Advances in Science and Technology Research Journal

Advances in Science and Technology Research Journal 2024, 18(1), 36–47
https://doi.org/10.12913/22998624/175863
ISSN 2299-8624, License CC-BY 4.0

Received: 2023.09.23
Accepted: 2023.12.26
Published: 2024.01.15

Acoustic-Based Drone Detection Using Neural Networks – A Comprehensive Analysis

Waldemar Paszkowski[1], Arkadiusz Gola[2*], Antoni Świć[3]

[1] Silesian University of Technology, Faculty of Organization and Management, ul. Roosevelta 26-28, 41-800 Zabrze, Poland

[2] Lublin University of Technology, Faculty of Mechanical Engineering, ul. Nadbystrzycka 36, 20-618 Lublin, Poland

[3] Lublin University of Technology, Faculty of Mechanical Engineering, ul. Nadbystrzycka 36, 20-618 Lublin, Poland

* Corresponding author's e-mail: a.gola@pollub.pl

ABSTRACT

The article presents and describes the implementation of research on the detection of a drone in an urban environment using of the sound features. The methods of drone detection were recognized on the basis of modeling and evaluation of the features of the audio and acoustic signal. The authors proposed the use of a neural network model for the needs of drone detection taking into account acoustic measurements in an anechoic chamber and in an urban environment. The final part presents the obtained results of the drone detection. For the purposes of detection, a neural network model was used in order to recognize the obtained images of the spectograms of sound sources.

Keywords: drone detection, acoustic mesasurements, signal, spectrogram, neural network.

INTRODUCTION

The information contained in the acoustic signal seems to be intuitive in detecting and tracking small flying objects. In the conducted research [1,2], the acoustic sensors used are characterized by the ability to detect small drones (quad-copters) at a range from 20 m [1] up to 600 m [3]. The authors of these studies point to the significant used of type of acoustic sensor and the method of processing the acoustic signal. In the application of sensors, microphone arrays [1,2] have some advantages over a single microphone. They allow not only to detect the presence of the drone, but thanks to the registration of the formed acoustic signal beam, it is possible to test the azimuth and elevation of the drone [4,5]. It is thus possible to identify the angle of its arrival location by means of two or more microphone arrays. In addition, due to beamforming, the signal-to-noise ratio is also increased, which in turn increases the range and indicator of correct detection and tracking. The authors [3] used the configuration for a microphone system in the shape of a tetrahedral pyramid with the microphones spaced approximately 27 cm apart. This arrangement of microphones in combination with the signal processing technique in the form of spectrograms allowed the detection of small drones with 99.5% probability for ranges smaller than 600 m. The propability of a drone being detected reliably depends not only on the type of acoustic sensors, but also on environmental factors such as wind force and background noise levels.

The detection of drones by means of acoustic signals is currently being investigated in two ways. The first is based on the use of one-dimensional information such as the magnitude or phase of an acoustic signal [6]. The second is based on a two-dimensional analysis such as Short Time Fourier Transform (STFT) or wavelets. In the second method, the neural networ is used as a machine learning method because it id suitable for studying a features in a 2D system. The neural network model is one of the basic approximations of image processing detection [7]. In one drone detection study, the authors used the STFT of an acoustic signal as a way to distinguish a flying object in an urban environment from motorized devices, with a drone-like harmonic characteristic. In this case, machine learning methods were used to classify the processed sound features. It should be noted that the efficiency of the applied neural network model is assessed according to the number of training epochs. In the experiment, lower false positives were observed for signals with a harmonic characteristic similar to a drone.

The formulated research problem of the project undertaken at the Silesian University of Technology was to develop a method of drone detection by processing the features of the acoustic signal, with the presence of other disturbing sources in the external environment. It was assumed that the problem would be solved by means of two designed measurement experiments: in an anechoic chamber and in an open space, with the use of specialized equipment for recording and processing the sound features. As part of the research was carried out the measurement and recording of an appropriate number of sound samples containing an acoustic signal from the drone as well as from other sound sources.

The collected data set was the starting point for the development of a method of detecting and identifying the location of the drone. The conducted measurement experiment in the anechoic chamber included the recording of sound samples containing only direct sounds of the drone. These samples constituted a set of pattern acoustic signals coming only from the drone.

REVIEW OF DRONE DETECTION METHODS

In research on the search for the relationship between an object in the environment which is a sound source and the parameters of the acoustic field can be helpful sound emission models. Sound sources of an urbanized environment are generally characterized by a random variability of acoustic emission, which can be investigated with the use of acoustic signals and audio signals. In general, signals from the emission of sound sources of machinery and equipment in an urban environment are generally non-stationary.

Among the research on the assessment of the quality of acoustic/audio signals, two categories of methods can be distinguished:

- subjective methods – consisting in assessing human hearing sensations,
- objective methods – based on the use of approximate mathematical models.
- In particular, the following can be distinguished among the objective methods:
- signal methods – most often the evaluated signal is compared with the original signal, without distortion (reference signal). There are also methods that do not include a reference signal,
- parametric methods – the sound quality is assessed on the basis of the knowledge about the applied processing technique and the knowledge of its parameters, which are the input arguments of the evaluation algorithm.

Digital methods of signal processing and analysis allow for the determination of many of its essential features and characteristics in the field of time, frequency and amplitude [8]. Along with the progress in the application of the methods of signal identification being the subject of the observed acoustic phenomenon is justified the description of simultaneous analysis of time and frequency parameters/features. Obtaining complete information about the features of the signal is of particular importance in view of the research area undertaken. The analysis of non-stationary spectral properties of an acoustic signal requires the use of windows that automatically narrow when analyzing high frequencies and are automatically expanded when analyzing low frequencies. The known methods of signal transformation allow for local analysis of the spectrum changing with time. Therefore, in tasks requiring the determination of local representations of the signal, time is functionally divided into sections where the signal can be considered stationary. The mathematical model of an acoustic/audio signal, in the case of its discretization, can be represented as a feature vector. Such a representation of the signal allows its digital form to be used for the purposes of advanced research.

The methods of assessing the acoustic/audio signal use advanced models and techniques of its digital processing. It was assumed that in order to detect the tested drone signal as a sound source, physical features will be modeled and represented in terms of domains. For the needs of signal features evaluation (Fig. 1), the following actions were adopted, i.e.:

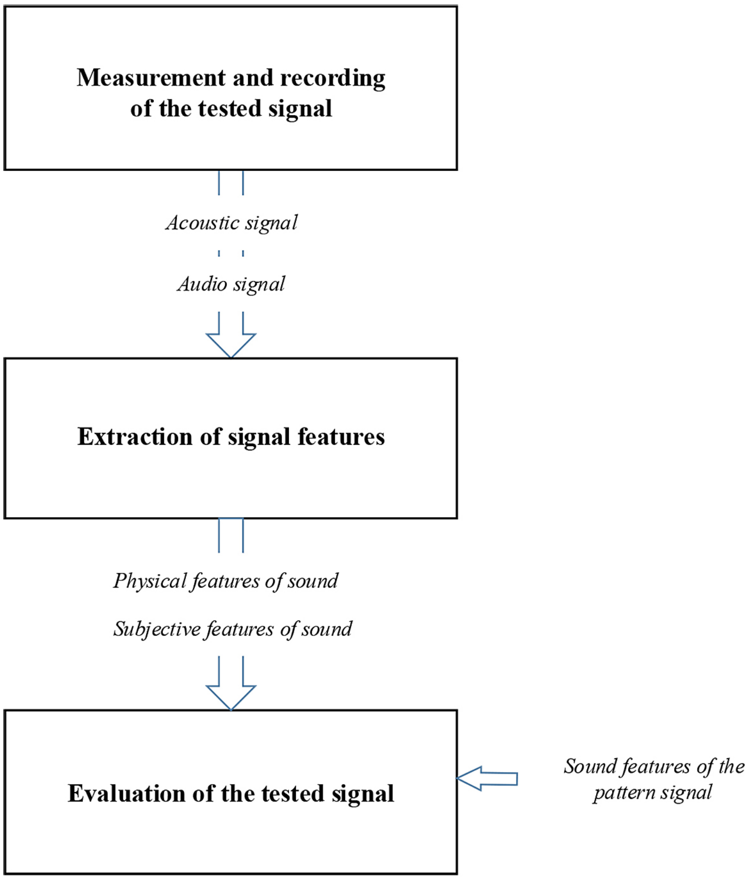


Figure 1. Scheme of processing the sound features of the tested signal for the needs of their assessment

- measurement and recording of the tested signal: it is the starting point for obtaining and processing information about the features of acoustic and audio signals,
- extraction of signal features: this step consists in identifying the physical features of the signal with respect to the selected domain of signal analysis,
- evaluation of the tested signal: consists in the use of methods and models that enable the evaluation of the features of the tested signal in relation to the sound features of the pattern signal.

Depending on the adopted purpose of the research, the physical and subjective features of sound may be analyzed. The subjective features of sound include loudness, sharpness, roughness, tonality, strength of sound fluctuation. It was assumed that for the purpose of the research undertaken, the sound pressure level as a physical features of sound will be analyzed. Sound pressure levels are commonly used in tasks on diagnostic evaluation of technical objects, or monitoring their condition. On the other hand, studies on the evaluation of acoustic impressions from natural and artificial sources on the human body use subjective features of sound.

The authors of this article based on the conducted literature and own research, selected methods analyzed of modeling and assessing the features of the acoustic/audio signal. The applied modeling methods for the diagnostic purposes of technical objects usually require an appropriate processing of acoustic signals in the time domain, frequency domain or both time and frequency [9]. The purpose of such signal processing is the possibility of obtaining, among others estimates characterized, by a high correlation with the monitored phenomenon. In this approach, were analyzed the following methods of assessing the acoustic/audio signal [10-13], i.e.:
- entropy,
- variance,
- neural network.

An audio signal is a representation of information that is directly related to an acoustic signal in the entropy method. This information can be represented by the physical features of the sound. Shannon based his quantitative information theory on the assumption that less likely states give more information. The measure of the amount of information should be such that its value increases along with the decreasing probability of the state to which the information relates. Taking into account the assumptions of information theory as a computational method, the Shannon entropy can be used to determine the assessment of the variability of individual physical vectors of signal characteristics at various levels of its decomposition. The key step of the applied computational procedure is to determine the optimal decomposition tree for the vector of each of the physical characteristics of the signal by means of wavelet packets. The purpose of the use of the entropy tree is to calculate the entropy of approximation and the details of the physical vectors of the signal features at successive levels of decomposition, which makes it possible to analyze their variability in the signal. The evaluation of the generated vector of the physical features of the signal can be performed with the use of binary trees. A tree optimization algorithm is used to evaluate the model of individual feature vectors representing a binary tree, e.i. best tree. This algorithm is based on the search for the optimal decomposition tree based on the minimum entropy criterion, which is the lowest value of the total entropy of all end nodes at a given decomposition level, in all possible tree structures. Entropy is associated with each tree node using a wavelet function and is computed at each decomposition level. In order to classify the tested signals, the Euclidean distance between the path length vectors of binary trees representing the feature vectors of the tested signal and pattern signals is used as a measure. The decisive criterion for the classification of the tested signal in relation to a given pattern signal may be the minimum total distance between two vectors representing the physical features of the signals.

The method of variance is one of the set of statistical methods generally used for comparing observations. It allows you to analyze and evaluate the results that depend on one or more factors acting simultaneously in a given experiment. One of its applications is singular classification statistical research. Then, the influence of only one classifying factor (controlled on many levels) is examined on the results of the observations. An important advantage of the analysis of variance is the interaction, i.e. the study of the combined effect of several factors. The degree of influence is assessed of one factor on the level of another factor [14].

For the purposes of the drone detection research, it was assumed that the analysis of the physical features of the signal between the classes of pattern signals could be performed with the use of statistical hypothesis verification. In this method, the statistical assessment of the significance of the differences between the variables may be statistically assessed by means of tests. The comparative evaluation of signals can be carried out by means of hypothesis verification. For the purpose of determining these assessments, the differences between the following statistical measures are calculated, i.e.:
- arithmetic means of sound features,
- variances of sound features,
- Pearson’s linear correlation coefficients of sound features.
- In the variance method, the signal evaluation can be as follows:
- determination of correlation coefficients between physical features for pattern signals and the tested signal,
- calculation of the resultant vector length of the determined coefficients of correlation of physical features of sound in individual classes,
- selection of the maximum value of the calculated lengths of the vectors of the correlation coefficients.

The proposed method of classification of the tested signals includes a comparison of the physical features of the pattern signal with the tested signal, by means of the vector length evaluation. The maximum value of the vector length representing the correlation coefficients of the physical features of the signal is assumed as the criterion determining the assignment of the tested signal to the model of the pattern signal.

According to the authors, the advantage the neural network method of over the entropy and variance methods is primarily the possibility of taking into account and modeling various measures of the features assessment including the acoustic signal. In the evaluation of the acoustic signal, the statistical significance of the occurrence of features is usually used, based on the correlation analysis performed. Based on the recognition of the possibilities and limitations of the use of acoustic signal evaluation methods over drone detection, the neural network method was selected for further research. The advantages of neural network models include their availability, simplicity of implementation and accuracy of calculations. It is possible to modify the algorithm in a simplified way with new data for training the network (e.g. if you want to detect a new type of drone with different characteristics of the signals emitted). Many publicly available materials provide information on the use of this method and the construction of appropriate network classifier layers.

RESEARCH PROCEDURE

The presented research procedure (Fig. 2) was used to carry out measurement experiments in an anechoic chamber and in an open space. The first stage of the research involved establishing the design assumptions. The standards specifying the conditions for conducting this type of research have been familiarized with in order to make the most detailed measurements. They concern, among others: the distance at which the microphones should be located from the sound source, measurement times and many other factors that should be met (such as: ensuring appropriate weather conditions, conducting research in undeveloped spaces, avoiding interference in the form of dominant other noise sources).

Prior to the acoustic measurements of the drone in the environment, the background acoustic level of other sources was measured. A background acoustic level of 33.7 dB(A) was recorded.

For performing environmental acoustic measurements, recommendations for meteorological conditions have been specified [15]. In particular, they include boundary conditions such as:

- temperature from -10°C to 50°C,
- humidity from 25% to 90%,
- average wind speed of up to 5 m/s,
- atmospheric pressure from 900 hPa to 1100 hPa.

Weather conditions were satisfactory when taking acoustic measurements of the drone in an outdoor environment. The measurements were conducted in sunny and windless weather at a temperature of 24 degrees. In accordance with the information provided in the standards [16,17], the required criteria were ensured in the tests. The requirements for the appropriate volume of the chamber in which the measurements were made were taken into account ($V = 202.2\ m^3$). Equal distances of 1 m from the source to the microphones were assumed. A Class 1 level meter with band-pass filters was used.

In accordance with the requirements concerning the principles and conditions of measurements [16,18], during their implementation, there should be no people or other objects outside the analysed object in the anechoic chamber. The combination of the measuring microphone with the sound level meter or with other instruments should also allow their placement outside the anechoic chamber. The microphone placed at the measuring point should be directed towards the tested source when taking measurements. The measuring tools used should have valid legalization certificates. Due to the specificity of the analysed source during the measurements taking place in the anechoic chamber, there was a drone operator and a person operating the measuring equipment.

In order to obtain representative results, the conditions such as the measurement time and the distance of the drone from the microphones were kept the same in both research environments. The measurement time of one sample was 1 min and the distance between the drone and the microphones was 3 m. Measurements were made using the Mavic 2 Pro Camera drone. It is characterized by multidirectional obstacle detection and its maximum speed can be up to 72 km/h. The arrangement of the measurement path was assumed with the drone located 3 m from the microphones. The array of microphones was directed immediately towards the noise source which was the

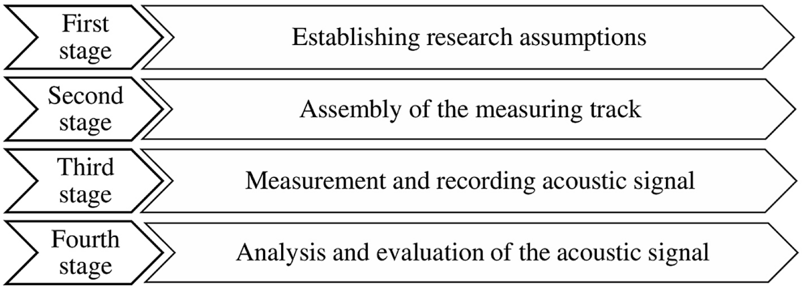


Figure 2. Stages of the applied research procedure

drone. The acoustic measurements were carried out for the case in which the drone was moving in a semi-spherical space in relation to the adopted microphone system. The second stage of the research was the assembly and testing of the measuring track (Fig. 3).

The input elements of the measurement path were a set of four microphones (1), which were characterized by an extremely flat frequency characteristic, ultra-high sound resolution and omnidirectional characteristic. Their sensitivity was 70 dB, with an impedance of 200 Ohm. They belong to the type of condenser microphones with a frequency response in the range of 20–20.000 Hz. These microphones were connected directly to the power supply system (2). Another element of the measuring circuit was an analog-to-digital converter with programmable filters and appropriate amplification (3). This element allowed to support four mono channels which were treated independently after the appropriate configuration on the hardware supporting the module. The frequency of bit clocks can be changed, as well as the position and parameters of the electronic components on the board as it is a highly modifiable digital circuit. The final element of the measurement path was a minicomputer (4), in this case BeagleBone Black which performed the tasks of: sound recording, frequency analysis and spectrogram processing, using a neural network and data acquisition.

For the needs of the measurements a rack has been assembled that allows the microphones to be located in a fixed position. For this purpose, were used aluminum strips with a length of 1 m and 0.5 m. Throughout the length of the slats, holes were made at a constant distance – every 0.5 m, (Fig. 4.). The adoption of fixed distances between the holes allowed for the installation of the microphones in a variable configuration, which made it possible to choose the optimal arrangement of the microphones. The assembled rack was attached to two microphone stands, also at a constant distance of 1.3 m from the ground. For a fixed configuration of microphones handles with

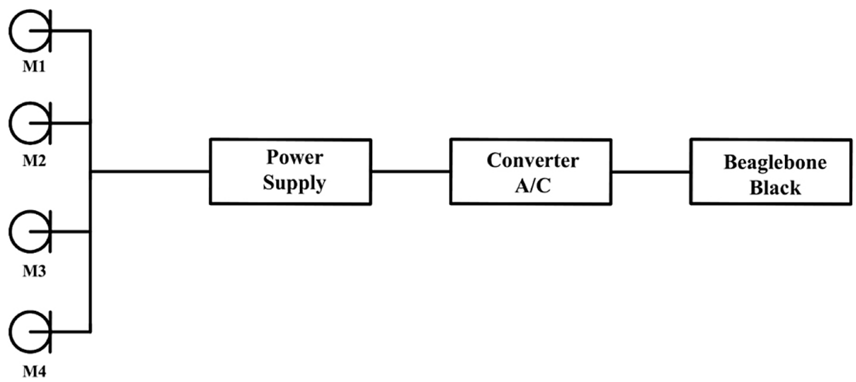


Figure 3. Elements of the assembled measuring track

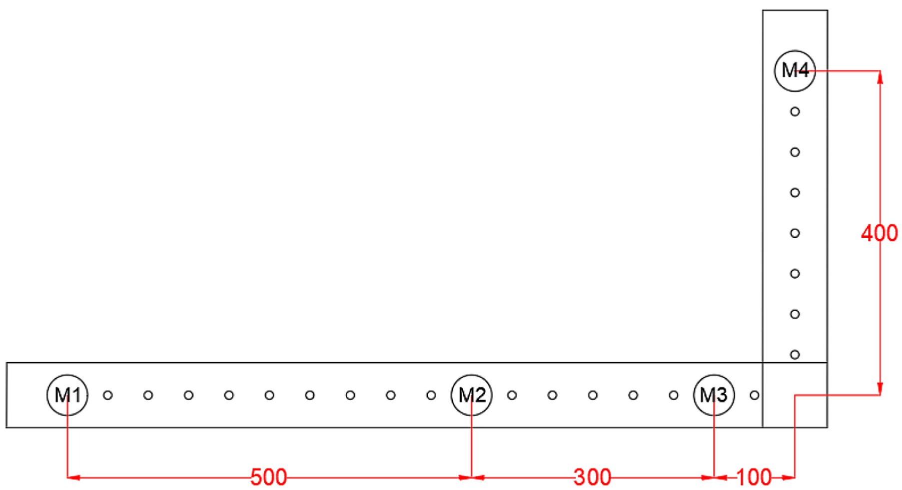


Figure 4. Arrangement of mounting strips with the adopted location of microphones

microphones are attached in the holes. The third stage consisted in measuring and recording the acoustic signal. At this stage of the research using the created measuring track, measurements were made in a confined space and in an open space. Before starting the drone measurements, the background noise level was measured. Its value was correspondingly in the anechoic chamber – at the level of 32 dB and in the open space 24 dB. Measurements in a confined space took into account only sounds coming directly from the drone (no reflections). On the other hand, in the open space (Fig. 5.), an acoustic signal was recorded including direct and reflected sounds from the drone and other sources.

The recording and processing of the acoustic signal was strictly defined. The signal from four microphones powered by the phantom power supply was sent to the PCM1864EVM module in the form of a differential XLR - RCA connection. Then, the signal from the module was transferred to the Beaglebone Black device in the form of three lines – signal, clock and synchronization. The PCM module acted as a master so it generated its own clock and timing. To support four mono channels, the clock and timing signals had to be modified so that the data would be received in TDM format. Appropriately modified driver, severely implemented in the kernel on the BBB (properly compiled), read the signal from three lines and distinguished between four mono channels. The fourth step involved the analysis and evaluation of the acoustic signal. The detection algorithm consisted of examining the signal from microphones in cooperation with the controller, recording samples of an appropriate length of time and performing frequency analysis. The result of data processing was a spectrogram. The spectrogram data was a set of input data to the neural network, which was designed to support the analysis and evaluation of data, taking into account the model data obtained from the anechoic chamber. If the data distribution image was sufficiently similar to the reference data, the BeagleBone generated this information by illuminating the LED. The selected neural network model was implemented using the Tensorflaw library [19] as a sequential model, i.e. a model having one input tensor and one output tensor. The model consists of ten sequential layers. The first layer is a scaling layer, aimed at scaling an image with a specific geometry and color depth to a form acceptable for input neurons and scaling the digital image value of the image record from the 0–255 range to the 0–1 range. Subsequently, the following weave and decimation layers are repeated three times. The first is the basic type of layer found in so-called convolutional networks. The second one reduces the size of the processed image by selecting the pixel with the highest value from the 2×2 pixel frame. Finally, a flattening layer and two "normal" neural network layers are used. The last layer in the neural network model contains only 2 neurons, which are also the output of the entire network. One of the neurons informs about the presence of a drone, and the other about its absence. A detailed description of the layers is shown in Table 1. The parameters of the neural

Figure 5. Photo from acoustic measurements in the park

network model described in Table 1 were selected experimentally, based on the simulations and analysis of the results. The method of data processing using the selected neural network model was as follows:

- Rescaling - scales the input data by multiplying it by a scaling factor. In our case, an image was given as the input, so the input values were in the range 0–255. A scaling factor of 1/255 was adopted, thanks to which data with values within the range 0–1 was obtained, which accelerated the process of training the neural network.
- Conv2D – building a 3×3 kernel and then convolutional with the input.
- MaxPooling2D – reduction of input value dimensions
- Flatten - converting multidimensional input data to a one-dimensional vector.

The recorded drone acoustic signal paths were identified on an ongoing basis. The description covered the physical parameters of the drone, such as: flight direction, speed, maneuvers, distance from the microphones. A selection was made for the recorded samples of samples of various sound sources in the external environment and for the drone itself. Samples with the registration of the drone also concerned the cases of its start-up before take-off and after its landing. Finally, the obtained acoustic signal samples were classified into two subsets, i.e. including the recording of the drone's operation and the measurement of the acoustic background. The acquired data from the acoustic signal was divided into 1-second fragments. Data structured in this way constituted an input set to the neural network model and were processed in order to generate a spectrogram and save it as a file with the extension ".jpg". The obtained images of the drone presence spectrograms show a continuous distribution of characteristic frequencies in the studied period of time. On the other hand, spectrograms containing other sound sources show the frequency distribution corresponding to different fragments of time (Table 2). The audio sampling frequency was 48 kHz. To prevent aliasing from occurring, the signal was passed through a low-pass filter suppressing frequencies above 6.5 kHz and then decimated. The decimation was based on using every third sample from the signal. Samples recorded in the chamber and samples from the open space constituted one input set for the neural network. This approach allowed to assess whether there is a probability of a signal emitted by a passing drone in the tested acoustic signal.

Table 1. Layers of the neural network model

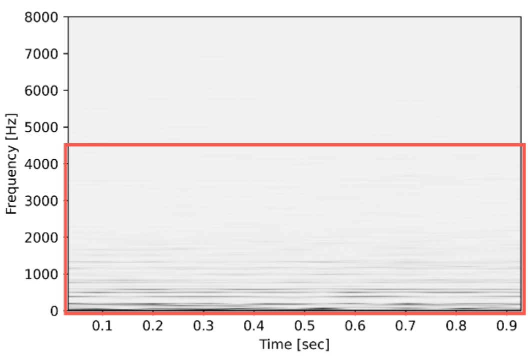
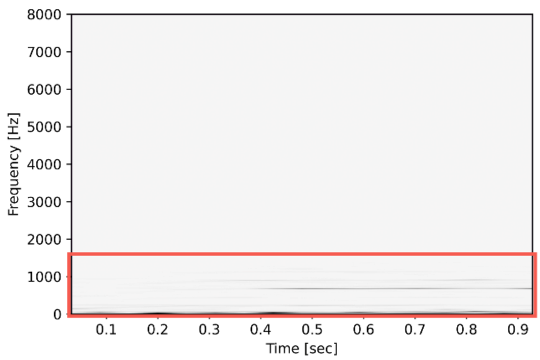
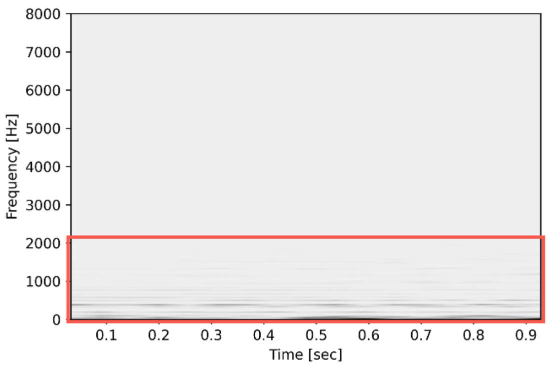
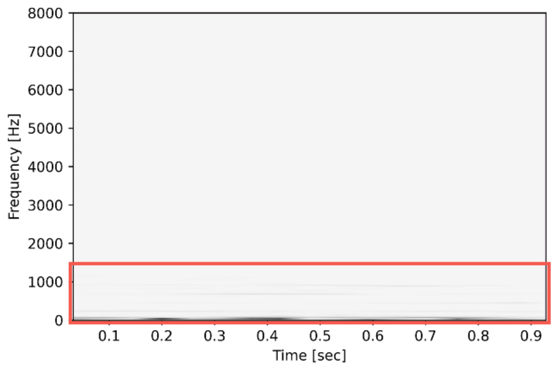
Model summary			
No.	Layer (type)	Output shape	Params
1	Rescaling	(None, 180, 180, 3)	0
2	Conv2D	(None, 178, 178, 32)	896
3	MaxPooling2D	(None, 89, 89, 32)	0
4	Conv2D	(None, 87, 87, 32)	9248
5	MaxPooling2D	(None, 43, 43, 32)	0
6	Conv2D	(None, 41, 41, 32)	9248
7	MaxPooling2D	(None, 20, 20, 32)	0
8	Flatten	(None, 12800)	0
9	Dense	(None, 128)	1638528
10	Dense	(None, 2)	258
Total params: 1,658,178 Trainable params: 1,658,178 Non-trainable params: 0			

In the research undertaken on drone acoustic signal classification, it was assumed as a trial that a shallow neural network would be suitable for pattern classification. The calculations were carried out in the Matlab 2021a environment using a neural network pattern recognition. A two-layer feed-forward network with hidden sigmoid neurons and softmax output neurons was used. The network is a two-layer anticipatory network with a sigmoidal transfer function in the hidden layer and a softmax transfer function in the output layer. The size of the hidden layer corresponds to the number of hidden neurons. The default layer size is 10. The number of output neurons is set to 2, which is equal to the number of classes specified in the response data. The network has two output neurons because there are two response values (classes) associated with each input vector. Each output neuron represents a class. When an input vector of the corresponding class is applied to the network, the corresponding neuron should generate a value of 1, and the other neurons should generate a value of 0. To prepare image collections from recorded acoustic signals, the tensor flow library was used [19]. Inputs for the network model were divided as follows:

- 70% for training,
- 15% to test whether the network generalizes and stop learning before over-fitting,
- 15% for independent testing of network generalization.

Training of the network continued until one of the stopping criteria was met. In this case, learning

Table 2. Selected spectograms containing the drone and other sound sources

The presence of a drone	Other sound sources
	
	

continued until the validation error increased sequentially for eight iterations ("Validation criterion met"). The computational accuracy in the eighth epoch was 97% with an error of 0.05 as a result of the applied neural network model (Fig. 7).

The signal evaluation consisted in processing the spectogram image by a previously prepared and trained neural classifier on the basis of reference training data. The classifier's output was a value from the set <0; 1> which determined the probability of the presence of a drone in the analyzed signal sample (Table 3).

The number of images of the set divided into two classes and used to train the network was 516. The adopted amount of data set for training was based on the recording time of the drone's acoustic signals. This time was 10 minutes and was the same for the implementation of measurements in the anechoic cabin and in the open space. The adopted sampling time of 1 second for the signals analyzed allowed the generation of such a training set. The obtained results from classifying the datasets at the 0.757 probability level (Table 3) for the drone presence signal in the park were considered satisfactory in this research. Attempts to increase the dataset for training the network and to modify the parameters of the neural network would help improve the obtained results.

DISCUSSION OF RESULTS AND CONCLUSIONS

The research on drone detection with the use of sound features was of a pilot nature. The recording of sound features with the use of spectograms seemed to be justified for the purposes of drone detection. The main advantage of a spectogram is the ability to record information in the time-frequency domain. As part of the research, the correctness of the drone detection in the external environment was obtained at the level of approximately 97% using the reference samples from the anechoic chamber using the neural network model. The obtained drone detection result in an urban environment was obtained on the basis of a set of pattern signals recorded in an anechoic chamber with the probability of its occurrence at the level of 0.997. A study provided by Casabianca and Zhang [20] evaluated the performance of ensemble deep learning models in acoustic identification of multi-rotor UAVs, which achieved

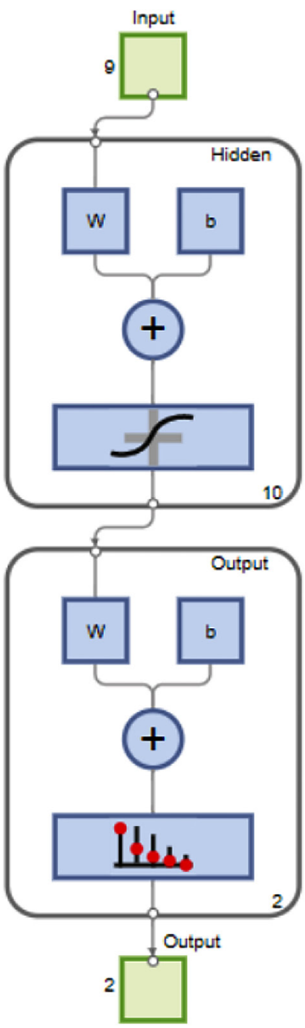


Figure 6. The neural network model used

classification accuracies of 94.7 and 91.0%, respectively, based on invisible, augmented and real test datasets. A study conducted on optimized single model architectures showed that the CNN and CRNN models outperformed the reconstructed RNN model with the highest performance. In this study, acoustic signals were transformed into mel spectrograms and CNNs were shown to be best suited for image processing tasks. The article [21] evaluates the illegal use of drones for malicious activities, proposing a novel approach that automates drone detection and identification processes using the drone's acoustic features with various deep learning algorithms. The experiments conducted in this work show that CNNs outperform both RNNs and CRNNs in detecting and identifying drones of known drone types.

The mentioned works split formatted audio files into multiple short (one-second) segments, specifying the intervals at which the audio clip will be segmented, allowing the deep learning algorithm to optimize model learning for real-time implementations for which the time required for detection and identification is critical. Based on the observations, it was concluded that one-second segmentation was sufficient. The results obtained in drone detection prestented in [22] paper contrast with those provided by Jeon at al. [23]. The authors noted that RNN obtained the best results compared to CNN, which does not confirm

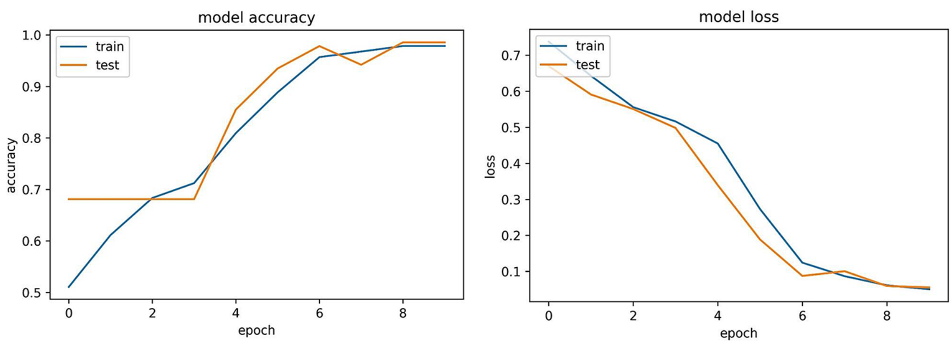


Figure 7. Accuracy characteristics of calculations for the adopted model

Table 3. Examples of the results of the probability of the drone occurrence in selected samples of acoustic signals (4 samples)

Name of sample	Type of sound source: drone/other source	The probability of the drone being present
Dron_280	Drone in an anechoic chamber	0.997
Traffic_jam_44	Sources of city traffic without the presence of a drone	0.449
Dron_park_78	A drone traveling in the park	0.757
Street_95	Other city sources without the presence of a drone	0.020

the results obtained in [22]. From the experiments conducted, it was found that CNN significantly outperformed RNN.

Due to the specificity of the research undertaken, the neural network model used: A two-layer feedforward network with hidden sigmoidal neurons and softmax output neurons should be treated as preliminary research. The authors of this article are aware of the accepted limitations of the study, such as the size of the input data sets, the selection of only the park area for model testing, or the model of the neural network itself.

The authors assume that further attempts will be made to obtain greater accuracy of the drone detection by using the optimization of the structure of the neural network model (e.g. changing the number of neurons). It should be noted that more neurons require more calculations and tend to over-fit the data. When the number of them is too large, it allows the network to solve more complex problems. More layers require more computation, but their use can result in the network solving complex problems more efficiently. Additionally, the use of descriptions of recorded samples of acoustic signals will allow for the creation of a more complex detection algorithm.

The use of drones in urban environments is becoming increasingly common. For this reason, there is a risk of drone collisions with other objects, terrorist threats, on the other hand, there is also a risk of privacy violations. While the obligation to register a drone equipped with a camera is regulated, there are no specific provisions regarding the violation of space on private land and the consequences of the resulting incidents.

The distance between the drone and the microphone is a limiting factor affecting detection performance. Therefore, in future work, it would be useful to investigate at what distances the drone will no longer be detected and how to optimize the placement of microphones , to ensure that the drone will always be detected within a certain radius. It is assumed that the commenced research will be continued in the field of analysis and evaluation:

- application of deep learning network models,
- the impact of the variability of the drone location parameters in relation to the microphone system,
- the influence of the variability of the arrangement of the microphone system on the obtained results,
- the importance of the diversity and development of urban infrastructure facilities on the obtained results of drone detection,
- detection of the drone in an environment strongly disturbed by other sound sources,
- the applicability of the recognized methods of processing the features of the acoustic and audio signals, as well as the accuracy of the results obtained.

Acknowledgment

The presented research results include the implementation of the project entitled “Development of a method of passive detection of drones with the use of acoustic signal processing”at the Silesian University of Technology in the PBL (Project Base Learning) formula.

REFERENCES

1. Birch G.C., Griffin J.C., M.K. Erdman: UAS detection, classification, and neutralization: Market survey, prepared by Sandia Nat’l. Labs, 2015.
2. Busset J. et al. Detection and tracking of drones using advanced acoustic cameras. Proc. SPIE Security Defence, vol. 9647, Toulouse, France, Oct. 2015, 1–8.
3. Benyamin M., Goldman G.H. Acoustic detection and tracking of a Class I UAS with a small tetrahedral microphone array. Army Research Lab. tech. rep. (ARLTR-7086), DTIC Doc., Sept. 2014.
4. Thibbotuwawa A., Nielsen P., Zbigniew B., Bocewicz G. Factors affecting energy consumption of unmanned aerial vehicles: an analysis of how energy consumption changes in relation to UAV routing. Advances in Intelligent Systems and Computing 2019; 853, 228-238, doi: 10.1007/978-3-319-99996-8_21.
5. Thibbotuwawa A., Bocewicz G., Nielsen P., Banaszak Z. UAV mission planning subject to weather forecast constraints. Advances in Intelligent Systems and Computing, 2020; 1004, 65-76, doi: 10.1007/978-3-030-23946-6_8.
6. Rabiner L.R., Schafer R.W. Digital processing of speech signals. Prentice Hall, 1978.
7. Shea K.O., Nash R.: An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458, 2015.
8. Matviykiv O., Klymkovych T., Bokla N., Lobur M., Melnyk M., Timofiejczuk A.: Simulation of acoustophoretic separation of microplastic particles in mkFluidic Lab-chip. 2020 IEEE XVITH International Conference on the Perspective Technologies

and Methods in Mems Design. Book Series: International Conference on Perspective Technologies and Methods in MEMS Design MEMSTECH, 123-126.

9. Shin J.-W., Lee Y.-J., Kim W.-H., Kwon Y., Yang H.- Y. Empirical study of drone sound detection in real-life environment with deep neural networks. In: Signal Processing Conference (EUSIPCO), 2017 25th European, pages 1858–1862. IEEE, 2017.
10. Jeon S., Shin J.-W., Lee Y.-J., Kim W.-H., Kwon Y., Yang. H.- Y. Empirical study of drone sound detection in reallife environment with deep neural networks. In: Signal Processing Conference (EUSIPCO), 2017 25th European, 1858–1862. IEEE, 2017.
11. Rutherford A. Introducing ANOVA and ANCOVA: A GLM Approach. SAGE Publications Ltd, London 2001.
12. Shannon C., Weaver W. The mathematical theory of communication. University Illinois Press, Urbana 1949.
13. Paszkowski W., Loska A. The use of neural network model in the assessment of annoyance of the industrial noise sources. Intelligent Systems in Production Engineering and Maintenance (ISPEM 2017). Book Series: Advances in Intelligent Systems and Computing, 637, 428-439. Springer Verlag 2018, doi: https://doi.org/10.1007/978-3-319-64465-3_41.
14. Loska A. Variant assessment of exploitation policy of selected companies managing technical network systems. Management Systems in Production Engineering 2015; 3(19), 179-188, doi: 10.12914/MSPE-13-03-2015.
15. Rozporządzenie Ministra Klimatu i Środowiska z dnia 7 września 2021 r. w sprawie wymagań w zakresie prowadzenia pomiarów wielkości emisji. Dziennik Ustaw Rzeczypospolitej Polskiej, poz. 1710.
16. PN-85/N-01333. Hałas. Dokładne metody określenia mocy akustycznej hałasu maszyn w komorze bezechowej i w otwartej przestrzeni.
17. PN-EN ISO 3746:2011. Akustyka. Wyznaczanie poziomów mocy akustycznej i poziomów energii akustycznej źródeł hałasu na podstawie pomiarów ciśnienia akustycznego. Metoda orientacyjna z zastosowaniem otaczającej powierzchni pomiarowej nad płaszczyzną odbijającą dźwięk.
18. Rozporządzenie Ministra Środowiska z dnia 30 października 2014 r. w sprawie wymagań w zakresie prowadzenia pomiarów wielkości emisji oraz pomiarów ilości pobieranej wody. Dz. U., poz. 1542.
19. https://www.tensorflow.org/tutorials/images/classification.
20. Casabianca P., Zhang Y.: Acoustic-based UAV detection using late fusion of deep neural networks. Drones 2021, 5(3), 54; https://doi.org/10.3390/drones5030054
21. Al-Emadi S., Al-Ali A. Audio-based drone detection and identification using deep learning techniques with dataset enhancement through generative adversarial networks. Sensors (Basel). 2021 Aug; 21(15): 4953. Doi: 10.3390/s21154953
22. Jeon S., Shin J., Lee Y., Kim W., Kwon Y., Yang H. Empirical study of drone sound detection in real-life environment with deep neural networks. Materiały z 25. Europejskiej Konferencji Przetwarzania Sygnałów 2017 (EUSIPCO); Kos, Grecja. 28.08 – 2.09.2017 r.