Report Date: 05/27/2022
To: ematson@purdue.edu, ahsmith@purdue.edu, lhiday@purdue.edu, and lee3450@purdue.edu
From: What is today's lunch?

- Ilmun Ku(mun90505@hufs.ac.kr)
- Seungyeon Roh(shtmddus99@konkuk.ac.kr)
- Gyeongyeong Kim(kky57389@sunmoon.ac.kr)
- Charles Taylor(taylo869@purdue.edu)

Summary

The mid-presentation manuscript was written and get feedback from Yaqin and Taylor. A rehearsal of mid-presentation was held and Yaqin was in attendance. Data Augmentation methods have been set. Practice for mid-presentation was conducted.

What 'What is today's lunch?' completed this week:

- Writing a mid-presentation manuscript.

Hello everyone our names are Ilmun Koo, Seungyeon Roh, Gyeongyeong Kim, and Charles Taylor. We are team "What is Today's Lunch" and we will now give our presentation. Today, Seungyeon Roh, I, will be presenting the first half while the second half of the presentation will be read by Gyeongyeong Kim.

Over the years as Unmanned Aerial Vehicle (UAV) technology has continued to advance it has become much easier and cheaper to be able to get ahold of a drone. Anyone today could easily order one or even buy one in their local store if they wanted to. However, while UAV accessibility has steadily grown, malicious activities have also become increasingly common. Especially UAVs with payloads that can easily be employed to endanger innocent civilians or Government Dignitaries with their airborne contents. These unknown packages could be potentially transporting harmful materials, explosives, etc. through public airspace.

In just 2019, a Houthi drone has been observed targeting senior Yemeni military officers and even exploded over a military base, killing six and wounding twelve. In another instance, a UAV was seen carrying a bottle with unidentified contents and even landed on the Japanese Prime Minister's living quarters. In 2018, Venezuela's President Maduro was attacked by two small drones carrying explosives. Therefore, in terms of anti-terrorism, it is worth correctly detecting UAV payloads in advance so that a Counterterrorism team or police can effectively take action to evacuate the area and reduce expected harm to the victims.

The purpose of our project is to build a payload detection system for UAVs using deep learning algorithms. The main point of our project is to improve upon the limitations of the prior study. In the prior study, Only Machine Learning (ML) algorithms were applied for the purpose of detecting UAVs carrying payloads, on the basis of the sound that they emit. Feature extraction alongside a combination of other features was utilized to process raw data to then be fed to the ML models. In this study, however, feature methods will be used to train and evaluate Deep

Learning (DL) algorithms as opposed to ML Algorithms. In the past, the study had been limited to using only the ML algorithms for detecting UAVs

This limitation was in place due to the available processing power when using a MacBook Air with only 8GB of RAM and the lack of data for utilizing deep learning algorithms. In this study, we will use MacBook Pro with 16GB of RAM which should provide enough computing power for the application of deep learning algorithms in order to classify the sounds that they emit. Furthermore, additional data is required for more consistent results. Since we have only 600 samples from the prior project, we are currently planning to go on a data collection trip in order to gather more samples. We even plan on employing data augmentation methods to create even more raw data for the algorithm to analyze. Data augmentation methods are pivotal for the smooth and continuous improvement of audio classification performance while utilizing these DL algorithms.

In the next section of the presentation by Gyeongyeong, he will cover features and deep learning algorithms, and data augmentation methods that we will use in our study. That concludes my presentation(seungyeon) and the presentation about the methodology by Gyeongyeong Kim will follow.

Hello everyone, I am Gyeongyeong Kim and today I will be presenting the methodology section of our project. First thing first, I would like to talk about the data collection method. In the prior study, data collection was conducted by utilizing several Raspberry Pis as microphones. However, Using Raspberry Pis was not as practical of a method as we'd like, due to the time and effort expended setting up every piece of equipment. However, employing laptops or cell phones as recording devices instead is a more cost-effective and easily accessible method that can be applied by anyone.

For this experiment, the setting for the data collection equipment is the same in each recording. There are also outside factors to consider, as the data collection environment is unfortunately restrained to only weather without rain or wind present**.** As you can see, the UAV, which in this case is a consumer-available drone, would be flying within a certain range of the laptop's microphone in order to record the sounds that the drone emits. Three separate classifications of sounds will be recorded; a drone without a payload, a drone with one payload, and a drone with two payloads. The payload weighs 560 grams each. Since the battery of the drone is limited and the battery dies even faster if a drone is carrying a payload, not much data can be collected in a single trip for data collection. For this reason, we have planned more than three data collection trips in order to gather a sizable amount of data for processing.

Also, Since the dataset is not enough for employing deep learning algorithms in the prior study, data augmentation will also be employed in order to secure a sufficient amount of data. At least 1000 samples are required for deep learning algorithms and normally 360 audio samples can be obtained by a single collection trip. Data augmentation is a method to create artificial data with some sort of distortion or transformation. It is commonly used to supplement a lack of data.

Furthermore, the augmentation of data could prevent overfitting before the deep learning model is trained. Overfitting refers to a status where a statistical model fits exactly against its training data. When this happens, the algorithm, unfortunately, cannot perform accurately against unseen data, which ends up defeating its purpose. The data obtained from data augmentation would then be split into segments about 5 to 10 seconds long each via a python script. A label will then be attached to each segment so that the segments can be utilized in the classifier model. And there will be three different kinds of labels; a drone without a payload, a drone with one payload, and a drone with two payloads. Afterward, feature extraction will be executed on the data. The reason why we use feature extraction is due to the fact that there is important information hidden in the audio data.

The performance of deep learning models with feature extraction has been shown to be better than the performance of those without. Through the help of feature extraction, the deep learning model will be fed more descriptive information than compared just raw audio files, in turn making the model perform better in data recognition. For this reason, Feature extraction methods such as Mel-Frequency Cepstral Coefficient(MFCC) or Mel Spectrogram will be conducted upon the data, even if it takes more time to do so. Once Feature extraction is finally complete, the deep learning model is ready to be trained. Convolutional Neural Networks (CNN) and Recurrent Neural Networks(RNN), which are typically applied in sound classification, will be employed in this experiment to prove their efficiency in UAV payload detection. A neural network is a series of algorithms that aims to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates. In this sense, neural networks refer to systems of neurons, either organic or artificial in nature. Essentially, RNN is made for the processing of time-series data. However, CNN was originally made to handle image data. Although it is widely studied and proved by researchers that CNN can be applied if audio data is transformed into image data. Surprisingly, the performance of CNN in sound classification turns out to be excellent. For this reason, RNN which can properly process time-series data, and CNN which has remarkable performance on sound classification will be utilized. Our future works are the same as follows. this experiment's data collection will be carried out in a manner that is relatively easy, and cost-effective. The collected data will then be taken through pre-processing procedures namely data augmentation and Feature extraction. Furthermore, deep learning algorithms will be analyzed for their efficiency in UAV payload detection through CNN and RNN. We especially appreciate the help of Dr.Matson and Yaqin.

- Selecting Data Augmentation methods
  - Pitch Scaling(Raw data augmentation)

    The process of changing the speed/duration of sound without affecting the pitch/frequency of sound. It is widely utilized for audio data augmentation and proved to be effective for improving model performance [15].

  - Time Stretching(Raw data augmentation)

    The process of changing the speed/duration of sound without affecting the pitch/frequency of sound. It is widely utilized for audio data augmentation and proved to be effective for improving model performance.

  - Time Masking(Spectrogram data augmentation)

    The vertical mask is made first and put then randomly into the spectrogram. A frequency channels $[f0, f0 + f)$ are masked. f is chosen from a uniform distribution from 0 to the frequency mask parameter F, and $f0$ is chosen from $(0, \nu - f)$ where ν is the number of frequency channels.

- Frequency Masking(Spectrogram data augmentation)

  t consecutive time steps [t0, t0 + t) are masked. t is chosen from a uniform distribution from 0 to the time mask parameter T, and t0 is chosen from [0, τ − t).

Things to do by next week

- Employing CNN and RNN for data from previous works.
- Writing papers including abstract, introduction, and literature review and Getting feedback.
- Going on a data collection trip with Mia.

Problems or challenges:

- There are datasets that Yaqin has collected. However, it is not large enough to apply Deep Learning algorithms. Therefore, it is required to collect more data or apply data augmentation to the original dataset.
- Payloads currently secured are too heavy for UAVs to carry. Payload weighs 560grams each. The battery dies fast when UAV carries payloads. It is required to get lighter payloads. Dr.Matson will take care of this issue by finding a lighter payload such as a hand grenade.

References

[1] M. al Kibsi, "Houthi drone targets senior yemeni officers, kills five soldiers," Available at https://www.aljazeera.com/news/2019/1/10/houthi-drone-targets-senior-yemeni-officers-kills-five=soldiers (2022/05/18).
[2] W. Ripley, "Drone with radioactive material found on japanese prime minister's roof," Available at https://www.cnn.com/2015/04/22/asia/japan-prime-minister-rooftop-drone/index.html (2022/05/18).
[3] E. Kelly, "Venezuela drone attack: Here's what happened with Nicolas Maduro," Available at https://www.usatoday.com/story/news/politics/2018/08/06/venezuela-drone-attack-nicolas-maduro-assassination-attempt-what-happened/913096002 (2022/05/18).
[4] Y. Wang, F. E. Fagiani, K. E. Ho, and E. T. Matson, "A feature engineering focused system for acoustic uav payload detection," the 14th International Conference on Agents and Artificial Intelligence (ICAART 2022), pp. 470–475, 2022.
[5] S. Wei, S. Zou, F. Liao et al., "A comparison on data augmentation methods based on deep learning for audio classification," 2019 2nd International Conference on Computer Information Science and Artificial Intelligence(CISAI 2019), pp. 1–8, 2020.
[6] S. Al-Emadi, A. Al-Ali, A. Mohammad, and A. Al-Ali, "Audio based drone detection and identification using deep learning," pp. 459–464, 2019.
[7] S. Jamil, M. Rahman, A. Ullah, S. Badnava, M. Forsat, S. S. Mirjavadi et al., "Malicious uav detection using integrated audio and visual features for public safety applications," Sensors, vol. 20, no. 14, p. 3923, 2020.
[8] Y. Wang, F. E. Fagian, K. E. Ho, and E. T. Matson, "A feature engineering focused system for acoustic uav detection," pp. 125–130, 2021.
[9] L. Shi, I. Ahmad, Y. He, and K. Chang, "Hidden markov model based drone sound recognition using mfcc technique in practical noisy environments," Journal of Communications and Networks, vol. 20, no. 5, pp. 509–518, 2018.

[10] V. Kartashov, V. Oleynikov, I. Koryttsev, S. Sheiko, O. Zubkov, S. Babkin, and I. Selieznov, "Use of acoustic signature for detection, recognition and direction finding of small unmanned aerial vehicles," pp. 1–4, 2020.
[11] U. Seidaliyeva, M. Alduraibi, L. Ilipbayeva, and A. Almagambetov, "Detection of loaded and unloaded uav using deep neural network," pp. 490–494, 2020.
[12] U. Seidaliyeva, M. Alduraibi, L. Ilipbayeva, and N. Smailov, "Deep residual neural network-based classification of loaded and unloaded uav images," pp. 465–469, 2020.
[13] F. Fioranelli, M. Ritchie, H. Griffiths, and H. Borrion, "Classification of loaded/unloaded micro-drones using multistatic radar," Electronics Letters, vol. 51, no. 22, pp. 1813–1815, 2015.
[14] D. Utebayeva, M. Alduraibi, L. Ilipbayeva, and Y. Temirgaliyev, "Stacked bilstm-cnn for multiple label uav sound classification," pp. 470–474, 2020.

[15] Kim, G., Han, D. K., & Ko, H. (2021), "Stacked bilstm-cnn for multiple label uav sound classification," pp. 470–474, 2020.