DOI: 10.61782/fa.2023.0049

COMPREHENSIVE DATABASE OF UAV SOUNDS FOR MACHINE LEARNING

Sebastian Kümmritz[1*] **Lothar Paul**[1]
[1] H2 Think gGmbH, Berlin, Germany

ABSTRACT

The use of Unmanned Aerial Vehicles (UAV) is steadily increasing. Besides the resulting benefits, there are also risks and dangers such as airspace violations or terrorist attacks, which require the development of effective drone defence systems. The realization of a drone defence system implies the following stages: Detection, Identification, Localisation and Neutralisation. In this paper, we address the drone detection and identification (classification) stage via acoustics using machine learning algorithms. A major problem with this approach is the lack of publicly available drone audio data. For this reason, we are building an extensive, open-access database consisting of both existing drone sounds and own drone recordings. This database contains drone sounds for all open drone classes from C0 (< 250 g) to C4 (< 25 kg).

Keywords: *UAV, drones, machine learning, data base*

1. INTRODUCTION

1.1 Background

The use of Unmanned Aerial Vehicles, frequently referred to as drones, is continually expanding. This offers enormous opportunities, but also leads to completely new dangers. An event often cited in this context and crucial for research in this area took place in December 2018 in the immediate vicinity of London Gatwick Airport, when more than 50 drones were spotted over a 15-hour period [1]. This resulted in 30 hours of flight disruption, costing £790,000. A similar event also took place at Frankfurt Airport on 9 May 2019 [2].

In addition to these disruptions to public infrastructure, there are a variety of other adverse events associated with UAVs, such as smuggling [3], disruption of events [4], invasion of privacy [5], assassination and terrorist attacks [6–8], and many more. To counter these growing threats, new strategies for detection, localisation and neutralisation are needed. The costs of an UAV detection system play a decisive role here. Existing solutions are not affordable for many users such as private individuals or public institutions.

1.2 UAV Detection Principles

The mainly used techniques for the detection of UAVs are RF-based techniques [9], image recognition [10], radar [11], infrared and acoustics [12]. All these methods offer different advantages and disadvantages, which are summarised in Tab. 1. The table is mainly taken from a paper by Park et al. [13], which gives a good overview of the current state of UAV detection.

While radar-based solutions cover a long range and are less sensitive to weather conditions, they are expensive, cannot be used everywhere due to regulations and are vulnerable to obstacles. RF-based methods allow for the localisation of the operator, but they are not applicable for unsupervised UAVs. Image recognition methods can be used to develop miniaturized solutions, but they are highly dependent on weather, which influences visibility and lighting conditions, as well as on the presence of obstacles. Acoustical sensors also offer the possibility of miniaturized solutions and they are very cheap, but they only offer a limited detection range. To compensate for the disadvantages of the individual methods, a combination of different techniques is often used [13].

*__Corresponding author__: *kuemmritz@h2think.org.*

Table 1: Summary of advantages and disadvanteges of different principles for UAV detection

Sensing principle	Advantages	Disadvantages
infrared	• weather-insensitive • long detection range	• low accuracy
RF	• obstacle-free • detect the operator	• not for unsupervised UAVs
radar	• weather-insensitive • long range	• high expense • regulations • vulnerable to obstacles
optic	• low expense • miniaturized • identification	• weather-insensitive • vulnerable to obstacles
acoustics	• miniaturized • cheap	• low detection range • low accuracy

1.3 Audio based UAV Detection with Machine Learning

This article focuses on the partial aspect of audio-based UAV detection. Although it only works at comparatively short distances, it does work in poor visibility conditions or when the view is obscured. Additionally, a big motivation in focusing on acoustics is rooted in its cost-effectiveness.

Automated detection of UAVs from acoustic data is possible with the help of machine learning algorithms. In [12], Dumitrescu et al. as well as Jeon et al. in [14] present different approaches such as the Gaussian Mixture Model (GMM), Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) to detect commercial UAVs. In [15] Li et al. focus on CNN and Shi et al. use Hidden Markov Models (HMM) in [16].

This paper does not discuss the advantages and disadvantages of these different machine learning techniques. Instead, we initially focus on the first step of each of these techniques: the data for the learning stage.

1.4 Lack of available audio data of UAVs

A basic prerequisite for machine learning is a sufficient amount of training data. This is precisely where currently a major challenge lies, as it is stated by Al-Emadi et al.: "There is a lack of publicly available data sets with acoustic recordings of UAVs, which are particularly important for training neural networks for UAV recognition and classification" [17]. Similarly, Utebayeva et al. complain in [18]: "The data preparation part also gathered UAV sounds from open sources. This stage was laborious due to the fact that there were not sufficient UAV sounds in known databases."

Resulting from these findings, the aim of this work is to fill the gap by creating a database of UAV sounds and making it available for public use. For that we build up a comprehensive data base consisting of gathered UAV sound and own recordings. Data for this study were meticulously collected through a comprehensive review of publicly available sources and by seeking input from fellow researchers in the field. The UAV recordings were obtained in two measurement campaigns in which we recorded a total of 16 different UAVs, at least two per category.

2. DATA BASE

2.1 Data sources

Aside from our own measurements, which represent about 85 % of all UAV sounds and approximately 70 % of the data base, we were able to include external, differently extended data from 52 sources. These include UAV sounds resected from 47 public domain UAV flight video clips from youtube and sound data released by the authors of thematically connected works on GitHub [17, 19, 20], though many of them, unfortunaley, could not be reached directly for some additional (sometimes indispensable) data labeling information [17,20]. Finally, we were able to include UAV sound data that were recorded by contacted partners [21] and institutions [22]. Together with our own recordings, 44 UAV types were covered so far, the data base containes 23.42 hours of raw UAV sounds and 3.22 hours of additional technical sounds actually.

2.2 UAV classes

Since certain UAV types are permitted in some airspaces, the UAV category is of particular importance. In 2017, extensive rules and regulations pertaining to the operation of UAVs were implemented in Germany for the first time [23]. Depending on the take-off weight of the UAV, these made it mandatory for the UAV to be marked, for the remote-controlling pilot to provide proof of knowledge and/or for the state aviation authority to issue an ascent permit. In 2021, the legal framework was harmonised at European level by EU Regulations (EU) 2019/945 and

(EU) 2019/947 [24] to establish common rules and standards. Since then, a risk-based approach has been followed, which does not differentiate between private and commercial applications. Tab. 2 summarises the classes defined in [24]. The classes C3 and C4 have the same operation ranges. They only differ in the systems installed. C3 requires altimeter, remote identification and geofencing, C4 does not.

Table 2: UAV classes according to [24].

Class	Weight	Allowed operation range	Operational Sub-Catregory
C0	< 250 g	Areas in which it cannot be ruled out that uninvolved persons may be flown over.	A1 - Near persons
C1	< 900 g	Areas in which it can be assumend that no uninvolved persons will be overflown.	A1 - Near persons
C2	< 4 kg	Areas in which a minimum horizontal distance of 30 m (5 m in low-speed modus) to uninvolved persons can be maintained.	A2 - safe distance to persons
C3\C4	< 25 kg	Areas in which no uninvolved persons are endangered and at least 150 m away from residental areas, industrial facilities, recreational facilities or similar.	A3 - far away from people

2.3 Data base structure

To tackle the labeling task depending on the classification approach we introduce a SQL data base. All recordings in this data base are categorized by the following classes:

- **(A)** Single UAV free-field
- **(B)** UAV outdoor (pure)
- **(C)** UAV outdoor (noise, disturbances)
- **(D)** Multiple (> 1) UAVs, free-field
- **(E)** Multiple (> 1) UAVs, outdoor (pure)
- **(F)** Multiple (> 1) UAVs, outdoor (noise, disturbances)
- **(G)** Mixed (UAV sounds overlapped with noises, babbles, . . .)
- **(H)** Background (different scenes without any UAVs)
- **(I)** Short period babble sound

Free-field recordings are achieved in a fully anechoic chamber. Two classes (H and I) do not represent UAV sounds. Instead, they comprise a collection of background sounds and disturbance data, which are useful for the creation of test datasets in the context of machine learning algorithms. If necessary, new content classes can be introduced easily. Each object contains the properties File ID, FileName, Sampling Rate, Replay-Time, Format, Directory Link, Origin, Drone Type, Uniformity Rotation, Distance, Signal Quality, Remarks and Weight. A truncated overview about the properties with corresponding subproperties can be seen in Fig. 1.

These content classes are represented in equally structured tables where each line entry (tuple) corresponds to a sound file of the appropriate content. Underlying binary files are not stored in the data base itself but can be accessed by means of stored links (directory paths).

Together with the recorded UAV type and file access information, multiple descriptive and classification attributes are – or may be - provided in the table line entries, that should enable a user to efficiently select and pick the data files he needs (by data base queries). This additional, file-related information, which represents the value of the data base consists of technical, file-, UAV-, source- and processing-related marks and pre-classifications.

Such pre-classifications usually result from the recording procedure and situation. So, operators may preclassify sounds into flight phases with increasing (takeoff, rising) or constant rotation speed or by the distance between microphones and UAVs.

The file related lines also contain technical information about the sound file itself (sampling rate, number of channel files, file format, replay time and signal quality) and indices pointing to tuples in other provided tables (see below). Additional non-formalised information or short remarks can be added in a designated field.

Finally, further file related, descriptive table entries are forseen to support classification procedures. For that aim, predefined sound data descriptors (own recordings, downloaded external, segmented or continuous etc.) or processing-status information (raw, filtered, resampled, used for teaching etc.), can be added and be used as search criteria in SQL queries.

The data base contains two other important tables which contain further descriptive information, that is not file-related but can be fetched using mandatory index en-

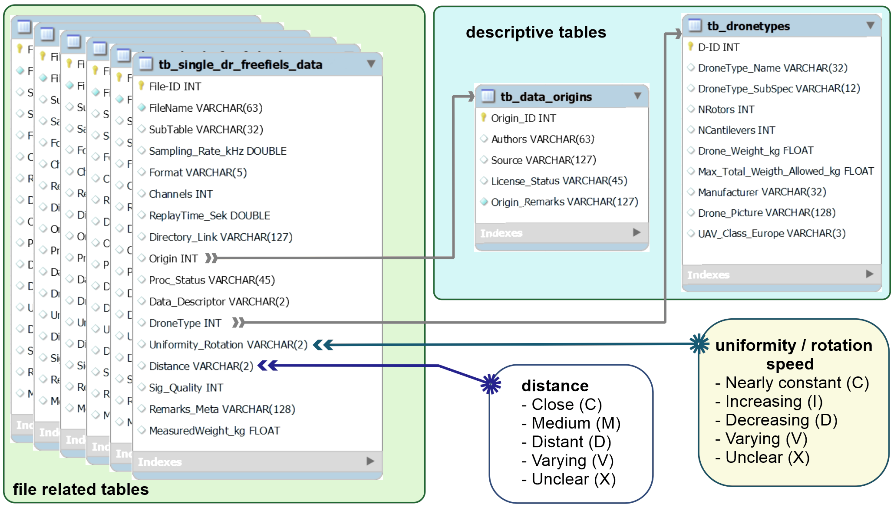

Figure 1: Structure of the database.

tries in the file tuples, one for the UAV types covered in the data bank and another for elaboration of the data sources.

The table tb_dronetype provides technical descriptive information about the UAVs (from product specification or description: UAV type, sub-type, number of rotors and cantilevers, UAV weight with and without payload), name of UAV manufacturers and optionally a picture link. It also contains the European UAV class that corresponds to the given weight values.

Further, the table tb_data_origins provides names of owners respectively authors of related reports or research works, related internet links, licensing status information and additional remarks on authorship and publications.

2.4 Data base access

The complete dataset for our database has been made publicly accessible via the Mobilitheksplattform of the Federal Ministry of Digital Affairs and Transport, Germany, and can be accessed through the link in the footnote [1]. However, it's important to note that the database is not hosted online. To utilize the data, users must download and host the database locally. Detailed instructions on how to set up and use the database can be found at the provided link.

[1] `https://mobilithek.info/offers/605778370199691264`

3. DATA BASE CONTENT

3.1 Overview

The goal of our work is to create a comprehensive, publicly accessible database of UAV sounds, which can be used as a training basis for own classifiers. The database was compiled through meticulous exploration of existing literature, outreach to other scholars in the field, and careful extraction of audio from YouTube, complemented by our own original UAV recordings. The yotube data were recorded with Audacity during their playback and then post-processed by normalizing the volume and removing of non-UAV sounds. Both, the collection and the own recordings, led to a total of 4924 UAV sound files with a total playback duration of 23.42 hours. By now we have samples of 44 different UAV models, 9 for class 0, 11 for class 1, 12 for class 2 and 13 for class 3/4.

The own recordings were carried out during two measurement compaigns, the first one in an anechoic chamber and the second one outdoor. During both compaigns the sound of a total of 16 different UAV models were collected, with at least two UAVs per class. An overview about all used UAVs can be found in Tab. 3.

3.1.1 Anechoic chamber

The first compaign took place in the anechoic chamber of the Institute of Acoustics and Speech Communication of the Technical University Dresden (Fig. 2). This room has a volume of 1000 m^3 what is big enough to let huge

Table 3: Summary of measured UAVs; In\Out indicates, if the corresponding UAVs have been measured indoor and\or outdoor.

UAV model	weight	Category	In	Out
Cartronic	15 g	C0	X	-
Potensic Firefly A20	25 g	C0	X	-
Emotion	93 g	C0	X	-
DJI Mavic Mini 3 Pro	247 g	C0	X	-
DJI Avatar	417 g	C1	X	-
DJI Mavic Air 2	566 g	C1	X	-
DJI Mavic Pro	673 g	C1	X	-
DJI Mavic 2 Pro	821 g	C1	X	-
DJI Mavic 3e	914 g	C2	X	X
Phantom 4	1,37 kg	C2	X	-
DJI Phantom 4 RTK	1,43 kg	C2	X	X
DJI Inspire 2	3,41 kg	C2	X	X
DJI Matrice 30 T	3,80 kg	C2	-	X
HP-X4 2020	5,49 kg	C3\C4	X	X
DJI M300	8,00 kg	C3\C4	X	X
HP-E616P-1	25,10 kg	C3\C4	-	X

UAVs fly. Measurements in an anechoic chamber are not influenced by acoustic influences from the environment such as noise or reflections. Reflections, especially from the ground, cause the UAV sound to overlap with itself in a short time interval, which can lead to serious deviations from the original signal. This type of interference is known in an acoustic context as the comb filter effect. The pure UAV sound recordings are, therefore, very valuable as a basis for data augmentation. Thus, by convolving the recordings with appropriate room impulse responses as well as mixing them with ambient noise, they can be used to create any real-world scenarios.

The measurements were conducted using five microphones, each distinct in its specifications:

1. WA101 PU Sound Intensity Probe by weles acoustics
2. AKG C451B
3. AKG C451B
4. Behringer ECM8000
5. BSWA MP253 microphone capsule with MA231 pre-amplifier

The WA101 PU Sound Intensity Probe was employed to record both sound pressure and velocity, despite the fact that sound velocity does not play a significant role in the detection and classification of UAVs. Both the first and

Figure 2: Measurements in anechoic chamber - ©by Fraunhofer IVI

fifth microphones are ICP measuring microphones, while the others are generally utilized in music production.

Taking into account the importance of cost-effectiveness in future experiments, our primary aim was to use these audio microphones as the main recording instruments. To ensure accuracy, microphones 1 and 2, as well as 4 and 5, were carefully arranged in close proximity to one another. This setup was designed to observe any potential differences or inconsistencies in signal quality between the different types of microphones. As part of the systematic setup, all the microphones were securely mounted on tripods at an approximate height of 1.5 meters to maintain a uniform recording environment.

Fig. 3 illustrates the frequency responses of the BSWA and Behringer microphones, presented in 1/6 octave steps for a randomly selected UAV measurement. As per the specifications in their respective data sheets, both these microphones are known to exhibit a nearly flat frequency response. This visualization effectively demonstrates that the frequency responses of these two microphones align substantially, further strengthening their reliability in such experiments. The figure clearly illustrates that the frequency responses of the two microphones are substantially similar.

We followed a consistent procedure for all UAV measurements, as illustrated in Fig. 4. The UAVs initiated their flight from a designated launch platform and then hovered for a minimum of 5 seconds near microphones 1 and 2. Following this, they traversed the room at an average altitude of 3 meters, passing by microphone 3. Once on the opposite side of the room, the UAVs hovered for another 5 seconds near microphones 4 and 5. The

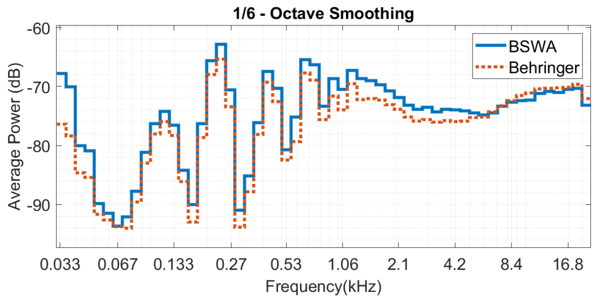

Figure 3: Comparison of the frequency response of Behringer and BSWA microphone

return flight mirrored the initial route, with the UAVs flying back past microphone 3 and pausing once more near microphones 1 and 2 for 5 seconds before finally landing back on the platform.

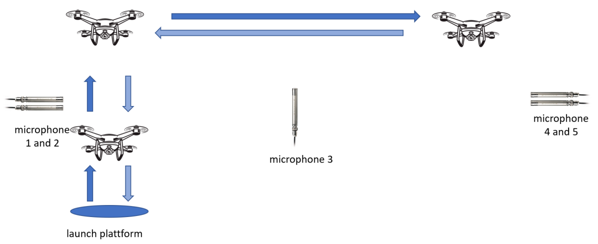

Figure 4: Schematic representation of the measurement process with take-off (1), hovering (2), flying in one direction (3), hovering in the oposite side of the room (4), flying back (5) and landing (6)

The two hovering locations were approximately 20 meters apart. Each traversal of this distance was conducted at a slow pace for the initial measurement cycle and at a faster speed for the subsequent round. The measurements incorporated two distinct flight speed scenarios. Although the exact speed wasn't quantified, it was broadly estimated that the UAVs flew at around 1-2 m/s during the slow flight scenario and 4-5 m/s during the faster flight scenario.

Each measurement was conducted at least twice to ensure data reliability. When conditions permitted, we conducted additional measurement cycles, sometimes with extra payloads. In one instance, the UAV propellers were alternated and the entire measurement cycle (slow, fast, without payload, with payload) was repeated.

For a more in-depth understanding, a video of these measurements is available on the website of our project partner, the Fraunhofer Institute for Transportation and Infrastructure [2] .

3.1.2 Outdoor

We carried out a second measurement campaign outdoors, with the intent to curate a dataset featuring the same UAVs but under varying environmental conditions. Such a dataset, though susceptible to the unpredictable elements of outdoor conditions, would prove beneficial for testing data robustness. These measurements were conducted at the test field of the Fraunhofer Institute for Transportation and Infrastructure Systems in Dresden.

Instead of conducting measurements with individual microphones as in our previous campaign, we carried out the current campaign using three microphone arrays, each comprising 24 microphones. Our measurement setup is illustrated in Figure 5, which displays both a photograph of the measurement situation with the arrays (composed of Behringer microphones) and a satellite view marking the positions of these arrays. The red circles indicate array locations, while the blue rectangle depicts the location of the pavilion that served as the base station.. An orange line outlining the approximate flight path of the UAVs is also included for reference.

Since our outdoor measurements incorporated a degree of redundancy, the database only includes sum signals of all microphones per array. This not only streamlined data reduction but also ensured enhanced signal quality due to the increased signal-to-noise ratio.

The microphone arrays employed in this experiment have a circular structure with a diameter of 34 cm. As a future plan, we aim to utilize the recorded data for tracking UAV trajectories by implementing simple direction-of-arrival beamformers. This aligns with our overarching goal of a comprehensive UAV detection and localization system.

4. CONCLUSION AND FUTURE WORK

Our examination of existing literature on audio-based UAV detection using machine learning algorithms has underscored a pronounced lack of suitable training data. To fill this gap, we have compiled a database, rich in diversity,

[2] `https://www.ivi.fraunhofer.de/de/forschungsfelder/fahrzeug-und-antriebstechnik/fahrzeug-und-verkehrssicherheit/luftverkehrssicherheit.html`

Figure 5: Upper image: An on-site photograph from the measurement campaign, featuring one of the used arrays in the foreground. Lower image: Google Maps view of the measurement location. The red circles indicate array positions, the orange line represents the approximate UAV flight path, and the blue rectangle outlines the location of the pavilion.

comprised of UAV sounds. It features recordings from 16 distinct UAV models, ensuring at least two models from each UAV class are represented.

The data compilation involved conducting two distinct measurement campaigns: an initial campaign within an anechoic chamber and a subsequent outdoor campaign. Further, we enriched the dataset with UAV recordings extracted from a range of sources, including digital platforms like YouTube, and valuable inputs from fellow researchers.

Collectively, our carefully curated database aggregates 23.42 hours of UAV recordings, covering 44 diverse UAV models, and serves as a robust resource for future machine learning algorithm training in UAV detection.

Developing a high-performing classifier necessitates the availability of ”good” data. This principle is echoed in a statement by Alexander Pretschner, Professor for Software and Systems Engineering at the Technical University of Munich, who asserts, “as long as a machine learning algorithm has enough examples of inputs, it finds the outputs. If the data are all clean, that’s useful, but they usually aren’t” [25]. Armed with this insight, our subsequent step is to leverage the ”good” data, primarily from the anechoic chamber measurements, and generalize them through suitable augmentation techniques such as pitching, modulation, and the addition of ambient noise or white noise.

The quality of this generalization can then be validated against the remaining data in the database, especially our own outdoor measurements. This verification process helps us establish the robustness of the classifier in varied real-world conditions.

Finally, as another aspect of our work but not central in this context, we intend to utilize our outdoor measurements to determine UAV trajectories using acoustic cameras. This initiative forms part of our ongoing efforts towards establishing a comprehensive UAV detection and localization system.

5. ACKNOWLEDGMENTS

This project was funded by the mFund programme of the Federal Ministry of Digital Affairs and Transport (Germany). We thank our project partners Susanne Günther and Vanessa Hilse of the Fraunhofer Institute for Transportation and Infrastructure Systems for the good and productive cooperation. We would also like to thank our colleague Ernst Swanepoel for his conscientious planning and execution of both measurement campaigns.

6. REFERENCES

[1] “Gatwick airport drone attack: Police have ’no lines of inquiry’,” *BBC News*, 09 2019.

[2] “143 flights cancelled at frankfurt airport due to drone sighting,” *The Local*, 05 2019.

[3] “’well-organised’ gang flew drones carrying drugs into prisons,” *BBC News*, 08 2018.

[4] “Man fined after flying drones over premier league stadiums,” *BBC News*, 09 2015.

[5] “San carlos woman says drone hovered near bedroom, wouldn’t go away,” *10 News*, 11 2017.

[6] M. Brocchetto, C. Dominguez, J. Sterling, and S. Pozzebon, “Venezuelan president survives apparent drone assassination attempt,” *CNN*, 08 2018.

[7] W. Ripley, "Drone with radioactive material found on japanese prime minister's roof," *CNN*, 04 2015.

[8] "Massachusetts man charged with plotting attack on pentagon and u.s. capitol and attempting to provide material support to a foreign terrorist organization," *Federal Bureau of Investigation*, 09 2011.

[9] P. Nguyen, M. Ravindranatha, A. Nguyen, R. Han, and T. Vu, "Investigating cost-effective RF-based detection of drones," in *Proceedings of the 2nd Workshop on Micro Aerial Vehicle Networks, Systems, and Applications for Civilian Use*, pp. 17–22, ACM.

[10] X. Shi, C. Yang, W. Xie, C. Liang, Z. Shi, and J. Chen, "Anti-drone system with multiple surveillance technologies: Architecture, implementation, and challenges," *IEEE Communications Magazine*, vol. 56, no. 4, pp. 68–74.

[11] B. Knoedler, R. Zemmari, and W. Koch, "On the detection of small UAV using a GSM passive coherent location system," in *2016 17th International Radar Symp. (IRS)*, pp. 1–4, IEEE.

[12] C. Dumitrescu, M. Minea, I. M. Costea, I. Cosmin Chiva, and A. Semenescu, "Development of an acoustic system for UAV detection," *Sensors*, vol. 20, no. 17, p. 4870.

[13] S. Park, H. T. Kim, S. Lee, H. Joo, and H. Kim, "Survey on anti-drone systems: Components, designs, and challenges," *IEEE Access*, vol. 9, pp. 42635–42659.

[14] S. Jeon, J.-W. Shin, Y.-J. Lee, W.-H. Kim, Y. Kwon, and H.-Y. Yang, "Empirical study of drone sound detection in real-life environment with deep neural networks," Publisher: arXiv Version Number: 1.

[15] S. Li, H. Kim, S. Lee, J. C. Gallagher, D. Kim, S. Park, and E. T. Matson, "Convolutional neural networks for analyzing unmanned aerial vehicles sound," in *2018 18th International Conference on Control, Automation and Systems (ICCAS)*, pp. 862–866, 2018.

[16] 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.

[17] S. Al-Emadi, A. Al-Ali, and A. Al-Ali, "Audio-based drone detection and identification using deep learning techniques with dataset enhancement through generative adversarial networks," *Sensors*, vol. 21, no. 15, p. 4953.

[18] D. Utebayeva, L. Ilipbayeva, and E. T. Matson, "Practical study of recurrent neural networks for efficient real-time drone sound detection: A review," *Drones*, vol. 7, no. 1, p. 26.

[19] K. J. Piczak, "ESC: Dataset for environmental sound classification," in *Proc. of the 23rd ACM international conference on Multimedia*, pp. 1015–1018, ACM.

[20] F. Svanstrom, C. Englund, and F. Alonso-Fernandez, "Real-time drone detection and tracking with visible, thermal and acoustic sensors," in *2020 25th International Conference on Pattern Recognition (ICPR)*, pp. 7265–7272, IEEE, 01 2021.

[21] P. Alloza, B. Vonrhein, and A. Movahed, "Sound localization of drones using an acoustic camera," in *QUIET DRONES International e-Symp. on UAV/UAS Noise*, 10 2020.

[22] J. Treichel and S. Körper, "Untersuchung der Geräuschemission von Drohnen," *www.ingenieur.de*, 01 2021.

[23] "EU-Regelungen für Drohnen," *Bundesministerium für Digitales und Verkehr*, 11 2021.

[24] "Easy access rules for unmanned aircraft systems (regulations (EU) 2019/947 and 2019/945)," *European Union Aviation Safety Agency*, 09 2022.

[25] A. Karp, J. Hiesserich, and P. Cipierre, *Von Artificial zu Augmented Intelligence: was wir von der Kunst lernen können, um mit Software die Zukunft zu gestalten*. Campus Verlag, 2023.