IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 20, 2020, accepted April 7, 2020, date of publication April 14, 2020, date of current version April 29, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2987930

A Fast and Accurate Recognition of ECG Signals Based on ELM-LRF and BLSTM Algorithm

FENGJUAN QIAO[1], BIN LI[1,2,3], YOUMEI ZHANG[1], HONGLI GUO[1], WEI LI[1], AND SHUWANG ZHOU[2,3]

[1]School of Mathematics and Statistics, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250353, China
[2]Shandong Computer Science Center (National Supercomputer Center in Jinan), Jinan 250041, China
[3]Shandong Artificial Intelligence Institute, Jinan 250041, China

Corresponding author: Bin Li (ribbenlee@126.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61973185, in part by the Development Plan of Young Innovation Team in Colleges and Universities of Shandong Province under Grant 2019KJN011, in part by the Shandong Province Key Research and Development Program under Grant 2018GGX103054, and in part by the Young Doctor Cooperation Foundation of Qilu University of Technology (Shandong Academy of Sciences) under Grant 2018BSHZ2008.

ABSTRACT Extreme learning machine based on local receptive fields (ELM-LRFs) is a very fast method that can be used for feature extraction and classification. Bidirectional long-short time memory network (BLSTM), a widely used type of recurrent neural network (RNN) architecture, has showed excellent performance in time series processing fields. In this paper, we combine the superiority of above algorithms and propose a fast and accurate hybrid deep learning model which is named DELM-LRF-BLSTM for ECG signal recognition. This model uses the segmented heartbeats as input and employs a deep ELM-LRF to gain significant local spatial features. Then we fed the features to a three-layer BLSTM and it can extract temporal features for ECG signal recognition. The combination of ELM-LRF and BLSTM can not only consider the local information in a heartbeat, but also consider the long-distance dependence between heartbeats. Experimental results on MIT-BIH Arrhythmia dataset show that the proposed DELM-LRF-BLSTM algorithm has high accuracy and sensitivity, up to 99.32% and 97.15% respectively, which verifies the effectiveness and feasibility of the model. Moreover, only 6.1 millisecond is needed for once heartbeat recognition operation. Due to it's high performance and low computational complexity, the proposed algorithm is feasible for practical use.

INDEX TERMS ECG recognition, local receptive fields, extreme learning machine, bidirectional long-short term memory, deep learning.

I. INTRODUCTION

Arrhythmia has been a common phenomenon characterized by cardiac conduction disorder, which may increase the risk of sudden death, and it can generally be manifested by electrocardiogram (ECG) [1]. The ECG collects electrical signals from external electrodes connected to the skin and can accurately capture the cardiac electrical activity over a period of time [2], [3]. The ECG has become an indispensable tool for clinicians to diagnose cardiovascular diseases and analyze pathology. Therefore, researches on ECG signals has received extensive attention [4]–[6], both in the computer field and biomedical field.

The associate editor coordinating the review of this manuscript and approving it for publication was Khalid Aamir.

Studies on ECG signals are generally divided into two parts: detection and classification. Studies on detection concentrate on QRS complex detection and determining heartbeats within the ECG data obtained for a certain period of time. The heartbeats detection methods mainly contain Kalman Filter [7], wavelet transform [8], digital filter-based [9] and threshold-based methods [10], etc. The classification of detected heartbeats is also an important step in the treatment of ECG signals. In the last decades, the rapid development of machine learning makes recognizing arrhythmia automatically from the ECG signals possible. The applications of deep learning methods have got satisfactory results in the analysis of biological signals and medical intelligent systems for auxiliary diagnosis. Acharya *et al.* [11] implemented a 9-layer convolutional neural network (CNN) on the

MIT-BIH Arrhythmia dataset and obtained 94.03% accuracy and 96.71% sensitivity. Yildirim *et al.* [12] used a 16-layer CNN model on the same dataset with performance of 95.2% accuracy and 93.52% sensitivity.

Moreover, long-short term memory network (LSTM) [13], [14] is an excellent variant model of recurrent neural network (RNN) and has showed superior performance in the process of time series. Due to its efficient sequence learning ability, many deformations of this architecture have been applied to ECG recognition tasks. Chauhan *et al.* [15] used LSTM networks to classify the normal and abnormal heartbeat with an accuracy of 96.45%. In order to further improve the classification performance on ECG signals, Shu *et al.* [16] combined LSTM with CNN to develop a new hybrid structure for processing ECG signals with variable length and achieved promising results with high recognition rate of 98.10%, sensitivity of 97.50%. Recently, Yildirim *et al.* [17] proposed a new approach for arrhythmia diagnosis based on bidirectional LSTM networks and convolutional auto-encoder network (CAE) and got 99.23% accuracy.

The above algorithms have achieved encouraging results. However, CNN and CNN-LSTM algorithms [18] have a common disadvantage: time consuming. These two networks need long time to adjust weights iteratively by using the stochastic gradient descent algorithm. To overcome this problem, Huang *et al.* [19] proposed a fast and non-finetuing ELM-LRF model. This model realizes the local connection between the input layer and the hidden layer by introducing the local receptive fields (LRFs) into ELM [20], which has the advantages of fast training speed, low computational complexity and high generalization ability. We utilize a deep ELM-LRF model to extract the features of ECG signals from deep levels. This deep ELM-LRF model is called DELM-LRF.

Single DELM-LRF model only considers the local relation in space and ignores the information in time domain. In view of this, with the help of LSTM's excellent sequence learning ability in the time domain, this paper combines the advantages of DELM-LRF and bidirectional LSTM (BLSTM) [21], and proposes a hybrid model, termed as DELM-LRF-BLSTM. The DELM-LRF-BLSTM model mainly contains two typical stages: feature extraction and sequence learning. In the first stage, we design a DELM-LRF model to extract spatial information of the ECG signal and its output is the input of next stage. In the second stage, BLSTM network is utilized to learn the temporal information and classify the ECG signal into five categories. Therefore,the proposed DELM-LRF-BLSTM algorithm is able to extract significant features both in space and time domain. We examined its performance on the well-known MIT-BIH Arrhythmia dataset. The results show that it not only improves the accuracy of classification, but also greatly reduces the training time.

The contributions of this study can be summarized as follows:

1) We propose a novel algorithm, called DELM-LRF-BLSTM, for ECG signals recognition by combining the advantages of DELM-LRF and BLSTM neural networks.
2) The DELM-LRF-BLSTM algorithm is fully automatic without noise filtering or manual feature extraction operations.
3) We examine the performance of the DELM-LRF-BLSTM algorithm on the MIT-BIH dataset. The results show that the proposed algorithm has high recognition accuracy and low computational complexity and can be used in practical applications.

The rest of this paper is arranged as follows. The structure of the proposed algorithm is described in Section 2. Experimental results are presented in Section 3. Discussions and comparison results with the-state-of-art methods are given in Section 4. Finally, Section 5 makes a conclusion for the proposed method.

II. METHODS

A. ELM-LRF AND DELM-LRF

In this part, the ELM-LRF neural networks are described concisely. More details could be found in [19]. ELM-LRF model is generally divided into two parts: feature extraction and classification.

In the feature extraction part, convolution operation and pooling operation are implemented. In the convolution layer, each element of the convolution kernels is a pseudorandom number depended on a probability distribution [22]. The convolution operation is performed using matrix multiplication for fast feature extraction. The number (m) and size ($w \times 1$) of the filters are artificially given according to experience, each filter is converted into a sparse Toeplitz matrix represented as W_1. Thus, the convolution operation is defined as:

$$C = W_1 X \tag{1}$$

where $X \in \mathbb{R}^{N\times J}$ denotes the input matrix with N samples, and each sample contains J sample points. $C \in \mathbb{R}^{N\times(J-w+1)}$ represents the feature map obtained after convolution operation.

In the pooling layer, Huang *et al.* [19] used square-root pooling operation in the ELM-LRF structure, which can be represented as:

$$P = \left(W_2 C^2\right)^{\frac{1}{2}} \tag{2}$$

where W_2 denotes the weights matrix in pooling layer, $P \in \mathbb{R}^{N\times(J-w+1)}$ represents the output of square-root pooling layer.

We use max pooling operation [23], [24] in this paper instead of square-root pooling to reduce the dimension of feature maps. This operation is described as:

$$P = \max\left(W_2 C\right) \tag{3}$$

The dimension of matrix P is determined by the scale d, described as $N \times ((J + w - 1)/d)$.

The classification stage can be regarded as a process of finding the closed-form solution-based output weights W_{out} by using least square algorithm [25]. The output weights W_{out} can be obtained according to (4) and (5).

1) if $N \leq (J + w - 1)/d$:

$$W_{out} = P^T \left(\frac{I}{\lambda} + PP^T \right)^{-1} Y \tag{4}$$

2) if $N \geq (J + w - 1)/d$:

$$W_{out} = \left(\frac{I}{\lambda} + P^T P \right)^{-1} P^T Y \tag{5}$$

where I denotes the identity matrix, λ is the penalty coefficient. Once the weights matrix W_{out} is determined, the model can be used for new data with the same distribution.

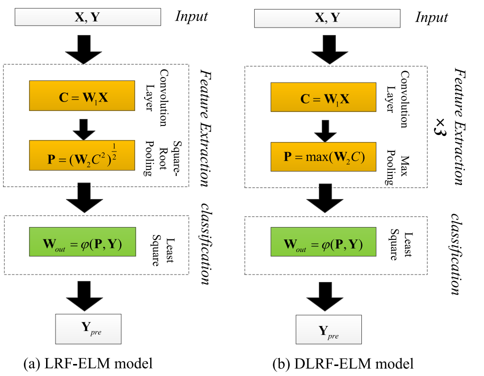

FIGURE 1. **(a) The ELM-LRF model described in [19] using a convolution layer and square-root pooling layer for feature extraction. (b) The proposed DELM-LRF using three alternately replaced convolution layer and max pooling layer for feature extraction.**

Next, we introduce a deep ELM-LRF (DELM-LRF) structure with three feature extraction stages to get spatial information of the ECG data in deep level. In the DELM-LRF structure, three feature extraction stages are cascaded with different convolution filters and pooling scales. Each feature extraction stage consists of a convolution layer and a max pooling layer. For simplicity, we only use random filters in the convolution layers. That means each point of convolution filter is randomly determined and no fine-tuning required. The output of the previous stage is the input of the next stage. By three-stage feature extraction operation, discriminative features were selected and fed into a least square layer for classification. In the whole training phase, only output weights matrix W_{out} is trained, weights in feature extraction stages such as W_1 and W_2 are randomly initialized without fine tuning. The comparisons of ELM- LRF with DELM-LRF is shown in Fig. 1. φ represents the least squares solution functions. X, Y denote the input sequences and target labels, respectively.

B. SEQUENCE LEARNING

1) LSTM

As a special recurrent neural network, LSTM has been widely used in different time series learning tasks by virtue of its

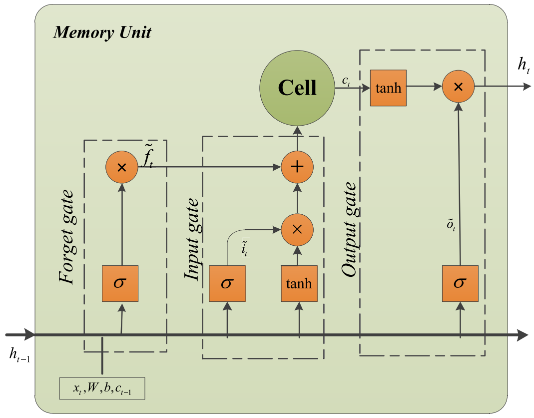

FIGURE 2. **Diagram of a basic LSTM memory unit.**

unique abilities in long-term dependency problems [26]. The most important component of a LSTM network are a set of memory units which consist of three gate structures and one cell state, as shown in Fig. 2.

The input gate controls which parts of the new information are saved to the cell state, and forget gate determines the reserved information of the historical cell state. The output gate controls which parts of the updated cell state are exported. The specific working process of a basic memory unit is expressed by the following formulas:

$$\begin{pmatrix} \tilde{f}_t \\ \tilde{i}_t \\ \tilde{o}_t \end{pmatrix} = \sigma \left(W \left[x_t, h_{t-1} \right] + b \right) \tag{6}$$

$$c_t = \tilde{f}_t \odot c_{t-1} + \tilde{i}_t \odot \tanh \left(W \left[x_t, h_{t-1} \right] + b \right) \tag{7}$$

$$h_t = \tilde{o}_t \odot \tanh \left(c_t \right) \tag{8}$$

where W and b are the weights matrix and bias vectors, respectively. $\tilde{f}_t$, $\tilde{i}_t$, $\tilde{o}_t$, c_t represent the outputs of forget gate, input gate, output gate and cell state at t time, respectively. x_t and h_t are input vector and hidden layer vector. σ denotes the logistic sigmoid function, and $\odot$ is the Hadamard product.

C. BLSTM

From the above knowledge, it can be obviously found that the memory cells of LSTM can only make use of previous information rather than subsequent information. The way to solve this problem is BLSTM [27], [28]. The BLSTM contains two inputs and two hidden layers and can be simply regarded as two independent LSTM structures. These different two inputs make use of the same sequence in opposite directions: forward and backward. The hidden layers connect the two inputs separately and their outputs were combined at each step. The outputs of second hidden layer at last step are connected to a dense layer and generate output information by a softmax function, which can be specifically expressed as (9).

$$P(y_{(i)} = j | x_{(i)}) = \frac{\exp(x^T w_j)}{\sum_{n=1}^{N} \exp(x^T w_n)} \tag{9}$$

where j is the categories of heartbeats, $x_{(i)}$ represents the i th sample of input, $P(y_{(i)} = j|x_{(i)})$ is the probability function to get the class label of the i th sample. N represents the total number of samples, w represent all the parameters used in our model.

Based on previous studies [29], it can be easily found that BLSTM outperforms LSTM. The prediction results of ECG signals are determined on a combination of previous and subsequent information. Therefore, we use BLSTM to get high recognition accuracy.

D. DELM-LRF-BLSTM

DELM-LRF can extract ECG signal features automatically with high efficiency. BLSTM network has also showed excellent generalization performance in problems related to time series. Meanwhile, ECG signals are continuous redundant time series with complex noises, which leads to great difficulties in feature extraction tasks. Therefore, we combine the above two algorithms and propose a new structure, called DELM-LRF-BLSTM, to process the ECG signals recognition problems. This algorithm consists of two stages: feature extraction and sequence learning. The diagram of DELM-LRF-BLSTM network is shown in Fig. 3.

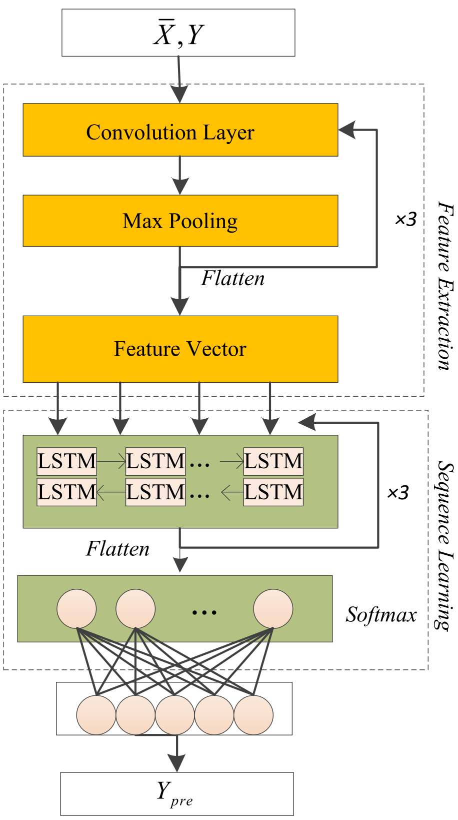

FIGURE 3. The diagram of DELM-LRF-BLSTM network. $\bar{X}$, Y, Y_{pre} denote the input heartbeats, corresponding true labels and predicted labels, respectively.

In the feature extraction stage, convolution and max pooling layers are alternately stacked to extract spatial features. This stage contains one input layer, three convolution layers with kernel size 4×1, three max pooling layers with pooling size 2×1 and one flatten layer. Among them, three convolution layers with kernel size 4×1 are used for learning high-level features. The weights of convolution kernel are randomly determined according to probability distribution and no fine tuning required. We also use three max pooling layers with pooling size 2×1 for dimension reduction and remaining the translation invariance of the feature. The output of last pooling layer are flattened and fed into the next stage.

In the sequence learning stage, we design a deep BLSTM network with three BLSTM layers. The input data of the first BLSTM layer are the output of the feature extraction stage and its output are fed into the next BLSTM layer. It is worth to note that a flatten layer and fully-connected layer are set behind three BLSTM layers. Similarly expressed in formula (9), softmax is selected as the activation function of fully-connected layer. The values of softmax function are the output of entire model, representing the probability of each category. The category labels are finally determined according to the probabilities.

III. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we verify the performance of the proposed DELM-LRF-BLSTM algorithm on ECG signals recognition problems.

All the simulations are carried out on a computer with a 3.70 GHz Intel(R) Core(TM) i7-8700K processor and a 32GB RAM. The MATLAB R2018a environment and Anaconda 3.5 are used for data preprocessing and algorithm implementation, respectively.

TABLE 1. Details of five beat types used in the experiments.

Beat Types	Number of Training set (80%)	Number of Testing set (20%)	Total
NSR	59970	14992	74962
LBBB	6454	1614	8068
RBBB	5803	1451	7254
VPC	5627	1407	7034
APB	2036	509	2545
Total	79890	19973	99863

A. DATA DESCRIPTION

In this paper, the well-known MIT-BIH arrhythmia [30] dataset is used for evaluating the performance of different algorithms. This dataset contains 48 records of ECG signals obtained from inpatients of Beth Israel Hospital (BIH) and each ECG signal contains more than 30 minutes and has been annotated by cardiovascular specialist. All of these ECG signals are segmented into 250 samples for training. The detailed description of five beat types used in this experiment is shown in Table 1. In this simulation, we totally use 99863 ECG signals including 74962 normal sinus rhythm (NSR), 8068 left bundle branch block(LBBB), 7254 right

TABLE 2. The formulas of several popular evaluation criteria.

	Formulas
Acc.	$\frac{TP+TN}{TP+FP+FN+TN}$
Pre.	$\frac{TP}{TP+FP}$
Rec.	$\frac{TP}{TP+FN}$
F1.	$2 \times \frac{Rec \times Pre}{Rec+Pre}$

bundle branch block(RBBB), 7034 ventricular premature contraction (VPC) and 2545 atrial premature beat (APB).

B. PRE-PROCESSING

The input data of the proposed method is a sequence of heartbeats. In order to obtain the heartbeats, we do the following steps:

Step 1: Segmenting the longer ECG data into a series of heartbeats based on QRS detection technique. Pan-Tompkins algorithm [31] is used in this work.

Step 2: Downsampling each heartbeat to a certain length containing 250 sample points.

Step 3: Normalizing the heartbeat data using (10).

$$\bar{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{10}$$

where x is the ECG data, x_{max} and x_{min} represent the maximum and minimum values of each sample attribute of ECG signal data.

C. EVALUATION CRITERIA

In order to evaluate the effectiveness of different methods, several popular evaluation criteria are calculated: accuracy (abbrev. Acc.), precision (abbrev. Pre.), sensitivity (abbrev. Sen.), and F1-score (abbrev. F1.). The computational formulas associated with these common metrics are illustrated in Table 2, where TP, TN, FP, FN and N stand for true positive, true negative, false positive, false negative and the number of whole samples, respectively.

D. RESULTS ON DELM-LRF

To compare the performance of different algorithms, we first obtain the experiment results on DELM-LRF network. The input are a sequence of normalized heartbeats and the output are corresponding labels. Table 3 shows the detailed information of each layer in this model.

In the feature extraction stage, the convolution kernel size is set to 20 × 1, 16 × 1, 5 × 1, respectively. Each point of convolution kernel is randomly determined according to certain probability distribution without manual fine-tuning. The pooling size is 3 × 1, 2 × 1, 3 × 1, respectively. Kernel size and pooling size are set using trial-and-error.

After the feature extraction stage, feature vectors with size 72 × 1 are obtained. Equation (4) and (5) are then used to get the output weights. The penalty coefficient λ is selected in array {1, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120,

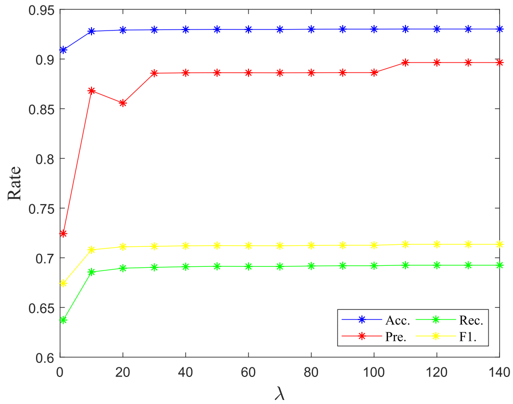

FIGURE 4. The curves of various evaluation criterion values varied with penalty coefficient.

130, 140, 150}. Figure 4 shows the changing curves of various average values of evaluation criterion varied with penalty coefficient. It is conspicuous that values of almost all evaluation criterion increase at an early stage as the coefficient values increase. The evaluation criterion values remain stable when the coefficient values are greater than 110. Thus, the threshold serves as the value of λ in the simulation results.

The performance of the DELM-LRF network is shown in Table 4. It is obvious that high performance is obtained in several classes except APB. The average precision, recall, and F1-score are 69.25%, 89.64%, and 71.35%, respectively. The APB class data achieve unsatisfactory results because of data deficiency. There are only 2545 heartbeats obtained from the dataset. The DELM-LRF network achieves an average classification accuracy of 93.18% on the training set and 93.02% on the testing set. It can be seen that although the classification accuracy is not superior to CNN [32]–[34], it costs a very little time in the training phase with 62.81s.

E. RESULTS ON DELM-LRF-BLSTM

It can be seen from the above experiment that the DELM-LRF is fast but the classification results are not satisfactory especially in the category of APB. In order to further improve the performance, a DELM-LRF-BLSTM is proposed for ECG recognition. The data illustrated in Table 1 are used for training and testing the DELM-LRF-BLSTM. The required parameters of this model are tuned using trial-and-error. The batch size is set to 20, the learning rate is 0.001 and dropout rate is 0.1. Adam optimizer is selected as cross entropy function. All results are obtained after 400000 iterations to ensure that the loss function reaches convergence.

In the first stage, input signals with size 250 per beat are fed into the proposed network. Feature are obtained by three alternately placed convolution layers with kernel size 17 × 1, 6 × 1, 5 × 1, respectively, and max pooling layers with pooling size 2 × 1. Features from different levels are efficiently extracted due to the strong spatial features learning ability of convolution operation. All features of each beat are flattened into a feature vector with 78 elements. Fig. 5 shows

TABLE 3. Detailed information of the DELM-LRF.

Layers	Layer Name	Kernel Size	Number of Filters	Output Shapes
0	Input	-	-	250×1
1	Convolution Layer	20×1	4	231×4
2	MaxPooling	3	4	77×4
3	Convolution Layer	16×1	8	62×8
4	MaxPooling	2	8	31×8
5	Convolution Layer	5×1	8	48×3
6	MaxPooling	3	8	24×3
7	Flatten	-	-	72
8	Output	-	-	5

TABLE 4. Classification performance of the DELM-LRF network on the testing set.

	Pre.	Rec.	F1.	Average Acc.
NSR	99.34%	93.51%	96.34%	93.02%
LBBB	86.49%	86.33%	86.41%	
RBBB	87.80%	94.86%	91.19%	
VPC	71.86%	93.52%	81.27%	
APB	0.78%	80.00%	1.54%	
Average	69.25%	89.64%	71.35%	

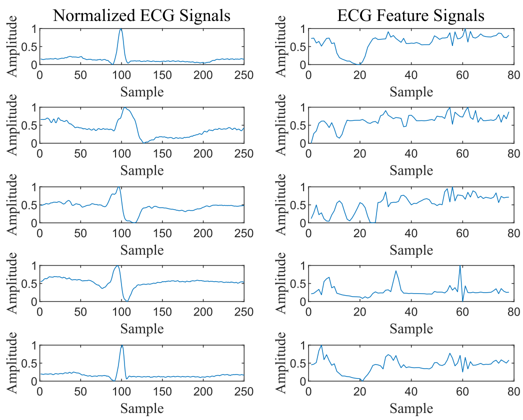

FIGURE 5. Waveforms of raw ECG signals (left) and their corresponding feature signals (right).

several waveforms of raw normalized ECG signals and their corresponding feature signals.

Since the sample points of the feature signals are greatly reduced compared with raw ECG signals, the training time of following sequence learning stage is going to decrease. The number of BLSTM units is equal to the dimension of input data which is 78. Three such BLSTM layers with the same number of BLSTM units are designed in our algorithm. The details of DELM-LRF-BLSTM structure used in this experiment is illustrated in Table 5.

The recognition results are obtained after 400000 training iterations. As we know, overfitting problem is a major problem in training phase of neural networks. We use dropout in this work since it is an effective measure to tackle the problem of overfitting in the process of training. Dropout means that some units are discarded at random with some certain rates to get rid of the redundant information. However, if the drop probability is too large, important information may be lost. For this purpose, a dropout of 0.1 is used in each

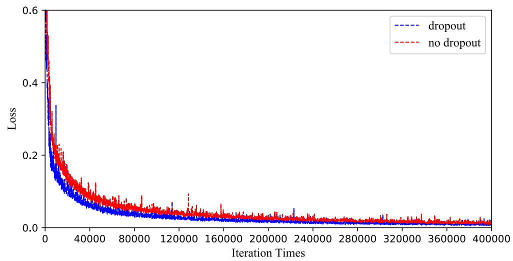

FIGURE 6. The performance curve of DELM-LRF-BLSTM with dropout and DELM-LRF-BLSTM without dropout in training phase.

BLSTM layer of the network in this paper. Figure 6 shows the comparison curves during training process for both DELM-LRF-BLSTM with and without dropout. We can observe that the DELM-LRF-BLSTM with dropout is able to converge after fewer iterations.

The evaluation criteria of DELM-LRF-BLSTM are summarized in Table 6. As observed from the table, the proposed algorithm achieved satisfactory results with 99.32% average accuracy, 98.80% average precision, 97.15% average recall, and 97.71% average F1-score. The performance increases in all categories based on the proposed algorithm, especially in APB. The precision, recall and F1-score increase from 0.78%, 80.00%, and 1.54% to 93.91%, 88.70%, and 91.23%, respectively.

F. COMPUTATIONAL COMPLEXITY

In this part, we estimate the time complexity of DELM-LRF-BLSTM theoretically. At first, we calculate the time complexity of feature extraction stage. It occurs due to the convolution operation. The complexity of each convolution layer is estimated as $O(M \cdot K \cdot i \cdot o)$, where M, K, i, o are the spatial size of the feature maps, the spatial size of the filters, the input channel and the output channel, respectively. The overall complexity of feature extraction stage is $O\left(\sum_{l=1}^{d} M_l \cdot K_l \cdot i_l \cdot o_l\right)$, where d is the number of convolutional layers, c_l is the spatial size of the filter in lth convolutional layer, i_l is the input channel of lth layer, o_l is the output channel of lth layer.

Next, we calculate the time complexity of sequence learning stage. For each time step, the time complexity of LSTM is $O(w)$, w is the number of tunable parameters [13]. BLSTM is considered as two different LSTM networks and the time

TABLE 5. The details Of DELM-LRF-BLSTM structure used in this experiment.

Layers	Layer Name	Kernel Size	Number of Filters	Tunable Parameters	Other Parameters
0	input	-	-	-	-
1	Convolution Layer	17×1	4	-	-
2	MaxPooling	2	4	-	-
3	Convolution Layer	6×1	8	-	-
4	MaxPooling	2	8	-	-
5	Convolution Layer	5×1	3	-	-
6	MaxPooling	2	3	-	-
7	Flatten	-	-	-	-
8	BLSTM	-	-	78 Unit	Dropout = 0.1
9	BLSTM	-	-	78 Unit	Dropout = 0.1
10	BLSTM	-	-	78 Unit	Dropout = 0.1
11	Flatten	-	-	-	-
12	Fully-connected	-	-	5 Unit	Activation= softmax

TABLE 6. The overall performances Of DELM-LRF-BLSTM.

	Pre.	Rec.	F1.	Average Acc.
NSR	99.47%	99.75%	99.61%	99.32%
LBBB	99.69%	99.57%	99.63%	
RBBB	99.72%	99.10%	99.41%	
VPC	98.72%	98.65%	98.68%	
APB	93.91%	88.70%	91.23%	
Average	98.30%	97.15%	97.71%	

TABLE 7. Computational Complexity comparison Of LSTM, BLSTM and DELM-LRF-BLSTM.

	Tunable Parameters	Complexity	Training Time
BLSTM	206076	≈ 8.15	$\approx$ 5.3h/epoch
LSTM	103038	≈ 4.08	$\approx$ 2.7h/epoch
DELM-LRF-BLSTM	25272	1	$\approx$ 0.7h/epoch

complexity per time step is equal to $O(W)$, W is the number of parameters used in BLSTM network.

The DELM-LRF-BLSTM complexity per time step can be calculated as the sum of the complexity of the feature extraction stage and the sequence learning stage: $O\left(\sum_{l=1}^{d}(M_l \cdot K_l \cdot i_l \cdot o_l) + W\right)$. Since there is no iteration used for parameter optimization in the feature extraction stage, the complexity for all the training process can be calculated as $O\left(\sum_{l=1}^{d}(M_l \cdot K_l \cdot i_l \cdot o_l) + W \cdot e\right)$, e is the number of epochs in sequence learning stage. Therefore, we can conclude that our algorithm has O complexity in the typical asymptotic notation.

Since there is no iteration required in feature extraction stage, it has little impact on the overall complexity. We ignore it when calculating the complexity of DELM-LRF-BLSTM. The complexity of DELM-LRF-BLSTM is estimated as $O(W)$ in both training and testing process, W is the number of parameters used in sequence learning stage. Table 7 compares the computational complexity of LSTM, BLSTM and DELM-LRF-BLSTM in our experiment. All three algorithms use the same number of hidden units and hidden layers. The parameters of BLSTM including batch size, learning rate and iteration times are set the same as the DELM-LRF-BLSTM. All experiments are conducted on a computer device with I7-8700K CPU at 3.70 GHz and there is no any acceleration technique such as parallel computation by GPUs to be used for faster speed.

From Table 7, we observe that the number of tunable parameters of DELM-LRF-BLSTM is much less than LSTM and BLSTM. Therefore, the DELM-LRF-BLSTM has the lowest computational complexity. The complexity of LSTM is about 4.08 times that of DELM-LRF-BLSTM. The complexity of BLSTM is about 8.15 times that of DELM-LRF-BLSTM.

Meanwhile, the DELM-LRF-BLSTM model takes the least training time. In our experiment, about 0.7 h is needed for one training epoch. The LSTM and BLSTM needs about 2.7 h and 5.3 h, respectively. Due to the large number of samples, the training time of DELM-LRF-BLSTM is still very long. However, once the model is trained well, it can be used for any groups without retraining steps. Therefore, the training time for DELM-LRF-BLSTM can be neglected in practice.

For practical use, we count the testing time of once operation, which begins from inputting single heartbeat to DELM-LRF-BLSTM and ending at outputting the recognition result. In our experiment, only 6.1 ms is needed for recognizing single heartbeat with a length of 250. Considering the pre-processing time, 2 s is enough. Therefore, the DELM-LRF-BLSTM takes very little time for heartbeat identification. It is real-time and feasible for practical use.

IV. DISCUSSION

In order to prove the superior performance of our proposed algorithm, Table 8 gives the comparison of the proposed algorithm with other state-of-the-art models in terms of recognition of ECG signals. To ensure the fairness of comparison, all results of comparison algorithms are obtained on the MIT-BIH Arrhythmia dataset and all algorithms are used to identify five types of heartbeats.

As shown in Table 8, Elhaj *et al.* [33] combined linear feature extractor (PCA+DWT) and nonlinear feature extractor (HOS+ICA). They obtained 98.91% recognition accuracy by using a SVM classifier. Kiranyaz *et al.* [34], Zubair *et al.* [35], and Acharya *et al.* [11] used CNN

TABLE 8. Comparison of the proposed algorithm with other state-of-the-art algorithms.

	Feature Extraction	Classification	Accuracy
Elhaj et al., 2016 [33]	PCA+DWT+HOS+ICA	SVM	98.91%
Kiranyaz et al., 2016 [34]	-	1D-CNN	99.00%
Zubair et al., 2016 [35]	-	CNN	92.70%
Acharya et al., 2017 [11]	-	9 layers CNN	94.03%
Yang et al., 2020 [36]	ICA	PCANet	98.63%
Yang et al., 2018 [37]	PCANet	SVM	97.08%
Li et al., 2016 [38]	WPE+RR	RF	94.61%
Martis et al., 2013 [39]	DWT+ICA	PNN	99.28%
-	-	BLSTM	97.26%
The proposed	**-**	**DELM-LRF**	**93.02%**
	-	**DELM-LRF-BLSTM**	**99.32%**

with different structures to classify ECG signals, achieving 99.00%, 92.70%, and 94.03% accuracy, respectively. Yang *et al.* [36] obtained 98.63% accuracy by using ICA and PCANet. Yang *et al.* [37] employed PCANet and SVM on the MIT-BIH Arrhythmia dataset and achieved 97.08% accuracy. Li *et al.* [38] combined WPE, RR and RF to identify ECG signals. They reached a recognition accuracy of 94.61%. Martis *et al.* [39] used DWT and ICA to extract features and PNN to classify them. The recognition accuracy obtained 99.28%. We use BLSTM network for identifying five categories of heartbeats and achieve 97.26% accuracy. In our study, the accuracy of the proposed DELM-LRF-BLSTM reaches 99.32%. It outperforms plain DELM-LRF and BLSTM, and it is better than most of the popular methods.

Feature extraction is an important step in deep neural networks. Discriminant features are able to improve the performance of networks. This work proposed an end-to-end learning structure without any hand-crafted feature extraction or additional process. This structure is able to automatically extract discriminant features from raw heartbeats and has showed remarkable success in the ECG recognition.

We examine the performance of the proposed structure on the MIT-BIH Arrhythmia dataset. The DELM-LRF can deeply extract the spatial features that affect the heartbeat category, and the BLSTM is appropriate for modeling temporal information of irregular trends in time series. The proposed DELM-LRF-BLSTM method achieves higher recognition accuracy than most of state-of-the-art algorithms. Therefore, the combination of DELM-LRF and BLSTM is valuable.

In addition, no parameters used in feature extraction stage is required training. The DELM-LRF can quickly extract features and reduce the input size of BLSTM stage. Compared with separate BLSTM, the combined model has lower computational complexity and higher recognition accuracy.

This paper presents a novel DELM-LRF-BLSTM algorithm. The algorithm can not only achieve higher recognition accuracy but also has faster recognition speed. The highlights of the proposed algorithm are summarized as follows:

1) The proposed algorithm is fully automated.
2) Noise filtering and manual features extraction are not required.
3) Discriminant features are extracted both in time and space domain.
4) High recognition accuracy is achieved.
5) The recognition speed is fast.

The potential limitations of our proposed algorithm are as follows:

1) The algorithm needs to detect the R peaks of the ECG signals.
2) It is assumed that each heartbeat has only one type of arrhythmia.
3) The proposed algorithm is not robust in detecting the APB from normal heartbeats.

V. CONCLUSION

In this paper, a new approach DELM-LRF-BLSTM is proposed by combining ELM-LRF and BLSTM. The spatial information of the ECG signals is first extracted by three alternately replaced convolution layers and three max pooling layers and then fed into the BLSTM network. This work achieved a good recognition performance on MIT-BIH Arrhythmia dataset with 99.32% accuracy and 97.15% sensitivity, which was superior to many other state-of-the-art methods. In addition, due to the low computational complexity and high speed for ECG recognition, it promotes the application of automatic diagnosis for arrhythmia in practice. In future work, we will further improve the recognition accuracy of ECG samples, including dimension reduction method, denoising or other techniques. Our work is expected to contribute to detecting arrhythmia timely and effectively and saving a large number of lives.

REFERENCES

[1] S. Min, B. Lee, and S. Yoon, "Deep learning in bioinformatics," *Briefings Bioinf.*, vol. 18, no. 5, pp. 851–869, Mar. 2017.

[2] A. Esmaili, M. Kachuee, and M. Shabany, "Nonlinear cuffless blood pressure estimation of healthy subjects using pulse transit time and arrival time," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 12, pp. 3299–3308, Dec. 2017.

[3] A. E. Dastjerdi, M. Kachuee, and M. Shabany, "Non-invasive blood pressure estimation using phonocardiogram," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Baltimore, MD, USA, May 2017, pp. 1–4.

[4] M. Kachuee, S. Fazeli, and M. Sarrafzadeh, "ECG heartbeat classification: A deep transferable representation," in *Proc. IEEE Int. Conf. Healthcare Informat. (ICHI)*, New York, NY, USA, Jun. 2018, pp. 443–444.

[5] V. Seena and J. Yomas, "A review on feature extraction and denoising of ECG signal using wavelet transform," in *Proc. 2nd Int. Conf. Devices, Circuits Syst. (ICDCS)*, Combiatore, India, Mar. 2014, pp. 1–6.

[6] C. Ye, M. T. Coimbra, and B. V. K. V. Kumar, "Arrhythmia detection and classification using morphological and dynamic features of ECG signals," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol.*, Buenos Aires, Argentina, Aug. 2010, pp. 1918–1921.

[7] Z. Zhang, Q. Yu, Q. Zhang, N. Ning, and J. Li, "A Kalman filtering based adaptive threshold algorithm for QRS complex detection," *Biomed. Signal Process. Control*, vol. 58, Apr. 2020, Art. no. 101827.

[8] C. Saritha, V. Sukanya, and Y. N. Murthy, "ECG signal analysis using wavelet transforms," *Bulg. J. Phys.*, vol. 35, no. 1, pp. 68–77, Feb. 2008.

[9] E. Pasolli and F. Melgani, "Active learning methods for electrocardiographic signal classification," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 6, pp. 1405–1416, Nov. 2010.

[10] V. S. Chouhan and S. S. Mehta, "Threshold-based detection of P and T-wave in ECG using new feature signal," *Int. J. Comput. Sci. Netw. Secur.*, vol. 8, no. 2, pp. 144–153, 2008.

[11] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, A. Gertych, and R. S. Tan, "A deep convolutional neural network model to classify heartbeats," *Comput. Biol. Med.*, vol. 89, pp. 389–396, Oct. 2017.

[12] Ö. Yıldırım, P. Pławiak, R.-S. Tan, and U. R. Acharya, "Arrhythmia detection using deep convolutional neural network with long duration ECG signals," *Comput. Biol. Med.*, vol. 102, pp. 411–420, Nov. 2018.

[13] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[14] Q. Lyu and J. Zhu, "Revisit long short-term memory: An optimization perspective," in *Proc. NIPS*, Montreal, QC, Canada, 2015, pp. 1–9.

[15] S. Chauhan and L. Vig, "Anomaly detection in ECG time signals via deep long short-term memory networks," in *Proc. IEEE Int. Conf. Data Sci. Adv. Analytics (DSAA)*, Paris, France, Oct. 2015, pp. 1–7.

[16] S. L. Oh, E. Y. K. Ng, R. S. Tan, and U. R. Acharya, "Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats," *Comput. Biol. Med.*, vol. 102, no. 1, pp. 278–287, Nov. 2018.

[17] O. Yildirim, U. B. Baloglu, R.-S. Tan, E. J. Ciaccio, and U. R. Acharya, "A new approach for arrhythmia classification using deep coded features and LSTM networks," *Comput. Methods Programs Biomed.*, vol. 176, pp. 121–133, Jul. 2019.

[18] P. Warrick and M. Nabhan Homsi, "Cardiac arrhythmia detection from ECG combining convolutional and long short-term memory networks," in *Proc. Comput. Cardiology Conf. (CinC)*, Rennes, France, Sep. 2017, pp. 161–460.

[19] G.-B. Huang, Z. Bai, L. L. C. Kasun, and C. M. Vong, "Local receptive fields based extreme learning machine," *IEEE Comput. Intell. Mag.*, vol. 10, no. 2, pp. 18–29, May 2015.

[20] G. Huang, G.-B. Huang, S. Song, and K. You, "Trends in extreme learning machines: A review," *Neural Netw.*, vol. 61, pp. 32–48, Jan. 2015.

[21] J. P. C. Chiu and E. Nichols, "Named entity recognition with bidirectional LSTM-CNNs," *Trans. Assoc. Comput. Linguistics*, vol. 4, no. 1, pp. 357–370, Dec. 2016.

[22] K. Jarrett, K. Kavukcuoglu, M. A. Ranzato, and Y. LeCun, "What is the best multi-stage architecture for object recognition?" in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Kyoto, Japan, Sep. 2009, pp. 2146–2153.

[23] M. M. Dos Santos, A. G. da Silva Filho, and W. P. Dos Santos, "Deep convolutional extreme learning machines: Filters combination and error model validation," *Neurocomputing*, vol. 329, pp. 359–369, Feb. 2019.

[24] D. Scherer, A. Meller, and S. Behnke, "Evaluation of Pooling Operations in Convolutional Architectures for Object Recognition," in *Proc. 20th ICANN*. Berlin, Germany: Springer, 2010, pp. 92–101.

[25] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: A new learning scheme of feedforward neural networks," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Budapest, Hungary, Jul. 2004, pp. 985–990.

[26] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling," in *Proc. INTERSPEECH*, Singapore, Sep. 2014, pp. 1–5.

[27] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," *Neural Netw.*, vol. 18, nos. 5–6, pp. 602–610, Jul. 2005.

[28] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM networks," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Warsaw, Poland, Jul. 2005, pp. 2047–2052.

[29] Ö. Yildirim, "A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification," *Comput. Biol. Med.*, vol. 96, pp. 189–202, May 2018.

[30] R. Mark and G. Moody. (May 1997). *MIT-BIH Arrhythmia Database*. [Online]. Available: http://ecg.mit.edu/dbinfo.html

[31] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vols. BME–32, no. 3, pp. 230–236, Mar. 1985.

[32] Z. Golrizkhatami and A. Acan, "ECG classification using three-level fusion of different feature descriptors," *Expert Syst. Appl.*, vol. 114, pp. 54–64, Dec. 2018.

[33] F. A. Elhaj, N. Salim, A. R. Harris, T. T. Swee, and T. Ahmed, "Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals," *Comput. Methods Programs Biomed.*, vol. 127, pp. 52–63, Apr. 2016.

[34] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 664–675, Mar. 2016.

[35] M. Zubair, J. Kim, and C. Yoon, "An automated ECG beat classification system using convolutional neural networks," in *Proc. 6th Int. Conf. IT Converg. Secur. (ICITCS)*, Prague, Czech Republic, Sep. 2016, pp. 1–6.

[36] W. Yang, Y. Si, D. Wang, and G. Zhang, "A novel method for identifying electrocardiograms using an independent component analysis and principal component analysis network," *Measurement*, vol. 152, Feb. 2020, Art. no. 107363.

[37] W. Yang, Y. Si, D. Wang, and B. Guo, "Automatic recognition of arrhythmia based on principal component analysis network and linear support vector machine," *Comput. Biol. Med.*, vol. 101, no. 1, pp. 22–32, Oct. 2018.

[38] T. Li and M. Zhou, "ECG classification using wavelet packet entropy and random forests," *Entropy*, vol. 18, no. 8, p. 285, Aug. 2016.

[39] R. J. Martis, U. R. Acharya, and L. C. Min, "ECG beat classification using PCA, LDA, ICA and discrete wavelet transform," *Biomed. Signal Process. Control*, vol. 8, no. 5, pp. 437–448, Sep. 2013.

FENGJUAN QIAO received the master's degree in mathematics from the Qilu University of Technology (Shandong Academy of Sciences), Jinan, China.

Her research has been concerned with scalable methods for intelligent diagnosis of arrhythmia, electrocardiogram classification with deep learning. She is interested in solving complex real-world problems arising from medical domain by use of various machine learning methods and deep learning methods.

BIN LI received the bachelor's, master's, and Ph.D. degrees in control science, operational research and cybernetics, and pattern recognition and intelligent system from Shandong University, China, in 2002, 2005, and 2012, respectively.

He is currently an Associate Professor with the School of Mathematics and Statistics, Qilu University of Technology (Shandong Academy of Sciences), China. His research interests include algorithms for neural networks, gait planning, and adaptive control of legged robots.

YOUMEI ZHANG received the Ph.D. degree in pattern recognition and intelligent system from Shandong University, China, in 2019.

She is currently working with the School of Mathematics and Statistics, Qilu University of Technology (Shandong Academy of Sciences), as an Instructor, China. Her research interests include crowd counting, facial expression recognition, and abnormal event detection.

HONGLI GUO received the master's degree in mathematics from the Qilu University of Technology (Shandong Academy of Sciences), Jinan, China.

Her research has been concerned with pattern recognition, image classification, and optimization of algorithms.

WEI LI received the master's degree in mathematics from the Qilu University of Technology (Shandong Academy of Sciences), Jinan, China.

His research involves the scalable method of intelligent diagnosis of arrhythmia, ECG classification, and deep learning. His research direction is to solve the problem of ECG signal classification by using deep learning methods and various integration strategies.

SHUWANG ZHOU received the master's degree from Shandong Normal University, in 2011.

He is currently an Assistant Researcher with the Shandong Computer Science Center (National Supercomputer Center in Jinan). His research interests include the Internet of Things and artificial intelligence.