2025 International Conference on Intelligent Computing
July 26-29, Ningbo, China
https://www.ic-icc.cn/2025/index.php

GFR: An Effective Plugin for Enhancing the ECG Classification Capability of Models

Xunde Dong[1], Yupeng Qiang[1*], Xiuling Liu[2], Yang Yang[1], Yihai Fang[3], and Jianhong Dou[4]

[1] The School of Automation Science and Engineering, South China University of Technology, Guangzhou 510641 China
[2] The Key Laboratory of Digital Medical Engineering of Hebei Province, College of Electronic and Information Engineering, Hebei University, Baoding, 071002, China.
[3] Zhujiang Hospital, Southern Medical University, Guangzhou, 510280, China
[4] General Hospital of the Southern Theater of the Chinese People's Liberation Army, Guangzhou, 510030, China
audxd@scut.edu.cn

Abstract. Electrocardiogram (ECG) serves as a crucial non-invasive diagnostic tool for monitoring clinical cardiac conditions. Remarkable progress has been achieved in deep learning-based ECG classification research. Generally, its overall architecture can be divided into three parts: the feature extraction layer, the feature fusion methods (typically concatenation and summation), and the multilayer perceptron (MLP) classification layer. In this paper, we propose a plugin Global Feature Refinement (GFR) module to enhance the performance of multi-branch models for ECG classification. The GFR plugin assigns weights to different branching features in a dynamic disease-aware manner to capture critical global information while emphasizing important features. Specifically, these dynamic weights are obtained through the integration, mapping, and scaling of global features. Finally, the weighted features are summed for ECG classification. Extensive experiments on three large-scale imbalanced datasets demonstrate that the GFR plugin, with less 6.2k additional parameters, improves the performance of eight models of different sizes to varying degrees. Specifically, the maximum improvement in F1 score and accuracy was 8.27% and 6.41%, respectively.

Keywords: electrocardiogram (ECG) classification, feature refinement, multi-branch networks.

1 Introduction

Electrocardiogram (ECG), as a non-invasive technique for monitoring the electrical activity of the heart, has been playing an indispensable role in clinical cardiological diagnosis. With the incidence of cardiovascular diseases increasing year by year, accurate analysis and classification of ECG signals is particularly important. In recent years, deep learning techniques have achieved remarkable results in the field of medical image

processing and signal analysis, providing a new research direction for automatic ECG classification.

Currently, the multi-branch model stands out as the most prevalently utilized model in ECG classification research. It principally encompasses the following three components: a multi-branch extraction layer for extracting temporal features, a fusion layer (either concatenating or adding features) for feature integration, and a fully-connected layer for classification.

However, existing multi-branch models typically fuse the extracted features through concatenation or summation. This approach may fail to achieve a sufficiently effective integration of features from different branches, consequently impinging on the model's performance. Moreover, as the depth of the model continues to increase, its complexity also rises significantly. Although a deeper model can enhance performance to a certain extent, effectively controlling the model's complexity to avoid overfitting and reduce the computational burden has emerged as a pressing issue that demands attention in current research.

For the above problems, the emergence of plugin modules provides a new solution. In this way, specific functional modules can be added flexibly without changing the original model architecture. This approach not only helps to simplify the design process of the model, but also enables the model configuration to be dynamically adjusted according to the needs of specific application scenarios, thus achieving higher adaptability and flexibility.

In the field of image processing, the application of plugin modules has achieved remarkable results. For example, the use of plugin modules such as Squeeze-and-Excitation Network (SEnet) [10], Convolutional Block Attention Module (CBAM) [25], and Channel Attention (CA) [8] has improved the accuracy of image classification.

Inspired by this and considering that most of the ECG studies are multi-branch networks, in this paper we propose the GFR plugin to capture key global information and enhance important branching features
by assigning weights to different branching features in a dynamic disease-aware manner.

The main contributions of this paper are summarized as follows:

1. This paper presents an effective plugin, GFR, designed to improve the performance of widely used multi-branch networks for ECG classification tasks.
2. GFR can be seamlessly integrated into multi-branch networks as a practical plug-and-play solution for improving ECG classification performance. Notably, this integration requires very few additional parameters.
3. Extensive experiments on three large unbalanced datasets, PTB-XL, CPSC2018 and Chapman, as well as on different multi-branch network models, have shown that GFR is effective.

The rest of the paper is organized as follows: In Section 2, we review related research work. Section 3 elucidates the proposed methodology in detail. In Section 4, we present the experimental procedures and in Section 5 we present the corresponding results. In Section 6, an elaborate discussion is provided. Finally, we conclude our findings in Section 7.

2 Related Work

2.1 Deep model architectures for ECG classification

In recent years, deep learning models for 12-lead ECG analysis have demonstrated significant technological evolution. Early studies primarily adopted single-branch network architectures, processing the 12-lead ECG as an integrated input for feature learning. For instance, Su et al. [23], Hao et al. [5], and Sakli et al. [21] achieved end-to-end classification by modifying classical networks such as Resnet and VGG. Deng et al. [2] innovatively combined time-frequency representations with Resnet152 to enhance temporal feature extraction, while Yao et al. [27] developed an attention-based time-incremental convolutional neural network, enhancing arrhythmia detection accuracy through dynamic modeling. Although these methods benefit from structural simplicity, they struggle to fully exploit the local discriminative features inherent in multi-lead signals.

To address the limitations of single-branch architectures, researchers proposed multi-branch network architectures that process individual leads through independent branches. Han et al. [4] pioneered a multi-lead residual network that synergizes local features with global representations. Pan et al. [17] advanced this approach by introducing a multi-task channel attention network, which incorporates cross-lead interaction mechanisms where attention modules adaptively calibrate branch contributions. He et al. [7] further enhanced robustness by applying residual shrinkage networks to myocardial infarction localization, effectively suppressing noise interference through soft thresholding. These studies demonstrated that multi-branch architectures can more precisely capture lead-specific characteristics.

With advancements in multi-view learning theory, researchers have partitioned the 12 leads into distinct views to construct multi-view multi-branch networks. Qiang et al. [18] divided the 12-lead into 5 views based on position, and achieved feature complementarity through multicore CNNs. Ma et al. [16] extended this framework by establishing cross-modal views between temporal signals and 2D spectrograms, leveraging cross-modal attention to enhance morphological feature extraction. Guan et al. [3] developed a multi-path fusion network that integrates 3D spatial and spectral views of ECG signals through bidirectional feature pyramids. These methods, empowered by multi-view feature interaction mechanisms, are propelling ECG classification from single-signal analysis toward a paradigm of multidimensional information synergy.

2.2 Plugins utilized in ECG classification

The attention module as a plugin has proved to be helpful for a variety of computer vision tasks such as image classification and image segmentation. One of the successful examples is Senet[10], which simply squeezes each 2D feature map to efficiently build inter-dependencies among channels. CBAM [25] further advances this idea by introducing spatial information encoding via convolutions with large-size kernels. In addition, CA [8] considers a more efficient way of capturing positional information and channel-wise relationships to augment the feature representations for mobile networks.

In recent years, inspired by their computer vision counterparts, attention plugins are increasingly integrated into ECG classification frameworks [1, 12, 15, 18, 20, 26, 29]. For example, Liu et al. [15] innovatively fused a two-channel squeezed-excited residual

neural network with expert features to propose a classification method for inter-patient heartbeats. In addition, Zhou et al. [29] designed an ECG data enhancement method based on Generative Adversarial Networks, which combines bidirectional long short-term memory networks and CBAM to further enhance the overall performance of ECG classification models. Jiang et al. [12] proposed a new CA-based two-branch convolutional neural network to efficiently capture features in the ECG domain by coordinating attention to adaptively assign attentional weights to key segments.

Different from these approaches that leverage expensive and heavy attention blocks, the GFR plugin proposed in this paper considers a more efficient way of assigning weights to different ECG branch features by means of dynamic disease-aware, which not only captures key global information but also significantly emphasizes important features.

3 Method

In the ECG classification task, a commonly utilized approach among researchers is the implementation of the 12-branch network (12BN) for classification, as depicted in Fig. 1. In this architecture, each ECG lead is individually fed into a specific branch of the model. After features are extracted in the backbone network, these features are passed to the feature fusion layer (typically using concatenation and summation operations) to form a comprehensive feature representation. Subsequently, the fused features are fed into the multi-layer perceptron (MLP) layer to perform classification tasks. Additionally, with the application of multi-view learning in the ECG field, new multi-view network (MVN) architectures are introduced. The general structure of this architecture is shown in Fig. 2. In contrast to the original multi-branch network, the input of every branch network in the MVN is no longer solely from a single lead. Instead, the 12-lead ECGs are divided into different views, with the signals of multiple leads corresponding to each view taken as inputs to the network branches. As explained in [7, 18], the study divides the 12-lead ECGs into five views, as illustrated in Fig. 3.

The two generic multi-branch network structures mentioned above consist of two crucial components: the feature extraction layer and the feature fusion layer (also referred to as the feature fusion method). Extensive research has been conducted on the feature extraction layer [4, 7, 13, 17, 19, 20, 27]. Typically, the feature extraction layer produces local features, higher-order features and abstract features. After the feature fusion method the features extracted from different branches of the network are fused and fed into the classification layer for classification. However, we believe that for multi-branch networks, the fusion method should be preceded by enhancing the global features of important branches. Therefore, it is crucial to “re-assign” these branch features prior to the fusion layer. In view of this, we propose an effective and practical global feature refinement plugin named GFR for multi-branch networks used in ECG classification. Its overall structure is depicted in Fig. 4. By further assigning weights to these branching features in a dynamic disease-aware manner before the fusion layer, important branching features can be enhanced and key global information can be

captured. This approach has the potential to significantly improve the classification ability of the classification layer.

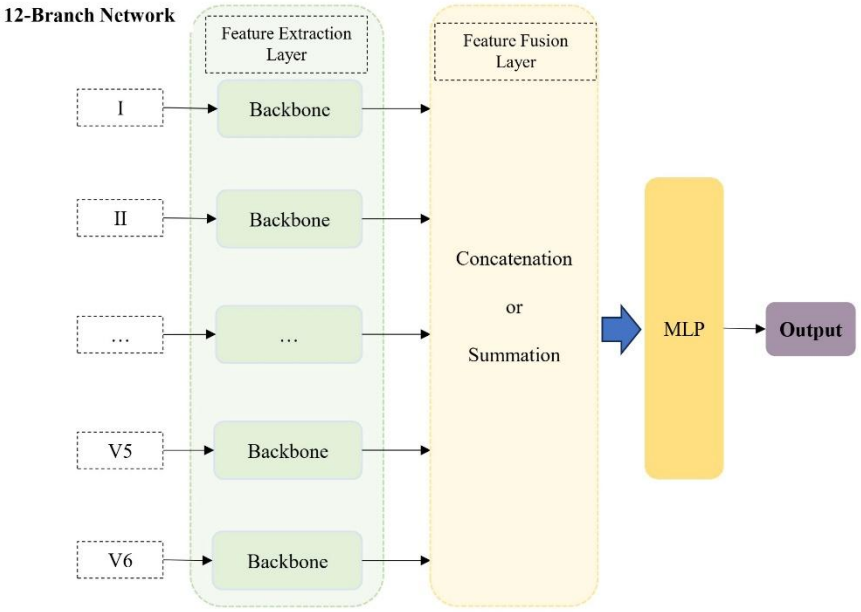

Fig. 1. Overall structure of 12BN framework.

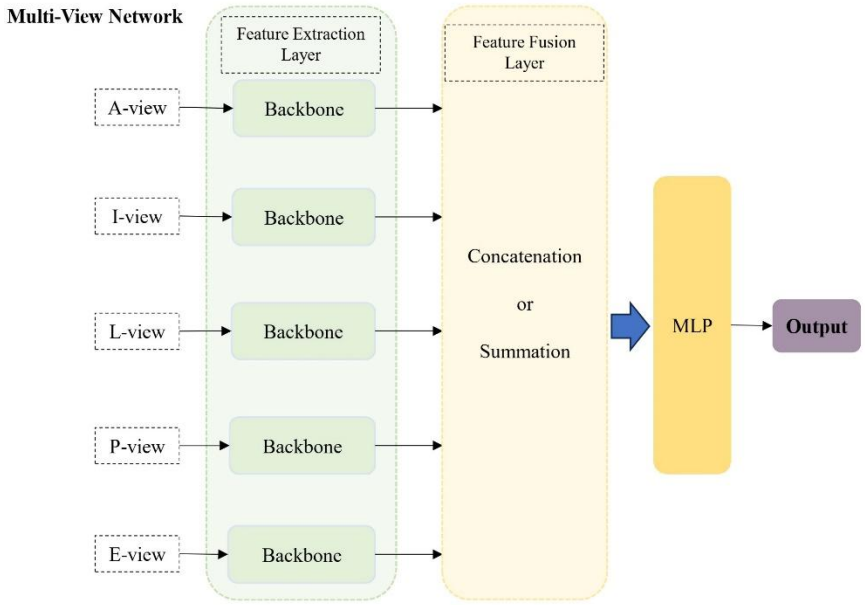

Fig. 2. Overall structure of MVN framework.

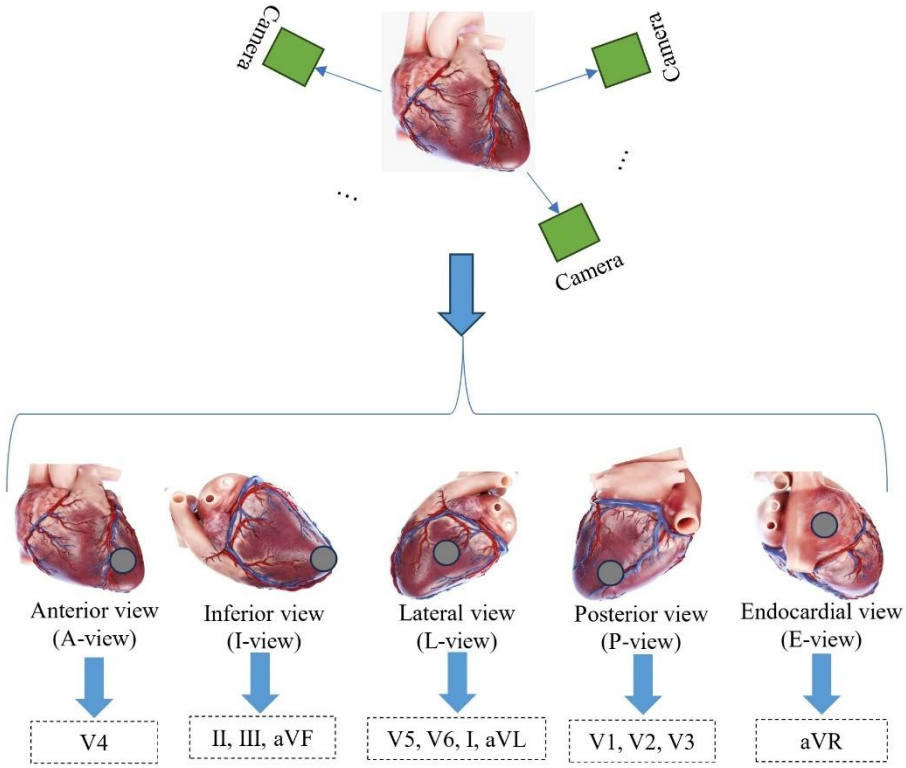

Fig. 3. Division of the 5 views of the 12-lead ECGs according to the correspondence of the different leads with different regions of the heart.

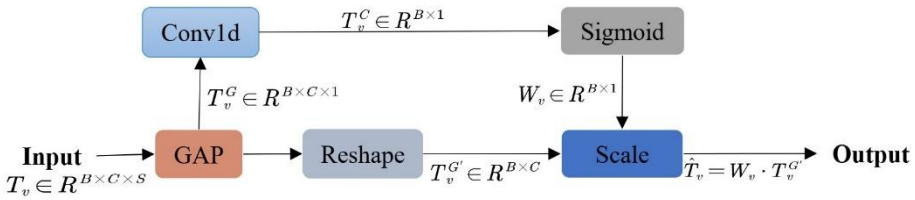

Fig. 4. Overall Structure of GFR.

To address the issue of channel dependency, we first consider the signals for each channel within the output features following the backbone. Within the backbone, each learnable convolutional layer operates on the original input signal using a local receptive field. Consequently, each unit of the output T_v (features obtained after the backbone of each branch, where v represents the v-th branch) is unable to exploit contextual information outside of its own region. This limitation becomes more pronounced in the lower layers of the network, where the receptive field size is comparatively small.

To address this problem, we propose a method that enhances each channel by integrating global temporal context from the entire sequence. This is achieved by applying Global Average Pooling (GAP) along the temporal dimension to compute a summarized representation for each channel. Specifically, we derive channel-wise statistics $T_v^G \in \mathbb{R}^{B\times C\times 1}$ by compressing T_v along the temporal dimension. The expression for T_v^G is defined as follows:

$$T_v^G = F_{GAP}(T_v)$$

where F_{GAP} represents the operation of GAP. Subsequently, to utilize the information extracted from the above processing, we perform a process of global channel embedding, which aims at learning channel information. In this paper, we employ a convolution operation to generate the global channel information. The vector $T_v^C \in R^{B\times 1\times 1}$ is generated by shrinking T_v^G along the channel dimension, and is computed using the following expression:

$$T_v^C = F_{conv}(T_v^G)$$

where F_{conv} represents the operation of convolution.

To leverage fully the information aggregated from previously performed operations, we reconstruct the features. The objective is to select an appropriate function to properly dynamically adjust the weights of the features in each branch. This function should satisfies the requirement of constraining the output within the range of 0 and 1, which helps avoid the problem of imbalanced weights (where one weight may excessively large or small). To fulfill this requirement, we employ a simple yet effective gating mechanism with a sigmoid function. More precisely, we apply a sigma activation function (σ) to the global channel information (T_v^C) for disease-aware dynamic learning of weights. Subsequently, we use the Scale operation to re-scale the reshaped branch feature ($T_v^{G'}$), thereby obtaining the weighted features $\hat{T}_v \in R^{B\times C}$ for each branch. The entire process can be mathematically expressed as follows:

$$W_v = \sigma(T_v^C)$$

$$\hat{T}_v = W_v \cdot T_v^{G'}$$

Finally, the enhanced features of each branch are fed into the subsequent fusion layer for feature fusion.

4 Experiments

4.1 Basic Setting

All the code used in this paper is implemented using PyTorch framework. The experiments were conducted on an Inter I7-13700K server equipped with NVIDIA RTX 4090 24GB GPU. To ensure comparability of the experimental results, the following parameters were set uniformly across all networks, with a batch size of 64, a fixed learning rate of 0.001 for 200 epochs, an Adam optimizer, and an early stop strategy. The other parameters remain the same as those used in the original papers.

4.2 Backbones

To assess the effectiveness of GFR, we employ two network frameworks: the 12BN and the MVN. We selectively utilized nine models (including backbones, fusion methods, and MLP layer) with distinct sizes, with the aim of ensuring both diversity and comprehensiveness in our experiments. The selected models encompassed Inception1d [11], Resnet50/101/152 [6], MCA-net [17], ATI-CNN [27], MobileNetV3 [9], VGG16 [22] and ECGMamba [19].

4.3 Datasets and Tasks

We performed a large number of inter-patient ECG classification experiments to validate whether GFR improves the overall model performance using three large unbalanced datasets, namely PTB-XL [24], CPSC2018 [14] and Chapman [28]. We conducted two multi-class classification experiments of the PTB-XL dataset: super-diag (5-class classification task) and MI localization (6-class classification task). A 9-class classification task was performed on the CPSC2018 dataset [14]. Two classification tasks were performed on the Chapman dataset [28]: a 4-class classification task and an 8-class classification task.

5 Results

In this paper, we employ the following metrics to assess the performance of models: accuracy (Acc), Area Under the Curve (AUC), and F1-score ($F1$). For the sake of clarity and simplicity in representation, we use the notation "Model-GFR" to indicate that the GFR is employed after the feature extraction process of the Model. For instance, "Inception1d-GFR" signifies that Inception1d serves as the model, and the GFR is utilized subsequently. Tables 1 to 3 present the experimental results on the PTB-XL, Chapman, and CPSC2018 databases. As shown in these Tables, the models of GFR used in this study have excellent overall performance. The following subsections analyze in detail the results obtained on each dataset using GFR as a feature refinement method in the multi-classification task scenario.

5.1 Experimental Results on the PTB-XL dataset

Table 1 clearly illustrates the performance metrics of different models in the experiments conducted for the multi-classification task on the PTB-XL dataset. Considering the performance metrics and the number of parameters of the various models, it is clear that the models using the GFR plugin exhibit superior overall performance.

As a case study for detailed explanation, we take the MI localization task as an example. Table 1 show that the introduction of the GFR plugin significantly improves the performance of multiple models in the task of MI localization. For example, on the MVN framework, the improved versions of GFR for all baseline models achieve improvements in the three metrics of AUC, F1 and Acc. Among them, Resnet50-GFR exhibits the most significant improvement, with its F1 increasing from 56.60% to 64.87% (+8.27% absolute improvement) and AUC increasing from 73.72% to 76.86% (+3.14%). Deep networks (e.g., the Resnet family) generally benefit from the GFR

plugin, with the AUC of Resnet101 and Resnet152 improving by 2.49% and 3.50%, respectively, on the MVN framework, and the number of parameters remaining essentially unchanged. On the 12BN framework, the GFR plugin remains effective. For example, the incorporation of GFR in Inception1d achieves a 3.06% AUC improvement and 2.26% F1 improvement.

Table1: Performance comparison on PTB-XL dataset for MI localization and super-diag tasks.

MI localization task								
Model	MVN				12BN			
	AUC(%)	F1(%)	Acc(%)	Param.	AUC(%)	F1(%)	Acc(%)	Param.
Inception1d	77.44	60.72	82.21	2.84M	75.42	58.50	82.61	5.67M
Inception1d-GFR	**79.42**	**64.99**	**84.31**	2.84M	**78.48**	**60.76**	**83.11**	5.67M
Resnet50	73.72	56.60	80.88	43.35M	75.77	57.36	81.45	86.66M
Resnet50-GFR	**76.86**	**64.87**	**83.35**	43.35M	**77.26**	**60.17**	**81.83**	86.66M
Resnet101	76.25	60.49	82.66	83.56M	77.36	60.45	81.51	167.09M
Resnet101-GFR	**78.74**	**62.85**	**83.04**	83.56M	**77.86**	**61.94**	**83.35**	167.09M
MCA-net	75.86	57.84	81.39	81.77k	78.19	62.62	82.08	164.26k
MCA-net-GFR	**77.46**	**59.36**	**81.57**	82.12k	**78.77**	**62.80**	**82.45**	165.04k
Resnet152	74.03	56.32	82.15	116.67M	77.89	61.53	82.59	233.33M
Resnet152-GFR	**77.53**	**60.19**	**82.61**	116.67M	**78.26**	**62.49**	**82.74**	233.33M
ATI-CNN	74.86	59.01	82.91	30.00M	75.86	59.15	81.70	59.99M
ATI-CNN-GFR	**78.72**	**62.61**	**83.04**	30.00M	**76.56**	**60.23**	**81.83**	59.99M
MobileNetV3	74.40	56.92	81.00	24.83M	72.35	52.40	79.99	49.60M
MobileNetV3-GFR	**75.23**	**57.03**	**82.00**	24.83M	**73.24**	**54.92**	**82.47**	49.60M
VGG16	75.66	57.68	80.50	31.89M	76.09	59.47	82.15	63.74M
VGG16-GFR	**78.01**	**62.36**	**83.38**	31.89M	**76.12**	**59.48**	**82.37**	63.74M
ECMamba	75.71	57.81	80.69	20.15M	76.29	59.61	82.32	48.40M
ECMamba-GFR	**75.87**	**58.02**	**81.02**	20.15M	**76.35**	**59.69**	**82.41**	48.40M
super-diag task								
Inception1d	92.73	73.96	59.45	2.84M	92.38	73.00	59.36	5.67M
Inception1d-GFR	**92.81**	**74.39**	**61.58**	2.84M	**92.68**	**74.27**	**62.70**	5.67M
Resnet50	90.31	72.04	58.08	43.35M	89.78	71.47	56.35	86.66M
Resnet50-GFR	**90.85**	**72.20**	**58.57**	43.35M	**92.48**	**72.73**	**62.42**	86.66M
Resnet101	90.87	71.68	56.77	83.56M	90.73	71.48	60.43	167.09M
Resnet101-GFR	**91.90**	**72.98**	**60.80**	83.56M	**92.02**	**72.89**	**61.43**	167.09M
MCA-net	91.84	73.06	59.13	81.77K	92.04	73.61	59.96	164.26k
MCA-net-GFR	**92.26**	**74.24**	**61.72**	82.12K	**92.29**	**73.95**	**60.10**	165.04k
Resnet152	90.58	72.64	58.24	116.67M	89.93	71.06	56.81	233.33M
Resnet152-GFR	**91.95**	**73.72**	**61.31**	116.67M	**92.00**	**73.83**	**61.08**	233.33M
ATI-CNN	91.37	72.32	**61.42**	30.00M	91.48	72.17	58.31	59.99M
ATI-CNN-GFR	**91.71**	**73.19**	61.28	30.00M	**91.82**	**73.79**	**58.76**	59.99M
MobileNetV3	90.97	72.02	58.43	24.83M	90.06	71.33	60.29	49.60M
MobileNetV3-GFR	**91.26**	**72.15**	**59.13**	24.83M	**90.57**	**71.43**	**61.39**	49.60M
VGG16	92.49	74.07	62.23	31.89M	91.78	72.88	62.09	63.74M
VGG16-GFR	**92.63**	**74.86**	**63.67**	31.89M	**92.14**	**73.91**	**62.24**	63.74M
ECGMamba	92.59	74.24	62.34	20.15M	91.86	72.96	62.16	48.40M
ECMamba-GFR	**92.68**	**74.36**	**62.41**	20.15M	**91.99**	**73.15**	**62.24**	48.40M

5.2 Experimental results on the Chapman dataset

Table 2 clearly illustrate the results in the eight-class and four-class classification tasks conducted on the Chapman dataset. From the observation of Table 2, it is evident that the models using the GFR plugin achieve the best overall performance in both ECG classification tasks on the Chapman dataset.

Table2: Performance comparison on Chapman dataset for eight-class and four-class classification tasks.

Model	eight-class classification task							
	MVN				12BN			
	AUC(%)	F1(%)	Acc(%)	Param.	AUC(%)	F1(%)	Acc(%)	Param.
Inception1d	88.46	77.66	90.68	2.84M	87.74	77.02	90.49	5.67M
Inception1d-GFR	**90.19**	**82.65**	**92.66**	2.84M	**90.83**	**82.91**	**93.88**	5.67M
Resnet50	88.70	79.98	92.56	43.35M	88.81	79.96	93.13	86.66M
Resnet50-GFR	**89.46**	**80.46**	**93.90**	43.35M	**92.09**	**84.39**	**93.69**	86.66M
Resnet101	88.17	**79.63**	92.84	83.56M	86.75	73.87	90.11	167.09M
Resnet101-GFR	**88.37**	79.59	**92.84**	83.56M	**89.84**	**80.25**	**93.13**	167.09M
MCA-net	84.51	71.00	88.14	81.77K	**85.52**	**71.37**	87.29	164.26k
MCA-net-GFR	**84.60**	**71.26**	**90.58**	82.12K	85.24	71.20	**89.28**	165.04k
Resnet152	85.90	73.65	90.77	116.67M	87.41	76.18	91.53	233.33M
Resnet152-GFR	**89.04**	**79.88**	**91.71**	116.67M	**90.39**	**81.72**	**92.75**	233.33M
ATI-CNN	88.77	79.67	92.47	30.00M	88.71	80.18	92.47	59.99M
ATI-CNN-GFR	**88.93**	**79.86**	**92.80**	30.00M	**89.69**	**80.25**	**93.34**	59.99M
MobileNetV3	**87.30**	**76.69**	92.28	24.83M	87.18	77.51	90.87	49.60M
MobileNetV3-GFR	87.29	76.46	**92.53**	24.83M	**87.90**	**78.24**	**92.00**	49.60M
VGG16	89.16	79.68	90.96	31.89M	86.92	76.35	91.81	63.74M
VGG16-GFR	**89.44**	**81.47**	**92.84**	31.89M	**88.39**	**79.33**	**92.28**	63.74M
ECGMamba	88.89	79.54	90.46	20.15M	86.76	76.26	90.87	48.40M
ECMamba-GFR	**88.97**	**79.68**	**90.78**	20.15M	**86.85**	**76.49**	**91.12**	48.40M
	four-class classification task							
Inception1d	97.79	96.67	97.05	2.84M	97.98	96.86	97.24	5.67M
Inception1d-GFR	**97.93**	**96.76**	**97.05**	2.84M	**98.13**	**97.13**	**97.43**	5.67M
Resnet50	97.82	96.14	97.05	43.35M	97.93	96.75	**97.14**	86.66M
Resnet50-GFR	**97.85**	**96.75**	**97.05**	43.35M	**97.98**	**96.82**	96.76	86.66M
Resnet101	97.59	96.31	96.67	83.56M	98.19	97.25	97.52	167.09M
Resnet101-GFR	**97.78**	**96.63**	**97.19**	83.56M	**98.25**	**97.31**	**97.53**	167.09M
MCA-net	96.67	94.96	95.43	81.77K	96.77	95.03	95.52	164.26k
MCA-net-GFR	**97.39**	**95.94**	**95.90**	82.12k	**97.25**	**95.86**	**96.29**	165.04k
Resnet152	97.38	96.05	97.33	116.67M	97.43	95.98	96.38	233.33M
Resnet152-GFR	**97.82**	**96.64**	**97.85**	116.67M	**98.04**	**97.08**	**97.33**	233.33M
ATI-CNN	97.42	95.93	96.29	30.00M	97.66	96.28	96.67	59.99M
ATI-CNN-GFR	**98.65**	**97.51**	**97.71**	30.00M	**98.57**	**97.44**	**97.71**	59.99M
MobileNetV3	97.46	96.05	96.48	24.83M	97.08	95.45	95.90	49.60M
MobileNetV3-GFR	**97.91**	**96.83**	**97.14**	24.83M	**98.23**	**97.21**	**97.52**	49.60M
VGG16	97.58	96.93	97.24	31.89M	97.72	96.54	96.86	63.74M
VGG16-GFR	**98.07**	**97.05**	**97.33**	31.89M	**98.10**	**96.96**	**97.45**	63.74M
ECGMamba	97.54	96.87	97.15	20.15M	97.64	96.49	96.79	48.40M
ECMamba-GFR	**97.89**	**97.01**	**97.34**	20.15M	**97.89**	**96.81**	**97.12**	48.40M

As a case study for detailed explanation, we take the eight-classification task as an example. In the eight-classification task, the GFR plug-in demonstrates significant performance gains. For example, Inception1d-GFR achieves a comprehensive breakthrough on the 12BN framework: a 5.89% (77.02% to 82.91%) improvement in F1, a 3.09% improvement in AUC, and a 3.39% improvement in accuracy with a negligible increase in the number of parameters. The deep network improvement is particularly prominent, with Resnet152-GFR improving F1 score by 6.23% (73.65% to 79.88%) and AUC by 3.14% on the MVN framework, which validates GFR's ability to feature refinement for complex classification tasks. GFR also exhibits strong potential for improvement on the 12BN framework. For example, Resnet50-GFR improves F1 by

4.43% and AUC by 3.28%. In particular, Resnet101-GFR had an F1 improvement of 6.38% (73.87% to 80.25%) at 12BN framework.

5.3 Experimental results on the CPSC2018 dataset

Table 3 clearly demonstrates the model performance in the 9-class classification task experiments on the CPSC2018 dataset. As shown in Table 3, the models using the GFR plugin achieve the best overall performance.

In the nine-classification task, the GFR plugin significantly improves the ability of the multi-model to discriminate between fine-grained ECG features. Inception1d-GFR achieves a cross-metric breakthrough in the 12BN framework: a 6.23% improvement in F1 (74.80% to 81.03%) and a 6.26% increase in accuracy (69.10% to 75.36%) while the number of covariates is maintained at 5.67M unchanged. The improvement of Resnet50-GFR on the same dataset is particularly impressive, with a 5.74% (74.55% to 80.29%) F1 improvement along with a 6.41% (70.70% to 77.11%) accuracy growth, validating the feature aggregation advantage of GFR for mid-layer convolutional networks. The deep network Resnet152-GFR showed an F1 improvement of 1.77% in the 12BN framework, but the MVN framework showed a slight decrease in accuracy of 0.37%.

Table3: Performance of various models on the 9-class classification task of the CPSC2018 dataset.

Model	MVN				12BN			
	AUC(%)	F1(%)	Acc(%)	Param.	AUC(%)	F1(%)	Acc(%)	Param.
Inception1d	95.30	77.36	74.20	2.84M	94.50	74.80	69.10	5.67M
Inception1d-GFR	**95.59**	**80.55**	**75.51**	2.84M	**95.63**	**81.03**	**75.36**	5.67M
Resnet50	96.07	78.59	**75.53**	43.35M	95.09	74.55	70.70	86.66M
Resnet50-GFR	**96.23**	**78.95**	75.50	43.35M	**96.57**	**80.29**	**77.11**	86.66M
Resnet101	96.04	79.27	77.73	83.56M	94.34	72.51	69.39	167.09M
Resnet101-GFR	**96.51**	**80.09**	**77.55**	83.56M	**96.15**	**78.93**	**77.41**	167.09M
MCA-net	92.99	71.51	64.72	81.77K	93.23	72.08	66.47	164.26k
MCA-net-GFR	**93.76**	**73.71**	**67.06**	82.12K	**93.74**	**72.62**	**66.76**	165.04k
Resnet152	95.81	78.89	**76.68**	116.67M	95.49	78.50	74.20	233.33M
Resnet152-GFR	**95.90**	**78.91**	76.31	116.67M	**96.06**	**80.27**	**78.86**	233.33M
ATI-CNN	95.73	79.11	77.70	30.00M	95.45	77.70	73.91	59.99M
ATI-CNN-GFR	**96.32**	**79.61**	**79.15**	30.00M	**96.20**	**82.21**	**79.45**	59.99M
MobileNetV3	95.98	75.66	73.34	24.83M	94.66	75.97	70.99	49.60M
MobileNetV3-GFR	**96.06**	**75.71**	**73.47**	24.83M	**96.19**	**78.94**	**76.23**	49.60M
VGG16	96.35	78.83	76.53	31.89M	95.36	77.55	72.89	63.74M
VGG16-GFR	**96.54**	**80.46**	**78.78**	31.89M	**96.51**	**81.22**	**78.57**	63.74M
ECGMamba	92.01	73.86	61.98	20.15M	94.97	76.59	71.36	48.40M
ECMamba-GFR	**92.32**	**73.98**	**62.13**	20.15M	**95.12**	**76.87**	**71.81**	**48.40M**

In the 12BN framework, GFR shows stronger optimization potential: the F1 improvement of ATI-CNN-GFR is 4.51% (77.70% to 82.21%) and the accuracy improvement is 5.54% (73.91% to 79.45%), which are both significantly higher than its gains in MVN (F1:+0.50%, Acc:+1.45%). The lightweight model MCA-net-GFR has an F1 gain of 2.20% and an accuracy gain of 2.34% in MVN, proving that GFR is still adaptable to small-scale networks. VGG16-GFR has an F1 gain of 1.63% and 3.67% respectively in the two frameworks (MVN and 12BN), showing its stable improvement ability for different multi-branch frameworks. In addition, MobileNetV3-GFR has only a

slight F1 increase of 0.05% in MVN, but achieves a F1 improvement of 2.97% in 12BN, suggesting that the model branching complexity affects the plugin utility to some extent.

Experiments demonstrate that the GFR plugin significantly enhances the model's capability to capture global discriminative features across diverse network frameworks. Notably, it introduces only marginal parameter increments (less than 0.8k) while maintaining computational efficiency, thereby avoiding substantial overhead to the baseline models.

6 Discussion

6.1 The impact of different plugin

In this section, we provide an in-depth discussion on the performance of the proposed GFR plugin in comparison with three popular plugins, namely, SEnet, CBAM, and CA, in multi-branch ECG classification models. The GFR plugin captures the key global information through GAP and convolutional layers to capture key global information and enhance important branching features through a weighting mechanism, whereas SEnet, CBAM, and CA focus on modelling inter-channel relationships, considering both channel and spatial information, and enhancing the attention mechanism using coordinate information, respectively.

Analyses based on experimental results showing in Table 4 demonstrate that the GFR plug-in exhibits significant performance advantages in the myocardial infarction localization task. Taking the Inception1d model as an example, under the MVN framework, compared with the three mainstream attentional mechanisms of SEnet, CBAM, and CA, GFR achieves a comprehensive breakthrough in key indicators: the AUC is improved by 1.12%, 1.42%, and 1.22%, and the F1 is improved by amounted to 3.59%, 3.79%, and 3.99%, and accuracy improved by 0.75%, 1.86%, and 0.86%, respectively. In the 12BN framework, the advantages of GFR are further highlighted: the AUC improvement extends to 2.13%, 2.38%, and 2.23%, and the F1 scores improve by 1.56%, 1.86%, and 1.76%, and a steady increase in accuracy. It is worth noting that GFR achieves the above performance breakthroughs while the number of parameters is only equivalent to 1/15 of the traditional attention mechanism (0.65k under the MVN framework vs. 10k in SEnet), showing excellent parameter efficiency, a feature that has important application value in model lightweight deployment scenarios.

In addition, we analyze the adaptations and limitations of each plugin in different application scenarios. the GFR plugin is particularly suitable for multi-branch models that require global feature refinement, and it especially excels in sequence data tasks such as ECG classification. While SEnet, CBAM and CA are more advantageous in tasks such as image processing, they may be insufficient in ECG classification which has complex temporal information.

Table 4: Comparison of model performance using different plugins in the MI localization task on the PTB-XL dataset. `+' represents the addition of different plugins to the base model, and `↑' represents the number of parameters added to the base model by adding different plugins.

Model	MVN				12BN			
	AUC(%)	F1(%)	Acc(%)	Param.	AUC(%)	F1(%)	Acc(%)	Param.
Inception1d	77.44	60.72	82.21	2.84M	75.42	58.50	82.61	5.67M
+SEnet	78.30	61.40	83.56	↑10k	76.35	59.20	82.78	↑30k
+CBAM	78.00	61.20	82.45	↑10k	76.10	58.90	82.76	↑20k
+CA	78.20	61.00	83.45	↑20k	76.25	59.00	82.98	↑40k
+GFR	**79.42**	**64.99**	**84.31**	↑0.65k	**78.48**	**60.76**	**83.11**	↑1.55k
MCA-net	75.86	57.84	81.39	81.77K	78.19	62.62	82.08	164.26K
+SEnet	76.50	58.10	81.13	↑2.90K	78.20	62.40	82.16	↑6.96K
+CBAM	76.40	58.30	81.27	↑2.57K	78.15	62.15	82.06	↑6.17K
+CA	76.70	58.50	**81.67**	↑8.40K	78.35	62.50	82.12	↑2.02K
+GFR	**77.46**	**59.36**	81.57	↑0.35k	**78.77**	**62.80**	**82.45**	↑0.78k
Resnet152	74.03	56.32	82.15	116.67M	77.89	61.53	82.59	233.33M
+SEnet	75.10	57.20	82.26	↑170k	78.15	61.90	82.68	↑400k
+CBAM	74.80	56.90	81.96	↑160k	78.02	61.55	82.26	↑400k
+CA	74.95	57.00	82.18	↑250k	78.12	61.60	82.62	↑600k
+GFR	**77.53**	**60.19**	**82.61**	↑2.57k	**78.26**	**62.49**	**82.74**	↑6.16k
ATI-CNN	74.86	59.01	82.91	30.00M	75.86	59.15	81.70	59.99M
+SEnet	75.74	59.31	82.78	↑170k	76.40	59.65	81.53	↑400k
+CBAM	75.61	59.52	82.46	↑160k	76.20	59.40	81.42	↑400k
+CA	75.82	59.87	82.93	↑250k	76.65	59.75	81.31	↑600k
+GFR	**78.72**	**62.61**	**83.04**	↑2.57k	**76.56**	**60.23**	**81.83**	↑6.16k

6.2 The impact of different components of GFR

To verify the rationality of the GFR design, we carried out ablation experiments to analyze the effects of the core size and activation function on the GFR performance. The experimental results are shown in Table 5. Our experimental results show that: when changing the kernel size in the convolution operation (e.g., using kernel sizes of 1, 3, and 5, which are denoted as w/k1, w/k3 and w/k5 in the Table 5), the model performance shows a decreasing trend, and the model achieves the optimal accuracy, F1 score, and other metrics on the ECG classification task when the kernel size is 1, which suggests that simple kernel sizes are able to achieve information retention, and too large is prone to cause model complexity to be increased; replacing the Sigmoid activation After replacing the Sigmoid activation function with common activation functions such as ReLU and Tanh (denoted as w/Relu and w/Tanh in the Table 5), the performance of the model decreases to different degrees, which verifies the importance and advantage of the Sigmoid activation function in generating reasonable weights.

Table 5: The impact of different components of GFR on model performance in the MI localization task on the PTB-XL dataset.

Model	MVN			12BN		
	AUC(%)	F1(%)	Acc(%)	AUC(%)	F1(%)	Acc(%)
Inception1d	77.44	60.72	82.21	75.42	58.50	82.61
w/k1	**79.42**	**64.99**	**84.31**	**78.48**	**60.76**	**83.11**
w/k3	76.85	59.12	81.93	74.60	57.32	81.45
w/k5	76.20	58.45	81.25	74.15	56.80	80.97
w/Relu	75.62	58.42	80.89	73.95	56.25	80.13
w/Tanh	75.30	57.91	80.45	73.60	55.80	79.85
MCA-Net	75.86	57.84	81.39	78.19	62.62	82.08
w/k1	**77.46**	**59.36**	**81.57**	**78.77**	**62.80**	**82.45**
w/k3	76.40	58.90	81.20	75.80	59.45	81.62
w/k5	75.75	58.10	80.55	75.25	58.90	81.03
w/Relu	75.10	57.85	80.10	74.85	58.30	80.45
w/Tanh	74.65	57.20	79.75	74.40	57.85	79.90
Resnet152	74.03	56.32	82.15	77.89	61.53	82.59
w/k1	**77.53**	**60.19**	**82.61**	**78.26**	**62.49**	**82.74**
w/k3	75.35	57.80	82.05	76.90	60.45	82.10
w/k5	74.70	56.95	81.30	76.25	59.80	81.55
w/Relu	74.15	56.40	80.85	75.80	59.25	81.10
w/Tanh	73.80	55.90	80.40	75.35	58.70	80.65
ATI-CNN	74.86	59.01	82.91	75.86	59.15	81.70
w/k1	**78.72**	**62.61**	**83.04**	**76.56**	**60.23**	**81.83**
w/k3	75.20	59.12	82.30	75.60	58.90	81.45
w/k5	74.55	58.25	81.65	75.05	58.15	80.90
w/Relu	74.00	57.80	81.10	74.60	57.60	80.35
w/Tanh	73.65	57.25	80.75	74.25	57.10	79.95

7 Conclusion

In this paper, we propose a plugin GFR to enhance the performance of multi-branch models for ECG classification. The GFR plugin assigns weights to different branching features in a dynamic disease-aware manner to capture critical global information while emphasizing important features. Specifically, these dynamic weights are obtained through the integration, mapping, and scaling of global features. Finally, the weighted features are summed for ECG classification. Extensive experiments on three large-scale imbalanced datasets demonstrate that the GFR plugin, with less 6.2k additional parameters, improves the performance of eight models of different sizes to varying degrees.

Although GFR has produced favorable results, certain limitations should be acknowledged. Firstly, there may be potential discrepancies between the experimental results and the real clinical setting due to the influence of accurate annotations by cardiologists. Therefore, further validation and testing of the method's practical utility in real clinical settings is necessary. Secondly, this paper focuses on inter-patient classification of three publicly available ECG datasets. It is important to note that the method may have limitations when dealing with other types of ECG analyses. Additionally, the method may face challenges in terms of interpretability, as it may be difficult to explain decisions to clinicians.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (No. 62450100); the Natural Science Foundation of Guangdong Province, China (No.

2024A1515220144); the National Natural Science Foundation of China (Nos. 82472110 and U20A20224).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Chen, J., Chen, T., Xiao, B., et al.: Se-ecgnet: multi-scale se-net for multi-lead ecg data. In: 2020 Computing in Cardiology. pp. 1–4. IEEE (2020)
2. Deng, M., Chen, K., Huang, D., et al.: An intelligent computer-aided diagnosis method for paroxysmal atrial fibrillation patients with nondiagnostic ecg signals. Biomedical Signal Processing and Control 88, 105683 (2024)
3. Guan, Y., An, Y., Guo, F., et al.: Mpfnet: Ecg arrhythmias classification based on multi-perspective feature fusion. In: International Symposium on Bioinformatics Research and Applications. pp. 85–96. Springer (2023)
4. Han, C., Shi, L.: Ml–resnet: A novel network to detect and locate myocardial infarction using 12 leads ecg. Computer methods and programs in biomedicine 185, 105138 (2020)
5. Hao, P., Yin, X., Wu, F., et al.: A novel feature fusion network for myocardial infarction screening based on ecg images. In: Image and Graphics: 11th International Conference, ICIG 2021, Haikou, China, August 6–8, 2021, Proceedings, Part II 11. pp. 547–558. Springer (2021)
6. He, K., Zhang, X., Ren, S., et al.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
7. He, Z., Yuan, S., Zhao, J., et al.: A robust myocardial infarction localization system based on multi-branch residual shrinkage network and active learning with clustering. Biomedical Signal Processing and Control 80, 104238 (2023)
8. 8. Hou, Q., Zhou, D., Feng, J.: Coordinate attention for efficient mobile network design. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 13713–13722 (2021)
9. Howard, A., Sandler, M., Chu, G., et al.: Searching for mobilenetv3. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 1314–1324 (2019)
10. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 7132–7141 (2018)
11. Ismail Fawaz, H., Lucas, B., Forestier, G., et al.: Inceptiontime: Finding alexnet for time series classification. Data Mining and Knowledge Discovery 34(6), 1936–1962 (2020)
12. Jiang, R., Fu, B., Li, R., et al.: A dual-branch convolutional neural network with domain-informed attention for arrhythmia classification of 12-lead electrocardiograms. Engineering Applications of Artificial Intelligence 139, 109480 (2025)
13. Jyotishi, D., Dandapat, S.: An attentive spatio-temporal learning-based network for cardio-vascular disease diagnosis. IEEE Transactions on Systems, Man, and Cybernetics: Systems (2023)
14. Liu, F., Liu, C., Zhao, L., et al.: An open access database for evaluating the algorithms of ecg rhythm and morphology abnormality detection. Journal of Medical Imaging and Health Informatics 8(7), 1368–1373 (2018)
15. Liu, J., Liu, Y., Jin, Y., et al.: A novel diagnosis method combined dual-channel se-resnet with expert features for interpatient heartbeat classification. Medical Engineering & Physics 130, 104209 (2024)

16. Ma, Z., Wang, J., Yue, J., et al.: A homologous and heterogeneous multi-view inter-patient adaptive network for arrhythmia detection. Computer Methods and Programs in Biomedicine 241, 107740 (2023)
17. Pan, W., An, Y., Guan, Y., et al.: Mca-net: A multi-task channel attention network for myocardial infarction detection and location using 12-lead ecgs. Computers in Biology and Medicine 150, 106199 (2022)
18. Qiang, Y., Dong, X., Liu, X., et al.: Mt-mv-kdf: A novel multi-task multi-view knowledge distillation framework for myocardial infarction detection and localization. Biomedical Signal Processing and Control 95, 106382 (2024)
19. Qiang, Y., Dong, X., Liu, X., et al.: Ecgmamba: Towards ecg classification with state space models. In: 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp. 6498–6505. IEEE (2024)
20. Qiang, Y., Dong, X., Yang, Y.: Automatic detection and localisation of myocardial infarction using multi-channel dense attention neural network. Biomedical Signal Processing and Control 89, 105766 (2024)
21. Sakli, N., Ghabri, H., Zouinkh, I.A., et al.: An efficient deep learning model to predict cardiovascular disease based on ecg signal. In: 2022 19th International Multi-Conference on Systems, Signals & Devices (SSD). pp. 1759–1763. IEEE (2022)
22. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
23. Su, K., Yang, G., Wu, B., et al.: Human identification using finger vein and ecg signals. Neurocomputing 332, 111–118 (2019)
24. Wagner, P., Strodthoff, N., Bousseljot, R.D., et al.: Ptb-xl, a large publicly available electrocardiography dataset. Scientific data 7(1), 154 (2020)
25. Woo, S., Park, J., Lee, J.Y., et al.: Cbam: Convolutional block attention module. In: Proceedings of the European conference on computer vision (ECCV). pp. 3–19 (2018)
26. Yang, Z., Jin, A., Li, Y., et al.: A coordinated adaptive multiscale enhanced spatio-temporal fusion network for multi-lead electrocardiogram arrhythmia detection. Scientific Reports 14(1), 20828 (2024)
27. Yao, Q., Wang, R., Fan, X., et al.: Multi-class arrhythmia detection from 12-lead varied-length ecg using attention-based time-incremental convolutional neural network. Information Fusion 53, 174–182 (2020)
28. Zheng, J., Zhang, J., Danioko, S., et al.: A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients. Scientific data 7(1), 48 (2020)
29. Zhou, F., Li, J.: Ecg data enhancement method using generate adversarial networks based on bi-lstm and cbam. Physiological Measurement 45(2), 025003 (2024)