Atrial Fibrillation Detection Using Deep Learning

Geoffrey Wong (1155109054)

Abstract

Deep ResNet and ResNeXt are proposed for the detection of atrial fibrillation. Both networks are evaluated based on the dataset provided by PhysioNet / CinC Challenge 2017. The dataset is preprocessed using FFT and converted to spectrograms. The best F-measure recorded is 0.813 with 3-Fold scaling.

1. Background

Atrial fibrillation (AF) is the most common arrhythmia diagnosed. In 2010, an estimated amount of 33.5 million individuals were affected [1]. Patients with AF experiences irregular and often rapid beating in the upper chambers of hearts (the atria). Some symptoms of AF include heart palpitations, shortness of breath and weakness [2]. Although AF is not life-threatening by itself, serious AF could develop blood clots and increase the risk of stroke by four to five times. The condition contributes to estimated 130 thousand deaths per year and the number is expected to rise with the aging of population [3].

2. Motivations

Electrocardiography (ECG) diagram is a technique that is used to monitor heart rate. Its data allows the detection of AF. However, AF detection remains a challenging problem due to its episodic nature. Severe symptoms are seldomly observed, especially for paroxysmal atrial fibrillation [4]. Therefore, automatic detection is needed to facilitate diagnosis of AF, which requires heart rate monitored over a long period of time and large amount of data generated from ECG.

3. Objective

Many previous studies of AF classification with ECG generally limits the scope to classification between AF and normal rhythms with clean and carefully selected data. Therefore, the objective of this project is to train a classifier that is comparable with the result of 2017 PhysioNet Challenge with the dataset provided, which is professionally labelled into 4 categories, including AF, normal, other rhythm and noise. The winner of 2017 PhysioNet Challenge obtained a F1 score of 0.83 [5]. Considering the scope of this project is limited to a course project lasted for one semester, the goal of this project is to train a classifier that obtains an F1 score equal to or above 0.8.

4. Methodology

a. Model Used

The model of choice for this project would be ResNet [6] and ResNeXt [7], both of which have achieved impressive result in image classification. Researches have demonstrated competitive result of using ResNet for classifying AF. Customized versions are created for both ResNet and ResNeXt by converting 2-D Convolutional Layers to 1-D Convolutional Layers so that they are compatible with 1-D time series data.

b. Fast Fourier Transform

The dataset provided by PhysioNet is in time domain format. The ECG recorded spans around 9 to 60 seconds with a 30Hz/s sampling frequency. However, some researches have shown that converting the time domain data into frequency domain format using Fast Fourier Transform (FFT) boosts the performance of the network. In this project, both 1-dimensional time domain data (Fig. 1) and 2-dimensional frequency domain data (Fig. 2) would be used for training and the result would be compared. The time domain data is converted to frequency domain using the *scipy.signal.spectrogram( )* function provided by the SciPy module in Python. The data is padded with zero before inputting into the network so it has uniform shape. (1-D shape: [18300] ; 2-D shape: [65, 568])

c. Unbalanced Data

The dataset provided by PhysioNet is unbalanced, with around 61% being Normal (*N*), 7% being Atrial Fibrillation (*A*), 30% being Other (*O*) and 2% being Noise (~). The network demonstrated tendency of focusing too heavily on classifying *N* and *O* data and ignoring *A* cases. Therefore, the training data is re-balanced by duplicating random entries of *A* and *O* so the number of entries of *N*, *A* and *O* are equal. The ~ data is left unchanged since the competition did not count this category into the result.

d. Search for Optimal Structure

The optimal structure of neural networks and suitable hyperparameters are searched via a methodology similar to that of genetic algorithm. 5 different models with slight differences are trained simultaneously. The successors of each batch are then randomly selected based on the maximum F1 score that the model obtained. Changes are then made to them to produce 5 offspring for the next iteration until peak performance is reached.

The highest F1 score achieved among all models is recorded as the final result.

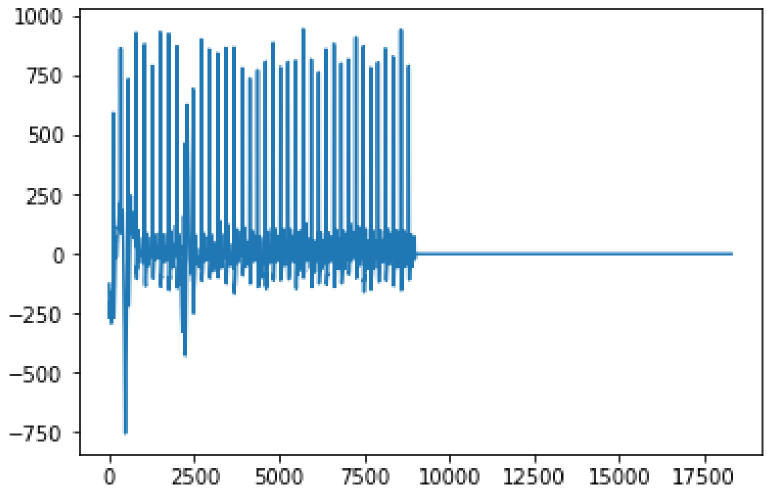

Fig. 1 Zero-padded time domain data

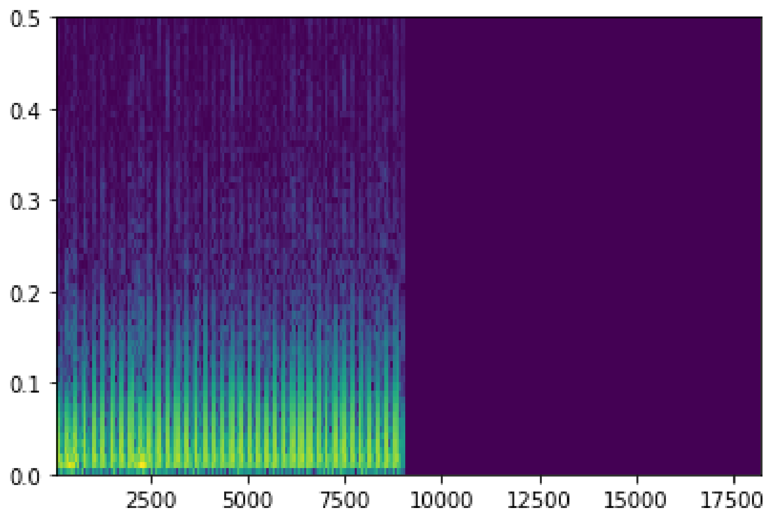

Fig. 2 Zero-padded frequency domain data

5. Results and Evaluations

The maximum F1 score for each model is recorded (Fig. 3). Since h5py, which is the preferred format for storing Keras model, has a limit to its header size, ResNeXt models cannot be saved.

Max F1 Score	ResNet	ResNeXt
Time Domain	0.779	0.766
Frequency Domain	0.813 (conv2d_114)	0.801

Fig. 3 Maximum F1 Score of different models

Frequency domain data has demonstrated better performance than that of time domain. To validate the result of the experiment, the 10 models with the same structure are trained using randomly selected training and testing data. The average F1 score is recorded (Fig. 4).

Average F1 score of 10 tries	ResNet	ResNeXt
Frequency Domain	0.7998	0.7742

Fig. 4 Average F1 Score of different models

The architecture that obtained the best performance is a 41-layers Residual Neural Network (40 Convolutional Layers and 1 FC layers). It is constructed using Residual Block (Fig. 5) and Down-scale Block (Fig. 6). The full structure and the hyperparameters are recorded below (Fig. 7).

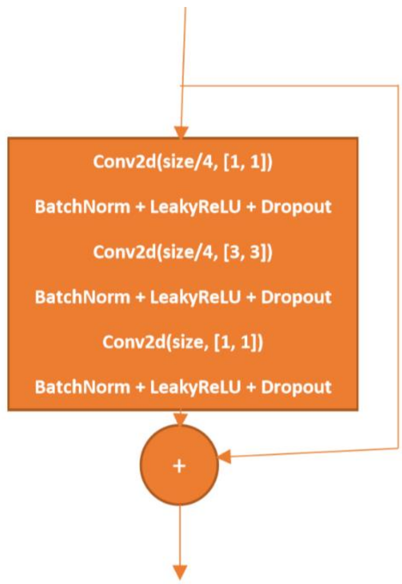

Fig. 5 Residual Block

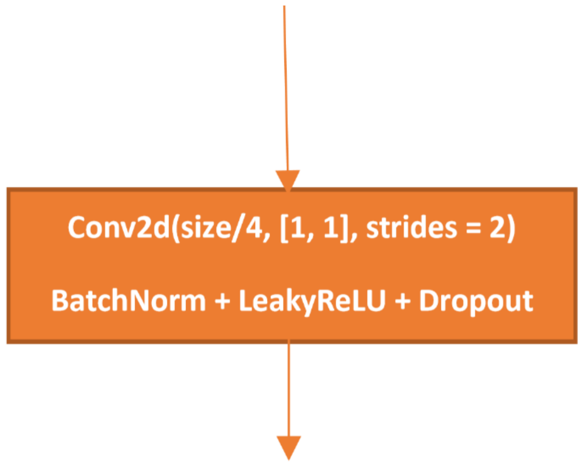

Fig. 6 Down-scale Block

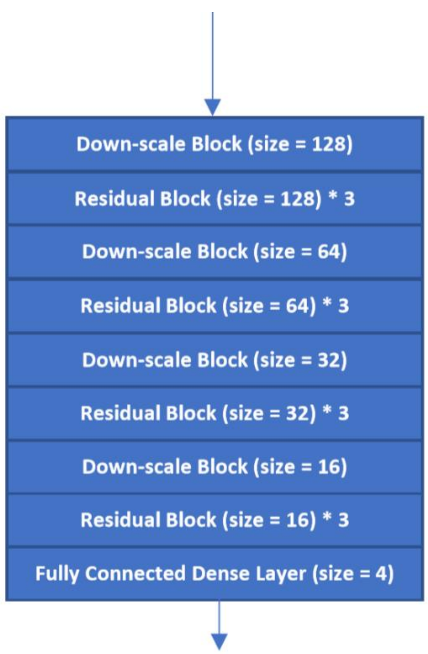

Fig. 7 Complete structure of the 41-layers ResNet Used. Here is the hyperparameters used for the network.

L2 Regularization = 0.001

Dropout = 0.13

Optimizer = Default ADAM of Keras

6. Discussion

During the evaluation, a pattern of misclassification is observed. A large portion of data belonging to the class of Other is misclassified as Normal (Fig. 8). The original plan was to investigate the reason behind these misclassifications using CycleGAN as part of the project [8]. However, the plan was ultimately stripped due to the lack of time. The code for the CycleGAN is also attached.

Label\ Preds	Normal	AF	Other	Noise
Normal	1524	14	129	5
AF	12	208	29	1
Other	199 (~30%)	65	550 (~60%)	12
Noise	31	15	15	33

Fig. 8 Classification grip

7. Conclusion

This project investigates the effect of time domain, frequency domain and the structure of models on the classification of Atrial Fibrillation data. The result shows that ResNet with frequency domain data has the best performance on classifying AF data within the scope of this project. The best F1 score achieved is 0.813.

8. Bibliography

[1] Patel NJ, Atti V, Mitrani RD, et al Global rising trends of atrial fibrillation: a major public health concernHeart 2018;104:1989-1990.

[2] Atrial fibrillation. National Heart, Lung, and Blood Institute. https://www.nhlbi.nih.gov/health/health-topics/topics/af/. Accessed March 12, 2019.

[3] Centers for Disease Control and Prevention. About multiple cause of death 1999–2011. CDC WONDER Online Database. 2014. Available from: http://wonder.cdc.gov/mcd-icd10.html.

[4] Faust, Oliver & Shenfield, Alex & Kareem, Murtadha & Tan, Ru San & Fujita, Hamido & Acharya, U Rajendra. (2018). Automated detection of atrial fibrillation using long short-term memory network with RR interval signals. Computers in Biology and Medicine. 102. 10.1016/j.compbiomed.2018.07.001.

[5] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals (2003). Circulation. 101(23):e215-e220.

[6] Khudorozhkov, Roman & Podvyaznikov, Dmitry. (2018). Benchmarks of ResNet Architecture for Atrial Fibrillation Classification.

[7] Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu: "Aggregated Residual Transformations for Deep Neural Networks", 2016; [http://arxiv.org/abs/1611.05431 arXiv:1611.05431].

[8] Jun-Yan Zhu*, Taesung Park*, Phillip Isola, and Alexei A. Efros. "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks", in IEEE International Conference on Computer Vision (ICCV), 2017.