Assignment 2

Mukul Kumar | 2017350

Spectrogram features

Window_size = 256, overlap = 64

Without noise:

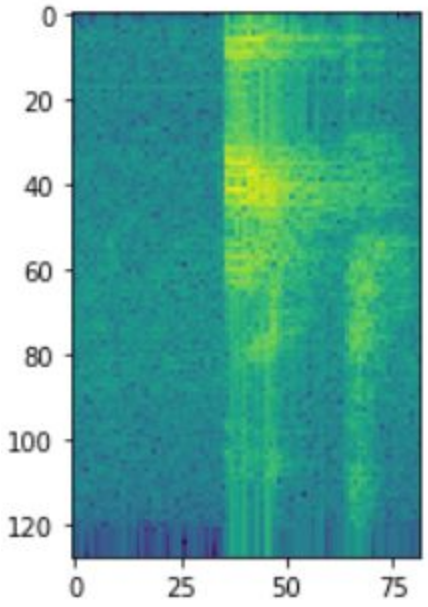
Accuracy : 24.72%

Precision : 0.25321648574600253

Recall: 0.22011949067862063

F1-score: 0.16673955002389637

Spectrogram sample plot



MFCC Features

Window_size = 400, overlap = 160, resolution = 512

a) Without noise

accuracy: 81.76%

precision: 0.8163

recall 0.8151

F1 Score: 0.8153

b) With white noise
   accuracy: 13.469999999999999%
   precision: 0.0375
   recall 0.1432
   F1 Score: 0.0538

c) With pink noise
   accuracy: 29.189999999999998%
   precision: 0.341
   recall 0.2975
   F1 Score: 0.2469

d) With running tap
   accuracy: 47.63%
   precision: 0.5772
   recall 0.4695
   F1 Score: 0.4479

e) With bike
   accuracy: 64.03%
   precision: 0.7115
   recall 0.6346
   F1 Score: 0.641

f) Doing the dishes
   accuracy: 58.46%
   precision: 0.599
   recall 0.5779
   F1 Score: 0.5685

g) dude_miowing
   accuracy: 79.11%
   precision: 0.7997
   recall 0.7876
   F1 Score: 0.7886

Results get worse due to adding of noise. The addition of noise would have helped if there was any overfitting, but there doesn't seem any and a validation accuracy of >81% was achieved, with a decent recall, precision, and F1-score.

There were different results observed by addition of different noises, this might be because of I've added only the first second of the noise to the original sound, therefore the duration for sound may differ, also it shows the effect of different noises on audio classification.

MFCC features gave very good results as compared to spectrogram results. Spectrogram feature calculation was also dramatically slow as compared to MFCC feature extraction.

Sample MFCC plot