Freesound General-Purpose Audio Tagging Challenge

https://www.kaggle.com/c/freesound-audio-tagging

Sound Classification With Deep Learning

41 Categories including "Acoustic_guitar", "Applause", "Bark", "Bass_drum"
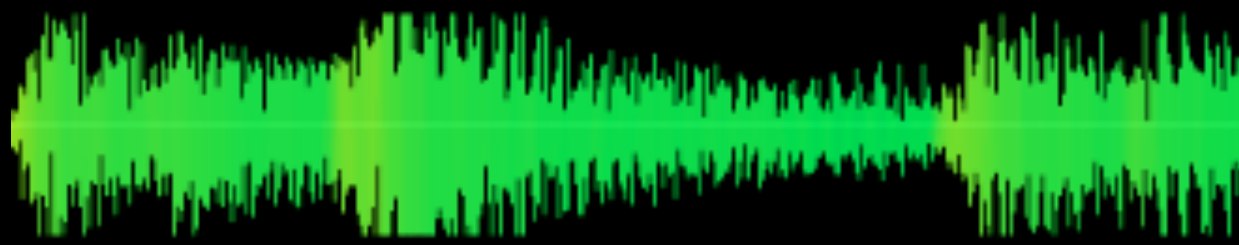
Data Set :
 The train data set includes ~9.5k samples unequally distributed among 41 categories. Each file is an audio file(.wav) and there are labels so that each file matches a label.
 The test data set includes ~9.4k samples and they have no label. We need to upload our answer to Kaggle to know our model's accuracy.

Problem :
 Labeling the test data using the model learned by train data.

Data Processing 1 : MFCC
 MFCC is the most generally used method for the sound processing. So we converted the data into arrays using MFCC (mfcc function in librosa module of python).

Data Processing 2 : STFT
 While MFCC is most commonly used in sound processing, passing a Mel filter filters out the frequency range that corresponds to the human voice, resulting in the loss of necessary data. So we also did it with STFT.

Deep Learning Model : CNN
 We trained the model using three convolutional layers and two fully connected layers.

Result (MFCC) : 75.4%
 The Accuracy when we converted the data using MFCC.

Result (STFT) : 63.7%
 The Accuracy when we converted the data using STFT.

Result (MFCC X STFT) : 82.3%
 The result of multiplying the predicted probability using MFCC and the predicted probability using STFT. (The accuracy is much better than in both cases)

Scoring was based on estimating three categories for each file, so it was 87.6%, slightly higher than 82.3%.

| 49 | ▲ 40 | **KYLius** | | 0.876 | 6 | now |

Your Best Entry ↑

Your submission scored 0.876, which is an improvement of your previous score of 0.823. Great job!

Tweet this!

Kim SeungHyuck

GitHub : https://github.com/smallerhand
Project : https://github.com/kyliusmethod/KYLius-method/tree/master/PROJECT2