ECG Heartbeat classification

Nguyen Cao Nguyen

February 2024

Introduction

Nowadays, ECG is widely used in cardiac health, especially in the diagnosis of cardiovascular diseases. However, manual analysis of ECG signals is difficult and time-consuming. To address this problem, many studies utilize machine learning to detect diseases. This report applied a deep-learning architecture to classify ECG Heartbeat to predict the diseases.

Datasets

The dataset used in this report is the PhysioNet MIT-BIH Arrhythmia downloaded from Kaggle. The dataset includes ECG recordings from 47 different subjects. Overall, the percentage of N and Q tend to occupy most of the dataset. As a consequence, I decided to reduce the number of the dataset, taking from the 47000th to the end. The new dataset is completed as seen below:
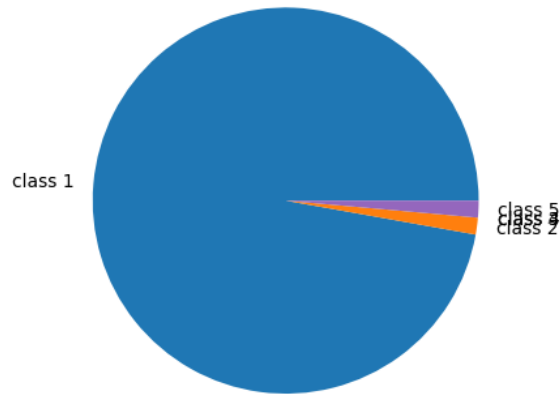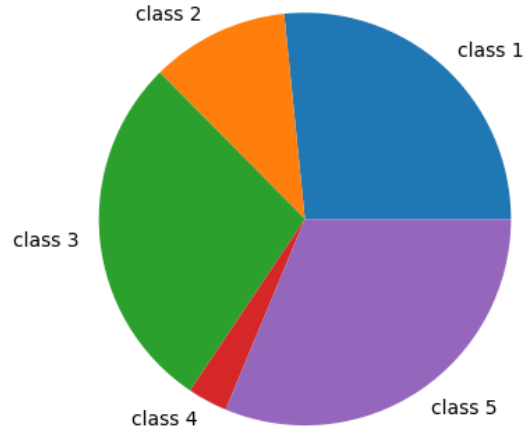

Figure 1: Data before equalizing

1

Figure 2: Data after equalizing

Method

This report utilized an existing architecture in another research. All convolution layers are 1-D convolution and apply 32 kernels of size 5. Max-pool layers in the architecture also have the same kernel size, with stride 2. After going through the first convolution layer, there are five residual blocks, including 2 convolution layers, a skip connection, and a pooling layer, connected by ReLu nonlinearities. The final part has 2 fully connected layers and a softmax layer.
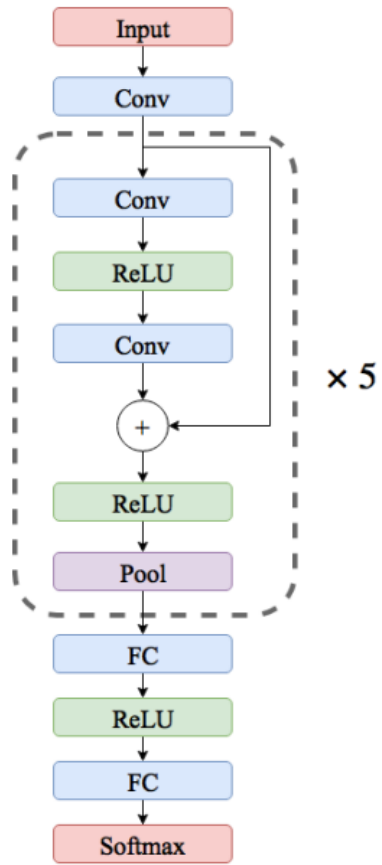
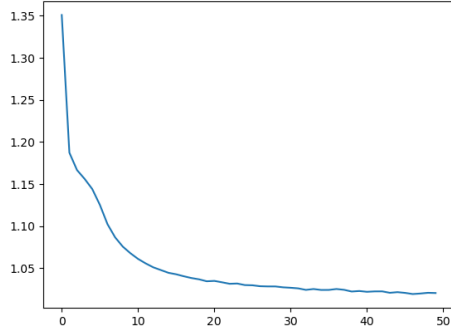Figure 3: Network architecture
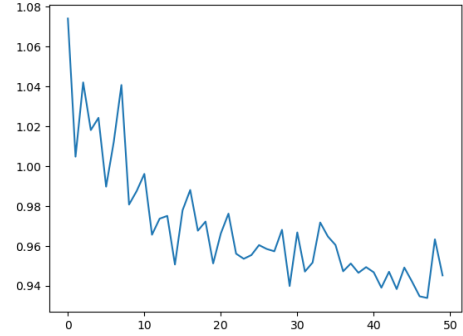
Result

Figure 4: train loss

Figure 5: valid loss

The proposed architecture achieved notable results. Both train loss and valid loss are seen to decrease significantly. The appropriate number of epochs should be around 25, where the loss of both tend to stop decreasing. Comparing with the test dataset, the model after going through 50 epochs achieved an accuracy of 88%. As can be seen from the confusion matrix of the test dataset, most of the objects in class 0,1,2,4 were correctly predicted. However, the model seemed to perform badly on the class 3. One of the potential reason is that the data in class 3 may not be sufficient for the model to learn its pattern.
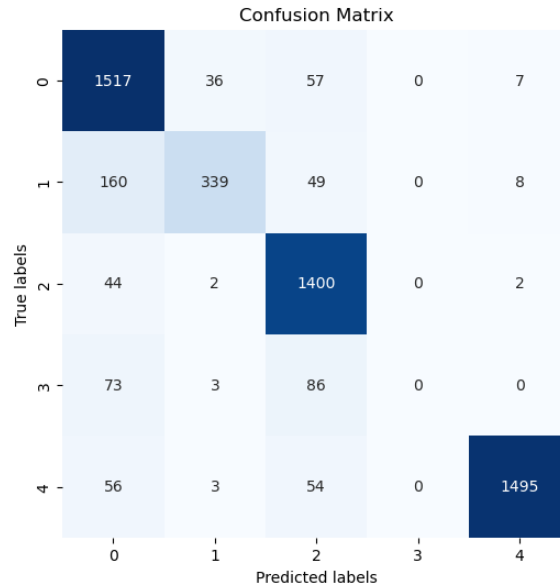

Figure 6: Confusion matrix

Conclusion

Overall, we could see that Residual Blocks handle the data of HeartBeat well. However, there are still works to be done with the data as well as the architecture to improve the overall performance, such as the class 3 prediction.