RESEARCH ARTICLE

Deep Regression Network With Sequential Constraint for Wearable ECG Characteristic Point Location

ZUO WANG[1,2], JINLIANG WANG[3], MINGYANG CHEN[3], WEI YANG[1,2], AND RONG FU[1,2]
[1]School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China
[2]Guangdong Provincial Key Laboratory of Medical Image Processing, Guangzhou 510515, China
[3]CardioCloud Medical Technology (Beijing) Company Ltd., Beijing 100084, China

Corresponding authors: Rong Fu (furong@smu.edu.cn) and Wei Yang (weiyanggm@gmail.com)

ABSTRACT Accurate location of characteristic points in wearable ECG signals may be a challenge due to the high noise. Taking the time sequence of waveforms and missing waveforms into account, we design a location regression network ECG_SCRNet, combined with the sequential constraints to accurately identify characteristic points of wearable ECGs. We add a classification head to determine whether there is a P-wave or a T-wave missing. This architecture ensures that the network considers both the time sequence of physiological waveform and class information to improve the accuracy in locating characteristic points. The proposed ECG_SCRNet was evaluated on a wearable dataset and the LUDB, achieving highly accurate results compared to other state-of-the-art methods. On the wearable dataset, the average *Sen, PPV* and *F1*score are 97.13%, 99.96%, and 99.51%. On the LUDB, the average *Sen, PPV* and *F1* score are 96.86%, 99.83%, and 98.97%. These results demonstrate that the proposed ECG_SCRNet has good flexibility and reliability when applied to signal characteristic point detection, and it is a reliable method for analyzing ECG signals in real time.

INDEX TERMS Wearable ECG, ECG characteristic points, heartbeat segmentation, CNN.

I. INTRODUCTION

Analysis of electrocardiogram (ECG) signals is one of the most important steps in the diagnosis of cardiac disorders [1]. Usually, cardiologists perform a visual inspection of the ECG in order to diagnose a patient, interpreting potential pathological deviations in the waveform. ECG signal characteristic points detection is the process of identifying characteristic points, such as onsets, peaks, and offsets of different ECG waveforms. Figure 1 shows an idealized ECG heartbeat with waves labeled including P, QRS and T-wave. Characteristic points detection can be performed directly on all available leads (multi-lead) or individual leads (single-lead). In order to achieve high diagnostic performance almost all the ECG analysis tools/software require information about the location and morphology of different segment

The associate editor coordinating the review of this manuscript and approving it for publication was Santosh Kumar.
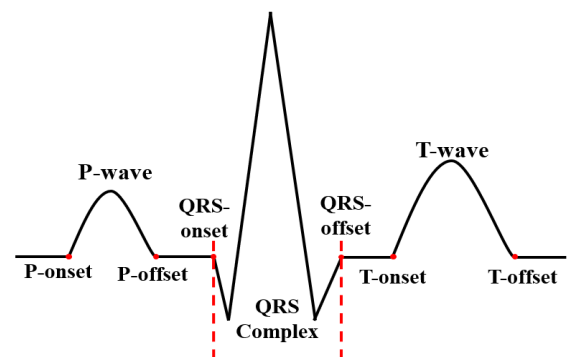
FIGURE 1. Idealized ECG heartbeat with characteristic points of different segment waveforms (P-QRS-T).

waveforms (P-QRS-T) in ECG records. Wave segmentation is of high medical interest, since clinically relevant intervals, amplitude values, or other features of the individual waves

and the segments in between can be derived and used for diagnostics.

In recent years, the advances in wearable technologies [2] and digital signal processing technologies have enabled the remote continuous monitoring of ECGs and human physical activities during daily life. There is an increasing demand for automatic applications working on wearable ECGs. Different from the ECG signal collected by the traditional electrocardiograph, various noise types that occupy the frequency band of the ECG signals, such as electromyogram (EMG) interference and motion artifacts, are simultaneously recorded by wearable devices. There are also human factors, such as the user's improper operation, which will also produce a heavy interference to the ECGs. Therefore, how to perform accurate and rapid detection in the presence of so much noise has become the focus of this research.

There are many methods to detect the characteristic points of ECG automatically. Digital signal processing algorithms using the wavelet transform and rule-based adaptive thresholds are often cited as state-of-the-art for ECG detection [3], [4], [5], reaching high precision and recall for the P, QRS and T waves. However, these methods require laborious rule adaptation when extended to morphologies outside the development dataset; moreover, these algorithms depend on the dataset. Their performance will be degraded as the dataset changed, and compromise their generalization.

Deep learning methods often show significant improvements in supervision tasks with respect to classical methods and can be successfully used to improve and automate heavy tasks in the field of medicine if sufficient training data are available [6]. Therefore, some deep learning methods have been applied to the classification of ECG signals to obtain detection results. The commonly used method is to use the encoder-decoder structure, like U-net [7], [8], [9], [10], [11], for waveform segmentation. Sodman et al. [12] utilized a novel convolutional neural network (CNN) with different kernel sizes for automatic P-waves, QRS complexes, and T-waves annotation on the QT database, which was comparable in performance to other state-of-the-art methods. Malali et al. [13] and Nurmaini et al. [14] used bidirectional long short-term memory (BiLSTM) network to classify ECG waveforms into three categories (P-waves, QRS complexes, and T-waves) and achieved an acceptable performance. Peimankar et al. [15] developed a CNN-2BiLSTM network for characteristic point detection of P-waves, QRS complexes, T-waves, and No waves (NW). This experiment indicated that combining CNN with LSTM layers achieved better performance than models with only CNN or LSTM layers. Xiaohong Liang et al. [16] proposed a standard dilated convolution module (SDCM) into the encoder path enabling the model to extract more useful ECG signal-informative features and a BiLSTM to obtain numerous temporal features. The feature sets of the ECG signals at each level in the encoder path were connected to the decoder part for multi-scale decoding to mitigate the information loss caused by the

pooling operation in the encoding process. However, these methods are applicable to routine ECGs collected at rest rather than wearable ECGs, which are often heavily affected by noise interference. And few studies have taken the time sequence of various waveforms into consideration. Therefore, using the aforementioned method to detect wearable ECG signals directly could result in detection errors that significantly impact the final index calculation. Furthermore, the method may perform poorly if a signal misses a P-wave or T-wave.

To solve these problems, we proposed a novel network, named as ECG_SCRNet, to detect wearable ECG signal characteristic points, which considers the time sequence of waveforms and utilizes a location regression architecture. This method can accurately detect the onsets and offsets of P, QRS complex, and T-waves. The main contributions of our methods are as follows:

(1) We propose an ECG signal characteristic points detection network based on directly predicting the location of waveforms that can reduce the parameters of the model.

(2) We consider the time sequence of waveforms to the detection results, a sequential constraint regression (SC-Regression) module is proposed to improve the accuracy of predictions.

(3) The proposed classification head is to predict whether there is the absence of waveform to assist the detection of waveform missing signals.

The rest of this paper is organized as follows: the methods and details of our network are introduced in Section II. The experimental results and discussion are given in Section III. Section IV is the discussion part. Finally, the conclusion is drawn in Section V.

II. MATERIALS AND METHODS
A. MATERIALS
In this paper, we utilized two datasets. The first is a private dataset consisting of ECG signals collected by a wearable ECG device developed by CardioCloud Medical Technology (Beijing) Co., LTD. Each signal has a duration of 15 seconds and a sampling frequency is 500Hz. The wearable 12-lead ECG dataset contains data from 767 users, covering 6 types of complex ECG patterns, including ST-segment change (STC), atrial fibrillation (AF), tachycardia, premature atrial contraction (PAC), premature ventricular contraction (PVC) and bundle branch block (BBB). Figure 2 displays two ECGs with missing waveform signals. The collected data is labeled by cardiologists from CardioCloud, which contains the position information of the starting and ending points of the P-wave, QRS complex and T-wave, as well as the R-wave peak. The signals are segmented into heartbeats using the R-wave peaks and are reshaped to the same length. We obtained a total of 4190 heartbeats, of which 759 are missing the P-wave and 80 are missing the T-wave. The last heartbeat of each piece of data is used as the test set, and the remaining
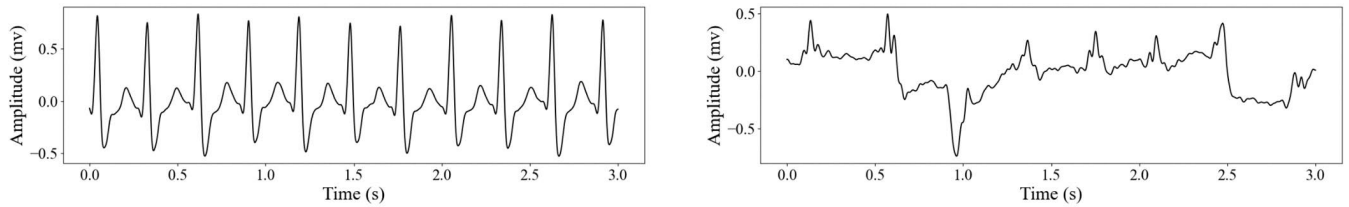
FIGURE 2. Two example ECG signals of the wearable dataset. Left shows the ECG of a patient with tachycardia without a significant P-wave. Right shows the ECG of a patient with AF and other arrhythmia symptoms, which P-wave and T-wave, which both P-wave and T-wave are not significant.
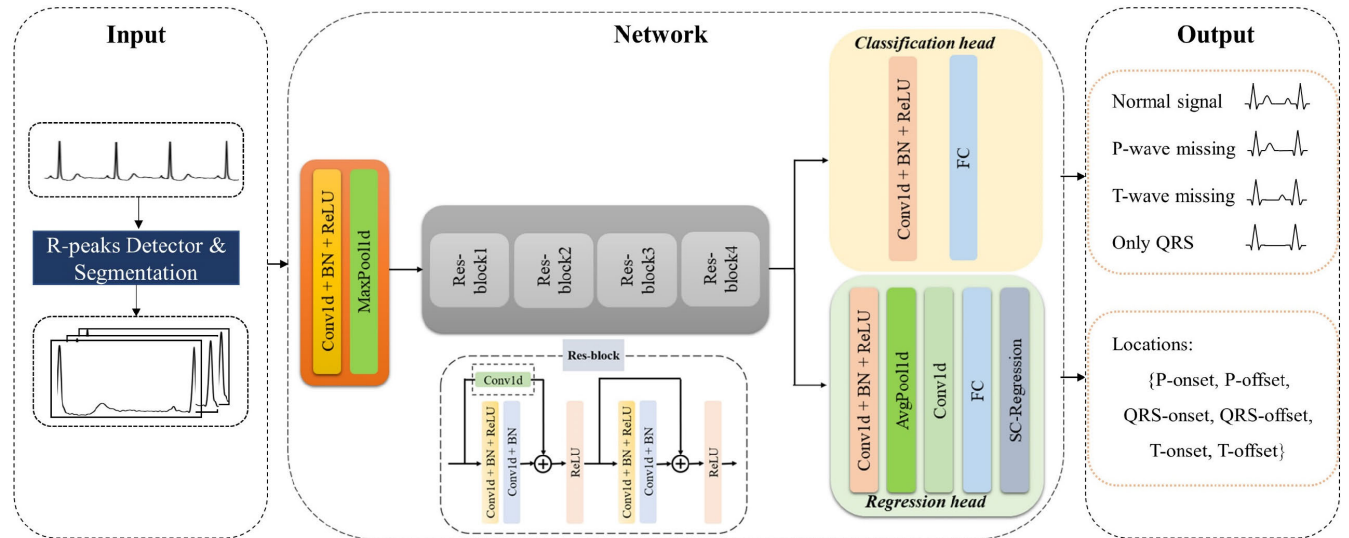
FIGURE 3. Flowchart of the proposed method. The input is 12-lead signal and only one lead is shown. The middle part displays the detailed composition of the ECG_SCRNet, which consists of a CNN encoder, a classification head and a regression head. Res-block means residual block and SC-Regression is the sequential constraint regression.

TABLE 1. Number of heartbeats for four waveforms in the wearable ECG dataset.

	Normal	P-wave missing	T-wave missing	Only QRS	Total
Training set	2768	593	50	12	3423
Test set	600	149	13	5	767
Total	3368	742	63	17	4190

heartbeats are used as the training set. The distribution of the records is shown in Table 1. To address noise interference in the wearable ECG data, we use various data amplification strategies, including random shielding, random lead shielding, and random lead flipping, to amplify the training set. Additionally, the data is normalized using the mean and standard deviation calculated over the entire dataset before being fed into the network for training.

The other dataset is the Lobachevsky University Database (LUDB) [17] established by Lobachevsky University, containing 200 pieces of ECG data from different subjects. The signal duration for each piece of data is 10 seconds, and the sampling frequency is 500Hz. Each lead of the ECG data in LUDB has markers including P, QRS, and T waves, and we select the marker on lead II as the label information.

For the neural network input, we used 12-lead ECG data, which were band-pass filtered using a fifth-order bandpass filter with a frequency range of 0.05-35Hz. This filter was used to remove baseline drift, power frequency interference, and other high-frequency noise, which could interfere with accurate analysis of the ECG data

B. METHODS

The proposed method for detecting characteristic points in ECG signals is illustrated in Figure 3. Initially, a 12-lead ECG signal is taken as input and segmented into individual heartbeats using an R-peak detection algorithm. These heartbeats are then fed into a neural network that has been designed to identify the locations of the P, QRS, and T-waves within the ECG signal, as well as to classify the type of heartbeat.

After obtaining the initial locations of the characteristic points, the results of the classification are used to modify these locations as necessary. This is done to account for any variations in the ECG signal that may be caused by different types of heartbeats, such as arrhythmias or other abnormalities.

1) HEARTBEAT SEGMENTATION

In our proposed method for ECG signal analysis, we first segment the ECG signal into individual heartbeats. To achieve
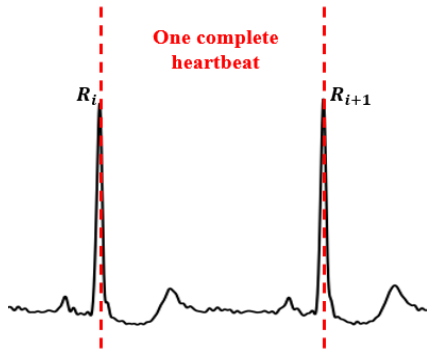
FIGURE 4. Diagrammatic representation of the proposed heartbeat segmentation strategy.

this, we define what constitutes a complete heartbeat and then use a segmentation algorithm to identify the start and end points of each heartbeat.

Unlike other segmentation methods that rely on physiological features, we prefer to use morphological features to segment the ECG signal. As shown in Figure 4, we directly extract the signal between two consecutive R-peaks as a single heartbeat, ensuring that there is no overlap between adjacent heartbeats. A complete heartbeat is defined as a set of sample points between two consecutive R-peaks, denoted as R_i and R_{i+1}, where R_i is the i-th R-peak. The advantage of such processing is that it ensures that there is only one characteristic point of each type within a heartbeat. We use a R-peaks detection method called HA-UNet, which was proposed by Tan et al. [18]. This method is based on a wearable ECG signals R-peaks detection network, which takes into account the characteristic features of individual heartbeats when detecting R-peaks.

By using a combination of our novel heartbeat segmentation strategy and the HA-UNet R-peaks detection method, we are able to accurately and efficiently detect individual heartbeats from the ECG signal, which is an essential step in our overall approach to ECG signal analysis.

2) NETWORK ARCHITECTURE

A multi-task learning framework [19] is used to address two tasks simultaneously, a classification task to determine whether there is a lack of waveform, and a regression task, which can directly map an input to locations.

The middle part of Figure 3 gives the architecture of the ECG_SCRNet. The encoding structure consists of a 1D Resnet backbone, which contains a Conv-modules (a convolutional layer, a batch normalization layer, and a ReLU activation layer), a max pooling layer and four residual blocks. The architecture of the residual block is also shown in Figure 3. The classification head is composed of a Conv-module and a fully connected layer, which classifies segmentations into four classes: normal, P-wave missing, T-wave missing and only QRS. For the regression head, after a Conv-module, we use an average pooling layer to average the feature maps. Then, a convolutional layer and a fully

connected layer are used to predict the intervals of each point. Finally, the proposed sequential constraint regression (SC-Regression) module is used to obtain the final locations of every waveform.

The final loss function is defined as following:

$$L_{Total} = L_{cls} + \lambda(L_{Inte} + L_{Loc}) \tag{1}$$

where L_{cls} is the loss function for the classification task, L_{Inte} represent the loss for the interval, L_{Loc} is for the location and λ is the weight of two tasks, which is automatically determined by the size of the parameters for each head [20].

Although there are 780 heartbeats missing P-wave, this is a relatively small fraction of the total number of normal heartbeats, and there are even fewer T-waves missing. Therefore, to account for the imbalanced data, we have chosen to use Focal loss [21] as the loss function, which is shown as following:

$$L_{cls} = -\alpha(1 - p_t)^\gamma log(p_t) \tag{2}$$

where α equalizes the positive and negative samples, γ is a hyperparameter which is set as two. This will allow us to place greater emphasis on correctly classifying the minority class (i.e., the heartbeats with missing P-waves) and reduce the impact of the majority class (i.e., the normal heartbeats) on the model's training. Overall, this should help to improve the model's performance in detecting heartbeats with missing P-waves, despite the imbalanced nature of the data. The loss function of L_{Inte} and L_{Loc} is to calculate the mean absolute error (MAE) between the prediction and truth of the interval and location.

3) SEQUENTIAL CONSTRAINT

The regression head of our proposed method is responsible for determining the start and end points of the different waveforms in each heartbeat. However, because there is a certain sequence to the appearance of waveforms within a heartbeat, directly regressing the locations can lead to errors in the sequence of points due to the presence of noise in the ECG signal. For example, the P-onset point may be incorrectly detected between the T-onset and T-offset points.

To address this issue, we proposed a sequential constraint regression (SC-Regression) method, which the location can be calculate as following:

$$L_{Loc} = Softmax(L_{Inte}) \times \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \tag{3}$$

For the regression task, we first calculate the interval between two adjacent points L_{Inte}. A Softmax activation layer is then applied to ensure that the interval is greater than zero. We then multiply L_{Inte} by a matrix in the formula above. This is achieved by summing the intervals separately to reflect the position of different waveform points while maintaining the natural time sequence. In other words, the output is

determined by the sum of interval, ensuring that the resulting sequence of points is in an increasing state. $L_{Inte} = [x_1, x_2, \ldots, x_n]$ is the interval set and the final locations $L_{Loc} = [x_1, x_1 + x_2, \ldots, \sum_{i=1}^{n} x_n]$. If the signal is lack of waveform, for example the P-wave is missing, we define that the P-wave is fused with the T-wave. The interval between T-offset and P-onset is zero. Similarly, the interval between P-onset and P-offset is zero and the location of P-onset and P-offset is equal to the location of T-offset.

However, if we sum the interval directly, the error of the lower point would be larger due to the accumulated error. Therefore, for the activation function, we choose Softmax activation, which is defined as following:

$$X_{Inte\_pred} = e^{z_i} \Big/ \sum_{c=1}^{C} e^{z_c} \qquad (4)$$

where z_i means the i-th prediction of interval and C is the number of the interval. After exponential transformation of interval, it will be greater than zero and the difference between each interval is increased. Moreover, the sum of all values is one, avoiding the influence of accumulated error and making the result more in line with the actual situation.

III. EXPERIMENTAL RESULT

A. EVALUATION METRICS

In this study, sensitivity (*Sen*), positive prediction value (*PPV*), and F1-score (*F1*) were used to evaluate the detection performance of the model for P, QRS, and T-waves. These metrics are defined as follows:

$$Sen = \frac{TP}{TP + FN} \times 100\% \qquad (5)$$

$$PPV = \frac{TP}{TP + FP} \times 100\% \qquad (6)$$

$$F1 = \frac{2 \times Sen \times PPV}{Sen + PPV} \times 100\% \qquad (7)$$

where TP is true positive, FP indicates false positive, and FN is false negative. According to the recommendation of the Association for the Advancement of Medical Instrumentation (AAMI) [22], [23], [24], when the heart rate is 70 beats per minute, if the absolute deviation between the algorithm detection result and the doctor's annotation is no more than 150ms, the algorithm detection result is considered correct. If the predicted value is within 150ms of the true value the doctor annotated, the result will be marked as a true positive (TP). Otherwise, the result is marked as a false positive (FP). If the algorithm fails to detect within the range of the doctor's annotation, the result will be marked as a false negative (FN). We get a better performance when the FP and TP are closer to zero.

B. EXPERIMENTAL SETUP

The experiments were implemented in Python and the deep learning framework was Pytorch. The optimizer chosen for model training was Adam. The batch size was set to 512 and

TABLE 2. Detection performance of P, QRS, and T characteristic points on wearable ECG dataset.

	TP	FN	FP	Sen (%)	PPV (%)	F1 (%)
P onset	600	7	28	99.85	95.54	97.17
P offset	599	7	29	98.84	95.38	97.08
QRS onset	766	0	1	100	99.87	99.93
QRS offset	767	0	0	100	100	100
T onset	752	1	4	99.87	99.47	99.67
T offset	747	1	9	99.87	99.81	99.34

TABLE 3. Detection performance of P, QRS, and T characteristic points on LUDB.

	TP	FN	FP	Sen (%)	PPV (%)	F1 (%)
P onset	1127	31	40	97.32	96.57	96.95
P offset	1123	31	44	97.31	96.23	96.77
QRS onset	1422	0	10	100	99.30	99.65
QRS offset	1432	0	0	100	100	100
T onset	1417	6	7	99.58	99.51	98.54
T offset	1413	6	11	99.58	99.23	99.40

the total number of epochs was set to 2000. The initial learning rate was set to 0.0001, and the learning rate was adjusted using a cosine annealing strategy.
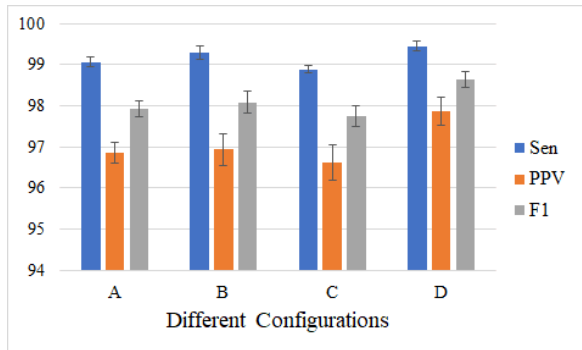
C. RESULTS

Table 2 displays the performance of the P, QRS, and T-waves onset and offset on wearable ECG dataset. The total number of P, QRS, and T-waves are 635, 767 and 757, with the respective TP for P-onset, P-offset, QRS-onset, QRS-offset, T-onset and T-offset being 600, 599, 766, 767, 752 and 747. The average *Sen*, *PPV* and *F1* of the three waveforms are 97.13%, 99.96%, and 99.51%. The performance obtained for QRS complex is better, followed by T-wave and P-wave. To better verify the effectiveness of the proposed ECG_SCRNet, we also conduct experiments of the data on the LUDB. Table 3 displays the performance of the P, QRS, and T-waves onset and offset on the LUDB. The total number of P, QRS, and T-waves are 1198, 1432 and 1430, with the respective TP for P-onset, P-offset, QRS-onset, QRS-offset, T-onset and T-offset being 1127, 1123, 1422, 1432, 1417 and 1414. The average *Sen*, *PPV* and *F1* of the three waveforms are 96.86%, 99.83%, and 98.97%. Likewise, our ECG_SCRNet gets best performance on QRS complex, followed by T-wave and P-wave.

D. ABLATION STUDY

To validate the effectiveness of our ECG_SCRNet, we conducted an ablation study on the wearable dataset. Figure 5 and Table 4 show the different configurations of experiments that we performed to verify the efficacy of our method. Table 4 provides the details of the different configurations

TABLE 4. The average P, QRS and T-wave F1-score of different configurations on the wearable dataset.

	Activation		SC-	Classification	F1 (%)		
	Sigmoid	Softmax	Regression	Head	P	QRS	T
Conf. A	✓			✓	94.92	98.61	95.68
Conf. B	✓		✓	✓	96.37	98.75	96.49
Conf. C		✓	✓		95.91	99.23	96.02
Conf. D		✓	✓	✓	**97.13**	**99.96**	**99.51**

FIGURE 5. The detection performance of various configurations on the wearable dataset.

TABLE 5. Detection performance of P, QRS, and T characteristic points on wearable ECG dataset with a tolerance window interval of 50ms.

|	TP	FN	FP	Sen (%)	PPV (%)	F1 (%)
P onset	570	7	58	98.79	90.76	94.61
P offset	587	7	41	98.82	93.47	96.01
QRS onset	764	0	3	100	99.61	99.80
QRS offset	764	0	3	100	97.61	99.80
T onset	694	1	62	99.86	91.80	95.66
T offset	699	1	60	99.86	92.06	95.80

TABLE 6. Detection performance of P, QRS, and T characteristic points on LUDB with a tolerance window interval of 50ms.

	TP	FN	FP	Sen (%)	PPV (%)	F1 (%)
P onset	1083	31	84	97.22	92.80	94.96
P offset	1080	31	87	97.21	92.54	94.82
QRS onset	1401	0	31	100	97.83	98.91
QRS offset	1414	0	18	100	98.74	99.37
T onset	1378	6	46	99.57	96.77	98.15
T offset	1391	6	33	99.57	97.68	98.62

and the average performance for P, QRS and T-wave. Figure 5 demonstrates the overall average performance.

The result demonstrates that our ECG_SCRNet (Conf. D) achieves the best average performance. Specifically, when compared to configurations of A, B and D, the best results are obtained when we use SC-regression and the Softmax activation layer. Furthermore, the results obtained with the classification head are better than those obtained without classification head, as seen when comparing Conf. C and Conf. D. It is also noteworthy that the best performance was achieved for the QRS complex, followed by the T and P-waves. Thus, our ECG_SCRNet is particularly effective in accurately locating the points of the QRS complex. Overall, the ablation study confirms the validity of ECG_SCRNet, and the results demonstrate that our ECG_SCRNet is more effective in accurately detecting the QRS complex, as well as the P and T-waves.

IV. DISCUSSION

In this study, we focus on signals that lacked some waveforms and consider the time sequence of waveforms to propose a novel ECG signal characteristic points located network (ECG_SCRNet) used the sequential constraint regression, matching a novel heartbeat segmentation strategy. Our proposed method was verified using two datasets. As seen in Figure 5, the addition of SC-regression and the Softmax activation layer resulted in improved results. Furthermore, the simple multi-task learning structure is found to improve the predictions compared to Conf. C, which do not include

a classification head. Table 3 demonstrated a better performance when LUDB data was not included in the training set, indicating that our method has good generalization.

Figure 6 (a)-(d) show the delineation result of different signals according to the points we detected. In Figure 6(a), the P-wave is missing in some heartbeats. Figure 6 (b) and Figure 6 (c) show performance for a noise level signal. While in Figure 6 (d), there is only QRS complex. The results are close to the doctor's annotation position and our proposed method can cope with data of various complex situations.

However, the tolerance window interval of 150ms is too loose for actual demand. Therefore, we reduced the tolerance to 50ms. Table 5 and Table 6 show the performance at the tolerance window interval of 50ms. Although we use a stricter tolerance, the result show better performance. The average *Sen, PPV* and *F1* are 99.56%, 94.21% and 96.95% on the wearable ECG dataset and 98.93%, 96.06% and 97.47% on LUDB.
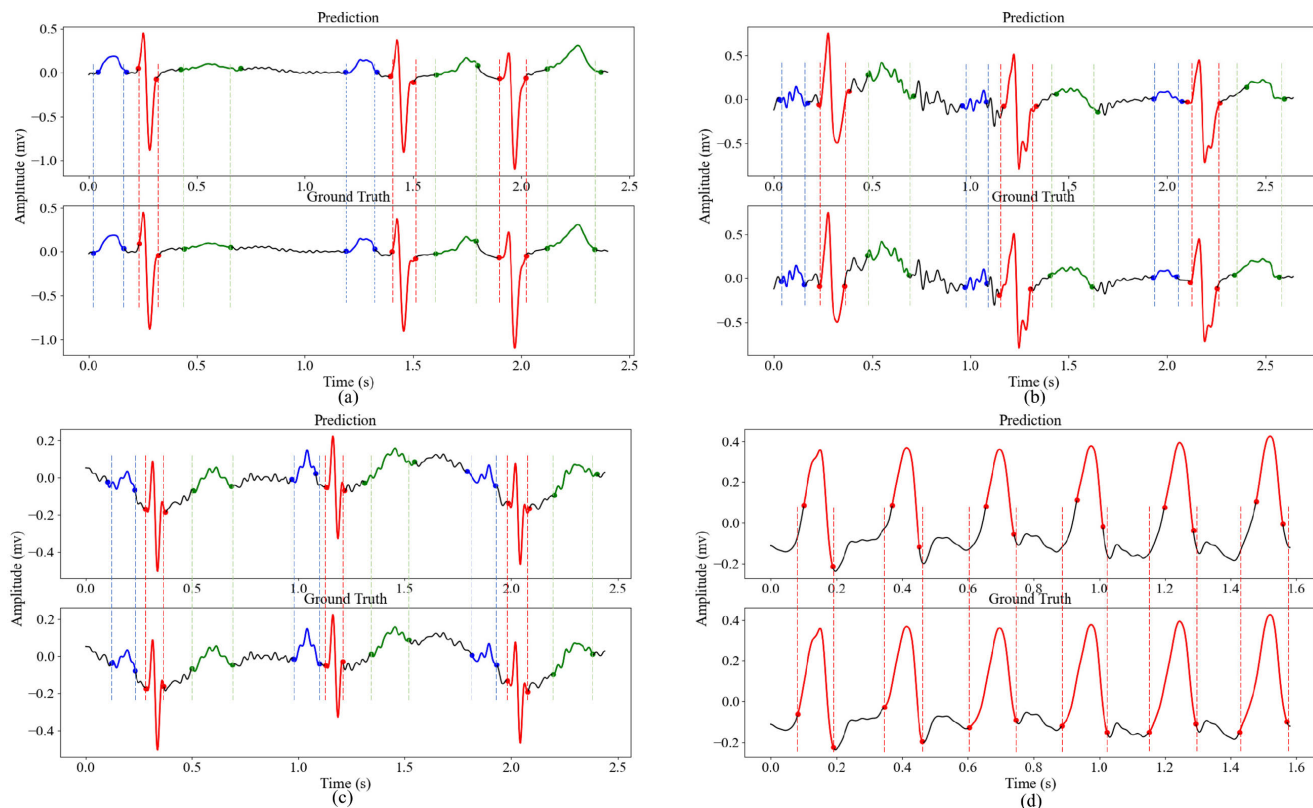
FIGURE 6. The heartbeat segmentation results of the wearable dataset. The blue part represents P-wave, the red part represents QRS complex and the green part represents T-wave.

TABLE 7. F1-score (%) of various methods on the wearable ECG dataset.

Methods	P onset	P offset	QRS onset	QRS offset	T onset	T offset
UNet [7]	93.69	94.45	98.89	98.94	96.98	96.40
3-UNet cascade [25]	94.42	95.28	97.56	98,42	97.10	97.71
3CNN_2BiLSTM [15]	92.27	93.56	94.77	96.80	94.40	94.66
ECG_SegNet [16]	95.82	95.76	98.40	98.69	97.83	97.70
ECG_SCRNet (Ours)	**97.17**	**97.08**	**99.93**	**100**	**99.67**	**99.34**

TABLE 8. F1-score (%) of various methods on LUDB.

| Methods	P onset	P offset	QRS onset	QRS offset	T onset	T offset
UNet [7]	92.70	98.83	97.86	97.81	97.01	96.15
3-UNet cascade [25]	93.37	93.60	98.82	98.82	97.18	95.46
3CNN_2BiLSTM [15]	91.45	94.44	94.31	94.56	95.40	94.50
ECG_SegNet [16]	94.72	94.76	98.89	98.86	97.91	97.15
ECG_SCRNet (Ours)	**96.95**	**96.76**	**99.63**	**99.59**	**98.54**	**99.38**

To further verify the validity of our method, we also evaluated it against other state-of-the-art methods. Table 7 and Table 8 display the P, QRS and T waveforms detection *F1* on the two datasets. As seen in Table 7 and Table 8, our method is superior to the other models for the three waveforms.

Figure 7 and Figure 8 show the average *F1* at different tolerance windows. Our method also obtains good results under different tolerance windows.

In summary, we used quantitative analysis to compare our method to other advanced models and found that our method
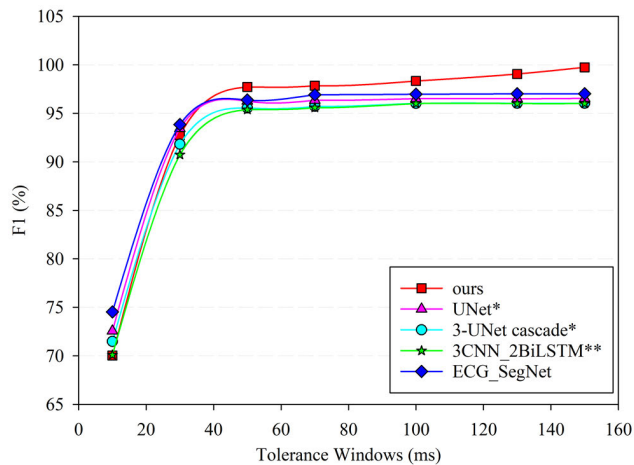
FIGURE 7. The average F1-score (%) of various methods with different tolerance windows on the wearable ECG dataset. *P < 0.05, **P < 0.01 vs ours.
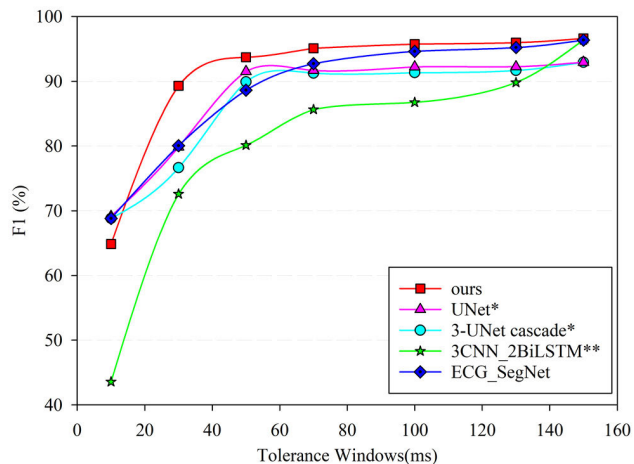
FIGURE 8. The average F1-score (%) of various methods with different tolerance windows on LUDB. *P < 0.05, **P < 0.01 vs ours.

yielded the best results on both datasets. Currently few wearable signal-based methods have been applied to the medical domain. By using a combination of neural network analysis and post-processing techniques, the proposed method is able to accurately and reliably detect the characteristic points of ECG signals, even in the presence of noise or other signal artifacts. This can be a valuable tool for clinicians and researchers working in the field of cardiology, as it allows them to quickly and accurately analyze ECG data and diagnose potential cardiac conditions.

However, the proposed method has three main limitations. First, we only labeled 767 signals chosen from various forms that we can't ensure that it contains all kinds of signals. As we known, deep learning is data-driven. If there is lack of the related data, the model will show a worse performance. Second, our method is highly dependent on the R-peaks detection algorithm. If the R-peaks in an ECG signal can't detect well, it will have a great impact on the

signal segmentation and then affect on the result of other waveforms. Third, because our method is basing on the time sequence of the waveforms and there are few signals which the waveforms would make a cross. This is contrary to our method that our method can't perform well.

In future studies, there are three limitations that need to be addressed. First, we need to work with clinicians to expand the dataset in order to increase the diversity of ECG signals. Second, we can incorporate effective data augmentation techniques to further improve the diversity of data or utilize unlabeled ECG signals to develop semi-supervised or unsupervised methods. Finally, we need to consider a more comprehensive strategy to eliminate the dependence on R-peaks detection and to handle all forms of signals. By addressing these limitations, we can further improve the performance and applicability of the proposed method.

V. CONCLUSION

Deep learning techniques have demonstrated improved performance over classical approaches for supervised tasks when provided with sufficient training data. One of the most challenging tasks in analyzing ECG waveforms is the accurate detection of the P, QRS, and T waves. In this paper, a deep learning approach was proposed that utilized a sequential constraint regression algorithm to directly predict the location of these three waveforms. By combining waveform classification with onset and offset prediction, the trained model was able to accurately determine the location of the waves. The model was then tested on two different test sets to evaluate its performance. The results showed that the proposed method performed well on both sets, with the F1 score of 99.51% and 98.97% on the wearable ECG dataset and the LUDB. The QRS complex showed the best performance followed by the P and T waves. These experimental results demonstrate that the proposed approach is feasible for analyzing all kinds of ECG signals.

DECLARATION OF COMPETING INTEREST

The authors have no relevant conflicts of interest to declare.

REFERENCES

[1] A. Mincholé, J. Camps, A. Lyon, and B. Rodríguez, "Machine learning in the electrocardiogram," *J. Electrocardiol.*, vol. 57, pp. S61–S64, Nov. 2019, doi: 10.1016/j.jelectrocard.2019.08.008.

[2] E. De Giovanni, T. Teijeiro, G. P. Millet, and D. Atienza, "Adaptive R-peak detection on wearable ECG sensors for high-intensity exercise," *IEEE Trans. Biomed. Eng.*, vol. 70, no. 3, pp. 941–953, Mar. 2023, doi: 10.1109/TBME.2022.3205304.

[3] J. P. Martinez, R. Almeida, S. Olmos, A. P. Rocha, and P. Laguna, "A wavelet-based ECG delineator: evaluation on standard databases," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 4, pp. 570–581, Apr. 2004, doi: 10.1109/TBME.2003.821031.

[4] S. Banerjee, R. Gupta, and M. Mitra, "Delineation of ECG characteristic features using multiresolution wavelet analysis method," *Measurement*, vol. 45, no. 3, pp. 474–487, Apr. 2012, doi: 10.1016/j.measurement.2011.10.025.

[5] C. Böck, P. Kovács, P. Laguna, J. Meier, and M. Huemer, "ECG beat representation and delineation by means of variable projection," *IEEE Trans. Biomed. Eng.*, vol. 68, no. 10, pp. 2997–3008, Oct. 2021, doi: 10.1109/TBME.2021.3058781.

[6] S. Sanchez-Martinez, O. Camara, G. Piella, M. Cikes, M. Á. González-Ballester, M. Miron, A. Vellido, E. Gómez, A. G. Fraser, and B. Bijnens, "Machine learning for clinical decision-making: Challenges and opportunities in cardiovascular imaging," *Frontiers Cardiovascular Med.*, vol. 8, Jan. 2022, Art. no. 765693, doi: 10.3389/fcvm.2021.765693.

[7] G. Jimenez-Perez, A. Alcaine, and O. Camara, "U-Net architecture for the automatic detection and delineation of the electrocardiogram," in *Proc. Comput. Cardiol. (CinC)*, Sep. 2019, pp. 1–4, doi: 10.22489/CinC.2019.284.

[8] G. Jimenez-Perez, A. Alcaine, and O. Camara, "ECG-DelNet: Delineation of ambulatory electrocardiograms with mixed quality labeling using neural networks," 2020, arXiv:2005.05236.

[9] G. Jimenez-Perez, A. Alcaine, and O. Camara, "Delineation of the electrocardiogram with a mixed-quality-annotations dataset using convolutional neural networks," *Sci. Rep.*, vol. 11, no. 1, p. 863, Jan. 2021, doi: 10.1038/s41598-020-79512-7.

[10] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2015, pp. 234–241.

[11] P. Sodmann and M. Vollmer, "ECG segmentation using a neural network as the basis for detection of cardiac pathologies," in *Proc. Comput. Cardiol.*, Sep. 2020, pp. 1–4, doi: 10.22489/CinC.2020.356.

[12] P. Sodmann, M. Vollmer, N. Nath, and L. Kaderali, "A convolutional neural network for ECG annotation as the basis for classification of cardiac rhythms," *Physiolog. Meas.*, vol. 39, no. 10, Oct. 2018, Art. no. 104005, doi: 10.1088/1361-6579/aae304.

[13] A. Malali, S. Hiriyannaiah, G. M. Siddesh, K. G. Srinivasa, and N. T. Sanjay, "Supervised ECG wave segmentation using convolutional LSTM," *ICT Exp.*, vol. 6, no. 3, pp. 166–169, Sep. 2020, doi: 10.1016/j.icte.2020.04.004.

[14] S. Nurmaini, A. E. Tondas, A. Darmawahyuni, M. N. Rachmatullah, J. Effendi, F. Firdaus, and B. Tutuko, "Electrocardiogram signal classification for automated delineation using bidirectional long short-term memory," *Informat. Med. Unlocked*, vol. 22, Jan. 2021, Art. no. 100507, doi: 10.1016/j.imu.2020.100507.

[15] A. Peimankar and S. Puthusserypady, "DENS-ECG: A deep learning approach for ECG signal delineation," 2020, arXiv:2005.08689.

[16] X. Liang, L. Li, Y. Liu, D. Chen, X. Wang, S. Hu, J. Wang, H. Zhang, C. Sun, and C. Liu, "ECG_SegNet: An ECG delineation model based on the encoder-decoder structure," *Comput. Biol. Med.*, vol. 145, Jun. 2022, Art. no. 105445, doi: 10.1016/j.compbiomed.2022.105445.

[17] A. I. Kalyakulina, I. I. Yusipov, V. A. Moskalenko, A. V. Nikolskiy, K. A. Kosonogov, G. V. Osipov, N. Yu. Zolotykh, and M. V. Ivanchenko, "LUDB: A new open-access validation tool for electrocardiogram delineation algorithms," *IEEE Access*, vol. 8, pp. 186181–186190, 2020, doi: 10.1109/ACCESS.2020.3029211.

[18] T. Huixin, "Heartbeat-aware convolutional neural network for R-peak detection of wearable device ECG data," *J. Southern Med. Univ.*, vol. 42, no. 3, pp. 375–383, 2022, doi: 10.12122/j.issn.1673-4254.2022.03.09.

[19] S. Ruder, "An overview of multi-task learning in deep neural networks," 2017, arXiv:1706.05098.

[20] R. Cipolla, Y. Gal, and A. Kendall, "Multi-task learning using uncertainty to weigh losses for scene geometry and semantics," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7482–7491, doi: 10.1109/CVPR.2018.00781.

[21] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 2, pp. 318–327, Feb. 2020, doi: 10.1109/tpami.2018.2858826.

[22] I. Sereda, S. Alekseev, A. Koneva, R. Kataev, and G. Osipov, "ECG segmentation by neural networks: Errors and correction," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2019, pp. 1–7, doi: 10.1109/IJCNN.2019.8852106.

[23] V. Moskalenko, N. Zolotykh, and G. Osipov, "Deep learning for ECG segmentation," 2020, arXiv:2001.04689.

[24] G. Chen, M. Chen, J. Zhang, L. Zhang, and C. Pang, "A crucial wave detection and delineation method for twelve-lead ECG signals," *IEEE Access*, vol. 8, pp. 10707–10717, 2020, doi: 10.1109/ACCESS.2020.2965334.

[25] F. Isensee, J. Petersen, A. Klein, D. Zimmerer, P. F. Jaeger, S. Kohl, J. Wasserthal, G. Koehler, T. Norajitra, S. Wirkert, and K. H. Maier-Hein, "nnU-Net: Self-adapting framework for U-Net-based medical image segmentation," 2018, arXiv:1809.10486.

ZUO WANG received the bachelor's degree in biomedical engineering from China Medical University. He is currently pursuing the master's degree in electronic information with the College of Biomedical Engineering, Southern Medical University, China. His research interests include ECG signal processing and deep learning.

JINLIANG WANG received the bachelor's degree in biomedical engineering from First Military Medical University (formerly), China, and the master's and Ph.D. degrees in biomedical engineering from Tsinghua University, China. He is currently the Co-Founder and the Executive Vice General Manager of CardioCloud Medical Technology (Beijing) Company Ltd., China. His research interests include wearable device development, intelligent analysis of medical signals, medical informatization, and internet medical.

MINGYANG CHEN received the bachelor's degree in software engineering from the Tianjin University of Technology. He is currently a Research and Development Engineer with CardioCloud Medical Technology (Beijing) Company Ltd., China. His research interests include intelligent analysis of medical signals, medical informatization, and internet healthcare.

WEI YANG received the bachelor's degree in industrial automation from the Wuhan University of Technology, China, the master's degree in control theory and control engineering from Xiamen University, China, and the Ph.D. degree in biomedical engineering from Shanghai Jiao Tong University, China. He is currently a Professor with the College of Biomedical Engineering, Southern Medical University, China. His research interests include medical image modality synthesis, intelligent analysis of medical images, intelligent analysis of medical signals, and radiomics. He is also a Council Member of the Guangdong Biomedical Engineering Society, China, the Standing Committee Member of the Medical Robotics and Artificial Intelligence Branch of the Guangdong Biomedical Engineering Society, China, and an Executive Member of the Medical Imaging Computing Seminar (MICS), China.

RONG FU received the first bachelor's degree in biomedical engineering and instrumentation from Southeast University, China, the second bachelor's degree in clinical medicine from Nanjing Medical University, China, and the Ph.D. degree in pathology from First Military Medical University, China. She is currently an Associate Professor with the College of Biomedical Engineering, Southern Medical University, China. Her research interests include color pathology image processing and smart medical intelligence product development. She also serves as a Board Member for the Guangdong Biophysical Association.

● ● ●