UDC 004.032.26.
MRSTI 28.23.15.

https://doi.org/10.55452/1998-6688-2025-22-4-219-226

[1]Semenyuk V.V.,
M.Tech.Sc., ORCID ID: 0000-0002-8580-7326,
e-mail: vvsemenyuk@ku.edu.kz
[1]*Kurmashev I.G.,
PhD, ORCID ID: 0000-0001-9872-7483,
*e-mail: ikurmashev@ku.edu.kz;
[2]Serbin V.V.,
PhD, Associate Professor, ORCID ID: 0000-0002-5807-3873,
e-mail: v.serbin@satbayev.university
[1]Kurmasheva L.B.,
MSc, ORCID ID: 0000-0001-5392-5873,
e-mail: lbkurmasheva@ku.edu.kz
[2]Moldagulova A.N.,
PhD, Professor, ORCID ID: 0000-0002-1596-561X,
e-mail: a.moldagulova@satbayev.university
[1]M. Kozybaev North Kazakhstan University, Petropavlovsk, Kazakhstan
[2]Satbayev University, Almaty, Kazakhstan.

DEVELOPMENT OF A REAL-TIME UAV RECOGNITION MODEL BASED ON YOLOV10 NEURAL NETWORK

Abstract

The paper deals with the development of a model for real-time recognition and classification of UAVs and birds based on the training of the YOLOv10 neural network. The research area is considered relevant in connection with the problems of UAV detection in the context of security, given their growing use in various fields. A dataset consisting of 6,255 images collected from proprietary archives and public resources is trained to train the model. The process of data annotation, augmentation and distribution was implemented using Roboflow.com service. The model was trained on NVIDIA GeForce RTX 4080 GPU using Ultralytics framework. Test results showed high recognition accuracy with mAP50 and mAP50-95 metrics exceeding previous versions of YOLO. The model demonstrates the ability for efficient object segmentation and tracking, which makes it promising for optoelectronic surveillance applications. The results of the study can be useful for developers of UAV and bird detection and classification systems, as well as for improving safety in various fields.

Keywords: neural networks, classification, recognition, optoelectronic surveillance channels, UAVs, YOLO, convolutional neural networks.

Introduction

Real-time recognition of unmanned aerial vehicles (UAVs) through surveillance cameras using neural networks and visual detectors is an urgent area for scientific research. This is driven by the need for public and state security, as UAVs are now being actively used for illegal activities including espionage, smuggling delivery, and combat raids. UAV recognition detection systems are designed to detect such threats in a timely manner and take appropriate measures to neutralize them. They also play an important role in protecting critical infrastructure such as airports, power plants and government buildings where there is a high risk of potential attacks.

The scientific novelty of the study lies in applying the YOLOv10 architecture – featuring dual label assignment, consistent matching metrics, and NMS-free design – to the task of real-time UAV and bird recognition. The work demonstrates improved mAP50-95 accuracy over previous YOLO versions on a custom dataset and shows the model's capability for efficient segmentation and tracking.

The originality stems from adapting and validating the newly introduced YOLOv10 mechanisms specifically for small, fast-moving airborne objects in optoelectronic surveillance systems.

Real-time drone recognition helps enforce flight regulations and control restricted areas. This is especially important when there are strict requirements and restrictions on the use of UAVs in certain areas, such as military bases and private territories. Modern neural networks and machine learning algorithms have the ability to process large amounts of data, which provides high accuracy recognition and opens new opportunities for developing effective monitoring systems. Such systems can be integrated with existing video surveillance systems, making them more affordable and efficient. The algorithms used as software modules for UAV detection and recognition in optoelectronic surveillance channels are YOLO, Faster R-CNN, and SSD. The authors found that single-stage YOLO detectors and two-stage Faster R-CNN algorithms compared to SSD provide better accuracy of UAV recognition, while the latter algorithm is also effective in target detection tasks [1]. However, according to the research results presented trained YOLO models outperform Faster RCNN in terms of accuracy metrics and speed (frames per second – FPS), which makes the former algorithm more promising for real-time UAV and bird recognition and classification tasks [2–4]. Considering the recent advances in the improvement of UAV recognition models based on the YOLO algorithm, we should mention the works [2–9]. The authors prepared a dataset in the form of UAV images for training the YOLOv2 neural network [2]. As a result of training, the model reached the maximum value of mAP50 (average accuracy at the threshold of intersection of bounding boxes 50%) – 0.75. The authors studied the features of UAV recognition by YOLOv3 neural network models [3]. By evaluating the qualitative performance of the trained experimental model, it is proposed to introduce densely connected modules to improve the inter-layer connectivity of convolutional neural networks, thereby improving the accuracy and FPS. As a result, the proposed model demonstrated mAP50-95 values of 0.36 and 60 FPS, which outperform the original YOLOv3 model by 0.03 and 24.45 FPS, respectively. The mAP values presented are not equivalent, as mAP50-95 defines the average accuracy at the 95% bounding box crossing threshold, which implies comparatively lower values compared to mAP50 [2, 3]. A better model of the YOLO architecture, YOLOv4, is used where the authors achieved mAP50 values of 0.75, which is consistent with the results of [2, 4]. However, the YOLOv4 models can be considered preferable to the YOLOv2 algorithm primarily due to its higher FPS, which is important for recognizing small objects such as UAVs in real time. YOLOv5 was used as a pre-trained neural network, which is currently considered to be the most popular model of the YOLO architecture for solving the problems of recognizing and classifying any objects from a user's dataset, with training and inference available on the CPU [5, 6]. The neural network from the first paper achieved a mAP50 value of 0.947, while the authors from trained the model to a mAP50 value of 0.904 [5–6]. The paper [7] presents the results of training more advanced versions of YOLO:

YOLOv5: mAP50 - 0.912;
YOLOX: mAP50 - 0.887;
YOLOv7: mAP50 - 0.524;
YOLOv7-tiny: mAP50 - 0.85;
YOLOv8 - 0.953.

However, the authors did not provide the performance of the mAP50-95 metric, which can provide more reliable information about the quality of UAV recognition and detection [7]. In 2023, Ultralytics presented an improved algorithm, YOLOv9 [8]. By incorporating programmable gradient information (PGI) and an efficient layer aggregation network into the architecture, this model will limit the loss of information in successive layers of a deep neural network. YOLOv9 like previous versions utilizes a non-maximal suppression (NMS) algorithm in the post-processing stage, designed to remove unnecessary object bounding boxes. Its use often discards useful predictions and increases computational cost, which is detrimental to the detection and recognition performance of small and maneuverable objects such as UAVs and birds. In this paper, a YOLOv10 model will be trained, which by incorporating new post-processing techniques from the DETR architecture, is hypothesized to increase the recognition and classification accuracy of UAVs and birds [9]. As an analytical accuracy parameter that guarantees the effectiveness of the developed model, mAP50-95 is chosen.

Materials and methods

A more advanced model for real-time recognition and classification of UAVs and birds is based on the YOLOv10 algorithm. YOLOv10 is a neural network based on two key principles: training the model without NMPs and designing based on efficiency and accuracy. The first concept is provided by integrating the advantages of neural networks of the DETR architecture, namely the use of dual label assignment and a consistent metric for matching predictions (Figure 1).

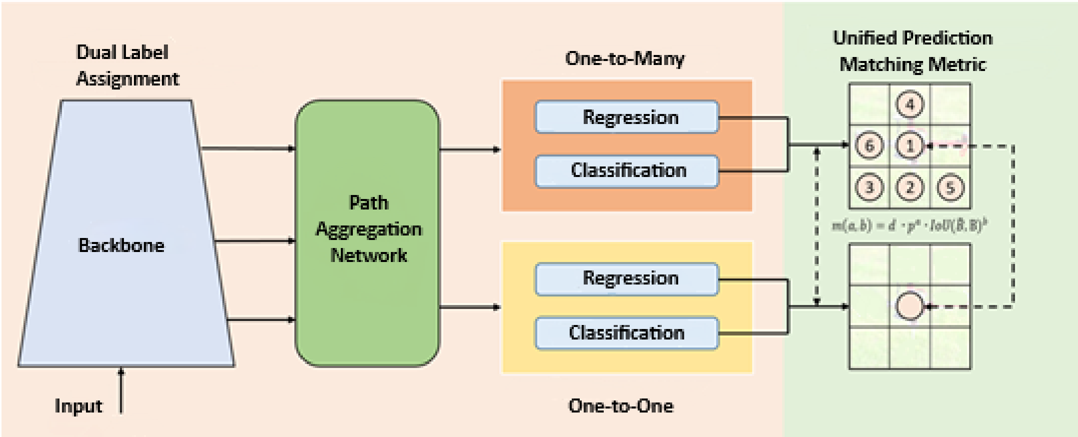


Figure 1 – Concept of the YOLOv10 model

The dual purpose of labels is defined by the architectural solution of using two Heads, where the first Head performs One-to-One matching (provides only one prediction to each benchmark from the user dataset without the use of NMPs) and the second Head performs One-to-Many matching. By utilizing the One-to-Many branch, comprehensive and tight control is provided to ensure that accuracy is maintained. Throughout the training phase, both Heads are used simultaneously, except for the Inference (prediction) phase. At this stage, only the One-to-One branch is used, this decision is made to reduce computational cost while maintaining the accuracy and efficiency of the model. To guarantee the coincidence of the forecasts of the two branches, a consistent metric (or matching metric) is used. The principle of operation of this metric is described by equation (1):

$$m(a,b) = d \cdot p^a \cdot IoU(\tilde{B},\mathrm{B})^b \quad (1)$$

where m is the corresponding accuracy metric (“One to One” or “One to Many”), d is the degree of localization of the prediction reference point within the instance, p is the classification score, a and b are the parameters defining the classification and localization tasks, respectively, IoU is the parameter defining the area of overlap between the predicted and actual frames, is the bounding boxes of the prediction and the instance, respectively. Matching the metrics of both branches ensures that the best samples are matched, thus comprehensive learning control is realized. Designing a performance-based model involves facilitating the classification channel by using two depth-separable convolutions followed by a 1x1 convolution (point convolution). The point convolution allows increasing the number of channels, while the depth convolution reduces the spatial dimensions, such a solution allows reducing the computational cost, as well as preserving more useful information in the downsampling stage. The accuracy orientation of the model is ensured by increasing the dimensionality of kernels in deep convolution layers, as well as by using the partial self-awareness module (PSM).

YOLOv10 supports 6 scales depending on the limitations of the computational resources of the hardware. The YOLOv10m version, which represents a medium version for universal use, is chosen for the experimental study. The dataset for training the neural network includes two classes of images: UAVs and birds. The dataset files in the form of images and videos are taken from the UAV’s own flight archive, as well as open-source resources Roboflow, Kaggle, Ultralytics, and GitHub. The stages of image annotation, augmentation and dataset distribution in the percentage ratio of 70/20/10 (training, validation and testing, respectively) are implemented in Roboflow.com service.

Training, validation and testing of the model (Figure 2) YOLOv10m is realized on the basis of AD103 graphics processor of NVIDIA GeForce RTX 4080 graphics card with support of CUDA Toolkit 12.1. The program code in Python language is implemented in PyCharm 2024 environment using Ultralytics framework, which allows using the open source resource of pre-trained YOLOv10 models. The following hyperparameters are set for training the neural network on the user set: number of epochs: 300; packet size: 16; learning rate: 0.001; momentum: 0.9; weight drop: 0.0005 and image size: 640. The trained YOLOv10m model has a file size of "best.pt" of 101 MB. In order to determine the FPS, the trained neural network is tested on inference using two test videos of UAV and bird flights.

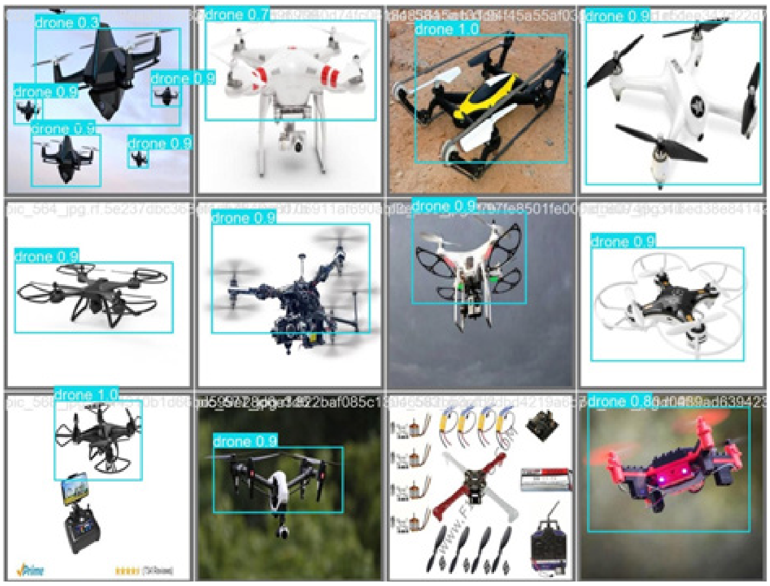
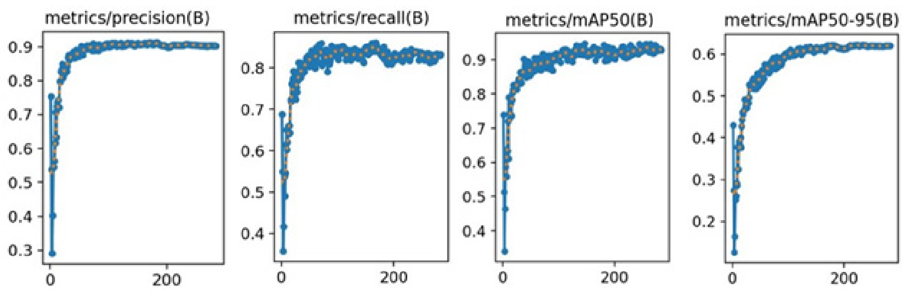


Figure 2 – Testing of the trained YOLOv10m neural network model

To compare the accuracy of the experimental neural network with the previous versions, the YOLOv8m and YOLOv9m models corresponding in purpose were also trained on the user dataset. The mAP50-95 metric was chosen as the comparative accuracy parameter to evaluate the ability of visual detectors to recognize and localize small objects.

Results

Figure 3a-d shows the accuracy metrics of the YOLOv10m model trained on a custom dataset of UAVs and birds.



a b c d

Figure 3 – Metrics of training results of the YOLOv10m neural network:
a - Precision; b - Recall; c - mAP50; d - mAP50-95

The frames of testing the trained model on the inferno using two videos of a DJI F450 UAV flight and a bird are shown in Figure 4a, b.

a b

Figure 4 – Inference frames of the trained YOLOv10m model:
(a) frame from the DJI F450 flight video; (b) frame of a bird (seagull) flight

Figure 5a, b shows the comparison plots of mAP50-95 accuracy metric and fast performance (FPS) of YOLOv8m, YOLOv9m and YOLOv10m neural networks.

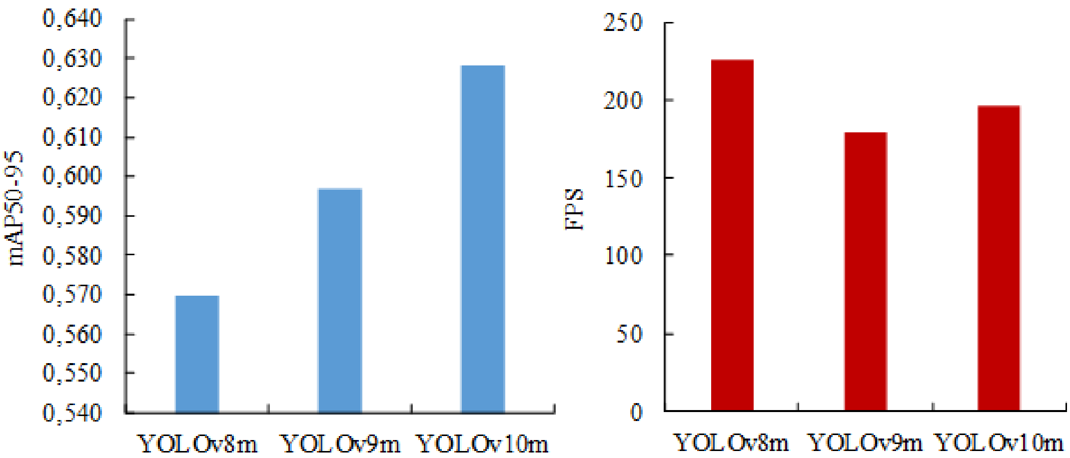


Figure 5 – Comparison plots of mAP50-95 accuracy metrics and FPS of YOLOv8m, YOLOv9m and YOLOv10m neural networks

As a result of the diagrams analysis we can conclude about the effectiveness of using the trained YOLOv10m model in the tasks of recognition and classification of UAVs and birds in optoelectronic surveillance systems.

Discussion

A dataset of 6255 UAV and bird images is prepared for training the experimental model of YOLOv10m neural network. After annotation and augmentation stages, the user dataset using Roboflow.com service is distributed in the percentage of 70/20/10 for training, validation and testing tasks, respectively. The training of experimental neural network resulted in the following maximum values of accuracy metrics (Figure 4a–d):

Completeness: 0.883;
Accuracy: 0.907;
mAP50: 0.953;
mAP50-95: by 0.628.

In terms of mAP50 metric, this model outperforms detectors from [2–7]. Comparing the mAP50-95 metrics (Figure 5a), YOLOv10m outperforms YOLOv8 by 0.058 and YOLOv9 by 0.031 in terms of accuracy, which proves the greater efficiency of the first model to recognize and localize small objects such as UAVs and birds in the far-field surveillance area. In terms of FPS, YOLOv10m is second only to the trained model YOLOv8m, which proves that the first model can be used in real-time in optoelectronic surveillance channels, combining the segmentation and tracking functions of small objects such as UAVs and birds.

Conclusion

1) In order to develop a new efficient model for UAV and bird recognition and classification, the YOLOv10m model is selected, which bypasses the Non-Maximal Suppression (NMS) algorithm in its architecture by introducing the concepts of dual label assignment, consistent metric for prediction matching and focus on accuracy and efficiency in design. A user dataset of 6,255 images is prepared to train the experimental models, of which 4,379 (70%) are specifically designed for training, 1,251 (20%) are oriented for validation and the remaining 625 (10%) are used in the testing phase of the neural network.

2) YOLOv10m training is implemented on the basis of AD103 GPU of NVIDIA GeForce RTX 4080 video card with CUDA Toolkit 12.1 support. As a result of evaluation of the obtained accuracy metrics, YOLOv10 outperforms the known visual detectors YOLOv2, YOLOv3, YOLOv4, YOLOv5, YOLOv7, YOLOX and YOLOv8 described by mAP50 [2–7].

3) To compare the efficiency of recognition and localization of small objects using the mAP50-95 metric, previous versions of the YOLO algorithm: YOLOv8 and YOLOv9 were trained on a user dataset. YOLOv10m was 9% more accurate than the former and 5% more accurate than the latter.

4) As a result of testing the trained YOLOv10m neural network on inference, it is found that this model is able to recognize and classify UAVs and birds in real time with high accuracy and efficiency. According to the high FPS values (average 196 FPS), this model is capable of segmenting and tracking localized objects, combining the classification task.

5) The results of this study will be useful to the developers of UAV and bird detection, recognition and classification systems.

Acknowledgment

The article includes the results of research carried out within the framework of grant funding under the project IRN AP19679009 funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan.

REFERENCES

1 Alkentar, Saad & Alsahwa, B. & Assalem, A. & Karakolla, D. Practical comparation of the accuracy and speed of YOLO, SSD and Faster RCNN for drone detection. Journal of Engineering, 27, 19–31 (2021). https://doi.org/10.31026/j.eng.2021.08.02.

2 Seidaliyeva, U., Alduraibi, M., Ilipbayeva, L., Almagambetov, A. 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, November 9–11, 2020), pp. 490-494. https://doi.org/10.1109/IRC.2020.00093.

3 Alsanad, Hamid & Sadik, Amin & Ucan, Osman & Ilyas, Muhammad & Bayat, Oguz. YOLO-V3 based real-time drone detection algorithm. Multimedia Tools and Applications, 81, 1–14 (2022). https://doi.org/10.1007/s11042-022-12939-4.

4 Singha, S., Aydin, B. Automated Drone Detection Using YOLO v4. Drones. September 2021. https://doi.org/10.3390/drones5030095.

5 Al-Qubaydhi, Nader & Alenezi, Abdulrahman & Alanazi, Turki & Senyor, Abdulrahman & Alanezi, Naif & Alotaibi, Bandar & Alotaibi, Munif & Razaque, Abdul & Abdelhamid, Abdelaziz & Alotaibi, Aziz. Detection of Unauthorized Unmanned Aerial Vehicles Using YOLOv5 and Transfer Learning. Electronics, 11, 2669 (2022). https://doi.org/10.3390/electronics11172669.

6 Aydin, B., Singha, S. Drone Detection Using YOLO v5, Eng. 4 (2023). https://doi.org/10.3390/eng4010025.

7 Zhai, X., Huang, Z., Li, T., Liu, H., Wang, S. YOLO-Drone: An Optimized YOLOv8 Network for Tiny UAV Object Detection. Electronics, 12, 3664 (2023). https://doi.org/10.3390/electronics12173664.

8 Li, Jun & Feng, Yongqiang & Shao, Yanhua & Liu, Feng. (2024). IDP-YOLOV9: Improvement of Object Detection Model in Severe Weather Scenarios from Drone Perspective. Applied Sciences, 14, 5277 (2023). https://doi.org/10.3390/app1412125277.

9 Muzammul, Muhammad & Algarni, Abdul & Ghadi, Yazeed & Assam, Muhammad. Enhancing UAV Aerial Image Analysis: Integrating Advanced SAHI Techniques with Real-Time Detection Models on the VisDrone Dataset. IEEE Access, pp. 1–1 (2024). https://doi.org/10.1109/ACCESS.2024.3363413.

[1]Семенюк В.В.,
т.ғ.м., ORCID ID: 0000-0002-8580-7326,
e-mail: vvsemenyuk@ku.edu.kz;
[1]*Курмашев И.Г.,
PhD, ORCID ID: 0000-0001-9872-7483,
*e-mail: ikurmashev@ku.edu.kz;
[2]Сербин В.В.,
PhD, қауымдастырылған профессор, ORCID ID: 0000-0002-5807-3873,
e-mail: v.serbin@satbayev.university;
[1]Курмашева Л.Б.,
магистр, ORCID ID: 0000-0001-5392-5873,
e-mail: lbkurmasheva@ku.edu.kz;
[2]Молдагулова А.Н.,
PhD, профессор, ORCID ID: 0000-0002-1596-561X,
e-mail: a.moldagulova@satbayev.university;

[1]М. Қозыбаев атындағы Солтүстік Қазақстан университеті, Петропавл қ., Қазақстан
[2]Satbayev University, Алматы қ., Қазақстан.

YOLOV10 НЕЙРОНДЫҚ ЖЕЛІСІ НЕГІЗІНДЕ НАҚТЫ УАҚЫТТА ҰШҚЫШСЫЗ ҰШУ АППАРАТТАРЫН ТАНУ МОДЕЛІН ӘЗІРЛЕУ

Аңдатпа

Мақалада YOLOv10 нейрондық желісін оқыту негізінде нақты уақыт режимінде ұшқышсыз ұшу аппараттары мен құстарды тану және жіктеу моделін әзірлеу қарастырылады. Зерттеу бағыты әртүрлі салаларда оларды қолданудың артуын ескере отырып, қауіпсіздік тұрғысынан ұшқышсыз ұшу аппараттарын анықтау мәселелерінің өзектілігімен байланысты. Модельді оқыту үшін жеке мұрағаттар мен ашық ресурстардан жиналған 6255 суреттен тұратын деректер жиынтығы дайындалды. Аннотациялау, деректерді күшейту және тарату процесі Roboflow.com қызметі арқылы жүзеге асырылды. Модельді оқыту ULTRALYTICS фреймворкін қолдана отырып, NVIDIA GeForce RTX 4080 графикалық процессорында жүргізілді. Тестілеу нәтижелері mAP50 және mAP50–95 көрсеткіштері бойынша YOLO желісінің алдыңғы нұсқаларымен салыстырғанда танудың жоғары дәлдігін көрсетті. Модель объектілерді тиімді сегментациялау және қадағалау қабілетін көрсетті, бұл оны оптикалық-электронды бақылау жүйелерінде қолдануға перспективалы етеді. Зерттеу нәтижелері ұшқышсыз ұшу аппараттары мен құстарды анықтау және жіктеу жүйелерін әзірлеушілерге, сондай-ақ әртүрлі салаларда қауіпсіздікті арттыруға пайдалы болуы мүмкін.

Тірек сөздер: нейрондық желілер, жіктеу, тану, оптикалық-электронды бақылау арналары, UAV, YOLO, конволюциялық нейрондық желілер.

[1]Семенюк В.В.,
м.т.н., ORCID ID: 0000-0002-8580-7326,
e-mail: vvsemenyuk@ku.edu.kz
[1]*Курмашев И.Г.,
PhD, ORCID ID: 0000-0001-9872-7483,
*e-mail: ikurmashev@ku.edu.kz
[2]Сербин В.В.,
PhD, ассоциированный профессор, ORCID ID: 0000-0002-5807-3873,
e-mail: v.serbin@satbayev.university
[1]Курмашева Л.Б.,
магистр, ORCID ID: 0000-0001-5392-5873,
e-mail: lbkurmasheva@ku.edu.kz
[2]Молдагулова А.Н.,
PhD, профессор, ORCID ID: 0000-0002-1596-561X,
e-mail: a.moldagulova@satbayev.university

[1]Северо-Казахстанский университет имени М. Козыбаева, г. Петропавловск, Казахстан
[2]Satbayev University, г. Алматы, Казахстан

РАЗРАБОТКА МОДЕЛИ РАСПОЗНАВАНИЯ БПЛА В РЕЖИМЕ РЕАЛЬНОГО ВРЕМЕНИ НА ОСНОВЕ НЕЙРОННОЙ СЕТИ YOLOV10

Аннотация

В статье рассматривается разработка модели распознавания и классификации БПЛА и птиц в режиме реального времени на основе обучения нейронной сети YOLOv10. Направление исследования считается актуальным в связи с проблемами обнаружения БПЛА в контексте обеспечения безопасности, учитывая их растущее использование в различных сферах. Для обучения модели подготовлен датасет, состоящий из 6255 изображений, собранных из собственных архивов и открытых ресурсов. Процесс аннотирования, аугментации и распределения данных был реализован с использованием сервиса Roboflow.com. Обучение модели проводилось на графическом процессоре NVIDIA GeForce RTX 4080 с использованием фреймворка Ultralytics. Результаты тестирования показали высокую точность распознавания с метриками mAP50 и mAP50-95, превышающими показатели предыдущих версий YOLO. Модель демонстрирует способность к эффективной сегментации и трекингу объектов, что делает ее перспективной для применения в системах оптикоэлектронного наблюдения. Результаты исследования могут быть полезны для разработчиков систем обнаружения и классификации БПЛА и птиц, а также для повышения безопасности в различных областях.

Ключевые слова: нейронные сети, классификация, распознавание, оптикоэлектронные каналы наблюдения, БПЛА, YOLO, сверточные нейронные сети.

Article submission date: 23.09.2025