АҚПАРАТТЫҚ-КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР / ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ / INFORMATION AND COMMUNICATION TECHNOLOGIES

DOI 10.54596/2958-0048-2025-2-175-183
UDK 629.735
IRSTI 28.23.37

DEVELOPMENT OF A CLASSIFICATION MODEL FOR UAVS AND BIRDS BASED ON THE YOLOV9 NEURAL NETWORK TO IMPROVE ANTI-DRONE SYSTEMS

Adilbekov A.[1*], Semenyuk V.[1], Proselkov A.[1]
[1]Manash Kozybayev North Kazakhstan University NPLC, Petropavlovsk, Kazakhstan*
[*]Corresponding author: alibekadilbek93@mail.ru

Abstract

The article presents the materials of the development of a model for classification and recognition of UAVs and birds based on the neural network of the YOLOv9 architecture in the optoelectronic channels of Anti-drone systems. To train the neural network, a dataset was prepared in the form of annotated images of UAVs and birds. The total number, taking into account augmentation, was 5265 images. The authors implemented training, verification and testing of neural networks in the Windows 11 operating system, in the Python 3.10.8 runtime environment and the Pycharm 2024 development environment. The training process was carried out on the basis of the AD103 graphics processor of the NVIDIA GeForce RTX 4080 video card with support for CUDA Toolkit 12.1. As a result of training the neural network, the following metrics were obtained: mAP50-95: 0.59; mAP50: 0.95; Recall: 0.89; Precision: 0.95. According to these indicators, the trained model outperforms the UAV and bird recognition and classification models trained on the basis of YOLOv2, YOLOv4, YOLOv5, YOLOv7 and YOLOX. The inference results on two videos with DJI Inspire 2 and DJI Mini 3 UAV flights showed FPS values of 131 and 119, respectively. It was found that, due to the obtained accuracy and FPS metrics, the trained YOLOv9 model can be used as a module for recognizing and classifying UAVs and birds in real time in the optoelectronic surveillance channels of Anti-drone systems.

Keywords: anti-drone, sensor fusion, deep learning, drones, YOLO, neural networks.

АНТИДРОН ЖҮЙЕЛЕРДІ ЖЕТІЛДІРУ ҮШІН YOLOV9 НЕЙРОНДЫҚ ЖЕЛІСІ НЕГІЗІНДЕ ҰҰА МЕН ҚҰСТАРДЫ ЖІКТЕУ МОДЕЛІН ӘЗІРЛЕУ

Адильбеков А.Е.[1*], Семенюк В.В.[1], Проселков А.В.[1]
[1]«Манаш Қозыбаев атындағы Солтүстік Қазақстан университеті» КеАҚ Петропавл, Қазақстан*
[*]Хат-хабар үшін автор: alibekadilbek93@mail.ru

Аңдатпа

Мақалада Антидрон жүйелерінің оптикалық-электрондық арналарында YOLOv9 архитектурасының нейрондық желісі негізінде ҰАА мен құстарды жіктеу және тану моделін әзірлеу бойынша материалдар ұсынылған. Нейрондық желіні оқыту үшін ҰАА мен құстардың аннотацияланған суреттері түрінде деректер жинағы дайындалды. Толықтыруларды қосқанда жалпы саны 5265 суретті құрады. Нейрондық желілерді оқыту, тексеру және тестілеу Windows 11 операциялық жүйесінде, Python 3.10.8 атқару ортасында және Pycharm 2024 әзірлеу ортасында. CUDA Toolkit 12.1 қолдауы бар NVIDIA GeForce RTX 4080 бейне картасының AD103 графикалық процессоры негізінде оқыту процессі жүзеге асырылды. Нейрондық желіні оқыту нәтижесінде келесі метрикалар алынды: mAP50-95: 0,59; mAP50: 0,95; Recall: 0,89; Precision: 0,95. Бұл көрсеткіштерде оқытылған модель YOLOv2, YOLOv4, YOLOv5, YOLOv7 және YOLOX нұсқаларында оқытылған ҰАА мен құстарды тану және жіктеу үлгілерінен асып түседі. DJI Inspire 2 және DJI Mini 3 ҰАА екі бейнесін шығару нәтижелері сәйкесінше 131 және 119 FPS

мәндерін көрсетті. Алынған дәлдік пен FPS көрсеткіштерінің арқасында үйретілген YOLOv9 моделін антидрон жүйелерінің оптикалық-электрондық бақылау арналарында нақты уақыт режимінде ҰАА мен құстарды тану және жіктеу модулі ретінде пайдалануға болатыны анықталды.
Кілт сөздер: антидрон, датчиктерді бірігу, терең оқыту, дрондар, YOLO, нейрондық желілер.

РАЗРАБОТКА МОДЕЛИ КЛАССИФИКАЦИИ БПЛА И ПТИЦ НА ОСНОВЕ НЕЙРОСЕТИ YOLOV9 ДЛЯ СОВЕРШЕНСТВОВАНИЯ АНТИДРОН СИСТЕМ

Адильбеков А.Е.[1*], Семенюк В.В.[1], Проселков А.В.[1]
[1]НАО «Северо-Казахстанский университет имени Манаша Козыбаева»*
Петропавловск, Казахстан
[*]Автор для корреспонденции: alibekadilbek93@mail.ru

Аннотация

В статье представлены материалы разработки модели классификации и распознавания БПЛА и птиц на основе нейронной сети архитектуры YOLOv9 в оптико-электронных каналах систем Антидрон. Для обучения нейронной сети был подготовлен набор данных в виде аннотированных изображений БПЛА и птиц. Общее количество, с учетом дополнения, составило 5265 изображений. Обучение, верификация и тестирование нейронных сетей осуществлялись в операционной системе Windows 11, в среде исполнения Python 3.10.8 и среде разработки Pycharm 2024. Процесс обучения осуществлялся на базе графического процессора AD103 видеокарты NVIDIA GeForce RTX 4080 с поддержкой CUDA Toolkit 12.1. В результате обучения нейронной сети были получены следующие метрики: mAP50-95: 0,59; mAP50: 0,95; Recall: 0,89; Precision: 0,95. По этим показателям обученная модель превосходит модели распознавания и классификации БПЛА и птиц, обученные на основе YOLOv2, YOLOv4, YOLOv5, YOLOv7 и YOLOX. Результаты вывода на двух видео с полетами БПЛА DJI Inspire 2 и DJI Mini 3 показали значения FPS 131 и 119 соответственно. Было установлено, что благодаря полученным показателям точности и FPS обученная модель YOLOv9 может быть использована в качестве модуля для распознавания и классификации БПЛА и птиц в реальном времени в оптико-электронных каналах наблюдения систем Антидрон.
Ключевые слова: антидрон, слияние датчиков, глубокое обучение, дроны, YOLO, нейронные сети.

Introduction

Currently, Anti-drone systems are actively used to solve problems of detection, classification and neutralization of UAVs (drones). Such a need is due to the growing number of incidents of using these devices for criminal purposes. Examples include violation of airport airspace, espionage, mass attacks on critical facilities for military purposes, delivery of prohibited items, and organization of failures in security systems. In this regard, the development of new methods for combating UAVs and the improvement of existing technologies for detection, classification and elimination of UAVs is relevant. As a result of the literature review of the Scopus and Web of Science databases, optoelectronic, acoustic, radio frequency, radar and combined (Sensor Fusion) methods are used to detect and classify UAVs, represented by the corresponding software and hardware solutions. Optoelectronic systems use cameras [1], laser sensors [2], thermal imagers [3] to accurately detect and track UAVs, but their effectiveness may be reduced in low visibility and lighting conditions. Radar systems [4, 5], on the other hand, operate based on radio waves and are capable of detecting objects in all weather conditions and at any time of day, although their accuracy in recognizing small objects may be limited. Radio frequency systems (RF based) [6, 7] detect UAV control and communication signals, which allows them to be detected even in the absence of visual contact, but they rely on the presence of radio signals and may encounter interference. Acoustic systems [8, 9] use microphones to detect sounds emitted by UAVs, which makes them useful in low visibility conditions, but their range and sensitivity may be limited by environmental noise. To improve detection accuracy and reliability, sensor fusion technology [10–12] is often used, which combines the data stream from different sensors. This allows for higher accuracy and reliability, although integration and data processing may be complex. This technology is implemented in modern Anti-drone systems, such as Elbit Systems ReDrone [13], DedroneRapidResponse [14] and others.

Accurate recognition and classification of UAVs relative to other objects is provided by the software component of the Anti-drone systems - artificial intelligence, which is represented by machine learning (ML) and deep leaning (DL) algorithms. In the optoelectronic channels of the Sensor Fusion systems, which are considered indispensable due to the accuracy of providing visual data, computer vision algorithms of the YOLO architecture are implemented. This algorithm, along with Faster R-CNN, SSD, RetinaNet and EfficientDet, is used as a visual detector of objects in real time. Due to key features such as single-stage processing, dividing images into a grid, joint prediction of different classes, high accuracy and speed, YOLO is more effective in solving problems of recognition and classification of objects, including UAVs. The authors [15–18] used various YOLO models to train neural networks on user datasets in the form of images of UAVs of different types, birds, etc. In [15], the authors trained the YOLOv4 model to recognize UAVs and birds, achieving the following average accuracy rates: mAP50 – 74.36%; precision – 0.95; Recall – 0.68; F1 – 0.79.

When tested on videos of two types of UAVs, DJI Phantom III and DJI Mavic Pro, the trained model achieved 20.5 and 19 FPS (frames per second), respectively, on inference. In [16], the authors used an earlier YOLOv2 model and achieved an mAP50 of 74.97%. The YOLOv5 model from [17] outperformed the previous model [16] by 15.4%. Higher-performance models of the YOLO architecture, such as YOLOX, YOLOv7, YOLOv8, are studied in [18]. YOLOv8 is a more advanced version of the previous models, thanks to new features and improvements implemented by the developers of Ultralytics. The new backbone network, anchorless detection head, and loss function contributed to high-quality training of the model with an mAP50 of 95.3%. The accuracy of UAV recognition and classification, which are characterized by the mAP50, mAP50-95, Precision and Recall metrics, are limited by the loss of information in successive layers of deep neural networks. This problem can be solved by implementing programmable gradient information (PGI) and the architecture of efficient layer aggregation network (GELAN).

In order to improve the model for recognizing and classifying UAVs in optoelectronic detection channels of Anti-drone systems by increasing the accuracy indicators, the following objectives must be completed within the framework of this study:

- Prepare a dataset of UAV and bird images for training the experimental YOLOv9 neural network model;
- Train the YOLOv9 neural network model to determine the accuracy indicators;
- Test the model on inference to determine FPS.

The results obtained allow us to draw a conclusion about the effectiveness of using the YOLOv9 neural network model for recognizing and classifying UAVs and birds.

Research methods

The training of the neural network model for UAV and bird classification will be based on the pre-trained YOLOv9 algorithm. This algorithm overcomes the shortcomings of methods for overcoming information loss, such as reversible architectures, masking modeling, and the concept of deep supervision, by implementing PGI. PGI (Fig 1) is based on gradient generation using an auxiliary reversible branch, which allows avoiding loss at semantic levels without additional computational costs.

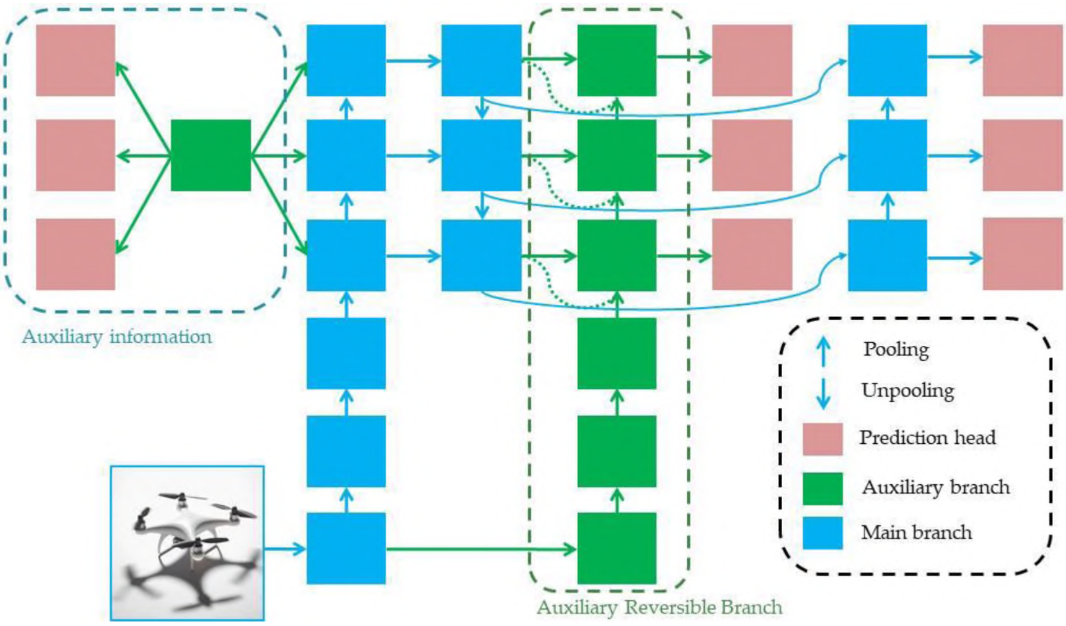


Figure 1. PGI Architecture diagram

The main structural components of PGI, in addition to the above-mentioned auxiliary reversible branch, are the main branch and multi-level auxiliary information. The main branch is used to organize logical inference, and multi-level auxiliary information solves the problems of error accumulation due to deep observation, which is essential for training lightweight models of the YOLOv9 architecture. The presence of GELAN in the YOLOv9 neural network ensures the creation of multi-scale feature maps for class prediction, optimization of parameters, computational complexity, accuracy and inference speed. This advantage is due to the combination of two neural network architectures CSPNet and ELAN, which provide gradient path planning. This solution allows users to update computing units for any logical output devices without significant performance losses.

To train the YOLOv9 model on custom datasets, the following steps must be completed:

- Prepare a dataset of images for two classes (UAVs and birds);
- Annotate classification objects in each image;
- Data augmentation;
- Distribute the dataset into three parts for training, validation, and testing;
- Export data for training in a special design environment;
- Select hyperparameters and start training.

To successfully train the neural network model, it is necessary to prepare a dataset for two classes of objects that will meet the criteria of data diversity (a variety of UAV and bird images representing different lighting conditions, weather conditions, angles and backgrounds), class balance, high image quality, image normalization to one size (e.g. 416 x 416; 640 x 640). Images for a custom dataset can be found in open sources such as Roboflow, Kaggle, Ultralytics, GitHub, and also use custom data (photos, videos) of UAV and bird flights. Annotation is a markup of images with special bounded boxes. This frame is defined by the coordinates of the upper left corner and lower right corner or the center coordinates and dimensions (width and height). For each image, an annotation file is created that contains information about the objects in the image. The format of the annotation file can be different, but usually YOLO uses a text format, where the class of the object and the coordinates of the bounding box are indicated for each object. For YOLOv9, information on annotated objects is stored in a separate ".txt" file. Roboflow.com is used as a service for loading, storing, processing and annotating (Fig 2) class objects. Augmentation technology is used to increase the variability and quantity of data. This technology is a technique for artificially increasing the size and diversity of a data set by applying various transformations to the original data. This technique is widely used in machine learning, especially in computer vision, to improve the quality of models and prevent overfitting.

a b

Figure 2. Annotating objects in Roboflow: a) UAVs; b) birds

In the case of images, augmentation involves applying various transformations. Geometrical transformations include rotating an image by a random angle, changing its scale, shifting it horizontally or vertically, flipping it horizontally or vertically, as well as bending and distorting it. Color transformations involve adjusting the brightness, changing the contrast, saturation, and color hues of an image. Also, various types of noise can be added to an image, blurring can be applied, or the image can be randomly cropped and resized back to its original proportions. In addition to these methods, there are other techniques such as excluding random square areas from an image (cutout) and mixing two images and their labels to create a new sample (mixup). Using data augmentation helps models become more robust to various types of distortions, which ultimately improves their ability to generalize to new data and leads to better performance on real data. Using this technology and the Roboflow interface, the Grayscale and Blur augmentation methods were applied. In addition to training, the model undergoes validation and testing stages, in connection with which the data set is distributed in percentage terms of 82/13/6 (training, validation and testing, respectively). A special annotation format YOLOv9 was selected to export the dataset to the neural network training project.

The training, verification, and testing of neural networks were implemented in the Windows 11 operating system, in the Python 3.10.8 runtime, and the Pycharm 2024 development environment. The training process was carried out on the AD103 GPU of the NVIDIA GeForce RTX 4080 video card with support for CUDA Toolkit 12.1. The program code was written and edited using the Ultralytics YOLOv8 framework, which contains YOLOv9 neural network models pre-trained on the COCO dataset. Among the five YOLOv9 models (YOLOv9t, YOLOv9s, YOLOv9m, YOLOv9c, YOLOv9e), YOLOv9c was selected for the experiment. This model was pre-trained on the COCO dataset with the following parameters: mAP50-95 - 0.53; mAP50 - 0.702; params (number of parameters) – 25.5 m (millions); FLOPs – 102.8. The following hyperparameters are set to train the neural network on the custom dataset: number of epochs: 100; batch size: 16; learning rate: 0.001; momentum: 0.9; weight decay: 0.0005 and image size: 640. The trained YOLOv9c model has a "best.pt" file size of 88 MB. The neural network was tested on inference using two test videos with DJI Inspire 2 and DJI Mini 2 UAV flights.

Results

1. Results of preparing a dataset for training the YOLOv9c neural network

Fig. 3 shows the interface of the prepared dataset in the Roboflow.com service.

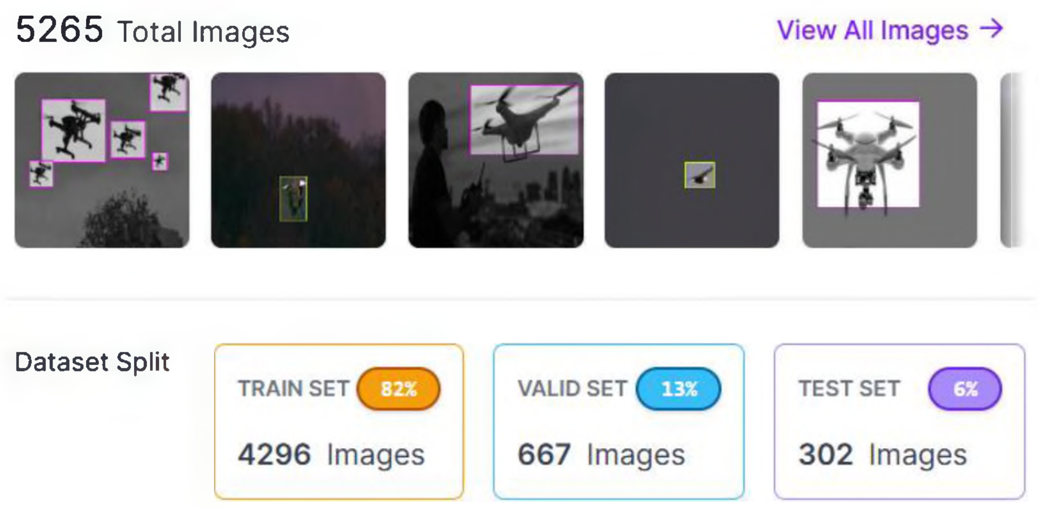


Figure 3. UAV and bird dataset prepared by Roboflow.com

The prepared dataset was used to train the neural network of the YOLOv9c architecture.

2. Results of training the YOLOv9c neural network on a custom dataset

Fig. 4 a-d shows the metrics of the training results of the YOLOv9c neural network model.

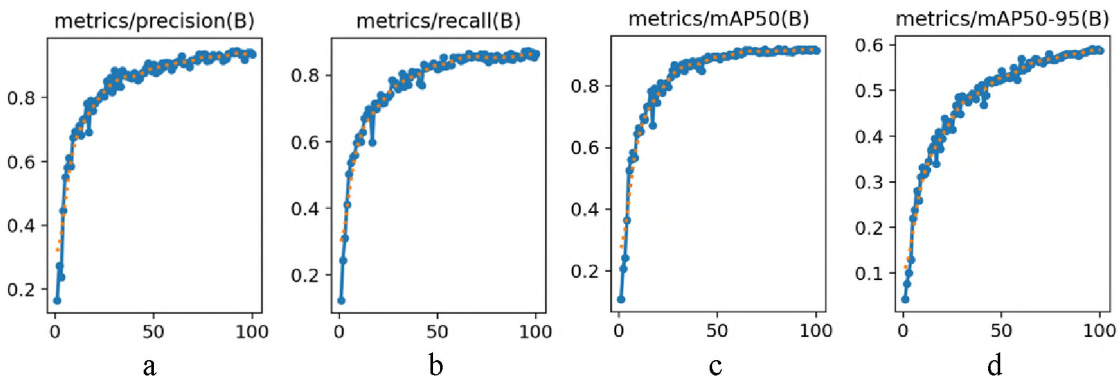
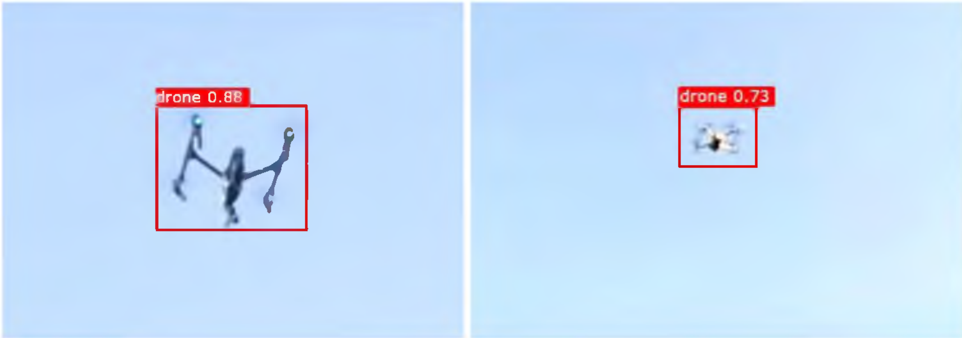


Figure 4. Metrics of the results of training the YOLOv9c neural network for 100 epochs (Ox-axis): a – Precision; b – Recall; c – mAP50; d – mAP50-95

The trained model was tested on inference using two videos to determine the average Latency (or FPS) value.

3. Inference test results

Fig. 5 a, b show frames of inference of the trained YOLOv9c neural network model on two videos with the flight of DJI Inspire 2 and DJI Mini 3, respectively.



a b

Figure 5. Frames from the inference of the trained YOLOv9c model: a) frame from the video of the DJI Inspire 2 flight; b) frame from the flight of the DJI Mini 3

The inference results are used to estimate the average FPS presented in Section 4 Discussion.

Discussion

To train the YOLOv9c neural network, a dataset of 5265 annotated images (Fig 3) was prepared in the Roboflow.com service. Of this set, 4296 images (82%) are intended for training, 667 images (13%) for validation, 302 files (6%) for testing.

As a result of training, the following maximum metric values were obtained (Fig 4 a-d):

– mAP50-95: 0.59;
– mAP50: 0.95;
– Recall: 0.89;
– Precision: 0.95.

The obtained mAP50 values exceed the accuracy of the trained model from [15, 16] by 20%; [17] by 4.6%. The YOLOv9s model corresponds to the trained YOLOv8 neural network presented in [18] in this indicator. Thus, the YOLOv9 architecture neural network models, along with YOLOv8, are able to improve the accuracy of recognition and classification of UAVs and birds through optoelectronic surveillance cameras. As a result of inference testing, the average Latency value for the first video was 7.6 ms and 8.4 ms for the second video. These indicators depend on the GPU characteristics and for less productive computing devices, it is recommended to use lighter versions such as YOLOv9t, YOLOv9s and YOLOv9m. The obtained Latency values prove the efficiency of using YOLOv9 as a basic software component for recognizing and classifying UAVs in optoelectronic surveillance channels of Anti-drone systems.

Conclusion

1) As part of the study of the possibility of using the YOLOv9 neural network to recognize and classify UAVs and birds using the Roboflow.com service, a dataset of 5265 annotated images was prepared.

2) Based on the AD103 graphics processor of the NVIDIA GeForce RTX 4080 video card with support for CUDA Toolkit 12.1, the YOLOv9c neural network was trained and metrics were obtained demonstrating a relatively high accuracy of UAV and bird classification in comparison with previous models of the YOLO architecture.

3) As a result of testing the trained YOLOv9c neural network on inference, average FPS values of 131 and 119 were obtained, which, along with high accuracy, proves the possibility of using this model as a module for recognizing and classifying UAVs and birds in real time in the optoelectronic surveillance channels of Anti-drone systems.

4) The results of this study will be useful to developers of Anti-drone systems.

Conflict of interest

The authors declare that they have no conflict of interest in relation to this research, whether financial, personal, authorship or otherwise, that could affect the research and its results presented in this paper.

Financing

This research is funded by the Committee for Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AR19679009).

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technologies when creating the current work.

References:

1. Mahdavi F., Rajabi R. (2020). Drone Detection Using Convolutional Neural Networks, 2020 6th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), Mashhad, Iran, 2020, pp. 1-5, DOI: 10.1109/ICSPIS51611.2020.9349620.
2. Hammer M., Borgmann B., Hebel M., Arens M. (2020). Image-based classification of small flying objects detected in LiDAR point clouds. Proceedings of SPIE - The International Society for Optical Engineering, 11410, 1141002. DOI: 10.1117/12.2557246.
3. Mebtouche N.E.-D., Baha N. (2022). Robust UAV detection based on saliency cues and magnified features on thermal images. Multimedia Tools and Applications. 82(13), pp. 20039-20058. DOI: 10.1007/s11042-022-14271-3.
4. Beasley P., Ritchie M., Griffiths H., Miceli W., Inggs M., Lewis S., Kahn B. (2020). Multistatic Radar Measurements of Uavs at X-Band and L-Band. IEEE National Radar Conference - Proceedings, 2020-September, art. no. 9266444. DOI: 10.1109/RadarConf2043947.2020.9266444.
5. Teo M.I., Seow C.K., Wen K. (2021). 5G Radar and Wi-Fi Based Machine Learning on Drone Detection and Localization. 2021 IEEE 6th International Conference on Computer and Communication Systems (ICCCS), Chengdu, China, 2021, pp. 875-880, DOI: 10.1109/ICCCS52626.2021.9449224.
6. Flak P., Czyba R. (2023). RF Drone Detection System Based on a Distributed Sensor Grid With Remote Hardware-Accelerated Signal Processing. IEEE Access. PP. 1-1. DOI: 10.1109/ACCESS.2023.3340133.
7. Yang S., Luo Y., Miao W., Ge C., Sun W., Luo C. (2021). RF Signal-Based UAV Detection and Mode Classification: A Joint Feature Engineering Generator and Multi-Channel Deep Neural Network Approach. Entropy. 23. 1678. DOI: 10.3390/e23121678.
8. Al-Emadi S., Al-Ali A., Al-Ali A. (2021) Audio-based drone detection and identification using deep learning techniques with dataset enhancement through generative adversarial networks. Sensors 2021, 21, 4953. DOI: 10.3390/s21154953.
9. Salman S., Mir J., Farooq M.T., Malik A.N., Haleemdeen R. (2021) Machine learning inspired efficient audio drone detection using acoustic features. In Proceedings of the 2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST), Islamabad, Pakistan, 12–16 January 2021; pp. 335–339. DOI: 10.1109/IBCAST51254.2021.9393232.
10. Svanström F., Alonso-Fernandez F., Englund C. (2021). A Dataset for Multi-Sensor Drone Detection. Data in Brief. 39. DOI: 10.1016/j.dib.2021.107521.
11. Semenyuk V., Kurmashev I., Lupidi A., Cantelli-Forti A. (2023). Developing the GoogleNet neural network for the detection and recognition of unmanned aerial vehicles in the Data Fusion System. Eastern-European Journal of Enterprise Technologies, 2(9-122), pp. 16–25. DOI: 10.15587/1729-4061.2023.276175.
12. Jajaga E., Rushiti V., Ramadani B., Pavleski D., Cantelli-Forti A., Stojkovska B., Petrovska O. (2022). An Image-Based Classification Module for Data Fusion Anti-drone System. Lecture Notes in

Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 13374 LNCS, pp. 422–433 DOI: 10.1007/978-3-031-13324-4_36.
13. Elbit Systems (2024). ReDrone. [online] Elbit Systems. Available at: https://elbitsystems.com/product/redrone/ [Accessed 22 June, 2024].
14. Dedrone (2024). DedroneRapidResponse: Multi-Layered Mobile Drone Detection Unit. [online] Dedrone. Available at: https://www.dedrone.com/solutions/dedrone-rapid-response [Accessed 22 June, 2024].
15. Singha S., Aydin B. (2021). Automated Drone Detection Using YOLO v4. Drones. September 2021. DOI: 10.3390/drones5030095.
16. Seidaliyeva U., Alduraibi M., Ilipbayeva L., Almagambetov A. (2020). Detection of Loaded and Unloaded UAV Using Deep Neural Network. In Proceedings of the 2020 Fourth IEEE International Conference on Ro-botic Computing (IRC), Taichung, Taiwan, 9–11 November 2020; pp. 490–494. DOI: 10.1109/IRC.2020.00093.
17. Aydin B., Singha S. (2023). Drone Detection Using YOLO v5. Eng. 4. DOI: 10.3390/eng4010025.
18. Zhai X., Huang Z., Li T., Liu H., Wang S. (2023). YOLO-Drone: An Optimized YOLOv8 Network for Tiny UAV Object Detection. Electronics. 12. 3664. DOI: 10.3390/electronics12173664.

Information about the authors:

A. Adilbekov – corresponding author, doctoral student, master of technical sciences, senior lecturer of the department of Energetic and Radioelectronics, Manash Kozybayev North Kazakhstan University NPLC, Petropavlovsk, Kazakhstan; e-mail: alibekadilbek93@mail.ru;
V. Semenyuk – master of technical sciences, senior lecturer of the Project office, Manash Kozybayev North Kazakhstan University NPLC, Petropavlovsk, Kazakhstan, e-mail: Evdimid@mail.ru;
A. Proselkov – master of technical sciences, specialist, Office of reception and recruitment, Manash Kozybayev North Kazakhstan University NPLC, Petropavlovsk, Kazakhstan; e-mail: proselkov96@gmail.com.