Avădănei Antonia, grupa 233

Surgical mask detection-Project Documentation

My project documentation includes:

1. Description of the machine learning
    1. Chosen feature set
    2. Model
    3. Details
2. The classification accuracy for the provided validation set
    1. The precision and recall
3. Comparation between my models

1.1 Chosen feature set

Taking into consideration that the audio data provided cannot be understood directly by any models, I used **librosa** library to extract the features and convert them into an understandable format. LibROSA is a python package for music and audio analysis. (https://librosa.github.io/librosa/)

```
for audio_file in audio_files:
    file = path + audio_file
    time_series, sampling_rate = librosa.load(file, res_type='kaiser_fast')
    mel_frequency = librosa.feature.mfcc(time_series, n_mfcc=200, sr=sampling_rate)

    scaled_mfccs = np.mean(mel_frequency.T, axis=0)
    feats.append(scaled_mfccs)
```

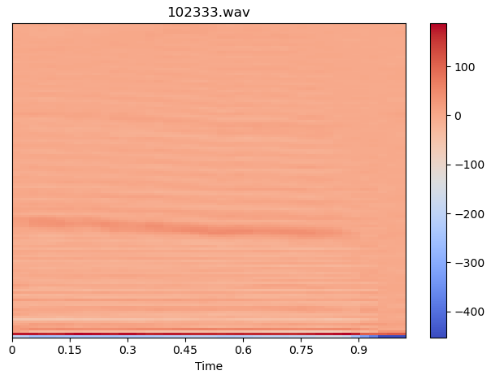
In order to load the audio files I used **librosa.load()** which has the following parameters:

- **path:** ->The location of the audio files that I want to load
- **res_type:str : kaiser_fast** default is **kaiser_best** but I used **kaiser_fast** in order to reduce the loading time
- +other default parameters

And returns:

- **y: np.ndarray [shape=(n,) or (2, n)]** ->audio time series
- **sr: 22050** ->sampling rate of *y*

Then I used **librosa.feature.mfcc()** which generate **mel-frequency cepstral coefficients** from the time series and returns a **np.ndarray [shape=(n_mfcc, t)]**



Visual representation of 102333.wav file

When n_mfcc was higher, so was the accuracy of the model. However, the learning time was also increasing, so I thought 200 is a honest value.

1.2 Model

I used a SVM with C=1 and kernel='linear' because I found some advantages:

1. 1. Training a SVM with a Linear Kernel is **Faster** than with any other Kernel.
2. 2. When training a SVM with a Linear Kernel, only the optimisation of the **C Regularisation** parameter is required. On the other hand, when training with other kernels, there is a need to optimise the γ parameter which means that performing a grid search will usually take more time.

1.3 Details

I noticed that when C was higher than one the model had lower accuracy on both validation files and test files, so I decided that C=1 is the best option. Also, I used linear SVM because the complexity of the 'rbf kernel' grows with the size of the training set.

2.1 Precision and recall

The F1 is interpreted as a weighted average of the precision and recall, where an F1 score reaches its best value at 1 and worst score at 0. The relative contribution of precision and recall to the F1 score are equal. The formula for the F1 score is:

F1 = 2 * (precision * recall) / (precision + recall)

F1_score for my SVM model is 0.7034284548129352

Accuracy_score is 0.692

3. Comparation between my models

At first I used a Knn model, using k=3, 10 and 30 and norm ‘l1’, but the accuracy on the validation test were between 0.50 and 0.60.