

Deep Learning Models for Arrhythmia Classification using Stacked Time-frequency Scalogram Images from ECG Signals

Parshuram N Aarotale¹ and Ajita Rattani²

¹ Wichita State University, Wichita, USA

² University of North Texas, Denton, USA

Abstract

Electrocardiograms (ECGs), a medical monitoring technology recording cardiac activity, are widely used for diagnosing cardiac arrhythmia. The diagnosis is based on the analysis of the deformation of the signal shapes due to irregular heart rates associated with heart diseases. Due to the infeasibility of manual examination of large volumes of ECG data, this paper aims to propose an automated AI-based system for ECG-based arrhythmia classification. To this front, a deep-learning-based solution based on a self-attention mechanism has been proposed for ECG-based arrhythmia classification. Twelve lead electrocardiograms (ECG) of length 10 sec from 45,152 individuals from Shaoxing People's Hospital (SPH) dataset from PhysioNet with four different types of arrhythmias were used. The sampling frequency utilized was 500 Hz. Median filtering was used to preprocess the ECG signals. For every 1 sec of ECG signal, the time-frequency (TF) scalogram was estimated, resulting in 10 TF scalograms for each ECG signal. The TF scalogram image was used as an input to pretrained convolutional neural network (CNN), 1D CNN, and 1D CNN-LSTM (Long short-term memory) models, for arrhythmia classification. The pre-trained models such as ResNet-50 obtained a test accuracy of about 98%, 1D CNN obtained 92.7% and CNN-LSTM obtained 95% in ECG-based arrhythmia classification. Our experimental results demonstrate the efficacy of the fine-tuned CNN and the CNN-LSTM model in the automatic classification of arrhythmia types from ECG signals.

1. Introduction

Cardiovascular disorders are a leading cause of death and are highly prevalent affecting 49.5% of people over the age of 20 in the USA alone [1]. An electrocardiogram (ECG) examines the electrical activity of the heart and aids in the diagnosis of cardiovascular disorders by detecting different irregularities in patterns. ECG images have traditionally been manually evaluated by medical professionals

to evaluate a patient's heart health. But as technology developed and due to infeasibility in the manual examination of large volumes of ECG data, a wide range of automated diagnostic tools appeared [2]. Early detection and an improved prognosis are the advantages of automatic cardiac abnormality identification using ECGs. Thus, improving the detection of arrhythmia and providing vital assistance to medical professionals [2]. An electrocardiogram (ECG) report is obtained using a Holter monitor and a typical 12-lead design made up of three limb leads, three pressurized limb leads, and six thoracic leads [3]. While 12-lead ECGs are often employed in such applications, using fewer leads allows for inexpensive, portable, and user-friendly point-of-care equipment [4]. Apart from conventional machine learning techniques like feed-forward neural networks, decision trees, and support vector machines[5–7], deep learning techniques have lately received significant attention in automatic cardiac abnormality identification using ECGs [8–11].

In this paper, we **propose** and assess the applicability of a deep learning-based solution for classifying 4 different classes of cardiac anomalies using stacked time-frequency scalogram images from 12-lead ECG data.

2. Related work

Wu et al. [12] proposed an efficient 1D-CNN framework consisting of 12 layers. Additionally, the authors implemented a denoising technique based on wavelet-based filtering. Niu et al. [3] proposed a novel adversarial deep-learning architecture for ECG classification. The authors enhanced the dataset size with the aid of domain adaptation for improved generalization. A new deep learning framework using a combination of CNN and bidirectional LSTM models was proposed by Liang et al.[11]. For ECG classification, Abdullah and Ani [13] proposed a combination of 1D CNN and long short-term memory (LSTM) based system. For the 10-class classification of arrhythmias, Abdalla et al. [14] used an 11-layer convolutional neural network (CNN) model. A bidirectional long-short-term memory network based on Recurrent Neural Network(RNN)

was proposed by Kim and Pyun [15] for the classification of ECG signals.

Alarsan [10] evaluated multiple machine-learning methods for the classification of heart rhythms. They used the MLlib, a machine learning library, to evaluate the machine learning models. The authors suggested that the random forest classifier obtained the highest accuracy. A hybrid CNN-LSTM model for categorizing arrhythmia was proposed by Liu et al. [16].

To effectively classify arrhythmia, Zhang et al. [17] used a 1D CNN model with a global average pooling layer. Convolutional neural networks were used by Mattila et al. [18] to classify arrhythmias for patients. For the ECG classification task, T Rahhal et al.[19] used a pre-trained VGGNet model. They transformed the signals into the time domain using a continuous wavelet transform. A bi-directional LSTM-based deep learning ECG classifier was proposed by Yildirim [20] for ECG classification.

3. Methods

3.1. Dataset

Twelve lead electrocardiograms (ECG) of length 10 sec from 45,152 individuals from Shaoxing People’s Hospital (SPH) dataset [21, 22] from PhysioNet with four different types of arrhythmias i.e., atrial fibrillation (AFIB), supraventricular tachycardia (ST), sinus bradycardia (SB), and sinus rhythm (SR) is used in this study. The sampling frequency utilized was 500 Hz for recording signals. The institutional review boards at Shaoxing People’s Hospital and Ningbo First Hospital approved the study and granted the waiver request to acquire informed consent.

3.2. Time frequency representations of ECG signals

All ECG signals are pre-processed using median filtering. Then, for every 1 sec of ECG signal, the time-frequency (TF) scalogram was estimated, resulting in 10 TF scalograms for each ECG signal. TF scalograms are calculated from each channel and stacked row-wise to obtain a single image from 12 channels for 1 sec, as shown in Figure 1. These stacked time-frequency scalogram images are then fed to pre-trained deep learning models such as ResNet50 [23], EfficientNetB0 [24] and MobilnetV2[25] for ECG Arrhythmia classifications. Additionally, these images are fed to the CNN model developed from scratch with self-attention and multi-head attention mechanism for ECG Arrhythmia classification. For 1D CNN models, scalogram images are wrapped to a 1D feature vector.

Table 1. The architecture of 2D CNN used in this work.

Layer Type	Output Dimensions	Parameters
Conv	223x223x32	2,432
Max-Pool	111x111x32	0
Conv	107x107x64	51,264
Max-Pool	53x53x64	0
FC (with Dropout)	500	9,012,500
FC	256	128,256
FC	64	16,448
FC (Output)	4	260

3.3. Pre-trained Deep Learning Models

This work utilizes different deep learning models such as ResNet50 [23], EfficientNetB0 [24] and MobilNetV2 [25] for ECG classification. These models are fine-tuned using TF scalogram images with a batch size of 32, a learning rate of 0.001, and Adam optimizer[26]. These models are pre-trained ImageNet models and perform well for Arrhythmia classification after fine-tuning.

3.4. Convolutional Neural Network (CNN) Architecture

For the purpose of cross-comparison, we have used a simple CNN-based classification model. This model consists of two convolutional layers, followed by max-pooling layers; then feature maps are flattened into a single vector, which is then passed through a series of fully connected layers of size 500, 256, and 64 neurons. This is followed by a dropout layer (Dropout rate of 0.5) and the final classification layer of size 4 neurons, corresponding to four classes.

The self- and multi-head attention is applied to the basic CNN to obtain attention-weighted features. To this front, the input feature map is converted into query (Q), key (K), and value (V) representations through independent convolutional layers, with Q and K being reduced to 1/8th of the input channels. Q and K are then used to compute attention weights, to prioritize relevant spatial areas. For 1D-CNN-LSTM architecture, the scalogram images are resized to a 1D vector for input to the CNN model. The convolutional layers are followed by a three-layer LSTM of size 128 hidden units, followed by a dropout and the final classification layer of 4 neurons.

3.5. Performance Metrics

The standard performance evaluation metrics i.e., precision, recall, accuracy, and specificity [27, 28], were employed to assess the performance of the deep-learning models for arrhythmia classification.

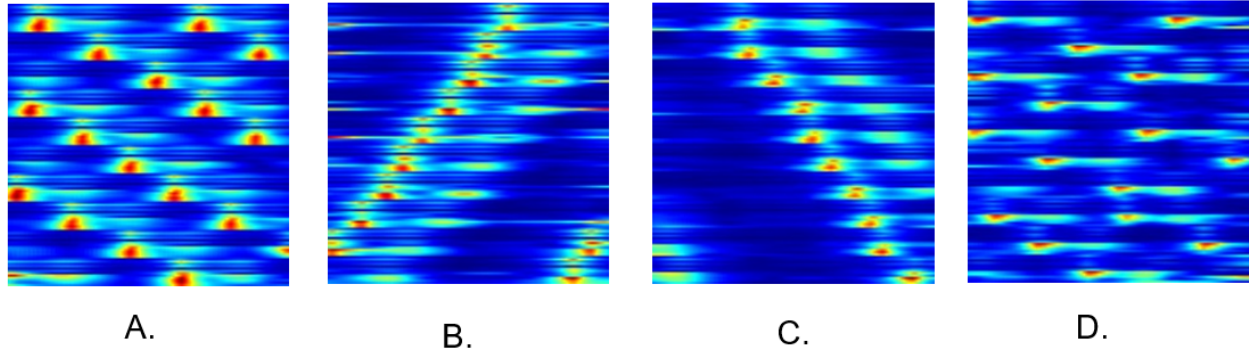


Figure 1. Time-frequency Scalogram stacked row-wise to obtain a single image from a 12 channel sample ECG signal for the four arrhythmia classes, A. supraventricular tachycardia (ST), B. sinus rhythm (SR), C. sinus bradycardia (SB), and atrial fibrillation (AFIB), respectively.

Table 2. Comparison of Deep-learning Model for ECG-based Arrhythmia Classification using Accuracy (A), Precision (P), Recall (P) and Specificity (S) when evaluated on the SPH test set.

Model	A(%)	P	R	S
ResNet 50 [23]	98.29	0.96	0.96	0.99
EfficientNet-B0 [24]	98.26	0.98	0.98	0.99
MobileNet-V2[25]	98.23	0.97	0.97	0.99
1D-CNN	92.78	0.84	0.83	0.94
1D-CNN-LSTM	95.07	0.90	0.89	0.97
Simple CNN	89.76	0.80	0.79	0.93
CNN-SA	92.48	0.81	0.81	0.94
CNN-MHA	91.52	0.79	0.78	0.93

4. Results and Discussion

Table 2 shows the precision, recall, accuracy, and specificity of different deep-learning models for arrhythmia classification. It can be seen that the best performance is obtained by the fine-tuned ResNet-50 model with precision, recall, specificity, and accuracy of 0.96, 0.96, 0.99, and 98.29, respectively. Overall, the fine-tuned version of the pre-trained models obtained better performance than the CNN models trained from scratch. Among the CNN models trained from scratch, 1D-CNN+LSTM obtained the best performance with precision, recall, specificity, and accuracy of 0.90, 0.89, 0.96, and 95.07, respectively. The least performance is obtained by the basic 2D CNN model without an attention mechanism.

A similar study in [29] used the SPH dataset and extracted the TF scalogram images from the ECG signals, separately for each channel, for training the Squeeze-Net model. The model obtained an accuracy score of 0.36 for arrhythmia classification using ECG signals. This suggests the efficacy of using stacked Scalogram images from all the leads together (obtaining around 98.29% accuracy) us-

ing our study for ECG-based arrhythmia classification.

5. Conclusions

In this study, we evaluated the efficacy of stacked time-frequency scalogram images from 12-lead ECG data for arrhythmia classification. Experimental investigation on the SPH dataset using different deep-learning architectures suggested the efficacy of the fine-tuned CNN models pretrained on ImageNet weights for arrhythmia classification. On cross-comparison with existing studies on deep-learning models based on a single TF scalogram image for each lead, stacked TF-based scalogram along with the fine-tuned SOTA CNN can obtain a significant performance improvement. As a part of future work, cross-dataset evaluation of the deep-learning models trained on the SPH dataset will be conducted. Further, the evaluation will be conducted on an extended list of arrhythmia classes using advanced deep-learning models based on self-supervised learning.

References

- [1] Virani SS, Alonso A, Aparicio HJ, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, Chamberlain AM, Cheng S, Delling FN, et al. Heart disease and stroke statistics-2021 update: a report from the american heart association. *Circulation* 2021; 143(8):CIR0000000000000950.
- [2] Rafi TH, Ko YW. Heartnet: Self multihead attention mechanism via convolutional network with adversarial data synthesis for ecg-based arrhythmia classification. *IEEE Access* 2022;10:100501–100512.
- [3] Niu L, Chen C, Liu H, Zhou S, Shu M. A deep-learning approach to ecg classification based on adversarial domain adaptation. In *Healthcare*, volume 8. MDPI, 2020; 437.
- [4] Wickramasinghe NL, Athif M. Multi-label cardiac abnormality classification from electrocardiogram using deep

- convolutional neural networks. In 2021 Computing in Cardiology (CinC), volume 48. IEEE, 2021; 1–4.
- [5] Soman T, Bobbie PO. Classification of arrhythmia using machine learning techniques. *WSEAS Transactions on computers* 2005;4(6):548–552.
- [6] Saboori R, Salehi AW, Vaidya P, Gupta G. Classification of arrhythmia using machine learning techniques. In *Innovations in Information and Communication Technologies (IICT-2020) Proceedings of International Conference on ICRIHE-2020, Delhi, India: IICT-2020*. Springer, 2021; 445–452.
- [7] Aziz S, Ahmed S, Alouini MS. Ecg-based machine-learning algorithms for heartbeat classification. *Scientific reports* 2021;11(1):18738.
- [8] Lyon A, Mincholé A, Martínez JP, Laguna P, Rodriguez B. Computational techniques for ecg analysis and interpretation in light of their contribution to medical advances. *Journal of The Royal Society Interface* 2018;15(138):20170821.
- [9] Ribeiro AH, Ribeiro MH, Paixão GM, Oliveira DM, Gomes PR, Canazart JA, Ferreira MP, Andersson CR, Macfarlane PW, Meira Jr W, et al. Automatic diagnosis of the 12-lead ecg using a deep neural network. *Nature communications* 2020;11(1):1760.
- [10] Alarsan FI, Younes M. Analysis and classification of heart diseases using heartbeat features and machine learning algorithms. *Journal of big data* 2019;6(1):1–15.
- [11] Liang Y, Yin S, Tang Q, Zheng Z, Elgendi M, Chen Z. Deep learning algorithm classifies heartbeat events based on electrocardiogram signals. *Frontiers in Physiology* 2020; 11:569050.
- [12] Wu M, Lu Y, Yang W, Wong SY. A study on arrhythmia via ecg signal classification using the convolutional neural network. *Frontiers in computational neuroscience* 2021; 14:564015.
- [13] Abdullah LA, Al-Ani MS. Cnn-lstm based model for ecg arrhythmias and myocardial infarction classification. *Adv Sci Technol Eng Syst* 2020;5(5):601–606.
- [14] Abdalla FY, Wu L, Ullah H, Ren G, Noor A, Mkindu H, Zhao Y. Deep convolutional neural network application to classify the ecg arrhythmia. *Signal Image and Video Processing* 2020;14:1431–1439.
- [15] Kim BH, Pyun JY. Ecg identification for personal authentication using lstm-based deep recurrent neural networks. *Sensors* 2020;20(11):3069.
- [16] Liu F, Zhou X, Wang T, Cao J, Wang Z, Wang H, Zhang Y. An attention-based hybrid lstm-cnn model for arrhythmias classification. In 2019 International Joint Conference on Neural Networks (IJCNN). IEEE, 2019; 1–8.
- [17] Zhang W, Yu L, Ye L, Zhuang W, Ma F. Ecg signal classification with deep learning for heart disease identification. In 2018 International Conference on Big Data and Artificial Intelligence (BDAl). IEEE, 2018; 47–51.
- [18] Takalo-Mattila J, Kiljander J, Soininen JP. Inter-patient ecg classification using deep convolutional neural networks. In 2018 21st Euromicro Conference on Digital System Design (DSD). IEEE, 2018; 421–425.
- [19] Al Rahhal MM, Bazi Y, Al Zuair M, Othman E, BenJdira B. Convolutional neural networks for electrocardiogram classification. *Journal of Medical and Biological Engineering* 2018;38:1014–1025.
- [20] Yildirim Ö. A novel wavelet sequence based on deep bidirectional lstm network model for ecg signal classification. *Computers in biology and medicine* 2018;96:189–202.
- [21] Zheng J, Guo H, Chu H. A large scale 12-lead electrocardiogram database for arrhythmia study (version 1.0.0). *PhysioNet 2022* Available online http://physionet.org/content/ecg_arrhythmia10_0 accessed on 23 November 2022;.
- [22] Zheng J, Chu H, Struppa D, Zhang J, Yacoub SM, El-Askary H, Chang A, Ehwerhemuepha L, Abudayyeh I, Barrett A, et al. Optimal multi-stage arrhythmia classification approach. *Scientific reports* 2020;10(1):2898.
- [23] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016; 770–778.
- [24] Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*. PMLR, 2019; 6105–6114.
- [25] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018; 4510–4520.
- [26] Kingma DP, Ba J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* 2014;.
- [27] Diker A, Cömert Z, Avcı E, Toğaçar M, Ergen B. A novel application based on spectrogram and convolutional neural network for ecg classification. In 2019 1st International Informatics and Software Engineering Conference (UBMYK). IEEE, 2019; 1–6.
- [28] Fawcett T. An introduction to roc analysis. *Pattern recognition letters* 2006;27(8):861–874.
- [29] Cornely AK, Carrillo A, Mirsky GM. Reduced-lead electrocardiogram classification using wavelet analysis and deep learning. In 2021 Computing in Cardiology (CinC), volume 48. IEEE, 2021; 1–4.

Address for correspondence:

Dr. Ajita Rattani
 Dept of Computer Science and Engineering,
 University of North Texas at Denton, USA.
ajita.rattani@unt.edu