MCA Assignment

Shagun Uppal (2016088)

February 2020

1 Spectogram

To compute spectrogram features for audio signals, we use a sliding window to divide in into frames, compute short-time discrete fourier transform across each and hence, get the spectrogram features as the output.

Below, are the precision and recalled as per the SVM trained on Spectrogram features. The parameters for the SVM are chosen as follows:

- C = 0.1
- kernel = *linear*

	Precision	Recall
Without noise	73.70	73.45
With noise	**76.79**	**76.28**

Table 1: Precision and Recall using Spectrogram features with and without background noise augmentation

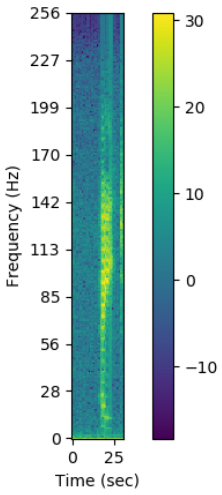

Figure 1: Sample plot for extracted spectrogram features from an audio file in the train set.

2 MFCC

For computing Mel-Frequency Cepstral Coefficients (MFCCs) for the given audio signals, we first amplify the higher frequencies (since they have smaller magnitudes), then split into discrete short-time frame signals, followed by windowing and fourier transform, thereby calculating power spectrum. The final mfcc features are mean-normalised. The two inbuilt functions used here are DFT and DCT.

Below, are the precision and recalled as per the SVM trained on MFCC features. The parameters for the SVM are chosen as follows:

- C = 0.5
- kernel = *polynomial*

	Precision	Recall
Without noise	86.79	85.36
With noise	**87.04**	**85.88**

Table 2: Precision and Recall using MFCC features with and without background noise augmentation

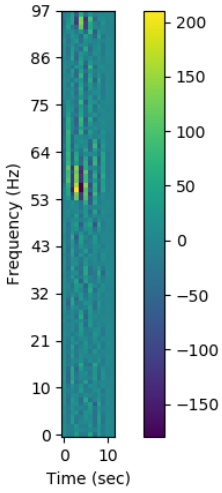

Figure 2: Sample plot for extracted mfcc features from an audio file in the train set.

3 Analysis

Effect of Noise: The effect of noise is a minimal enhancement in case of MFCC features, however ther is a significant increase of around 2-3% due to data augmentation of audios using background noise.

MFCC vs Spectrogram: In this case, MFCC performs better than Spectrogram as indicated by the precision. It also has other advantages as it has more compressed representations. In this work, the MFCC feature size is 97 x 12, whereas the Spectrogram features size being 256 x 31. Also, since MFCCs are a bit more decorrelated, they show better performance on linear models like SVMs.