Acceleration and Braking Recognition

Felipe Bombardelli

May 5, 2019

1 Introduction

The problem is identify

Your task is to write a script (in Python) which reads a single WAV file similar to what you can see in Testing data and returns predictions of braking or accelerating trams according to classes used in the dataset. The output should be in CSV format as described below. Please attach also a results visualization of your choice. The script should be executable on Unix and should contain only open-source libraries. Please, provide also a brief report about your thoughts, your approach, what other possibilities you considered, why did you decided for that one, and an estimation how well your model is going to perform on our testing data. email

Data are labeled as follows :

1. Negative - Checked
2. Accelerating - 1_New
3. Accelerating - 2_CKD_Long
4. Accelerating - 3_CKD_Short
5. Accelerating - 4_Old
6. Braking - 1_New - Skoda T15
7. Braking - 2_CKD_Long - Tatra KT8D5R.N2P
8. Braking - 3_CKD_Short - Tatra T6A5 one or two carriages
9. Braking - 4_Old - Tatra T3 and all its modifications, one or two carriages

id	Label	Samples	Average Duration
0	negative/checked	828	9.83
1	accelerating/1_New	493	3.57
2	accelerating/2_CKD_Long	169	3.57
3	accelerating/3_CKD_Short	74	3.56
4	accelerating/4_Old	410	3.38
5	braking/1_New	381	4.58
6	braking/2_CKD_Long	143	4.51
7	braking/3_CKD_Short	62	4.35
8	braking/4_Old	436	3.95

2 State of Art

There is no work about train recognizing by sound on the IEEEXplorer, but there some works about environmental sound recognition.

3 Program Structure

The code of solution was divided on 3 modules for a better organization and reuse of code. The first module called dataset is responsible by handle and normalize the dataset. The second, way_classic is responsible by training, evaluation and performance classifiers like KNN, Neural Network and SVM. And by last, way_cnn is responsible by training, evaluation and performance classifier using Convolution Neural Network (CNN).

3.1 Dataset

The dataset has

3.2 Data Normalize

Existem algumas maneiras para representar o som,

Time-Frequency Representations short-time Fourier transform (STFT) with linear and Mel scales, constant-Q transform (CQT) continuous Wavelet transform (CWT)

"In the domain of environmental sound it has been noted that time-frequency representations are especially useful as learn- ing features [14][15][16][17][18] due to the non-stationary and dynamic nature of the sounds." [1]

The sounds have different duration and it needs to divide the sound in little window with same size. So the descriptors on these windows have the same size and it possible to use them with a classifier.

4 Results

5 Discussion

6 Conclusion

References

[1] Muhammad Huzaifah. Comparison of time-frequency representations for environmental sound classification using convolutional neural networks. *CoRR*, abs/1706.07156, 2017.