Bidirectional Recurrent Neural Network and Convolutional Neural Network (BiRCNN) for ECG Beat Classification*

Pengwei Xie[1], Guijin Wang[1], Chenshuang Zhang[1], Ming Chen[1], Huazhong Yang[1], Tingting Lv[2], Zhenhua Sang[2], Ping Zhang[2]

***Abstract*— We propose a novel electrocardiogram (ECG) beat classification algorithm using a combination of Bidirectional Recurrent Neural Network (BiRNN) and Convolutional Neural Network (CNN) named as BiRCNN. Our model is an end-to-end model. The morphological features of each ECG beat is extracted by CNN. Then the features of each beat are considered in the context via BiRNN. The assessment on MIT-BIH Arrhythmia Database (MITDB) resulted in a sensitivity of 98.7% and a positive predictivity of 96.4% on average for the VEB class. For the SVEB class, the sensitivity was 92.8%, which was an over 6% promotion compared with the state-of-the-art method, and the positive predictivity was 81.9% on average. The results demonstrate the superior classification performance of our method.**

I. INTRODUCTION

Electrocardiogram (ECG) is a noninvasive measurement method widely used in the diagnosis and monitoring for cardiovascular disease. It reflects the electrical depolarization and repolarization patterns of the heart. Any disturbance of heart rate, or change in the morphological pattern, is an indiction of arrhythmia[1]. And automatic classification of arrhythmia by ECG has attracted wide attention for researchers. However, it is a challenging task since ECG signal shows significant variations for different patients under different conditions.

A wide range of automatic ECG classification methods have been proposed. These methods mainly contains two parts: feature extraction and classification. A variety of hand-crafted features has been extracted, inluding higher order statistics [2], the independent component analysis (ICA) based features [3], hermite transform coefficients [2][4][5], wavelet transform features [6][7][8] and temporal features [3][4][5][6][8]. However, these handcrafted features may not represent the underlying characteristics of ECG waveform and restrict the performance.

Traditional classifiers like support vector machines[2], optimum-path forest[4] and other frequently-used classifiers such as Bayesian classifier and decision trees are great on small data sets, but don't perform well on large data sets, which means they can't take full advantage of big data.

Deep learning methods have been applied in many fields and have great achievement, including the field of ECG beat classification. An adaptative implementation of 1D Convolutional Neural Network (CNN) was proposed in [9], which used CNN for feature extraction and classification and achieved good results. Although CNN is good for extracting interior morphological features, it does not utilize the information among beats. In [10], recurrent neural networks (RNN) and clustering technique were used to classify ECG beats. Clustering technique was used to obtain a representative training data set and beat morphology information was directly fed into RNN to get classification results. This method has achieved the state-of-the-art performance so far.

In this work, we propose a novel end-to-end beat classification model based on a combination of Bidirectional Recurrent Neural Network (BiRNN) and Convolutional Neural Network (CNN) named as BiRCNN. ECG waveform from lead I and lead II would be fed into CNN to extract interior morphological features automatically, then these features were put into BiRNN to learn the relations between the current beat and the adjacent beats. Besides, HRV (short for heat rate variance, measuring the length of RR interval) sequences of several beats would be put into BiRNN to extract temporal features. Finally, morphological features and temporal features would be considered synthetically to generate the classification result. The experimental results on the MIT-BIH Arrhythmia Database (MITDB) demonstrate that our proposed system achieves superior beat classification performance.

II. SYSTEM FRAMEWORK

The framework of our proposed system is shown in Figure 1. Firstly ECG data is preprocessed by denoising, data segmentation and grouping. Then the processed data is randomly selected as training data set, and three evaluation data sets are established. Finally our model is trained and then test on the evaluation data sets respectively.

A. Preprocessing

Firstly, we utilize a wavelet-based denoising method to filter the noise. ECG signals are decomposed into 9 scales using Dual-Tree Complex Wavelet Transform(DTCWT)[11]. Information in scale 3 to 8 is retained to reconstruct the signals, while the others are regarded as noise and removed.

Secondly, prior information about ECG signals would be used to segment data from each record. Annotations of R peaks are used to locate beats. We then use a window function *g(x)* to segment waveform of each beat. The window function g(x) is defined as

*This work is partially supported by the National Key Research Project of People's Republic of China No.SQ2017YFB140080-04 and No.2016YFC0900901 and the Self-determined Project of Tsinghua University No.20161080077.

[1]Department of Electronic Engineering, Tsinghua University, Beijing, China

[2]Beijing Tsinghua Changgung Hosipital, Tsinghua, Beijing, China

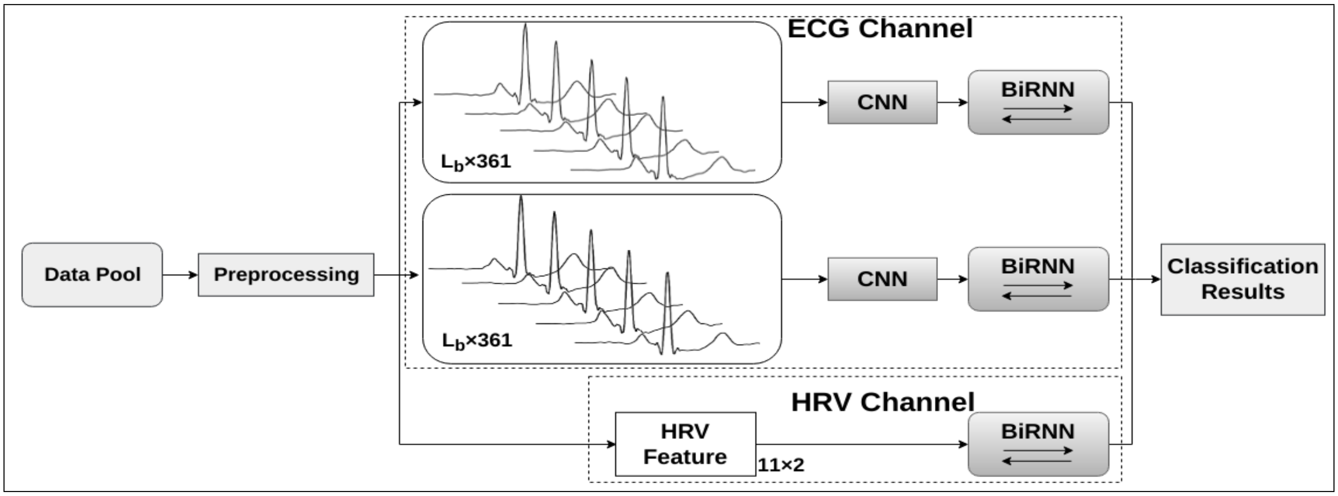

Fig. 1. System Framework.

$$g(x) = \begin{cases} s(100ms, max(x_R - 250ms, 0.7 \times RR_{left})), \\ x < x_R \\ 1 - s(100ms, min(x_R + 100ms, 0.5 \times RR_{right})), \\ x \geq x_R \end{cases}$$

where function *s(var,ave)* represents the *sigmoid* function, with *var* meaning the variance of the *sigmoid* and *ave* meaning the distance between the midpoint of the *sigmoid* and the origin. x_R is the R peak's position of the current beat. RR_{left} represents the RR interval between the current beat and the previous beat, and RR_{right} represents the RR interval between the current beat and the following beat.

Next, signal of each beat segmented by the window function is normalized as a 721-dim vector. The first 180 points and the last 180 points are discarded to reduce data dimension since these points are mostly zero and the remaining 361 points are enough to reflect the characteristics of each beat.

Lastly, when the 361-dim vector of each beat is extracted, we group it with its previous L_t beats and its following L_t beats, resulting as a $L_b \times 361$ matrix, where $L_b = 2 \times L_t + 1$. L_b is the number of beats for each beat group and it will be discussed later. The reason we group beats like this is that we expect our network to learn the features of beats in the context, which is better than considering the features of each beat separately.

B. Classification

We create an end-to-end model named BiRCNN to accomplish the beat classification task. There are two channels of our model: ECG channel and HRV channel. ECG channel learns the morphological features of ECG signal from two leads. HRV channel learns the temporal features. Final classification result is acquired from the features offered by both of ECG channel and HRV channel. More details about our model would be presented in the following part.

III. BiRCNN Model

BiRCNN is used for both feature extraction and classification of ECG beats. The network architecture of our model is shown in Figure 2.

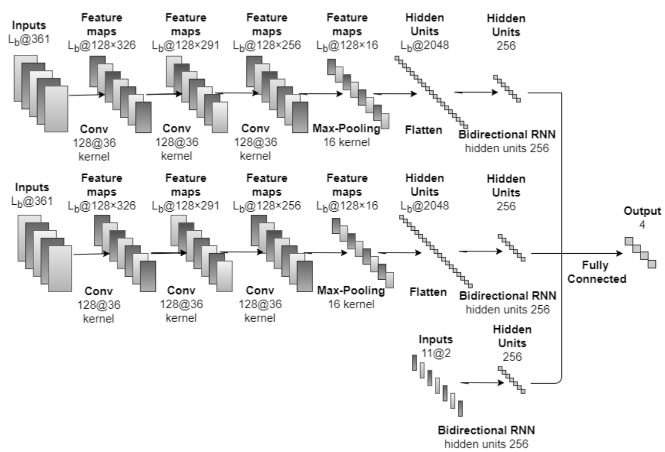

Fig. 2. Network Architecture of BiRCNN Model

A. ECG Channel

In ECG channel, two leads of signal would be processed separately with the same network configuration since some types are easier to be detected from one lead while the others are more suitable to be detected by the other lead.

Data of each beat group, which is arranged as a $L_b \times 361$ matrix, is fed into our networks. Signal of each beat group is processed stepped by time. In every timestep, data of each beat would be fed into 3 convolutional layers. Each convolutional layer is followed by a Rectified Linear Unit (ReLU) activation function. The feature maps are then fed into a max pooling layer and a dropout layer with a keep rate of 0.5. To achieve the superior result, we use 128 convolutional kernels with kernel size of 36 for each convolutional layer, and thus the output size of features is $L_b \times 16 \times 128$. Overall, the interior morphological features of each beat are extracted by CNNs.

The output features from the CNNs are then flattened to $L_b \times 2048$, in order to be input to the Bidirectional RNN (BiRNN). BiRNN, which is formed by concatenating a forwarding simple RNN and a backwarding simple RNN, would be utilized to learn the relations of features in the context. A dropout layer with keep rate of 0.25 would be followed after that to keep sparsity of the networks. As a result, the features of size 256 from each lead would be concatenated.

B. HRV Channel

In HRV channel, HRV sequences of size 11×2 are put into BiRNN to extract the temporal features. HRV of the current R peak and the next R peak is combined with its previous and its following 5 HRV values respectively as a 11-dim vector. HRV of the previous R peak and the current R peak with the surrounding 10 HRV values are combined as the other 11-dim vector. The vectors are normalized and fed into BiRNN to learn the temporal features. The feature size is 256. The reason we use the HRV channel is that some SVEBs have similar morphology with normal beats, but have abnormal RR intervals. Also, in [6], the authors draw the conclusion that the RR interval could improve ECG beat classification performance.

Finally, output features are concatenated to put into a 4-class softmax classifier to acquire the ultimate classification result. We implement the total network based on a deep learning library named Keras.

IV. RESULTS AND DISCUSSION

A. Experiment Setups

As recommended by AAMI[12], all beats should be classified into five types: normal (N), ventricular (V, VEB), supraventricular (S, SVEB), fusion of normal and ventricular (F) and unclassified (Q). In this paper, type Q is not considered since it is marginally represented in the available database and the key mission is to distinguish VEBs and SVEBs from others.

We establish our model and evaluate its performance on MIT-BIH arrhythmia database (MITDB), which is widely used in researches on ECG classification[2]-[6],[8]-[10],[13]. MITDB contains 48 records, each containing two-lead ECG signals for 30-min duration selected from 24-h recordings of 47 individuals. Continuous ECG signals are bandpass-filtered at 0.1-100 Hz and then digitized at 360 Hz. According to the AAMI convention, 4 paced records (102,104,107,217) are excluded from the training and test data sets.

The beat labels in the MITDB from 16 subtypes of 4 types are counted as in Table I. Data of type N (including subtypes of N,L,R,B) counts too much in the data set, so we randomly choose 8,000 pieces of data from those of type N, and then combined them with the data of the other 3 types as a new data set. Due to the fact that there are differences in the waveform among beats of different subtypes, we build the training set from one third of data randomly selected from each subtype in the new data set. The evaluation data sets are clarified in the following part.

TABLE I

NUMBERS OF BEAT LABELS IN MITDB

Label	N	L	R	B	A	a	J	S
Number	74528	8073	7257	0	2545	150	83	2
Label	e	j	n	V	E	r	F	None
Number	16	229	0	6903	106	0	803	0

For classification performance measured, four standard metrics are used: classification accuracy (*Acc.*), sensitivity (*Sen.*), specificity (*Spe.*) and positive predictivity (*Ppr.*). Using true positive (*TP*), false positive (*FP*), true negative (*TN*), false negative (*FN*), *Acc.*, *Sen.*, *Spe.*, *Ppr.* are defined as follows. *Acc.* is the ratio of the number of correctly classified patterns to the total number of patterns classified: *Acc.=(TP+TN)/(TP+TN+FP+FN)*. *Sen.* is the rate of correctly classified events among all events: *Sen.=TP/(TP+FN)*. *Spe.* is the rate of correctly classified nonevents among all nonevents: *Spe.=TN/(TN+FP)*. *Ppr.* is the rate of correctly classified events in all detected events: *Ppr.=TP/(TP+FP)*.

B. Results on MIT-BIH Arrhythmia Database

Our proposed algorithm is compared with four other existing algorithms[5][8][9][10], all of which comply with the AAMI standards. We evaluate the performance on MITDB in three evaluation data sets. Data Set 1 contains 11 records (200, 202, 210, 213, 214, 219, 221, 228, 231, 233, 234) for VEB detection, and another three records (212, 222, 232) are added for SVEB detection. Data Set 2 contains 24 records (200 and onward). Data Set 3 contains all 44 records (without paced records). The average classification results of 10 independent experiments and comparison with the three algorithms are shown in Table II.

Several observations could be made from Table II. First, compared with [5][8][9], our results are much better for all metrics. Second, the performance of our method is comparable with that of [10], which is the state-of-the-art method at present. And it is worth noting that we have a over 6% promotion in Sen. of SVEB compared to [10].

C. BiRNN Evaluation

In order to show the significance of our BiRNN model and to determine the optimal value of L_b(number of beats in each beat group), we train models over the network configuration with different L_b values. It should be noted that when L_b equals 1, it means there is just one beat in each group and BiRNN will be of no effect.

To find out the optimal value of L_b, we change it from 1 to 11 with interval of 2, and evaluate the models on data set 2 respectively. The results are shown in Figure 3, with *Sen.* and *Ppr.* of VEB and SVEB focused.

It shows that as L_b increases, *Sen.* of SVEB remains stable earlier and shows a downward trend then, while *Ppr.* of SVEB first rises and then drops quickly. And it is clear that the optimal value of L_b is 5, which is adopted in our

TABLE II

CLASSIFICATION RESULTS OF THE PROPOSED METHOD AND COMPARISON WITH FOUR ALGORITHMS FROM LITERATURE

Data Set	Methods	VEB				SVEB			
		Acc	*Sen*	*Spe*	*Ppr*	*Acc*	*Sen*	*Spe*	*Ppr*
Data Set 1	Jiang and Kong[5]	98.8	94.3	99.4	95.8	97.5	74.9	98.8	78.8
	Ince et al.[8]	97.9	90.3	98.8	92.2	96.1	81.8	98.5	63.4
	Kiranyaz et al.[9]	98.9	95.9	99.4	96.2	96.4	68.8	99.5	79.2
	Zhang et al.[10]	99.4	97.6	99.7	97.6	**98.7**	87.4	**99.4**	**89.4**
	Proposed	**99.6**	**98.4**	**99.7**	**98.0**	98.4	**92.9**	98.7	83.6
Data Set 2	Jiang and Kong[5]	98.1	86.6	99.3	93.3	96.6	50.6	98.8	67.9
	Ince et al.[8]	97.6	83.4	98.1	87.4	96.1	62.1	98.5	56.7
	Kiranyaz et al.[9]	98.6	95	98.1	89.5	96.4	64.6	98.6	62.1
	Zhang et al.[10]	**99.6**	97.5	**99.8**	**97.9**	**98.9**	86.7	**99.5**	**89.0**
	Proposed	99.4	**98.8**	99.5	95.7	98.7	**92.8**	98.9	81.8
Data Set 3	Ince et al.[8]	98.3	84.6	98.7	87.4	97.4	63.5	99.0	53.7
	Kiranyaz et al.[9]	99	93.9	98.9	90.6	97.6	60.3	99.2	63.5
	Zhang et al.[10]	**99.7**	97.1	**99.9**	**98.1**	**99.3**	85.9	**99.7**	**88.7**
	Proposed	99.6	**98.8**	99.6	95.5	99.1	**92.7**	99.3	80.2

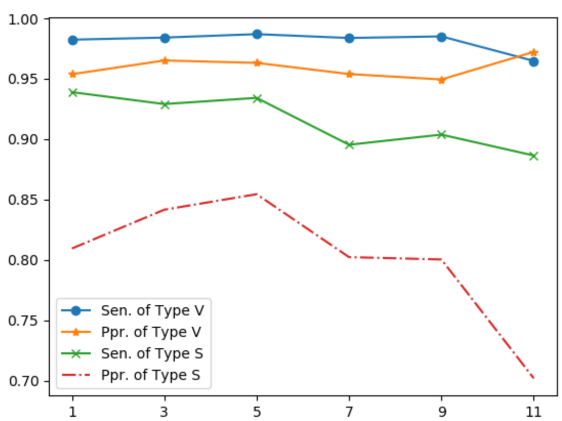

Fig. 3. Sen. and Ppr. of VEB and SVEB for Different Number of Beats in Each Beat Group

final model. In all, we can conclude that BiRNN is pretty important for improving the beat classification performance.

V. CONCLUSIONS

In this paper, we firstly propose a novel end-to-end ECG classification model based on Bidirectional Recurrent Neural Network(BiRNN) and Convolutional Neural Network(CNN) named as BiRCNN. We use CNN to extract interior morphological features of each beat, and the features of each beat are considered in the context via BiRNN. Temporal features are extracted by BiRNN. The experimental results obtained on the MIT-BIH Arrhythmia Database show that our proposed system achieves superior beat classification performance.

REFERENCES

[1] G. J. Taylor, *150 Practice ECGs: Interpretation and Review*. John Wiley & Sons, 2008.

[2] S. Osowski, L. T. Hoai, and T. Markiewicz, "Support vector machine-based expert system for reliable heartbeat recognition," *IEEE transactions on biomedical engineering*, vol. 51, no. 4, pp. 582–589, 2004.

[3] S. N. Yu and K. T. Chou, "Integration of independent component analysis and neural networks for ECG beat classification," *Expert Systems with Applications*, vol. 34, no. 4, pp. 2841–2846, 2008.

[4] E. J. S. Luz, T. M. Nunes, V. H. C. De Albuquerque, J. P. Papa, and D. Menotti, "ECG arrhythmia classification based on optimum-path forest," *Expert Systems with Applications*, vol. 40, no. 9, pp. 3561–3573, 2013.

[5] W. Jiang and S. G. Kong, "Block-based neural networks for personalized ECG signal classification," *IEEE Transactions on Neural Networks*, vol. 18, no. 6, pp. 1750–1761, 2007.

[6] O. T. Inan, L. Giovangrandi, and G. T. Kovacs, "Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features," *IEEE transactions on Biomedical Engineering*, vol. 53, no. 12, pp. 2507–2515, 2006.

[7] İ. Güler and E. D. Übeylı, "ECG beat classifier designed by combined neural network model," *Pattern recognition*, vol. 38, no. 2, pp. 199–208, 2005.

[8] T. Ince, S. Kiranyaz, and M. Gabbouj, "A generic and robust system for automated patient-specific classification of ECG signals," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 5, pp. 1415–1426, 2009.

[9] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 664–675, 2016.

[10] C. Zhang, G. Wang, J. Zhao, P. Gao, J. Lin, and H. Yang, "Patient-specific ECG classification based on recurrent neural networks and clustering technique," in *Biomedical Engineering (BioMed), 2017 13th IASTED International Conference on*. IEEE, 2017, pp. 63–67.

[11] I. W. Selesnick, R. G. Baraniuk, and N. C. Kingsbury, "The dual-tree complex wavelet transform," *IEEE signal processing magazine*, vol. 22, no. 6, pp. 123–151, 2005.

[12] *Testing and reporting performance results of cardiac rhythm and ST segment measurement*, Association for the Advancement of Medical Instrumentation American National Standard, Rev. ANSI-AAMI EC57:2012, 2012.

[13] J. S. Wang, W. C. Chiang, Y. L. Hsu, and Y. T. C. Yang, "ECG arrhythmia classification using a probabilistic neural network with a feature reduction method," *Neurocomputing*, vol. 116, pp. 38–45, 2013.