IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 27 May 2024, accepted 17 June 2024, date of publication 20 June 2024, date of current version 20 November 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3417344

RESEARCH ARTICLE

A Generative Elastic Network Method for R Peak Detection Suitable for Few-Shot Learning

NAN XIAO[1,2], KUN ZHAO[1], RUIZE MA[1,2], AND HAO ZHANG[1]
[1]Institute of Microelectronics, Chinese Academy of Sciences, Beijing 100029, China
[2]School of Integrated Circuits, University of Chinese Academy of Sciences, Beijing 100049, China

Corresponding author: Hao Zhang (haozhang@ime.ac.cn)

This work was supported by the National Key Research and Development Program of China under Grant 2023YFC2410602.

ABSTRACT R peak detection is fundamental to the analysis of long-term electrocardiogram (ECG) signals. Despite their significant success in R peak detection, neural networks based on statistical learning usual require more than 50% of all data for training. However, it is often difficult to provide such a high proportion of training data in practice. This paper proposes a novel R peak detection method based on Generative Elastic Network (GEN), which is suitable for few-shot learning. Utilizing the Lobachevsky University Database (LUDB), this method achieves an accuracy exceeding 99% by using less than 3% of the data for training and 14% for validation. It dramatically reduces the dependency on large volumes of data for training and validation, while preserving an accuracy level that is on par with existing methods.

INDEX TERMS Generative elastic network, electrocardiogram (ECG), R peak detection, few-shot learning.

I. INTRODUCTION

Electrocardiogram (ECG) signals, as crucial electrophysiological indicators, find extensive applications in the early diagnosis, treatment, and long-term monitoring of cardiovascular diseases (CVD). ECG records the electrical activities of the heart during each individual cardiac cycle, or heartbeat. By examining the morphological characteristics of each heartbeat and the rhythm between heartbeats, medical professionals can effectively analyze and assess the cardiac functional status. Accurate identification and segmentation of each individual heartbeat in long-term continuous ECG signals form the foundation for subsequent analysis and research.

A typical complete heartbeat, as illustrated in Figure 1, generally comprises the P wave, QRS complex, T wave, and occasionally a U wave following the T wave. However, due to the U wave's typically minimal amplitude and unclear physiological mechanism, studies on ECG often overlook it [1]. In practical scenarios, the P and T waves may also have small amplitudes, making them difficult to detect. Therefore, the decisive component for identifying an individual heartbeat is the QRS complex, specifically the R peak. Thus, R peak detection is crucial for the independent identification and segmentation of ECG heartbeats.

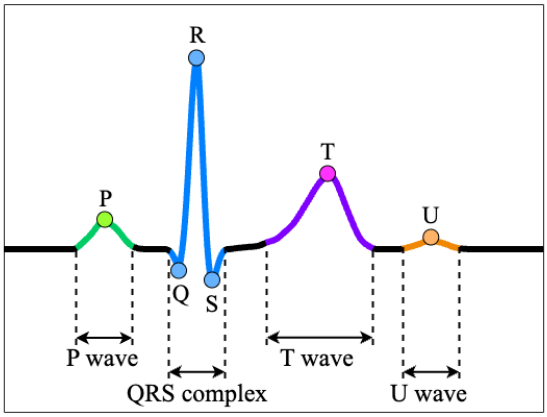


FIGURE 1. The components of an independent heartbeat.

In practical applications, ECG signals exhibit complex morphological, frequency, and amplitude characteristics, while also being subject to various interferences and noise, making R peak detection a challenging task. Traditional signal processing-based methods include mathematical morphology [2], digital filtering [3], wavelet transform [4], differentiator [5], and Hilbert transform [6], among others.

The associate editor coordinating the review of this manuscript and approving it for publication was M. Sabarimalai Manikandan.

When dealing with noise-disturbed and highly variable signals, these methods often result in instability and low accuracy. In the past decade, new methods and technologies based on deep neural networks have witnessed extensive development across various fields. In the realm of ECG R peak detection, techniques such as convolutional neural network (CNN) [7], Long Short Term Memory (LSTM) [8] and U-Net [9] have largely overcome the limitations of traditional methods, enabling more precise identification and extraction of features from the cardiac cycle in ECG signals, with overall accuracies even surpassing 99%. The majority of these methodologies are founded on the widely utilized MIT-BIH Arrhythmia Database [10] and the QT Database [11], both established in the previous century. The recordings within these databases comprise two-channel ECG signals, yet the corresponding leads for these channels are not consistent, with some even lacking explicit lead annotations. In contrast, Lobachevsky University Database (LUDB) [12], published in 2020, encompasses 200 records featuring comprehensive 12-lead signals, surpassing the MIT-BIH Arrhythmia Database and the QT Database in both quantity and quality of data. Recently, the detection of R peaks based on the LUDB has attracted considerable research interest. Matias et al. proposed a mask-based neural network approach for the point-by-point masked segmentation of time series to detect R peaks [13]. Han et al. integrated convolutional neural network (CNN), long short-term memory (LSTM) and ensemble learning in a novel method, employing a dynamic threshold adjustment strategy for decision-making [14]. Liang et al. designed a model based on an encoder-decoder structure, incorporating Standard Dilated Convolution Module (SDCM) to extract useful ECG signal features, and utilized bidirectional LSTM (BiLSTM) to capture temporal characteristics [15]. Furthermore, Chen et al. introduced a refined post-processing method for ECG waveforms based on a 1D-UNet architecture [16].

However, existing neural network models are based on statistical learning methods and typically require extensive synchronous data to extract statistically significant information [17]. These studies often focus solely on enhancing overall accuracy, for which more than 50% of the data is generally utilized for training purposes.

In practical scenarios, especially for rare diseases with scarce data, traditional neural networks trained on large datasets face significant challenges, prompting extensive research into few-shot learning (few-shot learning) [18]. The objective of few-shot learning is to enable network models to acquire new knowledge with only a minimal number of samples. To this end, researchers have proposed various strategies, such as meta-learning [19] methods. Model-Agnostic Meta-Learning (MAML) [20] and its variants such as first-order MAML (FOMAML) [21] optimize initial parameters through multi-task training to achieve rapid adaptation to new tasks. Additionally, metric-based learning strategies like Siamese Networks [22] and Matching Networks [23] classify by optimizing the similarity measures between samples, while Prototypical Networks [24] classify by computing the distance between class prototypes and test samples. Model-based approaches, such as generative models using Conditional Variational Auto-encoders (CVAEs) [25] and Generative Adversarial Networks (GANs) [26], support data augmentation and feature extraction by generating new sample data. Techniques such as AutoAugment [27] and transfer learning [28], are also employed to reduce the required number of training samples by transferring knowledge from large datasets. Despite these efforts, existing few-shot learning methods largely rely on additional optimization work and have not fundamentally resolved the dependency of neural networks on extensive data. Moreover, while these studies primarily focus on image processing, effective few-shot learning strategies for one-dimensional signal processing tasks, such as ECG R peak detection, remain relatively unexplored.

Therefore, we propose a novel R peak detection method based on Generative Elastic Network (GEN), which is naturally suitable for few-shot learning. Unlike neural networks based on statistical learning methods, GEN does not rely on statistical significance of data, but treats each sample equally: if existing knowledge can correctly classify the sample, it is deemed redundant and discarded; if not, it is considered a key sample for learning. Therefore, for a certain piece of knowledge (such as the waveform of a typical QRS complex), GEN only needs one key sample to learn it, instead of requiring a large number of redundant and repeated samples to learn like statistical learning methods. Experimental results based on the LUDB show that while the accuracy of the GEN-based R peak detection method proposed in this paper has reached the level of existing methods, the dependence on training and validation data has been significantly reduced from more than 50% of the total amount of data to 17%, which demonstrates its significant advantages in few-shot learning.

II. METHODS

Neural networks based on statistical learning methods typically fit the mapping relationship from the sample space to the feature space, i.e., the feature extraction process, by utilizing abundant training data, and eventually output prediction results through methods such as Softmax. The fitting of feature extraction needs to balance between overfitting and underfitting [29]. For this purpose, about more than 50% of the data is usually required for training, and only less than 40% of the data is used for testing, which is far from meeting the needs of few-shot learning.

On the other hand, statistical learning requires the presetting of numerous hyperparameters, and its training process relies on the back-propagation algorithm. After completing the calculation of all network parameters once, the network parameters are adjusted by propagating the loss function backward, completing one epoch. The training process of deep networks usually requires numerous epochs of trial

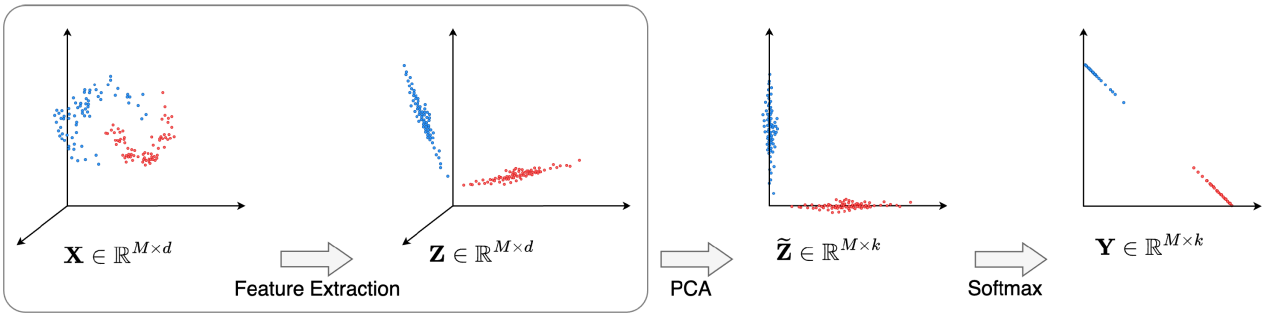


FIGURE 2. The transformation process from the sample set **X** to the label set **Y**.

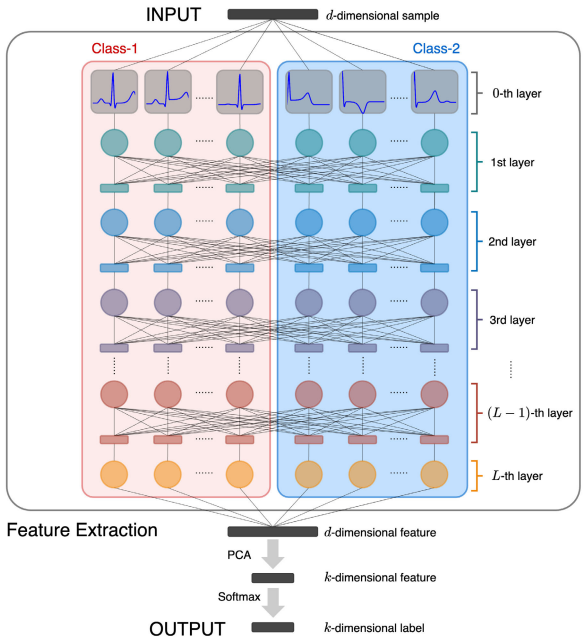


FIGURE 3. The architecture of GEN.

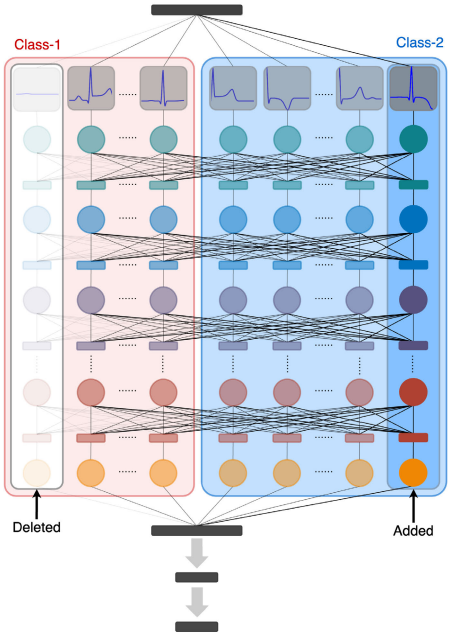


FIGURE 4. Implement error correction and expansion of GEN by deleting or adding a group of nodes.

and error before the loss function can gradually converge. Typically, training a neural network is challenging, lacking clear theoretical guidance and adjustment directions, relying solely on empirical trial and error.

Unlike traditional neural networks, Generative Elastic Network (GEN) is a novel model distinct from statistical learning methods. Our research on GEN has been submitted but has not yet been published.GEN is derived from the concept of **Spatial Coding Dimension** in information compression theory, which characterizes the density of a dataset in space. The goal of feature extraction is to minimize the Spatial Coding Dimension for data within the same class (i.e., to maximize the aggregation of data within classes), while maximizing the Spatial Coding Dimension between different classes (i.e., to maximize the separation between classes). We abstract the process of feature extraction as a mathematical optimization problem that aims to minimize the Spatial Coding Dimension. Through mathematical derivation, we ensure that it constitutes a convex optimization problem. Subsequently, the iterative solution process of convex optimization is expanded into a network architecture.

The architecture of GEN is computed through training rather than presetting, and the parameters of each node in the network only need to be calculated forward once without the need for repeated back-propagation. In other words, it only requires one epoch to complete the calculation. Moreover, the nature of convex optimization ensures that the network will converge strictly.

Essentially, GEN also seeks the mapping process from samples to features, but instead of directly fitting the mapping itself, it records the mapping process of **key samples** to their features. Finally, it outputs prediction results through Principal Component Analysis (PCA) and Softmax.

Figure 2 shows the binary classification process of M samples in a 3-dimensional space, where the data dimension $d = 3$, the number of classes $k = 2$. The original training set $\mathbf{X}$ is transformed into $\mathbf{Z}$ after L iterations of

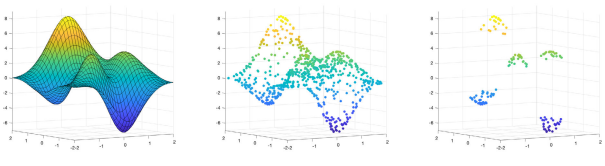


FIGURE 5. Global and local mapping relationships.(a) The statistical learning method fits the global mapping relationship; (b) Use many local mapping relationships to replace the global relationship; (c) Usually the data is sparsely distributed in space.

feature extraciton, where the positions in space gradually shift until they are distributed near two mutually orthogonal lines, each corresponding to a class. Subsequently, PCA is used to reduce the 3-dimensional space of $\mathbf{Z}$ down to the 2-dimensional space of $\widetilde{\mathbf{Z}}$, where the two axes correspond to the two orthogonal lines in the original 3-dimensional space. Finally, the 2-dimensional space of $\widetilde{\mathbf{Z}}$ is normalized using Softmax, outputting 2-dimensional classification labels $\mathbf{Y}$. In this process, each sample in the training set is considered as a **key sample** for learning.

Therefore, as shown in Figure 3, GEN computes a rectangular structure of $M \times L$, where M is the number of key samples, and L is the number of iterations of the convex optimization solution process corresponding to the network during training. Each column in GEN represents the process by which each key sample obtains the corresponding features through L transformations. A column in the network is the local mapping relationship between the key sample and its corresponding feature. The combination of many columns in the network is to accumulate many local mapping relationships and piece together an approximately global relationship.

The subsequent network validation process involves individually testing each sample in the validation set: if the existing network can correctly classify the sample, it is discarded; otherwise, the sample is considered a key sample. The training algorithm is then reapplied to compute the transformation process of this key sample within the L-layer network, and it is added as a new column to the network architecture. Additionally, through regular reviews, key samples that are prone to errors are identified and removed from the network. Figure 4 shows the changes in network architecture corresponding to the addition and deletion of key samples. Therefore, the key samples in GEN actually include all samples from the training set, as well as those selected from the validation set for expanded learning.

As shown in Figure 5, in many practical applications such as ECG R peak detection, the data in the mapping space is sparsely distributed. Reflecting the global relationship through the local mapping relationships of many key samples not only improves efficiency but also simplifies expansion and error correction. This learning method is essentially a continuous process of plugging gaps, where newly learned local relationships do not affect other existing local relationships, thereby avoiding the problem of resolving old errors but introducing new ones.

For GEN, the majority of the data can be used as the test set, while only a small portion is reserved for training and validation, which shows significant advantages for few-shot learning. Initially, GEN adds all training samples as key samples to the network architecture. Subsequently, the network undergoes validation. The essence of validation involves individually inputting samples from the validation set into the existing network for testing. If a sample passes the test, it is discarded; if it fails, the sample is identified and selected as a new training sample for secondary training. Importantly, this secondary training does not adjust the existing network nodes but adds new columns to the network. This learning mode closely resembles the human method of refining knowledge reserves by accumulating typical individual cases. That is, for new problems encountered, if they can be solved with existing knowledge, they are considered redundant; if not, they warrant individualized learning. This means that the network itself is elastic, and the addition or deletion of network nodes corresponds to the addition or deletion of recorded key samples and their mapping processes, as shown in Figure 4. Through the expansion and correction of the validation set, in cases of good data quality, GEN can even improve accuracy to 100%.

Neural networks based on statistical learning methods model the mapping process from data to labels using statistical information from a vast amount of data. The impact of an individual sample on the overall network training result is almost negligible, resulting in poor adaptability to small-sample datasets. In contrast, GEN do not model the mapping from data to labels per se, but instead treat training and validation samples as key samples, individually recording the mapping process for each sample. Each key sample can directly modify the network by adding or deleting a row of nodes without affecting other nodes.

On the other hand, neural networks based on statistical learning methods fundamentally seek to delineate the separation boundaries between different classes and subsequently determine through activation functions which side of the boundary a sample belongs to [30]. Each node within the network serves as a fitting parameter for this boundary. Conversely, non-statistical learning methods such as GEN do not attempt to fit separation boundaries. GEN identifies all necessary samples through training and validation processes to be learned as key samples. This approach embodies a lightweight learning model closer to human-like learning, where knowledge is expanded by accumulating typical ''problem cases''. For this type of learning, mastering one sample of a problem case suffices, unlike statistical learning methods that require a substantial amount of redundant data of the same type to derive statistically valuable information, which is particularly well-suited for few-shot learning.

Therefore, GEN is an entirely data-dependent network model, with the advantage of requiring minimal data quantities for learning, thus satisfying the demands of few-

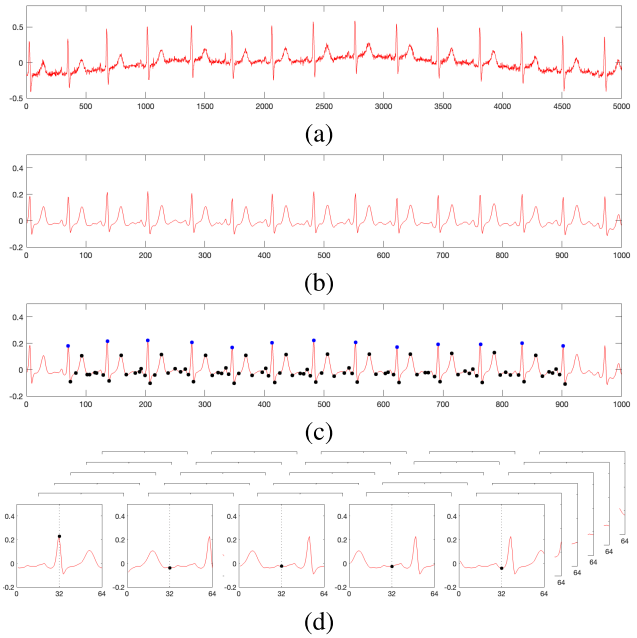


FIGURE 6. Data preprocessing. (a) 500Hz original signal of Record No. 142 lead I; (b) The signal after 1 to 15 Hz bandpass filtering and resampling to 100Hz; (c) Finding all inflection points, where the black points are the non-R points and the blue ones are the R peaks; (d) Take 64 sampling points centered on each inflection point as data samples.

shot learning. However, this brings about several challenges, namely high demands on the quality of data used for training and validation (i.e., key samples). On one hand, the labels of these samples must be accurate, as the network's capability to handle noise and outliers is limited. Any incorrect reference in a learned "problem case" could severely impact the learning outcomes. On the other hand, the samples must be distinct and representative, ensuring diversity among problem cases without redundancy.

III. DATASET

Existing databases that are widely applied include the MIT-BIH Arrhythmia Database and the QT Database. The MIT-BIH Arrhythmia Database comprises 48 recordings from 47 subjects collected between 1975 and 1979, while the QT Database aggregates 105 records from several datasets of the last century. Each record in these databases contains only two-channel ECG, and the corresponding leads for each channel are inconsistent. In fact, the morphological features of ECG signals vary significantly between different leads, which complicates the detection of the R peak. In contrast, the Lobachevsky University Database (LUDB), released in 2020, includes 200 records from 200 subjects collected between 2017 and 2018, encompassing both healthy participants and patients with various cardiovascular diseases. While the data diversity is richer, the total amount of data is almost six times that of QT database. More importantly, each record in LUDB includes a complete and independent set of 12 lead signals, which is crucial for both single-lead and multi-lead analysis.

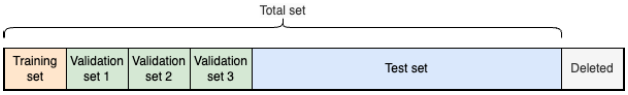


FIGURE 7. The records contained in each set are shown in Table 1, where the 24 records of the validation set are used for three times.

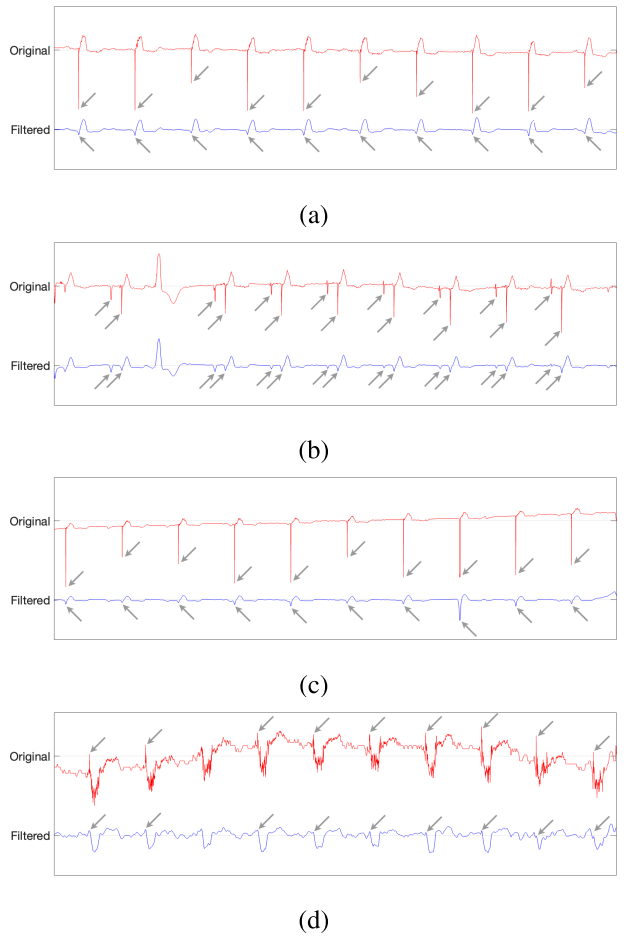


FIGURE 8. 4 deleted records with high-frequency pacemaker signals. (a) No.34; (b) No. 45; (c) No.74; (d) No.93.

Therefore, this study conducts experimental analysis based on LUDB.

The signals in LUDB are digitized at 500 Hz. The R peaks of QRS complexes are manually annotated by cardiologists for all records and independently for each lead, and all records received an expert classification by abnormalities [12]. This dataset contains a total of 58,429 annotated P waves, QRS complexes, and T waves.

A. DATA PREPROCESSING

ECG signals collected from different devices often vary in quality and sampling frequency. The noise present in ECG signals mainly includes electromyographic noise, power-line interference, and baseline drift [31]. Considering the generalization ability under complex conditions, we employ

TABLE 1. Divide 194 records in LUDB into training set, validation set and test set.

Set	Records	Number of records	Number of samples
Total	All records except 34, 45, 74, 93	196	16959
Training	10, 99, 101, 110, 111, 125	6	490(2.89%)
Validation 1	1, 8, 17, 35, 41, 69, 83, 90		
Validation 2	100, 108, 116, 117, 133, 143, 146, 148,	24	2344(13.82%)
Validation 3	160, 167, 168, 174, 175, 178, 189, 190		
Test	All other records	166	14125(83.29%)

the simplest bandpass filtering from 1 to 15 Hz and downsample to 100 Hz. Although this may introduce some distortion to the waveform, it does not affect R peak detection.

Considering that the R peak is the maximum amplitude in the R wave, representing one of the extreme points on the ECG signal, the detection of R peak is to classify all extreme points into two classes, namely R peak inflection points and non-R peak inflection points. These non-R peak inflection points include noise, P peak, T peak, etc. Therefore, we do not need to identify each sampling point in the signal one by one, but only need to pay attention to all inflection points and organize the data set for learning and prediction. We take each inflection point as the center and extract 31 and 32 sampling points backward and forward, respectively, forming a data segment of 64 sampling points as a data sample, organizing the dataset accordingly.

B. DATA PARTITION

For the 12-lead ECG signals in the LUDB, we choose lead I from each record excluding records number 34, 45, 74, and 93, as shown in Figure 8. The high-frequency pacemaker signals in these four records will be filtered out during preprocessing. The remaining 196 records are used as the data source, and the total of 16,959 samples are divided into training, validation, and test sets according to Table 1. The total set represents the union of these three sets and is also used as a part of test in the subsequent experimental section to verify whether the learned samples can be correctly identified. Besides, the validation set is divided into 3 subsets, each containing 8 records, which are provided to GEN in turn during the subsequent validation process. The schematic diagram of data partition is shown in Figure 7. Since GEN is entirely data-dependent, the selection of data used for training and validation is very important. The training and validation data were selected based on varied waveform morphologies of the records, rather than at random. In the following subsections, we will explain why these particular records were chosen.

IV. EXPERIMENTAL STUDY

For each sample consisting of 64 sampling points, if the 32nd point corresponds to an R peak, then that sample is labeled as an R peak. Conversely, if the central 32nd point is not an R peak (it could be noise, a P wave, a T wave, etc.), then the sample is labeled as non-R point. For instance, in Figure 6(d), the first sample on the left is labeled as an R peak, while the remaining four samples are not. This approach effectively transforms the detection of R peaks into a binary classification problem.

From the results of this binary classification, we calculated three quantitative results: true-positive (TP) when a R peak is correctly detected by the proposed algorithm, false-negative (FN) when a R peak is missed, and false-positive (FP) when a non-R point is detected as R peak. To evaluate the performance of the proposed detection algorithm, sensitivity (Sen), positive predictive value (PPV), F1 Score and accuracy(Acc) can be computed by using the following equations, respectively

$$Sen = \frac{TP}{TP + FN} \times 100\% \tag{1}$$

$$PPV = \frac{TP}{TP + FP} \times 100\% \tag{2}$$

$$F1 = \frac{2 \times Sen \times PPV}{Sen + PPV} \times 100\% \tag{3}$$

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \times 100\%. \tag{4}$$

Sensitivity, also known as the true positive rate, measures the model's ability to correctly identify positive instances. Positive Predictive Value (PPV), indicates the proportion of positive identifications that were actually correct, which is essential in scenarios where the consequences of false positives are significant, ensuring that predictions are reliable. F1 Score, which is the harmonic mean of Sensitivity and Positive Predictive Value, making it particularly suitable for imbalanced datasets. Accuracy, the most intuitive performance metric, reflects the overall correctness of the model across both classes.

A. INITIAL TRAINING

According to the data partition shown in Table 1, we initially utilize a training set composed of 490 samples from 6 records. These 6 records were selected due to their varied waveform morphologies, detailed as follows:

- Record No.10 displays a completely normal ECG signal.
- Record No.90 features generally small R-peak amplitudes and deep S waves.
- Record No.101 has a mostly normal ECG, but includes a waveform with a notably wide R wave.
- Record No.110 predominantly exhibits small R-peak amplitudes and has instances of inverted T waves; additionally, one of the R peaks is inverted.

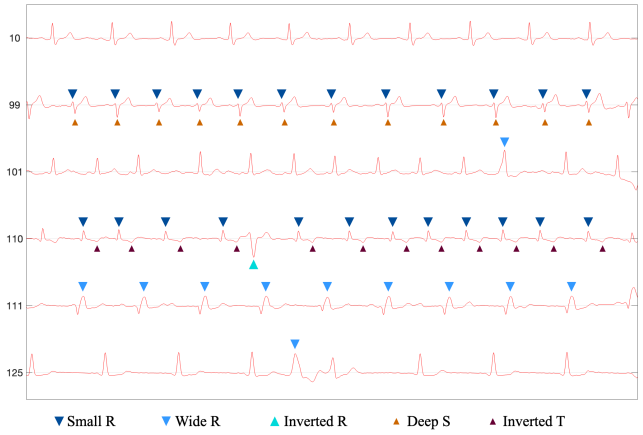


FIGURE 9. 6 records in the training set.

- Record No.111 is characterized by wide R waves.
- Record No.125 presents a distinctly wide R wave.

As shown in Figure 9, despite comprising only six records, the dataset showcases a diversity of waveforms, including small R waves, multiple forms of wide R waves, inverted R waves, deep S waves, and inverted T waves. For GEN, the diversity and representativeness of the data are crucial for effective learning.

During the training phase, GEN considers all samples as key samples for learning, thereby forming a network with **490 key samples** $\times$ **10 layers**. Initially, GEN does not filter out any data from the training set but instead learns from all samples, leading to 490 rows in GEN following the initial training.

Subsequently, we directly use the test set and total set for testing. The performance metrics for R peak detection are shown in Table 2. It can be observed that by training on less than 3% of the data, GEN achieves an accuracy of approximately 99%. This demonstrates a significant difference from traditional statistical learning methods, which require large amounts of data to acquire knowledge. GEN can learn quite rich and effective knowledge from a small number of key samples.

B. VALIDATION

According to the data partition shown in Table 1, validation was conducted three times, each based on records included in validation sets 1, 2, and 3.

First, we use validation set 1 consisting of 8 records totaling 808 samples to expand and correct GEN. This results in a network of **509 key samples** $\times$ **10 layers**, which has added 19 key samples compared to the initial trained GEN. This indicates that from the 8 records in validation set 1, 19 samples that had not yet been mastered were identified and retrained.

Next, we proceeded with validation set 2, composed of 742 samples from 8 records. This led to a network with **518 key samples** $\times$ **10 layers**, an increase of 9 key samples compared to the first validation, indicating that GEN identified and retrained 9 new unmastered samples from the 8 records in validation set 2.

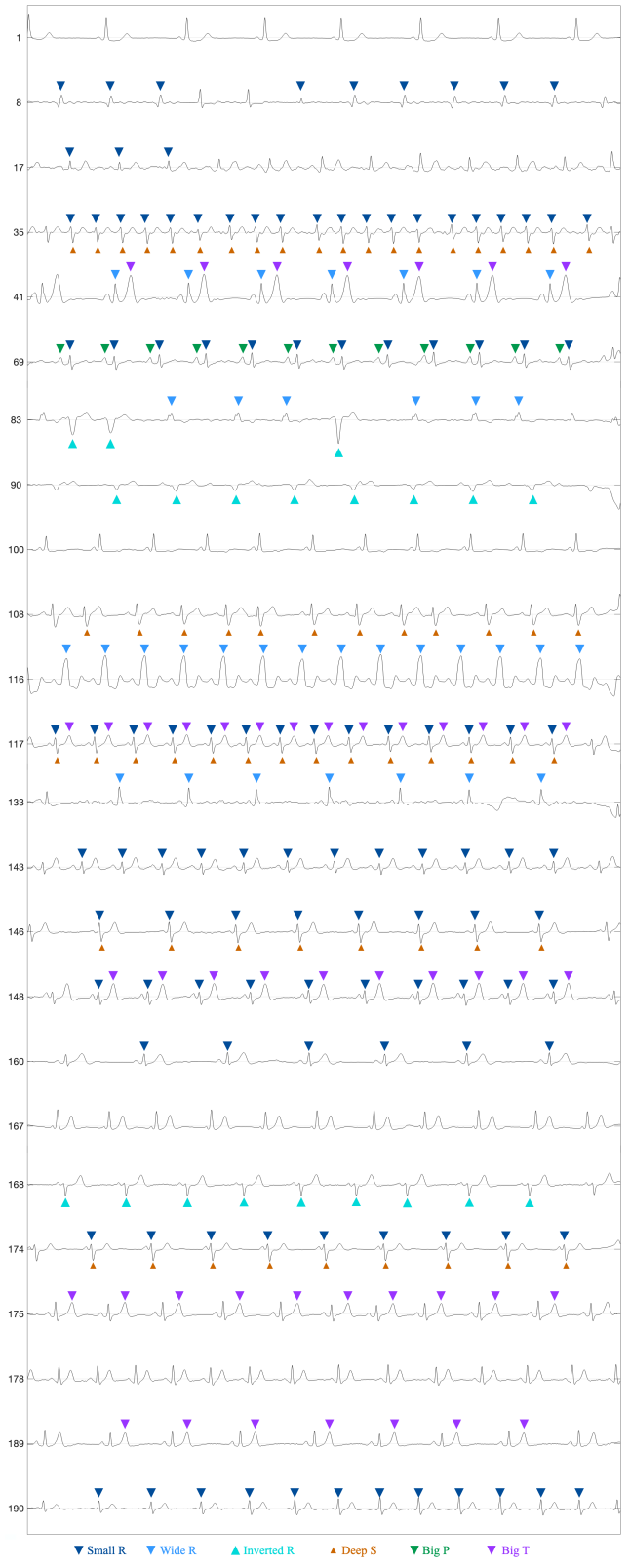


FIGURE 10. 24 records in the validation set.

TABLE 2. Test performance of GEN on the test set and total set after initial training.

Test on	TP	TN	FP	FN	Sen(%)	PPV(%)	F1(%)	Acc(%)
Test	1452	12592	31	50	96.67	97.91	97.29	99.43
Total	1663	15058	106	132	92.65	94.01	93.32	98.60

TABLE 3. Test performance of GEN on the test set after each validation.

Validation	TP	TN	FP	FN	Sen(%)	PPV(%)	F1(%)	Acc(%)
1	1479	12599	24	23	98.47	98.40	98.44	99.67
2	1499	12623	0	3	99.80	100	99.90	99.98
3	1502	12623	0	0	100	100	100	100

TABLE 4. Test performance of GEN on the total set after each validation.

Validation	TP	TN	FP	FN	Sen(%)	PPV(%)	F1(%)	Acc(%)
1	1749	15066	98	46	97.44	94.69	96.05	99.15
2	1780	15146	18	15	99.16	99.00	99.08	99.81
3	1795	15164	0	0	100	100	100	100

TABLE 5. Comparison result of proposed method with existing methods.

Method	Data partition (proportion)			Performance measures			
	Training	Validaiton	Test	Sen(%)	PPV(%)	F1(%)	Acc(%)
Proposed	**2.89%**	**13.82%**	**83.29%**	**100**	**100**	**100**	**100**
Chen (2023) [16]	80%	-	20%	100	100	100	100
Han (2022) [14]	70%	-	30%	100	100	100	100
Liang (2022) [15]	50%	20%	30%	99.50	98.27	98.88	97.73
Matias (2021) [13]	64%	16%	20%	100	99.70	99.85	99.81

Finally, using validation set 3, which also comprised 742 samples from 8 records, led to a network with **527 key samples** $\times$ **10 layers**, an increase of 9 key samples from the second validation, demonstrating that GEN identified and retrained 9 additional new unmastered samples from the records in validation set 3.

The test results obtained from the three validation show that the more data used for validation, the more complete the knowledge of R peak detection learned by GEN. That is, the local relationships used to replace the global relationships are continuously filled. The performance metrics are shown in Table 3 and 4. After the third validation with 13.82% data, the accuracy of the prediction has reached 100%, and there is no need to provide more data for validation. GEN has selected 37 ''unlearned'' new knowledge from these 24 records including 2344 validation samples for learning. The waveforms of these 24 records, selected due to their varied morphologies, are depicted in Figure 10 and include the following specifics:

- Big P Waves: Record No.69.
- Deep S Waves: Records No.35, 108, 117, 146, 174.
- Wide R Waves: Records No.41, 83, 116, 133.
- Small R Waves: Records No.8, 17, 35, 69, 117, 143, 148, 174.
- Inverted R Waves: Records No.83, 90, 168.
- Big T Waves: Records No.41, 117, 143, 146, 148, 160, 175, 189.

C. RESULTS AND DISCUSSION

Existing R peak detection methods based on statistical learning typically use over 50% of the data for training to achieve accuracies above 99%, while the specific requirements for validation data vary. However, the proportion allocated for testing does not exceed 30%. Table 5 provides a comparison between the GEN-based detection method proposed in this paper and current methods that also use LUDB for R peak detection. It is evident that the training data requirement for the GEN-based method drastically reduces from over 50% to just 2.89%, and the demand for validation data is only 13.82%. For GEN, such training and validation data represent the minimal dataset required to ensure accuracy reaches the desired level. Any additional training and validation data would be redundant for GEN and therefore would not impact the test results. Together, training and validation account for 16.71% of the total data, substantially increasing the proportion available for testing to 83.29%. Meanwhile, the performance indicators maintain the same level as existing methods. This demonstrates the significant advantages of the GEN-based R peak detection method in few-shot learning.

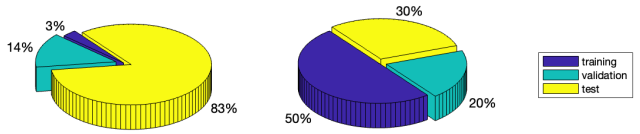


FIGURE 11. The pie chart on the left shows that the requirement for training and validation of the GEN-based method proposed in this paper is only about 17%. The pie chart on the right shows the requirement for training and validation of statistical learning methods is usually exceed 70%.

For a more intuitive comparison of the distribution ratios of training, validation, and testing sets, Figure 11 presents a clearer visual contrast.

Statistical learning methods typically utilize large datasets for training and validation, often employing random sampling techniques to maintain statistical consistency across the data. This approach ensures that the selected data is representative and consistent in statistical terms. However, GEN, being a fully data-dependent network model, has distinct requirements. Although it demands significantly less data for training and validation, it imposes stringent quality criteria on the data, encompassing two main aspects:

- Accuracy of Labels: GEN treats each sample equally, lacking the capacity to discern erroneous sample-label pairs during the training and validation processes. It operates under the assumption that all sample-label pairs utilized are accurate, which could lead to challenges if this assumption does not hold true.
- Representativeness and Diversity of Data: The samples used for training and validating GEN need to be as representative and diverse as possible to minimize redundancy. This reduces dependency on large data volumes. If the samples used are too similar, GEN may only learn a narrow scope of knowledge and fail to capture a comprehensive understanding of the dataset.

The selection of training and validation data thus becomes crucial. Currently, this process relies heavily on manual curation, which highlights the need to develop automated methods to replace manual selection. On the other hand, experimental results based on the LUDB indicate that the existing training and validation sets are sufficient, as performance metrics have already reached their optimum. Consequently, any additional data for training and validation would be considered redundant by the network and discarded. However, the determination of the appropriate sizes for the training and validation sets in this study relies on empirical trial and error, lacking a quantitative research on their optimal size selection.

V. CONCLUSION

The detection of R peaks is fundamental in analyzing long-term ECG signals. Neural networks based on statistical learning have achieved significant success in addressing R peak detection challenges. However, these methods typically require extensive datasets for training, often exceeding 50% of the total data volume to achieve accuracies above 99%, which poses a challenge for in few-shot learning. This paper introduces a novel GEN-based method for R peak detection. Experimental results using 16,959 samples from the LUDB database show that training with just 2.89% of the data can yield a testing accuracy of about 99%. Upon further validation with 13.82% of the data, accuracy escalates to 100%. These results indicate that the GEN-based R peak detection method proposed in this paper reaches the performance levels of existing methods while significantly reducing the demand for training and validation data from over 50% to less than 17%, thereby demonstrating considerable advantages in few-shot learning applications. Although this method still has shortcomings such as the need to manually select training and validation data, it also provides inspiration for similar few-shot learning tasks to reduce reliance on large-scale data when performing signal identification and analysis through network models.

REFERENCES

[1] B. Surawicz, "U wave: Facts, hypotheses, misconceptions, and misnomers," *J. Cardiovascular Electrophysiol.*, vol. 9, no. 10, pp. 1117–1128, Oct. 1998.

[2] S. Yazdani and J.-M. Vesin, "Extraction of QRS fiducial points from the ECG using adaptive mathematical morphology," *Digit. Signal Process.*, vol. 56, pp. 100–109, Sep. 2016.

[3] A. Kaur, A. Agarwal, R. Agarwal, and S. Kumar, "A novel approach to ECG R-Peak detection," *Arabian J. Sci. Eng.*, vol. 44, no. 8, pp. 6679–6691, Oct. 2018.

[4] N. Sharma, R. K. Sunkaria, and L. D. Sharma, "QRS complex detection using stationary wavelet transform and adaptive thresholding," *Biomed. Phys. Eng. Exp.*, vol. 8, no. 6, Nov. 2022, Art. no. 065011.

[5] C. Nayak, S. K. Saha, R. Kar, and D. Mandal, "An optimally designed digital differentiator based preprocessor for R-peak detection in electrocardiogram signal," *Biomed. Signal Process. Control*, vol. 49, pp. 440–464, Mar. 2019.

[6] M. A. Kumar and K. M. Chari, "Efficient FPGA-based VLSI architecture for detecting R-peaks in electrocardiogram signal by combining Shannon energy with Hilbert transform," *IET Signal Process.*, vol. 12, no. 6, pp. 748–755, Aug. 2018.

[7] A. Peimankar and S. Puthusserypady, "DENS-ECG: A deep learning approach for ECG signal delineation," *Expert Syst. Appl.*, vol. 165, Mar. 2021, Art. no. 113911.

[8] J. Laitala, M. Jiang, E. Syrjälä, E. K. Naeini, A. Airola, A. M. Rahmani, N. D. Dutt, and P. Liljeberg, "Robust ECG R-peak detection using LSTM," in *Proc. 35th Annu. ACM Symp. Appl. Comput.*, Mar. 2020, p. 1104.

[9] S. Vijayarangan, R. Vignesh, B. Murugesan, S. Preejith, J. Joseph, and M. Sivaprakasam, "RPnet: A deep learning approach for robust r peak detection in noisy ECG," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 345–348.

[10] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, Jun. 2001.

[11] P. Laguna, R. G. Mark, A. Goldberg, and G. B. Moody, "A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG," in *Proc. Comput. Cardiol.*, Sep. 1997, pp. 673–676.

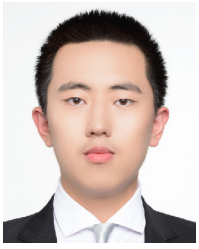
[12] A. I. Kalyakulina, I. I. Yusipov, V. A. Moskalenko, A. V. Nikolskiy, K. A. Kosonogov, G. V. Osipov, N. Y. Zolotykh, and M. V. Ivanchenko, "LUDB: A new open-access validation tool for electrocardiogram delineation algorithms," *IEEE Access*, vol. 8, pp. 186181–186190, 2020.

[13] P. Matias, D. Folgado, H. Gamboa, and A. Carreiro, "Time series segmentation using neural networks with cross-domain transfer learning," *Electronics*, vol. 10, no. 15, p. 1805, Jul. 2021.

[14] C. Han, W. Que, S. Wang, J. Zhang, J. Zhao, and L. Shi, "QRS complexes and t waves localization in multi-lead ECG signals based on deep learning and electrophysiology knowledge," *Expert Syst. Appl.*, vol. 199, Aug. 2022, Art. no. 117187.
[15] 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.
[16] Z. Chen, M. Wang, M. Zhang, W. Huang, H. Gu, and J. Xu, "Post-processing refined ECG delineation based on 1D-UNet," *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 104106.
[17] V. N. Vapnik, "An overview of statistical learning theory," *IEEE Trans. Neural Netw.*, vol. 10, no. 5, pp. 988–999, Sep. 1999.
[18] Y. Song, T. Wang, P. Cai, S. K. Mondal, and J. P. Sahoo, "A comprehensive survey of few-shot learning: Evolution, applications, challenges, and opportunities," *ACM Comput. Surveys*, vol. 55, no. 13s, pp. 1–40, Jul. 2023.
[19] T. Hospedales, A. Antoniou, P. Micaelli, and A. Storkey, "Meta-learning in neural networks: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 9, pp. 5149–5169, Sep. 2022.
[20] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *Proc. 34th Int. Conf. Mach. Learn.*, vol. 70, Aug. 2017, pp. 1126–1135.
[21] A. Nichol, J. Achiam, and J. Schulman, "On first-order meta-learning algorithms," 2018, *arXiv:1803.02999*.
[22] D. Chicco, "Siamese neural networks: An overview," in *Artificial Neural Networks*. New York, NY, USA: Humana, Aug. 2020, pp. 73–94.
[23] O. Vinyals, C. Blundell, T. Lillicrap, K. Kavukcuoglu, and D. Wierstra, "Matching networks for one shot learning," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 29, 2016, pp. 3637–3645.
[24] J. Snell, K. Swersky, and R. Zemel, "Prototypical networks for few-shot learning," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 4080–4090.
[25] J. Xu and H. Le, "Generating representative samples for few-shot classification," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 8993–9003.
[26] Z. Ren, Y. Zhu, Z. Liu, and K. Feng, "Few-shot GAN: Improving the performance of intelligent fault diagnosis in severe data imbalance," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–14, 2023.
[27] B. Demirel, O. B. Baran, and R. G. Cinbis, "Meta-tuning loss functions and data augmentation for few-shot object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 7339–7349.
[28] Q. Sun, Y. Liu, T.-S. Chua, and B. Schiele, "Meta-transfer learning for few-shot learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 403–412.
[29] A. Shrestha and A. Mahmood, "Review of deep learning algorithms and architectures," *IEEE Access*, vol. 7, pp. 53040–53065, 2019.
[30] S. R. Dubey, S. K. Singh, and B. B. Chaudhuri, "Activation functions in deep learning: A comprehensive survey and benchmark," *Neurocomputing*, vol. 503, pp. 92–108, Sep. 2022.
[31] S. K. Berkaya, A. K. Uysal, E. S. Gunal, S. Ergin, S. Gunal, and M. B. Gulmezoglu, "A survey on ECG analysis," *Biomed. Signal Process. Control*, vol. 43, pp. 216–235, May 2018.

NAN XIAO received the B.S. degree in computer science and technology from Guilin University of Electronic Technology, Guangxi, China, in 2015, and the M.S. degree in computer science and technology from Guangxi University, Guangxi, in 2016. He is currently pursuing the Ph.D. degree with the Institute of Microelectronics, Chinese Academy of Sciences, and the University of Chinese Academy of Sciences. His research interests include machine learning and physiological signal processing.

KUN ZHAO received the M.S. degree in control theory and engineering from Xi'an University of Architecture and Technology, Shaanxi, China, in 2015. She is currently an Assistant Research Fellow with the Institute of Microelectronics, Chinese Academy of Sciences. Her research interests include physiological signal processing algorithms and ultra-low power embedded systems.



RUIZE MA received the B.S. degree in microelectronics science and engineering from the National University of Defense Technology, Hunan, China, in 2022. He is currently pursuing the Ph.D. degree with the Institute of Microelectronics, Chinese Academy of Sciences, and the University of Chinese Academy of Sciences. His research interests include circuit design and machine learning.

HAO ZHANG received the M.S. degree in signal processing from Hunan University, Hunan, China, in 2001, and the Ph.D. degree in signal processing from South China University of Technology, Guangdong, China, in 2005. He is currently a Professor with the Institute of Microelectronics, Chinese Academy of Sciences. His research interests include machine learning and physiological signal processing and in vitro diagnostics.