ITM Web of Conferences **78**, 02011 (2025) https://doi.org/10.1051/itmconf/20257802011
CSEIT 2025

Research on Classification of Electrocardiogram Signals Based on Improved Frequency Sliced Wavelet Transform

Xuyang Su

Department of Integrated Circuit, Nanjing University, Nanjing, China

Abstract. Aiming at the problems such as poor adaptability of basis function, high noise sensitivity and feature redundancy in ECG classification, this paper proposes an adaptive feature extraction framework based on improved Frequency Slice Wavelet Transform (FSWT). It includes dynamic band adjustment model and morphological adaptive kernel function, which can be used to preprocess electrocardiograph (ECG) image data to make data have time-frequency characteristics and remove noise at the same time. The experiment was based on the MIT-BIH database for data preprocessing, and the model combined with convolutional neural network and recurrent neural network was trained. The traditional image data preprocessing methods such as image translation, rotation and cropping were compared with the methods adopted in this experiment. Finally, it is concluded that the improved frequency slice wavelet transform can improve the accuracy and anti-interference ability of the recognition of ECG signals based on convolutional neural network, and improve its robustness, which is suitable for mobile wearable ECG detection and other medical applications.

1 Introduction

1.1 Background

Early diagnosis of cardiovascular disease depends on accurate identification of arrhythmia features in electrocardiogram (ECG) signals. Although deep learning has made progress in ECG classification, its black box characteristics are difficult to meet the interpretability needs of clinical diagnosis, and its poor ability to recognize noise, resulting in low accuracy in signal recognition, which is difficult to fully meet clinical diagnosis requirements. As a classical signal processing method, time-frequency analysis provides a clear feature basis of physical significance for arrhythmia classification by combining time-frequency domain features, but faces three technical bottlenecks: (1) the fixed basis function of traditional wavelet transform is difficult to match the non-stationary characteristics of ECG; (2) The low-amplitude characteristics of P wave /T wave are easily submerged by baseline drift; (3)

s18863721958@163.com

The high-dimensional characteristics of time-frequency matrix lead to the risk of feature redundancy and overfitting.[1][4]

1.2 Challenges

Frequency slice wavelet transform (FSWT) can directly move both frequency domain and time domain Windows by freely defining frequency slice function, overcomes the shortcomings of wavelet transform, and shows better time-frequency resolution adjustment ability than STFT and CWT[2]. However, most of the existing studies use fixed parameter Settings: the use of global Gaussian window leads to the loss of QRS group details[2]; The linear frequency modulation slicing scheme is not stable in noisy environment.[4] These methods ignore the dynamic correlation between ECG waveform morphological features and frequency band parameters, which restricts the further improvement of classification performance.

1.3 Innovation of this paper

This paper proposes an improved framework of adaptive FSWT for ECG classification. The core innovations include: (1) dynamic parameter mapping mechanism: establishing a mathematical model of QRS slop-bandwidth-window function parameters to achieve feature-sensitive dynamic adjustment of time-frequency resolution;[8] (2) Driver kernel function design: Based on the difference in amplitude and frequency characteristics of each ECG band, segmented composite window function is constructed to improve the low frequency feature retention ability of P/T wave interval[3].

1.4 Contribution and breakthrough

The MIT-BIH arrhythmia database (in accordance with AAMI EC57 standard) was used in the study[7], covering 5 common arrhythmias such as ventricular premature beat (PVC) and atrial premature beat (APC). Compared with the existing methods[1], it is shown that this scheme can improve the recognition accuracy of each disorder by 0.3% and f1score by 1.1%, which can meet the needs of real-time monitoring. This study provides a new theoretical tool for the analysis of interpretable ECG and has important engineering alue for the research and development of mobile medical devices.

2. Methodology

2.1 ECG dataset

The MIT-BIH Arrhythmia Database[7], which was constructed using data collected collaboratively by the Massachusetts Institute of Technology and Beth Israel Hospital between 1975 and 1979, is one of three internationally recognized ECG databases that serve as clinical benchmarks. The database comprises measurement data for 47 subjects, with each record lasting approximately 30 minutes. Approximately 30 percent of these records exhibit abnormal heart rhythms. In total, the database includes 13 distinct types of heartbeats, as illustrated in Table 1.

Table 1 MIT-BIH Arrhythmia Database

Beat type	Abbreviation	Symbol
Normal Sinus	NOR	N

Rhythm		
Premature Ventricular Contraction	PVC	P
Atrial Premature Contraction	APC	A
Ventricular Ectopic Beat	VEB	V
Left Bundle Branch Block	LBB	L
Ventricular Fibrillation	VFW	W
Paced Beat	PB	B
Right Bundle Branch Block	RBB	R

When acquiring MIT-BIH data and converting it into 128×128×1 grayscale images, the data from the MIT-BIH database should first be saved as an npy file. Subsequently, the dataset is divided into training and testing sets in an 8:2 ratio. The npy file is then loaded into a numpy array, where the timing signals are transformed into images, followed by the generation of the output visualizations.

2.2 Improved FSWT algorithm design

2.2.1 Traditional FSWT design

Suppose p(ω) is the Fourier transform of the function p(t). For any $f(t) \in L^2(R)$, the frequency slice wavelet transform (FSWT) is directly defined in the frequency domain as:

$$W_{\mathrm{f}}(t,\omega,\sigma) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \hat{f}(u)\hat{p}^*(\frac{u-\omega}{\sigma})e^{iut}du \qquad (1)$$

Its bid sigma is either a constant or a function of ω, t, and u. The star "*" denotes the conjugation of the functions. Specifically, ω and t are referred to as the observed frequency and time, respectively, u represents the evaluated frequency, and p denotes the frequency slice function (FSF).

FSWT time-frequency window:

In time domain:

$$\left[\mathrm{t} - \frac{1}{\sigma}\Delta t_p, t + \frac{1}{\sigma}\Delta t_p\right] \qquad (2)$$

In the frequency domain:

$$\left[\omega - \sigma\Delta\omega_{\mathrm{p}}, \omega + \sigma\Delta\omega_p\right] \qquad (3)$$

Both the frequency domain and time domain windows can be directly adjusted, thereby overcoming the limitations of wavelet transform. Here, $\hat{p}$ represents the frequency slicing function (FSF) in the formula. The frequency window size of the FSF varies inversely with 1/κ, exhibiting gradual expansion in the high-frequency region but abrupt changes in the low-frequency region. For high-frequency signals, such as vibration, FSWT demonstrates excellent performance. Conversely, FWST performs inadequately for most low-frequency physiological signals. Additionally, there exists a unique constant scaling parameter κ that controls time-frequency resolution. In FWST, selecting an appropriate scale parameter κ poses a significant challenge.

2.2.2 The improved FSWT structure

If there is a signal f(t) and its Fourier transform is assumed to be $\hat{f}(u)$, the frequency domain of the MFSWT model can be expressed as:

$$\mathrm{w}_f(t,\omega)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\hat{f}(u)\hat{p}^*(\frac{u-\omega}{q(\hat{f}(u))})e^{iut}du \tag{4}$$

$$\hat{p}(x)=e^{-\frac{x^2}{2}} \tag{5}$$

$$\mathrm{q}=\delta+sign(\nabla\left|\hat{f}(u)\right|) \tag{6}$$

2.3 Dynamic band adjustment model

The transient feature sensitivity is defined as:

$$S(t)=\frac{1}{\Delta t}\left|\frac{dV(t)}{dt}\right|+\lambda\frac{d^2V(t)}{dt^2} \tag{7}$$

Where λ is the slope weight coefficient (experience value 0.3-0.6).
The dynamic adjustment formula of frequency band width is as follows:
Where λ is the slope weight coefficient (experience value 0.3-0.6).
The dynamic adjustment formula of frequency band width is as follows:

$$B(t)=B_0*\left[1+\tanh(\frac{S(t)}{S_{th}})\right] \tag{8}$$

S_{th} is the adaptive threshold, determined by sliding window statistics of 75% quantile of QRS interval S(t).[3]

2.3.1 Morphological adaptive kernel function

QRS kernel function (high transient feature) :

$$\omega_Q(t)=e^{-(\alpha t)2}*\cos(2\pi f_c t+\beta t^2) \tag{9}$$

β is the frequency modulation factor, which controls the frequency band inclination.[11]
P/T region kernel function (low amplitude slow variation feature) :

$$\omega_P(t)=\frac{1}{(1+(\lambda t)^{\frac{1}{2}})}*\sin c(kt) \tag{10}$$

The sidelobe decay rate is controlled by γ and k parameters.[3]
Parameter adaptive process:
(1) R wave detection to determine the QRS center location;
(2) Dynamically adjust β and γ based on waveform slope.
Construct the time-varying kernel function combination as:

$$\omega(t)=\sum_m\omega_Q(t-mT_Q)+\sum_n\omega_P(t-nT_P) \tag{11}$$

 https://doi.org/10.1051/itmconf/20257802011

Compared with traditional time-frequency analysis technology, MFSWT does not need to set complex parameters, and can be dynamically adjusted to display more signal time-frequency information.

2.4 Experimental model

In terms of the model architecture, this project first constructs a convolutional neural network based on AlexNet with 6 convolutional layers. Batch normalization is implemented after both the convolutional layers and fully connected layers to mitigate overfitting. Meanwhile, dropout with a probability of 0.5 is applied after the fully connected layers to enhance the model's generalization capability. Subsequently, SENet and GRU are embedded into the network: SENet is inserted at the end of each convolutional layer to introduce channel attention mechanisms, thereby improving accuracy and computational efficiency[6]; GRU is incorporated after all convolutional operations to fully capture temporal features in ECG signals[9]. The model structure is shown in Figure 1below:

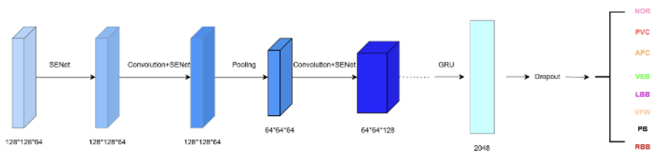

Figure 1 Model architecture

The code model in this experiment is essentially an eight-classification model. Since the original model is RGB images with three channels, this experiment adds transforms.Grayscale(1) to transform data into a single channel.

3 Results

3.1 Experimental Setup

The hardware platform used in this experiment is as follows: CPU: Intel Core i9-12900H (14 cores, 20 threads, up to 5.0 GHz); GPU: NVIDIA GeForce RTX 3070 Ti; Memory: 64GB DDR5-4800 (dual-channel). The software development environment is: Operating System: Windows 11 Pro 22H2 + WSL2 (Ubuntu 20.04 LTS). The core computing libraries used are: CUDA 11.7 + cuDNN 8.5.0, PyTorch 1.13.1 + TorchAudio 0.13.1, Python 3.8.15, all managed through Anaconda.1.1.3 Language, Spelling and Grammar: all papers must be written in UK English.. Papers that fail to meet basic standards of literacy are likely to be declined. Ensure that the paper is spell checked.

3.2 Evaluation index

Based on the original evaluation metrics of most deep learning models in this experiment, such as Accuracy, Precision, Recall, and F1 Score, a confusion matrix has been incorporated to visually illustrate the accuracy of model predictions. This addition aims to validate the reliability of the model for clinical diagnostic applications.

3.3 Results and Analysis

3.3.1 The effect of improved frequency slice wavelet transform in preprocessing ECG signals

After enhancing the frequency slice wavelet transform of electrocardiogram signals, the replication becomes clear, the spectral features become prominent, and they can be accurately identified by the model just like Figure2 and Figure3 .

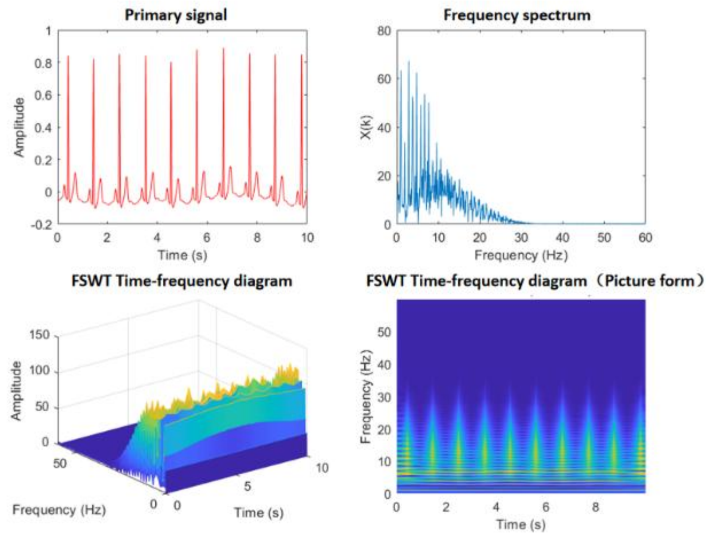

Figure 2 Improved FWST processing effect

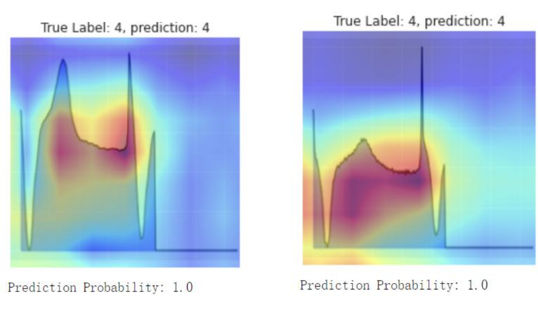

Figure 3 The result of time-frequency graph prediction after FSWT processing

3.3.2 Comparative analysis

In this experiment, Experimental Group 1 employed only the data augmentation method of image translation and rotation, whereas Experimental Group 2 simultaneously underwent frequency slicing wavelet transformation. Both groups utilized a unified training model with a batch_size set to 72. The learning rate η and learning decay rate γ were set to 1e-3 and 0.95, respectively, for 240 epochs of training (note: due to the model stabilizing after 150 epochs, training was halted at 150 epochs after implementing the improved frequency slicing wavelet transformation, but the results were still superior to those of Experimental

Group 1 without the transformation). Below are the results before (Figure 4) and after (Figure 5 and Figure 6) any data preprocessing and transformation:

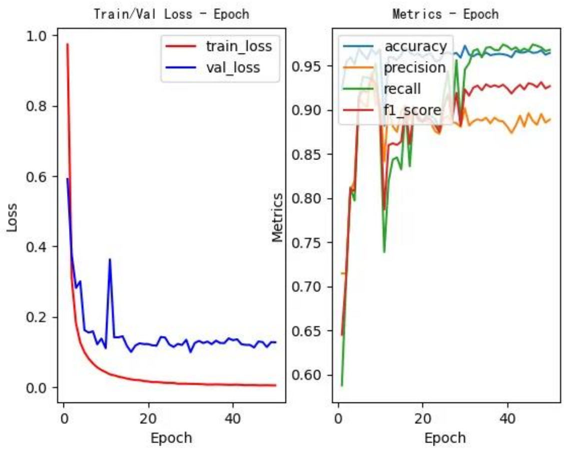

Figure 4 Training results of the control group without data preprocessing

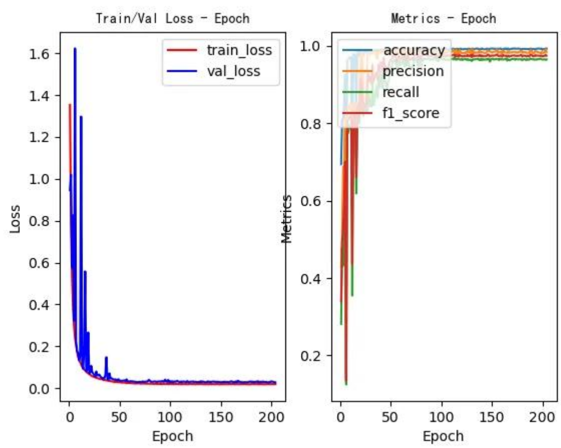

Figure 5 The results of the experimental group without frequency slicing wavelet transform

Table 2 The results of the experimental group without frequency slicing wavelet transform

Accuracy	0.9825
Precision	0.9803
Recall	0.9651
F1_score	0.9776

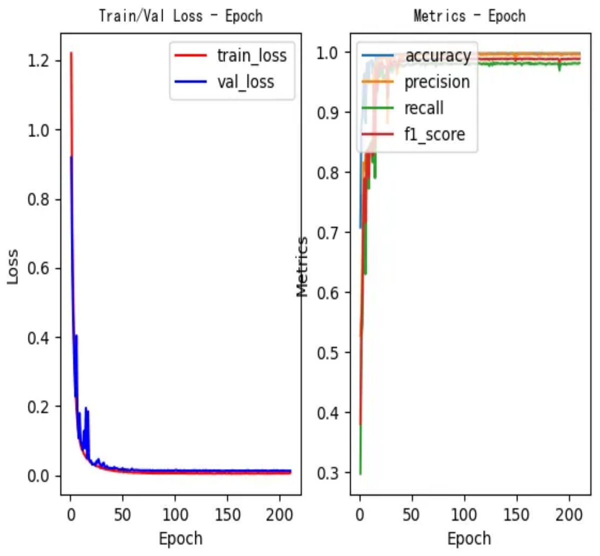

Figure 6 The results of the experimental group after frequency slice wavelet transform

Table 3 The results of the experimental group after frequency slice wavelet transform

Accuracy	0.9986
Precision	0.9989
Recall	0.9980
F1_score	0.9984

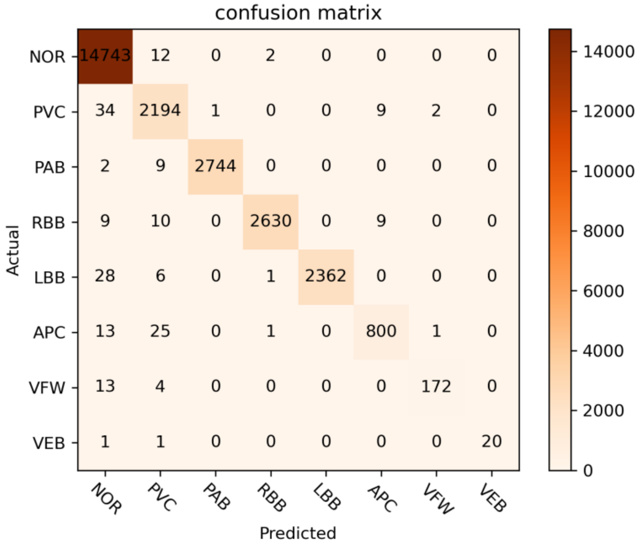

Figure 7 Training confusion matrix of the transformed experimental group

In the control group without data preprocessing, recall struggled to converge and exhibited overfitting. Therefore, it is necessary to perform data preprocessing in the experimental phase. By comparing experimental groups 1 and 2, significant improvements were observed in both prediction accuracy and F1 score, with increases of 0.0161 and 0.0208, respectively. Additionally, the training and prediction loss curves were smoother and more consistent. The improved frequency slice wavelet transform indeed facilitated the recognition of electrocardiogram (ECG) signals[1]. The confusion matrix clearly showed that the predicted data predominantly fell along the diagonal, indicating that this data processing method and convolutional neural network have demonstrated good performance in ECG signal recognition, which holds considerable significance for clinical applications[10].

4 Conclusion

This article introduces an improved frequency domain slice wavelet transform for data preprocessing. This method flexibly converts the obtained data into time-frequency maps for training, facilitating the full extraction of time-frequency features from ECG signals and enhancing the model's accuracy. A SENet-GRU hybrid neural network model is proposed. SENet assigns different weights to different channels based on their importance, thereby extracting multi-level features from ECG signals. As a recurrent neural network, GRU exhibits excellent performance in processing temporal signals like ECG, thus fully leveraging the temporal characteristics of these signals. The experimental method can be further extended to applications in wearable mobile medical devices and multi-modal medical image fusion. This xperiment also holds significance for future research, such as exploring ways to reduce the computational load of signal preprocessing and enhance efficiency[5].

References

1. Huang, J. et al., "ECG Arrhythmia Classification Using STFT-Based Spectrogram and CNN," IEEE Access, vol. 7, pp. 92871-92880, 2019.
2. Zhao, J. et al., "Frequency Slice Wavelet Transform for Mechanical Fault Diagnosis," Mechanical Systems and Signal Processing, vol. 115, pp. 260-293, 2019.
3. Singh, P. et al., "Deep Noise Suppression for ECG Signals Using Modified Wavelet Thresholding," IEEE Sensors Journal, vol. 22, no. 3, pp. 2638-2648, 2022.
4. Yan, Y. Z. H. et al., "Frequency Slice Wavelet Transform for Transient Vibration Analysis," Mechanical Systems and Signal Processing, vol. 23, no. 5, pp. 1474-1489, 2009.
5. Li, J. et al., "Gesture Myoelectric Signal Recognition Based on Improved Frequency Sliced Wavelet Transform and Convolutional Neural Network," Journal of Nanchang University (Engineering and Technology), vol. 43, no. 4, pp. 401-408, 2021.E. Chaudary et al., "EEG-CNN-Souping: Interpretable Emotion Recognition from EEG Signals," Computers and Electrical Engineering, vol. 123, 110189, 2025.
6. Moody, G. B. et al., "The MIT-BIH Arrhythmia Database on CD-ROM and Software for Use with It," Computers in Cardiology, pp. 185-188, 1990.
7. Wang, Y. et al., "Adaptive Time-Frequency Analysis Based on Parameter Optimization for Nonstationary Signal Processing," IEEE Transactions on Instrumentation and Measurement, vol. 70, 6003614, 2021.

8. Xu, X. et al., "Epileptic EEG Classification Using Frequency Slice Wavelet Transform and Graph Convolutional Network," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 30, pp. 2608-2617, 2022.
9. Chen, S. et al., "Generative Adversarial Networks for Electrocardiogram Augmentation," IEEE Journal of Translational Engineering in Health and Medicine, vol. 10, pp. 1-11, 2022.
10. Li, L. et al., "Sensitivity Analysis of Frequency Slice Wavelet Transform Parameters in Fault Diagnosis," Mechanical Systems and Signal Processing, vol. 168, 108693, 2022