2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Segment Origin Prediction: A Self-supervised Learning Method for Electrocardiogram Arrhythmia Classification

Chuankai Luo[1], Guijin Wang[1,*], Zijian Ding[1], Hui Chen[1], Fan Yang[1]

Abstract— **The automatic arrhythmia classification system has made a significant contribution to reducing the mortality rate of cardiovascular diseases. Although the current deep-learning-based models have achieved ideal effects in arrhythmia classification, their performance still needs to be further improved due to the small scale of the dataset. In this paper, we propose a novel self-supervised pre-training method called Segment Origin Prediction (SOP) to improve the model's arrhythmia classification performance. We design a data reorganization module, which allows the model to learn ECG features by predicting whether two segments are from the same original signal without using annotations. Further, by adding a feed-forward layer to the pre-training stage, the model can achieve better performance when using labeled data for arrhythmia classification in the downstream stage. We apply the proposed SOP method to six representative models and evaluate the performances on the PhysioNet Challenge 2017 dataset. After using the SOP pre-training method, all baseline models gain significant improvement. The experimental results verify the effectiveness of the proposed SOP method.**

I. INTRODUCTION

Arrhythmia can cause irregular rhythms and increase the risk of stroke and sudden cardiac death. With the popularization of single-lead devices such as AliveCor, the use of electrocardiogram (ECG) to diagnose arrhythmia can effectively reduce its mortality.

Due to the powerful feature learning capabilities of DNNs, some classical networks such as CNNs and RNNs have been designed to detect and classify arrhythmia. The CNN based model [1] obtained the best score in the PhysioNet Challenge 2017 competition [2]. Variant structures of ResNet [3][4] and structures of SE-block [5][6][7] are current trend in the field of arrhythmia classification. In addition, there are also variant structures of DenseNet [8][9], long short-term memory (LSTM) [10], and Convolutional Recurrent Neural Network (CRNN) [11] that achieved ideal results in the classification of arrhythmia.

Although the above-mentioned deep learning methods have made remarkable progress, the way to design a complex network structure to improve classification performance has encountered a bottleneck. At present, accurately labeled ECG datasets are generally small in scale (for example, the MIT-BIH Arrhythmia Database [12] consists of 47 subjects), causing these data-driven models to encounter overfitting problems in arrhythmia classification. As an effective method to improve the model's classification performance, self-supervised learning can alleviate the model overfitting problem by obtaining a suitable initialization weight in the pre-training stage.

Pioneered studies of self-supervised learning methods based on ECG data have shown the potential to improve model performances. Sarkar et al. [13] pioneered the effectiveness of self-supervised learning in the field of ECG-based emotion recognition. Cheng et al. [14] explored the effect of self-supervised learning in the classification of arrhythmia and proved that promoting subject invariance can improve classification performance. CLOCS [15] achieves strong generalization performance by designing patient-level self-supervised learning tasks with only a few labels. In general, both [14] and [15] prove the effectiveness of designing patient-level self-supervised learning tasks for arrhythmia classification.

In this paper, we propose a self-supervised method for arrhythmia classification named **SOP**, which stands for **S**egment **O**rigin **P**rediction. The SOP method allows the model to obtain a suitable initialization weight in the pre-training stage, thereby improving the classification effect of arrhythmia in the downstream stage. We design a data reorganization module, which allows the model to learn ECG features by predicting whether two segments are from the same original signal without using annotations. Further, by adding a feed-forward layer to the pre-training stage, the model can achieve better performance when using labeled data for arrhythmia classification in the downstream stage. We apply the proposed SOP method to six representative models and evaluate the performances on the PhysioNet Challenge 2017 dataset [2]. After using the SOP pre-training method, all baseline models gain significant improvement. In particular, some of the models achieve competitive performances compared with state-of-the-arts. We also find that the classification performance will be further improved if external data is introduced.

II. METHOD

As shown in Fig. 1, the proposed SOP method's overall framework consists of two parts: the pre-training stage and the downstream stage. Since self-supervised learning can obtain prior information without using annotations, the only input accepted in the pre-training stage is just data without labels. Similar to the self-supervised learning process in biological signals [16][14], the pre-training stage will treat the ECG signal as unlabeled data and fine-tune with data and labels in the downstream stage. In the pre-training stage, the unlabeled data will first pass through the data reorganization module to automatically generate labels and transform them

[1]Department of Electronic Engineering, Tsinghua University, Beijing, China.
*Corresponding author

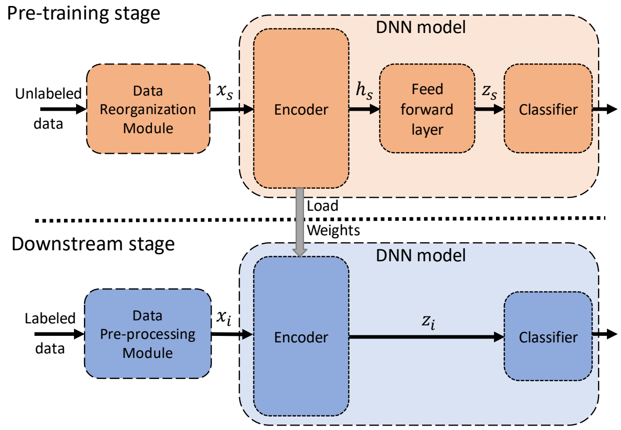

Fig. 1. The framework of the proposed SOP method.

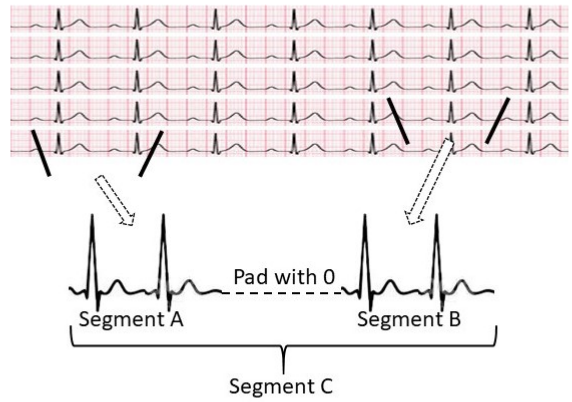

Fig. 2. Data reorganization module

into brand new labeled data x_s. And then x_s will be sent into the DNN model for training. We added an extra feed-forward layer to the model at this stage to enhance the model's characterization ability. The model weights obtained in the pre-training stage will be saved. In the downstream arrhythmia classification stage, the model will load the previously saved weights as the model's initialization and be fine-tuned on labeled data. We will explain some essential parts below.

A. Data Reorganization Module

According to the self-supervised learning process, unlabeled data needs to be reorganized to obtain automatically generated labels. Fig. 2 shows how our data reorganization module works.

- Data: First, we cut all the raw ECG signals into small segments with the same length. After that, for any segment A obtained after cutting, another segment B is randomly selected from the remaining segment set. A and B are concatenated into a brand-new segment C with some zeros padded in the middle.
- Label: We will generate a set of binary labels according to the way to generate segment C. If the two segments (A and B) in C are from the same original ECG signal, then the label of C is 1; otherwise, the label of C is 0.

After the data reorganization module, the unlabeled ECG signal will be converted into binary labeled data x_s so that the model can learn prior information through supervised learning.

B. Feed-forward Layer

Inspired by simCLR [17], a feed-forward layer composed of a fully connected layer and Relu activation is added to the model at the pre-training stage. The feed-forward layer can integrate the previously highly abstracted features after multiple convolutions and enhance the model's nonlinearity and learning ability. As shown in Fig. 1, the signal h_s will be converted into z_s, after the feed-forward layer (shown in Eq.1 where σ is a ReLU activation function).

$$z_s = g(h_s) = \sigma(Wh_s) \tag{1}$$

C. DNN Models

SOP method is model independent, so the DNN model in Fig. 1 can be replaced with any network. In order to verify the generality of SOP, we test six widely used and state-of-the-art models [18][19][11][20][5][1] in arrhythmia classification. Among them, [11] won first place in the CPSC 2018 competition [21]; All of the top five algorithms in the China ECG AI Contest 2019 competition [22] are all designed on the basis of [19][18]. [1] is currently widely recognized as the state-of-the-art model for the PhysioNet Challenge 2017 [2]; SE-block [20] with a large kernel size model [5] achieved second place in the PhysioNet 2020 competition [23].

D. Downstream Stage Framework

The task of the downstream stage is the arrhythmia classification problem that we actually need to solve.

The labeled data related to arrhythmia classification will pass through the data preprocessing module to obtain x_i as shown in Fig. 1. Unlike the data reorganization module, the data preprocessing module in the downstream stage has only denoising and random cropping processes.

The DNN model structure in the downstream stage will be consistent with the model we need to test, so there is no extra feed-forward layer. Besides, the output classes of the classifier will correspond to the actual number of classifications of arrhythmia.

After the model is built, the model weights obtained in the pre-training stage will be loaded as initial weights. A good model initialization can prevent the model from falling into a local saddle point during training, thereby improving the final classification effect. After the initial weight is obtained, the model will be fine-tuned on this basic.

III. EXPERIMENT

Here we lay out the protocol for our empirical studies to understand different design choices in our framework.

A. Database

The downstream dataset is what we actually need to classify arrhythmia on it.

The PhysioNet/CinC Challenge 2017 [2] published dataset is a large-scale, long-term single-lead arrhythmia classification dataset. Some arrhythmia researches [6][4][1][10] have been verified on it, which is very authoritative. It consists of 8528 recordings of single-lead ECG data, lasting between 9 s and 61 s. PhysioNet Challenge 2017 [2] are labeled with four classes: (1) normal sinus rhythm (N for short, 5076 records), (2) AF (A for short, 758 records), (3) alternative rhythm (O for short, 2415 records), and (4) noisy recordings (P for short, 279 records).

To evaluate algorithms' performance on the downstream stage, we fully follow the requirements of the competition organizing committee [2] and used F1 score to reflect the overall index. The final F1 is defined as:

$$F_1 = \frac{F_1N + F_1A + F_1O}{3} \tag{2}$$

The SOP pre-training stage dataset is not required to be consistent with the downstream stage's dataset, which means that we can introduce external data through the SOP method. The China Physiological Signal Challenge 2018 database [21] (including its hidden dataset provided by [23]) is a large open-source arrhythmia dataset released by the Seventh International Conference on Biomedical Engineering and Biotechnology. This dataset is currently the long-term (from 6 s to 60 s) and the authoritative largest dataset in the field of arrhythmia classification. CPSC 2018 dataset [21] consists of 10,330 12-lead ECG recordings. Each recording was sampled at 500 Hz.

B. Experiment Implementation Details

We conducted three sets of experiments: using the PhysioNet Challenge 2017 dataset [2] in the pre-training and downstream stage to explore the effects of the SOP method (*PhysioNet2017 experiment*), exploring the effect of combining external dataset to pre-training dataset (*CPSC2018+PhysioNet2017 experiment*), and exploring the effect of adding a feed-forward layer (*feed-forward layer effection experiment*).

PhysioNet2017 experiment is to verify whether the proposed SOP method can improve the model's classification performance. The pre-training and downstream datasets in this experiment are both PhysioNet Challenge 2017 [2]. For each ECG sample, we use the Discrete Wavelet Transform algorithm [24] to denoise. At the pre-training stage, all data has been removed from the label to simulate unlabeled signals. For the actual downstream arrhythmia classification task, except Hannun's model [1] was cut into 256 lengths as article's requirements, each signal was randomly cropped to 9000 lengths. We train at batch size 32 for 100 epochs at the downstream stage.

CPSC2018+PhysioNet2017 experiment. The pre-training stage dataset can be different from the downstream stage, which means that we can introduce external data through the proposed SOP method. Inspired by [17][25][15], the introduction of large-scale external data in the pre-training stage can further improve the final classification performance of the model. To verify this conjecture, this experiment uses the combination of CPSC 2018 [21] and PhysioNet Challenge 2017 [2] as the unlabeled pre-training dataset (CPSC 2018 dataset [21] is used to simulate the imported external data). The CPSC 2018 dataset [21] only keeps the data from the first lead, and the frequency is also reduced from 500hz to 300hz to consistent with the downstream dataset. Except for the dataset, all settings are the same as in the first experiment.

TABLE I

THE F1 SCORE OF EACH ARRHYTHMIA CLASSIFICATION MODEL ON THE PHYSIONET CHALLENGE 2017 DATABASE AFTER USING THE SOP PRE-TRAINING METHOD

Models	*Random init*	*After SOP method*	
		PhysioNet2017	CPSC2018 + PhysioNet2017
Huang et al.[18]	0.766	0.785	0.796
He et al.[19]	0.808	0.830	0.837
Chen et al.[11]	0.820	0.829	0.836
Hu et al.[20]	0.833	0.848	0.854
Zhao et al.[5]	0.849	0.856	0.861
Hannun et al.[1]	0.852	0.863	0.875

Feed-forward layer effection experiment is to verify the benefits of the extra feed-forward layer. This experiment will compare the downstream classification performance of whether there is a feed-forward layer in the pre-training stage. The pre-training and downstream datasets are both PhysioNet Challenge 2017 [2]

IV. RESULT

We show the performance on the experiment with the PhysioNet Challenge 2017 database [2] as the downstream dataset and the comparison results with several widely used models [18][19][11][20][5][1] in Table. I. The second column (Random init) shows the basic performance of models without using SOP method.

The results of *PhysioNet2017 experiment* show that all baseline models gain significant improvement by the SOP method. The F1 score of [1][20][5] methods exceeded or approached the 0.85 level. This experiment result demonstrates that the proposed SOP method can improve the model's classification performance by obtaining a suitable initialization weight that contains prior information of the ECG signal.

The *CPSC2018+PhysioNet2017 experiment* results show that the six models' classification performance has been further improved. Specifically, after using the SOP method, the F1 score of the previous state-of-the-art model [1] increased to 0.875. This experiment result demonstrates the introduction of external data can further improve the performance of

TABLE II
THE FEED-FORWARD LAYER'S EFFECT OF ARRHYTHMIA CLASSIFICATION AT PHYSIONET CHALLENGE 2017 DATASET

Models	*Without feed-forward layer*	*With feed-forward layer*
Huang et al.[18]	0.760	**0.785**
He et al.[19]	0.821	**0.830**
Chen et al.[11]	0.826	**0.829**
Hu et al.[20]	0.836	**0.848**
Zhao et al.[5]	0.852	**0.856**
Hannun et al.[1]	0.857	**0.863**

the model. It is worth noting that this experiment only uses the CPSC 2018 dataset [21] as an example of external data to obtain performance improvements. It is foreseeable that if more external data is used, performance may be further improved.

The results shown in Table. II prove the effectiveness of adding the feed-forward layer in the pre-training stage for downstream ECG arrhythmia classification tasks. After removing the feed-forward layer, the final classification effects of models will decrease.

V. CONCLUSIONS

In this paper, we propose a self-supervised pre-training method named SOP to enhance the ECG arrhythmia classification performance. The use of the data reorganization module and feed-forward layer allows the model to obtain prior information and suitable initialization weights in the pre-training stage, thereby improving model performance in the downstream stage. We also found that adding external data to the pre-training dataset can further improve the model's final classification effect. Without using additional annotations, all six widely used models have been improved. In particular, based on the original excellent performance model, we have achieved a 0.852 to 0.875 F1 score improvement, achieving state-of-the-art results on the PhysioNet Challenge 2017 database [2].

In theory, our SOP method can be used not only for single-lead arrhythmia diagnosis but also for diagnosing diseases of other multi-lead ECGs, which will have potential practical value in the future.

REFERENCES

[1] A. Y. Hannun, P. Rajpurkar, M. Haghpanahi, G. H. Tison, C. Bourn, M. P. Turakhia, and A. Y. Ng, "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nature medicine*, vol. 25, no. 1, pp. 65–69, 2019.

[2] G. D. Clifford, C. Liu, B. Moody, H. L. Li-wei, I. Silva, Q. Li, A. Johnson, and R. G. Mark, "Af classification from a short single lead ecg recording: the physionet/computing in cardiology challenge 2017," in *2017 Computing in Cardiology (CinC)*. IEEE, 2017, pp. 1–4.

[3] F. Andreotti, O. Carr, M. A. Pimentel, A. Mahdi, and M. De Vos, "Comparing feature-based classifiers and convolutional neural networks to detect arrhythmia from short segments of ecg," in *2017 Computing in Cardiology (CinC)*. IEEE, 2017, pp. 1–4.

[4] Z. Xiong, M. P. Nash, E. Cheng, V. V. Fedorov, M. K. Stiles, and J. Zhao, "Ecg signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network," *Physiological measurement*, vol. 39, no. 9, p. 094006, 2018.

[5] Z. Zhao, H. Fang, S. D. Relton, R. Yan, Y. Liu, Z. Li, J. Qin, and D. C. Wong, "Adaptive lead weighted resnet trained with different duration signals for classifying 12-lead ecgs."

[6] J. Zhu, Y. Zhang, and Q. Zhao, "Atrial fibrillation detection using different duration ecg signals with se-resnet," in *2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP)*. IEEE, 2019, pp. 1–5.

[7] J. Zhu, K. Xin, Q. Zhao, and Y. Zhang, "A multi-label learning method to detect arrhythmia based on 12-lead ecgs," in *Machine Learning and Medical Engineering for Cardiovascular Health and Intravascular Imaging and Computer Assisted Stenting*. Springer, 2019, pp. 11–19.

[8] J. Lai, Y. Chen, B. Han, L. Ji, Y. Shi, Z. Huang, W. Yang, and Q. Feng, "A densenet-based diagnosis algorithm for automated diagnosis using clinical ecg data," *Nan Fang yi ke da xue xue bao= Journal of Southern Medical University*, vol. 39, no. 1, pp. 69–75, 2019.

[9] X. Tao, H. Dang, D. Xiong, R. Liu, D. Liu, and F. Zhou, "Automated detection of atrial fibrillation based on densenet using ecg signals," in *Proceedings of the 2020 International Symposium on Artificial Intelligence in Medical Sciences*, 2020, pp. 69–73.

[10] T. Teijeiro, C. A. García, D. Castro, and P. Félix, "Arrhythmia classification from the abductive interpretation of short single-lead ecg records," in *2017 Computing in cardiology (cinc)*. IEEE, 2017, pp. 1–4.

[11] T.-M. Chen, C.-H. Huang, E. S. Shih, Y.-F. Hu, and M.-J. Hwang, "Detection and classification of cardiac arrhythmias by a challenge-best deep learning neural network model," *Iscience*, vol. 23, no. 3, p. 100886, 2020.

[12] G. B. Moody and R. G. Mark, "The impact of the mit-bih arrhythmia database." *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2002.

[13] P. Sarkar and A. Etemad, "Self-supervised ecg representation learning for emotion recognition," *IEEE Transactions on Affective Computing*, 2020.

[14] J. Y. Cheng, H. Goh, K. Dogrusoz, O. Tuzel, and E. Azemi, "Subject-aware contrastive learning for biosignals," *arXiv preprint arXiv:2007.04871*, 2020.

[15] D. Kiyasseh, T. Zhu, and D. A. Clifton, "Clocs: Contrastive learning of cardiac signals," *arXiv preprint arXiv:2005.13249*, 2020.

[16] K. Chaitanya, E. Erdil, N. Karani, and E. Konukoglu, "Contrastive learning of global and local features for medical image segmentation with limited annotations," 2020.

[17] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *International conference on machine learning*. PMLR, 2020, pp. 1597–1607.

[18] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[20] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7132–7141.

[21] F. Liu, C. Liu, L. Zhao, X. Zhang, X. Wu, X. Xu, Y. Liu, C. Ma, S. Wei, Z. He *et al.*, "An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection," *Journal of Medical Imaging and Health Informatics*, vol. 8, no. 7, pp. 1368–1373, 2018.

[22] The China ECG AI Contest 2019, http://mdi.ids.tsinghua.edu.cn.

[23] E. A. P. Alday, A. Gu, A. J. Shah, C. Robichaux, A.-K. I. Wong, C. Liu, F. Liu, A. B. Rad, A. Elola, S. Seyedi *et al.*, "Classification of 12-lead ecgs: the physionet/computing in cardiology challenge 2020," *Physiological measurement*, vol. 41, no. 12, p. 124003, 2020.

[24] R. J. Martis, U. R. Acharya, and L. C. Min, "Ecg beat classification using pca, lda, ica and discrete wavelet transform," *Biomedical Signal Processing and Control*, vol. 8, no. 5, pp. 437–448, 2013.

[25] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.