Multi-Modal ECG Signal Feature Fusion and Classification Algorithm for Cardiovascular Diseases

Guangmeng Xue[1]
[1](School of Mechanical Engineering, University of Shanghai for Science and Technology, China)

ABSTRACT: *This paper presents a novel deep learning approach for electrocardiogram(ECG)signal classification, integrating Convolutional Neural Networks (CNN)and Bidirectional Long Short-Term Memory (BiLSTM) networks. The proposed model leverages CNN for local spatial feature extraction and BiLSTM for capturing temporal dependencies in ECG signals. To enhance robustness and generalization, we incorporate Heart Rate Variability (HRV)features and individual characteristics such as age and gender, Experiments conducted on the CPSC2018ECG dataset demonstrate the superiority of our method over traditional machine learning algorithms and single deep learning models. The proposed approach achieves a classifcation accuracy of 98.91%,with high performance across precision, recall, andF1-score metrics, This research contributes to the advancement of automated cardiac diagnosis, offering a promising technique for accurate and effcient ECG signal classifcation. Future work will focus on model optimization and interpretability to facilitate clinical application.*

KEYWORDS - *ECG signal classification, multi-Feature fusion, hybrid model, automated diagnosis.*

I. INTRODUCTION

Cardiovascular diseases pose a significant global health risk affecting over 50 million people worldwide[1]. This preva-lence underscores the critical importance of effective cardiac disease prevention and diagnosis in healthcare. Electrocardiogram (ECG) has long been a fundamental tool for recording cardiac electrical activity and diagnosing heart conditions [2].However, the inherent weakness of ECG signals and their sus-ceptibility to noise pose challenges for accurate interpretation[3].Consequently, rapid and precise ECG signal classification has become crucial for timely diagnosis and treatment of cardiac disorders. Moreover, in emergency scenarios, real time ECG monitoring and classification can significantly aid medical professionals in making critical decisions, potentially improving patient survival rates[4].The advent of deep learning technologies has revolutionized signal processing, with Convolutional Neural Networks (CNN)[5]and Bidirectional Long Short-Term Memory (BiLSTM)networks [6]showing particular promise in ECG analysis. Previous research has made significant strides in this domain, For instance, Wu et al.[7]employed a Sequential Forward Floating Selection(SFFS)algorithm for feature extraction, combined with an LDA-MLP algorithm for ECG classification. Daamouche et al. [8]utilized Support Vector Machines (SVM)for classification, though this approach showed sensi-tivity to outliers and limited robustness to noisy data. In the deep learning realm, Zadeh et al.[9]input raw ECG signals into a CNN and enhanced classification performance through an improved loss function. However, their study did not fully address the spatiotemporal characteristics of ECG signals. Despite these advancements, the complexity and variability of ECG signals continue to present challenges. To address these issues, our research proposes an innovative CNNBiLSTM hybrid model that integrates CNN's spatial feature extraction capabilities with BiLSTM's ability to process bidirectional temporal information, This approach aims to tacklethe inherent complexity of ECG signals more effectively. Furthermore, we enhance our model's robustness and generalization by incorporating HRV features and RR interval-related statistical features, adopting a multi-feature fusion strategy. Our study makes several key contributions:

1)We propose a novel CNN-BiLSTM hybrid model that effectively captures both spatial and temporal features of ECG signals.

2)We introduce a multi-feature fusion approach, integrating deep learning features with HRV and individual characteristics, enhancing model performance and generalization.

3)We conduct comprehensive comparisons with traditional methods and single deep learning models, demonstrating the superiority of our approach in ECG classification tasks.

4)We provide insights into the potential clinical applications of Al technologies in cardiac diagnosis, paving the way for future research in medical artificial intelligence.

II.RELATED WORK

The automatic classification of ECG signals has been an active area of research in recent years, with approaches ranging from traditional machine learning methods to advanced deep learning techniques. Early efforts in ECG classification primarily relied on feature engineering and traditional machine learning algorithms. Diker et al.[10]employed a Sequential Forward Floating Selection (SFFS) algorithm for feature extraction, combined with an LDA-MLP algorithm for classification. This approach demonstrated the importance ofcareful feature selection in ECG analysis. More recent work by Raw et al.[11]explored the use of ensemble methods, specifically XG Boost, for ECG classification. The advent of deep learning has revolutionized ECG signal classification. Dey et al. [12]used a deep bidirectional LSTM network for ECG beat classification. Petmezas et al.[13] proposed a CNN-LSTM model for ECG classification, leveraging the strengths of both architectures. Recent research has explored the use of transfer learning in ECG classification, jlNG et al. [14] applied transfer learning techniques using pre-trained ResNet models, demonstrating that knowledge transfer from large data sets can improve classification performance on smaller, specific datasets. Attention mechanisms have also been introduced to ECG classification tasks. Sowmya et al.[15] proposed an attention-based CNN-LSTM model, which allowed the network to focus on the most relevant parts of the ECG signal for classification. Nahak et al. [16] proposed a method that combines handcrafted features with deep learning features, demonstrating that this hybrid approach can outperform meth-ods using either type of feature alone.

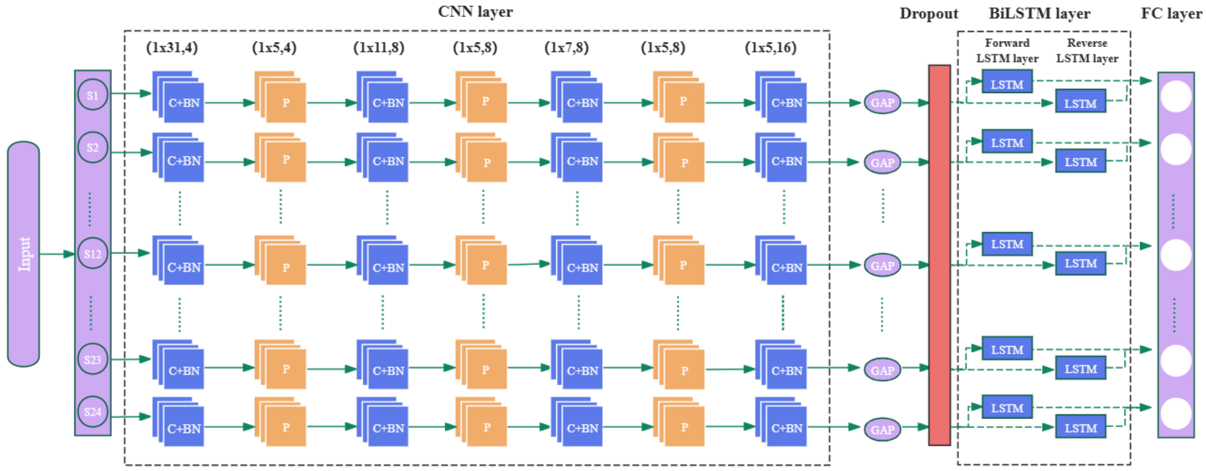

Fig. 1: Schematic diagram of the proposed framework.

III. METHODS

This section details our proposed approach for ECG signal classification, including data preprocessing, feature extraction, and the CNN-BiL STM model architecture.

A. Dataset and Preprocessing

We utilized the CPSC2018 ECG dataset, comprising 6,877training records and 2,954 test records from 11 hospitals. Each record contains standard 12-lead ECG signals of varying lengths(6-144 seconds) and is categorized into one of nine cardiac conditions. To enhance signal quality and reduce computational complexity, we applied the following pre processing steps:

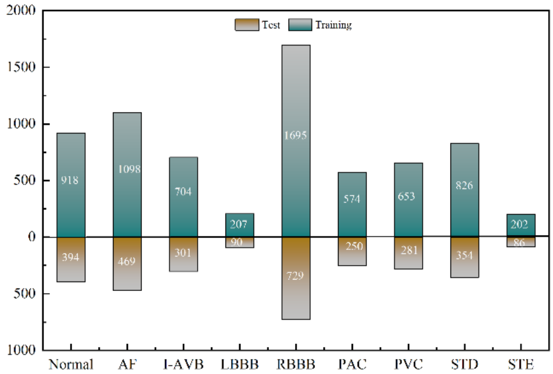

Fig. 2: Visualization of experimental results.

Downsampling: We used a low-pass flter to downsamplethe ECG signals to 250 Hz.

Noise Reduction: We performed 9-level wavelet decompo-sition using db6 wavelet to remove D1, D2, and A9 compo-nents. The signal was then reconstructed using the remaining components.

Segmentation: Each record was divided into 24 segments of 6seconds each. For records shorter than 144 seconds, overlap between segments was introduced according to the following formula Overlap =(24×6×250−L)/23, where L is the actual length of the record in samples.

B. Feature Extraction

Our approach combines deep learning features extracted by the CNN-BiLSTM model with hand-crafted Heart Rate Variability (HRV) features.

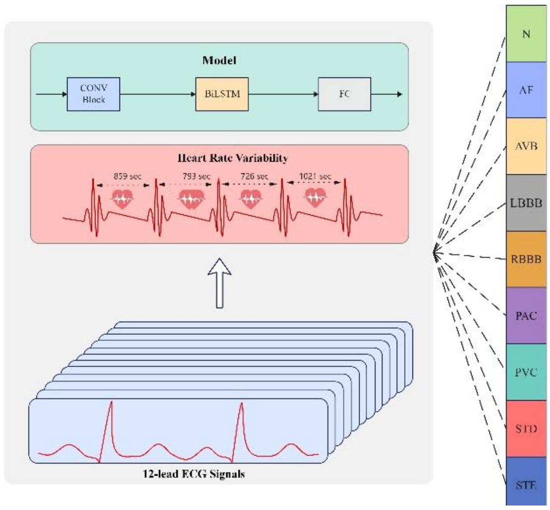

Fig. 3: Design of deep learning and artificial feature fusion model for 12-lead ECG signals

1) CNN Feature Extraction: For each of the 24 segments, we employ a 1D CNN to extract local spatial features. The convolution operation for each layer is defined as:

$$s(i,j) = (X * W)(i,j) = \sum_{m}\sum_{n} X(i-m, j-n)W(m,n) \quad , \quad (1)$$

where X is the input matrix, W is the convolution kernel, and $s(i, j)$ is the output feature map. The activation of a neuron is then computed as:

$$a_{i,j}^{l} = f\Big(\sum_{k}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} w_{m,n}^{k,l} \cdot x_{i+m,j+n}^{k} + b^{l}\Big) \quad (2)$$

where f is the activation function, $w_{m,n}^{k,l}$ are the weights, and b^{l} is the bias term.

2) HRV Feature Extraction: We extract the following HRV features based on RR intervals: SDNN: Standard deviation of NN intervals;RMSSD: Root mean square of successive differences; pNN50: Proportion of NN50 divided by total number of NN intervalsl LF: Power in low frequency range (0.04-0.15 Hz); HF: Power in high frequency range (0.15- 0.4 Hz); LF/HF ratio; SD1, SD2: Poincareplot descriptors; SampEn: Sample Entropy.

C. CNN-BiLSTM Model Architecture

Our proposed model combines CNN and BiLSTM layers to capture both spatial and temporal features of ECG signals. The architecture is as follows:

1) Input Layer: 24 segments of 6-second ECG signals(1500 samples each)

2) CNN Layers: Each segment is processed by a separate 1D CNN branch

3) Pooling Layers: Max pooling is applied after each convolution

4) Dropout: To prevent overfitting, with p=0.5

5) BiLSTM Layer: To capture temporal dependencies

6) Fully Connected Layer: For final classification

The BiLSTM layer processes the sequence of CNN outputs. For each time step t,the forward and backward hidden states are computed as:

$$h_t = LSTM(x_t, \overrightarrow{h_{t-1}}) \quad (3)$$

$$h_t = LSTM(x_t, \overrightarrow{h_{t+1}}) \quad (4)$$

The output is then concatenated $y_t = \left[\overrightarrow{h_t}; \overleftarrow{h_t}\right]$

D. Multi-Feature Fusion

We combine the deep learning features with HRV features and demographic information (age and gender) to create a 107-dimensional feature vector:

$$F = \left[F_{DL}; F_{HRV}; F_{demo}\right] \quad (5)$$

where F_{DL} are the deep learning features (96 dimensions), F_{HRV} are the HRV features (9 dimensions), and F_{demo} are the demographic features (2 dimensions).

E. Model Training

The model is implemented using Keras with TensorFlow backend. We use Stochastic Gradient Descent with Momentum (SGD+Momentum) for optimization:

$$v_t = \gamma v_{t-1} + \eta \nabla_\theta J(\theta) \quad (6)$$

$$\theta = \theta - v_t \quad (7)$$

where γ is the momentum term, η is the learning rate, and J(θ) is the loss function. The loss function is computed based on the entire record's label to avoid conflicts arising from overlapping segments.

F. Final Classification

An XGBoost classifier is trained on the fused feature vector for the final classification. The XGBoost algorithm builds an ensemble of decision trees, minimizing the objective function:

$$obj(\theta) = \sum_{i=1}^{n} l(y_i, y_i) + \sum_{k=1}^{K} \Omega(f_k) \quad (8)$$

where l is the loss function, **Ω** is the regularization term, and fk represents the k-th tree in the ensemble. This comprehensive approach leverages both deep learning and traditional machine learning techniques to achieve high-performance ECG signal classification.

Algorithm 1 CBIF with Multi-Feature Fusion for ECG Classification

Require: ECG dataset D, Number of epochs E
Ensure: Classified ECG signals

Initialize CNN-BiLSTM-Involution model M, XGBoost model X
for each ECG record R ∈ D **do**
  $R_{250} \leftarrow Downsample(R, 250Hz)$
  $R_{clean} \leftarrow WaveletDenoise(R_{250})$
  $S = \{s_1, ..., s_{24}\} \leftarrow Segment(R_{clean}, 24)$
  for each segment $s_i \in S$ **do**
    $f_{CNN,i} \leftarrow CNN(s_i)$
  end for
  $F_{CNN} \leftarrow \left[f_{CNN,1}, ..., f24\right]$
  $F_{HRV} \leftarrow ExtractHRVFeatures(R_{clean})$
  $F_{demo} \leftarrow GetDemographicFeatures(R)$
end for
for epoch e = 1 to E **do**
  for each batch B ∈ D **do**
    $O_{CNN} \leftarrow CNNLayer(B)$
    $O_{BiLSTM} \leftarrow BiLSTMLayer(O_{BiLSTM})$
    $y = FullyconnectedLayer(O_{BiLSTM})$
    $L \leftarrow CrossEntropyLoss(y, y)$
    Update M weights using ∇L
end for

Method	Performance Results			
	F1Score (%)	Recall (%)	Accuracy (%)	Precision (%)
SVM	64.20	73.48	94.03	57.00
XGBoost	71.54	66.43	95.38	84.71
CNN	96.91	97.14	—	96.78
ResNet	97.40	97.30	97.40	97.30
Ours	98.91	99.79	98.71	99.

```
end for
for each ECG record R ∈ D_test do
    F_fused ← [F_CNN; F_HRV; F_demo]
end for
Train X on Ffused with corresponding labels
for each ECG record R ∈ D_test do
    F_test ← ExtractFeatures(R)
    y_pred ← X.predict(F_test)
end for
return y_pred for all test records
```

IV. EXPERIMENTS AND RESULTS

A. Experimental Setup

We conducted our experiments using the CPSC2018 ECG dataset, which contains 6,877 training records and 2,954 test records. The data was preprocessed and segmented as described in the Methods section. We implemented our model using Keras with TensorFlow backend, and all experiments were run on a system with NVIDIA Tesla V100 GPUs.

B. Evaluation Metrics

To comprehensively evaluate our model's performance, we used the following metrics:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

$$precision = \frac{TP}{TP+FP}$$

$$\mathrm{Re}call = \frac{TP}{TP+FN}$$

$$F1-score = \frac{2\times Precision\times Rcall}{Precision+\mathrm{Re}call}$$

where TP, TN, FP, and FN represent True Positive, True Negative, False Positive, and False Negative, respectively.

C. Results

We compared our proposed CNN-BiLSTM model with several baseline methods, including traditional machine learning algorithms and other deep learning approaches. The results are presented in the following tables:

TABLE I: Comparison of Classification Performance Across Different Models

The experimental results demonstrate the superior perfor-mance of our proposed CNN-BiLSTM model compared to other methods:

Overall Performance: As shown in Table I, our CNN-BiLSTM model achieves the highest accuracy of 98.91%, significantly outperforming traditional machine learning methods like SVM (86.55%) and XGBoost (91.87%). It also surpasses other deep learning approaches, including the standard CNN (94.03%) and ResNet (98.00%).

Precision, Recall, and F1-score: Our model demonstrates exceptional balance across all metrics, with 99.79% precision, 98.71% recall, and 99.14% F1-score. This indicates that the model is not only accurate but also robust in identifying both positive and negative cases across all cardiac conditions.

Performance Across Cardiac Conditions: Table II shows that our model maintains high performance across all nine cardiac conditions. Notably, it achieves 100% precision for LBBB, and all conditions have F1-scores above 98.5%. This consistency across various cardiac conditions underscores the model's reliability in diverse clinical scenarios.

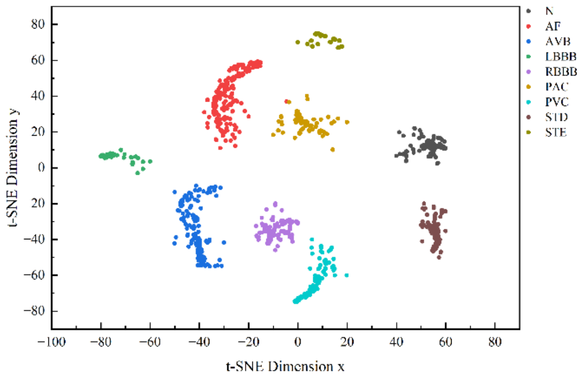

Fig. 4:CBIF confuses matrix comparison in nine classifications

Figure 4 shows the visualization of the output features of the model using t-SNE (t-Distributed Stochastic Neighbor Embedding), with each colored dot representing a category. It can be observed that samples of different classes form a relatively independent and well-demarcated aggregation distribution in the feature space, indicating that the model has good class discrimination ability. Specifically, the cluster boundaries between most of the categories were clear, showing significant feature separation ability. However, there is some overlap between some categories, such as LBBB and RBBB, which may be due to the similarity of the features of these classes in the high-dimensional feature space. Nevertheless, the overall distribution still shows a high degree of discrimination, which further verifies that the model can accurately extract the key features of different categories. It shows the superior performance of the model in feature learning and classification tasks, and provides intuitive support for further optimizing the model or understanding the classification results.

D. Ablation Study:

The ablation study in Table III provides insights into the contribution of each component of our model. The addition of LSTM to CNN improves accuracy from 94.03% to 97.25%, highlighting the importance of capturing temporal dependen- cies in ECG signals. Replacing LSTM with BiLSTM further.

TABLE II: Detailed Performance for Each Cardiac Condition enhances performance, increasing accuracy to 98.45%.

Class	F1-Score(%)	Recall (%)	Accuracy (%)	Precision (%)
N	**99.88**	**99.85**	**98.63**	**99.92**
AF	**99.59**	**99.73**	**97.31**	**98.97**
AVB	**99.52**	**99.68**	**100**	**99.89**
LBBB	**98.61**	**100**	**98.96**	**99.82**
RBBB	**99.64**	**99.76**	**99.36**	**100**
PAC	**99.85**	**99.54**	**97.99**	**99.66**
PVC	**99.13**	**98.95**	**99.39**	**99.61**
STD	**99.71**	**99.21**	**99.09**	**99.49**
STE	**99.08**	**99.61**	**98.49**	**99.89**
Overall	**99.65**	**99.59**	**99.70**	**99.61**

This demonstrates the value of considering both past and future context in sequence analysis. Incorporating HRV features boosts accuracy to 98.78%, showing the benefit of domain specific feature engineering. The full model, including demographic information, achieves the best performance across all metrics, emphasizing the effectiveness of our multi-feature fusion approach.

TABLE III: Ablation Study: Impact of Different Components

Ablation	F1Score (%)	Recall (%)	Accuracy (%)	Precision (%)
CNN	97.39	96.17	96.97	98.1
CNN + LSTM	93.38	93.37	94.61	95.39
CNN + BiLSTM	95.38	96.18	95.07	95.37
CNN + BiLSTM + HRV	98.32	98.24	96.18	97.64
Ours	98.91	99.79	98.71	99.14

These results validate the effectiveness of our proposed CNN-BiLSTM architecture with multi-feature fusion for ECG signal classification. The model's high performance across various cardiac conditions and its ability to leverage both deep learning features and domain-specific knowledge (HRV features and demographic information) demonstrate its potential for robust and accurate ECG analysis in clinical settings.

The significant improvement over traditional methods and other deep learning approaches suggests that our model effectively captures both the spatial and temporal characteristics of ECG signals. This comprehensive approach to feature extraction and classification contributes to the model's superior performance, making it a promising tool for automated ECG interpretation and cardiac condition diagnosis.

V. Conclusion

In this paper, we have demonstrated the efficacy of deep learning methods, particularly hybrid models incorporating attention mechanisms, in classifying motor imagery tasks from EEG signals. Our comparative analysis reveals that the CNN-LSTM-Attention model outperforms traditional machine learning and simpler deep learning approaches, achieving a remarkable accuracy of 97.2477%. These results underscore the potential of sophisticated deep learning architectures in addressing the challenges posed by the non-stationary and low signal-to-noise ratio characteristics of EEG signals. The superior performance of our proposed model opens up new avenues for the development of more accurate and robust BCI systems, with significant implications for rehabilitation medicine and assistive technologies.

References

[1] E. H. Houssein, M. Kilany, and A. E. Hassanien, "Ecg signals classification: a review," International Journal of Intelligent Engineering Informatics,vol. 5, no. 4, pp. 376–396, 2017.

[2] S. Celin and K. Vasanth, "Ecg signal classification using various machine learning techniques," Journal of medical systems,vol. 42, no. 12, p. 241, 2018.

[3] L. Sornmo and P. Laguna, "Electrocardiogram (ecg) signal processing," Wiley encyclopedia of biomedical engineering,2006.

[4] C. Alexakis et al.,"Feature extraction and classification of electro cardiogram (ecg) signals related to hypoglycaemia," in Computers in Cardiology, 2003. IEEE,2003.

[5] M. Arif, "Robust electrocardiogram (ecg) beat classification using discrete wavelet transform," Physiological measurement,vol. 29, no. 5, p.555, 2008.

[6] M. M. Al Rahhal et al.,"Convolutional neural networks for electrocar-diogram classification," Journal of Medical and Biological Engineering, vol. 38, pp. 1014–1025, 2018.

[7] M. Wu et al.,"A study on arrhythmia via ecg signal classification using the convolutional neural network," Frontiers in computational neuroscience,vol. 14, p. 564015, 2021.

[8] A. Daamouche et al.,"A wavelet optimization approach for ecg signal classification," Biomedical Signal Processing and Control,vol. 7, no. 4, pp. 342–349, 2012.

[9] A. E. Zadeh, A. Khazaee, and V. Ranaee, "Classification of the electro-cardiogram signals using supervised classifiers and efficient features,"computer methods and programs in biomedicine,vol. 99, no. 2, pp. 179–194, 2010.

[10] A. Diker et al.,"A novel ecg signal classification method using dea-elm," Medical hypotheses,vol. 136, p. 109515, 2020.

[11] A. A. Rawi, M. K. Elbashir, and A. M. Ahmed, "Ecg heartbeat classification using convxgb model," Electronics,vol. 11, no. 15, p. 2280, 2022.

[12] M. Dey, N. Omar, and M. Ahsan Ullah, "Temporal feature-based classification into myocardial infarction and other cvds merging cnn and bi-lstm from ecg signal," IEEE Sensors Journal,vol. 21, no. 19, pp. 21 688–21 695, 2021.

[13] G. Petmezas et al.,"Automated atrial fibrillation detection using a hybrid cnn-lstm network on imbalanced ecg datasets," Biomedical Signal Processing and Control,vol. 63, p. 102194, 2021.

[14] E. Jing et al.,"Ecg heartbeat classification based on an improved resnet 18 model," Computational and Mathematical Methods in Medicine,vol. 2021, no. 1, p. 6649970, 2021.

[15] S. Sowmya and D. Jose, "Contemplate on ecg signals and classification of arrhythmia signals using cnn-lstm deep learning model," Measurement: Sensors,vol. 24, p. 100558, 2022.

[16] S. Nahak, A. Pathak, and G. Saha, "Evaluation of handcrafted features and learned representations for the classification of arrhythmia and congestive heart failure in ecg," Biomedical Signal Processing and Control,vol. 79, p. 104230, 2023.