VPS VIBRATIONS IN PHYSICAL SYSTEMS

DOI: 10.21008/j.0860-6897.2024.2.09
Vibrations in Physical Systems, 2024, 35(2), 2024209

Residual convolutional neural network for continuous identification of aircraft noise

Michał BUKAŁA[1], Artur NOWOŚWIAT[1], Andrzej CHYLA[2]

[1] Silesian University of Technology
[2] SVANTEK Sp. z o.o., andrzej.chyla@gmail.com

Corresponding author: Michał BUKAŁA, email: michal.bukala@polsl.pl

Abstract To fully utilize the possibilities given by the aircraft noise monitoring system, it should quickly, automatically and accurately identify whether the limit-exceeding noise event is caused by the aircraft operation. Due to the often delayed access to airport operation logs, the system should operate with minimal or no non-acoustic data. The paper proposes the architecture of an aircraft noise detection method, meeting the above requirements and attempts to assess its effectiveness. Proposed approach involves using the residual convolutional neural network for solving the task. The network operates on 1/3 octave noise input data and determines the similarity of the input sound to the aircraft noise. The accuracy of the proposed method for a single data frame using real-life measurements exceeds 95% for a frame length of at least 30 seconds. Further work is progressing, focusing mainly on improving the quality of training data and refining the hyperparameters of the network.

Keywords: noise monitoring, aircraft noise, convolutional neural network, noise classification.

1. Introduction

Effective management of aircraft noise in the vicinity of airports and helipads is crucial for mitigating its impact on local communities and enhancing the overall aviation experience. Continuous aircraft noise monitoring systems play a pivotal role in this endeavour, as they provide essential data for optimizing various aspects of noise management, such as the utilization of approach and departure paths, distribution of aircraft types during the day, and adaptations to airport operating patterns and fleet compositions.

In recent years, advancements in machine learning and artificial intelligence have opened new avenues for automatic detection and identification of aircraft noise events. In this paper, we present a novel approach, the 'Residual Convolutional Neural Network for Continuous Identification of Aircraft Noise' (RCNN-CIAN), which aims to improve the accuracy and efficiency of identifying aircraft noise events in real-time.

Beyond the broader scope of noise management, the RCNN-CIAN system enables the timely indication of anomalous aircraft movements, offering valuable insights to airport authorities and local communities. By pinpointing such incidents promptly, appropriate actions can be taken to address any emerging concerns and maintain harmonious relations between airports and their neighbouring areas.

This paper outlines the architecture and design of RCNN-CIAN, its training process, and the evaluation results obtained from real-world aircraft noise data. The system's performance is analysed and compared to existing methods, demonstrating its efficacy in continuous aircraft noise identification.

By presenting this state-of-the-art convolutional neural network approach for aircraft noise detection, we aim to contribute significantly to the advancement of continuous noise monitoring systems, and ultimately facilitate more informed and data-driven decisions for noise management in aviation environments.

2. Methodology

Due to the increasing significance of aircraft noise management, numerous recent studies have been conducted to explore different approaches for the detection and classification of noise events. The current ISO standard [1] typically involves two stages: event extraction and the subsequent classification of the extracted noise events. Some authors have adopted spectral analysis and pattern recognition [2–4], while others have explored the application of neural network techniques [5]. The input data used for analysis may consist of various types of spectra, such as MFCC (Mel Frequency Cepstral Coefficients) or FFT (Fast Fourier

Transform), obtained either directly from the measurement equipment or derived from audio recordings of noise events in post-processing.

Despite varying methodologies and performance metrics, the reported efficiencies of these approaches are generally positive, surpassing simple, broadband classification [6]. However, it is essential to note that these methods often rely on noise events that have been previously extracted from the source signal using threshold overstepping analysis, which may introduce certain limitations and uncertainties.

A different approach, as proposed in studies involving quasi-continuous aircraft probability functions [7, 8], combines both event detection and classification into a single stage. This novel method is applied to one or multi-channel audio recordings, resulting in a significant data overhead. While promising, the computational requirements and data volume make it challenging to fully realize the potential of this approach.

In recent developments, artificial neural networks have been leveraged to achieve high classification accuracy without relying on sophisticated and data-expensive FFT or MFCC analyses [9]. Utilizing a convolutional neural network (CNN) with two convolutional layers, these study achieved a remarkable 97% classification accuracy by working with noise events of varying length. Subsequent tests on different training and test datasets [6] still demonstrated very promising performance, albeit slightly less accurate (90%). Remarkably, the focus was placed on the 1/3 octave spectra, simplifying the process and reducing computational complexity.

In the approach proposed here, a novel combination of the quasi-continuous aircraft probability function and a CNN working on the 1/3 spectral data, recorded by the measuring device at a 500 ms time resolution, is presented. The network architecture incorporates state-of-the-art techniques, such as introducing residual blocks and utilizing batch normalization, to fully harness the potential of this innovative approach.

By integrating the strengths of both the quasi-continuous aircraft probability function and the CNN, this research aims to advance the field of aircraft noise event detection and classification, offering a robust and efficient solution with the potential for real-time implementation and improved noise management strategies around airports and helipads.

2.1. Concept

Since the input data are in fact multiple time series a recurrent neural network architecture (RNN) using long short-term memory (LSTM) or gated recurrent unit (GRU) layers would seem ideal for the task. However, deeper analysis of the training data revealed numerous errors, originating from non-ideal human operators, who have been labelling the data based on a broadband sound levels aided with additional metadata (flight records and flight tracks). Therefore some rules to filter out the possibly outlying periods had to be devised, at the same time making time series analysis almost impossible. The decision to continue CNN research was backed up both by theoretical findings [10] and promising performance of such networks [9], tested in similar boundary conditions.

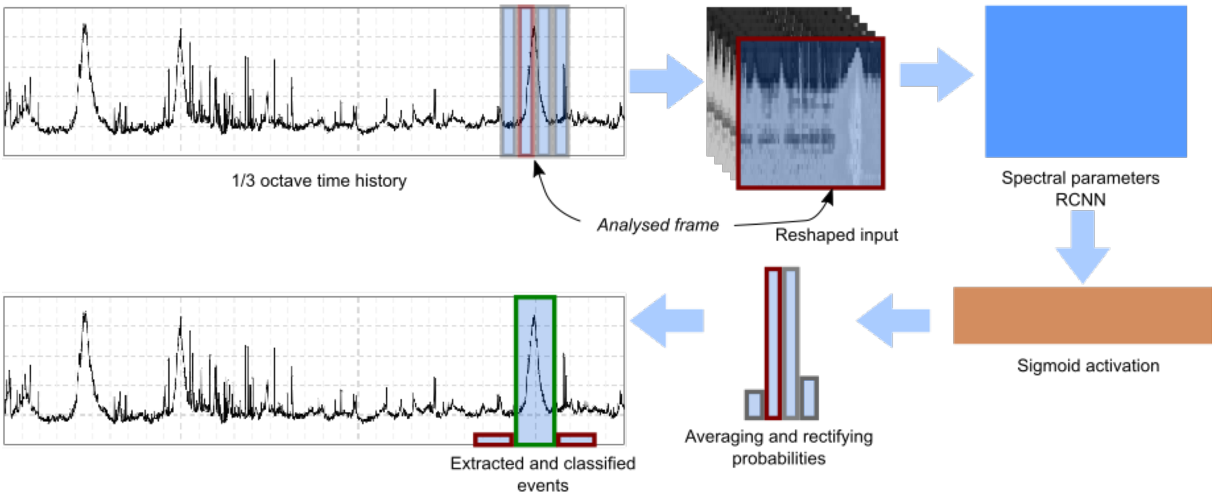

Figure 1. Concept of classification process.

The proposed concept for the detection and classification process (refer to Fig. 1) involves four stages of processing. The input data consists of time history information represented as 1/3 octave spectra. To facilitate the analysis, these spectra are divided into fixed-length, overlapping frames, forming T×S arrays. Here, T represents the frame length in samples, and S is equal to the number of 1/3 octave bands plus one A-weighted sound level sample.

These frames are then fed into a Convolutional Neural Network (CNN), which provides an uncalibrated probability score indicating whether each input frame contains an aircraft noise spectrum. Subsequently, a matrix of overlapping time-distributed probabilities is constructed and averaged.

To refine the results and determine the temporal boundaries of the aircraft noise events, a cutoff threshold is applied. This threshold converts the probability scores to rectified 0-1 values for each time step. The resulting values correspond to the presence or absence of an aircraft noise event at specific time intervals.

The proposed approach indeed highlights a crucial trade-off concerning the frame length (T) used in the method. This trade-off revolves around finding the optimal T value that balances two critical aspects: temporal resolution and single-frame classification accuracy.

On one hand, using a shorter frame length can enhance the temporal resolution of the method. This means that even short-duration aircraft noise events can be detected with greater precision. However, a shorter frame length results in fewer data points available for the neural classifier, potentially leading to reduced classification accuracy.

On the other hand, opting for a higher T value can provide more data for the neural classifier, which may lead to improved classification accuracy. However, this comes at the cost of having fewer frames that contain only aircraft noise. If the T value is too high, the network might struggle to identify instances where only aircraft noise is present, and this could impact the overall performance of the method.

The focus of this paper on researching and defining the optimal frame length is crucial in addressing this trade-off effectively. By exploring different frame lengths and evaluating their impact on temporal resolution and classification accuracy, the paper aims to find the sweet spot that offers the best compromise. This optimal frame length can then be used to achieve accurate detection and classification of aircraft noise events.

2.1. RCNN Architecture

The proposed network was primarily inspired by the ResNet architecture [11]. However, considering the limited size of the input data and the risk of overfitting with overly complex networks [10], the initial structure was simplified. The final architecture (see Fig. 2) comprised 7 processing stages, with 4 of them employing classic 2-dimensional convolutional layers. To address potential overfitting issues, dropout and batch normalization layers were incorporated [12, 13].

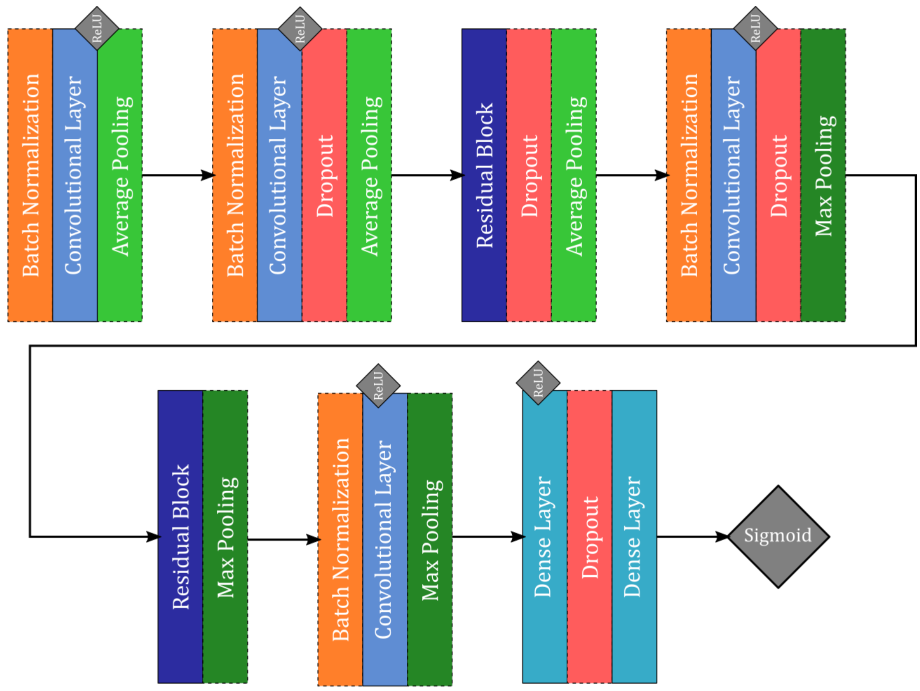

Figure 2. Proposed network architecture.

Stages 3 and 5 of the network were based on a residual block (see Fig. 3), which has been proven effective in deep learning tasks [14]. Finally, the last stage consisted of two dense layers with a sigmoid activation function to provide probability scores.

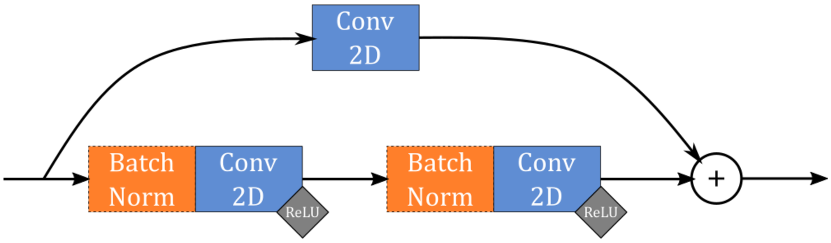

Figure 3. Residual block structure.

By streamlining the network architecture and adopting techniques to prevent overfitting, we aimed to strike a balance between model complexity and data limitations. The utilization of residual blocks further enhanced the network's ability to learn relevant features from the data. The network's output, in the form of probability scores, is essential for the continuous identification of aircraft noise events.

2.2. Training

The training data come from 6 measurement points spread around 2 civil airports in Poland, to maintain spectral variability, depending on the distance and angle of the measuring point from the runway threshold [15]. Data was gathered for 2 months in each point. Time history step was set to 500 ms. 1/3-octave spectrum was recorded. The data was manually labelled by the human operators, using additionally provided metadata (flight registers and ADSB flight tracks). Preprocessing involved dividing the continuous time histories into frames, each labelled as *containing only aircraft noise signal* (positive class) or *containing other types of noise* (negative class), as shown in Fig. 4.

Figure 4. Labelling convention.

Investigated frame lengths were: 90 s, 60 s, 45 s, 30 s, 15 s, 10 s and 5 s. Examples are shown in Figs. 5-6. Sound pressure level values were rescaled so that they fit in the ⟨0, 1⟩ range. At this stage of the research 100 000 frames for each frame-length were generated, 80% out of which was used as a training set, 10% as a validation set and 10% as a test set. Data was rebalanced, which means that frames count was equal for positive an negative class.

Preprocessing was implemented in Python on a local machine. The training process was performed on Google Colab platform using Adam Optimizer with learning rate of 0.001. Batch size was 128. Model performance for a single frame classification was assessed with F1 score.

a) b)

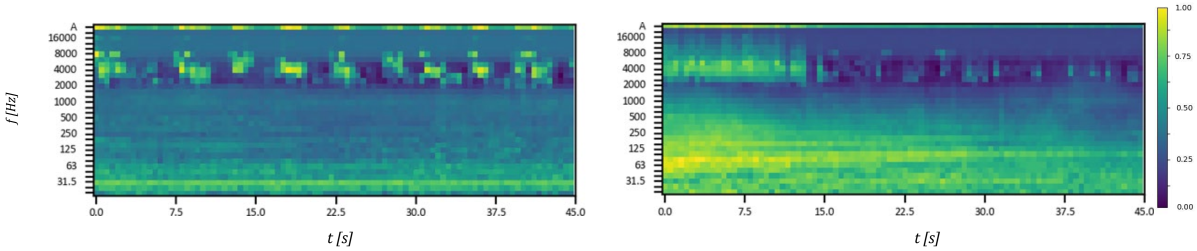

Figure 5. Visualisation of 45 s frames, with noise levels rescaled to 0 ÷ 1 range, containing: a) background noise, b) aircraft noise.

a) b)

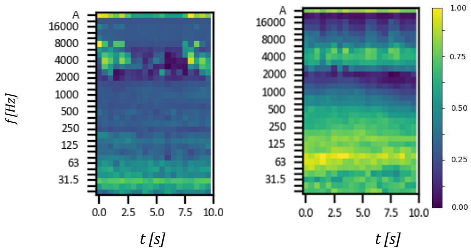

Figure 6. Visualisation of 10 s frames, with noise levels rescaled to 0 ÷ 1 range, containing: a) background noise, b) aircraft noise.

3. Results

Figure 7a illustrates the performance of the entire classification process, following the concept presented in Figure 1. The averaged probability scores exhibit limited variability for shorter frame lengths, indicating the network's difficulty in accurately identifying non-aircraft samples. Conversely, for the longest frame length, the probability scores never approach values close to 1, pointing to challenges in correctly identifying aircraft noise frames.

To determine the optimal cut-off thresholds for each tested frame length, real-life test data were utilized, and the results are displayed in Figure 8b. These thresholds are critical in refining the classification process and achieving accurate identification of aircraft noise events.

a) b)

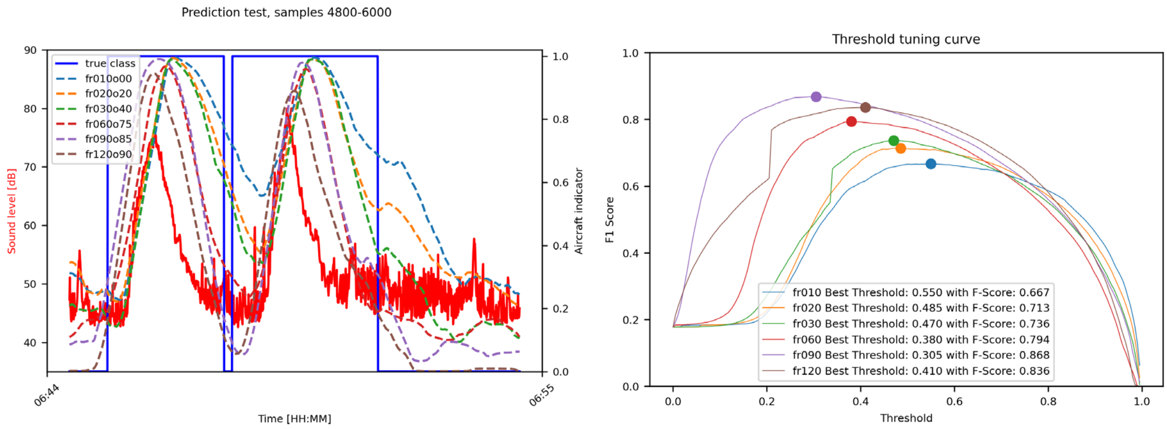

Figure 7. a) Prediction results, b) threshold tuning curves for prediction results.

Figure 8 shows the performance of whole classification process, according to concept presented in Fig. 1. Calculated F1 score was significantly different for various frame lengths, ranging from 0.667 for 5 s frame peaking with 0.868 for a 45 s frame which indicates 45 s frame length as optimal.

Final classification results are shown in Fig. 9. On the presented events it can be noticed that classification process based on 45 s frame is indeed closer to manually labelled events. 30 s frame classification however seems to perform better than a human operator, who in the presented example selected a noise-event time period with too high margin. Such mistakes do not influence the resulting value of L_{AE} but lead to decreased performance metrics of the neural network.

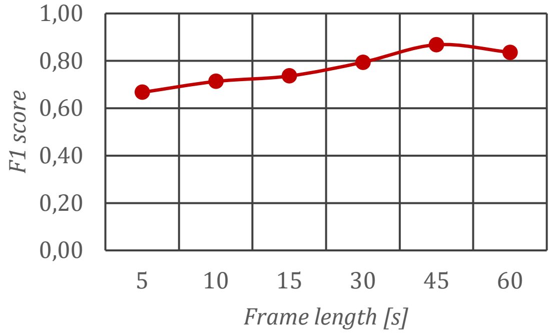

Figure 8. F1 score results for different frame lengths.

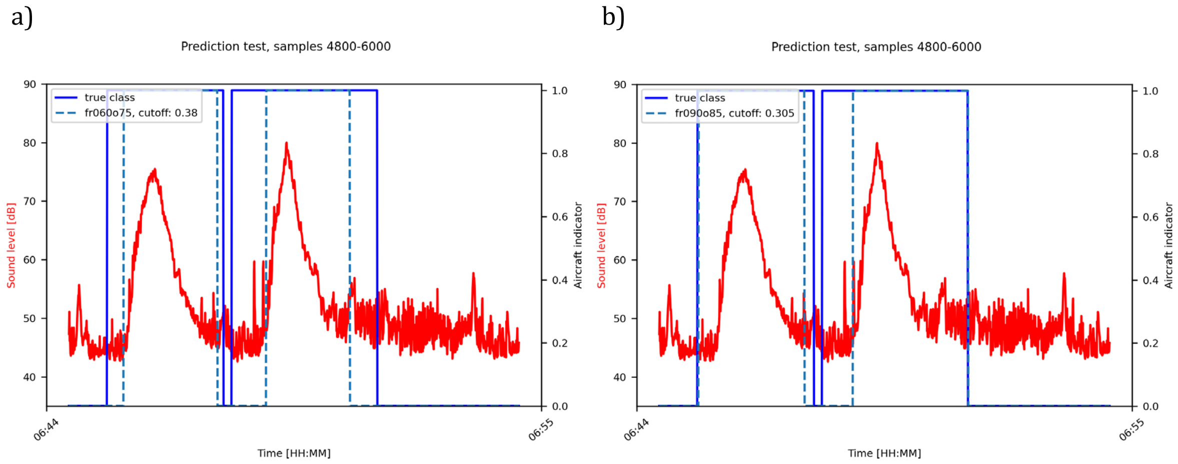

Figure 9. Aircraft noise events extracted by the network using input frame length: a) 30 s, b) 45 s.

4. Conclusions

This research introduces a novel approach to aircraft noise classification using a residual convolutional neural network (CNN). The proposed design encompasses the entire classification process, from data preprocessing to model output. The study demonstrates that the network architecture and the overall classification concept exhibit excellent performance in the testing environment.

One of the key findings of this research is the dependence of performance on the input frame length used in the classification process. The frame length, which refers to the duration of each input data segment fed into the CNN, plays a critical role in determining the accuracy of the classification results. The study identifies an optimal value for the classification frame length, striking a balance between temporal resolution and classification accuracy. By finding this optimal frame length, the research contributes to the effective detection and classification of aircraft noise events.

However, despite the encouraging results, the study also indicates the presence of input data errors, which can impact the classification performance. The researchers acknowledge the need for further research to address these errors and improve the accuracy of the classification process. Identifying and reducing input data errors are crucial steps in enhancing the reliability and robustness of the aircraft noise classification system.

The implications of this work are significant for the field of aircraft noise management. By leveraging a residual CNN and identifying the optimal frame length, the proposed design provides a valuable tool for real-time and continuous aircraft noise classification. The findings also underscore the importance of data quality and the potential challenges associated with input data errors. Addressing these challenges through further research will lead to more reliable and accurate noise classification, enabling better noise management strategies around airports and helipads.

In conclusion, this research presents a comprehensive approach to aircraft noise classification, showcasing the effectiveness of a residual CNN and the significance of selecting an appropriate frame length. The study's outcomes shed light on the interplay between temporal resolution and classification accuracy and emphasize the necessity of addressing input data errors. Overall, this work serves as a valuable contribution to the domain of aircraft noise management and lays the groundwork for future advancements in noise classification techniques.

Acknowledgments

This research was co-financed by the Ministry of Education and Science of Poland under the grant No RJO15/SDW/003_09.

Additional information

The authors declare: no competing financial interests and that all material taken from other sources (including their own published works) is clearly cited and that appropriate permits have been obtained.

References

1. ISO, International Standard ISO 20906:2009(E): Acoustics – Unattended monitoring of Aircraft sound in the vicinity of airports, Geneva, 2009
2. C. Rosin, B. Barbo; Aircraft Noise Monitoring: Noise Level Shape and Spectrum Pattern Recognition Applied to Aircraft Noise Detection; Proceedings of the Internoise 2010, Lisbon, Portugal, 13–16 June 2010
3. M. Márquez-Molina, L. Sánchez Fernández, S. Suárez-Guerra, L. Sánchez-Pérez; Aircraft take-off noises classification based on human auditory's matched features extraction; Applied Acoustics, 2014, 84, 83–90; DOI: 10.1016/j.apacoust.2013.12.003
4. L. Sanches Fernández, L. Sánchez-Pérez, J. Hernández; Monitoring System of Environment Noise and Pattern Recognition; International Journal of Energy and Environment, 2021, 15, 83–90; DOI: 10.46300/91012.2021.15.2
5. W. Wszołek, M. Kłaczyński; Recognition of aircraft noise in long-term environmental monitoring; Mechanics and Control, 2010, 29(4), 192–197
6. M. Bukała, A. Nowoświat, A. Chyla; Initial Assumptions for the System of Automatic Detection and Classification of Aircraft Noise Events.; Vibrations in Physical Systems, 2022, 33(2), 2022208; DOI:10.21008/J.0860-6897.2022.2.08
7. C. Asensio, M. Ruiz, and M. Recuero; Real-time aircraft noise likeness detector, Applied Acoustics, 2009, 71(6), 539–545; DOI: 10.1016/j.apacoust.2009.12.005
8. M. Genescà, J. Romeu, T. Pàmies, A. Sánchez; Real time aircraft fly-over noise discrimination; Journal of Sound and Vibration, 2009, 323, 112–29; DOI: 10.1016/j.jsv.2008.12.030
9. N. Heller, D. Anderson, M. Baker, B. Juffer, N. Papanikolopoulos; Convolutional Neural Networks for Aircraft Noise Monitoring, 2018, arXiv:1806.04779; DOI: 10.48550/arXiv.1806.04779
10. F. Chollet; Deep Learning with Python; Manning Publications Co., 2017
11. K. He, X. Zhang, S. Ren, J. Sun; Deep Residual Learning for Image Recognition, 2015, arXiv:1512.03385; DOI: 10.48550/arXiv.1512.03385
12. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov; Dropout: Simple way to prevent neural networks from overfitting; Journal of Machine Learning Research, 2014, 15(1), 1929–1958
13. S. Ioe, C. Szegedy; Batch normalization: Accelerating deep network training by reducing internal covariate shift, 2015, arXiv:1502.03167; DOI: 10.48550/arXiv.1502.03167
14. K. He, X. Zhang, S. Ren, J. Sun; Deep Residual Learning for Image Recognition, 2015, arXiv:1512.03385; DOI: 10.48550/arXiv.1512.03385
15. M. Bukała; Zastosowanie wskaźników oceny izolacyjności akustycznej ścian zewnętrznych budynków ze względu na hałas lotniczy; In: Współczesne zagadnienia z inżynierii lądowej – prace naukowe doktorantów (in Polish); I. Pokorska-Silva, K. Gromysz (Eds.); Wydawnictwo Politechniki Śląskiej, 2020