Knowledge-Based Systems 227 (2021) 107187

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Deep learning in ECG diagnosis: A review ☆

Xinwen Liu [a], Huan Wang [b,*], Zongjin Li [a], Lang Qin [a]

[a] Glasgow College, University of Electronic Science and Technology of China, Chengdu, China
[b] School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu, China

ARTICLE INFO

Article history:
Received 2 November 2020
Received in revised form 28 May 2021
Accepted 30 May 2021
Available online 2 June 2021

Keywords:
Electrocardiogram
Deep learning
Stacked auto-encoders
Deep belief network
Convolutional neural network
Recurrent neural network

ABSTRACT

Cardiovascular disease (CVD) is a general term for a series of heart or blood vessels abnormality that serves as a global leading reason for death. The earlier the abnormal heart rhythm is discovered, the less severe the sequela and the faster the recovery. Electrocardiogram (ECG), as a main way to detect the electrical activity of heart, is a very important harmless means of predicting and diagnosing CVDs. However, ECG signal has characteristics of complex and high chaos, making it time-consuming and exhausting to interpret ECG signal even for experts. Hence, computer-aided methods are required to relief human burden and reduce errors caused by tiredness, inter- and intra-difference. Deep learning shows outstanding performance on ECG classification studies recent few years. Its hierarchical architecture enables higher-level features obtained and its strong ability to feature extraction contributes to classification project. Latest studies can achieve higher accuracy and efficiency than manual classification by experts. In this paper, we review the existing studies of deep learning applied in ECG diagnosis according to four typical algorithms: stacked auto-encoders, deep belief network, convolutional neural network and recurrent neural network. We first introduced the mechanism, development and application of the algorithms. Then we review their applications in ECG diagnosis systematically, discussing their highlights and limitations. Our view about future potential development of deep learning in ECG diagnosis is stated in the final part of this paper.

1. Introduction

Cardiovascular disease (CVD) is a collective term for disorders related to the heart or blood vessels. According to the statistics provided by American Heart Association in 2019, CVDs had become a global dominant cause for death. More than 17.6 million death was caused in 2016 and it is estimated that this figure will reach 23.6 million in 2030 [1]. CVDs may cause blockage of blood vessels and formation of blood clots, which can lead to cerebral or cardiac ischemic necrosis, giving rise to stoke and myocardial infarction. Due to the long-term poor blood pumping of the heart, all organs in the body may be congested and deprived of oxygen, suffering from different degree of damage [2].

Electrocardiogram (ECG) is one of the most commonly used tools for clinical diagnosis in cardiovascular health due to its simplicity, low cost and non-invasive nature. For example, for the patients who suffer from acute heart failure, only in 7.5% of patients have normal ECG [3]. ECG is an object of analysis that detected by an ECG machine that can reflect electric potential change generated by the heart during each cardiac cycle. The heart contracts in a rhythmical manner with regular excitement of myocardium, pumping blood throughout the body. In the process myocardium contraction, slight current is generated by heart and conducted to body surface, causing potential changes in each part of the body. An ECG is be obtained by measuring the potential change through the electrodes from different parts of the body tissue, and recorded with an electrocardiograph or a vector electrocardiograph. In this way, the abnormal rhythm and activity of heartbeat can be shown. Hence, heart diseases or diseases that damage myocardial function can be diagnosed, including arrhythmia, myocardial infarction, coronary heart disease and part of medical co-morbidities such as diabetes and high blood pressure. ECG can also serve as a predictor of coronary heart disease, cardiovascular disease and congestive heart failure. Early detection of such diseases is necessary, because some of them are associated with increased risk of stroke or even sudden-death. There are studies demonstrating that ECG is of importance in predicting both short- and long-term outcomes. For example, for patients suffering from myocardial infarction, the sooner is the abnormal heart rhythm detected, the greater is the chance of avoiding threats of life and recovery [4]. In addition to basic traditional diagnosis and monitoring, ECG is currently employed in

☆ This work was supported by the Innovation Fund of Glasgow College, University of Electronic Science and Technology of China and in part by the Sichuan Science and Technology Program under Grant 2020108.

* Corresponding author.
E-mail addresses: xinwenliu.carrie@qq.com (X. Liu), wh.huanwang@gmail.com (H. Wang), lizongjin.alan@qq.com (Z. Li), qinlang51@126.com (L. Qin).

https://doi.org/10.1016/j.knosys.2021.107187
0950-7051/

telemedicine and home-care monitoring. Home-care monitoring plays a role in patients who suffered from cardiovascular disease, providing timely diagnosis with low price [5]. The telemedicine framework developed by Costa and Oliveira [6] reached 2600 examinations/month in their first 6 months and the most remote client is 85 km away.

Therefore, fast and accurate ECG diagnosis is necessary clinically. However, ECG signal has characteristics of high noise and high complexity, making it time-consuming and labor-intensive to identify certain diseases. Another problem is the individual variability [7], for example, every slight movement produces a baseline drift which is a low-frequency interference for ECG signal and the magnitude of electric potential measured changes with the placement of electrodes [8]. Furthermore, signal interpretation is a tiresome and complicated task, so there is probability of subjective uncertainty and human error in the process of analysis even for experts trained for years. Therefore, it is of great value to attach importance to the research and development of computer-aided method. Computer-supported analysis is able to analyze ECG signal in a more accurate and faster way without difference caused by inter-operators and operator-specific.

Computer-aided ECG interpretation system was first developed in 1960s [9]. The fully automatic system of the traditional intelligent algorithm for ECG interpretations includes three main steps: data preprocessing, feature extraction and classification. Data is denoised and padded or cut to make them into signal segments of the same length in the preprocessing stage. The feature extraction phase is the key part for ECG signal classification. Features can be extracted from the morphology of the ECG signal in the time and frequency domain or directly from the heart rhythm. Finally, the signals are classified into different types of heartbeat or disease according to the features extracted. The ultimate goal is to design algorithms that with high accuracy, efficiency and robustness and are able to reduce doctors' burden.

Deep learning, as a computer-aided method with strong ability to feature extraction, managed to achieve high accuracy in ECG signal classification [10]. Deep learning is achieved by building hierarchical artificial neural networks [11]. The simple non-linear modules of each layer allow deep learning with great advantage in processing complex non-linear signals such as ECG signal. Passing through each layer, more abstractive and high-level features can be gained [12], which contributes to the high classification accuracy. As a result, compared with conventional machine learning, deep learning has better ability to representation learning of intricate data with large samples. LeCun et al. [13] concluded that deep learning discovers intricate structure in large datasets by using the backpropagation algorithm to indicate how a model should change its internal weight values that are used to compute the representation in each layer. This enables deep learning to have fine fault tolerance and prevent the misjudgment caused by overfitting. Deep learning can automatically complete the task of feature extraction and classification by imitating the general-propose learning of human brain while these need human engineers to design in the past. In this way can it learn the implicit knowledge that was only mastered by experts in the past, which means human burden can be greatly reduced. What is more, the progress of Center Processing Unit (CPU) and Graphic Processing Unit (GPU) performance reduces the training and execution time dramatically [14]. This allows deep learning to have large sets of data trained and to apply more complexed algorithms, giving it greater development potential.

Started from the use of greedy layer-wise pretraining, stacked auto-encoders (SAE) and deep belief network (DBN) serve as early typical methods in deep learning field. Following this, convolutional neural network (CNN) becomes the most popular algorithm with the success in visual recognition and are extended to various fields. Recurrent neural network (RNN) is recursive networks popular for its outstanding performance in time series data processing, which is quite suitable for ECG signal classification task. Long short-term memory (LSTM) outperforms traditional RNN in long-term dependence, becoming a more common method. A number of deep learning methods have been applied to feature extraction and classification in ECG interpretation. SAE is an unsupervised way to extract features by encoding and decoding the input ECG segments. DBN can either works as SAE unsupervised or serve as a classifier in supervised manner. Both of these two methods often fine-tuned with active learning so that the important data can be focused. Since ECG is 1-D signal, it is input directly into 1-D CNN or transformed into image and processed by 2-D CNN. As for RNN, ECG is usually processed as time-series signal. There are networks combining CNN and RNN to learn both space and time information, which is proved to be efficient in ECG classification tasks. The concept of SAE is executed with LSTM and CNN by changing auto-encoder into LSTM and convolutional layer while remaining the procedure of encoding and decoding of the network. Most of deep learning methods have achieved accuracy which is higher than manual classification with little or no human assistance, reducing the burden on professionals efficiently. Compared with traditional methods, deep learning can process raw ECG signals directly without the requirement of preliminary feature extraction, allowing higher efficiency and simpler steps for usage. However, there is limitation on dealing with imbalanced input due to deep learning's strong dependence on input, while some kinds of heartbeat samples exist less in reality. Also, the high complexity and large amount of calculation makes the application in wearable devices still a difficulty. Moreover, ECG collected from reality always accompanied with noise and the preprocess of denoising requires much calculation resources. Given these existing problems, more robust models with less parameters should be constructed. In addition to the theoretical progress, the application of methods in reality should receive more attention. There have been some reviews focusing on ECG signal and deep learning. Murat et al. [15] only focused on heartbeat classification task. Hong et al. [16] went through deep learning's application in all field of ECG signal more than disease diagnosis. Faust's review contains researches on a range of physiological signals where only a small space is used to introduce ECG [17]. Here we can see that no comprehensive review of deep learning's application in ECG diagnosis has been carried out, while this direction has practical significance and development potential. Hence, we adopt the position that a review in studies of deep learning method applied in ECG diagnosis is necessary. In this paper, state-of-the-art studies are reviewed in a systematic way and their characteristics are highlighted. Meanwhile, an overview of deep learning on ECG diagnosis is illuminated by pointing out problems and potential development.

To explain the importance of deep learning in the field of ECG analysis, this paper is organized as follows. The basic knowledge of ECG signal and typical cardiovascular diseases is represented in Section 2 and the common-used databases are introduced in Section 3. In Section 4, the theoretical background of deep learning in is introduced and the relevant researches are presented in Section 5. To make it clear, we present the studies according to 4 classic deep learning architectures: SAE, DBN, CNN and RNN. Section 6 discussed the limitation and future opportunities of deep learning in ECG diagnosis, key issues of deep learning and other network architectures that are promising in ECG field. Finally, a brief conclusion is drawn in Section 7.

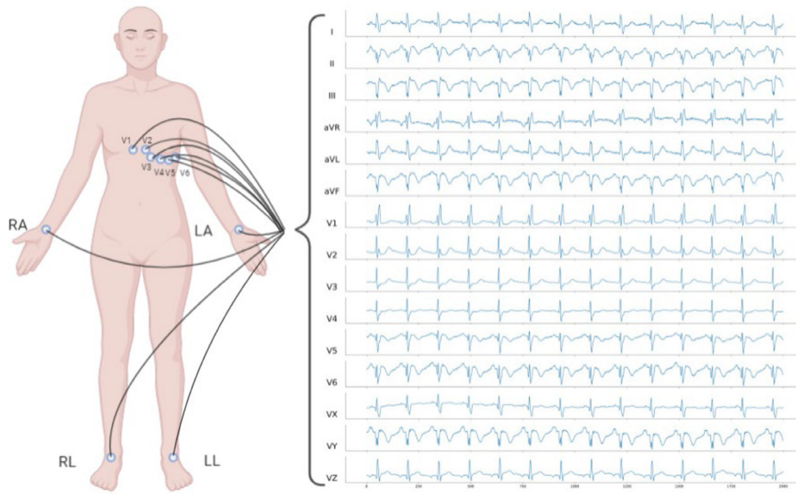

Fig. 1. Schematic diagram of the 12-lead ECG: 4 electrodes on limbs (RA, LA, RL, LL) and 6 electrodes on chest (V1–V6)..

2. ECG signal

The heart contracts in a rhythmical manner with regular excitement of myocardium, pumping blood throughout the body. In the process myocardium contraction, slight current is generated by heart and conducted to body surface, causing potential changes in each part of the body.

For ECG diagnosis, it is 12 leads that utilized most widely, in which 10 electrodes are applied and one of them serves as a reference to others (usually is the right leg) [18]. The configuration of 12 leads can be seen in Fig. 1 (four electrodes are placed on right leg (RL), right arm (RA), left leg (LL) and left arm (LA) respectively and six on the chest (V1–V6)). The limb leads reflect the condition of the cardiac vector plane and the chest lead system reflects the horizontal one. A 12-lead ECG can provide enough information for the diagnosis of various diseases, and can also be used to determine the location of the abnormality. However, most of the data that is used to test the performance of computer-aided methods is 2-lead ECG, for it is enough for diagnosis of certain disease and ECG beats classification [19].

An ECG of one cardiac cycle contains three waves, two intervals and two segments. Three significant waves denote the heart's three different electrical phenomenon in one heartbeat cycle: P wave (atrial depolarization), QRS complex wave (ventral depolarization) and T wave (repolarization) [20]. The length of two intervals (PR interval and QT interval) indicates the time it takes heart to complete the corresponding electrical change. Fig. 2 shows the constitution of a standard heartbeat. According to the abnormal of ECG signal, most heart diseases can be diagnosed. Typical heart problems are myocardial infarction, coronary artery disease and multiple types of arrhythmia. Arrhythmia can be divided into sixteen types, including tachycardia, bradycardia, supra-ventricular arrhythmia, ventricular arrhythmia, atrial fibrillation and so on, but most researches preferred not to classify it into so many categories. In the case of multiple leads, the specific location of myocardial infarction can be illuminated by ECGs.

3. Database

A few studies run their networks on datasets of more than one database to prove validity of the methods, which also make it easier to compare experiment results horizontally. Most of the studies we reviewed used the following databases. The specific parameters of commonly used databases are listed in Table 1.

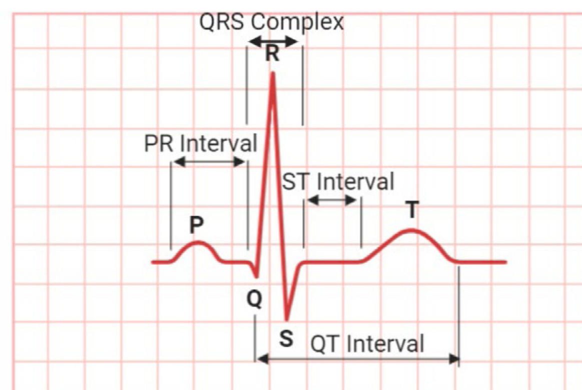

Fig. 2. The constitution of a single heartbeat, including P wave, QRS complex and T wave.

MIT-BIH Database is the most widely used database. It consists of 9 sub-databases of different diseases, among which the most popular one is the MIT-BIH Arrhythmia Database [21]. A collection of 48 fully annotated 30-minute two-lead ECGs is accessible in MIT-BIH Arrhythmia Database. MIT-BIH Arrhythmia Database can be divided into two groups. The first group contains representative samples of arrhythmia waveforms and artifact that might encountered during diagnosis. The second group includes complex ventricular, junctional, and supraventricular arrhythmias and conduction abnormalities.

European ST-T Database and Long Term AT Database both focus on the abnormality of ST segments and T waves. European ST-T Database was developed by European Society of Cardiology. Each subject was diagnosed with myocardial ischemia and each ECG record is two-hour long. Long-Term ST Database is composed of 86 records of paroxysmal or sustained atrial fibrillation, lasting for 21 to 24 h.

St.-Petersburg Institute of Cardiological Technics 12-lead Arrhythmia Database is collected from the patients who are diagnosed or suspected with coronary artery disease. This database is annotated by algorithm automatically and corrected manually.

PhysioNet Computing Cardiology Challenge 2017 dataset contains 8528 under-recognized short-term ECG recordings focusing on rhythm level: normal, AF, other, and noise. In addition, 3658 records are for private scoring. The data was sampled by an AliveCor healthcare device at 300 Hz. PhysioNet Computing

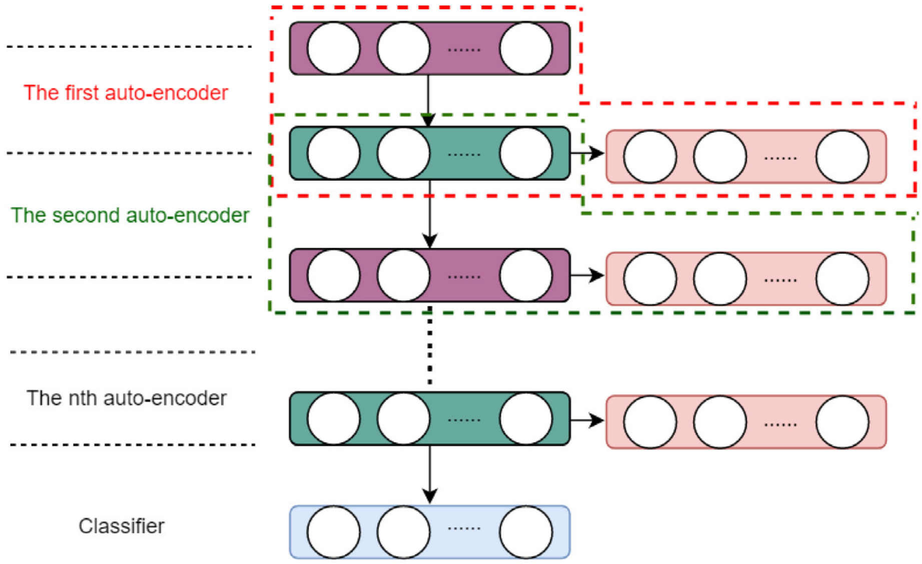

Fig. 3. The structure of SAE. Hidden layer of the first auto-encoder (in the red dotted frame) serves as input layer of the second one (in the green dotted frame). Pink blocks here refer to output layers of each auto-encoder..

Cardiology Challenge 2017 dataset is the second most popular dataset in recent few years and the newest one among all databases we listed in Table 1.

4. Theoretical background of deep learning

Deep learning is a subordinate branch of machine learning. It can learn and construct intrinsic features from neurons in numerous hidden layers of a neuron network. Deep learning network is developed based on neural network. Warren McCulloch and Walter Pitts came up with the concept of artificial neural network and the mathematical model of artificial neuron, thus opening the era of artificial neural network research. Following this, Frank Rosenblatt proposed perceptron [22], which is the origin algorithm of neural network proposed to deal with binary classification problem. Minsky and Papert [23] found two significant drawbacks of perception: XOR loop cannot be handled and the unaffordable calculation amount. After that, the development of neural networks experienced a long period of stagnation until the back propagation algorithm was raised by Rumelhart et al. [24]. This algorithm solved the problem of XOR, making it possible to train network of multiple layers. At the beginning of the 21st century, with the help of GPU and distributed computing, the computing power of computers has been greatly improved. Under this condition, Hinton and Salakhutdinov [25] effectively trained a deep belief network using greedy layer-wise pretraining. This technique was later extended to different neural networks. Since then, the revival of interest in deep learning has taken place. This renewal started with unsupervised feature learning including stacked auto-encoders (SAE) and deep belief network (DBN). Following these, convolutional neural network (CNN) came back into sight with AlexNet [26] and recurrent neural network (RNN) regained attention following CNN. Deep learning has been widely applied to multiple tasks, including semantic segmentation [27,28], object detection [29] and background subtraction [30,31]. Great results have been achieved in these application fields by deep learning though with limitation, showing its outstanding potential.

4.1. Stacked auto-encoders

Stacked auto-encoders (SAE) is one type of deep neural network composed of multiple layers of auto-encoders and a classifier, whose structure is shown in Fig. 3. In order to gain more powerful representation through complicated relationships, there are two other variants of auto-encoder applied in ECG classification: denoising auto-encoder and sparse auto-encoder. Vincent et al. [32] came up with denoising auto-encoder in order to obtain robust feature from a corrupted input and play the role of denoising and applied it to handwritten digit recognition and image recognition. Makhzani and Frey [33] proposed k-sparse auto-encoder which can extract deep features hierarchically while prevent the network from overfitting efficiently, and tested its effectiveness in handwritten digit recognition and object detection.

Auto-encoder is a single unit of SAE composed of one input layer, one hidden layer and one output layer. It encodes the unlabeled input, keeping the lower dimensional representation in the hidden layer, and then decodes the extracted features to reconstruct the input. An auto-encoder is trained by minimizing the error between the input and the reconstruction at the output layer.

For input x, extracted feature c and the reconstructed signal y can be represented by:

$$c = \theta(w_e * x + b_e) \tag{1}$$

$$y = \theta(w_d * c + b_d) \tag{2}$$

where $w_e = w_d = w$, denoting the encoding and decoding weight metric respectively. b_e and b_d represent the bias vectors. $\theta(\cdot)$ is the activation function.

SAE use layer-wised training method. Each auto-encoder is initialized before the neural network is trained. The first auto-encoder utilizes the input to predict itself, finding the initial parameter. Following this, the hidden layer of each auto-encoder also serves as the input layer of the successive auto-encoder. In this way, all the auto-encoders are trained sequentially. After pre-training, training set with labels is used for fine-tuning. By back-propagation, the cost function is minimized and the parameters are fine-tuned.

4.2. Deep belief network

Deep belief network (DBN) is a type of generative neural network composed of unsupervised simple learning modules [25]. Similar with SAE, DBNs are constructed by stacked restricted Boltzmann machines and a classifier as shown in Fig. 4. The

Table 1
Summary of databases, including number and duration of records, sampling frequency, number of leads and heartbeat types.

Database	No. of records	Duration	Sampling frequency	No. of leads	Diagnosis type
MIT-BIH Arrhythmia Database	48	30 min	360 Hz	2	Represent samples of arrhythmia
MIT-BIH supraventricular arrhythmia database (SVBD)	78	30 min	360 Hz	2	Supraventricular arrhythmias
European ST-T database	90	2 h	250 Hz	2	ST segment changes
Long-term ST database	86	21–24 h	–	2	ST segment changes
St.-Petersburg Institute of Cardiological Technics 12-lead Arrhythmia Database	75	30 min	257 Hz	12	Coronary artery disease
PhysioNet computing in cardiology challenge 2017	8528 train, 3658 test	30 s	300 Hz	1	Atrial fibrillation

connections between the top two layers are undirected, which form associative memory. Connections between the other lower layers are directional. A series of improvements based on the original RBM structure have been made. Mnih et al. [34] proposed Conditional Restricted Boltzmann Machines (CRBMs) to deal with multi-label classification and image denoising. Elfwing et al. [35] proposed expected energy-based RBM is to use negative expected energy instead of negative free energy to calculate the output, which greatly improves the RBM learning performance in handwritten digit recognition and object recognition.

Restricted Boltzmann machine (RBM) is one type of bi-directionally connected stochastic neural network that can predict the probability distribution over its input [36]. A RBM is a bipartite graph composed of one visible layer and a hidden one. Visible layer and hidden layer are designed to receive input and extract features respectively [37]. The training purpose of RBM is to retain the probability distribution of input maximumly. An energy function E(v, h) is define:

$$\mathrm{E}(v,\ h)\ =\ v'Wh-\ b_v'v-b_h'h \tag{3}$$

where v and h represents the activation matrix of visible layer and hidden layer respectively and v' is the transpose of v. b_v and b_h are the bias of them. W denotes the weight that the visible units associated with the hidden units. Thus, the joint probability distribution between the hidden units and visible units can be given by the energy function:

$$p(v,h)=\frac{\exp(-E(v,h))}{Z} \tag{4}$$

where

$$Z=\sum \exp(-E(v,h)) \tag{5}$$

is a partition function.

Due to the special structure of RBM, each hidden unit is conditionally independent when the state of visible layer is given and vice visa:

$$p(h|v)=\prod_i p(h_i,v) \tag{6}$$

$$p(v|h)=\prod_i p(v_i,h) \tag{7}$$

After pretraining, backpropagation is used to optimize parameters in RBMs.

There are two steps including pre-training and fine-tuning executed when running a DBN. In the pre-training process of a DBN, hidden units of previous RBM also serves as the visible units of the next one. Each RBM is trained sequentially and greedily, utilizing contrastive divergence (CD) to learn the initial weight of each RBM. Then backpropagation is used to fine-tune the discriminative performance by using labeled data.

DBN has an advantage over model completion for corrupted data while it is hard to be well-trained, making it difficult to achieve high accuracy in classification task.

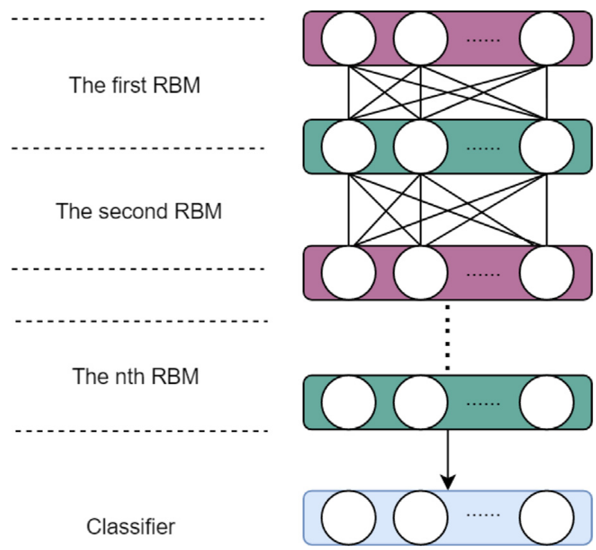

Fig. 4. The structure of DBN, where hidden layer of the first RBM serves as the visible layer of the second RBM. Every point between two adjacent layers is connected.

4.3. Convolutional neuronal network

Convolutional neural network (CNN) is a type of feedforward neural network with hierarchal structure [38]. Instead of fully connected layers of standard neural networks, CNN proposes learning filters that applies operation to each sub-region of input. In terms of network structure, one CNN typically consists of convolutional layers, pooling layers and fully connected layers [39]. Many typical deep learning networks are composed of CNN. Alexnet [26] introduces a new deep structure and dropout method, improving the performance of LeNet-5 by deepen the network. Then VGG-net [40] is proposed with smaller convolution kernel and deeper levels and its performance in image recognition is similar with that of human. Later, DenseNet [41] and ResNet [42] and, each introduces dense connection and group convolution respectively, and show good performance in object detection.

In convolutional layer, a convolution of each sub-region of input with filter kernel is computed, extracting features from the input from the previous layer. For the jth feature map in the lth layer,

$$c_j^l=\theta(\sum_{i\in M_j} x_i^{l-1}*w_{ij}^l+b_j^l) \tag{8}$$

where θ is the activation function and M_j represents the connectivity between c_j^l and feature maps from the previous layer. w_{ij}^l is the weight (or kernel) for the jth feature map and ith filter index and b_j^l is the corresponding bias.

Pooling layer is always placed behind convolutional layer and performs a down-sampling operation to a block. It is applied to reduce the features' size and select the most representative part by highlighting or suppressing features. In each small block of data, a single corresponding output is produced. After the convolutional layers and pooling layers, the obtained features of each sub-region are flattened into a one-dimensional vector as the input of fully connected layer. In this part, input data is mapped into different classes. Then a backpropagation with labeled data and learning rate is used to update parameter with the cost function.

4.4. Recurrent neuronal network

Recurrent Neural Network (RNN) is a type of recursive neural network that performs recursion in the evolution direction of the sequence and all nodes (or memory cells) are connected in a chain. Each node in the network is placed the sequence, in which the output of neuron is utilized for its input, thus making RNN able to obtain output dependent on its previous computation. This is shown to be very useful in processing time series data [43].

However, the effect of traditional RNN weakens when facing long-term dependencies. Hence, long-short term memory (LSTM) is proposed by Hochreiter and Schmidhuber [44] to overcome this problem. Each node in traditional RNN is replaced with memory cell in LSTM. The core idea behind the LSTM is updating memory in memory cells continuously, allowing useful information stored and redundant one abandoned. Each input needs go through three gates: the forget, input and output gate as shown in Fig. 5.

For the input x_t and the output at the previous time step h_{t-1},

$$f_t = \partial(w_f[h_{t-1}, x_t] + b_f) \tag{9}$$

$$i_t = \partial(w_i[h_{t-1}, x_t] + b_i) \tag{10}$$

where f_t represent the forget gate and i_t is the input gate. ∂ is the sigmoid activation function that can return a value between 0 and 1 to decide whether the information is stored. Then a new memory $\widetilde{c_t}$ is added and the original memory c_{t-1} is updated to c_t:

$$\widetilde{c_t} = \tanh(w_c[h_{t-1}, x_t] + b_c) \tag{11}$$

$$c_t = f_t c_{t-1} + i_t \widetilde{c_t} \tag{12}$$

Finally, the output h_t is computed on the base of the output gate o_t:

$$o_t = \partial(w_o[h_{t-1}, x_t] + b_o) \tag{13}$$

$$h_t = o_t * \tanh(C_t) \tag{14}$$

LSTM is the most popular RNN algorithm that used in ECG diagnosis models. By combining forward LSTM and backward LSTM, bidirectional LSTM (Bi-LSTM) was prposed to determine any point in the sequence by information from both future and past, which is tested to be effective in phonemes [45]. Following Bi-LSTM, Cho et al. [46] proposed Gated Recurrent Unit (GRU) and applied it to WMT'14 translation task, which can perform a similar function with LSTM while requiring less computation source. Fig. 5. The structure of a single LSTM module where h_t and x_t refers to the output and input of LSTM respectively.

5. Researches

In this section, deep learning methods of classifying ECGs are introduced according to their algorithm. In each part, a table is listed to show the research work's application, Deep Learning (DL) algorithm, the database used for the study, and the study results. The studies requiring more explanation are further explained following the table.

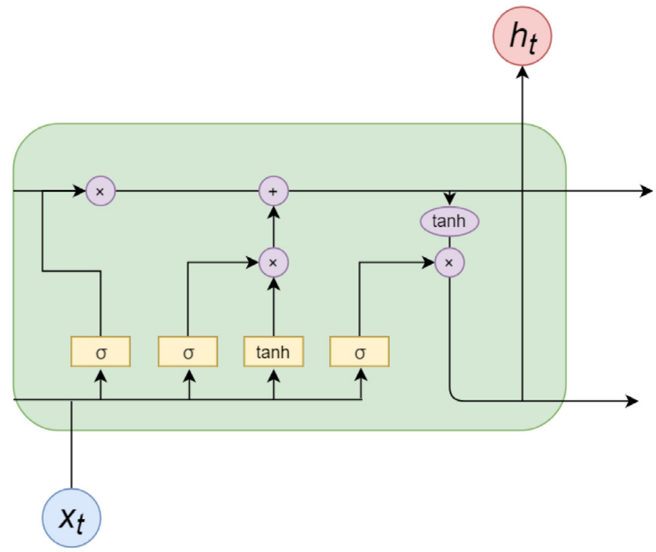

Fig. 5. The structure of a single LSTM module, where h_t and x_t refers to the output and input of LSTM respectively.

5.1. Stacked auto-encoders

The studies on SAE on classification tasks of ECG are listed in Table 2. SAE are often applied with sparsity constraint or denoising function. Active learning (AL) is introduced in some researches [47,48]. Besides, the ideas of auto-encoder are utilized by CNN and LSTM which are used to construct the encoding and decoding parts [52,53].

AL is frequently applied to the fine-tune process of ECG classification studies. AL criteria is frequently used to select the most ambiguous and relevant samples for experts to label for each iteration and the labeled data is put back to augment the training set in next iteration. This reduces experts' burden by only requiring label of the most valuable part. Rahhal et al. [47] used entropy and BT as metrics respectively in interpatient heartbeat classification and made a comparison between the two criteria. Hanbay [50] applied entropy as AL criteria in arrhythmia classification. For this study, the eigenvalue magnitude of ECG and weighted interval features are combined into a feature vector, serving as the input for the stacked denoising auto-encoders. In addition to AL, a training method divided into global training and patient-specific training is used for long-term ECG diagnosis [48]. To add, the input for this lab is the time–frequency spectrogram generated by MFSWT.

Different types of SAE can be combined in one network to complete different tasks. Nurmaini et al. [54] used SDAE to denoise the original signal, followed by SAE to extract features from the denoised data. DNN is utilized to fine-tune the whole model. Good performance is achieved in both accuracy and sensitivity, though the efficiency is low due to its long training time.

By realizing the idea of auto-encoder on CNN, Oh et al. [55] proposed a modified U-net applied in abnormal six kinds of heartbeat classification, where samples are compressed and upsampled for 3 times. However, this approach is limited by its speed due to the large amount of calculation and subtle changes of ECG waves can result in mistakes. Hou et al. [53] simply applied the SAE's structure by displacing both the encoding and decoding layer with LSTM modules.

SAE can achieve good results in ECG diagnosis, but a higher efficiency is generally acquired. Denoised stacked auto-encoders show advantage in building a robust model. As an unsupervised learning method, it is less dependent on the given label, which

Table 2
Models based on stacked auto-encoders, including methods that used stacked auto-encoders and its idea.

Author	Application	DL algorithm	Database	Result	Remark
Rahhal et al. [47]	Heartbeat classification (interpatient)	Stacked denoising auto-encoders with sparsity constraint	MIT-BIH Arrhythmia Database, SVBD Database, INCART	Accuracy 99.7%	Fine-tuned with AL;
Luo et al. [48]	Heartbeat classification and monitoring	Stacked denoising auto-encoders	MIT-BIH Arrhythmia Database	Accuracy 97.5%	Global training and local training
Yang et al. [49]	Arrhythmia classification	Stacked sparse auto-encoders	MIT-BIH Arrhythmia Database	Accuracy 99.5% Sensitivity 99.22% positive predictivity 99.37%	–
Hanbay [50]	Arrhythmia classification	Stacked denoising auto-encoders	MIT-BIH Arrhythmia Database	Accuracy 99.5% for VEB, 99.8% for SVEB	Fine-tuned with AL; Eigenvalue magnitude as input
Farhadi et al. [51]	Atrial fibrillation classification	Stacked sparse auto-encoders	MIT-BIH Arrhythmia Database	Accuracy 99.5%	–
Oh et al. [52]	Arrhythmia classification	Modified U-net	MIT-BIH Arrhythmia Database	Accuracy 97.2%	–
Hou et al. [53]	Arrhythmia classification	LSTM-based auto-encoder	MIT-BIH Arrhythmia Database	Accuracy 99.74% Sensitivity 99.35% Specificity 99.84%	TP for minority classes are poor due to data imbalance
Nurmaini et al. [54]	Contaminated arrhythmia classification	Stacked denoising auto-encoders, stacked auto-encoders, DNN	MIT-BIH Arrhythmia Database; MIT-BIH noise stress test database	Accuracy, sensitivity, specificity and F1-score of 99.34%, 93.83%, 99.57% and 91.44%	SDAE denoise; SA extract features; DNN fine-tune

means their model can be more generalized. Up to this point, we have completed the discussion on ECG diagnosis of unsupervised learning methods. In next two parts models based on supervised learning will be reviewed.

5.2. Deep belief network

Deep belief network (DBN) can be used for both supervised and unsupervised learning. Compared with stacked auto-encoders, there are less studies focusing on the application of DBN on ECG diagnosis. In addition to ECG analysis, DBN is applied to false alarm reduction task and shown to be efficient. Brief discussion is done in this section and more details of the studies are shown in Table 3.

Multiple models can be used in one single classification task. Altan et al. [56] proposed a multistage classification system composed of four different structured DBN modules. Each of N, S, V, F type heartbeat is separated when data passing through each module. For this model, SODP features of ECG signal serves as input. The network proposed by G et al. [57] is composed of three models: Primary Model, S Model, and V Model. Similar with [56], primary Model is trained to classify N, V, S, F type heartbeat while the purpose of S Model and V Model is just to separate S and V type respectively. DBN is also fine-tuned with active learning (AL) in this study. By classifying S and V type individually, the problem of data imbalance caused by less S and V type heartbeat can be relieved.

Altan et al. [58] pays more attention to pre-process in CAD diagnosis. Empirical Mode Decomposition is used to extract Instinct Mode Function (IMF) from 15-second short-term ECG segments. Then instantaneous frequency features are obtained by applying Hilbert Transform (HT) to IMF. DBN is tested to be the classifier that can show the best performance. Mathews et al. [59] tried to use simple features and lower sampling rate at 114 HZ.

It is worth mentioning that DBN is used to differentiate acceptable ECG segments from poor-quality ECGs and achieves significant results [60]. The ECG signal is contaminated by electrode motion artifact signal and a DBN is trained to judge the quality of ECGs by an established SNR threshold. The accuracy of noisy signal can recover to 81% from 58.7% while the value is 87% for clean data, which managed to reduce false alarm efficiently.

5.3. Convolutional neural network

CNN is widely applied in ECG diagnosis tasks in recent few years and outstanding performance has been achieved. In some researches, CNN is used in conjunction with RNN modules, such as LSTM, and we attribute such neural networks to the next section (RNN applications), so they will not be discussed in this section. Studies included in this section is listed in Table 4.

Besides the use of global and local training for long-term monitoring [62], a TDCNN a Tuned Dedicated CNN (TDCNN) for wearable devices is developed [63]. A General CNN (GCNN) trained with interpatient information is stored by the devices, followed by the fine-tune with patient-specific ECG. Compared with the previous studies, the storage space required by this approach is only 1/5 of Kiranyaz et al. [62] and the training time is shortened efficiently.

Since ECG is one-dimensional signal, 1-D CNN can be used without special process. By transforming ECG into time–frequency spectrogram, 2-D CNN can also be applied. In this way, representative models of 2-D CNN such as VGG can be realized in ECG signal tasks [64–66]. Apart from common mathematic conversion method such as Short Time Fourier Transform [64] and Frequency Slice Wavelet Transform [66], coupled convolution is also able to perform the same function to transform ECG into 2-D signal [65]. Coupled convolutional structure can help in a better fitting capability as well [67].

It can be seen that inter-patient diagnosis requires a more complicated network and more training data [64,68]. Rajpurkar et al. [64] proposed a 34-layer network trained by a self-constructed database containing more than 30,000 unique patients. To overcome overfitting caused by the depth of network, shortcut is used between blocks. The result of this study is compared with average cardiologist performance, and is shown to be better in most rhythm. For the Multiple-Feature-Branch CNN (MFB-CNN) [68] where each of 12-lead signal corresponds to each branch, the accuracy for patient-specific model is lower in both detection and location compared with class-based model. Although the accuracy for patient-specific model is lower in both detection and location compared with class-based model, it is proved that MFB-CNN can limit the impact caused by interpatient variability. Niu et al. [69] uses RR interval to represent temporal

Table 3
Models based on deep belief network, including disease diagnosis and false alarm reduction..

Author	Application	DL algorithm	Database	Result	Remark
Altan et al. [56]	Arrhythmia classification	DBN-based multistage algorithm	MIT-BIH Arrhythmia Database	Accuracy 96.10%	SODP features as input
Altan et al. [58]	Coronary artery classification	DBN	24-hour filtered ECG signal from PhysioNet	Accuracy 98.05% Sensitivity 96.02% Specificity 98.88%	15-second ECG segments as input
Mathews et al. [59]	Arrhythmia classification	DBN	MIT-BIH Arrhythmia Database	Accuracy 93.63% for SVEB and 95.87% for VEB at 114 Hz	–
G et al. [57]	Heartbeat classification	DBN	MIT-BIH Arrhythmia Database,MIT-BIH SVBD Database	Accuracy 97.5% for SVEB and 98.6% for VEB at 114 Hz	Fine-tuned with AL
Taji et al. [60]	Atrial fibrillation (noisy and clean)	DBN	MIT-BIH AFDB Database	For SNR = −20 dB, accuracy rise from 58.7% to 81%	To reduce false alarm rate
Song et al. [61]	Heartbeat classification	DBN	MIT-BIH Arrhythmia Database	Accuracy 98.49%	–

features, concentrating with signal symbolic metric as input. This contributes to a more comprehensive study of model and performs well in inter-patient arrhythmia diagnosis.

Besides some popular area such as arrhythmia classification and Atrial fibrillation, CNN algorithm is used in various aspects in ECG diagnosis: shockable and non-shockable ventricular arrhythmias [70], myocardial infarction detection [71] and localization [68]. CNN can also help in more precise classifications, for example, 14 classes irregular heart rhythm [64] and 17 classes arrhythmia heartbeat type. However, since the learning of CNN is quite dependent on data, CNN models can easily be led astray by imbalanced data or incorrect data labels.

5.4. Recurrent neural network

RNN has an advantage in dealing with time series data, which enable network to show good performance in ECG classification task. In this section, all RNN based method will be listed in Table 5 and discussed in the following passage.

Apart from used alone, RNN modules are commonly combined with other deep learning algorithm. To be precise, LSTM module is frequently used to analysis sequential features that extracted by CNNs [52,81,82,82,83] and this method is proved to be efficient. It worth mentioning that Bidirectional LSTM is able to learn from both previous context and future one by processing data in two directions [84]. This contributes to long-term diagnosis. Guo et al. [85] proposed a model with GRU module and CNN, focusing on inter-patient arrhythmia diagnosis. Since GRU cell only has two gates (an update gate and a reset gate), its computational efficiency is higher than normal LSTM while it can play the same function. Yao et al. [82] combined CNN-LSTM models with attention modules. Attention modules can locate the most informative part, which reduce the number of parameters by half and improve its efficiency. Due to the compatibility of attention module, varied-length ECG segments can be handled. [83] designed a comprehensive LSTM-CNN model that can deal with several kinds of CVDs including coronary artery, myocardial infarction and congestive heart failure. DELM-LRF [86] is also introduced to highlight the abnormal segment of ECG.

Different ECG features are used to enable the network to learn more comprehensive information of signal. In this situation, blended models are developed with RNN so that both representative and sequential features can be learned in one network. Wang et al. [87] proposed an updatable RNN network with a morphology part and a temporal part, extracting features from morphological vector and Premature-or-Escape-Flag (PEF) respectively. For each iteration the training set is updated with AL. Saadatnejad et al. [88] proposed a blended network made up of two separate RNN-based models (model α and model β). Different from each model corresponding to one feature, both RR interval and wavelet features serves as the input for each model. Model α combines two extracted features in the final fully-connected layer while model β concatenates them at first as the input. Qiao et al. [86] proposed a DELM-LRF-BLSTM network on atrial fibrillation prediction. DELM-LRF is composed of three feature extraction stages, each of which consists of a convolution layer and a max pooling layer. In this network, the square layer of DELM-LRF is replaced by Bi-LSTM and the number of tunable parameters is proved to be much smaller than single LSTM.

LSTM is shown to be useful in improving efficiency as a part of the models and help network learn ECG's contexts. Though good performance can be achieved in pure stacked LSTM networks [14,75,89,90], computational cost is high and training time is long because its large time span brings about a large number of parameters.

6. Discussion

We have analyze literature reports that apply deep learning to ECG diagnosis systematically. Therefore, in this section, current limitations and future potential directions of both deep learning in ECG diagnosis and deep learning method itself will be discussed.

6.1. Limitation and opportunities of deep learning in ECG diagnosis

Data imbalance is a common and important issue. Due to the characteristics of ECG, some kinds of heartbeat are harder to be recorded. Classes that are trained with less samples can show an obvious inaccuracy compared with other classes. A few studies has come up with ideas dealing with this problem, including special network structures [57,88], special focus on minor classes [47,50], [14,57,76,81] and loss function [80]. However, in some researches, the minority is reused for the balance of data. This may give rise to high accuracy in certain record being mistaken as state-of-the-art.

We consider the high complexity in deep neural networks as a main obstacle for the technology to be carried out in practice. Especially for ECG monitoring of wearable devices, high efficiency is of great importance. Besides basic method such as attention module [82], shortcut [64,93] that can reduce the number of parameters, lightweight blended models [86,88] are designed to relieve the burden. Compression of model can also be used at price of small decrease in results [94].

Table 4
Models based on convolutional neural network, hybrid models with other 3 algorithm are not included.

Author	Application	DL algorithm	Database	Result	Remark
Kiranyaz et al. [62]	Arrhythmia classification and monitoring	1-D CNN	MIT-BIH Arrhythmia Database	Accuracy 97.05%	–
Acharya et al. [70]	Shockable and non-shockable ventricular arrhythmias	1-D CNN	MIT-BIH Arrhythmia Database; MIT-BIH VFB Database; CUBD	Accuracy 93.18% Sensitivity 95.32% Specificity 91.04%	–
R. et al. [72]	Arrhythmia classification	1-D CNN	MIT-BIH Arrhythmia Database	Accuracy 94.03% for original data; 93.47% for clean data	–
Acharya et al. [73]	Myocardial infarction detection	1-D CNN	PTB Database	Accuracy 93.53% for original data; 95.22% for clean data	–
Acharya et al. [74]	Coronary artery classification	1-D CNN	Fantasia (for Normal); St.-Petersburg INCART 12-lead arrhythmia (for CAD)	accuracy 94.95%, sensitivity 93.72%	2 CNN networks for 2-second segments and 5-second one
Acharya et al. [20]	Arrhythmia classification	1-D CNN	MIT-BIH Arrhythmia Database; MIT-BIH VFB Database; CUDB	accuracy, sensitivity, and specificity of 92.50%, 98.09%, 93.13% for 2-second segments; 94.90%, 99.13%, 94.90% for 5-second segments	2 CNN networks for 2-second segments and 5-second one
Rajpurkar et al. [64]	Irregular heart rhythm classification (14 classes)	2-D CNN (34 layers)	Self-constructed dataset of 30,000 unique patients	Accuracy 77.6% in sequence level, 80.9% in set level; exceed cardiologist performance	Add shortcut between models; for varied length segment
Yıldırım et al. [75]	Arrhythmia classification (17 classes)	1-D CNN	MIT-BIH Arrhythmia Database	Accuracy 91.33% 0.015s per sample	–
Li et al. [63]	Arrhythmia classification and monitoring	Tuned Dedicated CNN (TDCNN)	MIT-BIH Arrhythmia Database	Accuracy 96.89%	Only trained GCNN is stored in devices instead of data
Zhai and Tin [65]	Arrhythmia classification	2-D CNN	MIT-BIH Arrhythmia Database	Accuracy 97.3% Sensitivity 89.3% Specificity 98.0% for S beats	S beat set are pre-selected by network; Coupled-convolution structure
Liu et al. [68]	Myocardial infarction detection and localization	Multiple Feature Branch CNN (MFB-CNN)	PTB Database	Detection and localization accuracy: 99.95% and 99.81% for class-based; 98.79% and 94.82% for patient-specific	12 branches for 12-lead ECG
Jiang et al. [76]	Heartbeat classification	CNN context-feature module (CTFM)	MIT-BIH Arrhythmia Database; European ST-T Database; MIT-BIH ST Change Database	Accuracy 98.4%	Borderline-SMOTE (BLSM) and 2 phase training also used for imbalanced data
Baloglu et al. [77]	Myocardial infarction detection	1-D CNN	PTB Database	Accuracy 99.78%	–
Huang et al. [66]	Arrhythmia classification	2-D CNN	MIT-BIH Arrhythmia Database	Accuracy 99.0%	–
Wang [78]	Atrial fibrillation classification	CNN Elman neural network	MIT-BIH AF Database	Accuracy 9.4% Sensitivity 97.9% Specificity 97.1%	–
X. and H. [67]	Heartbeat classification	1-D CNN	MIT-BIH Arrhythmia Database	Accuracy 99.43%	Coupled-convolution structure
D. et al. [79]	Atrial fibrillation detection based on non-standardized single-lead ECG	1-D CNN	Self-collected patched-based ECG	Accuracy 93.1% Sensitivity 93.1% Specificity 93.4%	–
Romdhane et al. [80]	Heartbeat classification	CNN	INCART Datasets	Accuracy 98.41% Precision 98.37% F1 98.38%;	Focal loss for data imbalance
Niu et al. [69]	Interpatient arrhythmia classification	MPCNN	MIT-BIH Arrhythmia Database	Accuracy 96.4% F1 76.6% for SVEB, 89.7% for VEB	Signal symbolic matrix concentrate RR interval matrix as input

Another problem is that ECG includes noise from a number of sources in reality. General resolution used in researches is to denoise data with pre-processing [48,70,84,94], which augments computational cost and reduces efficiency. Hence, more robust models should be built. In addition to this, since experts may make mistakes in the manual labeling process, label noise problem should also be taken into consideration. Unsupervised or weakly supervised learning can also be a way to explore in this aspect.

For future development of deep learning in ECG diagnosis field, the construction of a standard and comprehensive database

Table 5
Models based on recurrent neural network, including CNN-LSTM hybrid networks.

Author	Application	DL algorithm	Database	Result	Remark
Yildirim [91]	Arrhythmia classification	LSTM	MIT-BIH Arrhythmia Database	Accuracy 99.39%	–
Tan et al. [81]	Coronary artery classification	LSTM, CNN	Fantasia; St.-Petersburg INCART 12-lead arrhythmia	Accuracy 99.85%	–
Guo et al. [85]	Interpatient arrhythmia classification	GRU, CNN	MIT-BIH Arrhythmia Database	F1 score 61.25 for SVEB, 89.75 for VEB	–
Faust et al. [17]	Atrial fibrillation detection	LSTM	MIT-BIH Atrial Fibrillation Database	Accuracy 99.77% with blindfold validation	–
Faust et al. [89]	Arrhythmia classification	LSTM, CNN	MIT-BIH Arrhythmia Database	Accuracy 98.10% Sensitivity 97.50% Specificity 98.70%	–
Wang et al. [87]	Heartbeat classification	LSTM	MIT-BIH Arrhythmia Database;St.-Petersburg INCART 12-lead arrhythmia; MIT-BIH SVDB Database	Accuracy 99.9% Sensitivity 99.8%% Specificity 99.9%	Fine-tuned with AL; Morphological vector and PEF as input
Andersen et al. [84]	Atrial fibrillation detection and monitoring	LSTM, CNN	MIT-BIH Arrhythmia Database; MIT-BIH AF Database; MIT-BIH NSR Database	Accuracy 97.80% Sensitivity 98.98% Specificity 96.95%	Postprocessing
Yao et al. [82]	Arrhythmia classification	LSTM, CNN Attention module	1st China Physiological Signal Challenge	PPV 82.6%, Recall 80.1%, accuracy 81.2%	Halved parameter amount; for varied length segment
Sadaj [92]	Inter- and intra-patient heartbeat classification	LSTM-based auto-encoder; CNN	MIT-BIH Arrhythmia Database	Accuracy:99.53% for inter-patient, 99.92% for intra-patient	Test in inter- and intra-2 paradigms
Saadatnejad et al. [88]	Arrhythmia classification	RNN	MIT-BIH Arrhythmia Database	Accuracy 99.3% for VEB; Accuracy 98.6% for SVEB	Blend model
Qiao et al. [86]	Arrhythmia classification	DELM-LRF-BLSTM	MIT-BIH Arrhythmia Database	Accuracy 99.32% Sensitivity 97.15%	–
Sun et al. [90]	Atrial fibrillation prediction	LSTM	Long-term AF Database; AF terminal challenge Database	Accuracy 92% 92% F-score	Avoid the gradient explosion
Ping et al. [93]	Atrial fibrillation	LSTM, CNN	Cardiology Challenge 2017 Dataset	sped up by 38% F1 score 89.55%	Shortcut added between models
Lih et al. [83]	CAD, Myocardial infarction, congestive heart failure	LSTM, CNN	St.-Petersburg INCART 12-lead arrhythmia, PTB Database, BIDMC CHF Databases, Fantasia Databases	Accuracy 98.51% Sensitivity 97.89% Specificity 99.3% Positive predict 97.3%	–

should be attached more importance. Studies focusing on different aspects were implemented in different databases with different number of leads and durations, making it hard to compare among researches fairly. To be precise, MIT-BIH database, the most commonly used dataset, was collected over 40 years ago. Recently built single-lead PhysioNet database is of high quality, but only focus on short-term ECG. Therefore, the emergence of high-quality long-term dataset will certainly stimulate relevant studies on ECG.

Many symptoms and signs of cardiovascular disease are shown in other physiological data besides ECG, so other vital information, such as respiratory rate and blood pressure, can be used in combination with ECG in diagnosis. For example, Xu et al. [95] predicts patient moralities in ICU with ECG, real-time vital signs and interventions. However, different physiological information varies largely in sampling frequency, which can be considered as multimodal data. Hence, how to utilize multimodal data simultaneously to improve researches' results can serves as a potential study opportunity.

6.2. Limitation and key issues of deep leaning methods

In spite of the rapid development, there are still limitation and open issues for deep learning methods. We will discuss this from three aspect: architecture and visualization, generalization and regularization, robustness and stability and optimization.

The characteristics of architecture includes the number, size and type of the layers [31]. By choosing a proper architecture, good expressiveness can be obtained. Poor interpretability is a key issue of architecture. To start with, there is few systematic ways to help adjust architecture in experiments, making it requires a number of tests to find an efficient model. Moreover, ramifications of incorrect result can be serious in some area such as autonomous driving and medical aspects, so it is almost intolerant for uncertainty. Therefore, it is important to interpret how deep learning works. This can be realized by mathematic justification and visualization. Mathematic theory behind can help to justify the properties of deep learning [96]. Visualizing a network mainly has two ways: visualizing strongest feature map activations and training images to maximize certain neuron's activation. Techniques are introduced to extract more details in images feature maps, throwing light on how network utilize knowledge [97].

The generalization property of deep learning includes the generalization ability for different inputs and different tasks. Since no basic principles can guide the design of these learning systems, most of networks are only designed and fine-tuned for an specific application with high accuracy [27] or only tested on un-standard dataset [28], which limit their reproducibility and generality. Generalization can also express in form of overfitting. Overfitting is caused by over detailed study that model learns from the training data, which would negatively affect the generalization performance of the model. The mainstream method dealing with this problem includes early-stopping, adding weight penalties (L1 and L2 regularization) and dropout [98]. Many studies recently are focused on explaining the underpinnings of dropout [99] and developing dropout that can eliminate the number of parameters to optimized without diminishing its effect [100].

When the training set presents problems such as significant complex noise, abnormal point intrusion and imbalanced categories, deep learning's effectiveness is often not guaranteed [101]. This is the robustness problem of deep learning. Robustness problem can affect the results of the network in practical application because of scale variation, light change and some difference between training data and practice [102]. Salehinejad and Valaee [100] proposed that the model with better accuracy often has a worse the robustness and that the structure of the model has a greater impact on robustness compared with the size of the model. Hence, it is a key problem that how to balance accuracy and robustness of network and how to design a structure with good robustness [103]. Instability problem in training will cause exploding or vanishing gradients, preventing network from further training. Techniques that can help develop stable architecture for deep network are another key issue in deep learning field [104,105]. It is proved that some approaches efficient in improving stability may influence robustness negatively [105]. Therefore, how to design a robust network via stable training is also a key issue [106]. Universal adversarial perturbation is a general way to outline the potential breaches of network security, so it is also a key field to develop well-generalized adversarial perturbation [107].

Optimization concerns the way of training the network, with a general problem of non-convexity and a challenge of efficient convergence. It is possible for optimization algorithms to stuck in non-global minima, preventing network achieving a better result. To solve this problem, a few studies are carried out, focusing on finding the condition where any local optimum is global optimum [108,109] and the property of local minima in deep learning [110]. The challenge of efficient convergence is caused by internal covariate shift and normalization is a common way to limit it. Researches on normalization aim at accelerating and stabilizing the learning process, offering a better convergence [111].

6.3. Potential architectures illuminating the future of ECG diagnosis

In addition to architectures have been applied in ECG diagnosis field, there are other architectures that can help to open the perspective.

Generative Adversarial Nets (GAN) was proposed for generative models' estimation by using adversarial process [112] with its corresponding evaluation metric, the Inception score [113]. It trains one generator model (G) that captures the data distribution and one discriminator model (D) that determine the probability of data from dataset instead of G simultaneously. G is trained to GAN can derive more realistic and more suitable generated samples for training purposes through continuous self-discrimination, achieving great success image processing. GAN can help in ECG denoising tasks. In fact, Singh et al. [113] have shown that a single GAN model can be effective in several noise conditions. GAN is also used widely in data argumentation. Hence, it can also be considered as a possible way to deal with the problem of imbalance data by applying data argumentation to the minor classes.

Fussy deep neural network (FDNN) is the combination of fussy learning and deep neural network. Neural and fussy system can be combined in sequential ways by applying fussy logic to input or output, or in joint framework by fusing fussy and neural features [114]. Fussy learning can help to overcome the network's sensitivity to uncertainty in raw data, which is one of the major drawbacks of deep learning. Hence, FDNN has an advantage of good noise resistance, improving performance on vague and noisy data. FDNN has been tried in healthcare application and time-series prediction tasks [115], but few study on ECG has been carried out. Due to the physiological signals' nature characteristics of heterogeneous, FDNN may help to enlighten the problem caused by interpatient difference in ECG diagnosis.

Probabilistic Neural Networks (PNN) model was proposed based on Bayesian decision theory [116]. PNN is shown to be easy-trained, fast convergence and robust to adversarial examples, which makes it suitable for real time processing. Additionally, it can easily be carried out in hardware since the number of neurons of each PNN layer remains constant. These characteristics are useful when designing wearable ECG systems. Probabilistic Process Neural Networks (PPNN), one extended model of PNN with fewer parameters, is tested to be effective in 12-lead ECG signals [117]. Besides, PNN was also developed to assess the prediction uncertainty which is not exhibited in most deep learning methods [118]. Uncertainty assessment can verify the network's reliability in some work that requires strict correctness.

7. Conclusion

In this paper, existing deep learning studies on ECG diagnosis classification tasks were reviewed and summarized systematically. We reviewed relevant studies from perspective of data, basic algorithm and models. State-of-the-art studies were reviewed according to the deep learning algorithm they used and their distinguishing features were highlighted with discussion. We found that deep learning has achieved good performance in various cardiovascular diseases, though there is still limitation and challenges. Our contribution can be expected in two aspects. Firstly, deep learning researches on ECG diagnosis classification are given a comprehensive and systematical review. Furthermore, the problems and limitations in current researches and future potential opportunities are highlighted. This can serve as a reference for further studies.

CRediT authorship contribution statement

Xinwen Liu: Literature review, Figures and tables, Writing - original draft, Writing - review & editing. **Huan Wang:** Writing - review & editing. **Zongjin Li:** Basic knowledge summary. **Lang Qin:** Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Innovation Fund of Glasgow College, University of Electronic Science and Technology of China and in part by the Sichuan Science and Technology Program under Grant 2020108.

References

[1] E.J. Benjamin, P. Muntner, A. Alonso, M.S. Bittencourt, C.W. Callaway, A.P. Carson, et al., Heart disease and stroke statistics-2019 update: A report from the American heart association, Circulation 139 (10) (2019) e56–e528.

[2] Y. Yu, D.I. Chasman, J.E. Buring, L. Rose, P.M. Ridker, Cardiovascular risks associated with incident and prevalent periodontal disease, J. Clin. Periodontol. 42 (1) (2015) 21–28.

[3] Pishoy Gouda, Paul Brown, McAlister, A. Finlay, et al., Insights into the importance of the electrocardiogram in patients with acute heart failure, Eur. J. Heart Fail. (2016).
[4] M. Bernard Lown, M. Michael D. Klein, M. Philip I. Hershberg, Coronarv and precoronarv care, Am. J. Phys. Med. 46 (5) (1969) 705–724.
[5] S.S. Lobodzinski, A.A. Jadalla, Integrated heart failure telemonitoring system for homecare, Cardiol. J. 17 (2) (2010) 200–204.
[6] C. Costa, J.L. Oliveira, Telecardiology through ubiquitous internet services, Int. J. Med. Inform. 81 (9) (2012) 612–621.
[7] G. de Lannoy, D. Francois, J. Delbeke, M. Verleysen, Weighted conditional random fields for supervised interpatient heartbeat classification, IEEE Trans. Biomed. Eng. 59 (1) (2012) 241–247.
[8] A. Dupre, S. Vincent, P.A. Iaizzo, Basic ECG theory, recordings, and interpretation, in: P.a. Iaizzo (Ed.), Handbook of Cardiac Anatomy, Physiology, and Devices, 2005, pp. 199–201.
[9] P. Hubert, A. R Pipberg, S. F Arms, Automatic screening of normal and abnormal electrocardiograms by means of a digital electronic computer, Proc. Soc. Exp. Biol. Med. (106) (1961) 2–130.
[10] F. Murat, O. Yildirim, M. Talo, U.B. Baloglu, Y. Demir, U.R. Acharya, Application of deep learning techniques for heartbeats detection using ECG signals-analysis and review, Comput. Biol. Med. 120 (2020) 103726.
[11] I. Goodfellow, Y. Bengio, A. Courville, Deep learning, 2016.
[12] F.H.S.D. Silva, Deep learning for corpus callosum segmentation in brain magnetic resonance images, 2018.
[13] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.
[14] Y. Wang, C. Ding, Z. Li, G. Yuan, S. Liao, X. Ma, et al., Towards ultra-high performance and energy efficiency of deep learning systems: An algorithm-hardware co-optimization framework, 2018.
[15] F. Murat, Z. Yldrm, M. Talo, U.B. Baloglu, U.R. Acharya, Application of deep learning techniques for heartbeats detection using ECG signals- analysis and review, Comput. Biol. Med. 120 (2020).
[16] S. Hong, Y. Zhou, J. Shang, C. Xiao, J. Sun, Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review, Comput. Biol. Med. (2020) 103801.
[17] O. Faust, Y. Hagiwara, T.J. Hong, O.S. Lih, U.R. Acharya, Deep learning for healthcare applications based on physiological signals: A review, Comput. Methods Programs Biomed. (2018) S352133162.
[18] Satija Udit, Ramkumar Barathram, Manikandain M. Sabarimalai, A review of signal processing techniques for electrocardiogram signal quality assessment, IEEE Rev. Biomed. Eng. 11 (2018) 36–52.
[19] I. Tomasic, R. Trobec, Electrocardiographic systems with reduced numbers of leads synthesis of the 12-lead ECG, IEEE Rev. Biomed. Eng. (7) (2014) 126–142.
[20] U.R. Acharya, Y.H.S.L. Oh, M.A.J.H. Tan, A. Gertych, R.S. Tan, A deep convolutional neural network model to classify heartbeats, Comput. Biol. Med. (89) (2017) 859–869.
[21] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, et al., Physiobank, physiotoolkit, and physionet : Components of a new research resource for complex physiologic signals, Circulation 101 (23) (2000) e215–e220.
[22] F. Rosenblatt, The Perceptron - A Perceiving and Recognizing Automaton, Cornell Aeronautical Laboratory Report, Vol. 85, 1957.
[23] M. Minsky, S. Papert, Perceptrons, vol. 84, 1969.
[24] D. Rumelhart, G. Hinton, R. Williams, Learning representations by back propagating errors, Nature 323 (1986) 533–536.
[25] G.E. Hinton, R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks, Science 313 (July) (2006) 504–507.
[26] A. Krizhevsky, I. Sutskever, G. Hinton, Imagenet classification with deep convolutional neural networks, in: Proc. NIPS, 2012, pp. 84–90.
[27] Y. Guo, Y. Liu, T. Georgiou, M.S. Lew, A review of semantic segmentation using deep neural networks, Int. J. Multimed. Inf. Retr. (2017).
[28] A. Garcia-Garcia, S. Orts, S. Oprea, V. Villena Martinez, P. Martinez-Gonzalez, J. Rodríguez, A survey on deep learning techniques for image and video semantic segmentation, Appl. Soft Comput. 70 (2018).
[29] Z. Zhao, P. Zheng, S. Xu, X. Wu, Object detection with deep learning: A review, 2018.
[30] T. Minematsu, A. Shimada, H. Uchiyama, R. Taniguchi, Analytics of deep neural network-based background subtraction, J. Imaging 4 (2018) 78.
[31] T. Bouwmans, S. Javed, M. Sultana, S. Jung, Deep neural network concepts for background subtraction: A systematic review and comparative evaluation, 2018.
[32] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P.A. Manzagol, Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion, J. Mach. Learn. Res. 11 (12) (2010) 3371–3408.
[33] A. Makhzani, B. Frey, K-sparse autoencoders, Comput. Sci. (2013).
[34] V. Mnih, H. Larochelle, G.E. Hinton, Conditional restricted Boltzmann machines for structured output prediction, 2012.
[35] S. Elfwing, E. Uchibe, K. Doya, Expected energy-based restricted Boltzmann machine for classification, Neural Netw. 64 (2015) 29–38.
[36] A. Fischer, C. Igel, An introduction to restricted Boltzmann machines, 2012.
[37] H. Larochelle, M. Mandel, R. Pascanu, Y. Bengio, D. Lee, Learning algorithms for the classification restricted Boltzmann machine, J. Mach. Learn. Res. 13 (1) (2012) 643–669.
[38] L.O. Chua, CNN : A paradigm for complexity, 1998.
[39] Y. Lecun, Y. Bengio, Convolutional networks for images, speech, and time series, 1998.
[40] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, Comput. Sci. (2014).
[41] F. Iandola, M. Moskewicz, S. Karayev, R. Girshick, K. Keutzer, Densenet: Implementing efficient convnet descriptor pyramids, 2014, Eprint Arxiv.
[42] K. He, X. Zhang, S. Ren, J. Sun, Deep Residual Learning for Image Recognition, IEEE, 2016.
[43] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar, P. Muller, Deep learning for time series classification: A review, Data Min. Knowl. Discov. 33 (4) (2019) 917–963.
[44] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780.
[45] S. M., K.P. K., Bidirectional recurrent neural networks, IEEE Trans. Signal Process. 45 (11) (1997) 2673–2681.
[46] K. Cho, B.V. Merrienboer, C. Gulcehre, D. Ba Hdanau, F. Bougares, H. Schwenk, et al., Learning phrase representations using rnn encoder-decoder for statistical machine translation, Comput. Sci. (2014).
[47] M.M.A. Rahhal, Y. Bazi, H. AlHichri, N. Alajlan, F. Melgani, R.R. Yager, Deep learning approach for active classification of electrocardiogram signals, Inform. Sci. 345 (2016) 340–354.
[48] K. Luo, J. Li, Z. Wang, A. Cuschieri, Patient-specific deep architectural model for ECG classification, J. Healthc. Eng. 2017 (2017) 1–13.
[49] J. Yang, Y. Bai, F. Lin, M. Liu, Z. Hou, X. Liu, A novel electrocardiogram arrhythmia classification method based on stacked sparse auto-encoders and softmax regression, Int. J. Mach. Learn. Cybern. 9 (10) (2018) 1733–1740.
[50] K. Hanbay, Deep neural network based approach for ECG classification using hybrid differential features and active learning, IET Signal Process. 13 (2018).
[51] J. Farhadi, G. Attarodi, J. Dabanloo, M. Mohandespoor, M. Eslamizadeh, Classification of atrial fibrillation using stacked auto encoders neural networks, in: Proc. Classification of Atrial Fibrillation Using Stacked Auto Encoders Neural Networks, 2018.
[52] S.L. Oh, E.Y.K. Ng, R.S. Tan, U.R. Acharya, Automated beat-wise arrhythmia diagnosis using modified U-net on extended electrocardiographic recordings with heterogeneous arrhythmia types, Comput. Biol. Med. (2018).
[53] B. Hou, J. Yang, P. Wang, R. Yan, LSTM based auto-encoder model for ECG arrhythmias classification, IEEE Trans. Instrum. Meas. (2019) 1.
[54] S. Nurmaini, A. Darmawahyuni, A. Noviar, S. Mukti, M. Rachmatullah, F. Firdaus, et al., Deep learning-based stacked denoising and autoencoder for ECG heartbeat classification, Electronics 9 (2020).
[55] S.L. Oh, E.Y.K. Ng, R.S. Tan, U.R. Acharya, Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats, Comput. Biol. Med. 102 (2018) 278–287.
[56] G. Altan, Y. Kutlu, N. Allahverdi, A multistage deep belief networks application on arrhythmia classification, Int. J. Intell. Syst. Appl. Eng. 6 (2016) 222–228.
[57] S. G, K.P. T, K.K. V, Classification of ECG beats using deep belief network and active learning, Med. Biol. Eng. Comput. 56 (10) (2018) 1887–1898.
[58] G. Altan, N. Allahverdi, Y. Kutlu, Diagnosis of coronary artery disease using deep belief networks, Eur. J. Eng. Nat. Sci. 1 (2) (2017) 29–36.
[59] S.M. Mathews, C. Kambhamettu, K.E. Barner, A novel application of deep learning for single-lead ECG classification, Comput. Biol. Med. 99 (2018) 53–62.
[60] B. Taji, A.D.C. Chan, S. Shirmohammadi, False alarm reduction in atrial fibrillation detection using deep belief networks, IEEE Trans. Instrum. Meas. 67 (5) (2018) 1124–1131.
[61] L. Song, D. Sun, Q. Wang, Y. Wang, Automatic classification method of arrhythmia based on discriminative deep belief networks, J. Biomed. Eng. 36 (3) (2019) 444–452.
[62] S. Kiranyaz, T. Ince, M. Gabbouj, Real-time patient-specific ECG classification by 1-D convolutional neural networks, IEEE Trans. Biomed. Eng. 63 (3) (2016) 664–675.
[63] Y. Li, Y. Pang, J. Wang, X. Li, Patient-specific ECG classification by deeper CNN from generic to dedicated, Neurocomputing 314 (2018) 336–346.
[64] P. Rajpurkar, A.Y. Hannun, M. Haghpanahi, C. Bourn, A.Y. Ng, Cardiologist-level arrhythmia detection with convolutional neural networks, 2017, pp. 1707–1836.
[65] X. Zhai, C. Tin, Automated ECG classification using dual heartbeat coupling based on convolutional neural network, IEEE Access 6 (2018) 27465–27472.

[66] J. Huang, B. Chen, B. Yao, W. He, ECG arrhythmia classification using stft-based spectrogram and convolutional neural network, IEEE Access PP (99) (2019) 1.
[67] X. X., L. H., ECG heartbeat classification using convolutional neural networks, IEEE Access 8 (2020) 8614–8619.
[68] W. Liu, Q. Huang, S. Chang, H. Wang, J. He, Multiple-feature-branch convolutional neural network for myocardial infarction diagnosis using electrocardiogram, Biomed. Signal Process. Control 45 (2018) 22–32.
[69] J. Niu, Y Tang, Z Sang, W. Zhang, Inter-patient ECG classification with symbolic representations and multi-perspective convolutional neural networks, IEEE J. Biomed. Health Inform. 24 (5) (2020) 1321–1332.
[70] U.R. Acharya, H. Fujita, S.L. Lo, U. Raghavendra, J.H. Tan, M. Adam, et al., Automated identification of shockable and non-shockable life-threatening ventricular arrhythmias using convolutional neural network, Future Gener. Comput. Syst. (2017) S167739X–S17315248X.
[71] U.R. Acharya, H. Fujita, S.L. Lo, Y. Hagiwara, J.H. Tan, M. Adam, Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals, Inform. Sci. 415–416 (2017) 190–198.
[72] A.U. R., O.S. L., H. Y., A deep convolutional neural network model to classify heartbeat, Comput. Biol. Med. (89) (2017).
[73] U.R. Acharya, H. Fujita, O.S. Lih, M. Adam, J.H. Tan, C.K. Chua, Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network, Knowl.-Based Syst. 132 (sep.15) (2017) 62–71.
[74] U.R. Acharya, H. Fujita, O.S. Lih, Y. Hagiwara, J.H. Tan, M. Adam, Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network, Inform. Sci. 405 (2017) 81–90.
[75] Ö. Yıldırım, P. Pławiak, R. Tan, U.R. Acharya, Arrhythmia detection using deep convolutional neural network with long duration ECG signals, Comput. Biol. Med. 102 (2018) 411–420.
[76] J. Jiang, H. Zhang, D. Pi, C. Dai, A novel multi-module neural network system for imbalanced heartbeats classification, Expert Syst. Appl. X (2019).
[77] U.B. Baloglu, M. Talo, O. Yildirim, R.S. Tan, U.R. Acharya, Classification of myocardial infarction with multi-lead ECG signals and deep CNN, Pattern Recognit. Lett. 122 (2019) 23–30.
[78] J. Wang, A deep learning approach for atrial fibrillation signals classification based on convolutional and modified elman neural network, Future Gener. Comput. Syst. 102 (2019).
[79] L. D., B. Y, S. Y, Z. X, M. C., Non-standardized patch-based ECG lead together with deep learning based algorithm for automatic screening of atrial fibrillation, IEEE J. Biomed. Health Inform. (2020) 1.
[80] T.F. Romdhane, H. Alhichri, R. Ouni, M. Atri, Electrocardiogram heartbeat classification based on a deep convolutional neural network and focal loss, Comput. Biol. Med. 123 (2020) 103866.
[81] J.H. Tan, Y. Hagiwara, W. Pang, I. Lim, S.L. Oh, M. Adam, et al., Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals, Comput. Biol. Med. 94 (2018) 19–26.
[82] Q. Yao, R. Wang, X. Fan, J. Liu, Y. Li, Multi-class arrhythmia detection from 12-lead varied-length ECG using attention-based time-incremental convolutional neural network, Inf. Fusion 53 (2019).
[83] O.S. Lih, V. Jahmunah, T.R. San, E.J. Ciaccio, T. Yamakawa, M. Tanabe, et al., Comprehensive electrocardiographic diagnosis based on deep learning, Artif. Intell. Med. 103 (2020) 101789.
[84] R.S. Andersen, A. Peimankar, S. Puthusserypady, A deep learning approach for real-time detection of atrial fibrillation, Expert Syst. Appl. 115 (2019) 465–473.
[85] L. Guo, G. Sim, B. Matuszewski, Inter-patient ECG classification with convolutional and recurrent neural networks, Biocybern. Biomed. Eng. 39 (2019).
[86] F. Qiao, B. Li, Y. Zhang, H. Guo, W. Li, S. Zhou, A fast and accurate recognition of ECG signals based on ELM-LRF and BLSTM algorithm, IEEE Access PP (2020) 1.
[87] G. Wang, C. Zhang, Y. Liu, H. Yang, D. Fu, H. Wang, et al., A global and updatable ECG beat classification system based on recurrent neural networks and active learning, Inform. Sci. 501 (2018).
[88] S. Saadatnejad, M. Oveisi, M. Hashemi, LSTM-based ECG classification for continuous monitoring on personal wearable devices, IEEE J. Biomed. Health Inform. 24 (2) (2020) 515–523.
[89] O. Faust, A. Shenfield, M. Kareem, T.R. San, H. Fujita, U.R. Acharya, Automated detection of atrial fibrillation using long short-term memory network with RR interval signals, Comput. Biol. Med. 102 (2018) 327–335.
[90] L. Sun, Y. Wang, J. He, H. Li, D. Peng, Y. Wang, A stacked LSTM for atrial fibrillation prediction based on multivariate ECGs, Health Inf. Sci. Syst. 8 (1) (2020) 19.
[91] Ö. Yildirim, A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification, Comput. Biol. Med. 96 (2018) 189–202.
[92] M. Sadaj, Inter- And Intra- Patient Ecg Heartbeat Classification for Arrhythmia Detection: A Sequence to Sequence Deep Learning Approach, in: 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019.
[93] Y. Ping, C. Chen, L. Wu, Y. Wang, M. Shu, Automatic detection of atrial fibrillation based on CNN-LSTM and shortcut connection, Healthcare 8 (2020) 139.
[94] S. Han, H. Mao, W.J. Dally, Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding, in: Proc. ICLR, 2016.
[95] Y. Xu, S. Biswal, S. Deshpande, K. Maher, J. Sun, Raim: Recurrent attentive and intensive model of multimodal patient monitoring data, 2018, pp. 2565–2573.
[96] R. Vidal, J. Bruna, R. Giryes, S. Soatto, Mathematics of deep learning, 2017.
[97] F. Wang, H. Liu, J. Cheng, Visualizing deep neural network by alternately image blurring and deblurring, Neural Netw. 97 (2017) 162–172.
[98] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: A simple way to prevent neural networks from overfitting, J. Mach. Learn. Res. 15 (1) (2014) 1929–1958.
[99] P. Mianjy, R. Arora, R. Vidal, On the implicit bias of dropout, 2018.
[100] H. Salehinejad, S. Valaee, Ising-dropout: A regularization method for training and compression of deep neural networks, 2019.
[101] B. Sengupta, K. Friston, How robust are deep neural networks? 2018.
[102] D. Su, H. Zhang, H. Chen, J. Yi, P.Y. Chen, Y. Gao, Is robustness the cost of accuracy? – A comprehensive study on the robustness of 18 deep image classification models, 2018.
[103] I. Diakonikolas, D. Kane, A. Stewart, Robust learning of fixed-structure Bayesian networks, 2016.
[104] B. Chang, L. Meng, E. Haber, L. Ruthotto, D. Begert, E. Holtham, Reversible architectures for arbitrarily deep residual neural networks, 2017.
[105] Eldad Haber, Lars Ruthotto, Stable architectures for deep neural networks, Inverse Problems (2018).
[106] S. Zheng, Y. Song, T. Leung, I. Goodfellow, Improving the robustness of deep neural networks via stability training, in: Proc. Computer Vision & Pattern Recognition, 2016.
[107] S.M. Moosavi-Dezfooli, A. Fawzi, O. Fawzi, P. Frossard, Universal adversarial perturbations, in: Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2017.
[108] B.D. Haeffele, R. Vidal, Global optimality in neural network training, in: Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2017.
[109] M. Nouiehed, M. Razaviyayn, Learning deep models: Critical points and local openness, 2018.
[110] C. Yun, S. Sra, A. Jadbabaie, A critical view of global optimality in deep learning, 2018.
[111] S. Malladi, I. Sharapov, Fastnorm: Improving numerical stability of deep network training with efficient normalization, in: Proc. ICLR 2018 Conference, 2018.
[112] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, et al., Generative adversarial networks, Adv. Neural Inf. Process. Syst. 3 (2014) 2672–2680.
[113] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, X. Chen, Improved techniques for training gans, 2016.
[114] Y. Deng, Z. Ren, Y. Kong, F. Bao, Q. Dai, A hierarchical fused fuzzy deep neural network for data classification, IEEE Trans. Fuzzy Syst. (2017).
[115] R. Das, S. Sen, U. Maulik, A survey on fuzzy deep neural networks, ACM Comput. Surv. 53 (3) (2020) 1–25.
[116] Specht, Probabilistic neural networks for classification, mapping, or associative memory, in: Proc. IEEE International Conference on Neural Networks, 1988.
[117] F. N., X. S, L. Y, L. K., A probabilistic process neural network and its application in ECG classification, IEEE Access 7 (2019) 50431–50439.
[118] J. Gast, S. Roth, Lightweight probabilistic deep networks, in: Proc. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018.