UDC 629.7.014-519.058.6:528.72:629.331 **doi: 10.32620/aktt.2024.3.08**

Rostyslav TSEKHMYSTRO, Oleksii RUBEL, Vladimir LUKIN

National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine

STUDY OF THE DEPENDENCE OF ACCURACY IN VEHICLES SEARCH ON THE SIZE OF THE OBJECT USING UAV IMAGES

Digital images are increasingly used to analyze different types of objects and their localization and classification. There are many areas for using this information; it is often employed for surveillance systems, automatic driving of vehicles, or exploration of new territories. At the same time, there are a fairly large number of neural networks that allow implementation of this functionality by training them using data sets of various types and classifications. Often, data sets created with the help of unmanned aerial vehicles are frequently used for research tasks. Such datasets allow the recognition of various types of objects without direct access to them, which allows safe exploration of different territories. The use of unmanned aerial vehicles is quite common nowadays, especially in the fields of photography and videography. Many photographers use unmanned aerial vehicles to take pictures of landscapes and use automatic tracking systems for movement. Automatic movement systems and object search systems are quite sensitive to the size of the object and the quality of the search algorithm. Because of the wide applicability of this task, as well as the small amount of initial data, the topic of our work is the study of the dependence of the accuracy of localization and classification of objects on their area in images obtained using unmanned aerial vehicles. The ***main subject*** *of this study is the quality of neural networks that allow obtaining information about objects, as well as research by obtaining statistical data and a test set of data on the dependence of detection accuracy on the size of the object. The* ***goal of this study*** *was to obtain statistics on the accuracy of localization and classification depending on the size of the object and to determine the accuracy thresholds using the obtained statistics. The* ***task of this study*** *is to train common neural networks with an open architecture on a set of data obtained using unmanned aerial vehicles and to determine their characteristics, particularly the dependence of recognition accuracy on the size of the object. The* ***expected result*** *of the work is the threshold values of the size of the object, which are permissible for a sufficiently accurate classification and localization of objects, as well as the metrics of the quality of the work of the studied neural networks. Because of this work,* ***conclusions*** *are given that reflect the threshold values of object sizes, on which the recognition accuracy depends.*

Keywords: *object localization; YOLO v5; SSD; FasterRCNN; vehicle classification; unnamed aerial object.*

Introduction

Motivation

The scope of object recognition and localization tasks is quite wide. These tasks arise in many areas of research and practice. Recognition systems can be of different types and use different devices to acquire images. Most systems work with conventional images obtained with the help of optical systems [1]. But in some areas, images acquired by radar [2], sonar [3], or infrared [4] systems are also employed.

More and more often, unmanned aerial vehicles are used to explore the Earth surface [5]. They allow acquiring images of the Earth surface without direct access. But obtaining high-quality images using unmanned aerial vehicles (UAV) is quite a challenge. When using them, there are problems with the weight of the device, its energy independence and energy efficiency, and its design [6]. The use of high-resolution sensors is quite problematic because they require a large amount of energy, which negatively affects the weight of the UAV. By using efficient image processing using neural networks for object localization [7], it is possible to detect even small objects in high definition (HD) images.

At the same time, the dependence of object detection accuracy on the size of these objects in the photo remains unknown. Obtaining such dependences for different types of objects will allow designing unmanned aerial vehicles and the corresponding equipment for the required tasks, knowing the required sensor parameters in advance.

State-of-the-art

Analyzing studies on the dependence of localization accuracy on the size of the object, it becomes clear that the use of UAV images is quite new, but is becoming widespread. Analyzing our previous study [8] intended on obtaining the dependence of people localization on UAV images, we have already got some statistics. This publication is aimed at confirming or refuting, as well as expanding the conclusions made in the previous work.

The studies of other authors [9, 10] show mostly the general accuracy of neural networks for object localization. The information provided in these studies shows that neural networks can locate transport with a fairly high accuracy.

The speed of neural networks for localization and classification also plays an important role [11, 12]. With sufficient performance, high speed allows neural networks to be used in real-life devices, making such systems more productive. The speed of a neural network also affects the ability to process images in real time. Fast neural networks can now be utilized to control vehicles and, in the case of UAVs, for their control and guidance[13].

Also, when studying available publications [14, 15], a set of metrics was obtained to display the accuracy of the neural network for localization and classification. So, to determine the accuracy of localization, intersection over union (IoU) [16] is often used. When evaluating the classification accuracy, Precision and Recall [17] are employed.

Objectives and the approach

Our work is devoted to determining the dependence of the accuracy of localization and classification of the object, which will be combined into a single accuracy parameter, on the size of this object expressed in pixels. To do this, several neural networks were trained, these are YOLO v5 [18], YOLO v8 [19] and FasterRCNN [20]. These neural networks are quite widely exploited in the considered sphere. For trained neural networks, the metrics specified in the previous subsection will be calculated. These metrics will allow assessing the quality of the trained model and compare the result with the results of the metrics given in other studies. Using the obtained weights of neural networks, localization and classification will be evaluated on a test dataset having pre-created annotations, and, based on the obtained data, the localization accuracy statistics will be built depending on the size of the object. Then, conclusions will be drawn.

1. Train and test data

One of the most important aspects in the training of a neural network is the choice of a training dataset. The quality of annotations in the training dataset affects the accuracy of the neural network, as well as the size of objects that will be localized. That is why it is important to study the training dataset before using it.

Investigating the work of other authors [21], the VisDrone dataset was chosen, which has a sufficient number of images for training. Analyzing this data set, statistics were obtained on the distribution of objects by size and classes. In total, the dataset has 6 classes, these are people, cars, buses, trucks, bicycles, tricycles. The main thing for our task is the amount of data in categories related to transport.

In general, the selected dataset is concealed from 7019 images, of which 6471 images belong to the training sample and 648 to the validation one. The training part contains 343205 objects, of which about 65% is different types of transport. Figure 1 shows the distribution of the number of objects in each of the dataset classes in the training sample, and Figure 2 shows the distribution of the number of objects for each of the classes in the validation part of the dataset. When analyzing the number of objects, it is noticeable that most of the objects in the dataset belong to the category of cars, which meets our requirements when choosing a data set.

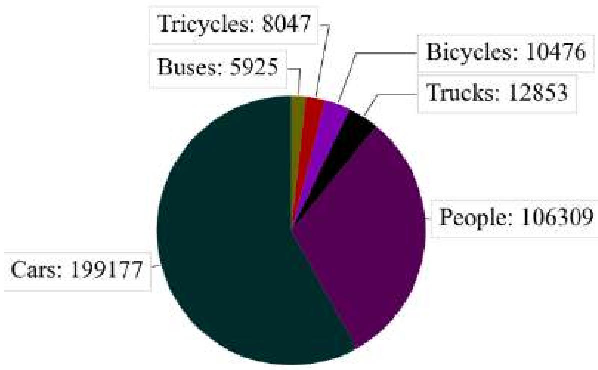


Fig. 1. Distribution of object categories in the training subset in VisDrone dataset

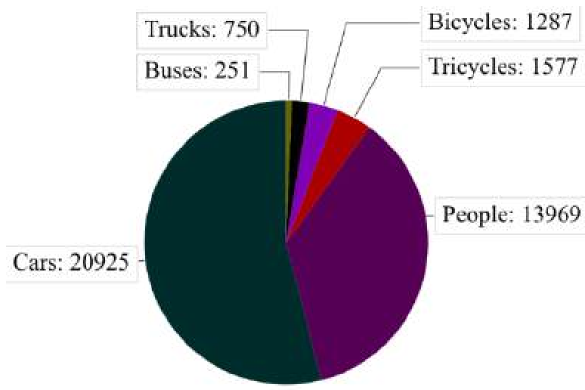


Fig. 2. Distribution of object categories in the validation subset in VisDrone dataset

The size of objects also has an important role in learning, as noted above, so the statistics of the distribution of objects relative to their size was also summarized. To determine this characteristic, the size of objects for all annotations in the dataset was calculated and the distribution of the number of objects from their size was built. The obtained statistics are presented in Figure 3. Given the data obtained, it is noticeable that the main part of the objects has a size of up to 300 pixels, and the median distribution is a value of 50 pixels. The statistics obtained fully meets the needs for research.

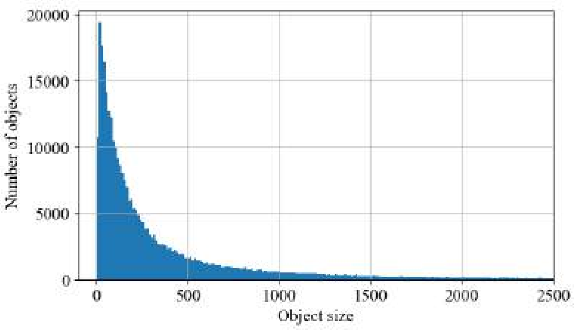


Fig. 3. Distribution of object sizes
in training dataset

To determine the metrics of the accuracy of the trained models, a data set (named UAVDT [22]) was selected consisting of 4000 images, which are divided into 3 categories. Given that the study focuses on vehicle localization and classification, the dataset was chosen to contain the largest number of vehicle-related categories. In the selected dataset, 130168 objects are annotated, which are divided into 3 classes: cars, trucks and buses. For the dataset, statistics on the distribution of annotated objects by categories were built, which is presented in Figure 4. Analyzing the data obtained, it is noticeable that the bulk of the objects are cars.

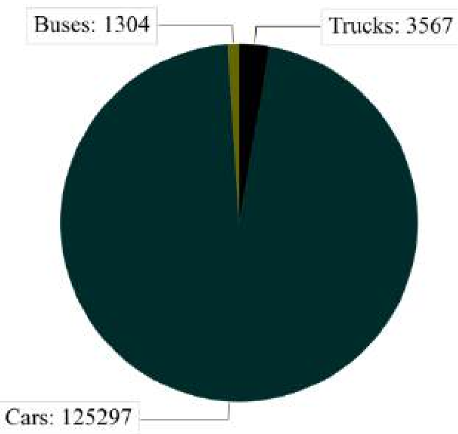
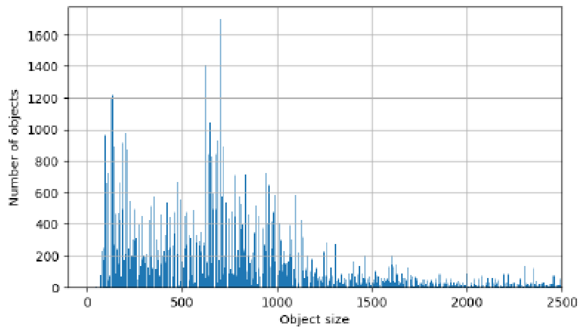
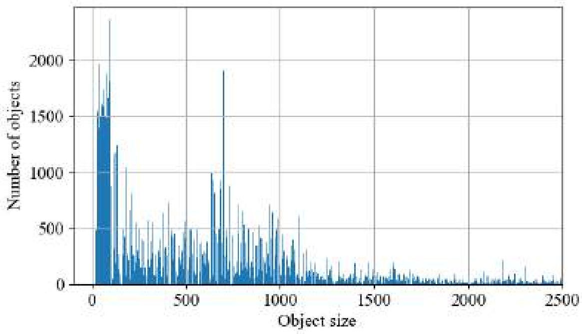


Fig. 4. Distribution of object categories
in test dataset

Also, for the selected data set, statistics on the distribution of objects by size are constructed, in accordance with the training part, which is presented in Figure 5,a. When analyzing the results obtained, it is noticeable that the median of the sizes distribution is slightly to the right of the test sample and has a value close to 650 pixels, this differs significantly from the expected parameters of the data set. Considering this factor, the data were annotated using labelImg marking program. During the labeling process, 40,000 small objects were added to obtain more detailed statistics when analyzing the trained model. The obtained statistical data for the test dataset, after additional data collection, are presented in Figure 5,b. The median of the statistics obtained is in the region of 50 pixels, which corresponds to the training dataset and fully satisfies our needs.



a



b

Fig. 5. Distribution of objects size
in test dataset with default annotation (a)
and with added annotations (b)

2. Neural network training

To study the influence of the size of the object on the accuracy of its localization and segmentation, several architectures of neural networks were chosen, which should correspond to the following characteristics:

- popularity at the time of the study;
- prevalence of use;
- the ability to scale or change backbones, layers.

Many architectures correspond to these characteristics, but we chose YOLO v5, YOLO v8, FasterRCNN. YOLO v5 has been a widely studied and used architecture for recent years. YOLO v8 is the latest architecture and continuation of a series of models from ultralitics. FasterRCNN is a fairly classical neural network architecture for localizing and classifying objects. In view of these factors, we can expect that the given comparison will be quite objective.

2.1. YOLO v5

To train the model with the YOLO v5 architecture, the optimal implementation of the model in time and quality was chosen - YOLO v5 small (hereinafter YOLO v5s). Also, the entire learning process was conducted using the official implementation and infrastructure from ultralitics. Training continued through 300 epochs, each of which was validated on a validation dataset. Indicators of loss functions and metrics obtained for the validation dataset were used to assess the quality of training during its implementation.

In the learning process, mean squared error (MSE) (1) [23] was used as a regression loss function to localize object frames:

$$MSE = \frac{1}{n}\sum_{i=0}^{n}(Y_i - \hat{Y}_i)^2 , \qquad (1)$$

where n is a length of samples,

Y is a true vector,

$\hat{Y}$ is a predicted vector.

Binary cross-entropy (BCE) loss (2) [24] was used to display the loss function of objectivity, which reflects the confidence of the model that there is an object in the frame:

$$l_n = -w_n[y_n \log\sigma(x_n) + (1-y_n)\log(1-\sigma(x_n))], \qquad (2)$$

where x_n is the input classification vector,

y_n is the true classification vector,

$\sigma(x_n)$ is the probability of each class.

Cross-Entropy loss (3) [25] was also used as a loss function for classification:

$$l_n = -\sum_{c=1}^{C} w_c \log \frac{\exp(x_{n,c})}{\sum_{i=1}^{C}\exp(x_{n,i})} y_{n,c}, \qquad (3)$$

where x_n is the true classification data for the object,

y_n is the classification vector provided by this method,

w_c is the weight for each class,

C is the number of classes.

For training, the stochastic gradient descent (SGD) [26] optimizer was used, the parameter values of which are the following: lr = 0.01, momentum = 0.937, 16 images per batch. Graphs obtained during the training of the model are presented in Figure 6, for loss of objectivity (Figure 6, a), for loss of classification (Figure 6, b) and for loss of localization (Figure 6, c). Analyzing the obtained data, it is noticeable that, by the end of the training, the model reached a plateau and the loss function almost does not change.

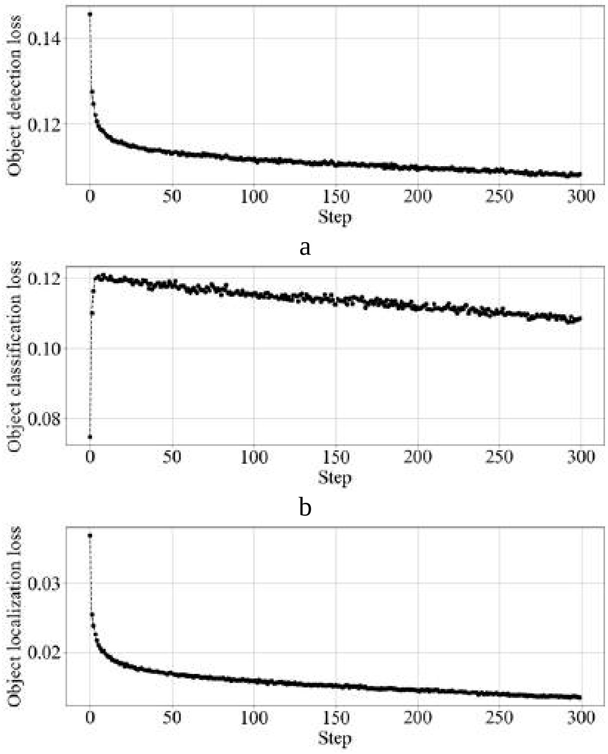


Fig. 6. Loss functions for YOLO v5 training process, which include objectness loss (a), object localization loss (b) and localization loss (c)

2.2. YOLO v8

YOLO v8 is a continuation of the development of the YOLO architecture, therefore, a small version was also chosen to obtain comparable results with the version described in the previous part. The training was also conducted using an official repository from ultralitics. BCE loss is used as a loss function for classification, which is calculated for each class, the implementation is presented in the previous subsection. As a loss function for localization, the so-called boundary box (BBOX) loss is used, which is the inverse of the intersection over union (IoU) metric, which reflects the localization accuracy of the predicted frames relative to the annotated ones:

To calculate the loss of objectivity, the third loss function is used - Distribution Focal Loss (DFL) [27]. It reflects the accuracy of localization and confidence of the neural network, which is an object in the frame (objectness).

For training, the stochastic gradient descent (SGD) optimizer was used, the parameters of which are the following: lr = 0.01, momentum = 0.9, 16 images per batch. In accordance with YOLO v5, YOLO v8 also employed 300 epoches. Figure 7 shows graphs of loss functions as described above, so that Figure (7,a) shows the classification loss function, Figure (7,b) - the localization loss function, and the data for the DFL loss function are given in Figure (7, c). Analyzing the data obtained, it is noticeable that the loss functions at the end of training have a rather low value.

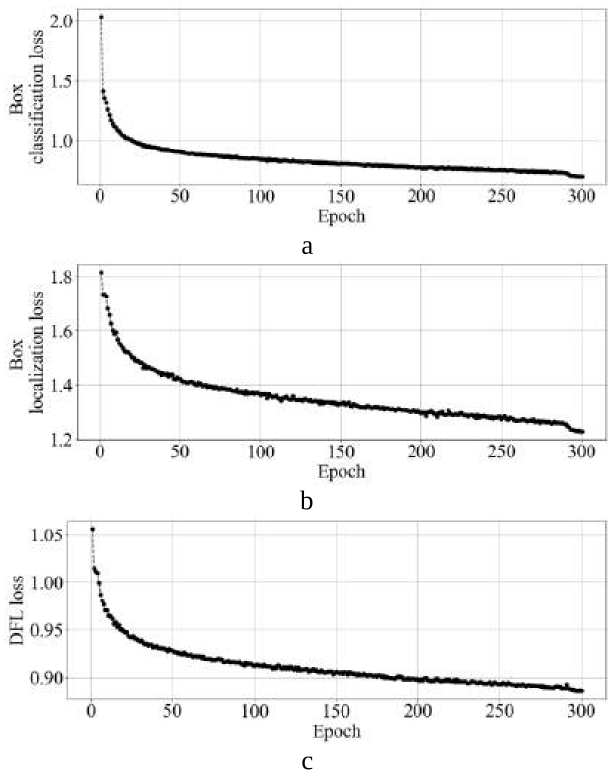


Fig. 7. Loss functions for YOLO v8 training process, which include classification loss (a), object localization loss (b) and DFL loss (c)

2.3. Faster RCNN

Faster RCNN architecture is based on RCNN architecture, which has three levels. The first level is the extraction of features from an image using one of the neural networks (backbones), for example, it can be Mobilenet, Resnet, Retina or other architectures. The speed and accuracy of the resulting neural network depends on the network used. The second level that receives feature input from the first level is called the region proposal network (RPN), which offers a set of regions given the value in the features. The third level is called region of interest (ROI) pooling, dedicated to localization and classification of regions proposed at level 2. As information used for classifications and definition of objectivity, an array of features obtained at the first level is utilized. Thus, the reuse of features allows speeding-up the model.

Taking into account the information about the structure of the neural network, we chose a fast model using mobilenet v3 as a backbone. This version of the network is quite fast and has good localization accuracy indicators.

For Faster RCNN training, smoothed L1 (4) loss function is used to calculate localization loss for RPN and for localization loss at model output:

$$l_n = \begin{cases} \dfrac{0.5(x_n - y_n)^2}{beta}, & \text{if } |x_n - y_n| < beta, \\ |x_n - y_n| - 0.5 * beta, & \text{otherwise}, \end{cases} \tag{4}$$

where x_n is the true information about the object,

y_n is the predicted information about the object,

beta is the threshold of change between L1 and L2 loss (hyperparameter, non-negative, default is 1.0).

The loss function for classification is cross-entropy loss, as already described for other models. For training, the Adam optimizer was used, with the parameters lr = 0.001, weight_decay=0.0005. The batch size during the training was 16 images. Validation was carried out once in an epoch, in total, the training lasted 300 epochs, according to the previous models. The results obtained for the loss functions are shown in Figure 8, for localization losses in RPN (Figure 8, a), total localization losses (Figure 8, b) and classification losses (Figure 8, c).

3. Performance of trained networks

To check the quality of the models, metrics were used, which are the most often exploited in localization and classification tasks. They allow displaying the accuracy of localization and classification for models, as well as obtaining the information about the threshold that is used to filter out objects with a low score.

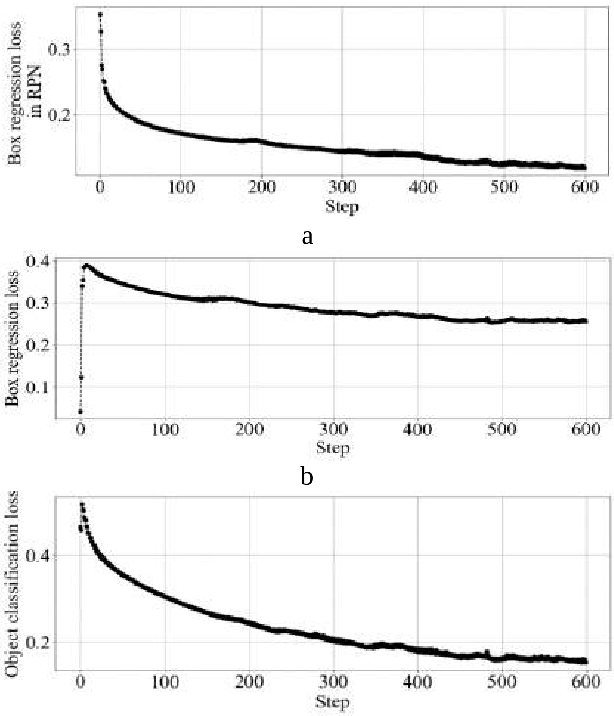
As a basic localization metric, we used intersection over union (IoU), which allows obtaining accurate parameters of the intersection of predicted and annotated blocks. The IoU calculation algorithm can be visually represented as shown in Figure 9. The shaded area is an indicator of the IoU metric, an annotated frame is highlighted in green, and a model is provided in red.

Precision (5) and Recall (6) metrics were used to calculate classification accuracy, the calculation formulas of which are presented below:

$$P = \frac{TP}{FP + TP}, \tag{5}$$

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

$$R = \frac{TP}{FN + TP}. \qquad (6)$$



a b c

Fig. 8. Loss functions for Faster RCNN training process, which include box regression loss in RPN (a), overall box regression loss (b) and classification loss (c)

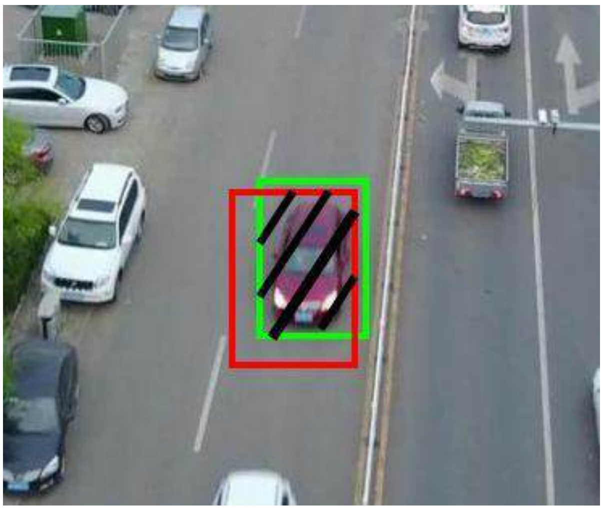


Fig. 9. Visual represented of IoU

To calculate the above parameters, the distribution indicators of the provided blocks are required. Among them, true positive rates (TP) display the number of correctly provided frames for the class for which it is calculated, true negative rates (TN) display the number of frames that have been annotated for other classes, against which it is calculated, false positive rates (FP) - the number of frames provided by the network, but not annotated in the dataset, false negative rates (FN) - the number of frames annotated in the dataset, but not provided by the network.

Validation was carried out for all data from the validation part of the dataset, the data from which are in no way involved in the training of the neural network. In view of this, network comparison is objective and can be relied to select a model to use. Validation results are presented in Table 1 for all neural networks.

Table 1
Precision, recall, and IoU metrics for the trained models

Model	Precision	Recall	IoU
YOLO v5	0.540	0.928	0.672
YOLO v8	0.527	0.667	0.683
Faster RCNN	0.528	0.627	0.645

Analyzing the obtained metrics, it is noticeable that the models from the YOLO architecture work are on about the same level, but YOLO v5 is still preferable in terms of the considered metrics. Faster RCNN also has a small gap from competitors, but the resulting metrics are quite consistent with the needs for localization tasks.

4. Vehicle localization study

To conduct research on the characteristics of the localization of transport, the framework was predicted using the neural networks described in the previous part. All the obtained results were saved in a file and the value of the localization accuracy metric was obtained relative to each of the predicted objects. Comparison of the predicted frames with annotated ones was carried out by counting the IoU metric, as a result, the frames passing the threshold were selected for comparison. For each of the compared frames, the area was calculated, based on the data provided by the network. These data are approximate, because they use only the provided frame, and the object may not occupy the entire area of this frame. Given this, the result may have a bias.

First, let us analyze the data obtained using the YOLO v5 neural network, because, as shown above, it has the best characteristics. The obtained dependence of work accuracy is shown in Figure 10. Analyzing the results obtained, it is noticeable that the best accuracy (more than 0.8) is achieved with an object size of 265 pixels, the threshold is highlighted with a red dotted line. The average accuracy, which is sufficient to localize most types of objects (more than 0.6), is achieved with the object size of more than 100 pixels, highlighted in the Figure with a yellow dotted line. Also, objects that are localized with less accuracy, but can be taken into account when using a neural network, were assigned to the zone from 0.4 to 0.6 (highlighted by a red dotted line). The size of this type of objects should exceed 75 pixels. Objects having a smaller size can be recognized by the neural network, but have a low level of accuracy and confidence of the neural network, respectively, when lowering the threshold, a large number of incorrect objects can be localized.

Conducting research for the objects provided by the YOLO v8 neural network, we got the dependence that is depicted in Figure 11. Analyzing the graph, it is noticeable that, with respect to the results for YOLO v5, we have a slightly biased dependence. Thus, sufficient accuracy is achieved with the object size of 310 pixels, which is significantly larger than for the previous model. The object size has also increased, which is necessary for sufficient classification accuracy. For YOLO v8, it is 130 pixels. The size of the object for uncertain classification has not changed significantly, it is 80 pixels.

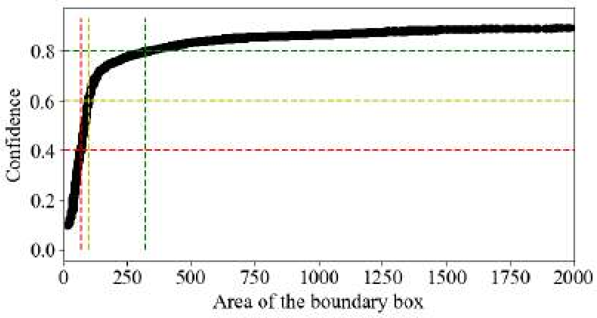


Fig. 10 Dependence of localization accuracy and classification on boundary box size for YOLO v5

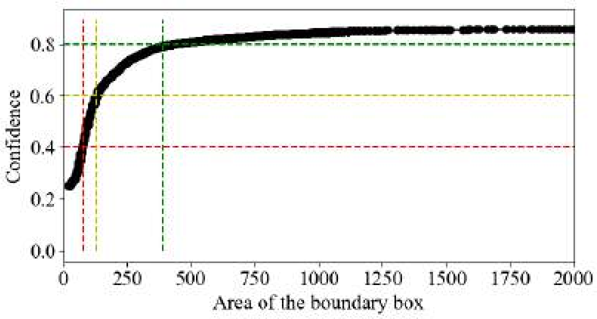


Fig. 11 Dependence of localization accuracy and classification on boundary box size for YOLO v8

For Faster RCNN, the dependence of classification accuracy on the object area was also calculated, but in the case of this neural network, the statistics are much worse than for the previous ones. The results are shown in Figure 12. Analyzing the statistics obtained, it is noticeable that high accuracy is achieved with the object area of 1350 pixels, which is a fairly high value and considerably larger than for the previous models. Also, the level of sufficient accuracy is achieved with the object size of 920 pixels, which is too large. A low level of accuracy is achieved at 550 pixels, which is much higher than the level of high accuracy for previous models.

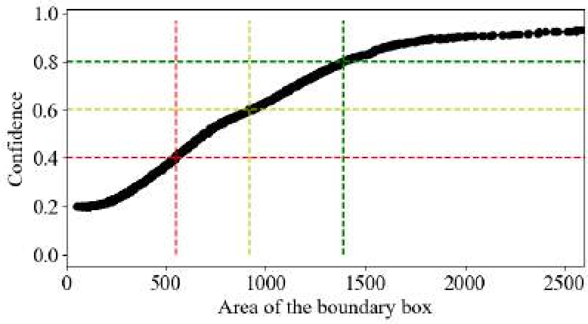


Fig. 12 Dependence of localization accuracy and classification on boundary box size for Faster RCNN

Taking into account the obtained statistics and analyzing the results, we can understand that Faster RCNN has worse results compared to other models. To reduce the variation in results, the results of Faster RCNN will not be used in determining the overall score. Summarizing the overall assessment, one can determine the average threshold of high accuracy of localization and classification, it ranges from 260 to 320 pixels, depending on the used neural. Thus, it can be determined that the average value of the object area, which is sufficient to achieve high localization and classification accuracy, is about 290 pixels. Also, the analysis determines that the size of the object for sufficient accuracy of localization and classification should be from 100 to 130 pixels, which is about 0.03% of the size of the studied image, which has size 640x640 pixels. The area value for minimum localization and classification accuracy is 75 pixels and is approximately the same for models with YOLO architecture. Objects having an area of less than 75 pixels are recognized with low accuracy, which makes it impossible to classify them.

Conclusions

As a result of the research, the statistics of the accuracy of localization and classification of objects were determined depending on their size. The parameter of estimation accuracy was chosen as an indicator, which reflects the confidence of the neural network in the predicted object. Also, for trained neural networks, localization accuracy metrics and classification accuracy metrics such as IoU, Precision and Recall were determined.

Analyzing the results, it can be noted that the Faster RCNN architecture is quite good in terms of metrics, but it has much worse accuracy in localizing small objects, so its results were not used. YOLO v5 and YOLO v8 have similar indicators for work quality metrics, as well as for the indicators of the studied dependence.

When analyzing the dependencies of the accuracy of localization and classification of objects on their area in pixels, several threshold values of the areas of objects can be distinguished:

- with an area of the object from 290 pixels, it can be calculated that the object is localized and classified with great accuracy (larger than 0.8);
- when the area of the object is more than 130 pixels - the object can be localized with sufficient accuracy, which is more than 0.6;
- to achieve the minimum permissible accuracy (more than 0.4), the minimum permissible size of the object is 75 pixels;
- if the area of the object is less than 75 pixels, the object is localized and classified with low accuracy, which leads to unreliable results.

Authors' contributions: conception – **Oleksii Rubel**, **Rostyslav Tsekhmystro**; methodology – **Oleksii Rubel**, **Rostyslav Tsekhmystro**; problem formulation – **Oleksii Rubel**; analysis – **Rostyslav Tsekhmystro**; model development – **Rostyslav Tsekhmystro**; software – **Rostyslav Tsekhmystro**; validation – **Vladimir Lukin**, **Oleksii Rubel**; analysis of results – **Rostyslav Tsekhmystro**, **Oleksii Rubel**, **Vladimir Lukin**; visualization – **Rostyslav Tsekhmystro**; writing – **Rostyslav Tsekhmystro**; revision and editing – **Oleksii Rubel**, **Vladimir Lukin**.

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

The study was conducted without financial support.

Data availability

The manuscript contains no associated data.

Use of Artificial Intelligence

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

All authors have read and agreed to the published version of this manuscript.

References

1. Bondžulić, B., Stojanović, N., Lukin, V., Stankevich, S. A., Bujaković, D., & Kryvenko, S. Target acquisition performance in the presence of JPEG image compression. *Defence Technology*, 2023. DOI: 10.1016/j.dt.2023.12.006.

2. Wang, L., Tang, J., & Liao, Q. A Study on Radar Target Detection Based on Deep Neural Networks. *IEEE Sensors Letters*, 2019, vol. 3, no. 3, article no. 7000504. DOI: 10.1109/LSENS. 2019.2896072.

3. Yu, S. Sonar Image Target Detection Based on Deep Learning. *Mathematical Problems in Engineering*, 2022, vol. 2022, article no. 5294151. DOI: 10.1155/2022/5294151.

4. Zhao, M., Li, W., Li, L., Hu, J., Ma, P., & Tao, R. Single-Frame Infrared Small-Target Detection: A survey. *IEEE Geoscience and Remote Sensing Magazine*, 2022, vol. 10, no. 2, pp. 87-119. DOI: 10.1109/MGRS.2022.3145502.

5. Cox, T. H., Somers, I. & Fratello, D. J. *Earth Observations and the Role of UAVs: A Capabilities Assessment.* NASA, 2007. 346 p.

6. Boukoberine, M. N., Zhou, Z. & Benbouzid, M. A critical review on unmanned aerial vehicles power supply and energy management: Solutions, strategies, and prospects. *Applied Energy*, 2019, vol. 255, paper no. 113823. DOI: 10.1016/j.apenergy.2019.113823.

7. Xiuling, Z., Huijuan, W., Yu, S., Gang, C., Suhua, Z. & Quanbo, Y. Starting from the structure: A review of small object detection based on deep learning. *Image and Vision Computing*, 2024, vol. 146, paper no. 105054. DOI: 10.1016/j.imavis.2024.105054.

8. Tsekhmystro, R., Rubel, O. & Lukin, V. Investigation of the effect of object size on accuracy of human localisation in images acquired from unmanned aerial vehicles. *Aerospace Technic and Technology*, 2024, no. 2, pp. 83-90. DOI: 10.32620/aktt.2024.2.09.

9. Tang, G., Ni, J., Zhao, Y., Gu, Y. & Cao, W. A Survey of Object Detection for UAVs Based on Deep Learning. *Remote Sensing*, 2024, vol. 16, no. 1, article no. 149. DOI: 10.3390/rs16010149.

10. Yamani, A., Alyami, A., Luqman, H., Ghanem, B. & Giancola, S. Active Learning for Single-Stage Object Detection in UAV Images. *2024 IEEE/CVF Winter Conference on Applications of Computer Vision*

(WACV), 2024, pp. 1849-1858. DOI: 10.1109/WACV57701.2024.00187.

11. Agrawal, P., Girshick, R. & Malik, J. Analyzing the Performance of Multilayer Neural Networks for Object Recognition. *Computer Vision – ECCV 2014*, 2014, vol. 8695, pp. 329–344. DOI: 10.1007/978-3-319-10584-0_22.

12. Ren, J. & Wang, Y. Overview of Object Detection Algorithms Using Convolutional Neural Networks. *Journal of Computer and Communications*, no. 10, pp. 115–132. DOI: 10.4236/jcc.2022.101006.

13. Khan, S. A., Lee, H. J. & Lim, H. Enhancing Object Detection in Self-Driving Cars Using a Hybrid Approach. *Electronics*, 2023, vol. 12, no. 13, article no. 2768. DOI: 10.3390/electronics12132768.

14. Padilla, R., Netto, S. L. & da Silva, E. A. B. A Survey on Performance Metrics for Object-Detection Algorithms. *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 2020, pp. 237-242. DOI: 10.1109/IWSSIP48289.2020.9145130.

15. Tsekhmystro, R., Rubel, O. & Lukin, V. Study of methods for searching and localizing objects in images from aircraft using convolutional neural networks. *Radioelectronic and Computer Systems*, 2024, vol. 109, no. 1, pp. 87-98. DOI: 10.32620/reks.2024.1.08.

16. Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., & Savarese, S. Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 658-666. DOI: 10.48550/arXiv.1902.09630.

17. Ting, K. M. Precision and Recall. *Encyclopedia of Machine Learning*, 2010. 781 p. DOI: 10.1007/978-0-387-30164-8_652.

18. Lian, J., Yuhang, Y., Linhui, L., Zhenghao, W. & Yafu, Z. Small Object Detection in Traffic Scenes Based on Attention Feature Fusion. *Sensors*, 2021, no. 9, article no. 3031. DOI: 10.3390/s21093031.

19. Rahman, S., Rony, J. H., Uddin, J. & Samad, M. A. Real-Time Obstacle Detection with YOLOv8 in a WSN Using UAV Aerial Photography. *Journal of Imaging*, 2023, vol. 9, no. 10, article no. 216. DOI: 10.3390/jimaging9100216.

20. Girshick, R. Fast R-CNN. *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1440-1448. DOI: 10.1109/ICCV.2015.169.

21. Zhu, P., Wen, L., Du, D., Bian, X., Fan, H., Hu, Q., & Ling, H. Detection and Tracking Meet Drones Challenge. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, vol. 44, no. 11, pp. 7380-7399. DOI: 10.1109/TPAMI. 2021.3119563.

22. Du, D., Qi, Y., Yu, H., Yang, Y., Duan, K., Li, G., Zhang, W., Huang, Q., & Tian, Q. The Unmanned Aerial Vehicle Benchmark: Object Detection and Tracking. *European Conference on Computer Vision (ECCV)*, 2018, vol. 128, pp. 1141-1159. DOI: 10.1007/s11263-019-01266-1.

23. Wallach, D. & Goffinet, B. Mean squared error of prediction as a criterion for evaluating and comparing system models. *Ecological Modelling*, 1989, vol. 44, pp. 299-306. DOI: 10.1016/0304-3800(89)90035-5.

24. Ruby, U., Theerthagiri, P., Jacob, J. & Vamsidhar, Y. Binary cross entropy with deep learning technique for Image classification. International Journal of *Advanced Trends in Computer Science and Engineering*, 2020, vol. 9, no. 4, pp. 5393-5397. DOI: 10.30534/ijatcse/2020/175942020.

25. Mao, A., Mohri, M. & Zhong, Y. Cross-Entropy Loss Functions: Theoretical Analysis and Applications. *Proceedings of the 40th International Conference on Machine Learning*, 2023, vol. 202, pp. 23803-23828. DOI: 10.48550/arXiv.2304.07288.

26. Robinson, H. & Monro, S. A Stochastic Approximation Method. *The Annals of Mathematical Statistics*, 1951, vol. 22, no. 3, pp. 400-407. DOI: 10.1214/aoms/1177729586.

27. Li, X., Wang, W., Wu, L., Chen, S., Hu, X., Li, J., Tang, J. and Yang, J. Generalized focal loss: learning qualified and distributed bounding boxes for dense object detection. *Proceedings of the 34th International Conference on Neural Information Processing Systems (NIPS '20)*, 2020, article no. 1763. DOI: 10.48550/arXiv.2006.04388.

Надійшла до редакції 29.04.2024, прийнята до опублікування 17.06.2024

ДОСЛІДЖЕННЯ ЗАЛЕЖНОСТІ ТОЧНОСТІ ПОШУКУ АВТОМОБІЛІВ ВІД РОЗМІРІВ ОБ’ЄКТА ЗА ЗОБРАЖЕННЯМИ БПЛА

Р. В. Цехмистро, О. С. Рубель, В. В. Лукін

Цифрові зображення все частіше використовуються для задач аналізу різних типів об’єктів, їх локалізації та класифікації. Сфер використання цієї інформації досить багато, часто її використовують для систем спостереження, автоматичного водіння транспорту чи дослідження нових територій. Разом з цим

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

існує досить велика кількість нейронних мереж, що дозволяють реалізувати даний функціонал шляхом їх навчання використовуючи набори даних різних типів та класифікацій. Дуже часто для задач дослідження використовують набори даних, що зроблені за допомогою безпілотних літальних апаратів. Такі набори даних дозволяють розпізнавати різного роду об’єкти, не маючи до них прямого доступу, що дає змогу безпечно досліджувати різні території. Використання безпілотних літальних апаратів досить поширено в наш час, особливо в задачах фотографії та відео, багато фотографів використовують безпілотні літальні апарати для отримання знімків краєвидів, а також системи автоматичного стеження для переміщення. Системи автоматичного переміщення, а також системи пошуку об’єктів, досить чутливі до розміру об’єкта, а також якості алгоритму пошуку. Саме через широку застосовуваність даної задачі, а також малої кількості початкових даних, **темою** роботи є дослідження залежності точності локалізації та класифікації об’єктів від їх площі на знімках, що отримані за допомогою безпілотних літальних апаратів. Основну **увагу привернуто** до якості роботи нейронних мереж, що дозволяють отримувати інформацію про об’єкти, а також дослідження за допомогою отримання статистичних даних та тестовому наборі даних залежності точності детектування від розміру об’єкта. **Метою роботи** є отримання статистичних даних щодо точності локалізації та класифікації в залежності від розміру об'єкта, а також визначення порогів точності з використанням отриманих статистичних даних. **Задача роботи -** це навчання розповсюджених нейронних мереж з відкритою архітектурою на наборі даних, що отримані за допомогою безпілотних літальних апаратів, та визначення їх характеристик, зокрема залежності точності розпізнавання від розміру об’єкта. Очікуваний **результат роботи -** це порогові значення розміру об’єкта, що допустимі для достатньо точної класифікації та локалізації об’єктів, а також метрики якості роботи досліджуваних нейронних мереж. В результаті роботи буде наведено **висновки**, що відображають порогові значення розмірів об’єктів, від якого залежить точність розпізнавання.

Ключові слова: локалізація об’єктів; YOLOv5s, SSD, FasterRCNN, класифікація техніки, безпілотні літальні апарати.

Цехмистро Ростислав Вікторович – асп. каф. інформаційно-комунікаційних технологій ім. О. О. Зеленського, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

Рубель Олексій Сергійович – канд. техн. наук, доц. каф. інформаційно-комунікаційних технологій ім. О. О. Зеленського, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

Лукін Володимир Васильович – д-р техн. наук, проф., зав. каф. інформаційно-комунікаційних технологій ім. О. О. Зеленського, Національний аерокосмічний університет ім. М.Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

Rostyslav Tsekhmystro – PhD Student of the Department of Information and Communication Technologies named after O. O. Zelensky, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine,
e-mail: rostyslav.tsekhmystro@gmail.com, ORCID: 0000-0002-1515-7065.

Oleksii Rubel – Candidate of Technical Science, Associate Professor at the Department of Information-Communication Technologies named after O. O. Zelensky, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine,
e-mail: s.rubel@khai.edu, ORCID: 0000-0001-6206-3988.

Vladimir Lukin – Doctor of Technical Sciences, Professor, Head of the Department of Information-Communication Technologies named after O. O. Zelensky, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine,
e-mail: v.lukin@khai.edu, ORCID: 0000-0002-1443-9685.