Received March 9, 2018, accepted April 27, 2018, date of publication May 8, 2018, date of current version June 5, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2833841

Automated ECG Classification Using Dual Heartbeat Coupling Based on Convolutional Neural Network

XIAOLONG ZHAI[1] **AND CHUNG TIN**[1,2,3,4]
[1]Department of Mechanical and Biomedical Engineering, City University of Hong Kong, Hong Kong
[2]Centre for Biosystems, Neuroscience, and Nanotechnology, City University of Hong Kong, Hong Kong
[3]Centre for Robotics and Automation, City University of Hong Kong, Hong Kong
[4]CityU Shenzhen Research Institute, Shenzhen 518057, China

Corresponding author: Chung Tin (chungtin@cityu.edu.hk)

This work was supported in part by the National Natural Science Foundation of China, China, under Grant 31671103, in part by the Research Grants Council of Hong Kong SAR under Project CityU138613 and Project CityU11213717, in part by the City University of Hong Kong under Project 7004456 and Project 7004604, and in part by the Croucher Foundation of Hong Kong under Project CityU 9500014.

ABSTRACT A high performance electrocardiogram (ECG)-based arrhythmic beats classification system is presented in this paper. The classifier was designed based on convolutional neural network (CNN). Single channel ECG signal was segmented into heartbeats in accordance with the changing heartbeat rate. The beats were transformed into dual beat coupling matrix as 2-D inputs to the CNN classifier, which captured both beat morphology and beat-to-beat correlation in ECG. A systematic training beat selection procedure was also proposed which automatically include the most representative beats into the training set to improve classification performance. The classification system was evaluated for the detection of supraventricular ectopic beats (SVEB or S beats) and VEB using the MIT-BIH arrhythmia database. Our proposed method has demonstrated superior performance than several state-of-the-art detectors. In particular, our proposed CNN system has improved sensitivity and positive predictive rate for S beats by more than 12.2% and 11.9%, respectively, over these top performing algorithms. Our proposed CNN classifier with an automatic training beats selection process has shown to outperform the previous methods. The classifier is also a personalized one by combining training set from a common pool and a subject-specific set of ECG data. Our proposed system provides a reliable and fully automatic tool for detection of arrhythmia heartbeat without the need for manual feature extraction or expert assistant. It can potentially be implemented on portable device for the long-term monitoring of cardiac arrhythmia.

INDEX TERMS Convolutional neural network (CNN), ECG classification, arrhythmia, patient-specific.

I. INTRODUCTION

Electrocardiography (ECG) has become a promising source for the study of structure and function of the heart due to its low cost, ease of use, high efficiency and non-invasiveness. It reveals the electrophysiological pattern of depolarization and repolarization of the heart muscles during each heartbeat. The term ''arrhythmia'' refers to any changes from the normal sequences of electrical impulses in the heart which causes abnormal heart rhythms. Arrhythmias can be completely harmless or life-threatening. It may lead to tachycardia or even sudden cardiac arrest. In the research of arrhythmia detection, heartbeat classification based on ECG signal has become a valuable and promising techniques for early warning of arrhythmias. However, variation in ECG signals can be significant among different subjects. Under different circumstances, the morphologies and rhythms produced by the same symptoms of arrhythmia can be quite different as well [1], [2]. Experienced cardiologists can distinguish abnormal heartbeats from normal sinus rhythms easily by observing the ECG. However, this is still a challenging task for a computer to perform automatically due to the variation in ECG signals and the differences of recording environment. For a healthy subject, the morphology and rhythm can be quite variable even in a short period of time [1].

Numerous methods have been proposed for generic heartbeat classification using ECG signal based on techniques such as discrete wavelet transform [3], [4], feature selection [5]–[7], hidden Markov models (HMM) [8] and mixture of experts method [9], [10]. In [3], morphological features (wavelet and independent component analysis) and dynamic features (RR interval) are combined to give a set of more comprehensive features. These methods require certain amount of priori knowledge of the signals or they need expert input frequently. These limit the application of the method and higher variations may be encountered when classifying new subjects' ECG signals. Furthermore, integrity of ECG components, such as P, Q, R, S and T waves, may also be required for these algorithms. However, for arrhythmia, these ECG components may not always be well defined and their extraction becomes ambiguous.s

To circumvent the limitation of these methods that require manual feature selection, some researchers turn their sight to convolutional neural network (CNN) and several CNN based approaches have been proposed for ECG classification recently [10]–[13]. CNN was proposed by LeCun *et al.* [14] in 1990 and it has emerged as one of the most powerful machine learning approaches in recent years. With the great help of rapidly developing graphics process unit (GPU) technology, CNN shows advantages on both accuracy and efficiency in image recognition [15], audio classification [16] and semantic identification [17]. Recent studies have also shown great potentials of CNN in dealing with biomedical applications, such as animal behavior classification [18], histopathological diagnosis [19], protein structure prediction [20] and electromyography (EMG) pattern recognition [21]. Recent studies have also found promising applications of CNN in time series bio-signal such as ECG [10]–[13]. CNN framework has a clear advantage of making use of large training dataset for improving classification performance. For instance, Rajpurkar *et al.* [13] showed a classifier for atrial fibrillation trained from ~30,000 patients' ECG data that can outperform the average cardiologist performance. Even with a smaller training set (hundreds of beats), several recent studies have shown improved performance of cardiac arrhythmia detection with CNN [10]–[12]. These improvements may be due to the feature learning capability of CNN. In contrast to many conventional method, no manual or explicit feature extraction or feature selection will be required. These processes may otherwise lead to loss of information in the data at various stages. Thus, we also took advantages of the feature learning capability of CNN in this study. We explored the 2-D approach for ECG classification with CNN. Using a 2-D encoded ECG inputs, not only the continuous waveforms, but also the relationships of various ECG components in adjacent heartbeats can be readily captured by the convolutional filters. Hence, as compared to these 1-D algorithms, the 2-D approach may boost the performance of the CNN classifier in terms of both network capacity and regularity.

In this paper, we first proposed a 2-D CNN classifier for heartbeat classification. To supply a feature-rich input to the CNN network, three adjacent heartbeats in ECG were transformed into a 2-D coupling matrix that capture morphology of single heartbeat as well as temporal relationship in adjacent beats. Then, we proposed an automatic selection process to include the most representative beats into the training set to improve classifier performance, as opposed to a randomly selected set used in other studies. Our proposed method was tested using publicly accessible MIT-BIH arrhythmias database and compared with previous work following AAMI recommendations [22].

II. METHODOLOGY

A. DATABASE

The MIT-BIH arrhythmia database [1] was used in this study. It is a publicly accessible database which has been widely used to evaluate performance of different ECG-based heartbeat classification algorithms [5], [7], [10]–[12], [23], [24]. The MIT-BIH arrhythmia database comprises of 48 records of ECG collected from 47 subjects (record 201 and 202 came from the same male subject) and each record contains 30 minutes long two-channel ECG signals which were filtered using a 0.1 to 100Hz bandpass filter and digitized at 360 Hz. Each beat has been labeled by two cardiologists independently and the timing of the R peaks (or the local extremum) for each beat is also given. The labels have been continuously updated in the past few decades. The data were handled according to AAMI ECAR-1987 recommended practice [22] for evaluation of our classification algorithm. The same data handling procedure was used in [11], [23], and [24] and our results can be compared to them directly. Four records (#102, #104, #107 and #217) were excluded because they contain paced heartbeats. Based on the characteristics and symptoms of the subjects, the 44 records used can be divided into two groups [1]. The first group (includes 20 records with labels # starting with 1) serve as the representative sample of a variety of waveforms and artifacts that an arrhythmia detector might encounter in routine clinical use. The remaining 24 records (records with labels # starting with 2) include complex ventricular, junctional and supraventricular arrhythmias and conduction abnormalities. Complexity in rhythm, QRS morphology variation or signal quality in some of these records will present significant challenge to arrhythmia detectors. According to the AAMI recommendation [22], each ECG beat can be classified as N (beats originating in the sinus mode), S (supraventricular ectopic beats), V (ventricular ectopic beats), F (fusion beats) or Q (unclassifiable beats). To train a subject specific classifier, the training data consisted of two parts, a common part and a subject specific part. The common part of training data was selected from the first group (record # started with 1) and is used for all testing subjects (second record group, record # started with 2). Following the AAMI recommended practice, at most 5 minutes of recordings from a subject were used for classifier training purpose. So the subject specific part of training data included the heartbeats from the first 5 minutes of the ECG recording of each testing subject. The remaining 25 minutes of the record was used for testing.

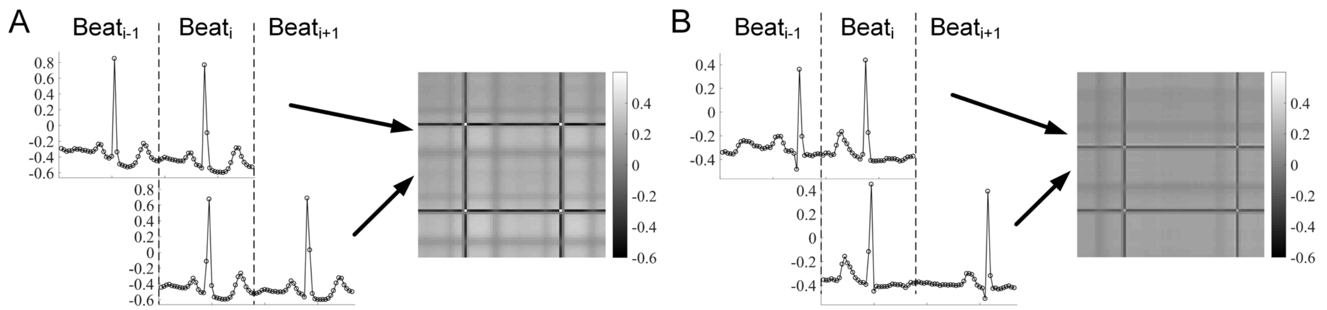

FIGURE 1. Dual-beat coupling matrix computed from two pairs of adjacent beats. A. Coupling matrix of beats originating from the sinus mode. B. Coupling matrix from a supraventricular ectopic beat.

B. ECG SIGNAL PREPROCESSING

In this study, we used the modified lead II channel only from the database. Automatic ECG classification is particularly useful for portable or wearable device and it is expected that few channel number (even single channel) would be found in these devices. Hence, we developed our algorithm to handle small channel number of ECG. The timing of each heartbeat has been labeled for the corresponding R peak in the database. Hence, we can directly obtain the R-R intervals for each beat in our segmentation. Nevertheless, numerous robust methods have already been available for R peak detection [25], [26] and algorithm for this is beyond the scope of current study.

1) HEARTBEAT SEGMENTATION

The morphology of each heartbeat is critical for classification of arrhythmia. Some studies segmented the ECG signals into equal length at preprocessing [11]. However, the heartbeat rates may vary significantly among different subjects and over time, and hence each beat is of different length. For this reason, the heartbeat rate, and hence the beat length, should be considered dynamically over time and specifically for each subject. The beat was segmented such that it was centered around the R peak.

For each record in the first group (the common part), the segmentation length is computed as the average R-R interval over $\pm$T sec from the current beat. Hence, the length for each beat is variable. In this study, we chose T = 10sec such that the average R-R interval is calculated over 20sec. For each record in the second group (the subject specific part, both training and testing beats), the beats are segmented using the R-R interval averaged over the first 5 minutes of the record. As such, all the beats in the second group for each subject have equal length. Using average R-R interval over every 20sec yielded similar results in our study.

2) SEGMENT LENGTH SCALING

In order to unify the different segment length due to variation in heartbeat rate for inputting into the classifier, the segments are first scaled into the same length. If the length (in sample number) of an extracted beat is N_i, we will first up-sample it, by interpolation, by a factor equals to the input size, M, designed for the CNN. Hence, the length of the beat becomes $N_i \times M$. Then, the mean values for every N_i samples are calculated such that eventually, we can obtain a segment of length M for beat with any original length of N_i.

C. INPUTS FOR CLASSIFIER

The dual-beat coupling matrix, which integrates both beat waveform and beat-to-beat correlation was used as input in this study. Heartbeat arrhythmia can be analyzed not only by single beat morphology, but also beat-to-beat correlation. Hence, we took into account a series of three adjacent beats in formulating the input for the classifier. Two pairs of heartbeat are extracted following the same segmentation principle as described above. The first pair consists of the current beat and the previous beat, denoted as the column vector:

$$Dual\_beat_{i-1,i} = Beat_{i-1}[L] \ldots Beat_{i-1}[L], \\ Beat_i[1], \ldots\ldots Beat_i[k], \ldots Beat_i[L] \quad (1)$$

where $Beat_i[k]$ is the k^{th} sample of ECG signal in time of the current beat. The second pair consists of the current beat and the next beat (denoted as the column vector $Dual\_beat_{i,i+1}$ in similar way). All three beats have the same length extracted based on the current beat ($Beat_i$). These two dual-beat vectors are scaled to the required input size M respectively, as described above. Then we compute the coupling matrix (CM), with size M$\times$M, as follow,

$$CM = \lfloor Dual\_beat_{i-1,i}[1], \ldots, Dual\_beat_{i-1,i}[M] \rfloor \\ \bullet \left[Dual\_beat_{i,i+1}[1], \ldots, Dual\_beat_{i,i+1}[M] \right]^T \quad (2)$$

Two examples of the resultant coupling matrix with size M$\times$M are shown in Fig. 1.

The coupling matrices integrate both morphological and rhythmic information of the ECG into a single input. For instance, in Fig. 1, the four high intensity points (white) correspond to the four R peaks of the two pairs of heartbeat in concern. The intensity reflects the relative amplitude of R peaks in these the beats, while the relative position of these points reflects the rhythm of this segment of ECG. Fig. 1A shows a segment of ECG originated from sinus mode, while Fig. 1B shows a segment around a supraventricular ectopic beat. In Fig. 1A, the four white points distribute more evenly

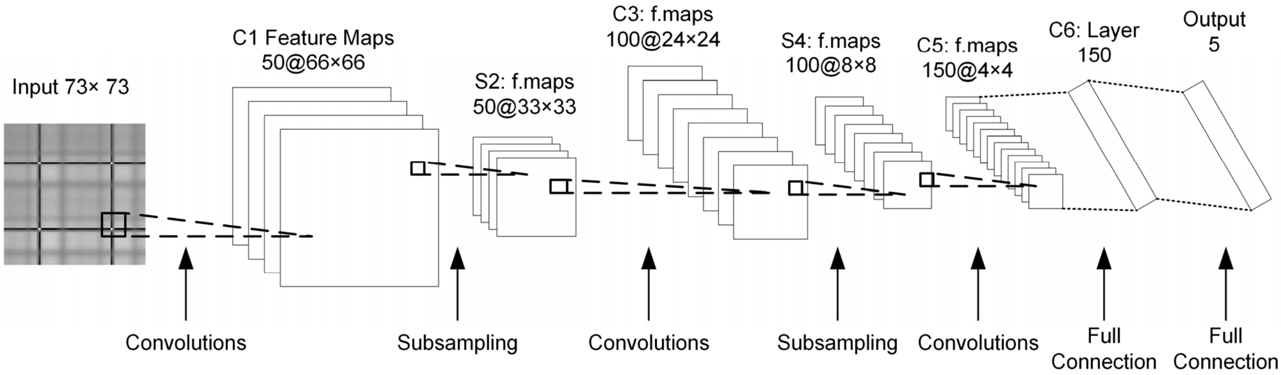

FIGURE 2. Schematic for proposed CNN classifier.

over the matrix and it corresponds to a regular rhythm. Other parts of the coupling matrix provide further information about the segments such as the morphology and rhythm of P, Q, S and T waves.

D. CNN CLASSIFIER

Fig. 2 shows a schematic for our CNN classifier. This CNN classifier has a similar form to the one used for MNIST handwritten digit database [27]. Our CNN model contains 3 convolutional layers, 2 sub-sampling layers (1 maximum sub-sampling layer and 1 average sub-sampling layer), 1 fully connected layer with dropout and a softmax loss layer. Rectified linear units (ReLU) is used as activation function. Different input sizes were tested for the best performance. The convolution filter sizes were adjusted accordingly. An open source MATLAB toolbox MatConvNet was used to implement the CNN classifier [28].

E. CLASSIFICATION PROCEDURES

In this study, we followed the AAMI recommendation [22] using the same procedure as in [11], [23], and [24] to evaluate our classification system. The training data consists of two parts, a common part and a subject-specific part. The common part comes from the first record group (record# from 100 to 124) and remains the same for all testing subjects. The common part consists of 245 representative beats, including 75 N beats, 75 S beats, 75 V beats, 13 F beats and 7 Q beats (more details below). The subject-specific part includes all the heartbeats from the first 5 minutes of ECG recordings of the testing subject. Beats from the remaining 25 minutes were used for testing.

F. SELECTION OF TRAINING BEATS

In previous studies [11], [23], [24], the training beats were usually randomly selected from the dataset. However, among S beats in particular, the morphology can vary significantly from beat to beat and some S beats were easily misclassified as N beats. As such, previous studies have suffered from a relatively low sensitivity and/or positive predictive rate for S beats [11], [23], [24]. Hence, more care is required in selecting the training beats to overcome these problems.

Here, we proposed a procedure for selecting the more representative S beats for the training set without manual input. In the first group records (the common part), S beats exist only in 12 (out of 20) records (#100, 101, 103, 108, 112, 113, 114, 116, 117, 118, 121 and 124). We randomly select 75 N beats and 75 S beats respectively from each of these 12 records as training set for a preliminary classifier (same structure as Fig. 2) and test on 100 N beats selected from each of the remaining 8 records, hence a total of 800 beats. The prediction accuracy is noted for these 75 S beats in the training set. We repeat the procedure for 200 times using different set of 75 S beats for training the preliminary classifier. After that, we average the accuracy for each S beat selected over these 200 simulations and the 75 S beats with the highest accuracy are picked as the training data in the common part for training the actual classifier. The procedure is shown in Fig. 3. The idea is to include the S beats that have significantly different morphology than the N beats (hence high prediction accuracy for N beats in this selection process) in the training set for the final classifier such that the resultant classifier will be less prone to confusion between S beats and N beats. We also compared the performance using the 75 S beats with the lowest average accuracy as well randomly selected S beats.

G. ASSESSMENT INDICATORS

To evaluate the performance of the proposed classifier, we used four statistical indicators in this study, which have been commonly used in previous work [5]–[7], [9]–[12], [23], [24], [29]. They are classification accuracy (Acc), sensitivity (Sen), specificity (Spe) and positive predictive rate (Ppr). The classification Acc measures the overall performance of the proposed method on all valid heartbeats. However, as the number of different types of beats varies, Sen, Spe and Ppr are less biased in assessing the classifier performance.

The four statistical indicators can be calculated as follow,

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$Sen = \frac{TP}{TP + FN} \tag{4}$$

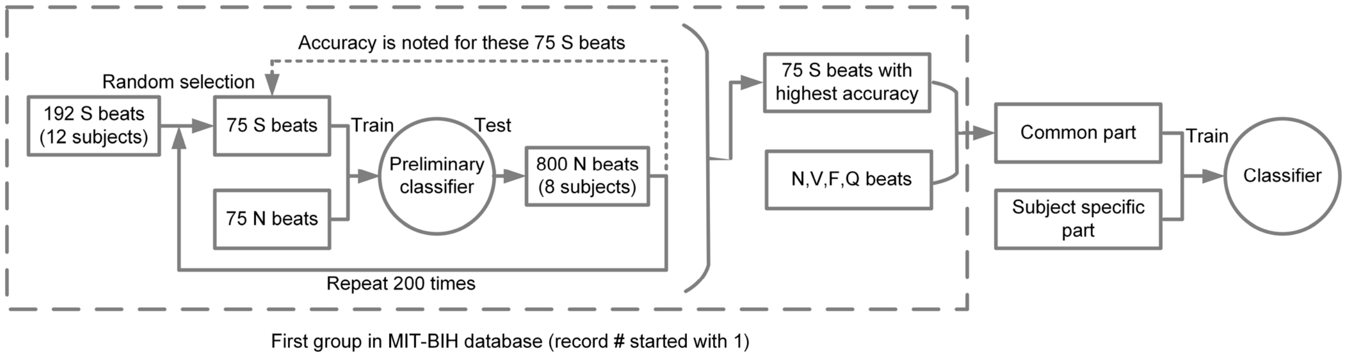

FIGURE 3. Selection of training beats.

$$Spe = \frac{TN}{TN + FP} \quad (5)$$

$$Ppr = \frac{TP}{TP + FP} \quad (6)$$

where TP is true positive, TN is true negative, FP is false positive and FN is false negative. We here evaluated the classification performance for S and V beat detection as in several previous studies [5]–[7], [9]–[12], [23], [24], [29].

TABLE 1. Confusion matrix of the ECG beat classification results for the 24 test records in the MIT-BIH arrhythmia database.

Ground Truth	Classification result				
	N	S	V	F	Q
N	40877	593	171	191	32
S	357	1797	163	20	3
V	171	36	4510	83	8
F	88	0	37	487	0
Q	4	1	2	0	1

III. RESULTS

Table 1 shows the confusion matrix for all testing beats in the 24 test records in the MIT-BIH arrhythmia database using input size of 73. It is shown that some of the S beats are misclassified as N beats, but the number is significantly lower than the previous best result by Kiranyaz *et al.* [11] using 1-D CNN.

To determine the effect of input size on the performance, we tested with six different input sizes. We started with input size of 28 which is the input resolution used in MNIST database. The results are shown in Table 2. Table 2 shows that Sen and Ppr of S beats are particularly sensitive to the input size. Acc and Spe for both S and V beat classifications are generally high at >95% and change only slightly with different input size. For better visualization, we plotted the Sen and Ppr of S and V beats for different input size in Fig. 4. It is clear that at small input size of 28, the Sen and Ppr for both S and V beat detection are the lowest. This is likely due to the low resolution of input to capture the necessary information of the original ECG signal. Yet, Sen and Ppr for V beat remind relatively high (>80%), probably because V beats are usually well distinguished from other beat type in waveform and even a lower input resolution is sufficient to reveal the differences. Fig. 4 shows that Ppr of S beats drops with larger input size. This is likely because the network structures were not adjusted properly for each individual input size. The oversized convolutional kernel may not give good performance. Fig. 4 suggests that an input size of 73 provided relatively high performance in both Sen and Ppr of V beats and S beats in our network. Hence, we used input size of 73 as the preferred choice in this study.

TABLE 2. VEB and SVEB classification performance with different input sizes.

Input size	VEB				SVEB			
	Acc	Sen	Spe	Ppr	Acc	Sen	Spe	Ppr
28	96.7	81.6	98.4	84.1	96.4	54.1	98.5	64.3
43	97.3	91.4	97.9	82.5	97.7	66.2	99.3	82.6
58	98.5	92.8	99.1	92.0	97.9	72.1	99.1	80.6
73	98.6	93.8	99.2	92.4	97.6	76.8	98.7	74.0
88	98.5	90.9	99.3	93.4	97.5	78.1	98.4	70.9
103	98.3	90.1	99.2	92.5	97.0	76.8	98.0	66.0

As we discussed above, the selection of training beats is critical for the classifier performance. Here, we compared the actual classifier performance using training S beats that were denoted as the most representative beats and those denoted as the least representative beats by our automatic selection process (see Method Section F "Selection of training beats"). We also compared with the results using randomly selected training beats. Input size of 73 was used in all cases. Simulation for each case was repeated 10 times. The results are shown in Table 3. The selection process significantly affected Ppr of S beat detection. With the most representative S beats, mean Ppr for S beats is the highest with the smallest standard deviation. Hence, the proposed selection process, which is fully automatic, not only improves the Ppr for S beats, but also increases the stability of the classifier from sampling the training set.

Table 4 summarizes the top performance of our proposed classifier as compared to several other methods that followed AAMI ECAR-1987 recommended practice. It is shown that our classifier generally performed among the top in all indicators. It is worth noting that our approach has significantly

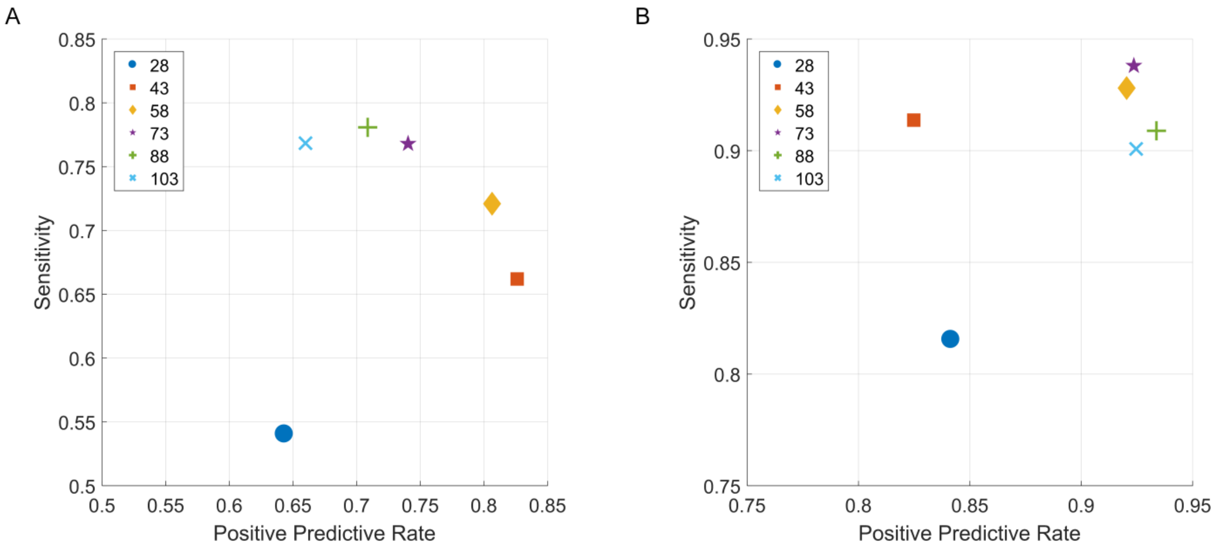

FIGURE 4. Sensitivity and positive predictive rate using different input sizes for (A) S beats and (B) V beats.

TABLE 3. Comparison of different selection methods of training beats. The results are presented as mean ± SD.

Selection methods	VEB			
	Acc	Sen	Spe	Ppr
Most Rep. S	98.7±0.1	93.2±0.6	99.3±0.1	93.4±0.6
Least Rep. S	98.6±0.1	93.4±0.4	99.2±0.1	92.5±0.6
Randomly Selected S	98.5±0.3	92.7±1.8	99.1±0.3	92.0±2.3
Selection methods	SVEB			
	Acc	Sen	Spe	Ppr
Most Rep. S	97.6±0.1	77.0±0.5	98.7±0.1	73.9±1.3
Least Rep. S	96.2±0.2	79.2±0.7	97.1±0.2	57.4±2.0
Randomly Selected S	96.3±0.9	78.6±1.4	97.2±0.9	58.2±7.9

TABLE 4. VEB and SVEB classification performance of the proposed method and comparison with former studies (24 testing records).

Methods	VEB				SVEB			
	Acc	Sen	Spe	Ppr	Acc	Sen	Spe	Ppr
Jiang [23]	98.1	86.6	99.3	93.3	96.6	50.6	98.8	67.9
Ince [24]	97.6	83.4	98.1	87.4	96.1	62.1	98.5	56.7
Kiranyaz[11]	98.6	95.0	98.1	89.5	96.4	64.6	98.6	62.1
Proposed	98.6	93.8	99.2	92.4	97.5	76.8	98.7	74.0

improved Sen and Ppr of S beats up to >70%. These are the two indicators that suffered from low performance in the other methods.

The AAMI also recommends that the problem of V beat and S beat detection can be considered separately. According to their recommendation, for V beat detection, the testing dataset contains 11 records (#200, 202, 210, 213, 214, 219, 221, 228, 231, 233, and 234) and for S beat detection, the testing dataset contains 14 records (#200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, and 234). We evaluated our classifier using the same procedure and the results are summarized and compared with several other studies in Table 5. Our proposed method again performed among the top for almost all indicators. Although our Ppr for S beats is lower than the other two methods, our method could still achieve >70% Ppr without sacrificing the Sen (85.3% vs. 74.9% and 68.8%). Hence, we showed that our proposed method has a generally high performance regardless of the classification procedure.

TABLE 5. VEB and SVEB classification performance of the proposed method and comparison with former studies (11 recordings for testing VEB detection and 14 recordings for testing SVEB detection.)

Methods	VEB				SVEB			
	Acc	Sen	Spe	Ppr	Acc	Sen	Spe	Ppr
Hu [9]	94.8	78.9	96.8	75.8	N/A	N/A	N/A	N/A
Jiang [23]	98.8	94.3	99.4	95.8	97.5	74.9	98.8	78.8
Ince [24]	97.9	90.3	98.8	92.2	96.1	81.8	98.5	63.4
Kiranyaz[11]	98.9	95.9	99.4	96.2	96.4	68.8	99.5	79.2
Proposed	99.1	96.4	99.5	96.4	97.3	85.3	98.0	71.8

IV. CONCLUSIONS

We have proposed a CNN-based framework for heartbeat classification using dual-beat ECG coupling matrix. This 2-D encoded dual-beat coupling matrix of ECG is an effective representation of both heartbeat morphology and rhythm. We have also proposed an automatic selection procedure for picking the most useful training beats systematically to boost the classification performance. Our proposed method was tested using the MIT-BIH arrhythmia database. We showed that our 2-D CNN-based classifier can offer 12.2% higher Sen and 11.9% higher Ppr respectively for S beat detection when compared with previous 1-D CNN based method [11]. Furthermore, performance metrics for VEB detection are all well above 90%. These results support that our 2-D CNN framework could be a useful tool for automatic heartbeat classification without explicit ECG feature extraction.

The recent decade has witnessed the tremendous advancement of CNN and other deep learning frameworks in computer vision, audio recognition, artificial intelligence and so on. The application of CNN in biomedical signal has also attracted more attention. Recently, three studies have implemented CNN for ECG-based heartbeat classification using the MIT-BIH arrhythmia database [10]–[12]. ECG signal is a time signal and often a single channel was used in these previous studies, and hence a 1-D input. However, the relationship among different ECG components in adjacent heartbeats may also be useful indicators of presence of arrhythmia. Inter-beat features have shown to be useful in boosting classification performance of ECG [30]. Nevertheless, such relationships are difficult to convey using 1-D input. Here, we proposed to use dual ECG beat coupling matrix as inputs for the CNN classifier, which are 2-D presentation of the beat morphology and beat-to-beat correlation of the 1-D ECG signal. While including more beats in forming the input may be advantageous, we have tested our proposed system using five adjacent heartbeats. However, no improvement was observed (Sen of S beats indeed dropped by $\sim$12%) and it is suggested that our system design will need specific optimization in handling five adjacent beats. Nevertheless, we have verified that using these 2-D inputs can boost the classifier performance over using 1-D inputs, and hence 1-D convolution filter, particularly in terms of Sen and Ppr of S beats detection (Tables 4 and 5). Although only a few hundreds of beats were used in training the CNN in order to balance the samples of different beat types, our CNN system emerges as a powerful tool of feature learning and demonstrated improved performance over several previous classifiers.

We have also proposed a systematic procedure for selecting the more representative S beats to improve the classifier performance. Al Rahhal *et al.* [10] also performed an iterative procedure to identify the uncertain beats to add to the training set but they required a human expert to relabel them at each iteration for fine tuning the classifier in order to reach a high performance. Instead, our proposed method is fully automatic. We have verified that our method can improve the performance and stability of the classifier especially for S beat Ppr by reducing the misclassification of N beat to S beat. This is particularly important for those subjects whose ECG beats in the first 5min of record do not contain S beats, and hence not found in the subject specific training set. Although annotating clinical signal is costly and can only be performed by expert, we believe that with a bigger common pool database in addition to the MIT-BIH database, the Sen and Ppr of S beats can be further improved to match the level of other indicators. Rajpurkar *et al.* [13] collected an extensive set of ECG records (64,121 records from 29,163 patients) and showed that a deep CNN classifier for detection of atrial fibrillation trained from them can outperform the average cardiologist performance. This should motivate the establishment of more extensive public ECG database for advancing the related classification technology.

The popularity of CNN and other deep learning frameworks in various applications have led to more efficient computational tools which have essentially improved the speed of the training process and eased the complication of implementation. In this study, with the help of NVIDIA CUDA®Deep Neural Network library [31], our CNN classifier can be effectively parallelized with NVIDIA GTX 980m graphic card. The average training time for each subject is only 134s. These advantages make the CNN a more flexible platform for heartbeat classification using ECG signal. While the network structure used in this study mainly inspired from those in image recognition, further work will investigate improved network for ECG *per se*. Future work will examine the robustness of the classification system in long-term use.

REFERENCES

[1] R. Mark and G. Moody, *MIT-BIH Arrhythmia Database Directory*. Cambridge, MA, USA: MIT, 1988.

[2] K. Minami, H. Nakajima, and T. Toyoshima, "Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 2, pp. 179–185, Feb. 1999.

[3] C. Ye, B. V. K. V. Kumar, and M. T. Coimbra, "Heartbeat classification using morphological and dynamic features of ECG signals," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 10, pp. 2930–2941, Oct. 2012.

[4] 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, 2013.

[5] P. de Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 7, pp. 1196–1206, Jul. 2004.

[6] T. Mar, S. Zaunseder, J. P. Martínez, M. Llamedo, and R. Poll, "Optimization of ECG classification by means of feature selection," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 8, pp. 2168–2177, Aug. 2011.

[7] T. Teijeiro, P. Félix, J. Presedo, and D. Castro, "Heartbeat classification using abstract features from the abductive interpretation of the ECG," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 2, pp. 409–420, Mar. 2017.

[8] R. V. Andreao, B. Dorizzi, and J. Boudy, "ECG signal analysis through hidden Markov models," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 8, pp. 1541–1549, Aug. 2006.

[9] Y. H. Hu, S. Palreddy, and W. J. Tompkins, "A patient-adaptable ECG beat classifier using a mixture of experts approach," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 9, pp. 891–900, Sep. 1997.

[10] M. M. Al Rahhal, Y. Bazi, H. AlHichri, N. Alajlan, F. Melgani, and R. R. Yager, "Deep learning approach for active classification of electrocardiogram signals," *Inf. Sci.*, vol. 345, pp. 340–354, Jun. 2016.

[11] 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.

[12] U. R. Acharya, H. Fujita, O. S. Lih, Y. Hagiwara, J. H. Tan, and M. Adam, "Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network," *Inf. Sci.*, vol. 405, pp. 81–90, Sep. 2017.

[13] P. Rajpurkar, A. Y. Hannun, M. Haghpanahi, C. Bourn, and A. Y. Ng. (2017). "Cardiologist-level arrhythmia detection with convolutional neural networks." [Online]. Available: https://arxiv.org/abs/1707.01836

[14] Y. Le Cun *et al.*, "Handwritten digit recognition with a back-propagation network," in *Proc. Adv. Neural Inf. Process. Syst.*, 1990, pp. 396–404.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[16] G. Hinton *et al.*, "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, Nov. 2012.

[17] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, Apr. 2016.

[18] U. Stern, R. He, and C.-H. Yang, "Analyzing animal behavior via classifying each video frame using convolutional neural networks," *Sci. Rep.*, vol. 5, Sep. 2015, Art. no. 14351.

[19] G. Litjens *et al.*, "Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis," *Sci. Rep.*, vol. 6, May 2016, Art. no. 26286.
[20] S. Wang, J. Peng, J. Ma, and J. Xu, "Protein secondary structure prediction using deep convolutional neural fields," *Sci. Rep.*, vol. 6, Jan. 2016, Art. no. 18962.
[21] X. Zhai, B. Jelfs, R. H. M. Chan, and C. Tin, "Self-recalibrating surface emg pattern recognition for neuroprosthesis control based on convolutional neural network," *Frontiers Neurosci.*, vol. 11, p. 379, 2017, doi: 10.3389/fnins.
[22] *Recommended Practice for Testing and Reporting Performance Results of Ventricular Arrhythmia Detection Algorithms*. Arlington, VA, USA: Association Advancement Medical Instrumentation, 1987.
[23] W. Jiang and S. G. Kong, "Block-based neural networks for personalized ECG signal classification," *IEEE Trans. Neural Netw.*, vol. 18, no. 6, pp. 1750–1761, Nov. 2007.
[24] T. Ince*, S. Kiranyaz, and M. Gabbouj, "A generic and robust system for automated patient-specific classification of ECG signals," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 5, pp. 1415–1426, May 2009.
[25] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 3, pp. 230–236, Mar. 1985.
[26] C. Li, C. Zheng, and C. Tai, "Detection of ECG characteristic points using wavelet transforms," *IEEE Trans. Biomed. Eng.*, vol. 42, no. 1, pp. 21–28, Jan. 1995.
[27] Y. LeCun, C. Cortes, and C. J. Burges. (2010). *MNIST Handwritten Digit Database*. AT&T Labs. vol. 2. [Online]. Available: http://yann.lecun.com/exdb/mnist
[28] A. Vedaldi and K. Lenc, "MatConvNet: Convolutional neural networks for MATLAB," in *Proc. ACM Conf. Multimedia*, 2015, pp. 689–692.
[29] M. Llamedo and J. P. Martinez, "An automatic patient-adapted ECG heartbeat classifier allowing expert assistance," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 8, pp. 2312–2320, Aug. 2012.
[30] Z. Zhang, J. Dong, X. Luo, K.-S. Choi, and X. Wu, "Heartbeat classification using disease-specific feature selection," *Comput. Biol. Med.*, vol. 46, pp. 79–89, Mar. 2014.
[31] S. Chetlur *et al.* (2014). "cuDNN: Efficient primitives for deep learning." [Online]. Available: https://arxiv.org/abs/1410.0759

XIAOLONG ZHAI received the B.Eng. degree in communication engineering from Tongji University in 2012 and the M.Sc. degree in electrical and electronic engineering from The University of Hong Kong in 2014. He is currently pursuing the Ph.D. degree in mechanical and biomedical engineering with the City University of Hong Kong. His current research interests include neuroprosthetic systems, machine learning, and biomedical signal processing and analysis.

CHUNG TIN received the B.Eng. degree from The University of Hong Kong in 2002 and the M.Sc. and Ph.D. degrees from the Massachusetts Institute of Technology in 2004 and 2011, respectively, all in mechanical engineering. He has been an Assistant Professor with the Department of Mechanical and Biomedical Engineering, City University of Hong Kong, since 2012. His research interests include computational neuroscience, sensorimotor learning, and neuroprosthetic systems. He was a recipient of the Croucher Foundation Scholarship, Hong Kong, the American Heart Association Predoctoral Fellowship, and the Early Career Award from the Research Grant Council, Hong Kong.