Research Article

An Automatic System for Atrial Fibrillation by Using a CNN-LSTM Model

Fengying Ma¹, Jingyao Zhang¹, Wei Chen^{2,3}, Wei Liang¹, and Wenjia Yang^{2,3}

¹School of Electrical Engineering and Automation, Qilu University of Technology (Shandong Academy of Sciences), Jinan, China

²School of Mechanical Electronic & Information Engineering, China University of Mining and Technology-Beijing, Beijing 100083, China

³School of Computer Science & Technology, China University of Mining and Technology, Xuzhou 221116, China

Correspondence should be addressed to Wei Chen; chenwdavior@163.com and Wei Liang; dzhlw0918@qlu.edu.cn

Received 18 June 2020; Revised 24 July 2020; Accepted 27 July 2020; Published 28 August 2020

Guest Editor: Chi-Hua Chen

Copyright © 2020 Fengying Ma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Atrial fibrillation (AF) is a common abnormal heart rhythm disease. Therefore, the development of an AF detection system is of great significance to detect critical illnesses. In this paper, we proposed an automatic recognition method named CNN-LSTM to automatically detect the AF heartbeats based on deep learning. The model combines convolutional neural networks (CNN) to extract local correlation features and uses long short-term memory networks (LSTM) to capture the front-to-back dependencies of electrocardiogram (ECG) sequence data. The CNN-LSTM is fed by processed data to automatically detect AF signals. Our study uses the MIT-BIH Atrial Fibrillation Database to verify the validity of the model. We achieved a high classification accuracy for the heartbeat data of the test set, with an overall classification accuracy rate of 97.21%, sensitivity of 97.34%, and specificity of 97.08%. The experimental results show that our model can robustly detect the onset of AF through ECG signals and achieve stable classification performance, thereby providing a suitable candidate for the automatic classification of AF.

1. Introduction

Heart disease is the leading cause of human death, and the number of deaths due to cardiovascular diseases accounts for a large proportion of the total number of deaths worldwide [1]. Most cardiovascular diseases are often accompanied by arrhythmia. Among them, atrial fibrillation (AF) is the most common persistent arrhythmia. In our country, there are more than 10 million people suffering from AF. Its incidence increases with age, but in recent years, its incidence has shown an increasing trend in people of younger age groups [2–4]. Simultaneously, a series of complications related to AF, such as stroke, heart failure, and other diseases, also lead to high morbidity and mortality [5–7]. However, AF also shows strong unpredictability, and capturing AF signals in real time is difficult [8, 9]. Electrocardiogram (ECG) detection technology forms an important basis for AF diagnosis [10–12]. Therefore, the application of automatic detection technology for

diagnosing AF is necessary. Consequently, machine learning has significantly contributed to the development of real-time monitoring of AF, and timely intervention in the effective detection of AF can avoid serious consequences caused by an exacerbation of the disease [13].

With developments in information technology and artificial intelligence technology, the automatic classification of AF has made great progress. The traditional ECG classification algorithm is composed of two parts: feature extraction and classifier. First, principal component analysis, latent Dirichlet allocation, and other methods are designed for feature extraction and then placed into a support vector machine or random forest, among others, for classification in the classifier. A complete ECG signal is shown in Figure 1 where AF is mainly seen by the disappearance of the *p* wave or irregular RR intervals. The RR interval is the amount of time change between two *R* waves, and the feature recognition method based on the RR interval shows high accuracy. Because the wave value of the *R* wave has the largest

peak in the ECG signal, locating it is easy. However, the low amplitudes of the *P* and *T* waves make it challenging for them to be detected, and the feature extraction algorithm is still not mature enough. With developments in machine learning, traditional models have some insurmountable defects. First, traditional algorithms need to design feature extraction methods to extract useful information and combine machine learning algorithms for classification. This process may be accompanied with the loss of some information. When the extracted features cannot fully reflect the data, the classification results may appear to have larger errors. Second, it mainly relies on considerable prior expert knowledge and sufficient biomedical signal processing capabilities. On this basis, designing a good classifier algorithm is also necessary, but achieving optimal results is difficult.

Unlike traditional machine learning algorithms, deep learning-based methods have the ability to mine complex relations, and useful features of data and have been widely researched and applied in the automatic classification of AF. Xu et al. proposed a framework that combined an improved frequency slice wavelet transform and a convolutional neural network (CNN) for automatic AF beat recognition and achieved good performance [14]. Wei et al. constructed a synchronous feature of each heartbeat of an ECG signal through a recursive complex network and subsequently used CNN to detect AF by analyzing the eigenvalues of the recursive complex network [15]. Andersen et al. proposed an end-to-end model combining CNN and recurrent neural networks to classify ECG signals as AF or a normal sinus rhythm [11]. Pourbabae et al. developed a deep learning machine to screen and identify patients with paroxysmal AF [16]. Dang et al. proposed a model that uses a CNN-BLSTM network to diagnose arrhythmia and uses ECG signals to automatically detect AF, achieving relatively good results [17]. Several researchers have shown that combining the deep learning features with the classifier will significantly improve the performance of the system and make the classification results more ideal. Although the above research can effectively solve the classification problem of AF, we can observe that various neural networks have the ability to extract complex nonlinear characteristics from the original data without human intervention; however, learning the thinking mechanism of the ECG signal features with high accuracy required for monitoring is still a difficult task. CNN and long short-term memory networks (LSTM) are very efficient for feature extraction of ECG signals, and this superiority is applied herein to the AF detection algorithm.

In this study, we proposed a new diagnosis method for AF named CNN-LSTM, which can automatically detect AF from ECG signals. The contributions of this study are as follows:

- (i) We propose an automatic recognition method named CNN-LSTM that uses heartbeat features as input datasets to automatically identify AF in an ECG signal.
- (ii) CNN has advantages in image processing, while LSTM can compensate for the shortcomings of CNN in the context sequence. Therefore, the

combination of CNN and LSTM can effectively improve accuracy in the field of AF recognition.

- (iii) The use of multiscale signals representing the AF characteristics as the input of the network reduces computing resources. The design can be used to extract multiscale features and improve the generalization ability of the network model, and this study provides a high-precision classification method to meet the real-time monitoring needs of AF.

In the following sections, we provide a detailed experimental process and verify the performance of the method in the open access database. This automated method can analyze a large amount of data in a short time while ensuring high accuracy; thus, it may become a practical tool for providing real-time monitoring for patients and reducing the work pressure on doctors.

Section 1 of this article deals with the background of current research on AF and related research algorithms that have been implemented. Section 2 covers the data source and related network structure required for the experiment. Section 3 presents the experimental details, results, and analysis of the results. Section 4 presents the concluding statements and prospects for future work.

2. Material and Methods

2.1. Description of Dataset. Our experimental research is based on data from the MIT-BIH Atrial Fibrillation Database, which is publicly available from PhysioNet [18, 19]. This database includes 25 long-term ECG Holter records from different subjects (mainly paroxysmal attacks). It contains two ECG signal channels with AF annotations. The sampling rate of this database is 250 Hz, and these records also include beat notes manually marked by expert clinicians.

We preprocess the ECG signal to train and evaluate the automatic AF prediction method based on the CNN-LSTM model. As the duration of the AF recording in these data is different from the normal recording duration, all ECG signals are divided into the same duration and data balance is performed to better apply the learning of the model. After segmentation, 960,000 short-term ECG segments were obtained, comprising 480,000 segments of AF records and 480,000 segments of normal records. Figure 2 shows a comparison of the AF signal and the normal signal. The ECG signal segment is divided into a training set and a test set. To better detect the classification effect of the model, the signal processing and segmentation of the dataset are random in nature.

2.2. Networks. This section first reviews the CNN and LSTM network models, which are closely related to the model structure proposed herein. Then, our proposed research model is put forward, and the structure, parameters, and mathematical expressions of the model are described in detail.

FIGURE 1: A complete ECG signal.

FIGURE 2: Comparison of AF signal and N signal.

In this section, we describe the network structure model proposed in this paper, which mainly includes two convolutional layers, one LSTM layers, fully connected layers, and other computing operations.

2.2.1. Convolutional Neural Network. CNN is a feedforward neural network, and it mainly includes an input layer, a convolutional layer, a pooling layer, and an output layer. Its special network structure has great advantages in feature extraction and learning, especially in the field of image recognition, and thus it can achieve great success. Its structure is shown in Figure 3.

The CNN is connected to the input layer through a convolution kernel. The convolution kernel performs dot multiplication through a sliding window to achieve multi-scale feature extraction. Simultaneously, the weight-sharing mechanism of the convolution layer makes it more effective for feature extraction, greatly reducing the number of free variables that need to be learned. Subsequently, we add a pooling layer after the convolutional layer to reduce the feature matrix and network complexity. Because the input

ECG signals are one-dimensional time series, we use one-dimensional convolution in the convolution layer, as shown in Figure 4.

Before the data training, we normalized the data. The convolutional layer extracts features from the original input. The output of the a -th neuron of the one-dimensional convolutional layer is shown in the following equation:

$$O_a = \delta \left(\sum_{j=1}^n W_j X_{a-j+n} + b \right). \quad (1)$$

The input sequence is $X_l (l = 1, 2, \dots, n)$, where W denotes a matrix of weight coefficients, b is an offset coefficient, and n is the number of convolution kernels. Then, the result of the convolution is input into an activation function δ (in this case ReLU), and then the result of the convolution layer is fed back to the pooling layer.

2.2.2. Long Short-Term Memory Network. LSTM is a special recurrent neural network. LSTM is suitable for applications such as natural language processing [20, 21] and biomedical

FIGURE 3: CNN structure.

signal processing [22, 23]. LSTM improves the standard RNN model and adds a gate mechanism. It overcomes the problems of gradient disappearance, gradient explosion, and length dependence of traditional RNNs. The hidden layer of LSTM comprises an input gate, a forget gate, and an output gate. The structure is shown in Figure 5.

The input of the LSTM hidden layer includes not only the input X_t of the current sequence but also the state C_{t-1} of the hidden layer at the previous time; then, the output vector h_{t-1} and the output h_t and C_t of the current state are calculated. The key part of LSTM is what information we will discard from the cell state C_{t-1} at the last moment and how much information can be transferred to the current state C_t . This decision is made through a forget door. The next step is to decide how much new information is added to the next state. Finally, we determine the output value based on the current C_t and h_t . The update method of LSTM is as follows:

$$\begin{aligned} f_t &= \delta(W_f[h_{t-1}, X_t] + b_f), \\ i_t &= \delta(W_i[h_{t-1}, X_t] + b_i), \\ j_t &= \tanh(W_c[h_{t-1}, X_t] + b_c), \\ O_t &= \delta(W_o[h_{t-1}, X_t] + b_o), \\ h_t &= O_t * \tanh(C_t), \end{aligned} \quad (2)$$

where C_t is the state information of the memory unit, j_t is the accumulated information at the current moment, W is the weight coefficient matrix, b is the bias term, σ is the sigmoid activation function, and $\tanh$ is the hyperbolic tangent activation function.

2.2.3. Proposed Architecture. Neural networks have their own unique feature learning method. The CNN model converts the original input into a fixed-length vector representation through convolution kernels, sliding windows, and pooling to capture local features in the input, but the original data arrive. The dependency relation is difficult to learn, and LSTM can better understand the content of the input information through the memory unit that can compensate for the defects of the CNN. Therefore, a CNN-LSTM deep learning model is proposed to achieve the automatic classification of AF in this paper [24, 25].

Figure 6 shows the proposed CNN-LSTM network architecture. After inputting the ECG signal, the convolutional layer and the pooling layer in the CNN first extract local features and subsequently enter the hidden layer of LSTM to obtain optimal feature representation. Finally, the nonlinear function softmax in the fully connected layer is classified into the corresponding categories. CNN adds some processing such as normalization that can avoid overfitting and speed up training.

3. Experimental Results

3.1. Performance Evaluation. To estimate the performance of heartbeat classification, the performance of the model is usually evaluated with accuracy, specificity, and sensitivity [26–28]. They are defined as follows:

$$\begin{aligned} \text{sensitivity} &= \frac{TP}{TP + FN} \times 100\%, \\ \text{specificity} &= \frac{TN}{TN + FP} \times 100\%, \\ \text{accuracy} &= \frac{TN + TP}{TN + FP + FN + TP} \times 100\%, \end{aligned} \quad (3)$$

where TP denotes the number of correctly classified AF signals; FP denotes the number of incorrectly classified AF signals; TN is the number of correctly classified N signals; and FN is the number of misclassified N signals. The results of the experimental study are presented in the next section.

3.2. Implementation Details and Results. The experiment herein uses the TensorFlow neural network framework. Before the experiment started, the data labels were converted into corresponding one-hot vectors. The experiment is based on an equal number of AF records and normal records for training and testing, and the signal processing of the dataset is random. During the experiment, parameter optimization was performed, the batch size was set to 128, the learning rate was 0.01, multiple iteration training was performed, and the Adam updater was used to update the weights to obtain the best classification results. Table 1 lists the relevant parameters of the experimental network.

FIGURE 4: One-dimensional convolutional neural network structure.

FIGURE 5: LSTM hidden layer structure diagram.

Figure 7 shows the loss and accuracy curves of the CNN-LSTM model. It indicates performance change in the training set as the number of iterations increases. The network continues to converge, and the model does not appear to be overfitting.

Figure 8 shows the receiver operating characteristic curve (ROC) curve of the test set. The abscissa of the curve is the false positive rate, and the ordinate is the true positive rate. AUC denotes the area under the ROC curve. The AUC realized by the model was 0.97. The closer the AUC value is to 1, the better is the performance of the model.

Figure 9 shows the confusion matrix of the test set, which is used to measure the accuracy of a classifier. The upper left corner is TP, the upper right corner is FP, the lower left

corner is FN, and the lower right corner is FP. We can convert the result of the quantity in the confusion matrix to a ratio between 0 and 1 to facilitate standardized measurements.

As mentioned above, the CNN-LSTM model achieved an overall classification accuracy of 97.28% on the test set, with a sensitivity of 97.51% and a specificity of 97.06%.

In this study, by combining the deep learning model of CNN and LSTM, i.e., using CNN and LSTM to extract the characteristics of the ECG, the model can automatically extract features and achieve higher accuracy.

Table 2 shows a series of scientific studies based on ECG signals in the MIT-BIH AF database. It mainly includes three evaluation indices: accuracy, sensitivity, and specificity.

FIGURE 6: CNN-LSTM network structure.

TABLE 1: The parameters of the model.

Network layer type	Kernel size	Strides
Input layer	—	—
Con1d	17	1
Pooling1d	5	5
Con2d	7	1
Pooling2d	5	5
LSTM	—	—
Fully connected layer	—	—
Softmax	—	—

FIGURE 7: Loss and accuracy curves.

FIGURE 8: ROC curves.

FIGURE 9: Confusion matrix of test set.

TABLE 2: Comparison with previous work on the MIT-BIH AF database.

Author	Database	Features	Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)
Xu et.al. [15]	AFDB	MESWT	CNN	85.85	79.05	89.99
Wei et.al. [14]	AFDB	RCN	CNN	94.59	94.28	94.91
Andersen et.al. [11]	AFDB	RRI	LSTM + CNN	87.4	98.6	86.4
	MITDB NSRDB					
Dang et.al. [17]	AFDB	RR	CNN-BLSTM	96.59	99.93	97.03
		P-QRS-T	CNN-LSTM	94.07	94.25	92.73
Proposed model	AFDB	Deep features	CNN-LSTM	97.21	97.34	97.08

Through a comparison, we can observe that our proposed CNN-LSTM network model has improved on the input signal of the model and network structure compared with other deep learning methods and achieved good results.

4. Conclusion

In this study, we conducted an in-depth study of the ECG classification algorithm and constructed a network

combining CNN and LSTM. This network can extract the characteristics of ECG signals and classify them. Compared with traditional ECG classification methods, our proposed CNN-LSTM network structure used the MIT-BIH AF database and achieved a high classification accuracy. The experimental results confirm that our proposed CNN-LSTM network is effective for the automatic detection and classification of AF. In addition, this method occupies fewer computing resources and can theoretically achieve real-time performance, thereby contributing to the development of wearable ECG detection devices. Our future research may involve the use of a model that classifies AF tasks under nonfixed scale inputs to achieve further optimization of the neural network.

Data Availability

The data used to support the findings of this study have not been made available because the data also form part of an ongoing study. The original data of the study can be obtained at <https://physionet.org/>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the Shandong University Undergraduate Teaching Reform Research Project (approval number: M2018X078) and the Shandong Province Graduate Education Quality Improvement Program 2018 (approval number: SDYAL18088). This study was also partially supported by the Major Science and Technology Innovation Projects of Shandong Province (grant no. 2019JZZY010731).

References

- [1] H. Wang, N. Mohsen, A. Christine et al., "Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a Systematic Analysis for the Global Burden of Disease Study 2015," *The Lancet*, vol. 388, no. 10053, pp. 1459–1544, 2016.
- [2] C. R. C. Wyndham, "Atrial fibrillation: the most common arrhythmia," *Texas Heart Institute Journal*, vol. 27, no. 3, pp. 257–267, 2000.
- [3] C.-H. Chen, F. Song, F.-J. Hwang, and L. Wu, "A probability density function generator based on neural networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 541, Article ID 123344, (2020).
- [4] V. Markides and R. J. Schilling, "Atrial fibrillation: classification, pathophysiology, mechanisms and drug treatment," *Heart*, vol. 89, no. 8, pp. 939–943, 2003.
- [5] L. Mainardi, Sornmo, and S. Cerutti, *Understanding Atrial Fibrillation: The Signal Processing Contribution*, Morgan Claypool Publishers, San Rafael, CA, USA, 2008.
- [6] E. J. Benjamin, P. A. Wolf, R. B. D'Agostino, H. Silbershatz, W. B. Kannel, and D. Levy, "Impact of atrial fibrillation on the risk of death," *Circulation*, vol. 98, no. 10, pp. 946–952, 1998.
- [7] C.-H. Chen, "A cell probe-based method for vehicle speed estimation," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E103.A, no. 1, pp. 265–267, 2020.
- [8] C.-H. Chen, F.-J. Hwang, and H.-Y. Kung, "Travel time prediction system based on data clustering for waste collection vehicles," *IEICE Transactions on Information and Systems*, vol. E102.D, no. 7, pp. 1374–1383, 2019.
- [9] J. R. Mehall, R. M. Kohut Jr., E. W. Schneeberger, W. H. Merrill, and R. K. Wolf, "Absence of correlation between symptoms and rhythm in "Symptomatic" atrial fibrillation," *The Annals of Thoracic Surgery*, vol. 83, no. 6, pp. 2118–2121, 2007.
- [10] S. George, I. Rodriguez, D. Ipe, P. T. Sager, I. Gussak, and B. Vajdic, "Computerized extraction of electrocardiograms from continuous 12-lead holter recordings reduces measurement variability in a thorough QT study," *The Journal of Clinical Pharmacology*, vol. 52, no. 12, pp. 1891–1900, 2012.
- [11] R. S. Andersen, A. Peimankar, and S. Puthusserypady, "A deep learning approach for real-time detection of atrial fibrillation," *Expert Systems with Applications*, vol. 115, pp. 465–473, 2019.
- [12] S. Asgari, A. Mehrnia, and M. Moussavi, "Automatic detection of atrial fibrillation using stationary wavelet transform and support vector machine," *Computers in Biology and Medicine*, vol. 60, pp. 132–142, 2015.
- [13] Z. Yao, Z. Zhu, and Y. Chen, "Atrial fibrillation detection by multi-scale convolutional neural networks," in *Proceedings of the 2017 International Conference on Information Fusion IEEE*, Xi'an, China, July 2017.
- [14] X. Xu, S. Wei, C. Ma, K. Luo, L. Zhang, and C. Liu, "Atrial fibrillation beat identification using the combination of modified frequency slice wavelet transform and convolutional neural networks," *Journal of Healthcare Engineering*, vol. 2018, Article ID 2102918, 8 pages, 2018.
- [15] X. J. Wei, C. Zhang, M. Liu et al., "Atrial fibrillation detection by the combination of recurrence complex network and convolution neural network," *Journal of Probability and Statistics*, vol. 2019, Article ID 8057820, 9 pages, 2019.
- [16] B. Pourbabaei, M. J. Roshtkhari, and K. Khorasani, "Deep convolutional neural networks and learning ECG features for screening paroxysmal atrial fibrillation patients," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, no. 12, pp. 2095–2104, 2018.
- [17] H. Dang, "A novel deep arrhythmia-diagnosis network for atrial fibrillation classification using electrocardiogram signals," *IEEE Access*, vol. 7, pp. 2169–3536, 2019.
- [18] A. L. Goldberger, L. A. N. Amaral, L. Glass et al., "PhysioBank, PhysioToolkit, and PhysioNet," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
- [19] G. Moody, "A new method for detecting atrial fibrillation using rr intervals," *Computers in Cardiology*, vol. 10, pp. 227–230, 1983.
- [20] W. De Mulder, S. Bethard, and M.-F. Moens, "A survey on the application of recurrent neural networks to statistical language modeling," *Computer Speech & Language*, vol. 30, no. 1, pp. 61–98, 2015.
- [21] H. Palangi, L. Deng, Y. Shen et al., "Deep sentence embedding using long short-term memory networks: analysis and application to information retrieval," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 4, pp. 694–707, 2016.
- [22] O. Faust, A. Shenfield, M. Kareem, T. R. San, H. Fujita, and U. R. Acharya, "Automated detection of atrial fibrillation using long short-term memory network with RR interval

- signals,” *Computers in Biology and Medicine*, vol. 102, pp. 327–335, 2018.
- [23] P. Cao, X. Li, K. Mao et al., “A novel data augmentation method to enhance deep neural networks for detection of atrial fibrillation,” *Biomedical Signal Processing and Control*, vol. 56, Article ID 101675, 2020.
 - [24] J. Lu, L. A. Hendricks, M. Rohrbach et al., “Long-term recurrent convolutional networks for visual recognition and description,” *IEEE Transactions on Pattern Analysis Machine Intelligence*, vol. 39, no. 4, pp. 677–691, 2017.
 - [25] W.-P. Xiong, T.-C. Li et al., “Research on partial least squares method based on deep confidence network in traditional chinese medicine,” *Discrete Dynamics in Nature and Society*, vol. 2020, Article ID 4142824, 10 pages, 2020.
 - [26] F. Ma, J. Zhang, W. Liang et al., “Automated classification of atrial fibrillation using artificial neural network for wearable devices,” *Mathematical Problems in Engineering*, vol. 2020, Article ID 9159158, 6 pages, 2020.
 - [27] C. Zhang, J. He et al., “A crash severity prediction method based on improved neural network and factor Analysis,” *Discrete Dynamics in Nature and Society*, vol. 2020, Article ID 4013185, 13 pages, 2020.
 - [28] D. M. W. Powers, “Evaluation: from precision, recall and ffactor to roc, informedness, markedness correlation,” *Journal of Machine Learning Technologies*, vol. 2, pp. 37–63, 2011.