POSTER: AcousticPrint: Acoustic Signature based Open Set Drone Identification

Harini Kolamunna*, Junye Li§, Thilini Dahanayaka*, Suranga Seneviratne*, Kanchana Thilakaratne*, Albert Y. Zomaya*, and Aruna Seneviratne§
University of Sydney* University of New South Wales§

ABSTRACT

Malicious or improper use of drones can pose significant privacy and security threats in both civilian and military settings. There are many situations where it requires to detect the presence of a drone and identify the exact model to be used in applications such as law enforcement depending on the size and capabilities of different models. Nonetheless, this remains a challenging task, especially in low visibility, limited access, or hostile environments. In this paper, we propose to use acoustic signatures to identify the make and the model of drones. We achieved 94% accuracy in a closed set scenario and 80% accuracy in a more challenging open set scenario.

CCS CONCEPTS

• **Computing methodologies** → **Feature selection**; **Neural networks**.

KEYWORDS

Drones, Acoustic fingerprinting, LSTM, Drone Audio Dataset

ACM Reference Format:
Harini Kolamunna*, Junye Li§, Thilini Dahanayaka*, Suranga Seneviratne*, Kanchana Thilakaratne*, Albert Y. Zomaya*, and Aruna Seneviratne§. 2020. POSTER: AcousticPrint: Acoustic Signature based Open Set Drone Identification. In *13th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec '20), July 8–10, 2020, Linz (Virtual Event), Austria.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3395351.3401700

1 INTRODUCTION

Drones are becoming widely available and are benignly used in multiple applications such as cinematography, surveying, and legal goods delivery. Nonetheless, they are also being used for reconnaissance, invading personal or secure spaces, harming targeted individuals, smuggling drugs and contraband, or creating public disturbances. For example, recently, departures at the Heathrow airport were temporarily suspended after reports of a drone sighting [11]. In another example, a weaponized drone hovered over a public gathering in Venezuela and dropped explosives targeting high profile personnel and the general public [12]. As such, identifying drones in different environments to assist the decisions on whether or not to contain the drone, is a necessity.

WiSec '20, July 8–10, 2020, Linz (Virtual Event), Austria

ACM ISBN 978-1-4503-8006-5/20/07.
https://doi.org/10.1145/3395351.3401700

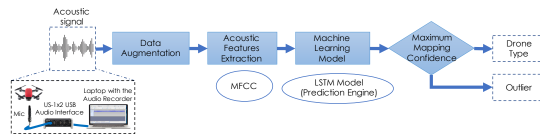


Figure 1: System Architecture.

Multiple works demonstrated the possibility of detecting drones using various forms of data such as video, RF, thermal imaging, radar, and WiFi [10]. Several recent works shows that drones can be also detected using acoustic signals [3, 5, 9]. Also, the early work by Al-emadi et al. [2] showed the feasibility of acoustics based drone classification by using two dornes. In contrast to these, we focus on identifying the drone type among a larger number of drone classes in realistic open world conditions. Moreover, we demonstrate the usage of both experimentally collected and online sourced data.

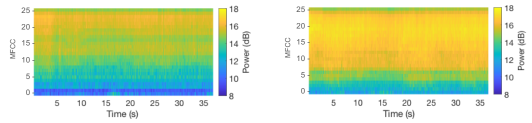
In this paper, we show that the combined acoustic signal primarily generated by the propellers, motor, and the mechanical vibrations of the body has a sufficiently unique signature and can be used to identify the drone type. Also, we extend the model to incorporate the capability of deciding whether the incoming sound signal is from an authorized device by eliminating non-drone sounds.

2 METHODOLOGY

We show the schematic overview of our system architecture in Figure 1 and describe various steps in the processing pipeline below.

Data collection: We captured the acoustic signals emanating from the drone sampled at 44.1kHz, using a high-quality directional microphone (RODE NTG4 shotgun), when the drones are flying around at 20m above ground and within a 50m radius. This experiment was conducted in a park in 3 sessions over 2 days when there is not much other activities happening. We collected 3-4 minutes of data at each session for 5 classes of drones (*Parrot Bebop 2*, *DJI Mavic Pro*, *DJI Phantom4 Advanced*, *DJI Spark*, *DJI Matrice 100*). Also, we used YouTube videos of flying drones to extract the clear audio of drones flying around and used them to enhance the data set. There, we extracted audios for 4 drones (*Autel EVO*, *DJI Inspire2*, *JME*, and *DJI MavicAir*) using 3 different videos for each class. These audios were split into three parts; 60% for *training* and 20% each for *validation* and *test* sets. For each class, we allocated a single audio to be used to get only the *testing* samples.

Data augmentation: As the operational conditions of the drones can vary, we augmented the *training* and *validation* samples with amplitude scaling and frequency warping [6]. Each sample is scaled 5 times along in the time axis and in amplitude by two separate values selected from a uniform distribution, U(0.8, 1.2), that gives a six fold increment.



(a) MFCC of DJI Spark (b) MFCC of DJI Phantom 4 Adv

Figure 2: MFCC analysis

MFCC feature extraction: We segmented the signal in to 25ms frames with 15ms overlap. For each frame, we calculated 26 Mel Frequency Cepstral Coefficients (MFCCs) that are commonly used in audio analysis. Next, we aggregated 20 frames each to form windows that will be subsequently used as time steps in our classifier. In Figure 2 we show MFCCs of two different drone classes to highlight the differences in spectral characteristics.

LSTM model: We used a Long Short-Term Memory (LSTM) network as our classifier since it has been shown promising results in audio signal classification tasks [8]. The LSTM architecture we used consists of 20 time steps, 2 stacked LSTM layers. We used a hidden state size of 32. At test time, we made predictions for 20 consecutive windows and calculated the average probability vector. Out of that we selected the drone class having the highest probability.

Open set data: To extend the ability of the classifier to identify *Unknown* classes we added a background class, i.e. *Known-Unknowns (KU)*, consists of data for; 1) other mechanical sounds such as *vehicles*; 2) similar acoustic signals such as *humming of bees*; 3) common non-mechanical sounds such as *human voice*, and 4) *calm environment*. We experimentally collected data for the *calm environment* and scraped audio signals from YouTube videos for the other data types. To test our model's performance under *Unknown* classes, we used data scraped from [13]. This consists of 13 vehicle sounds and 2 drone sounds that were not used in training. All the data we used in this work is publicly available at our github repository [1].

Training and testing: We tested two models 1) **M1:** closed-set model that contain only the 9 *known* drone, 2) **M2:** open-set model in which the training and validation is done with *known* and *KU* classes and *testing* with *known*, *KU*, and *unknown* classes.

3 RESULTS

The confusion matrix for the closed-set (M1) prediction is shown in Figure 3 where the overall accuracy is 94%. Also, the accuracy of the online scraped dataset is comparable to the experimentally collected dataset, indicating that drone-specific acoustic signatures are still preserved even in the processed data available in online sources. This is an important observation because it is not realistic to collect acoustic samples from all commercially available drones through experimentation. Scraping online data allows to expand the classifier to cover many drones.

With the *KU* class addition in the open-set model (M2), the *Drone* and *KU* prediction accuracy remained as high as in M1 (92% and 90% respectively). The *Unknown* motor vehicle and drone sounds gave 70% accuracy, resulting an overall accuracy of 80%. However, the *Unknown* drone sounds were predicted to be as one of the known *Drone* classes. On the one hand, this is a desirable result since it indicates that the classifier can make the drone vs. non-drone decision accurately. One the other hand, it also indicates that there is always the possibility of some unknown class being identified as one of the known drones and further model improvements are necessary to have much tighter decision boundaries.

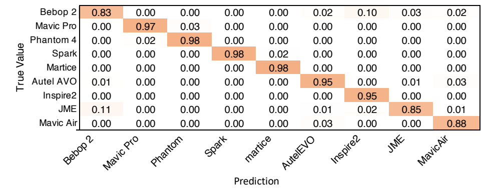


Figure 3: Normalized confusion matrix (closed-set).

4 CONCLUSION AND FUTURE WORK

Our results showed that there are unique 'AcousticPrints' associated with drones and a prediction model can be developed to achieve high closed-set accuracy. We also demonstrated that the data can be sourced from online sources without affecting the performance, which important as it is not practical to collect data from all drone types in all possible flying scenarios. Finally, with careful selection of signals to train an additional background class, we showed that it is possible to enable more realistic open-set predictions.

Our work can be extended in multiple ways. For example, MFCC filterbank scaling, which is originally designed for voice recognition, can be modified to better capture the signature variants in higher frequencies as drones have dominant energy components in higher frequencies. Also, it is possible to try more advanced open-set classification techniques to improve the robustness of the model [4, 7]. Moreover, although we might have addressed Doppler effect through indirect data augmentation and data collection as we collected traces of drones flying around freely, further studies are required to understand the full effect of Doppler effect. Finally, the effectiveness of a model purely trained from online sourced data to identify drones in real-time can be evaluated.

ACKNOWLEDGMENTS

This project was financially supported by NSW Defence Innovation Network and NSW State Government (DINPP-2018-04-16).

REFERENCES

[1] AcousticPrint. [n.d.]. https://github.com/AcousticPrint/AcousticPrint.
[2] S. Al-emadi, A. Al-ali, A. Mohammad, and A. Al-ali. 2019. Audio Based Drone Detection and Identification using Deep Learning. In *IWCMC'19*. IEEE, 459–464.
[3] M. Z. Anwar, Z. Kaleem, and A. Jamalipour. 2019. Machine Learning Inspired Sound-Based Amateur Drone Detection for Public Safety Applications. *IEEE Transactions on Vehicular Technology* 68, 3 (2019), 2526–2534.
[4] A. Bendale and T. E. Boult. 2016. Towards Open Set Deep Networks. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 1563–1572.
[5] A. Bernardini, F. Mangiatordi, E. Pallotti, and L. Capodiferro. 2017. Drone detection by acoustic signature identification. *Electronic Imaging* 17, 10 (2017).
[6] J. Chauhan, J. Rajasegaran, S. Seneviratne, A. Misra, A. Seneviratne, and Y. Lee. 2018. Performance Characterization of Deep Learning Models for Breathing-Based Authentication on Resource-Constrained Devices. *IMWUT* 2, 4 (2018).
[7] Akshay Raj Dhamija, Manuel Günther, and Terrance Boult. 2018. Reducing Network Agnostophobia. In *NIPS*. 9157–9168.
[8] A. Graves, A. Mohamed, and G. Hinton. 2013. Speech recognition with deep recurrent neural networks. In *ICASSP*. 6645–6649.
[9] S. Jeon, J. Shin, Y. Lee, W. Kim, Y. Kwon, and H. Yang. 2017. Empirical study of drone sound detection in real-life environment with deep neural networks. In *EUSIPCO*. 1858–1862.
[10] NCC. [n.d.]. Drones Detect, identify, intercept, hijack. http://tinyurl.com/j8u44sx.
[11] BBC News. [n.d.]. Heathrow airport: Drone sighting halts departures. https://www.bbc.com/news/uk-46803713.
[12] BBC News. [n.d.]. Venezuela President Maduro survives 'drone assassination attempt' . https://www.bbc.com/news/world-latin-america-45073385.
[13] Soundsnap. [n.d.]. . https://www.soundsnap.com/.