure 7. The architecture of the CNN-1.

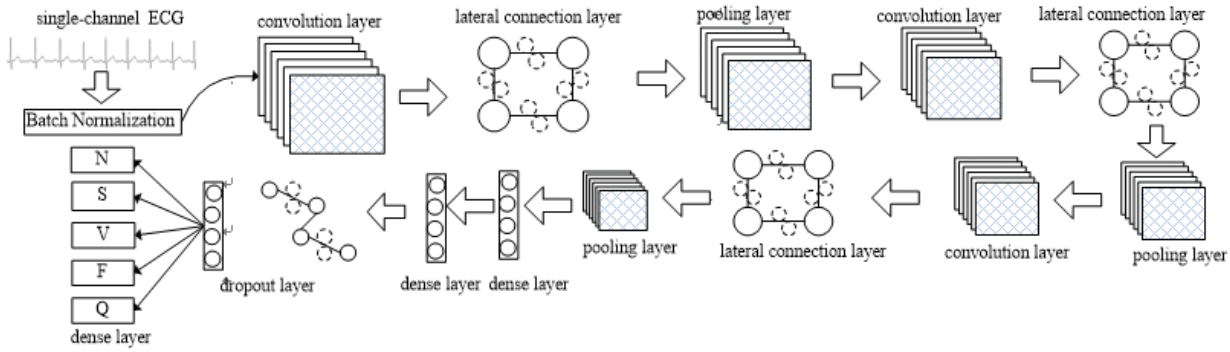


Figure 8. The architecture of the CNN-2.

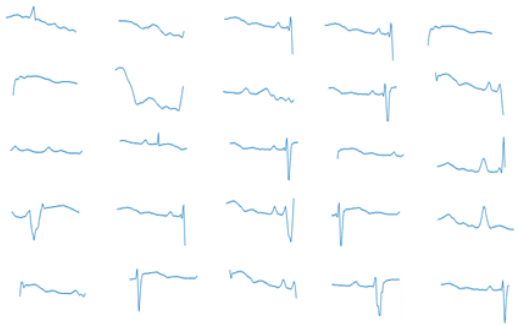


Figure 9. Learned features from CNN-1.

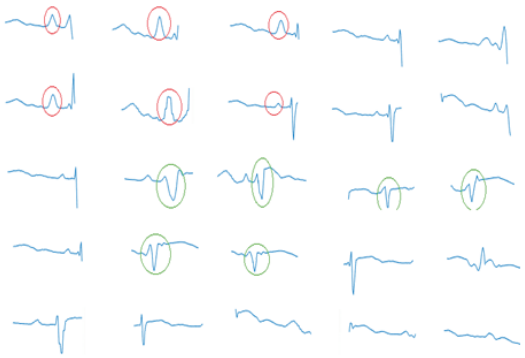


Figure 10. Learned features from CNN-2

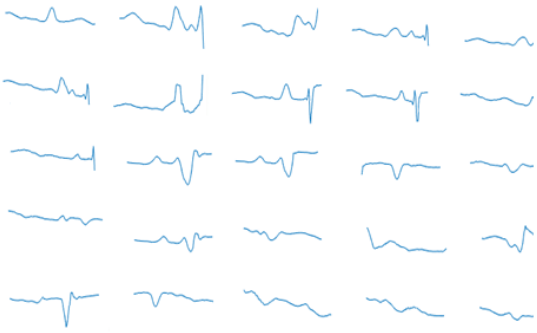


Figure 11. Learned features from LCAN.

From Figure 9, we can see that learned features of CN N-1 are irregular and scattered. The reason maybe that the CNN-1 is a "black box" model. This model learns features through backpropagation (BP) algorithm. Therefore, the training process of model is irregular and difficult to understand. Compared with Figure 9 and 10, we observe that the learned features of CNN-2 are more concentrated, which is reminiscent of the arrangement of neurons in the V1 region of the human visual cortex. The semantic arrangement expression significantly enhances the interpretability and performance of the proposed model. Furthermore, as shown in Figure 10, certain filters (indicated by red and green circles) exhibit a distribution that is similar to the Mexican hat function. These diverse Mexican hat functions can extract different multiscale features [60], which is advantageous for arrhythmia classification.

4.3 Compared With Existing Models

In order to prove the usefulness of the proposed model in this paper, a comparative results based on the MIT-BIH Dataset are listed in Table 6. These studies include the residual network-based model framework proposed by Hannun et al. [63], the transformer-based deep neural network proposed by Hu et al. [64] and the network model combining RNN network with attention proposed by Mousavi et al. [65]. The purpose of the comparison is not to prove that the proposed model is superior to the existing methods, but to explore the feasibility of the unsupervised method for arrhythmia detection. From table 6, we can see that Ref. [64] has obtained the best performance with TAC 99.5% and F1 0.94. The reasons for excellent performance are that multi-head attention, self-attention and supervised training algorithm were used in this paper. Compared with LCAN, the structure of model in Ref. [64] is relatively complex. Table 6 also shows that Ref. [66] has achieved good performance in identifying eleven classed of ECG events. A preprocessing step was used in Ref. [66], which not only model the peak characteristics of ECG, but also preserve signal structure and morphology. Compared with existing models, the proposed method also obtains a better classification performance for arrhythmia detection. Therefore, it is a useful decision-support tool for automatic arrhythmia detection. More importantly, our method is unsupervised learning. It is impossible to score labeled data from portable wearable ECG devices. Hence, unsupervised method is critical for these devices.

5 Conclusion

In this paper, for the problems of dependence labeled data and poor interpretability of deep learning model in the classification of ECG events, we carried out relevant research. Finally, a novel lateral connection network with learnable semantic feature

Table 6. Performance comparison with other studies.

Method	Feature extraction	classes	TAC	F1
[63]	Deep learning	5	98%	0.897
[64]	Deep learning	4	99.5%	0.94
[65]	Deep learning	5	97.6%	0.858
[67]	Deep learning	5	94%	—
[68]	Handcrafted features	2	99%	—
[69]	Deep learning	4	97.37	—
[70]	Handcrafted features	5	99.5%	—
[71]	Deep learning	5	99.4%	0.958
[66]	Deep learning	11	97.63%	0.926
LCAN	Deep learning	5	**98%**	**0.92**

representation is proposed. Figure 11 shows that the proposed model can learn semantic feature representation, which resembles the arrangement of neurons in the V1 region of the human visual cortex. This learned semantic feature representation helps improve network interpretability. Table 5 indicates that the proposed model obtains excellent performance with a KP of 0.95, an F_1 of 0.92, and a TAC of 98%. Results evidence that the proposed model and training method are useful and effective. The proposed method does not need to design features. It can learn features from raw ECG signals, which can be adapted to other time serial data, such as those derived from EEG and EMG. Moreover, the proposed method is unsupervised, making it highly suitable for application in home sleep monitoring equipment.

In this study, we propose a novel unsupervised network and explore its availability. From Table 6, we can see that the classification performance of LCAN is not the best. The promising results will motivate continued exploration. Future work include Eq. (1) improving the performance of classification of LCAN by adding a recursive operation that can use contextual information; Eq. (2) Classification of ECG signal fragments containing multiple classes, and Eq. (3) testing the efficiency of LCAN using other physiological signals.

Acknowledgments

This work was supported in part by the Henan provincial key science and technology research projects under Grants NOS. 232102210074 and 232102210076, in part by the National Science Foundation of China under Grants 61973177, in part by the Henan Key Laboratory of Smart Lighting, in part by Henan International Joint Laboratory of Behavior Optimization Control for Smart Robots, in part by the programme of Henan Innovative Research Team of Cooperative Control in Swarm-based Robotics, in part by the Joint Postgraduate Training Base Grant YJS2022JD45, in part by the Award Plan for Tianzhong Scholars of Huanghuai University in 2019, in part by Zhumadian Artificial Intelligence & Medical Engineering Technical Research Centre, in part by Zhumadian industrial innovation and development research major project under Grant 2020ZDA06.

References

[1] E. J, M. J. Blaha, S. E. Chiuve, Benjamin. Heart disease and stroke statistics-2017 update: a report from the american heart association, Circulation, vol. 135, pp. E646-E646, 2017.

[2] X. Li, C. Wu, J. Lu, et al., Cardiovascular risk factors in China: a nationwide population-based cohort study, The Lancet Public Health, vol. 5, pp. e672-e681, 2020.

[3] A. Isin, S. Ozdalili, Cardiac arrhythmia detection using deep learning, Procedia Computer Science, vol. 120, pp.268-275, 2017.

[4] Z. Yldrm, P. Pawiak , S. T. Ru , et al., Arrhythmia detection using deep convolutional neural network with long duration ECG signals, Computers in Biology and Medicine, vol.102, pp.411-420, 2018.

[5] A. Mb, B. Tt, B. Sd, D. Rstc, et al., Automated arrhythmia detection with homeomorphically irreducible tree technique using more than 10,000 indi-

vidual subject ECG records, Information Sciences, vol. 575, pp.323-337, 2021.

[6] A. Y. Hannun, P. Rajpurkar, M. Haghpanahi, et al., Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network, Nature medicine, vol. 25, pp. 65-69, 2019.

[7] C. Vimal, B. Sathish, Random forest classifier based ECG arrhythmia classification, International Journal of Healthcare Information Systems & Informatics, vol. 5, pp. 1-10,2009.

[8] V. N. Pham, H. L. Tran, Electrocardiogram (ECG) circuit design and using the random forest to ECG arrhythmia classification, Lecture Notes in Networks and Systems, DOI:https://doi.org/10.1007/978-3-031-22200-9_54, 2023.

[9] A. H. Khandoker, M. Palaniswami, C. K. Karmakar, Support vector machines for automated recognition of obstructive sleep apnea syndrome from ECG recordings, IEEE Transactions on Information Technology in Biomedicine, vol. 13, 37-48, 2009.

[10] E. H. Houssein, I. E. Ibrahim, N. Neggaz, et al., An efficient ecg arrhythmia classification method based on manta ray foraging optimization, Expert Systems with Applications, DOI: https://doi.org/10.1016/j.eswa.2021.115131,2021.

[11] Y. W. Hau, H. W. Lim, C. W. Lim, et al., P204 automated detection of atrial fibrillation based on stationary wavelet transform and artificial neural network targeted for embedded system-on-chip technology, European Heart Journal, DOI: 10.1093/ehjci/ehz872.075, 2020.

[12] M. Alfaro-Ponce, I. Chairez, R. Etienne-Cummings, Automatic detection of electrocardiographic arrhythmias by parallel continuous neural networks implemented in FPGA, Neural Computing and Applications, vol. 31, 363-375, 2017.

[13] M. I. Owis, A. H. Abou-Zied, A. B. M. Youssef, et al., Study of features based on nonlinear dynamical modeling in ECG arrhythmia detection and classification, IEEE transactions on biomedical engineering, vol. 49, pp. 733-736, 2002.

[14] B. Venkataramanaiah , J. Kamala, ECG signal processing and KNN classifier-based abnormality detection by VH-doctor for remote cardiac healthcare monitoring, Soft Computing, vol.24, 17457-17466,2020.

[15] T. Tuncer, S. Dogan, P. Plawiak, et al., Automated arrhythmia detection using novel hexadecimal local pattern and multilevel wavelet transform with ECG signals, Knowledge-Based Systems, vol. 186, pp. 1-19, 2019 .

[16] D. A. Coast, R. M. Stern, G. G. Cano, et al., An approach to cardiac arrhythmia analysis using hidden Markov models, IEEE Transactions on Biomedical Engineering, vol. 37, pp. 826-836, 2002.

[17] A. Sadoughi, M. B. Shamsollahi, E. Fatemizadeh, et al., Detection of Apnea bradycardia from ECG signals of preterm infants using layered hidden markov model, Annals of Biomedical Engineering, vol. 49, pp. 2159-2169, 2021.

[18] C. Angermueller, T. Pärnamaa, L. Parts, et al., Deep learning for computational biology, Molecular Systems Biology, DOI: 10.15252/msb.20156651, 2016.

[19] Y. Lecun, Y. Bengio, G. Hinton, Deep learning, Nature, vol. 521, pp.436-444 ,2015.

[20] Y. H. Awni, R. Pranva, H. Masoumeh, et al., Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network, Nature Medicine, vol. 25, pp.65-69,2019.

[21] C. Uraab, L. Shu, A. Yh, et al., (2017). A deep convolutional neural network model to classify heartbeats. Computers in Biology and Medicine, 89(2017)389-396.

[22] L. Fiorina, B. Lefebvre, C. Gardella, et al., Smartwatch-based detection of atrial arrhythmia using a deep neural network in a tertiary care hospital, Europace, DOI: https://doi.org/10.1093/europace/euac053.563, 2022.

[23] F. Uslu, M. Varela, G. Boniface, et al., LA-Net: A Multi-task deep network for the segmentation of the left atrium, IEEE transactions on medical imaging, 41 (2)(2021)456-464.

[24] X. Fan, Q. Yao, Y. Cai, et al., Multiscaled fusion of deep convolutional neural networks for screening atrial fibrillation from single lead short ECG recordings, IEEE Journal of Biomedical and Health Informatics, 22(6)(2018)1744-1753.

[25] H. M. Lynn, S. B. Pan, P. Kim, A deep bidirectional GRU network model for biometric electrocardiogram classification based on recurrent neural networks, IEEE Access, 7(2019)145395-145405.

[26] R. S. Andersen, A. Peimankar, S. Puthusserypady, A deep learning approach for real-time detection of atrial fibrillation, Expert Systems with Applications, 115(2019)465-473.

[27] S. Mousavi, F. Afghah, A. Razi, et al., ECGNET: Learning where to attend for detection of atrial fibrillation with deep visual attention, In 2019 IEEE EMBS International Conference on Biomedical & Health Informatics 2019, pp.1-4.

[28] E. Choi, M. T. Bahadori, J. Sun, et al.. Retain: An interpretable predictive model for healthcare using reverse time attention mechanism, In advances in neural information processing systems, 2016, pp. 3504-3512.

[29] F. Wu, B. B. Liao, Y. H. Han Interpretability for Deep Learning, Aero Weaponry, 2019(1) 39-46.

[30] P. W. Koh , P. Liang, Understanding black-box predictions via influence functions, 70(2017) 1885–1894.

[31] Q. Zhang, Y. N. Wu, S. C. Zhu, Interpretable convolutional neural networks, In: 2018 IEEE/CVF conference on computer vision and pattern recognition, DOI: 10.1109/CVPR.2018.00920, 2018.

[32] Z. C. Lipton, The mythos of model interpretability, Communications of the ACM, 61(10)(2018)36-43.

[33] Z. Hu, Z. Yang, X. Liang, et al., Toward controlled generation of Text, arXiv:1703.00955, 2017.

[34] A. Mahendran, A. Vedaldi, Understanding deep image representations by inverting them, In 2015 IEEE conference on computer vision and pattern recognition, DOI: 10.1109/CVPR.2015.7299155, 2015.

[35] A. Dosovitskiy, T. Brox, Inverting visual representations with convolutional networks, In 2015 IEEE conference on computer vision and pattern recognition, DOI: 10.1109/CVPR.2016.522, 2016.

[36] R. C. Fong, A. Vedaldi, Interpretable explanations of black boxes by meaningful perturbation, In 2017 IEEE international conference on computer vision, DOI: 10.1109/ICCV.2017.371, 2017.

[37] R. R. Selvaraju, M. Cogswell, A. Das, et al., Grad-cam: visual explanations from deep networks via gradient-based localization, In 2017 IEEE international conference on computer vision, DOI: 10.1109/ICCV.2017.37, 1, 2017.

[38] Q. Zhang, R.Cao, F. Shi, et al., Interpreting CNN knowledge via an explanatory graph, In Thirty-second AAAI conference on artificial intelligence, pp. 4454-4463, 2018.

[39] W. Hong, X. Yunchao, L. Dawei, et al., Rapid identification of X-ray diffraction patterns based on very limited data by interpretable convolutional neural networks, Journal of Chemical Information and Modeling, vol. 60, pp.2004-2011, 2020.

[40] D. Bau, B. Zhou, A. Khosla,et al., Network dissection: Quantifying interpretability of deep visual representations, In 2017 IEEE conference on computer vision and pattern recognition, DOI: 10.1109/CVPR.2017.354, 2017.

[41] P. Koh, P. Liang, Understanding black-box predictions via influence functions, In proceedings of the 34th international conference on machine learning, pp.1885-1894, 2017.

[42] H. Asanuma, Recent developments in the study of the columnar arrangement of neurons within the motor cortex, Physiological Reviews, vol. 55,pp. 143-156, pp.1975.

[43] E. Kandel, J. schwartz, T. Jessell, et al., Principles of neural science, 5th ed. New York, USA: McGraw-Hill, 2012.

[44] D. H. Hubel, T.N. Wiesel, Ferrier lecture: functional architecture of macaque monkey visual cortex, Proceedings of the Royal Society of London, vol. 198, pp. 1-59, 1977.

[45] K. Tanaka, Columns for complex visual object features in the inferotemporal cortex:clustering of cells with similar but slightly different stimulus selectivities, Cerebral Cortex, vol. 13, pp. 90-99, 2003.

[46] A. Angelucci, C. Bressloff P, Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons, Progress in Brain Research, vol. 154, pp.93-120, 2006.

[47] V. A. Lamme, H. Supèr, H. Spekreijse, Feedforward, horizontal, and feedback processing in the visual cortex, Current Opinion in Neurobiology, vol. 8, pp. 529, 1998.

[48] T. Kohonen, Self-organized formation of topologically correct feature maps, Biological Cybernetics, vol. 43, pp. 59-69,1982.

[49] G. Bebis, M. Georgiopoulos, N.V. Lobo, et al., Using self-organizing maps to learn geometric hash functions for model-based object recognition, IEEE Transactions on Neural Networks, vol. 9, 560-570, 1998.

[50] A. Y. Ng, Sparse autoencoder, CS294 A Lecture notes 72, 2011.

[51] G.B. Moody, R.G. Mark, The impact of the MIT-BIH arrhythmia database, IEEE Engineering in Medicine and Biology Magazine, vol. 20, pp. 45-50,2001.

[52] R. Mark, Aami-recommended practice: testing and reporting performance results of ventricular arrhythmia detection algorithms, in: Association for the Advancement of Medical Instrumentation, Arrhythmia Monitoring Subcommittee, AAMI ECAR, 1987.

[53] D. J. Felleman, D. Essen, Distributed hierarchical processing in the primate cerebral cortex, Cerebral Cortex, vol.1, pp. 1-47,1991.

[54] Z. P. Lo, M. Fujita, B. Bavarian, Analysis of neighborhood interaction in Kohonen neural networks, In IEEE fifth international proceedings parallel processing symposium, pp. 246-249,199,.

[55] Z. P. Lo, Y. Yu, B. Bavarian, Analysis of the convergence properties of topology preserveing neural networks. IEEE Transactions on Neural Networks, vol. 4, 207-220,1993.

[56] J. M. Zhang, Y. Wu, A new method for automatic sleep stage classification, IEEE Transactions on Biomedical Circuits and Systems, vol. 11, pp. 1097-1110, 2017.

[57] S. F. Liang, C. E. Kuo, Y. H. Hu, et al., Automatic stage scoring of single-channel sleep EEG by using multiscale entropy and autoregressive models, IEEE Transactions on Instrumentation and Measurement, vol. 61, pp. 1649-1657, 2012.

[58] J. Cohen, A coefficient of agreement for nominal scales, Educational & Psychological Measurement, vol. 20, pp. 37-46, 1960.

[59] H. Zhang, C. M. Cartwright, M. S. Ding, et al., Image feature extraction with various wavelet functions in a photorefractive joint transform correlator, Optics Communications, vol. 185, pp. 277-284, 2000.

[60] https://cs231n.github.io/understandingcnn.

[61] ahilS.,http://medium.com/towards-data-sci ence/experiences-with-a-new-kind-of-convolution-dfe603262e4c, 2017.

[62] M. Hammad, A. M., Iliyasu, A. Subasi, et al., A multitier deep learning model for arrhythmia detection, IEEE Transactions on Instrumentation and Measurement, DOI: 10.1109/TIM.2020.3033072, 2020.

[63] H. Rui, C. Jie, Z. Li, A transformer-based deep neural network for arrhythmia detection using continuous ECG signals, Computers in Biology and Medicine, DOI: https://doi.org/10-.1016/j.compbiomed.2022.105325, 2022.

[64] S. Mousavi, F. Afghah, F. Khadem, U.R. Acharya, Ecg language processing (elp): a new technique to analyze ecg signals, Computer Methods and Programs in Biomedicine, vol. 202, pp.105959, 2021.

[65] A. Chandrasekar, D. D. Shekar, A. C. Hiremath, et al., Detection of arrhythmia from electrocardiogram signals using a novel gaussian assisted signal smoothing and pattern recognition, Biomedical Signal Processing and Control, DOI: https://doi.org/10.1016/j.bspc.2021.103469, 2022.

[66] C. Uraab, L. Shu, A. Yh, et al., A deep convolutional neural network model to classify heartbeats, Computers in Biology and Medicine, vol. 89, pp. 389-396, 2017.

[67] M. Hammad, A. Maher, K. Wang, et al., Detection of abnormal heart conditions based on characteristics of ECG signals, Measurement, vol. 125, pp. 634-644, 2018.

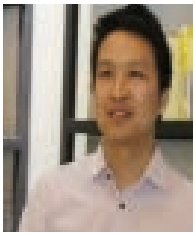
[68] M. Amrani, M. Hammad, F. Jiang, et al., Very deep feature extraction and fusion for arrhythmias detection, Neural Computing and Applications, vol. 30, pp. 2047-2057, 2018.

[69] R. Li, X. Zhang, H. Dai, et al., Interpretability analysis of heartbeat classification based on heartbeat activity's global sequence features and BiLSTM-attention neural network, IEEE Access, vol. 7, pp. 109870-109883, 2019.

[70] J. Lv, Q. Ye, Y. Sun, et al., Heart-darts: classification of heartbeats using differentiable architecture search, In 2021 International Joint Conference on Neural Networks, DOI: 10.1109/IJCNN52387.2021.9534184, 2021.

Junming Zhang received the Ph.D. degree in Computer Science from Tongji University, Shanghai, China, in 2018. His research interests include machine learning, biomedical signal processing, and automatic sleep staging.
https://orcid.org/0000-0003-1836-3478

Ruxian Yao received the M.Sc. degree in Software Engineering from Tongji Un-iversity, Shanghai, China, in 2009. His research interests include object detection and localization, video retrieval of illuminated area targets, and behavioral feature recognition related to illuminated areas and populations.
https://orcid.org/0000-0002-1314-528X



Jingfeng Gao received the Ph.D. degree in Computer Science from Tokyo University of Agriculture and Technology, Tokyo, Japan, in 2013. His research interests include biometric feature recognition, natural language understanding, machine learning and pattern recognition.
https://orcid.org/0009-0004-1560-1068

Gangqiang Li received the Ph.D. degree in Information and Communication Engineering from Shenzhen University, Shenzhen, China, in 2021. His research interests include new energy generation prediction and control, distributed network data injection attack detection.
https://orcid.org/0000-0002-7363-1060

Haitao Wu received the Ph.D. degree in Computer Software and Theory from WuHan University, WuHan, China, in 2015. His research interests include software design, cloud computing and optimization problems.
https://orcid.org/0009-0001-8689-7386