Modern programming technologies in the tasks of identification and classification of military aircraft using machine learning algorithms

Dmytro Uhryn[a], Artem Karachevstev[a], Taras Terletskyi[b], Oleh Kaidyk[b], Mariya Talakh[a], Viktor Ilin[a], Volodymyr Bogachuk[c], Oleksandr Kaduk[c], Zinagul Suranchiyeva[d], Zbigniew Omiotek[e]
[a]Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine; [b]Lutsk National Technical University, Lutsk, Ukraine; [c]Vinnytsia National Technical University, Vinnytsia, Ukraine; [d]Kazakh National Women's Teacher Training University, Almaty, Kazakhstan; [e]Lublin University of Technology, Lublin, Poland

ABSTRACT

This article addresses the development of an intelligent military aircraft identification system using artificial intelligence, machine learning, and deep self-learning technologies to enhance national security and military efficiency. The system aims to automatically and accurately recognize and classify aircraft in images, offering advantages over traditional methods such as higher productivity, speed, accuracy, and the elimination of human error. The importance of deep learning solutions for threat detection and operational efficiency is emphasized. Modern visual data-based object recognition methods and tools are analysed. The methodology includes collecting and preprocessing data, developing a high-precision recognition system based on Yolov8, annotating objects with Roboflow, and creating training, validation, and testing subsets in the yolo format. The paper details the dataset formation process and presents satisfactory results in fast recognition of military aircraft with high classification accuracy. A comparative analysis of Yolov8, R-CNN, and GPT-4 models shows Yolov8's superiority in prediction accuracy and performance. The article describes the model management system for adjusting hyperparameters, selecting object categories, and initiating the training and forecasting process. Testing results demonstrate Yolov8's optimality for military aircraft identification, achieving accurate target identification in complex situations using advanced deep learning algorithms.

Keywords: intelligent military aircraft identification, artificial intelligence, machine learning, deep learning, Yolov8, object recognition, national security, classification accuracy.

1. INTRODUCTION

Computer vision systems are now widely and successfully used across various industries, extracting information from digital images to perform actions and provide recommendations. These systems are prevalent in energy, utilities, manufacturing, automotive, retail, and healthcare, addressing problems such as obstacle avoidance in trajectory planning. Thanks to advances in neural network algorithms, the area of application is constantly growing[19,20]. One significant application of deep learning is in military aircraft identification. Traditional methods, often relying on human operators, are time-consuming, error-prone, and susceptible to manipulation. To overcome these challenges, an advanced, automated system leveraging artificial intelligence (AI), machine learning (ML), and deep learning (DL) technologies is essential. Such a system can automate the identification process, offering greater accuracy and speed, and significantly reducing human error. Algorithms trained on extensive data can recognize and classify military equipment with high precision, aiding in military planning and national security.

The goal is to develop a highly accurate and efficient object recognition system based on the YOLOv8 architecture. This system must process and analyse large volumes of visual data, including public datasets and high-quality satellite and aerial imagery. It should demonstrate superior performance compared to existing models like R-CNN and GPT-4 in terms of prediction accuracy and operational efficiency, even under challenging conditions such as poor lighting, adverse weather, and partial occlusions.

*d.ugryn@chnu.edu.ua

Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments 2024,
edited by Ryszard S. Romaniuk, Andrzej Smolarz, Waldemar Wójcik, Proc. of SPIE
Vol. 13400, 134000K · © 2024 SPIE · 0277-786X · doi: 10.1117/12.3054877

Furthermore, the system should include a model management framework for flexible adjustment of hyperparameters, selection of object categories, and seamless initiation of training or prediction processes using pre-trained models. This adaptability is crucial for meeting various operational requirements and ensuring continuous improvement. In summary, there is a pressing need for a state-of-the-art AI-based identification system to significantly enhance the accuracy and efficiency of military aircraft recognition. This advancement will bolster national security and operational effectiveness, highlighting the growing role of AI and ML in defence and beyond.

2. MATERIAL AND METHODS

In recent years, technologies related to computer vision have attracted a lot of attention from researchers. Traditional object detection methods based on manually created features and surface architecture are already outdated and do not meet modern requirements. In digital processing, an image can be represented as a two-dimensional function $f(x, y)$, where x and y are coordinates on the plane. The value of the function f at any point given by a pair of coordinates (x, y), corresponds to the color at that point. For efficient digital image processing, it is appropriate to use the image shape. In this case, each pixel of the image (raster element) is characterized by the following parameters:

$$f(x, y) = [x, y][R, G, B], \tag{1}$$

where $[x, y]$ are the pixel coordinates, and $[R, G, B]$ are the color components in the additive color model. Thus, a digital image is defined by a raster matrix, where each element corresponds to a pixel with specified coordinates and colour values:

$$f(x, y) = \begin{bmatrix} [x_{11}, y_{11}][R_{11}, G_{11}, B_{11}] \dots [x_{1n}, y_{1n}][R_{1n}, G_{1n}, B_{1n}] \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ [x_{m1}, y_{m1}][R_{m1}, G_{m1}, B_{m1}] \dots [x_{mn}, y_{mn}][R_{mn}, G_{mn}, B_{mn}] \end{bmatrix}. \tag{2}$$

Richter et al. provides a detailed overview of deep learning object detection frameworks that address various problems such as occlusion, obstacles, and low resolution with different levels of modification based on R-CNN (convolutional neural networks based on regions)[1]. Through numerous changes, the object detection algorithm has improved in terms of speed and accuracy. Thanks to the tireless efforts of numerous researchers, deep learning algorithms are rapidly evolving and delivering higher performance in object detection. Applications such as pedestrian recognition, medical imaging, robotics, driverless cars, and facial recognition can help save labour in many industries. Another papers provide a basic overview of object detection methods, including two classes of object detectors[2-5]. The two-stage detector uses R-CNN, Fast R-CNN, and even Faster R-CNN algorithms, while the one-stage detector uses YOLO v1, v2, v3, and SSD.

Two-stage detectors focus on accuracy, while the main advantage of single-stage detectors is speed. The authors present also an improved version of YOLO, called YOLO v3-Tiny, and compares it with the previous version[6,7]. Compared to the classical approach, modern data processing methods use advanced artificial intelligence and machine learning algorithms such as Yolov8, R-CNN, and GPT-4. Yolov8, the latest version of the famous "you look only once" algorithm, effectively improves the speed and accuracy of real-time object detection. R-CNN (Region-based Convolutional Neural Networks) significantly improves recognition accuracy by analysing specific areas of the image. GPT-4 not only efficiently processes and generates text, but also analyses images, expanding the capabilities of multimodal data analysis. This feature makes GPT-4 a particularly valuable and powerful tool for analysis and identification.

In order for a neural network to perform object detection and recognition tasks, it must be trained first. A training algorithm or learning is a method or mathematical model for adjusting the parameters of a neural network by modelling the input environment. This is achieved by adjusting the weights and thresholds (shifts) of the network. From the algorithmic point of view, training a neural network consists in selecting 1 specific model from the set of allowed models. Among the most well-known training algorithms are the gradient descent method and its variants, the Broyden-Fletcher-Goldfarb-Channo family algorithm, the Levenberg-Marquardt algorithm, etc. He also optimized gradient descent algorithms such as adaptive motion estimation (Adam), adaptive sub gradient methods for online learning and stochastic optimization (Adgrad), and Nesterov acceleration[8,-10].

It is an algorithm that uses a trained neural network model to produce predictions or conclusions for a test data set. This stage requires significantly fewer computing resources than training and is error-free, so it can be run on a graphics processing unit (GPU). Inference algorithms are usually divided into three stages: matching input data, selecting relevant features, and performing calculations to obtain predictions. The type of neural network training algorithm depends on the nature of the input and output data. Currently, there are 3 main types of training: reinforcement learning, supervised learning (with teacher), and unsupervised learning (without teacher)[11-14].

Supervised learning is a process in which pre-labelled data sets are fed to the input of a neural network. Each instance of data is fed to the network's input and processed, and the result is compared to the target value, which is the desired output of the network. Then, according to certain rules, the error is calculated, which leads to the adjustment of the weighting factor in the network. The rules for changing weights are determined by the chosen learning algorithm. The vector of training samples is sequentially input into the input data and the error is calculated, and for each vector there is also a sub-branch of supervised learning - partial supervised learning (semi-supervised learning), and both labelled and unlabeled data are used for training until the total error is equal to the entire training sample reaches an acceptably low level. In this case, a small, labelled sample is used to initially train the model, after which it is used to label the remaining unlabelled data. This approach can improve the accuracy of object search and recognition when the input sample is small and the object class is as close as possible[15-18]. The supervised learning approach is widely used in the task of searching and recognizing objects in images. This is due to the fact that a large, labelled dataset (an image with a specified object) is available for such tasks, and a fully connected layer of the neural network is used. The fully connected layer allows the neural network to detect complex relationships between the pixels of the input image and the categories of recognized objects. The availability of large, labelled datasets provides a sufficient number of examples for efficient training with network weight correction by minimizing the error in the training set. Figure 1 shows the basic modern architecture of neural networks used to solve object search and recognition tasks.

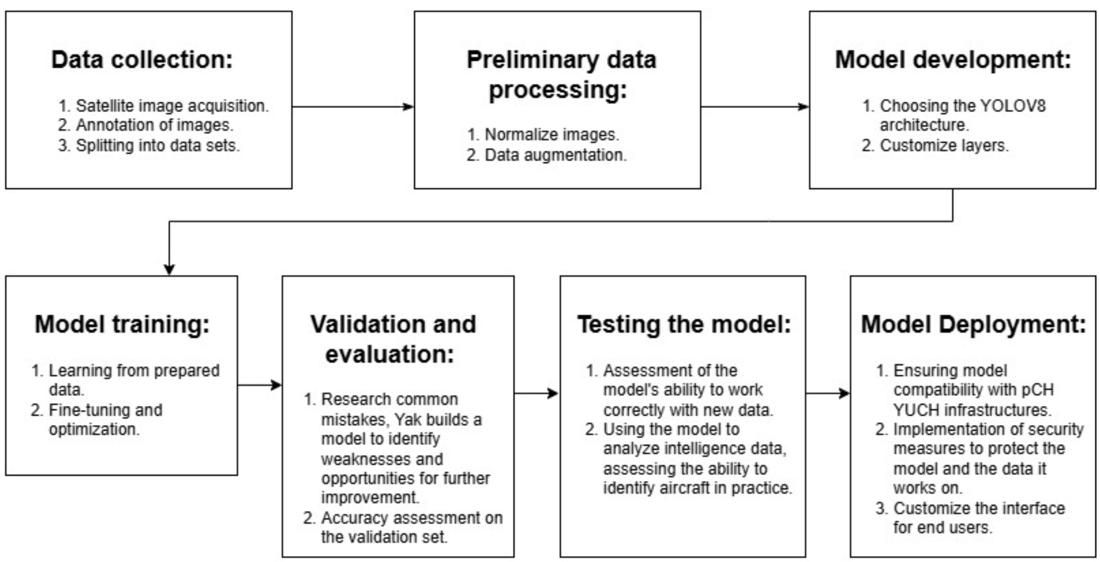


Figure 1. Types of deep neural networks used to recognize and detect objects in an image.

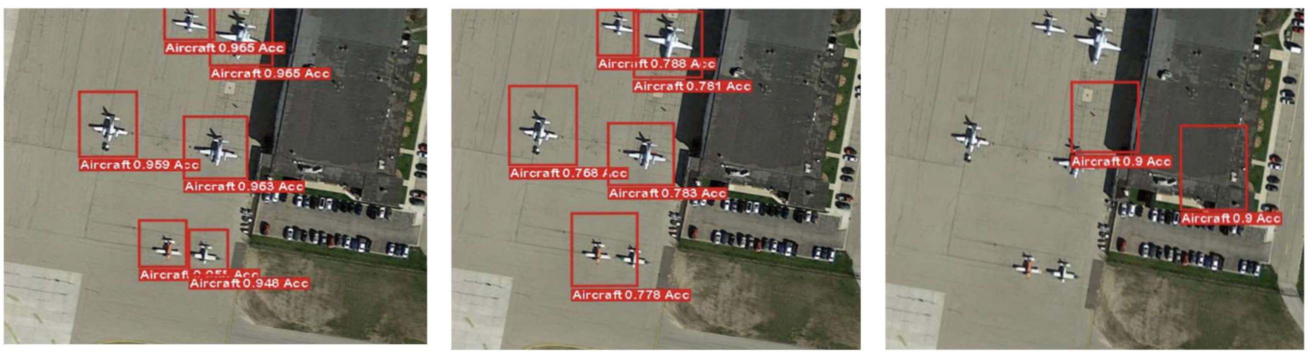
3. RESULTS AND DISCUSSION

To ensure the diversity and volume of data, the use of publicly available datasets and open sources was considered. An important selection criterion was the quality and relevance of the images to the task of identifying military aircraft. The basis for the first dataset was the Military Aircraft Detection Dataset published by Kaggle. This dataset contains a wide range of images of military aircraft and provides various data for training models. After selecting the main source, the next step was the process of collecting and adapting the data to the requirements of the specific task. Additional analysis of the images was performed to assess their suitability and compliance with established standards. The next step was to select the most representative images and annotate them for use in model training. The images used can be acquired using satellite technology or from aircraft that guarantee high resolution.

The analysed data is carefully prepared before being used in the model training process. This preparation included checking the quality of the images, labelling, and annotations (Fig. 2), as well as splitting the entire dataset into training sets, validating the confidence, and testing the model. The Roboflow resources were used to automatically generate annotations and corresponding text files that display the coordinates of objects in the images in the Kaggle dataset. This tool helps to create a pair of text files corresponding to the selected coordinates in the image. The last step in the process of preparing a dataset is to create the dataset itself, dataset.yaml. This file contains the relative path to the image and the corresponding text file (label) with annotations. It also contains information about the number and names of classes that the network should identify during training. These class names correspond to the objects detected in the input image.

(a) (b)

Figure 2. Illustration of image annotation and highlighting techniques: a – annotation of the desired element in the image, b – highlighting airplane polygons in the image.

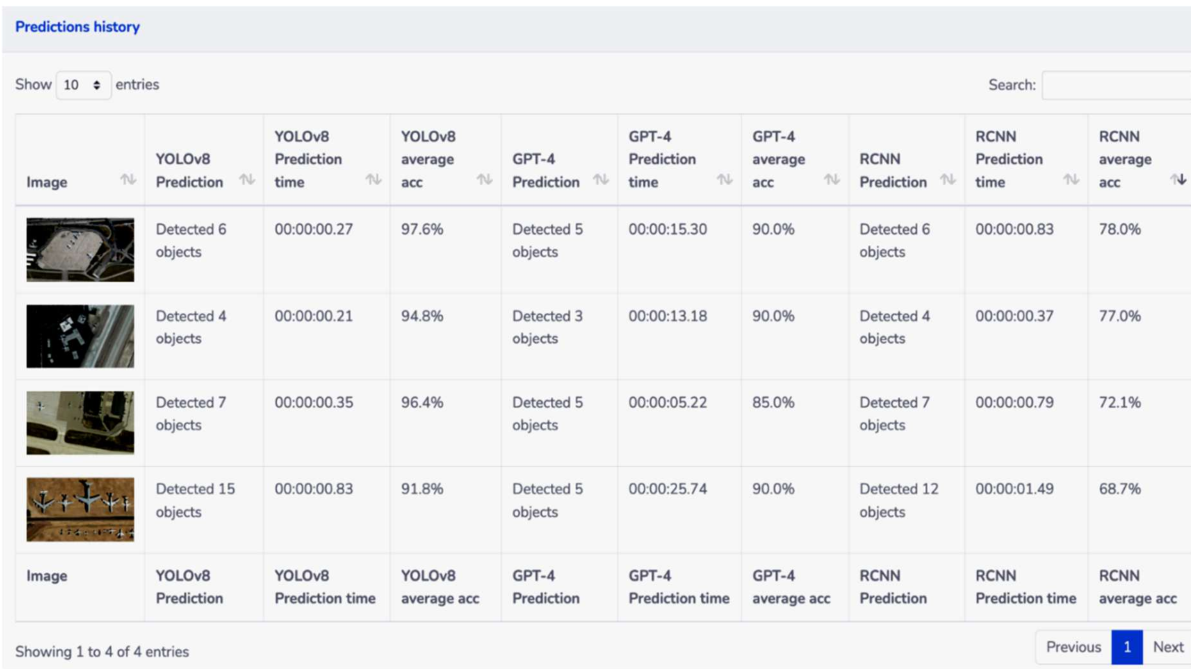


(a) (b) (c)

Figure 3. The result of recognition and classification of the trained: a – YOLOv8 model, b – R-CNN model, c – GPT-4 model.

Based on the analysis of Fig. 3, we can conclude that the model successfully managed to recognize all the aircraft in the image. The system identified 7 objects with an average accuracy of 95.5% (Fig. 3a). From the analysis of Fig. 3 shows that the R-CNN model did not successfully solve its task by failing to recognize one aircraft. The prediction accuracy of the R-CNN model is significantly lower than that of YOLOv8 and amounts to 77.2% (Fig. 3c). During the research, the next step is to activate the object recognition function in the selected image (Fig. 3). The numbers next to the object names indicate the probability of correct recognition and correct classification of these objects. At the same time, other images used to test this model also showed positive recognition results and were saved in the model's multimedia catalogue. Based on the developed model of the military air object identification and classification system, a table of predictions for the YOLOv8, R-CNN, and GPT-4 models was created using machine learning (Fig. 4). From this table, it can be concluded that YOLOv8 provides the most accurate predictions, and in addition, the model has a significant speed advantage over other models.

Model management allowed us to create and train our own model for classifying military aircraft. The system allows users to improve the model training hyperparameters at their own discretion, as well as the data category programs that are used. After that, the model training process is started by sending a POST request to the server. The model is trained in the background, allowing users to continue working with the system. Before starting to train the model, you can manually configure the hyperparameters and select the most appropriate category of the dataset using a POST request to the server. The model training process takes place in the background. A check is installed on the training page to prevent possible problems with the training model. In addition to training the model, you can use a pre-trained model to make predictions. Based on the analysis of the classification results presented above, we can conclude that the pre-trained YOLOv8 model successfully performs the task. This model is specially designed to accurately detect and classify objects in different environments, and it has demonstrated high accuracy in identifying target objects. One of the key factors in YOLOv8's performance is its ability to reliably recognize objects even in conditions that can be difficult for other models, such as low light, different weather conditions, or partial obscuration of an object. This is achieved thanks to the deep learning and advanced algorithms behind YOLOv8.



Predictions history

Show 10 entries Search:

Image	YOLOv8 Prediction	YOLOv8 Prediction time	YOLOv8 average acc	GPT-4 Prediction	GPT-4 Prediction time	GPT-4 average acc	RCNN Prediction	RCNN Prediction time	RCNN average acc
	Detected 6 objects	00:00:00.27	97.6%	Detected 5 objects	00:00:15.30	90.0%	Detected 6 objects	00:00:00.83	78.0%
	Detected 4 objects	00:00:00.21	94.8%	Detected 3 objects	00:00:13.18	90.0%	Detected 4 objects	00:00:00.37	77.0%
	Detected 7 objects	00:00:00.35	96.4%	Detected 5 objects	00:00:05.22	85.0%	Detected 7 objects	00:00:00.79	72.1%
	Detected 15 objects	00:00:00.83	91.8%	Detected 5 objects	00:00:25.74	90.0%	Detected 12 objects	00:00:01.49	68.7%
Image	YOLOv8 Prediction	YOLOv8 Prediction time	YOLOv8 average acc	GPT-4 Prediction	GPT-4 Prediction time	GPT-4 average acc	RCNN Prediction	RCNN Prediction time	RCNN average acc

Showing 1 to 4 of 4 entries Previous 1 Next

Figure 4. Table of forecast history for YOLOv8, R-CNN and GPT-4 models.

4. CONCLUSIONS

The development of AI-driven identification systems for military aircraft is crucial for national security and operational efficiency. These systems leverage artificial intelligence, machine learning, and deep learning to accurately and rapidly recognize and classify aircraft from large datasets, including images. Unlike traditional methods, AI-based systems enhance performance, speed, and accuracy while minimizing human error, making them essential for effective military operations and threat detection. This paper reviews modern object recognition techniques, focusing on a deep learning model based on YOLOv8. The study outlines a process for data collection and preprocessing, including using public datasets and high-quality satellite imagery. The dataset was meticulously prepared with quality checks, annotations, and splitting into training, validation, and test sets. The YOLOv8 model, tested with this dataset, achieved high accuracy and fast recognition of military aircraft. Comparative analysis of YOLOv8, R-CNN, and GPT-4 shows YOLOv8 offers superior prediction accuracy and performance. A model management system was developed to adjust hyperparameters and manage object categories. YOLOv8 demonstrated exceptional efficiency in detecting and classifying objects, even under challenging conditions, proving it to be the most effective model for military aircraft identification.

REFERENCES

[1] Richter, C. P., Matic, A. I., Wells, J. D., Jansen, E. D. and Walsh, J. T., "Neural stimulation with optical radiation," Laser Photonics Rev. 5(1), 68-80 (2011).

[2] Zhao, Z., Zheng, P., Xu, S. and Wu, S., "Object detection with deep learning: A review," Computer Vision and Pattern Recognition 1, 1-21 (2019).

[3] Adarsh, P., Rathi, P. and Kumar, M., "YOLO v3-Tiny: Object Detection and Recognition using one stage improved model," 6th International Conference on Advanced Computing and Communication Systems (ICACCS), 687-694 (2020).

[4] Wu, M., Yue, H. and Wang, J., "Object detection based on RGC mask R-CNN. IET Image Processing," 14, 1502-1508 (2020).

[5] Aguilera, L. and Jacobson, D., "Deep Learning CNN Implementation on Packed Malware for Cloud Cross Domain Solution Filters," In 2022 International Conference on Data Science and Its Applications (ICoDSA). IEEE, 192-197 (2022).

[6] Bonaccorso, G., "Mastering Machine Learning Algorithms: Expert techniques for implementing popular machine learning algorithms, fine-tuning your models, and understanding how they work," Packt Publishing Ltd, (2020).

[7] Masooma, M., "ANN vs CNN vs RNN: Neural Networks Guide," https://levity.ai/blog/neural-networks-cnn-ann-rnn, (November 2022).

[8] Mayank, M., "Convolutional Neural Networks, Explained," https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939, (August 2020).

[9] Bochkovskiy, A., Redmon, J., Sinigardi, S., Hager, T., Jaled, J., et al., "AlexeyAB/darknet:YOLOv4 (version yolov4). Zenodo," (2021), https://doi.org/10.5281/zenodo.562267.

[10] Elgendy, M., "Deep learning for vision systems (1st ed.)," (2020).

[11] Wang, C. Y., Liao, H. Y. M., Wu, Y. H., Chen, P. Y., Hsieh, J. W. and Yeh, I. H., "CSPNet: A new backbone that can enhance learning capability of CNN," In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, 390-391 (2020).

[12] Kin-Yiu, W., "Implementation of Scaled-YOLOv4 using PyTorch framework (v1.0.0)," Zenodo, (2021), https://doi.org/10.5281/zenodo.5534091.

[13] Jocher, G., Stoken, A., Chaurasia, A., Borovec, J., Xie, T., et al., "Ultralytics/YOLOv5:v6.0–YOLOv5n 'Nano' models. Roboflow integration. TensorFlow export, OpenCV DNN support (v6.0)," Zenodo, (2021), https://doi.org/10.5281/zenodo.5563715.

[14] Yilmaz, A., Demircali, A. A., Kocaman, S. and Uvet, H., "Comparison of Deep Learning and Traditional Machine Learning Techniques for Classification of Pap Smear Images," (2020).

[15] Nguyen, N. V., Rigaud, C. and Burie, J. C., "Semi-supervised Object Detection with Unlabelled Data. In VISIGRAPP (5: VISAPP)," 289-296 (2019).

[16] Kim, J. A., Sung, J. Y. and Park, S. H., "Comparison of Faster-RCNN, YOLO, and SSD for real-time vehicle type recognition, "In 2020 IEEE international conference on consumer electronics-Asia (ICCE-Asia)," 1-4) (2020).

[17] Kuznetsova, A., Rom, H., Alldrin, N., Uijlings, J., Krasin, I., Pont-Tuset, J. and Ferrari, V., "The open images dataset v4: Unified image classification, object detection, and visual relationship detection at scale," International Journal of Computer Vision 128(7), 1956-1981 (2020).

[18] Nepal, U. and Eslamiat, H., "Comparing YOLOv3, YOLOv4 and YOLOv5 for autonomous landing spot detection in faulty UAVs," Sensors 22(2), 464 (2022).

[19] Dzierżak, R. and Omiotek, Z., "Application of Deep Convolutional Neural Networks in the Diagnosis of Osteoporosis," Sensors 22, 8189 (2022).

[20] Grądz, Ż., Wójcik, W., Gromaszek, K., Kotyra, A., Smailova, S., Iskakova, A., Yeraliyeva, B., Kumargazhanova, S. and Imanbek, B., "Application of Fuzzy Neural Networks in Combustion Process Diagnostics," Energies 17(1), 212 (2024).