1DCNN-TRSNet: a Hybrid End-to-End Arrhythmia Classification Deep Network Based on Transformer

Yongtao Chen, Sen Qiu, Zhelong Wang and Hongyu Zhao

September 8, 2023

1DCNN-TRSNet: A Hybrid End-to-End Arrhythmia Classification Deep Network Based On Transformer

1st Yongtao Chen
Dalian University of Technology
School of Control Science and Engineering
Dalian, China
cyt3212021@163.com

2ndSen Qiu
Dalian University of Technology
School of Control Science and Engineering
Dalian, China
qiu@dlut.edu.cn

3rd Zhelong Wang
Dalian University of Technology
School of Control Science and Engineering
Dalian, China
wangzl@dlut.edu.cn

4th Hongyu Zhao
Dalian University of Technology
School of Control Science and Engineering
Dalian, China
zhy.lucy@hotmail.com

***Abstract*—Intelligent recognition of arrhythmias using ECG signals is crucial in the diagnosis and prevention of heart diseases. However, traditional methods for detecting arrhythmias rely on manual analysis by specialized physicians, which can be subjective and time-consuming. In this paper, we propose a novel end-to-end 1DCNN-Transformer hybrid deep learning model for the automatic recognition of arrhythmia types. Our proposed model employs a Transformer-based architecture to aggregate local contextual spatial features from a 1DCNN and incorporates temporal information in the high-level abstract features of the convolutional network output using positional encoding. The proposed deep learning model comprises a Transformer-Encoder and Transformer-Decoder that utilizes a multi-headed self-attentive mechanism to couple spatiotemporal features from different time segments and filter useful feature information. Each module converts input feature information into a higher-level abstract output, enabling the model to learn a complex abstract transformation function directly from the original ECG signal. Finally, the projected output is mapped onto the arrhythmia label space. Experimental results on the MIT-BIH Arrhythmia Database demonstrate that our proposed 1DCNN-Transformer network achieves excellent performance, with an overall average recognition accuracy of 99.46%, across five categories of arrhythmia signals: normal beat (N), right bundle branch block beat (R), left bundle branch block normal beat (L), premature ventricular contraction (V), and atrial premature beat (A).**

***Index Terms*—Arrhythmia, Classification, Transformer, Encoder-Decoder.**

ACCORDING the World Health Organization (WHO), arrhythmias are a significant group of cardiovascular diseases and account for more than 31% of deaths worldwide, with their incidence increasing. As arrhythmias may be asymptomatic and exhibit features such as painlessness and insignificant myocardial ischemia, they can lead to stroke or sudden death, making long-term monitoring necessary for selected patients [1]. Electrocardiogram (ECG) is a widely used and cost-effective non-invasive technique for recording changes in the electrical activity of the human heart, which has become a key tool for physicians in the diagnosis of cardiac arrhythmias. However, manual visual analysis of irregular ECGs by trained professional cardiologists is time-consuming, subjective, and expensive, and does not meet the clinical requirements for real-time diagnosis of arrhythmias. Therefore, the development of automated intelligent systems that can assist in the accurate and efficient detection and diagnosis of arrhythmias is essential.

It is widely recognized that the current traditional machine learning-based methods for arrhythmia classification are typically divided into four fundamental steps: preprocessing, feature extraction, feature selection, and classification [2].During the preprocessing stage, several techniques have been utilized by researchers such as wavelet transform, Fourier decomposition, high-pass filtering methods, Kalman filtering, empirical modal decomposition, and others [3]- [7] to remove various artifacts, including myoelectric noise, power line interference, baseline drift, and other types of noise from raw ECG signals. Once the ECG signals have been pre-processed, they are then segmented to extract heartbeats of a specific length. To achieve high-accuracy classification, it is crucial to obtain high-quality features from these segmented heartbeats. Several types of features can be extracted from the segmented heartbeats, including morphological features such as p-wave, QRS wave, and T-wave duration and amplitude [8] [9], time-frequency features [10] [11], and nonlinear features, such as fuzzy entropy [12]. These features can be obtained using various transformations, such as discrete wavelet transform [13], complex wavelet transform [14], variable modal decomposition [15], empirical modal decomposition [16], and others. P. Sharma [17] employed the Linear Adaptive Sine-

(Corresponding author: S. Qiu.)

This work was jointly supported by the National Natural Science Foundation of China under Grant No. 62272081, No. 61803072 and No. 61903062, Natural Science Foundation of Liaoning Province, China under Grant 2021-MS-111, in part by the Fundamental Research Funds for the Central Universities, China under Grant No. DUT22YG128. The authors would like to express their thanks to these funding bodies.

Cosine Algorithm (LA-SCA) to optimize feature extraction from ECG signals, which were then input into both deep neural networks (DNN) and support vector machines (SVM) for ECG signal identification. In another study, Dias et al. [18] extracted various features from each beat, based on RR interval, morphological values, and higher-order statistics. The combination of these features was then fed into a linear discriminant classifier to identify arrhythmia classification. S. Celin [19] initially applied low-pass and high-pass filters to eliminate high-frequency and redundant noise from the extracted data. They then used statistical parameter features and applied several classifiers, including SVM, Adaboost, ANN, and Naive Bayes, to identify the extracted features.Traditionally, machine learning approaches involve extracting local features from a fixed time window, which may lead to the loss of temporal information. In recent years, deep learning methods have been increasingly used for arrhythmia classification. For instance, depthwise separable CNN [20], Deep LSTM [21], and CNN-LSTM [22] architectures, which are based on convolutional networks (CNN) and recurrent neural networks (RNN), have shown promising results in this field. These methods can capture both spatial and temporal features from raw ECG signals, which can significantly improve the accuracy of arrhythmia classification.Recent studies [23]- [25] have suggested that deep learning models for ECG signals can benefit from incorporating attention mechanisms. These mechanisms can dynamically capture the differences between individual frames in the signal and assign different weights based on their importance. By doing so, attention-based models can improve recognition results and optimize the accuracy of arrhythmia classification. Therefore, the use of attention-based models is an active area of research in ECG signal processing, with promising results reported in recent studies.

Recently, the transformer architecture [29] has shown great success in processing time series data by relying entirely on the multi-headed self-attention mechanism [26]- [28]. This method enables the transformer to capture global dependencies directly from the input sequence without using convolution or recursion operations. As a result, transformer models can learn long-term dependencies more efficiently and have shorter path lengths than traditional CNN and RNN models. This makes transformers particularly suitable for ECG signal processing, where long-term dependencies between signal frames are critical for accurate arrhythmia classification. In this paper, we propose an end-to-end 1DCNN-Transformer (1DCNN-TRS) model for arrhythmia classification. Our model combines the advantages of both 1D convolutional neural networks (1DCNN) and transformers, allowing for effective feature extraction while retaining temporal information. The main contributions of our work are as follows.

1) To investigate long-term dependencies and inter-segment interactions commonly present in ECG signals, we proposed a model that combined 1DCNN, Transformer Encoder, Transformer Decoder, and MLP to automatically predict arrhythmia categories. The 1DCNN extracts spatial features within local contextual information while the Transformer models the long-term dependence using attention mechanisms. Furthermore, the Transformer Decoder captures inter-segment interactions present in the ECG signal. These components were successfully combined to create an effective model for predicting arrhythmia categories with high accuracy.
2) We propose a deep network consisting of an 8-layer encoder and a 1-layer decoder driven by ECG temporal and spatial feature data and present the 1DCNN-TRS framework. A multi-headed self-attentive machine is used to couple the spatiotemporal features between different segments of the ECG signal and filter the input feature vector i.e. by assigning greater weight to the more important features for ECG abnormality identification.

I. Proposed method

In this section, we introduce the proposed 1DCNN-TRS framework, which employs a multi-headed attention mechanism to address long-term dependencies. The model is an end-to-end structure, as depicted in Fig. 1. Specifically, the original heartbeat signal is first fed into the 1DCNN module, which extracts local contextual space features.To incorporate positional information, we utilize positional encoding before feeding the input into the multilayer Transformer-Encoder. The encoding block then extracts spatiotemporal features of the ECG signal sequence. Subsequently, the multilayer decoder generates the next predicted element by utilizing the context information provided by the encoding block and the previously predicted element. Lastly, we employ an inference sub-network with a fully connected layer to predict the human ECG state using the high-level abstract feature representation learned by the encoder. All components in the 1DCNN-TRS framework are jointly trained as a whole, optimizing the model to learn the best possible features and dependencies for predicting ECG arrhythmias accurately.

A. Transformer overview

Transformer networks has proven to be highly effective for a variety of tasks beyond natural language processing, including ECG signal analysis. In the Transformer framework, each encoder computes attention using queries, keys, and values from the output of the previous encoder layer, incorporating residual networks and layer normalization to improve model stability. The decoder utilizes both the contextual information provided by the encoder and the predicted previous element to generate a prediction for the next element in the output sequence. This approach facilitates effective learning of dependencies between input features and enables accurate prediction of ECG abnormalities. In the cross-attention layer, the query is derived from the output of the previous decoder layer, while the keys and values are obtained from the output of the entire encoder block. In computing the decoder multi-head mask self-attention, all three (query, keys, and values) are derived from the output of the preceding decoder layer. This approach

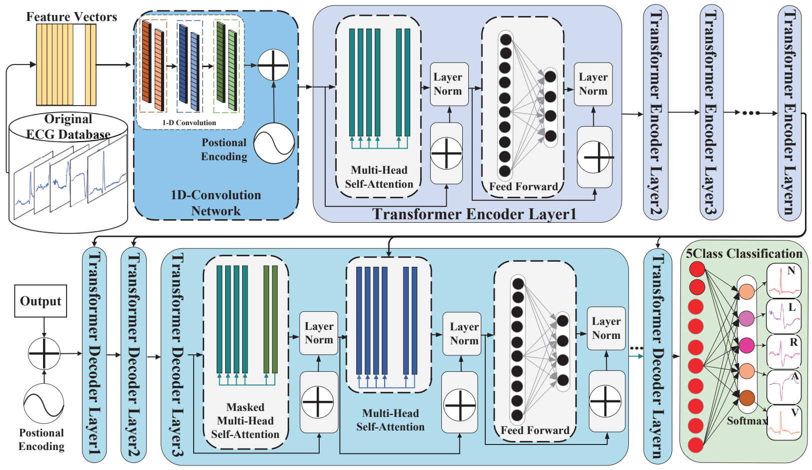

Fig. 1: 1DCNN-TRS framework for arrhythmia recognition

allows the model to attend to relevant information in both the input and decoded sequence, facilitating highly accurate predictions of ECG abnormalities. Unlike traditional encoder-decoder structures, the Transformer network relies solely on multi-headed self-attention rather than convolutional and recursive layers, allowing for greater flexibility and improved accuracy in the sequence-to-sequence learning tasks.

B. 1D-CNN and Positional Encoding

Since ECG signals are one-dimensional temporal sequences, a one-dimensional temporal convolutional network is used to encode the temporal information of the input feature sequence and generate a feature sequence $\hat{X} \in \mathbb{R}^{n \times d}$, as in Eq. (1), where k represents the convolutional kernel size. Since the Transformer-encoder does not have recursive operations and the Transformer network does not know the position information of each element in the sequence, positional encoding is added to incorporate positional information of each sequence component in A. This can be done using either relative or absolute position encoding techniques, as described in [29]. Position encoding (PE) is used to introduce time-dependent information to the convolutional sequence $\hat{X}$. This is accomplishedusing sin and cos delta functions of different frequencies, which are added to the input features of each element in the sequence. The position embedding matrix $P \in \mathbb{R}^{n \times d}$ for a convolutional sequence is defined as follows: Position encoding (PE) of the convolutional sequence $\hat{X}$ using sin and cos delta functions of different frequencies is used to introduce time-dependent information. The position embedding matrix of a convolutional sequence is defined as the matrix $P \in \mathbb{R}^{n \times d}$:

$$\hat{X} = \mathrm{Con}\, v1d(X, k) \tag{1}$$

$$\begin{aligned} P_{i,2j} &= \sin\left(\frac{i}{10000^{2j/d}}\right) \\ P_{i,2j+1} &= \cos\left(\frac{i}{10000^{2j+1/d}}\right) \end{aligned} \tag{2}$$

where $i \in [1, \ldots, n], j \in \left[0, \ldots \frac{d}{2}\right)$, the final transformer model input feature sequence $Y = \hat{X} + P \in \mathbb{R}^{n \times d}$is obtained.

C. Multiple Self-Attention Mechanism

As an attention mechanism, self-attention reveals meaningful contextual information by identifying long and short-distance dependencies in a sequence, ultimately achieving efficient allocation of information processing resources. Multi-headed self-attention is utilized to establish multiple complex contextual connections among sequence elements, resulting in a more comprehensive global representation. This is achieved

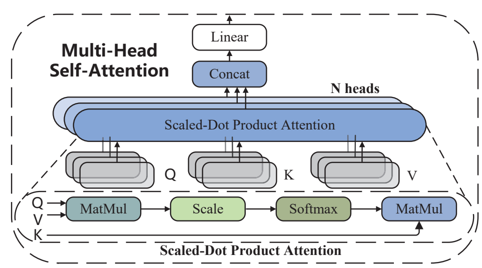

Fig. 2: Multi-Head Self-Attention.

through partitioning the attention into multiple subhead spaces, allowing for the learning of multiple sets of distinct linear projections. These self-attention heads perform dot product operations in parallel, enabling the network to capture a diverse range of relationships and features between sequence elements. Fig.2 illustrates the process of the multi-head self-attention mechanism. Let denote a sequence of N elements, where the feature dimension of each element is D. Each self-attentive head learns three sets of defined weight matrices by: $\mathrm{W}_i^Q \in \mathbb{R}^{d_{\text{mod }del} \times d_q}$, $\mathrm{W}_i^K \in \mathbb{R}^{d_{\text{model}} \times \times d_k}$, $\mathrm{W}_i^V \in \mathbb{R}^{d_{\text{madd }d} \times d_v}$. In each subhead space, the self-attention operation is performed by the following steps.

First, in the n^{th}head self-attention layer, the input sequence is projected Q_i, K_i, V_i through the three sets of linear transformations of the above-mentioned weight matrices, and then$S_i = Q_i.K_i$ calculates the attention weight sum : of each element in the sequence. This attention score determines the importance of other features when the features are encoded and learned at the position sequence. In addition, to improve the stability of the gradient iterations, the attention score S_i: is scaled to:$\overset{\wedge}{S_i}= S_i/\sqrt{d_k}$. Then, by means of a softmax function, $\hat{S}_i$ is converted into a probability between 0 and 1:$P_i = soft\max(S_i)$. Finally, the output of the embedded weighted attention value is calculated: The single-headed scaled dot product attention process can be expressed as shown in the following equation. Compared with the recursive operations in traditional recurrent neural networks, the attention mechanism can be computed in parallel, and each step of computation no longer depends on the result of the previous step. Therefore, the attention mechanism has fewer parameters and higher computational efficiency.

$$head_i = Att(Q_i, K_i, V_i) = softmax(Q_i.K_i/\sqrt{d_i}).V_i \quad (3)$$

Accordingly, the multi-head attention mechanism can be defined as the following equation, h is the number of heads, and$\mathrm{W}^O \in \mathbb{R}^{d_v \times d_{\mathrm{model}}}$.

$$Multihead(Q, K, V) = Concat(head_1, ..., head_h)W^{\mathrm{O}} \quad (4)$$

D. Residual Connection and Layer Normalization

In each sublayer of the Transformer, there is a residual connection and a normalization layer to avoid gradient vanishing caused by the deep network. The output can be expressed as follows:

$$output_{Res} = LayerNorm(input + Sublayer(input)) \quad (5)$$

II. EXPERIMENTS AND RESULT

In this section, we present our experiments on the publicly available MIT-BIH Arrhythmia Database [30], which is widely used in the field of human cardiac arrhythmia recognition.To evaluate the performance of our proposed 1DCNN-TRS model, we compare it with commonly used end-to-end time series recognition models, including 1D-CNN, LSTM, CNN-LSTM, Attention-LSTM, and 1DCNN-TE. Additionally, we compare our model with feature engineering-based SVM classifiers to determine if our proposed approach, based on single-lead ECG signals, outperforms the general model in terms of recognition accuracy.

A. Data preprocessing and heartbeat interception

The experimental dataset was obtained from the BIH Arrhythmia Laboratory [30] and comprises 48 records from 47 individuals collected between 1975 and 1979. Each record was sampled at a frequency of 360Hz for approximately 30 minutes, resulting in over 10,900 heartbeats being collected.Furthermore, the majority of heartbeats were analyzed and labeled by two expert cardiologists with a high degree of precision, enabling accurate and reliable analysis for our study.

The normal ECG signal is composed of P, Q, R, S, and T waves [31]. Physicians diagnose cardiovascular disease by analyzing changes in individual ECG waveforms [31], such as altering QRS wave amplitude height and duration, ST-segment height, p-wave, and T-wave when abnormal. However, a raw ECG contains an overwhelming amount of information, making it difficult to detect the precise location of the largest and sharpest beating QRS wave and dividing the raw ECG into multiple consecutive regular heartbeats.The Pan-Tompkins algorithm [32] has been widely used for r-peak detection, and a similar algorithm is employed in this study to identify QRS waves. A total of 250 sample points were extracted before and after the R peak of the original ECG signal, thus resulting in each record having a length of 250. The dataset consisted of five categories of ECG signals: N (Normal beat), V (Premature Ventricular Contraction), R (Right Bundle Branch Block beat), L (Left Bundle Branch Block beat), and A (Atrial Premature Beat). There were 3471, 6993, 6212, 4779, and 2545 records for each category respectively, resulting in a total of 24000 heartbeats in the dataset. The dataset was randomly divided into a training set and a test set, with the ratio of the two being set to 7:3.

B. Evaluation metrics

We chose popular metrics for evaluating arrhythmia classification performance, such as F1-score, Precision (PR), Recall (RE), and Accuracy (ACC).

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

$$PR = \frac{TP}{TP + FP} \quad (7)$$

$$\mathrm{RE} = \frac{TP}{TP + FN} \quad (8)$$

$$F1 = \frac{2 \times PR \times RE}{PR + RE} \quad (9)$$

Where TP, TN, FP, and FN represent true positive, true negative, false positive, and false negative for the specific class being evaluated.

C. Experimental Setup

Our proposed approach has been implemented in the PyTorch environment. For the 1-D CNN utilized in the input embedding subnetwork, we stacked 2 layers with a convolutional kernel size of 3 to increase the ECG feature dimension to 20. We also applied a dropout of 0.1 to prevent overfitting. Additionally, we constructed a transformer model with eight layers of transformer-encoder and one layer of transformer-decoder. Finally, the loss function is constructed using CrossEntropyLoss and the proposed model is trained using the Adam optimizer with a learning rate of 1e-4. The batch size was set to 200, and the number of epochs was set to 150. We trained and evaluated the model on a NVIDIA GeForce RTX 3080 GPU with 16 GB of RAM. In our study, we compared the proposed 1DCNN-Transformer with five different deep learning methods, particularly the Transformer Encoder model. Based on the experimental results, the 1DCNN-Transformer-based method outperformed all other methods and improved the classification of ECG anomalies.

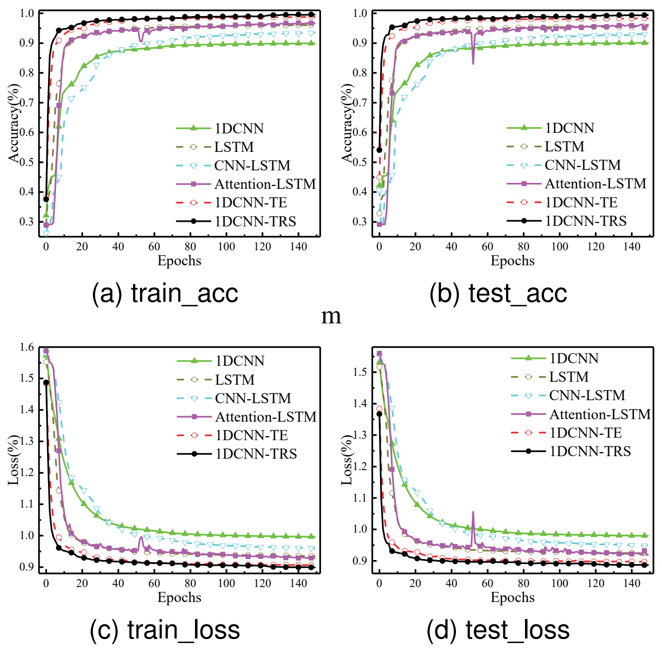

Fig. 3: training and test accuracy cures of 1DCNN, LSTM, CNN- LSTM, Attention-LSTM,1DCNN-TE,1DCNN-TRS.

Deep learning models, when trained with sufficient ECG datasets, have the ability to learn all previously identified manual features as well as unknown but important features. This enables them to improve the recognition rates of ECG without the use of fixed coding features.

1) convolution neural network. This intelligent and automatic recognition of ECG species by CNNs is now widely used in the field [33].
2) long short-memory networks. The Long Short-Term Memory (LSTM) network [34] is designed to excel in processing continuous time-series data. In addition to constructing a deep bidirectional LSTM network, we also developed a 1DCNN-LSTM model that promotes cross-learning between 1DCNN and bidirectional LSTM networks. Moreover, we created an Attention-LSTM network that assigns varying degrees of importance to the output of the bidirectional LSTM network using a self-attention mechanism.
3) Transformer Encoder. In addition to the models mentioned above,we also compared the performance of our proposed algorithm with that of the 1DCNN-TE model, which lacks the Transformer Decoder structure.

D. Experimental results

We conducted comparison experiments and obtained the loss and accuracy curves for the six deep learning classifiers, which are presented in Fig. 3. The 1DCNN-TRS and 1DCNN-TE models achieved overall average classification accuracies of 99.46% and 98.26%, respectively. Our proposed 1DCNN-TRS method achieved 95% accuracy for recognizing abnormal ECGs in the test set after only seven iterations. The curve of

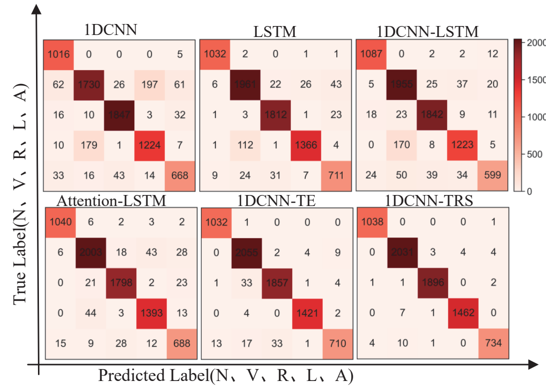

Fig. 4: The overall confusion matrixes of ECG beats classification for six classifiers.

this method converges rapidly and smoothly, outperforming the 1DCNN-TE, LSTM, and 1DCNN models as well as other models. Fig. 4 illustrates the confusion matrix, which presents a detailed distribution of the recognition results for multiple categories from the testing phase using the six classifiers. To assess the generalization ability of the model, we evaluated the performance of different types of diseases using the Receiver Operating Characteristics (ROC) of subjects with different classifiers [35]. Fig. 5 displays the ROC curves of N, V, R, L, and A for multiple classifiers, and we also calculated the area under the ROC curve (AUC). Notably, the macro-average AUC of 1DCNN-TRS is 99.93 %.

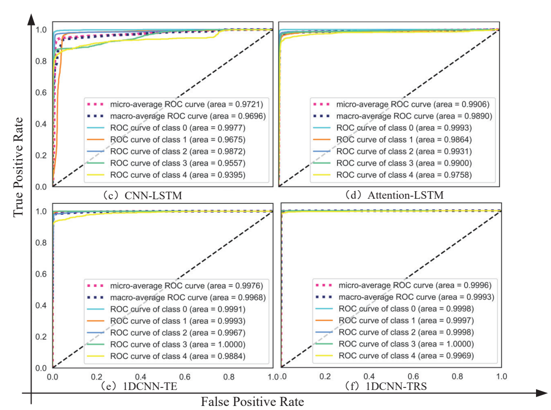

Fig. 5: Receiver operating characteristic curves of six models.

To better understand the performance of the 1DCNN-TRS model, we used the unsupervised visualization technique t-distributed stochastic neighbor embedding (t-SNE) [36] to map the high-dimensional vectors learned by the second Dense layer of the LSTM, 1DCNN-TE, and 1DCNN-TRS models to a two-dimensional space. This allowed us to visualize abstract features on the MIT-BIH dataset, as displayed in Figure 6. The outcomes demonstrate that the 1DCNN-TRS model clusters more closely and can differentiate between class A and class V more accurately. This is attributed to the smaller gaps between the same classes and the greater separation of different classes. Based on these results, we can deduce that the abstract features

TABLE I: **Comparison between the proposed methodology and different state-of-the-art methods that use multiclass**

Reference	Type of ECG beats	Feature	Classifiers	Acc(%)
Zengetal.,2022 [1]	5(N,V,L,R,P)	TQWT	CNN-LSTM	97.20
Liuetal.,2022 [37]	5(N,V,L,R,A)	Raw data	LSTM, Autoencoder	98.57
Mengetall.,2022 [38]	3(N,S,V)	Raw data	transformer Encoder	99.32
Xiaetal.,2023 [39]	4(N,S,V,F)	Raw data	TCGAN	94.69
Muratetal.,2020 [40]	5(N,V,L,R,A)	Raw data	Deep BiLSTM	99.00
Huangetal.,2019 [41]	5(N,V,L,R,A)	TFP	2D-CNN	99.00
Ohetal.,2018 [42]	5(N,V,L,R,A	Raw data	CNN+LSTM	98.10
Yildirimal.,2018 [43]	4(N,L,R,P)	DWT	DBLSTM-WS	99.39
Lial.2016 [44]	5(N,V,L,R,A)	KICA+DWT	SVM optimized by GA	98.8
Our study	5(N,V,L,R,A)	Raw data	1DCNN-TRS	99.46

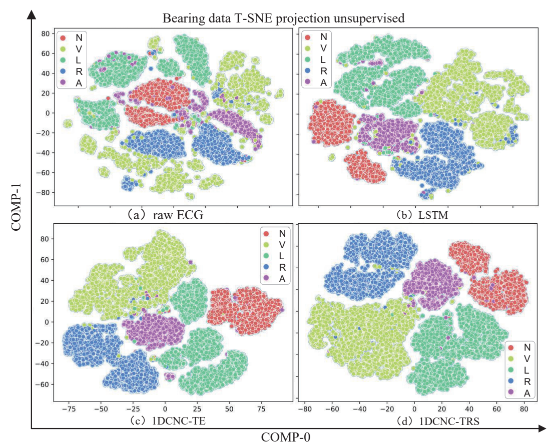

Fig. 6: (a) t-SNE of all raw ECG set's 250-dimensional representation;t-SNE of the second dense layer's eighteen-dimensional representations of the ECG signals all set by the (b) LSTM (c) 1DCNN-TE and (d) 1DCNN-TRS.

transmitted by our proposed method can offer a meaningful separation for the multiclassification task.

III. Discussion

Various methods have been utilized for multi-category arrhythmias classification, such as CNN-LSTM [1], LSTM Autoencoder [37], Lightweight Deformation Encoder (LDE) [35], Deep LSTM [36], 2DCNN [37], CNN+LSTM [38], DBLSTM-WS [39], and SVM optimized by GA model [40]. These techniques are based on either raw data or selected features, and their accuracy rates range from 97.20%to 99.39%, as presented in Table 1. This study introduces a new method, 11DCNN-TRS, which achieves higher classification accuracy than previous techniques. Our approach utilizes a complete Encoder-Decoder structure along with a multi-head attention mechanism to seamlessly integrate temporal and spatial features in ECG beats. These outcomes emphasize the potential of our proposed End-to-End mode for automated ECG beats classification and demonstrate its superiority over existing approaches in the field.

Using our comparative evaluation approach, we have demonstrated the efficacy of our proposed 1DCNN-TRS methodology in precisely identifying different kinds of heartbeats. Our approach utilizes 1DCNN to extract local features from ECG signals and employs positional encoding to capture temporal information from the abstract high-level features provided by the encoding blocks of 1DCNN. Our deep neural network architecture consists of an 8-layer encoder and a 1-layer decoder that leverage both temporal and spatial features of ECG data to facilitate the learning of complex transformation functions for accurate classification.In our transformer layers, we utilize multi-head self-attention mechanism to couple temporal and spatial features across various ECG segments. The mechanism filter out irrelevant noise information and assign more weight to important features for arrhythmia recognition.By means of successive transformations, each module progressively abstracts input information to higher levels of representation, eventually allowing the network to precisely predict the corresponding class label from raw ECG data. As a result, our proposed approach represents a significant progress in the field of ECG-based arrhythmias classification.

IV. Conclusion

Obtaining an early diagnosis of arrhythmia types is critical in reducing the incidence of cardiovascular events. In this study, we introduce a new end-to-end 1DCNN-TRS deep learning model that uses a Transformer Encoder and Transformer Decoder composition to classify multiclass arrhythmias while simultaneously learning spatiotemporal features from various time segments of ECG signals. We trained and evaluated our model on the MIT-BIH arrhythmia dataset, achieving an impressive overall average recognition accuracy of 99.46% for the five different categories of N, V, L, R, and V raw ECG. To further assess the performance of our proposed model, we compared it with various other machine learning methods such as 1DCNN, LSTM, 1DCNN-LSTM, LSTM-Attention, and 1dCNN-TE. Moreover, we compared our model with advanced methods that have been proposed in recent years, and our experimental results indicate that our model outperforms these methods in terms of recognition accuracy. In future work, we intend to expand our proposed method to cover multimodal physiological signals such as ECG signals, EEG signals, inertial motion signals, and EMG signals. Our ultimate objective is to offer clinicians a comprehensive tool to facilitate the evaluation of human body states and the diagnosis of related diseases.

REFERENCES

[1] Zeng W, Su B, Chen Y, et al. Arrhythmia detection using TQWT, CEEMD and deep CNN-LSTM neural networks with ECG signals[J]. Multimedia Tools and Applications, 2022:1-29.

[2] Tan, Jen, Hong, et al. Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network[J]. Information Sciences: An International Journal, 2017, 405:81-90.

[3] Yang Y, Wei Y S. New Threshold and Shrinkage Function for ECG Signal Denoising Based on Wavelet Transform[C]// International Conference on Bioinformatics & Biomedical Engineering. IEEE, 2009.

[4] Tripathi P M, Kumar A, Komaragiri R, et al. A novel approach for real-time ECG signal denoising using Fourier decomposition method[J]. Research on Biomedical Engineering, 2022, 38(4):1037-1049.

[5] Censi F, Calcagnini G, Triventi M, et al. Effect of high-pass filtering on ECG signal on the analysis of patients prone to atrial fibrillation[J]. Annali dell'Istituto superiore di sanita, 2009, 45(4):427-431.

[6] Vullings R, Vries B D, Bergmans J. An Adaptive Kalman Filter for ECG Signal Enhancement[J]. IEEE Trans Biomed Eng, 2011, 58(4):1094-1103.

[7] Zhao Z, Yi L, Lu Q. Adaptive Noise Removal of ECG Signal Based On Ensemble Empirical Mode Decomposition[J]. InTech, 2011.

[8] Yochum M, Renaud C, Jacquir S. Automatic detection of P, QRS, and T patterns in 12 leads ECG signal based on CWT[J]. Biomedical Signal Processing& Control, 2016, 25(Mar.):46-52.

[9] Ye C, Kumar B V, Coimbra M T. Heartbeat Classification Using Morphological and Dynamic Features of ECG Signals[J]. Biomedical Engineering IEEE Transactions on, 2012, 59(10):p.2930-2941.

[10] Rahman Q,Tereshchenko L, Kongkatong M, et al. Utilizing ECG-Based Heartbeat Classification for Hypertrophic Cardiomyopathy Identification[J]. IEEE Transactions on NanoBioscience, 2015, 14(5):505-512.

[11] Lin C H. Frequency-domain features for ECG beat discrimination using grey relational analysis-based classifier[J]. Computers & Mathematics with Applications, 2008, 55(4):680-690.

[12] He H, Tan Y. Automatic pattern recognition of ECG signals using entropy-based adaptive dimensionality reduction and clustering[J]. Applied Soft Computing, 2017, 55:238-252.

[13] Rai H M, Trivedi A, Shukla S. ECG signal processing for abnormalities detection using multi-resolution wavelet transform and Artificial Neural Network classifier[J]. Measurement Journal of the International Measurement Confederation, 2013, 46(9):3238-3246.

[14] Np A, Ms B, Sj C. Design and implementation of a robust noise removal system in ECG signals using dual-tree complex wavelet transform[J]. Biomedical Signal Processing and Control, 63.

[15] Dora C, Biswal P K. Correlation-based ECG Artifact Correction from Single Channel EEG using Modified Variational Mode Decomposition[J]. Computer Methods and Programs in Biomedicine, 183:105092-105092.

[16] Hasan N I, Bhattacharjee A. Deep Learning Approach to Cardiovascular Disease Classification Employing Modified ECG Signal from Empirical Mode Decomposition[J]. Biomedical Signal Processing and Control, 2019, 52(July 2019):128-140.

[17] Sharma P, Dinkar S K. A Linearly Adaptive Sine-Cosine Algorithm with Application in Deep Neural Network for Feature Optimization in Arrhythmia Classification using ECG Signals[J]. Knowledge-based systems, 2022(Apr.22):242.

[18] Dias, Felipe MeneguittiMonteiro, Henrique L. M.Cabral, Thales WulfertNaji, RayenKuehni, MichaelLuz, Eduardo Jose da S. Arrhythmia classification from single-lead ECG signals using the inter-patient paradigm[J]. Computer Methods and Programs in Biomedicine: An International Journal Devoted to the Development, Implementation and Exchange of Computing Methodology and Software Systems in Biomedical Research and Medical Practice, 2021, 202(1).

[19] Celin S, Vasanth K.ECG Signal Classification Using Various Machine Learning Techniques[J]. Journal of Medical Systems, 2018, 42(12).

[20] Degirmenci M, Ozdemir M A, Izci E, et al. Arrhythmic Heartbeat Classification Using 2D Convolutional Neural Networks. 2020.

[21] Bahrami M, Forouzanfar M. Sleep apnea detection from single-lead ECG: A comprehensive analysis of machine learning and deep learning algorithms[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-11.

[22] Rai H M, Chatterjee K. Hybrid CNN-LSTM deep learning model and ensemble technique for automatic detection of myocardial infarction using big ECG data[J]. Applied Intelligence:1-19.

[23] Lu X, Wang X, Zhang W, et al. An End-to-End Model for Ecg Signals Classification Based on Residual Attention Network[J]. Social Science Electronic Publishing.

[24] Wang Y, Zhou G, Yang C. Inter-patient Heartbeat Classification Using Modified Residual Attention Network with Two-phase Training and Assistant Decision[J]. IEEE Transactions on Instrumentation and Measurement, 2022.

[25] 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[J]. Information Fusion, 2020, 53:174-182.

[26] Xia Y, Xu Y, Chen P, et al. Generative adversarial network with transformer generator for boosting ECG classification[J]. Biomedical Signal Processing and Control, 2023, 80: 104276.

[27] Meng L, Tan W, Ma J, et al. Enhancing dynamic ECG heartbeat classification with lightweight transformer model[J]. Artificial Intelligence in medicine, 2022, 124: 102236.

[28] Lu P, Wang C, Hagenah J, et al. Improving Classification of Tetanus Severity for Patients in Low-Middle Income Countries Wearing ECG Sensors by Using a CNN-Transformer Network[J]. IEEE Transactions on Biomedical Engineering, 2022.

[29] Vaswani A, Shazier N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.

[30] [31] Goldberger A L, Amaral L A N, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals[J]. circulation, 2000, 101(23): e215-e220.

[31] Gupta V, Mittal M. A novel method of cardiac arrhythmia detection in electrocardiogram signal[J]. International Journal of Medical Engineering and Informatics, 2020, 12(5): 489-499.

[32] Pan J, Tompkins W J. A real-time QRS detection algorithm[J]. IEEE transactions on biomedical engineering, 1985 (3): 230-236.

[33] Chen C, Hua Z, Zhang R, et al. Automated arrhythmia classification based on a combination network of CNN and LSTM[J]. Biomedical Signal Processing and Control, 2020, 57: 101819.

[34] Yildirim O, Baloglu U B, Tan R S, et al. A new approach for arrhythmia classification using deep coded features and LSTM networks[J]. Computer methods and programs in biomedicine, 2019, 176: 121-133.

[35] Wang H, Shi H, Chen X, et al. An improved convolutional neural network based approach for automated heartbeat classification[J]. Journal of medical systems, 2020, 44: 1-9.

[36] Kim N, Seo W, Kim J, et al. WavelNet: A novel convolutional neural network architecture for arrhythmia classification from electrocardiograms[J]. Computer Methods and Programs in Biomedicine, 2023, 231: 107375.

[37] Liu P, Sun X, Han Y, et al. Arrhythmia classification of LSTM autoencoder based on time series anomaly detection[J]. Biomedical Signal Processing and Control, 2022, 71: 103228.

[38] Meng L, Tan W, Ma J, et al. Enhancing dynamic ECG heartbeat classification with lightweight transformer model[J]. Artificial Intelligence in medicine, 2022, 124: 102236.

[39] Xia Y, Xu Y, Chen P, et al. Generative adversarial network with transformer generator for boosting ECG classification[J]. Biomedical Signal Processing and Control, 2023, 80: 104276.

[40] Murat F, Yildirim O, Talo M, et al. Application of deep learning techniques for heartbeats detection using ECG signals-analysis and review[J]. Computers in biology and medicine, 2020, 120: 103726.

[41] Huang J, Chen B, Yao B, et al. ECG arrhythmia classification using STFT-based spectro-gram and convolutional neural network[J]. IEEE access, 2019, 7: 92871-92880.

[42] Oh S L, Ng E Y K, San Tan R, et al. Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats[J]. Computers in biology and medicine, 2018, 102: 278-287.

[43] Yildirim Ö. A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification[J]. Computers in biology and medicine, 2018, 96: 189-202.

[44] Li H, Yuan D, Wang Y, et al. Arrhythmia classification based on multi-domain feature extraction for an ECG recognition system[J]. Sensors, 2016, 16(10): 1744.