Identifying Electrocardiogram Abnormalities Using a Handcrafted-Rule-Enhanced Neural Network

Yuexin Bian†, Jintai Chen†, Xiaojun Chen, Xiaoxian Yang‡, Danny Z. Chen‡, and Jian Wu‡

Abstract—A large number of people suffer from life-threatening cardiac abnormalities, and electrocardiogram (ECG) analysis is beneficial to determining whether an individual is at risk of such abnormalities. Automatic ECG classification methods, especially the deep learning based ones, have been proposed to detect cardiac abnormalities using ECG records, showing good potential to improve clinical diagnosis and help early prevention of cardiovascular diseases. However, the predictions of the known neural networks still do not satisfactorily meet the needs of clinicians, and this phenomenon suggests that some information used in clinical diagnosis may not be well captured and utilized by these methods. In this paper, we introduce some rules into convolutional neural networks, which help present clinical knowledge to deep learning based ECG analysis, in order to improve automated ECG diagnosis performance. Specifically, we propose a Handcrafted-Rule-enhanced Neural Network (called HRNN) for ECG classification with standard 12-lead ECG input, which consists of a rule inference module and a deep learning module. Experiments on two large-scale public ECG datasets show that our new approach considerably outperforms existing state-of-the-art methods. Further, our proposed approach not only can improve the diagnosis performance, but also can assist in detecting mislabelled ECG samples.

Index Terms—ECG classification, deep learning, rule inference

1 INTRODUCTION

Electrocardiogram (ECG) is a type of commonly-used test in clinical practice for diagnosing patients suffered from cardiac abnormalities. Over 300 million ECG records are produced worldwide each year [1], which are a heavy burden for manual ECG diagnosis. For example, in China, around 250 million individuals take ECG tests each year, but only around 36,000 proficient doctors are engaged in analyzing such ECG data [2]. Thus, there is a clear gap between the supply and demand in clinical ECG diagnosis.

Recently, artificial intelligence (AI) methods have been revolutionizing various tasks [4], [5], [6], [7], [8], including diagnosis practices [9], [10], [11], [12], [13], [14] and provided effective assistance in automated signal analysis. Compared with doctors’ (manual) analysis of ECG, such methods can make cardiac abnormality diagnosis more efficient. Among them, many machine learning methods were proposed for automated ECG diagnosis. In early years, various traditional methods employing decision trees [15], SVM [16], random forests [17], and Bayesian networks [18] were applied to classify ECG signals, but did not yield satisfactory performances. Recently, deep learning approaches have drastically improved performances of various recognition tasks, including automatic ECG diagnosis [10], [13], [19], [20], [21], [22]. Deep learning methods for ECG can be roughly divided into three types, graph based [23], [24], [25], recurrent neural network (RNN) based [26], [27], [28], and convolution based [13], [14], [21], [29] methods. Specifically, graph based neural networks aim to capture the dependencies among cardiac abnormalities, since many ECG cases belong to multiple categories (abnormality types). RNN based methods treat ECG signals as time series and perform temporal feature extraction. Convolution based methods process ECG data as a special case of images, and often achieve better results [21].

In clinical practice, clinicians often analyze ECG records in two main aspects: (1) experienced doctors may analyze whether the shapes of some key segments are normal, including P waves, T waves, and QRS complexes; (2) clinicians quantitatively analyze voltages and duration of certain waves. However, the existing neural network based
Related work

Automated diagnosis has recently witnessed a rapid progress due to the fast development of deep neural networks (DNNs). In particular, many efforts have been dedicated to extending deep neural networks and designing DNN models for ECG classification. Hannun et al. [21] first developed a DNN to classify 12 arrhythmia classes and demonstrated that an end-to-end DNN can classify a broad range of distinct arrhythmias from single-lead ECGs with high diagnostic performance. Many improvements were based on it. Wang et al. [23] added graph attention networks to capture class dependencies. Luo et al. [22] combined bi-directional long short term memory (LSTM) with DNN to capture temporal features, dividing ECG data into 9 classes. However, these methods did not take into account whether a DNN correctly focuses on the key information in ECG signals to detect the corresponding classes, which might yield inferior performance in identifying some diseases.

On the other hand, many researchers attempted to utilize clinical knowledge in automated ECG analysis. Zhang et al. [30] proposed a disease-specific feature selection method to select ECG features and classify ECGs into five types. Xu et al. [31] used handcrafted rules in morphological classification for ST segments, and obtained more detailed and better results than the previous neural network methods [32], suggesting possible clinical significance. Jin et al. [33] combined a rule inference method and a convolutional neural network (CNN) to classify ECG data into normal and abnormal classes; they used some statistics (mean and variances) to depict the heart rate for diagnosis. Sannino et al. [34] detected peaks and waves to extract ECG temporal features, and leveraged a neural network to classify normal and abnormal ECG cases. In [35], Mondéjar et al. used support vector machine (SVM) to cope with temporal and morphological information (e.g., RR intervals of ECGs, wavelets, and several amplitude values). However, they simply constructed some rules or fed hand-crafted features to the models, resulting in sub-optimal performances. We argue that it would be more effective to combine rule-based features and neural networks. In addition, these methods were only presented to classify limited kinds of abnormalities (binary classification in most cases) and could not identify multiple diseases simultaneously, which did not meet the needs in clinical practice.

The major contributions of this work are as follows:

- We construct a rule-enhanced module to help promote neural networks, by providing rules according to diagnostic knowledge for ECG analysis. This design aims to provide clinical interpretation for arrhythmia with a higher consistency with experts’ attention on ECG.
- Our model surpasses current state-of-the-art methods on two large-scale public ECG classification datasets, verifying the effectiveness of the hand-crafted rules we use.
- In a case study, we show that our model is able to assist in detecting mislabeled samples, which is possibly beneficial to some practical tasks including corrupted label correction and AI-assisted annotation.

2 RELATED WORK

Automated diagnosis has recently witnessed a rapid progress due to the fast development of deep neural networks (DNNs). In particular, many efforts have been dedicated to extending deep neural networks and designing DNN models for ECG classification. Hannun et al. [21] first developed a DNN to classify 12 arrhythmia classes and demonstrated that an end-to-end DNN can classify a broad range of distinct arrhythmias from single-lead ECGs with high diagnostic performance. Many improvements were based on it. Wang et al. [23] added graph attention networks to capture class dependencies. Luo et al. [22] combined bi-directional long short term memory (LSTM) with DNN to capture temporal features, dividing ECG data into 9 classes. However, these methods did not take into account whether a DNN correctly focuses on the key information in ECG signals to detect the corresponding classes, which might yield inferior performance in identifying some diseases.

On the other hand, many researchers attempted to utilize clinical knowledge in automated ECG analysis. Zhang et al. [30] proposed a disease-specific feature selection method to select ECG features and classify ECGs into five types. Xu et al. [31] used handcrafted rules in morphological classification for ST segments, and obtained more detailed and better results than the previous neural network methods [32], suggesting possible clinical significance. Jin et al. [33] combined a rule inference method and a convolutional neural network (CNN) to classify ECG data into normal and abnormal classes; they used some statistics (mean and variances) to depict the heart rate for diagnosis. Sannino et al. [34] detected peaks and waves to extract ECG temporal features, and leveraged a neural network to classify normal and abnormal ECG cases. In [35], Mondéjar et al. used support vector machine (SVM) to cope with temporal and morphological information (e.g., RR intervals of ECGs, wavelets, and several amplitude values). However, they simply constructed some rules or fed hand-crafted features to the models, resulting in sub-optimal performances. We argue that it would be more effective to combine rule-based features and neural networks. In addition, these methods were only presented to classify limited kinds of abnormalities (binary classification in most cases) and could not identify multiple diseases simultaneously, which did not meet the needs in clinical practice.
Compared with the aforementioned methods, our proposed method does not simply feed ECG features to classifiers. In contrast, we construct different rules for different abnormalities based on clinical knowledge, and fuse the rule-based outputs and multi-label predictions of a deep learning model.

3 METHODOLOGY

In clinical practice, clinicians often analyze ECG signals by focusing on certain particular segments when determining whether specific abnormalities are observed. For instance, an ECG signal is determined to be a case of “low QRS voltage” if the amplitudes of all the QRS complexes in the limb leads are $<0.5$ mV, and the amplitudes of all the QRS complexes in the precordial leads are $<1.0$ mV. Deep neural networks have been shown to have advantages in capturing shape patterns, leading to breakthroughs in image recognition. Also, motivated by clinical practice, we think rule-based methods can well complement deep neural networks for ECG analysis.

Based on the above motivation, we design a new hand-crafted-rule-enhanced neural network (HRNN) for identifying ECG abnormalities (see Fig. 2 for an overview of HRNN). We seek to leverage the power of a convolutional neural network (CNN) for automatic feature extraction, while employing a rule-based module to provide some key information that has been proven to be useful in clinical ECG diagnosis. To better combine these two components, we construct a super learner to fuse what the neural network learns and what the rule inference module provides.

In what follows, we describe the deep learning module in Section 3.1, the rule inference module in Section 3.2, and the proposed super learner in Section 3.3. The training strategy of HRNN is discussed in Section 3.4.

3.1 Automatic Feature Extraction by a Deep Neural Network

As previous research has demonstrated that CNNs can well process ECG signals [13], [21], we develop an end-to-end CNN to cope with ECG signals for arrhythmia classification.

An overview of HRNN is shown in Fig. 2, in which the architecture of the CNN part (denoted by $f_{\text{CNN}}$) is specified in Fig. 3. This CNN architecture is modified from ResNet-34.
Specifically, the input ECG \( l_d R \) is 2100 (ms) for replacement by other architectures. I, II, III, V1-V6, aVL, aVR, in the A means \( f \) is the in the lead \( l \). The batch normalization layers and rectified linear unit (ReLU) layers are placed as in the original ResNet-34, which are not shown in Fig. 3.

The input of deep learning module includes three items: the raw ECG signal, the patient’s age (encoded as a one-hot vector of size 10), and the patient’s gender (a scalar). The deep learning module outputs a probability vector for possible abnormality classes. Specifically, the input ECG signal is specified as \( x \in \mathbb{R}^{w \times d} \), where \( w \) and \( d \) are the length of the signal and the number of leads, respectively. The output of \( f_{\text{CNN}} \) is \( z \in \mathbb{R}^{w \times d} \), which is further fed to a global average pooling layer for global semantics abstraction. The global average pooling is utilized to compress \( z \in \mathbb{R}^{w \times d} \) into \( z' \in \mathbb{R}^{d} \). Because the patient’s age and gender also affect the diagnosis results (according to clinicians’ viewpoint), we concatenate the age feature vector and gender feature with \( z' \) and make the prediction (denoted by a vector \( h \in \mathbb{R}^{10} \)) for abnormality categories via a fully-connected layer with a sigmoid function (see Fig. 2). We define the scalar feature of gender as: (i) if the gender is missing, it is converted to 0; (2) if the gender is “male”, it is 1; (3) if the gender is “female”, it is 2. For the feature vector of age, we assume that the patient’s age is smaller than 100, and encode the numerical age as a 10-dimensional vector. If the age is missing, it is converted to 0; (2) if the gender is “male”, it is 1; (3) if the gender is “female”, it is 2. For the feature vector of age, we assume that the patient’s age is smaller than 100, and encode the numerical age as a 10-dimensional one-hot vector, \( \text{len} \text{age} \), where \( \text{len} \text{age} \) is in the range of \( [0, 100] \), and encode the numerical age as a 10-dimensional one-hot vector, \( \text{len} \text{age} \).

In Table 1, we compute \( t(s) \) as in Eq. (1), where \( \text{len}(s) \) is the number of the points recorded in the ECG signal, and \( \text{sr}(s) \) is the sampling rate of the signal:

\[
t(s) = \frac{\text{len}(s)}{\text{sr}(s)} \times 1000 \text{(ms)}
\]

We compute \( A(s|l) \) as in Eq. (2), in which \( v(l) \) means the baseline voltage of \( s \) in the lead \( l \), which is presented as a vector of the same length as \( s \). \( t(s) \) and \( v(l) \) are the sampling rate of the signal:

Fig. 4 gives an example for visualizing the classical display format of ECG signals. Some abnormalities can be detected by measurements on a single ECG recording, while other abnormalities become apparent only by observing several leads and such diagnosis can be relatively complicated. As discussed in Section 1, neural networks might ignore some clinically useful information, and thus it is our desire to introduce the assistance of rule-based methods.

In this design, we propose a rule inference module, which performs some handcrafted rules on ECG signals to introduce certain clinical information. This rule inference module consists of a “Segmentation and Delineation” part and a “Rule Inference” part. Our “Segmentation and Delineation” part processes the ECG signals to obtain the cardiac cycles and key segments of ECG (see Fig. 5). First, we apply a band-pass (3-50Hz) filter to filter the ECG signal, so as to deburr the signal and remove the signal offsets. Second, we process R peak detection and ECG segmentation. There are methods available in BioSPPy [38] to process the R peak detection and ECG segmentation. In this paper, we follow the work in [39], and utilize the first-order information and second-order information of the filtered ECG signal to delineate the P wave, QRS complex, and T wave. In experiments, we find that the “Segmentation and Delineation” part works well, and consequently, the exact cardiac cycles and key segments are obtained (see Fig. 6).

After obtaining the cardiac cycles and key segments, we formulate rules for 15 ECG abnormalities according to criteria in the previous work [37], [40], [41]. The rules are somewhat simplified, but still maintain a high degree of consistency with experts’ attention on ECGs. The cardiac abnormalities measured by these rules and the corresponding rule formulas are reported in Table 1, in which \( t(s) \) means the duration of a segment \( s \), \( A(s|l) \) means the amplitude of a segment \( s \) in the lead \( l \), \( l \in \{I, II, III, V1-V6, aVL, aVR, aVF\} \), \( v(s|l) \) indicates the voltage value of a segment \( s \) in the lead \( l \), which is presented as a vector of the same length as \( s \). \( t(s) \) and \( v(l) \) are the sampling rate of the signal.

After obtaining the cardiac cycles and key segments, we formulate rules for 15 ECG abnormalities according to criteria in the previous work [37], [40], [41]. The rules are somewhat simplified, but still maintain a high degree of consistency with experts’ attention on ECGs. The cardiac abnormalities measured by these rules and the corresponding rule formulas are reported in Table 1, in which \( t(s) \) means the duration of a segment \( s \), \( A(s|l) \) means the amplitude of a segment \( s \) in the lead \( l \), \( l \in \{I, II, III, V1-V6, aVL, aVR, aVF\} \), \( v(s|l) \) indicates the voltage value of a segment \( s \) in the lead \( l \), which is presented as a vector of the same length as \( s \). \( t(s) \) and \( v(l) \) are the sampling rate of the signal.

1. This example is borrowed from [37].
regard the 50th percentile of the voltage value of a segment from Toff to Pon as the baseline amplitude. For simplicity, $v(l)$ could be computed as 0.

$$A(s(l)) = \max((v(s(l)) - v(l)) \times u(mV)),$$

if $s$ is an "upper arch" - shaped segment

$$= \min((v(s(l)) - v(l)) \times u(mV)),$$

if $s$ is a "down bend" - shaped segment

where $u$ is the unit voltage defined in the dataset. One can see that the rules define some abnormalities in a precise quantification way. In the training and inference phases, the rule inference module is performed on ECG signals without back-propagation, and yields predictions for various abnormalities. If an ECG signal meets a rule formula, then the rule inference module returns "1" for this abnormality category, and if not, the rule inference module returns "0". The predictions for all the abnormalities are concatenated into a vector, and are forwarded for final prediction.

### 3.3 A Super Learner for Prediction Fusion

To combine the predictions provided by the rule inference module and the deep learning module, it is desirable to model the dependencies between these two methods and fuse their outputs for the final cardiac abnormality identification. Here we treat both the rule inference module and the deep neural network as meta learners, and introduce a super learner to fuse their predictions.

The concept of "super learner" was first proposed in [42], which is a weighted combination of the predictions of the meta learners. In our design, the input to the super learner is two prediction vectors produced by the deep learning module, $h_{DL} = [h_1, h_2, \ldots, h_N]$, and by the rule inference module, $l_{RULE} = [l_1, l_2, \ldots, l_N]$, where $N$ is the number of abnormalities, and the value of the $i$th element indicates the predicted probability of the $i$th abnormality (or normality). By fusing these two predictions, the proposed Super Learner produces the final prediction vector, $\hat{y} = [\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_N]$. Formally, the operation of the Super Learner is defined in Eq. (3), where $w$ is a learnable weight vector of size $N$, $s(w)$ is the sigmoid function, and "." denotes the element-wise dot multiplication:

$$\hat{y} = h_{DL} \cdot s(w) + l_{RULE} \cdot (1 - s(w))$$

In particular, if the rule inference module considers fewer than $N$ kinds of abnormalities, a mask vector is constructed as $mask = [m^l]_l^N$ ($m^l \in \{0, 1\}$). If the $i$th abnormality (or normality) is predicted by the rule inference module, then the $i$th element $m^l = 1$; otherwise, $m^l = 0$. The learnable weight vector $w$ is of the same length as the prediction vector produced by the rule inference module, and its elements corresponding to the predictions that are not provided by the rule inference module are padded with zeros (requiring no gradients) to align with $h_{DL}$. The prediction vector $l_{RULE}$ is padded in the same way. Thus, $w$, $l_{RULE}$, and $h_{DL}$ all have the identical size of $N$. The masked mechanism for the final prediction is performed as in Eq. (4), where $mask$ indicates to invert every element in the $mask$ vector, and "." denotes the element-wise multiplication:

$$\hat{y} = mask \cdot (h_{DL} \cdot s(w) + l_{RULE} \cdot (1 - s(w))) + mask \cdot h_{DL}$$

### TABLE 1

| ECG Abnormalities                   | Rule Formulas                                                                 |
|-------------------------------------|-----------------------------------------------------------------------------|
| Poor R-wave progression$^1$        | (a) $A(R\mid V1) > A(R\mid V2) > A(R\mid V3) > A(R\mid V4)$                  |
|                                    | (b) $A(R\mid V2) > 0 \& A(R\mid V3) > 0 \& A(R\mid V1) + A(R\mid V2) + A(R\mid V3) < 0.2mV$ |
| Arrhythmia                          | STD($t(PP)$) > 120ms$^2$                                                   |
| Tachycardia                         | heart rate > 120                                                           |
| Bradycardia                         | heart rate < 60                                                            |
| Right axis deviation (RAD)$^3$      | $-2 \times A(QRS\mid III) < A(QRS\mid I) < 0 \& A(QRS\mid III) > 0$          |
| Left axis deviation (LAD)           | $A(QRS\mid I) > 0 \& A(QRS\mid III) < A(QRS\mid I)$                        |
| Low QRS voltage                     | $A(QRS\mid l) < 0.5mV, l \in \{I, II, III\}$ or $A(QRS\mid l) < 1mV, l \in \{V1, V2, V3\}$ |
| QT prolongation                     | $t(QT) > 0.4s \& t(QT)/(heart rate)^{1/2} > 0.43$                           |
| Clockwise rotation (CR)             | 0.9 < $A(R\mid V1)/A(S\mid V1) < 1.1 \& 0.9 < A(R\mid V2)/A(S\mid V2) < 1.1$ |
| Counterclockwise rotation (CCR)     | $A(R\mid l)/A(S\mid l) < 1, l \in \{V1, V4\}$                             |
| First degree atrioventricular block | $t(PR) > 200ms$                                                            |
| Abnormal Q waves                    | $A(Q\mid l) > 1/4 \times A(R\mid l), l \in \{II, III, aVF\}$ or $t(Q) > 40ms$ |
| T wave change (T change)            | $A(T\mid l) < 1/10 \times A(R\mid l)$ or $A(T\mid l) > 0.5mV, l \in \{I, II, V2-V6\}$ |
| Right atrium enlargement (RAB)      | $A(P\mid l) > 0.15mV, l \in \{V1, V2\}$ or $A(P\mid l) > 0.25mV, l \in \{II, III, aVP\}$ |
| Left ventricular high voltage (LVHV)$^4$ | \begin{align*}
(a) & A(R\mid V5) > 2.5mV \& A(R\mid V6) > 2.5mV \\
(b) & A(R\mid V5) + A(S\mid V1) > 4.0mV \text{ if } (gender = male) or > 3.5mV \text{ if } (gender = female) \\
(c) & A(R\mid l) > 1.5mV or A(R\mid s VL) > 1.2mV or A(R\mid a VF) > 2.0mV \\
(d) & A(R\mid l) + A(S\mid III) > 2.5mV
\end{align*}

$^1$($a$ or $b$) is sufficient to detect the Poor R-wave progression.

$^2$STD denotes standard deviation computing.

$^3$A(QRS(I)) = A(QR) + A(R(I)) + A(S(I))$

$^4$(a) or (b) or (c) or (d) is sufficient to detect the left ventricular high voltage.
3.4 Model Training

The entire framework HRNN is specified using a newly proposed loss function $\mathcal{L}$ in the training phase. Since abnormal ECG cases are usually much fewer in the real world data (the datasets), the deep neural network part might fall into a pattern collapse and indiscriminately return “zeros” (indicating the “normality” category). Thus, in designing the loss function, we manage to use the predictions generated by the rule inference module to guide the predictions of the deep learning module, because the rule inference module often performs better on the categories with fewer cases in the datasets.

Assume that the ground truth of an ECG record is $y = \{y_i\}_{i=1}^N$, $y_i \in \{0, 1\}$. When $y_i = 1$, it means that the ECG record is with the abnormality of category $i$, and $y_i = 0$ is for normality of category $i$. The weighted cross-entropy loss is defined by:

$$L(y, \hat{y}) = \sum_{i=1}^{N} w_i \cdot y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i)$$  \hspace{1cm} (5)$$

where $w_i$ is the class weight of category $i$ computed as $w_i = \frac{M_i}{M}$ ($M$ is the total number of all ECG cases in the dataset and $M_i$ is the number of cases belonging to category $i$), and $\hat{y}$ is the prediction probabilities obtained by the Super Learner. The predictions of the rule inference module are also used to guide the prediction of HRNN, and its total loss function is defined by:

$$\mathcal{L} = L(y, \hat{y}) + \lambda L(l_{\text{RULE}}, \hat{y})$$  \hspace{1cm} (6)$$

where $L$ is the weighted binary cross-entropy in Eq. (5), $l_{\text{RULE}}$ is the output probabilities generated by the rule inference module, and $\lambda$ is an importance parameter.

4 EXPERIMENTS

4.1 Datasets

We use the datasets for the first and second rounds of the contest of the Hefei Hi-tech Cup ECG Intelligence Competition\(^2\) for a multi-label classification task of 55 classes and 34 classes, respectively. In clinical ECG diagnosis, clinicians often give detailed analysis with the pathogenesis and types of abnormalities. The datasets we use cover comprehensive ECG disease classes that are commonly found in clinical application scenarios. The dataset used in the first round of the contest (called “TianChi ECG dataset-1”) contains 24,106 samples, and the second round dataset (called “TianChi ECG dataset-2”) contains 20,096 samples. Each sample has 8 leads (I, II, V1, V2, V3, V4, V5, V6). Each sample was recorded in 10 seconds with 500 Hz sampling frequency, and the unit voltage is $4.88 \times 10^{-3}$ millivolts. For these two datasets, the ECG input for $f_{\text{CNN}}$ is with a shape of $5000 \times 12$. Since the standard 12-lead ECG is the most commonly-used format in ECG analyses [43], we computationally added the other four leads to the original datasets following Eq. (7). An example of the input ECG is shown in Fig. 7.

The 55 cardiac abnormalities (or normalities) of TianChi ECG dataset-1 and the 34 cardiac abnormalities (or normalities) of TianChi ECG dataset-2 are shown in Table 2.

4.2 Experimental Setting

We use the Adam optimizer [44] with default hyper-parameters. The base learning rate (base_lr) was initialized as $10^{-4}$, and we gradually warm-up [45] the first 2 epochs. The learning rate scheduler is set as the cosine scheduler with a weight decay of $10^{-6}$. The batch size is 32, and the number of epochs is 60.

We conduct a comprehensive evaluation, and compare our approach with four state-of-the-art ECG signal classification models, including 1D Transformer-XL [7], SE-ECGNet [46], 1D ResNet-34 [47], and MLWGAT [23], on the two TianChi ECG datasets. We adopt the same data pre-processing method and configurations for all these methods.

We develop all the models using PyTorch, and run all the experiments on an NVIDIA GTX 2080Ti 64GB GPU machine.

4.3 Evaluation Metrics

Following conventional settings [21], [23] and taking patients’ concern into account, we report the average per-class recall (CR), average per-class F1 (CF1), average overall recall (OR), and average overall F1 (OF1) for performance evaluation. For each ECG sample, the labels are predicted as positive if the probabilities for them are larger than 0.5. Generally, the average overall recall (OR) and average per-class recall (CR) are relatively more important for ECG abnormality detection in the clinical setting since neglecting a disease is much more harmful for the patient.
4.4 Experimental Results

We present comparison results with the state-of-the-art methods. In addition, we provide some running examples for showing that our model can assist detecting mislabelled samples.

4.4.1 Comparison With State-of-the-art Methods

The results of our comparative evaluation experiments are summarized in Tables 3 and 4. We report the OF1, CF1, OR, and CR of the four known models evaluated on the two TianChi ECG datasets.

In Tables 3 and 4, we highlight the best result for each metric in bold. One can observe from the results that our model achieves the best performance on CF1, OR, and CR. Meanwhile, HRNN achieves a 50.7% average per-class F1 (CF1) score and outperforms the state-of-the-art performance by 1.2%. Somehow, with the highest OF1 score and overall recall score among all the five methods. With $\lambda = 1$, HRNN achieves a 95.2% overall recall score and a 63.2% average per-class recall score, outperforming the state-of-the-art performance by over 4% and 13%, respectively. Meanwhile, HRNN achieves a 50.7% average per-class F1 (CF1) score. For TianChi ECG dataset-2, with $\lambda = 0$, HRNN achieves the highest OF1 score and overall recall score among all the five methods. With $\lambda = 1$, HRNN achieves a 95.2% overall recall score and a 63.2% average per-class recall score, outperforming the state-of-the-art performance by over 5% and 16%, respectively. Meanwhile, HRNN achieves a 50.7% average per-class F1 (CF1) score and outperforms the state-of-the-art performance by 1.2%. Somehow, with $\lambda = 1$, our overall F1 (OF1) scores are lower than the four models. This is probably mainly due to the presence of more false-positive (FP) samples. False positive is a result indicating that a given condition occurs when it actually does not. If we detect FP samples, it means that our model outputs 1 for a specific disease but the ground truth of this disease is 0. However, we will show below that our identification of some of the FP samples, it means that our model outputs 1 for a specific disease but the ground truth of this disease is 0.

TABLE 2

The Record Number and Proportion of Each Class in the Two Tianchi ECG Datasets

| Categories                     | # of cases (proportion %) | ECG dataset-1 | ECG dataset-2 |
|--------------------------------|---------------------------|---------------|---------------|
| Sinus rhythm                   | 16918 (70.18%)            | 9536 (47.45%) |
| Bradycardia                    | 3372 (13.99%)             | 5272 (26.23%) |
| Tachycardia                    | 2100 (8.34%)              | 4910 (24.43%) |
| T change                       | 3241 (14.19%)             | 3490 (17.37%) |
| RAD                            | 1005 (4.38%)              | 1131 (5.628)  |
| LAD                            | 1137 (4.72%)              | 1128 (5.613)  |
| Arrhythmia                     | 924 (3.83%)               | 904 (4.498%)  |
| VBP                            | 971 (4.03%)               | 571 (2.841%)  |
| RBBB                           | 392 (1.63%)               | 556 (2.767%)  |
| CRBKB                          | 1109 (4.60%)              | 423 (2.105%)  |
| LVHV                           | 4326 (17.95%)             | 415 (2.062%)  |
| APB                            | 1470 (6.10%)              | 316 (1.572%)  |
| ST-T change                    | 2111 (8.76%)              | 299 (1.488%)  |
| ST change                      | 2962 (12.31%)             | 287 (1.43%)   |
| IAVB                           | 282 (1.17%)               | 142 (0.707%)  |
| IRBBB                          | 199 (0.83%)               | 126 (0.627%)  |
| Atrial fibrillation            | 1217 (5.05%)              | 120 (0.597%)  |
| NS-ST                          | 78 (0.32%)                | 64 (0.318%)   |
| CCR                            | 34 (0.14%)                | 61 (0.304%)   |
| Abnormal Q waves               | 53 (0.22%)                | 52 (0.259%)   |
| Labb                           | 106 (0.44%)               | 36 (0.179%)   |
| NS-T                           | 125 (0.52%)               | 35 (0.174%)   |
| CR                             | 30 (0.12%)                | 35 (0.174%)   |
| RAE                            | 24 (0.10%)                | 32 (0.159%)   |
| RVR                            | 229 (0.95)                | 29 (0.144%)   |
| LBBB                           | 18 (0.08%)                | 27 (0.134%)   |
| CLBBD                          | 20 (0.08%)                | 27 (0.134%)   |
| Short PR interval              | 55 (0.23%)                | 24 (0.119%)   |
| Early repolarization           | 37 (0.15%)                | 22 (0.109%)   |
| LAD                            | 74 (0.31%)                | 16 (0.088%)   |
| Poor R-wave progression        | 19 (0.08%)                | 16 (0.088%)   |
| NS-SSST                        | 61 (0.25%)                | 16 (0.088%)   |
| Fusion wave                    | 18 (0.08%)                | 8 (0.04%)     |
| QRS low voltage                | 1543 (6.40%)              | 3 (0.015%)    |
| Normal ECG                     | 4171 (17.30%)             | -             |
| Critical ECG                   | 1911 (7.92%)              | -             |
| Abnormal ECG                   | 1061 (4.40%)              | -             |
| LHV                            | 432 (1.79%)               | -             |
| QT prolongation                | 101 (0.42%)               | -             |
| Differential conduction        | 75 (0.31%)                | -             |
| Atrial fibrillation            | 42 (0.17%)                | -             |
| Intraventricular aberrant conduction | 36 (0.15%)            | -             |
| Bigeminy                       | 27 (0.11%)                | -             |
| Ventricular premature beats    | 33 (0.14%)                | -             |
| Abnormal repolarization        | 29 (0.12%)                | -             |
| uAPB                           | 25 (0.10%)                | -             |
| Cor Pulmonale                  | 27 (0.11%)                | -             |
| SVR                            | 30 (0.12%)                | -             |
| Short series of atrial tachycardia | 21 (0.09%)            | -             |
| RVE                            | 18 (0.08%)                | -             |
| Atrioventricular conduction delay | 10 (0.04%)            | -             |
| Bifascicular block             | 20 (0.08%)                | -             |
| NS-IB                          | 17 (0.07%)                | -             |
| NS-ID                          | 16 (0.07%)                | -             |
| P pulmonale                    | 17 (0.07%)                | -             |

Some of the categories use abbreviations. The comparison table of full names and abbreviations is given in Appendix.

TABLE 3

Comparison With the Four State-of-the-art Methods on TianChi ECG Dataset-1

| Method              | OF1    | CF1    | OR     | CR     |
|---------------------|--------|--------|--------|--------|
| 1D Transformer [7]  | 0.7731 | 0.2814 | 0.7185 | 0.2554 |
| 1D ResNet-34 [47]   | 0.8679 | 0.4988 | 0.8632 | 0.5030 |
| SE-ECGNet [46]      | 0.8651 | 0.5024 | 0.8545 | 0.4680 |
| MLWGAT[23]          | 0.8401 | 0.4259 | 0.8267 | 0.3959 |
| HRNN1               | 0.8691 | 0.4855 | 0.8653 | 0.4852 |
| HRNN2               | 0.6312 | 0.5071 | 0.9065 | 0.6318 |

1 With $\lambda$ in Eq. (6) set to 0.  
2 With $\lambda$ in Eq. (6) set to 1.

TABLE 4

Comparison With the Four State-of-the-art Methods on TianChi ECG Dataset-2

| Method              | OF1    | CF1    | OR     | CR     |
|---------------------|--------|--------|--------|--------|
| 1D Transformer [7]  | 0.8950 | 0.3123 | 0.8601 | 0.2773 |
| 1D ResNet-34 [47]   | 0.9038 | 0.4686 | 0.8872 | 0.4598 |
| SE-ECGNet [46]      | 0.9019 | 0.4780 | 0.8940 | 0.4620 |
| MLWGAT[23]          | 0.8401 | 0.4259 | 0.8267 | 0.3959 |
| HRNN1               | 0.9104 | 0.4655 | 0.9018 | 0.4619 |
| HRNN2               | 0.7224 | 0.5001 | 0.9520 | 0.6402 |

1 With $\lambda$ in Eq. (6) set to 0.  
2 With $\lambda$ in Eq. (6) set to 1.
better ability to perform multi-label classification, thus this decrease in OF1 score could be acceptable.

We also display the Recall Score (sensitivity) for each class of Tianchi ECG dataset-2 in Fig. 8. HRNN yields relatively high sensitivity in identifying different ECG abnormalities compared to the two state-of-the-art methods. In addition, we can see that there are some categories which are not included in our rule inference module, and their sensitivities still get improved (e.g., AF, NS-ST, AQ, ST-change, etc.). It might be because our model gives confidence in the rule inference module, and then back-propagation pushes the deep learning module to pay more attention to the performance of the other categories. This evidence shows that our fusion of the rule inference and deep learning network outputs is effective.

Notably, we observe that 1D ResNet-34 can also attain competitive performances on TianChi ECG Dataset-1 (see Table 3), but as Fig. 1 shows, the model cannot focus on the segment associated with the corresponding ECG abnormality. This might be because the deep learning model is effective for the ECG cases with obvious abnormalities, but for those abnormalities that are inconspicuous (e.g., those abnormalities that can only be observed with partial voltages and intervals), the deep learning model fails to capture accurate information. This phenomenon suggests that incorporating clinical knowledge into deep learning may provide a good potential to obtain performance improvements.

### 4.4.2 Ablation Studies

We conduct experiments to verify the effect of introducing age and gender information to the neural network. We also conduct experiments to verify the contributions of the rule inference module, where we remove the meta-learner of the "Rule Inference" on the basis of HRNN and find out how the model performs. Tables 5 and 6 show that the information of age and gender benefits our model, which is consistent with the previous medical observations. For instance, human's heart rate decreases with age, and the left atrial hypertrophy is also observed to be related to gender. The results in Tables 5 and 6 also show that our rule inference module facilitates ECG abnormality identification with a clear margin (shown by the indicators OR and CR scores).

In addition, we observe that equipping the rule inference module is almost free of cost for inference. Hence, it is promising to equip other neural networks with the rule inference module, which can improve performances at low cost.

#### 4.4.3 Detection of Mislabelled Samples

Our rule inference module is able to detect mislabelled samples, as reviewed and verified by senior certified ECG clinicians. In the following examples, dash lines delineate P wave in the ECG signals, dot lines for QRS complex, and dash-dot lines for T wave. We also add blue to the signals to illustrate the sections focused by senior certified ECG clinicians.

- Low QRS voltage: In Fig. 9, we can see that QRS amplitude is less than 5mV in limb leads, which meets the requirement according to Table 1.
- Right atrium enlargement (RAE): In Fig. 10, Right atrium enlargement produces a peaked P wave (P pulmonale) with amplitude > 0.25mV in the inferior leads (II, III, and AVF). However, the sample in Fig. 10 does not have a peaked P wave.
- T wave change (T change): In Fig. 11, the amplitude of the T wave is less than 1/10 of the R peak in the wave in the V1 lead.
- Right axis deviation (RAD): The sample shown in Fig. 12 meets the requirement of the corresponding rule in Table 1.
- Left axis deviation (LAD): The sample shown in Fig. 13 meets the requirement of the rule for Left axis deviation in Table 1.

### TABLE 5

| Method          | OF1  | CF1  | OR   | CR   |
|-----------------|------|------|------|------|
| HRNN.dl1 (original) | 0.8661 | 0.4973 | 0.8579 | 0.4853 |
| HRNN.dl2        | 0.8741 | 0.5062 | 0.8743 | 0.4854 |
| HRNN (λ = 0)    | 0.8691 | 0.4855 | 0.8653 | 0.4852 |
| HRNN (λ = 1)    | 0.6312 | 0.5071 | 0.9065 | 0.6318 |

1Only deep learning module in HRNN without coding information of gender and age.
2Only deep learning module in HRNN with coding information of gender and age.

### TABLE 6

| Method          | OF1  | CF1  | OR   | CR   |
|-----------------|------|------|------|------|
| HRNN.dl1 (original) | 0.9033 | 0.4667 | 0.8890 | 0.4581 |
| HRNN.dl2        | 0.9132 | 0.4718 | 0.8981 | 0.4537 |
| HRNN (λ = 0)    | 0.9104 | 0.4655 | 0.9018 | 0.4619 |
| HRNN (λ = 1)    | 0.7224 | 0.5001 | 0.9520 | 0.6402 |

1Only deep learning module in HRNN without coding information of gender and age.
2Only deep learning module in HRNN with coding information of gender and age.
C15

Counterclockwise rotation (CCR): In Fig. 14, the R wave of the V5 or V6 lead appears on the V2, V3, and V4 leads; this sample could be diagnosed as CCR.

5 CONCLUSION AND FUTURE WORK

In this paper, we proposed a new Handcrafted-Rule-enhanced Neural Network, HRNN, to classify arrhythmia diseases using ECG signals. To our best knowledge, this is the first approach that combines handcrafted rule methods and neural networks for multi-label ECG classification. Experiments showed that HRNN achieves the highest overall recall score and the highest average per-class recall score on both the two TianChi ECG datasets, and outperforms state-of-the-art methods with clear margins (e.g., over 4% in the overall recall score and 13% in the average per-class recall score). Meanwhile, our model attains the highest average per-class F1 score among state-of-the-art methods. Experiments also demonstrated that HRNN is able to help identify mislabelled samples, showing good potential for some practical tasks including corrupted label correction and AI-assisted annotation.

There are several possible improvements and extensions to HRNN that we would like to pursue in future work: (1) encoding rules into neural networks and training rule modules with back-propagation; (2) leveraging graph neural networks to model the dependencies among different abnormalities and utilizing such dependencies in automatic ECG diagnosis.

REFERENCES

[1] H. Holst, M. Ohlsson, C. Peterson, and L. Edenbrandt, “A confident decision support system for interpreting electrocardiograms,” Clin. Physiol., vol. 19, pp. 410–418, 1999.

[2] F. Pan, “Yuancheng xindian jiance manzu duoyanghua jiankang yiliao fuwu xuqi-fang Zhongguo yuancheng xindianzhenduan xuezu fuzhuren Wangxinkang jiaoshou [Remote ECG monitoring meets the needs of diversified health and medical services-interview with Professor Wang Xinkang, the vice chairman of the Chinese Remote ECG Diagnostics Group],” China Medical Herald, 2017.
BIAN ET AL.: IDENTIFYING ELECTROCARDIOGRAM ABNORMALITIES USING A HANDCRAFTED-RULE-ENHANCED NEURAL NETWORK 2443

[3] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual explanations from deep networks via gradient-based localization,” in Proc. Int. Conf. Comput. Vis., 2017, pp. 618–626.

[4] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 770–778.

[5] K. Wei, M. Yang, H. Wang, C. Deng, and X. Liu, “Adversarial fine-grained composition learning for unseen attribute-object recognition,” in Proc. Int. Conf. Comput. Vis., 2019, pp. 3740–3748.

[6] K. Wei, C. Deng, and X. Yang, “Lifelong zero-shot learning,” in Proc. Int. Joint Conf. Artif. Intell., 2020, pp. 351–357.

[7] Z. Dai, Z. Yang, Y. Yang, J. Carbonell, Q. V. Le, and R. Salakhutdinov, “Attentional neural models beyond a fixed-length context,” 2019, arXiv:1901.02860.

[8] K. Wei, C. Deng, X. Yang, and D. Tao, “Incremental zero-shot learning,” IEEE Trans. Cybern., early access, Sep. 30, 2021, doi: 10.1109/TCYB.2021.3110369.

[9] S. Kiranyaz, T. Ince, and M. Gabbouj, “Real-time patient-specific ECG classification by 1-D convolutional neural networks,” IEEE Trans. Biomed. Eng., vol. 65, no. 3, pp. 1156–1165, 2018.

[10] B. Pyakylai, N. Kazachenko, and N. Mikhailovsky, “Deep learning for ECG classification,” in Proc. J. Phys.: Conf. Ser., IOP Publishing, 2017, doi: https://doi.org/10.1088/1742-6596/913/1/012004.

[11] S. Saadatnejad, M. Oveis, and M. Hashemi, “LSTM-based ECG classification for continuous monitoring on personal wearable devices,” IEEE J. Biomed. Health Inform., vol. 24, no. 2, pp. 515–523, Feb. 2020.

[12] J. Chen, X. Zheng, J. Yu, D. Z. Chen, and J. Wu, “Flow-Mixup: A trans-formational LSTM network model for ECG signal classification,” in Proc. IEEE Int. Conf. Bioinf. Biomed., 2021, pp. 534–541.

[13] J. Chen, H. Yu, R. Feng, D. Z. Chen, and J. W. GG, “Flow-Mixup: Classifying multi-labeled medical images with corrupted labels,” in Proc. IEEE Int. Conf. Bioinf. Biomed., 2020, pp. 917–924.

[14] J. Chen, X. Zheng, H. Yu, D. Z. Chen, and J. Wu, “Electrocardio-gram Panoramic Viewing: Long new ECG views with self-supervision,” 2021, arXiv:2105.06293.

[15] J. Rodriguez, A. Goni, and A. Illarramendi, “Real-time classifica-tion of ECGs on a PDA,” IEEE Trans. Inf. Technol. Biomed., vol. 9, no. 1, pp. 23–34, Mar. 2005.

[16] S. Celin and K. Vasanth, “ECG signal classification using vari-ational LSTM network model for ECG signal classification,” Proc. IEEE Int. Conf. Bioinf. Biomed., vol. 46, no. 4, pp. 1232–1240, Apr. 2019.

[17] D. Batra, “Grad-CAM: Visual explanations from deep networks and explainability technique for 12-lead electrocardiogram interpretation,” in J. Amer. Medical Assoc. Cardiol., vol. 25, pp. 65–69, 2019.

[18] Q. Yao, R. Wang, X. Fan, and J. Liu, “Multi-class rhythm abnormalities from 12-lead varied-length ECG using attention-based time-1incremental convolutional neural network,” Inf. Fusion, vol. 53, pp. 174–182, 2020.

[19] Z. Zhang, J. Dong, X. Luo, K.-S. Choi, and X. Wu, “Heartbeat classifica-tion using disease-specific feature selection,” Comput. Biol. Med., vol. 46, pp. 79–89, 2016.

[20] S. M. Mathews, C. Kambhamettu, and K. E. Barner, “A novel LMS algorithm for ECG signal preprocessing and KNN classifier,” IEEE Trans. Inf. Technol. Biomed., vol. 10, no. 1, pp. 23–34, Mar. 2005.

[21] B. Hou, J. Yang, P. Wang, and R. Yan, “LSTM-based auto-encoder model for ECG arrhythmias classification,” IEEE Trans. Instrum. Meas., vol. 69, no. 4, pp. 1232–1240, Apr. 2019.

[22] J. Gao, H. Zhang, P. Lu, and Z. Wang, “An effective LSTM recur-vent network to detect arrhythmia on imbalanced ECG dataset,” J. Biomed. Inf., vol. 75, 2017.

[23] Q. Yao, R. Wang, X. Fan, and J. Liu, “Multi-class arrhythmia detection from 12-lead varied-length ECG using attention-based time-incremental convolutional neural network,” in Proc. IEEE Int. Conf. Bioinf. Biomed., 2020, pp. 1365–1374.

[24] S. Dalu and N. Pawar, “Detection and classification of QRS and ST segment using WNN,” J. Inst. Comput. Sci. Netw., vol. 1, no. 3, 2012.

[25] L. Jin and J. Dong, “Classification of normal and abnormal ECG recordings using a lead convolutional neural network and rule inference,” Sci. China Inf. Sci., vol. 60, 2017, doi: 10.1007/s11432-016-9047-6.

[26] S. G. Sannino and G. De Pietro, “A deep learning approach for ECG-based heartbeat classification for arrhythmia detection,” Future Gen. Comput. Syst., vol. 86, pp. 446–455, 2018.

[27] V. Mond ejar-Guerra, J. Novo, J. Rouco, M. G. Penedo, and M. Ortega, “Hourbeat: Heartbeat classification fusing temporal and morphological information of ECGs via ensemble of classifiers,” Biomed. Sign. Process. Control, vol. 47, pp. 41–48, 2019.

[28] J. Hu, L. Shen, and G. Sun, “Squeeze-and-excitation networks,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2018, pp. 7132–7141.

[29] S. G. Wagner, Marriott’s Practical Electrocardiography. Baltimore, MD, USA: Williams & Wilkins, 2001.

[30] C. Carreiras et al., “BioSPPy: Biosignal processing in Python,” 2015. [Online]. Available: https://github.com/PIA-Group/BioSPPy/.

[31] Y. Bian, “ECG processing: Segmentation and delineation for ECG signals in Python,” 2021. [Online]. Available: https://github.com/alwaysbyx/ecg-processing/.

[32] S. Kurisu et al., “Poor R-wave progression and myocardial infarct size after anterior myocardial infarction in the coronary intervention era,” Int. J. Cancer Heart Vascul., vol. 7, pp. 106–109, 2015.

[33] T.C. Chan, W.J. Brady, R.A. Harrigan, J.P. Ornato, and P. Rosen, “ECG in emergency medicine and acute care,” 2005. [Online]. Available: https://books.google.com/books?id=mKSFGqAACAAJ.

[34] M. J. Van der Laan, E. C. Bolley, and A. E. Hubbard, “Super learner,” Statist. Appl. Genetics Mol. Biol., vol. 6, no. 1, 2007, doi: 10.2202/1541-026X.106.

[35] A. H. Ribeiro et al., “Automatic diagnosis of the 12-lead ECG using a deep neural network,” Nature Commun., vol. 11, 2020, Art. no. 1760.

[36] D. P. Kingma and J. Ba, “Adam: A method for stochastic opti-mization,” 2014, arXiv:1412.6980.

[37] P. Goyal, P. Dollár, et al., “Accurate, large minibatch SGD: Training ImageNet in 1 hour,” 2017, arXiv:1706.02677.

[38] H. Zhang, W. Zhao, and S. Liu, “SE-ECCNet: A multi-scale deep residual network with squeeze-and-excitation module for ECG signal classification,” in Proc. IEEE Int. Conf. Bioinf. Biomed., 2020, pp. 2685–2691.

[39] C. Brito, A. Machado, and A. L. Sousa, “Electrocardiogram beat-classification based on a ResNet network,” Stud. Health Technol. Inform., vol. 264, pp. 55–59, 2019.

Yuexin Bian received the BS degree from Zhe-jiang University, Hangzhou, China, in 2021. She is currently working toward the PhD degree with the Department of Electrical and Computer Engineer-ing, University of California San Diego, San Diego, CA. Her research interests include machine learn-ing, optimization for control, data mining and signal processing and analysis. She contributes to this work in Zhejiang University.
Jintai Chen received the BS degree in applied statistics from the Zhongnan University of Economics and Law, Wuhan, China. He is currently working toward the PhD degree with the College of Computer Science, Zhejiang University, Hangzhou, China. His research interests include deep learning, machine learning and data mining, especially on the computer vision, data mining, and medical intelligence. He has authored some papers at prestigious international conferences and journals, such as, ICML, AAAI, IJCAI, CVPR, MICCAI, and IEEE Transactions on Medical Imaging, and served as reviewers for MICCAI, AAAI, IJCAI, IEEE Transactions on Neural Networks and Learning Systems, and Journal of Biological Systems.

Xiaojun Chen received the MS degree in applied informatics from the Liverpool University, Liverpool, UK, in 2017. His research interests include machine learning, data mining, deep learning and computer vision. He has authored some papers at conferences and journals, such as, Journal of the American College of Cardiology and IEEE Access.

Xiaoxian Yang received the PhD degree in management science and engineering from Shanghai University, Shanghai, China, in 2017. She is currently an assistant professor at Shanghai Polytechnic University, China. Her research interests include business process management, formal verification, wireless network, and mobile health. Her researches are supported by the National Natural Science Foundation of China (NSFC), Natural Science Foundation of Shanghai (NSFS), CERNET Innovation Project, and Foundation of Science and Technology Commission of Shanghai Municipality. She has published more than 20 papers in academic journals such as IEEE Transactions on Intelligent Transportation Systems, ACM Transactions on Multimedia Computing, Communications, and Applications, IEEE Transactions on Computational Social Systems, ACM Transactions on Internet Technology, Future Generation Computer Systems, Mobile Networks and Applications, International Journal of Software Engineering and Knowledge Engineering, Future Generation Computer Systems, International Journal of Distributed Sensor Networks, Remote Sensing, and international conferences such as CollaborateCom, SEKE and IDEAL. She obtained two patents and three registered software copyrights in China, involving Wireless Network, Workflow Management, and Formal Verification. She had participated in organizing international conferences and workshops, such as CollaborateCom 2018, ChinaCom 2019 and Mobicase 2019. She also worked as guest editor for Mobile Networks and Applications and WINE, and served as reviewers for IEEE Transactions on Industrial Informatics, IEEE Transactions on Intelligent Transportation Systems, ACM Transactions on Multimedia Computing, Communications, and Applications, ACM Transactions on Internet Technology, Wiley Transactions on Emerging Telecommunications Technologies, Future Generation Computer Systems, Peer-to-Peer Networking and Applications, Journal of Biomedical And Health Informatics, Wireless Networks, Computer Network, etc.

Jian Wu (Member, IEEE) received the PhD degree in computer science and technology from Zhejiang University, Zhejiang, China, in 1998. He is a CFF member, CCF TCSC member, CCF TCAPP member and member of the “151 Talent Project of Zhejiang Province”. He is recently the director of Research Centre of Zhejiang University and vice-president of National Research Institute of Big Data of Health and Medical Sciences of Zhejiang University. His research interests include medical artificial intelligence, service computing, and data mining.

For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/csdl.