Electrocardiogram Classification by Modified EfficientNet with Data Augmentation

Naoki Nonaka[1], Jun Seita[1]

[1] Medical Sciences Innovation Hub Program, RIKEN, Tokyo, Japan

Abstract

The standard 12-ECG are widely used in the diagnosis of arrhythmias and other cardiac disorders. Early and correct diagnosis of cardiac abnormalities can improve treatment results. However, manual interpretation of the electrocardiogram (ECG) is time-consuming and difficult to be scaled. Thus, automatic detection and classification of cardiac abnormalities can assist physicians in the diagnosis of the growing number of ECGs recorded.

In recent years, deep neural networks (DNNs) have shown significant improvement in variety of tasks, including ECG classification. In this study, we attempt to classify 12-ECG Physionet Challenge 2020 data using DNN model. We adopt EfficientNet model, which achieved state-of-the-art result with ImageNet classification task, and modify model for ECG classification. During the training, we adopt data augmentation for ECG to improve the robustness of the model.

With proposed method, we achieve score of 0.581 (Team name: NN-MIH).

1. Introduction

The standard 12-ECG are widely used in the diagnosis of arrhythmias and other cardiac disorders. If an ECG can be read accurately and cardiac abnormalities can be detected early, there is a good chance of improving the prognosis. However, reading an ECG requires a highly trained professional to perform the task and is time consuming. The ability to automatically classify ECGs and detect abnormalities would aid physicians in diagnosis and allow them to handle more ECGs.

Deep neural networks (DNNs) have driven substantial advances and demonstrated dramatic improvement of state of the art in tasks like image recognition, machine translation and speech recognition [1–5]. DNNs have been applied to ECG classification tasks and several works have shown DNNs can detect irregular ECGs without the needs of feature engineering [6, 7] with result exceeding cardiologist.

Thus, in this study, we attempt to classify 12 leads ECG Physionet Challenge 2020 data using DNN model. The DNN model has been actively studied in the field of image processing. EfficientNet [8] shows high accuracy in classification tasks using Image Net, which is a typical dataset in image processing. Therefore, in this study, as a DNN model to classify ECGs, we use a modified EfficientNet model to handle ECGs. The main ideas of our approach are the following three points.

- Use modified EfficientNet for ECG classification
- Relabel samples to overcome label and short length signal data mismatch
- Apply data augmentation to improve robustness of model.

The rest of the paper is organized as follows: the overview and preprocessing of data is described in Section 2, the model architecture used for ECG data classification is described in Section 3. In Section 4, we described data augmentation used during model training, in Section 5 we describe details of training setting. In Section 6 we explain results obtained and study is concluded at Section 7.

2. Preprocess of data

We first describe sequence of preprocessing applied to the provided datasets. As shown in Table 1, six datasets with different collection conditions and lengths was provided for the Physionet Challenge 2020. In order to train model using all datasets provided, we aligned datasets. Specifically, we aligned the frequencies, normalized the range of amplitudes, and finally aligned the series lengths to be equal. The frequencies were aligned to 100 Hz, so that only values at each $100/f$ point were recorded according to the frequency f of each dataset. Subsequently, the amplitudes of each sample were normalized. To exclude the effect of extremely large values of amplitude, we normalized using the maximum and minimum values after excluding the upper and lower 0.1% of points. Finally, each sample was divided into 10-second intervals in order to align the series length between samples. If sample was less than 10 seconds length, zero padding was performed to align the length of the sample.

Table 1. Overview of datasets.

Dataset	Number of samples	Length of original data	Sampling frequency (Hz)	Number of aligned samples
CPSC	6,877	10 - 150 sec	500	10,502
CPSC Extra	3,453	10 - 100 sec	500	5,242
St. Petersburg	75	30 min	257	13,320
PTB	549	30 - 130 sec	1000	5,895
PTB-XL	21,873	10 sec	500	21,837
Georgia	10,344	10 sec	500	10,292

3.   Model architecture

In this study, an modified version of EfficientNet [8], a DNN model that achieves high accuracy in image classification, was used to classify ECGs. DNNs are known to be highly accurate in a wide range of tasks such as image processing [1, 2] and natural language processing [3, 4], and previous studies have reported that DNNs are also accurate in ECG classification[6, 7]. Most of the DNN-based ECG classification methods used so far have used ResNet-based classification models. On the other hand, in the area of image processing, where DNNs have been actively studied, EfficientNet [8] has achieved better classification accuracy than ResNet [2] with fewer parameters in the ImageNet classification task, which is a typical benchmark dataset for image classification. Preliminary experiments were conducted to compare ECG classification models based on ResNet and EfficientNet, and the results showed that the ResNet model required about 10 times more parameters to achieve the same level of accuracy as EfficientNet. Thus, in this study, EfficientNet modified to handle one-dimensional ECG data, was used for the classification task.

The EfficientNet used for classification is described in this paper. The overview of the whole network and the individual blocks are shown in Figure 1. The model consists of two parts: the EfficientNet backbone part, which extracts features from the input ECGs, and the prediction block, which predicts each class label from the extracted features. The EfficientNet backbone part is based on the Mobilenetv3 [9] that is repeated 16 times. The backbone part receives 12 lead $\times$ sample length input data and outputs the features in $12,800$ dimension. The prediction block takes the $12,800$ dimensional features from the backbone data and outputs the predictions of binary class labels. Since one prediction block predicts one class label, we prepared a number of prediction blocks corresponding to the number of class labels to predict. The EfficientNet backbone and prediction block was jointly trained.

4.   Augmentation

In this study, we adopt DNNs to classify ECG data. In general, a large amount of data is required to train a DNN model, and the larger the number of data, the better the model can be trained. Therefore, we used data augmentation to increase the diversity of data by adding perturbation to the data and to train robust models. In the case of image classification, data augmentation is performed by operations such as flipping, padding and zooming. In ECG data, augmentation such as padding and changing amplitude is also possible. Recently, it has been shown that AutoAugment [10], which use reinforcement learning to find the best combination of multiple data expansion methods, and RandAugment [11], which randomly combine multiple data augmentation methods, can improve the accuracy of image classification without changing the structure of the model. Therefore, in this study, an ECG classification model was trained using a random combination of augmentation methods [12] on ECG data.

The additional hyper parameters of augmentation is magnitude of each augmentation, M and number of augmentation to apply in each batch, N. For each augmentation method, strength of augmentation was split into 20 steps. Based on grid search of M and N, we chose M as 8 and N as 3.

5.   Training setting

In this section, we explain training settings for relabeling and classification model.

5.1.   Relabeling model

We describe the training conditions for a DNN model that performs relabeling on a split sample to align the series length. The Georgia and PTB-XL datasets were used to train the relabeling model. Among the data contained in the two datasets, only samples that met the following conditions were constructed for training the relabeling model.
- The series length of the original data is 10 seconds
- Labels assigned are not assigned to the same sample for Sinus rhythm and other diseases

As a result of selecting samples that satisfied these conditions, there were $6,604$ cases of sinus rhythm and $9,855$ cases of not sinus rhythm. The dataset was divided into three datasets (train, valid and test sets) in the ratio of 8:1:1, and the parameters in the train set were used for training. The model with an EfficientNet backbone and
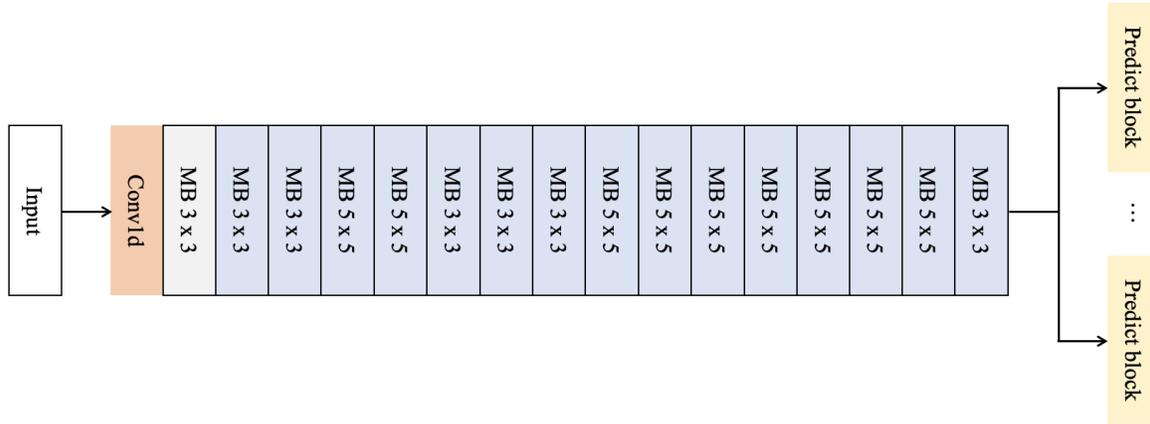
(a) Over all architecture of classification model.
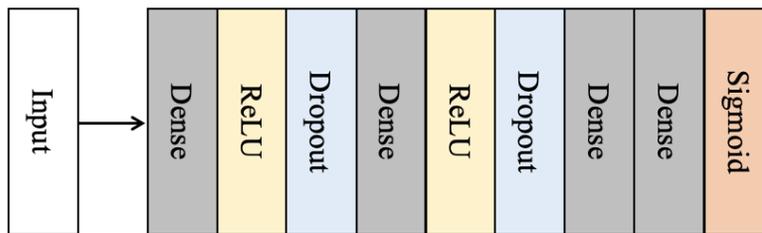
(b) Architecture of prediction block.

Figure 1. Model architecture of ECG classification model. Figure (a) shows overall architecture. MB block is MobileNetV3 [9] structure. 3×3 stands for 3×3 kernel size and 5×5 stands for 5×5 kernel size.

a prediction block was used for the training. Optimization was done using Adam [13], and the learning rate was set to 0.0001. The training batch size was set to 64 and the gradient clipping value was set to 1.

5.2. Classification model

We describe the training conditions for a model that predicts disease labels from a series-length aligned sample. The model was trained as a multi-label classification task that determines in binary whether each given disease occurs for a single sample. All six datasets provided were used for training, and the frequency and series length of each sample were aligned for training. For the PTB and St. Petersburg datasets, newly assigned labels were used by the relabeling model, while the original labels were used for the other datasets. The DNN model was designed to have an EfficientNet backbone that takes a 12-lead ECG as input and a prediction block corresponding to the number of classes to be scored. Twenty-four class labels were used to predict 24 classes, which consisted of 27 class labels that were treated as the same label among the scored classes. ECG augmentation was used to train the model, and optimization was done using Adam [13] with a learning rate of 0.001. The batch size during training was set to 512 and the gradient clipping value was set to 10.

6. Result and discussion

In this section we explain result of relabeling and classification of ECG data.

First, we relabeled dataset to mitigate the label mismatch between split data and original label. We trained deep neural network model to relabel split data. The training was carried out by giving binary labels for relabeling, either sinus rhythm or other binary labels. As a result, the accuracy of the training was 91.5%. The trained relabeling model was used to relabel the series length samples. The relabeling model was applied to the samples that were judged to be of sinus rhythm, regardless of the label of the original data. Samples judged not to be of sinus rhythm were given the same label as the label given to the original data. Relabeling was applied on all six datasets provided and used to train the classification model.

Subsequently, model was trained on a length aligned ECG data to classify cardiovascular abnormalities. In order to improve the robustness of the model being trained, ECG augmentation was applied during training. The training was performed on a relabeled dataset. The trained model was evaluated for the agreement between the original labels and the predictions for each sample before aligning the series length. As shown in Figure 2, if the original data were split, the predictions were summarized for each

class label for each sample after the split, and the largest of the predictions was the final prediction. Our model achieved score of 0.581.
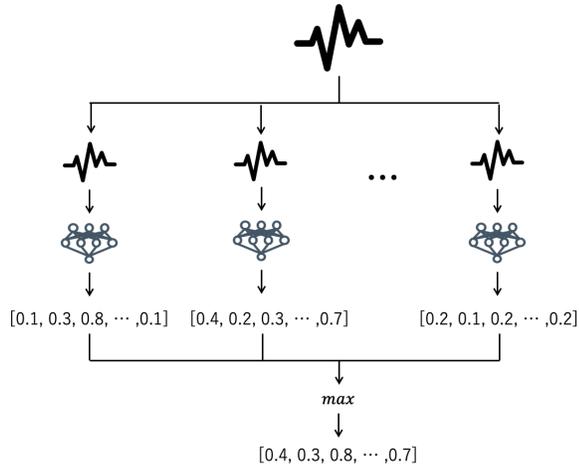

Figure 2. Overview of integrating prediction labels for validation data. We evaluate split samples and subsequently calculate max value for each class label.

7. Conclusion

In this study, EfficientNet, which has been reported to achieve high accuracy in image classification, was improved to handle ECG data, and data augmentation was applied to ECG data to train it. To deal with the possibility of label discrepancies between the original and the split data, we trained the relabeling model and relabeled the data. We trained the relabeled data with EfficientNet, and achieved a score of 0.581.

References

[1] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2015; 1–9.

[2] 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.

[3] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv14090473 2014;.

[4] Devlin J, Chang MW, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv181004805 2018;.

[5] Graves A, Mohamed Ar, Hinton G. Speech recognition with deep recurrent neural networks. In 2013 IEEE international conference on acoustics, speech and signal processing. IEEE, 2013; 6645–6649.

[6] Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, Ng AY. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nature medicine 2019; 25(1):65.

[7] Attia ZI, Noseworthy PA, Lopez-Jimenez F, Asirvatham SJ, Deshmukh AJ, Gersh BJ, Carter RE, Yao X, Rabinstein AA, Erickson BJ, et al. An artificial intelligence-enabled ecg algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. The Lancet 2019;394(10201):861–867.

[8] Tan M, Le QV. Efficientnet: Rethinking model scaling for convolutional neural networks. arXiv preprint arXiv190511946 2019;.

[9] Howard A, Sandler M, Chu G, Chen LC, Chen B, Tan M, Wang W, Zhu Y, Pang R, Vasudevan V, et al. Searching for mobilenetv3. In Proceedings of the IEEE International Conference on Computer Vision. 2019; 1314–1324.

[10] Cubuk ED, Zoph B, Mane D, Vasudevan V, Le QV. Autoaugment: Learning augmentation policies from data. arXiv preprint arXiv180509501 2018;.

[11] Cubuk ED, Zoph B, Shlens J, Le QV. Randaugment: Practical automated data augmentation with a reduced search space. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2020; 702–703.

[12] Nonaka N, Jun S. Data augmentation for electrocardiogram classification with deep neural network. submitted to arXiv preprint 2020;.

[13] Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv14126980 2014;.

Address for correspondence:

Jun Seita

Nihonbashi 1-chome Mitsui Building, 15th floor 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan

jun.seita@riken.jp