REAL-WORLD ON-BOARD UAV AUDIO DATA SET FOR PROPELLER ANOMALIES

Sai Srinadhu Katta [*], *Kide Vuojärvi* [*], *Sivaprasad Nandyala,*
Ulla-Maria Kovalainen, Lauren Baddeley

Secure Systems Research Centre (SSRC) - Technology Innovation Institute (TII), Abu Dhabi, UAE
{sai, kidex, siva, ulla-mariax, lauren}@ssrc.tii.ae

ABSTRACT

Detecting propeller damage in Unmanned Aerial Vehicles (UAV) is a crucial step in ensuring their operational resilience and safety. In this work, we present a novel real-world audio data set of propeller anomalies, and use several deep learning models to classify the damage. This data set consists of more than 5 hours of audio recordings, covering all configurations of intact and broken propellers in a UAV quadcopter. A microphone array was mounted onto a UAV, and numerous autonomous indoor missions were flown. Our on-board setup has provided clean audio recordings containing little background noise. We have developed classification models for this data set, using different deep learning architectures: Deep Neural Networks (DNNs), Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM), and Transformer Encoder (TrEnc). We conclude that the TrEnc outperforms other architectures, having 11k parameters, .57M Flops, 98.30% accuracy, .98 precision, and .98 recall. Finally, we make our data set publicly available here[⊙].

***Index Terms*—** UAV, Propeller Anomaly, Audio Data Set, Deep Learning

1. INTRODUCTION

During recent years, technological advancements have increased the regular usage of UAVs (also known as drones). They are being used in many different scenarios and contexts, such as military missions, agricultural projects, search and rescue teams, and as part of delivery systems [1, 2]. As such, the safety of a UAV has become even more important in order to protect not only itself, but also its surrounding environment from any damage.

UAVs have to navigate safely, often in a dynamic and partly unknown environment. This brings many challenges to overcome, some of which can lead to damages or degradation of different parts of the UAV [3, 4, 5, 6, 7]. UAV's mainly consist of a frame, flight and motor controllers, motors, transceivers, propellers, and batteries. One of the most important subsystems in a UAV is the propulsion, and any failure, cut, damage, wear from environmental factors or malfunction of a UAV's propellers components can cause unsteady flight or even flight failure [8, 9, 10, 11]. In the interest of the mission as well as the safety of the UAV and its environment, the UAV damage should be detected to avoid any accidents or crashes, which can be done through the use of deep learning algorithms [8, 12, 13, 14, 15].

In general, computer vision could be used to solve the problem of identifying propeller anomalies by placing cameras on-board a UAV but this is prone to false positives, sensitive to different light conditions and requires significant processing [16]. To overcome the above limitations, one novel approach is to use acoustic sensors on-board the UAV to detect a difference between the sound of the propellers before and after an anomaly. The use of acoustic sensors has multiple advantages - they are inexpensive, less prone to hacking compared with other UAV sensors like GPS, camera, etc. [17], have good coverage, and only require a small amount of configuration compared to other sensors. We use on-board microphones to collect a large real-world clean audio data set of a UAV in different autonomous flying scenarios with a varied number of broken propellers, which is then analysed by several different deep learning architectures to create a system that detects propeller anomalies. The main contributions of this paper are:

1. We obtained audio data with little background noise, for the full mission from take-off to landing, by flying real-world scenarios with a UAV that we equipped with an on-board microphone.
2. To the best of our knowledge, this is the largest real-world audio data set for propeller anomaly detection in UAVs. Our data set contains over $5h$ of autonomous flight audio, and covers all possible configurations of broken and intact propellers in a 4 propeller UAV.
3. We have adapted DNN, CNN, LSTM, and Transformer Encoder architectures for the classification task on our data set, and verified their efficacy.

This paper is organised as follows. Section 2 contains a literature study related to the propeller anomalies, including limitations of the previous works. Section 3 gives the detailed steps followed in collecting the audio data set, including the set up and configurations used in UAV. Section 4 describes the

[*] Equal contribution.
[⊙] https://github.com/tiiuae/UAV-Propeller-Anomaly-Audio-Dataset

classifiers used in our work, and all the hyperparameters used. In Section 5, the performance of the learned models is verified, and the experimental results are discussed. Finally, Section 6 concludes the paper and describes future research plans.

2. LITERATURE REVIEW

In recent times, anomaly detection in UAVs has become a major trend and an important research topic [6, 10, 18, 19]. The majority of the works have used an external device (smartphone) to measure the noise emitted by a UAV, and built classification models to detect an unbalanced propeller blade in the UAV. We have used on-board microphone. The sizes of the previous audio data sets have also been small, less than half an hour [1, 14, 20]. Ours is more than 5hrs of audio data. The very few that have collected on-board audio haven't used the UAV in realistic flight scenarios (i.e. the UAV has been kept in one location, attached to a tripod) [9]. To overcome these drawbacks, we deploy a microphone array on-board a UAV, enabling the collection of clean propeller audio data of autonomous flights with minimal background noise.

Deep learning architectures are shown to give state-of-the-art (SOTA) results in various tasks in the fields of Computer Vision (CV), Natural Language Processing (NLP) and Audio Signal Processing (ASP). As DNN, CNN, LSTM, and Transformer are some of the commonly and widely used architectures [12, 13, 14, 15, 19, 21], we adapt these for use with our data set.

3. OUR FRAMEWORK FOR DATA COLLECTION

Most of the existing state-of-the-art has focused on collecting data off-board the UAV, see Section 2. Contrarily, our approach uses a UAV-mounted microphone array to collect the audio data of real-world flights with realistic flight paths, enabling better model training and evaluation. We use *real-world* to depict that the UAV was flying autonomously for the whole mission, which consisted of take-off, landing, and visiting multiple waypoints in indoor settings.

We recorded audio with a Holybro X500 UAV quadcopter using different configurations of damaged propellers which can affect the UAV's flying capabilities. A propeller is a fundamental element that determines the performance of the device, as well as the intensity of the UAV's noise. The propeller can have different amounts of blades, usually ranging from two to five, and the length and the angle of the blades can also vary. The analysed UAV is a quadcopter, having 4 propellers, each with 2 blades (see Figures 1 (left) and 2).

3.1. Audio Data Description

To record the flight audio, we mounted a ReSpeaker USB Microphone Array on top of the Holybro (see Figure 1 (right)). After some preliminary testing to reduce the noise from the turbulent air flow of the propellers, we decided to completely block the air flow to the microphones with small pieces of tape. This reduced the microphones into contact microphones, receiving the conducted audio signal via the UAV's frame. The array was mounted onto a 3D-printed mounting plate, which in turn was attached to the UAV frame with bolts.

For propellers, we used the S500 V2 Propellers that can be either *intact* or *broken* (see Figure 2). A propeller was considered intact if it had no structural issues (with minor scratching being acceptable), and broken if it had structural damage. As flying with broken propellers may be hazardous, we tested a few different types of cut propellers. In this experiment, we cut approximately $13mm$ from one of the propeller blades to make it broken. We chose this cut as a compromise between safety and audibility.

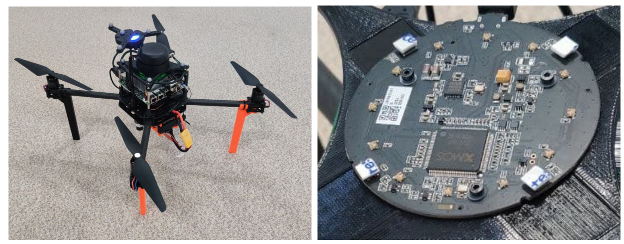

Fig. 1. Holybro X500 used in the experiments (left), ReSpeaker Mic Array mounted on the UAV (right)

Fig. 2. Broken (top, right blade) and intact (bottom) propellers

3.2. Scenarios with Different Propeller Configurations

We recorded audio for four different types of propeller configurations. The scenarios are given as follows:

1. Four intact propellers. This scenario acts as the baseline for all the other recordings, and is the most common scenario in the real world.
2. One broken propeller. Each combination of one broken and three intact propellers contributed 1/4 of the data for this configuration.
3. Two broken propellers. Each combination of two broken and two intact propellers contributed 1/6 of the data for this configuration.
4. Three or four broken propellers. Each configuration of three broken propellers contributed 1/6 of the data for this configuration, while the four broken propellers cases contributed 2/6. These scenarios were combined into one class due to their relative similarity.

3.3. Flight Routes for UAV Audio Data Set Collection

During the data collection, the UAV was flown inside a $15m$ by $7m$ room that is $4m$ high. Each flight consisted of a take-off, three or five waypoints, and a landing. While the UAV received the waypoints from us, the UAV flew autonomously from one waypoint to the next.

We chose to use four $1.5m$ by $0.5m$ unconnected areas from within the room to add some variance to the flight paths. These areas were located in the corners of a $7m$ by $2.5m$ rectangle. Each waypoint was selected randomly from one of these areas, and the altitude of each waypoint was selected randomly between $1m$ and $2m$. This method made every flight unique, while yielding a range of different angles and distances the UAV had to fly.

The flight paths were divided into two general shapes which we call *squares* and *figure eights*. In the figure eight flights, the waypoints were selected such that every other waypoint had the drone flying diagonally through the room. In the square flights, the drone never flew diagonally. Both of these shapes were flown in two directions: squares were flown both clockwise and counter-clockwise, and with figure eights the diagonal movement could happen either on the odd or even numbered waypoints.

Whenever we started flights of a new class, the UAV took off from a constant position within the room. In consecutive flights of the same class, the UAV took off from the previous landing position to diversify the take off location. Each flight consisted of an odd number of waypoints to ensure the take off and landing never happened in the same area.

For each scenario listed in 3.2, we flew 24 square flights in both directions, as well as 12 figure eight flights in both directions. This amounts to over $1h$ of audio for each scenario, yielding a total of over $5h$ of audio data.

4. UAV PROPELLER ANOMALY CLASSIFIERS

Deep learning architectures are shown to achieve state-of-the-art in various tasks in the fields of CV, NLP, and ASP. This section gives an overview of the commonly used deep learning architectures used in this work.

4.1. End-to-End Deep Learning Classifiers Architecture

The goal is to minimise the number of parameters and floating point operations per second (FLOPS), while maximising the performance in terms of accuracy, precision, and recall. The Mel-frequency cepstral coefficient (MFCC) features are extracted from the audio data and then fed to the model architecture variant. This helps to keep the end-to-end model smaller and more efficient in comparison to the spectrogram without degrading the performance. We kept the goal in mind while designing the model architecture variants. Figure 3 shows the end-to-end architecture of our work.

4.1.1. Deep Neural Network (DNN)

The DNN is made of fully-connected layers and non-linear activations. The input to the DNN is the flattened MFCC features, which feeds into a stack of hidden fully-connected layers. At the output is a linear layer followed by a softmax layer generating the output probabilities of the classes. We created different variants of this architecture by changing the hidden dimension [22].

4.1.2. Convolutional Neural Network (CNN)

CNNs exploit the local temporal and spectral correlation in the features via 2D convolution. The input to the CNN is the MFCC features, which feeds into a stack of convolution layers. At the output is a linear layer followed by a softmax layer generating the output probabilities of the classes. We created different variants of this architecture by changing the number of filters.

4.1.3. Long Short-Term Memory (LSTM)

LSTMs are known to model long term dependencies and are shown to work very well on various sequence modelling tasks. The input to the LSTM is the MFCC features, and the whole flattened output sequence is fed to a linear layer followed by softmax for output probabilities of the classes. We created different variants of this architecture by changing the hidden dimension.

4.1.4. Transformer Encoder (TrEnc)

Transformers are shown to be the fundamental block for SOTA on various sequence modelling tasks across various domains. In this work we use only the Transformer Encoder part. The input is the MFCC features, and the whole output sequence is fed to a linear layer followed by softmax for output probabilities of the classes. We created different variants of this architecture by changing the feedforward dimension inside the Transformer Encoder.

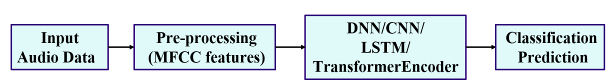

Fig. 3. Architecture used for the proposed work

5. EXPERIMENTS AND RESULTS

In this section, the effectiveness of the proposed approach for detecting propeller anomalies is evaluated. Figure 4 shows the spectrograms that are extracted from the audio data, and all pairs of classes are easily distinguishable.

5.1. Training Description

The data set is split into 80% for training, 10% for validation and 10% for testing. The audio files are broken into chunks of 1 second, and the number of classes is 4, as elaborated in 3.2. The loss function used is cross-entropy loss. For the MFCC feature selection, the window size is $40ms$, the stride is $30ms$ and 20 bins. Adam [23] is used with the learning rate 1e-3 for first 15000 steps followed by 1e-4 for the next 3000 steps. The

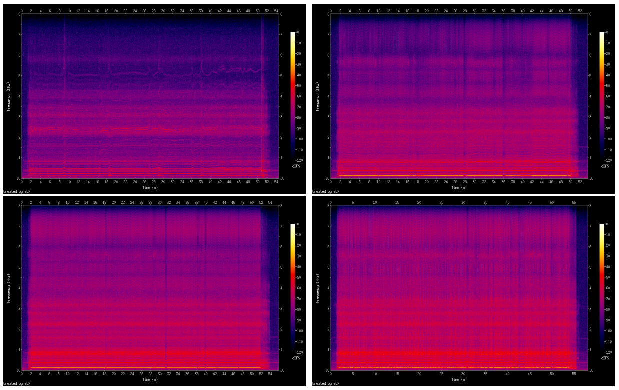

Fig. 4. Spectrograms for flights with 0 (top-left), 1 (top-right), 2 (bottom-left), and 4 (bottom-right) broken propellers.

batch size is of 128. We present the results of the test set in 5.2.

5.2. Results and Discussion of Classifiers' Performance

The DNN (see 4.1.1) variants have 3 hidden layers, where each hidden layer has the same number of dimensions. The number of dimensions are varied as 64, 128, and 256, see Table 1. DNN 256 variant achieves the best accuracy, precision, and recall of 98.0231%, .9793, and .9795, respectively. The DNN, in comparison to all the other architectures, has the least number of FLOPS but has the highest number of parameters.

The CNN (see 4.1.2) variants have 3 convolution layers of 3x3 and the number of filters are varied as 2-4-8, 4-8-16, and 8-16-32, see Table 2. CNN 8-16-32 variant achieves the best accuracy, precision, and recall of 98.2427%, .9818, and .9816, respectively. The CNN, in comparison to all the other architectures, has a relatively high number of both parameters and FLOPS.

The LSTM (see 4.1.3) variants are unidirectional 1 layer models, and the hidden dimension values are varied as 32, 64, and 128, see Table 3. LSTM 128 variant achieves the best accuracy, precision, and recall of 98.6271%, .9858, and .9856, respectively. The LSTM, in comparison to all the other architectures, has a relatively low number of parameters but the number FLOPS are on the higher end.

The Transformer Encoder (see 4.1.4) variants are 1 layer models with 2 attention heads and feed-forward dimension values are varied as 32, 64, and 128, see Table 4. Transformer Encoder 128 variant achieves the best accuracy, precision, and recall of 98.2976%, .9823, and .9821, respectively. When considering all the metrics (parameters, FLOPS, accuracy, precision, and recall), the Transformer Encoder outperforms all the other architectures.

From Tables 1-4 we can see that for all the architecture variants, as the model size increases, the accuracy, precision, and recall performance metrics improved as well. In terms of accuracy, precision, and recall, CNN variants performed better than the DNN variants, and the Transformer Encoder variants performed better than the CNN variants, with the LSTM variants performing the best.

DNN	Param	FLOPS	Acc	Prec	Rec
64	51k	.1M	97.4739	.9736	.9737
128	118k	.24M	97.6936	.9759	.9763
256	302k	.6M	98.0231	.9793	.9795

Table 1. Comparisons for DNN, each variant is of 3 layers.

CNN	Param	FLOPS	Acc	Prec	Rec
2-4-8	32k	.57M	97.6387	.9753	.9756
4-8-16	65k	2.1M	97.9132	.9782	.9785
8-16-32	133k	8M	98.2427	.9818	.9816

Table 2. Comparisons for CNN, each variant has 3x3 conv.

LSTM	Param	FLOPS	Acc	Prec	Rec
32 dim	13k	1.9M	98.0780	.9804	.9796
48 dim	23k	3.7M	98.1878	.9811	.9809
64 dim	35k	6.1M	98.6271	.9858	.9856

Table 3. Comparisons for LSTM variants

TrEnc	Param	FLOPS	Acc	Prec	Rec
2, 32	7k	.32M	98.0231	.9798	.9792
2, 64	8.5k	.4M	98.1329	.9806	.9804
2, 128	11k	.57M	98.2976	.9823	.9821

Table 4. Comparisons for Transformer Encoder, the first column represent number of heads and feedforward dimension respectively for each variant.

6. CONCLUSIONS AND FUTURE WORK

One of the most exciting areas in the aerospace industry today is UAVs and they have been used in a wide range of applications. UAV propeller damage detection is of utmost importance for the resiliency and safety of the drone. Recording the audio data of a drone with an on-board microphone is a promising approach to detect propeller anomalies. We gathered over $5h$ of real-world audio data for propeller anomaly detection, and showed that deep learning models can be used to detect this type of damage that has occurred to a UAV. Our results show that overall Transformer Encoder performs the best, with a lower number of parameters and FLOPS, but higher values of accuracy, precision, and recall.

For future work, the audio data set could be expanded to test the models in further anomalous scenarios, e.g. varying the extent of the propeller damage, performing malicious attacks, or flying outdoors. The models could also be tested on different UAV platforms to assess the transferability and generalizability of the models.

Acknowledgements: We thank Manuel Segarra-Abad and Johannes Pirhonen for their help in the experiments and discussions. This work was supported by the Secure Systems Research Centre (SSRC) - Technology Innovation Institute (TII), Abu Dhabi, U.A.E.

7. REFERENCES

[1] Kolamunna Harini and thilini et al., "Droneprint: Acoustic signatures for open-set drone detection and identification with online data," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, pp. 1–31, 2021.

[2] George K Fourlas and George C Karras, "A survey on fault diagnosis methods for uavs," *2021 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 394–403, 2021.

[3] Sang-ho Ko Lee Jun-yong, Won-tak Lee and Hwa suk Oh, "Fault classification and diagnosis of uav motor based on estimated nonlinear parameter of steady-state model," *Int J Mech Eng Robot Res*, pp. 22–31, 2020.

[4] Alessandro Benini and et al., "Fault detection of a vtol uav using acceleration measurements," *18th European Control Conference (ECC)*, pp. 3990–3995, 2019.

[5] A. Andria, G. and et al., "Design and performance evaluation of drone propellers," *5th IEEE International Workshop on Metrology for AeroSpace (MetroAeroSpace)*, pp. 407–412, 2018.

[6] Wendy A. Okolo Banerjee, Portia and Andrew J. Moore, "In-flight detection of vibration anomalies in unmanned aerial vehicles," *Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems*, vol. 3, 2020.

[7] Jangwon Jung Seo Yoojeong, Beomhui Jang and Sungbin Im, "Uav detection using the cepstral feature with logistic regression," *2018 Tenth International Conference on Ubiquitous and Future Networks*, pp. 219–222, 2018.

[8] Zhu Chen Liu, Wansong and Minghui Zheng, "Audio-based fault diagnosis method for quadrotors using convolutional neural network and transfer learning," *2020 American Control Conference*, pp. 1367–1372, 2020.

[9] Liu Hao and Yunfang Ren et al., "Drone detection based on an audio-assisted camera array," *2017 IEEE Third International Conference on Multimedia Big Data (BigMM)*, pp. 402–406, 2017.

[10] Giuseppe Ciaburro Iannace Gino and Amelia Trematerra, "Fault diagnosis for uav blades using artificial neural network," *Robotics*, vol. 8, pp. 59, 2019.

[11] Emiliano Pallotti Bernardini Andrea, Federica Mangiatordi and Licia Capodiferro, "Drone detection by acoustic signature identification," *Electronic Imaging*, pp. 60–64, 2017.

[12] Ashish Vaswani and Noam Shazeer et al., "Attention is all you need," *Advances in neural information processing systems*, pp. 5998–6008, 2017.

[13] Shaoqing Ren He Kaiming, Xiangyu Zhang and Jian Sun, "Deep residual learning for image recognition," *IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[14] Sara Abdulla Al-Ali Al-Emadi and Abdulaziz Al-Ali, "Audio-based drone detection and identification using deep learning techniques with dataset enhancement through generative adversarial networks," *Sensors*, p. 4953, 2021.

[15] Chun Fui Liew Takehisa Yairi Khan Samir and Richard McWilliam, "Unsupervised anomaly detection in unmanned aerial vehicles," *Applied Soft Computing*, vol. 83, 2019.

[16] Martinez-Carranza Cabrera-Ponce, Aldrich and Caleb Rascon, "Detection of nearby uavs using a multi-microphone array on board a uav," *International Journal of Micro Air Vehicles*, vol. 12, pp. 1–10, 2020.

[17] Newcastle University, "Are your sensors spying on you?," *www.ncl.ac.uk/press/articles/archive/2017/04/sensors/*.

[18] Manukyan Anush and Yves Le Traon et al., "Uav degradation identification for pilot notification using machine learning techniques," *2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA)*, pp. 1–8, 2016.

[19] Saman Zonouz Sadhu, Vidyasagar and Dario Pompili, "On-board deep-learning-based unmanned aerial vehicle fault cause detection and identification," *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5255–5261, 2020.

[20] Amr Mohammad Al-Emadi, Sara Abdulla Al-Ali and Abdulaziz Al-Ali, "Audio based drone detection and identification using deep learning," *2019 15th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 459–464, 2019.

[21] Naveen Suda Liangzhen Lai Zhang, Yundong and Vikas Chandra, "Hello edge: Keyword spotting on microcontrollers," *arXiv preprint*, p. arXiv:1711.07128, 2017.

[22] Yongbao Ma Wan Tongtang, Yi Zhou and Hongqing Liu, "Noise robust sound event detection using deep learning and audio enhancement," *2019 IEEE international symposium on signal processing and information technology (ISSPIT)*, pp. 1–5, 2019.

[23] Diederik P. Kingma and Jimmy Ba, "Adam: A method for stochastic optimization," *arXiv preprint*, p. arXiv:1412.6980, 2017.