Report Date: 05/13/2022
To:

- ematson@purdue.edu
- ahsmith@purdue.edu
- lhiday@purdue.edu
- lee3450@purdue.edu
- wang4070@purdue.edu

From: TN

- Eunyoung Bang (yeong35@kangwon.ac.kr)
- Yeongmin Seo (dudals1003@cu.ac.kr)
- Jeongyoun Seo (201810773@sangmyung.kr)

Summary

- The research goal was refined with mia.
- The dataset collection was started.

What TN completed this week

- The project goal was discussed with Mia.
    - An article that deep learning architecture and audio feature extraction was summarized for study.
    - It is hard to constantly drive a UAV and the gap between velocity(5mph, 7mph, and 9mph) is too short. Our goal is to classify UAV velocity as slow and fast. What is fast or slow is subjectively decided. A boundary of fast or slow is 10mph.
    - The project uses one or two kinds of UAVs to collect the dataset. Too many kinds of UAVs are hard to predict the speed and hard to collect the dataset. The audio dataset length is 20min for each kind of UAV. 10min is slow velocity, another 10min is fast velocity.
    - A laptop or cell phone is enough to record the dataset. If a cell phone is used to collect the dataset, a python code to split audio 10 seconds will be needed.
- Dataset collection
    - A python code to collect dataset automatically cut to 10 seconds is written by Eunyoung.
    - A speed gun to measure UAV velocity was bought from Amazon.
- A paper draft was confirmed by Minji.
    - An abstract has unnecessary sentences. The Abstract must be clear.
    - The paper is written in casual words. casual words will be changed to academic words.
    - “A UAV's price decreases and a UAV is available.” is subjective. Evidence of this sentence is needed.
- Read paper
    - A gear stick position and car’s velocity are predicted using the sound of the car's engine. MFCC, zero-crossing rate, and spectral centroid are used to feature extraction. zero-crossing rate and spectral centroid will be used in our project to feature extract [1].
- Read article
    - Zero crossing rate presents a degree of smooth sound. The spectral centroid presents a center of a sound. If audio has a high note, spectral centroid will have a high value.

Spectral centroid is calculated as the mean of frequency [2]. Two feature extraction methods will be supported to model training.

- A filter bank is an array of band-pass filters that separate input signals into several components. In addition, with the filter bank applied, the feature that discarded unnecessary information related to voice recognition and left only important characteristics is MFCC.
  Mel-scaled filter banks are used when machine learning algorithms are not sensitive to highly correlated inputs, and MFCC is used when machine learning algorithms are sensitive to correlated inputs.
  Since MFCC will be used for the sound processing of UAVs, this article describing the detailed process along with the code has been adopted [3].
- The author used 'libROSA' to show how to extract audio features for analysis and the 'libROSA' was used by Mia (Yaqin Wang) in her paper [4]. Author showed how to extract MFCCs, Chromagram, Spectral Centroid and how to build the feature vector also. The Chromagram is helpful when you analyze a music file to classify the note when the notes which are exactly one octave apart are perceived as particularly similar [5].

Things to do by next week

- Novelty and problems in a readme.md will be filled.
- Mid presentation ppt and script will be finished.
- The dataset will progress noise reduction and labeling to fast or slow.

Problems or challenges:

- How to drive a UAV with constant speed?
- Our research goal is to classify UAV velocity as fast or slow. Controlling the UAV velocity in fast range or slow range is a challenge.

References

[1] H. V. Koops and F. Franchetti, "An ensemble technique for estimating vehicle speed and gear position from acoustic data," in *Proc. 2015 IEEE Int. Conf. on Digital Signal Processing (DSP)*, 2015, pp. 422-426, doi: 10.1109/ICDSP.2015.7251906.

[2] S. Doshi, "Music Feature Extraction in Python," medium.com. https://towardsdatascience.com/extract-features-of-music-75a3f9bc265d (accessed May. 10, 2022).

[3] Speech Processing for Machine Learning: Filter banks, Mel-Frequency Cepstral Coefficients (MFCCs) and What's In-Between, haythamfayek, last modified 05/09, 2022, accessed Apr 21, 2016, https://haythamfayek.com/2016/04/21/speech-processing-for-machine-learning.html

[4] Y. Wang, F. E. Fagian, K. E. Ho and E. T. Matson, "A Feature Engineering Focused System for Acoustic UAV Detection," 2021 Fifth IEEE International Conference on Robotic Computing (IRC), 2021, pp. 125-130, doi: 10.1109/IRC52146.2021.00031.

[5] B. Athina. "Audio signal feature extraction for analysis." Medium.com. https://athina-b.medium.com/audio-signal-feature-extraction-for-analysis-507861717dc1 (accessed May. 9, 2022).