BULLETIN of Cherkasy State Technological University

Journal homepage: https://bulletin-chstu.com.ua/en

Vol. 29 No. 3. 2024

UDC 004.932.4
DOI: 10.62660/bcstu/3.2024.21

Classification of military equipment based on computer vision methods

Oleh Basystiuk*

Assistant
Lviv Polytechnic National University
79000, 12 Stepan Bandera Str., Lviv, Ukraine
https://orcid.org/0000-0003-0064-6584

Zoriana Rybchak

PhD in Technical Sciences, Associate Professor
Lviv Polytechnic National University
79000, 12 Stepan Bandera Str., Lviv, Ukraine
https://orcid.org/0000-0002-5986-4618

Dmytro Betsa

Student
Lviv Polytechnic National University
79000, 12 Stepan Bandera Str., Lviv, Ukraine
https://orcid.org/0009-0006-6582-8900

Abstract. Means of high-altitude reconnaissance, in particular satellites, reconnaissance drones and aviation complexes, are the most common means for solving the tasks of search and detection of targets. This work focuses on improving the process of finding and identifying targets by implementing an automatic search system using artificial intelligence, with a special emphasis on the use of this technology in drones, under conditions of limited computing resources. The purpose of the work was to create a machine learning model that would localise and classify military equipment using images obtained from unmanned aerial vehicles. Machine learning models used to localise objects in images based on CNN, ResNet, Fast CNN, EfficientDet and YOLO approaches are the research methods. Various computer vision approaches, based on convolutional networks, to localise and classify military equipment in images obtained from unmanned aerial vehicles have been investigated. The approach based on the YOLO8 method has proved to be the most effective one. The generalised precision of the proposed model of image segmentation technique is 70%, and the classification precision is close to 90%, the inference time of the proposed model is less than 400 milliseconds. The system takes an image as input and returns the input image with the found military equipment. In addition, the YOLO8 (nano, small, medium) methods have been tested in the problem of equipment identification and classification in images from unmanned aerial vehicles. The approach proves to be effective and has the potential for further application as well as improvement with larger sets. The system can be used in practice to optimise the search for targets, thus simplifying the task for the operator of unmanned aerial vehicles. Also, in the case of further refinement and optimisation for specific hardware resources, it has the potential for implementation in the real defence sector. Potentially, this solution can become an important tool for military intelligence and other related industries, where precise identification of objects in real-time images is important. The implementation of such systems can significantly increase the efficiency and speed of response in various scenarios of the use of unmanned aerial vehicles

Keywords: localisation; identification; labelling; convolutional neural network; image analysis

Article's History: Received 03.07.2024; Revised 16.09.2024; Accepted 21.10.2024

Suggested Citation:
Basystiuk, O., Rybchak, Z., & Betsa, D. (2024). Classification of military equipment based on computer vision methods. *Bulletin of Cherkasy State Technological University*, 29(3), 21-30. doi: 10.62660/bcstu/3.2024.21.

*Corresponding author

Copyright © The Author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/)

INTRODUCTION

The problem of pattern identification in images is relevant in many industries, including security, medicine, automotive industry and military technology. The identification of military equipment, in particular, has not been sufficiently investigated and described, because it was not of such public and scientific interest, besides, it was a completely non-trivial task. But since the 2000s, humanity has witnessed the rapid technological evolution of object detection and its profound impact on the entire field of computer vision, and therefore object detection, as one of the most fundamental and complex problems of computer vision, is given great attention of scientists.

The article by Z. Zou *et al*. (2023) examines this fast-growing field of research in detail in light of technical evolution spanning more than a quarter of a century (from the 1990s to 2022). The study covers a number of topics, including milestones of detector historical development, datasets, metrics, fundamental building blocks of an identification system, acceleration techniques, and state-of-the-art identification techniques in use. The use of camouflage and active countermeasures during reconnaissance and obtaining results about the equipment are factors that further complicate detection. The ability to precisely identify and classify equipment can be critical to tactical and strategic decisions.

The current state of development of computer vision technologies demonstrates significant achievements due to the rapid development of machine learning and artificial intelligence technologies. J. Liang (2020) has noted that the introduction of residual blocks into the ResNet network allows the network to be further deepened without suffering from performance degradation, i.e. avoiding overtraining. This approach simplifies the training of very deep networks, making possible to build models with hundreds or even thousands of layers, which was previously impractical due to the problem of vanishing gradient. This breakthrough has become a fundamental technique in many complex image identification tasks and has influenced the development of many subsequent deep learning architectures.

It also makes significant progress in solving the problem of pattern identification in images. Recent research focuses on the development of new algorithms and models that can provide high precision and efficiency of object identification in various conditions. In particular, A. Bochkovskiy *et al*. (2020) have focused on the study of activation functions, namely batch normalisation, Weighted-Residual-Connections (WRC), Cross-Stage-Partial-connections (CSP), Cross mini-Batch Normalisation, CmBN), Self-Adversarial-Training (SAT), and have achieved results with an average precision (AP) of 43.5% and an AP50 of 65.7% on the MS COCO dataset, while maintaining real-time processing speed of about 65 frames per second on the Tesla V100.

The article by Z.-Q. Zhao *et al*. (2019) considers an approach to object detection based on deep learning and describes in detail their network architecture, learning strategies, and optimisation functions, in particular, the application of You Only Look Once (YOLO) approaches, covering typical common architectures of object detection, modifications and enhancements to improve performance. In addition, the article considers specific tasks of object detection in the image, in particular, using video. The experimental analysis compares different methods, drawing meaningful conclusions and proposing promising directions and tasks for future research in the field of object detection and neural network-based learning systems.

The work by D. Zhang *et al*. (2024) presents a new imaging paradigm for detecting camouflaged objects using RGB-D images. To accomplish this task, the authors have created a CODD dataset derived from existing RGB-D visible object detection datasets using image-to-image transformation techniques (Yang *et al.*, 2023). CODD matches the diversity and complexity of widely used datasets for camouflaged object detection. Researchers G.-P. Ji *et al*. (2023) propose a selection strategy to ensure high image quality, taking into account the structural similarity between images before and after transformation, based on gradient-induced transition, which is a soft grouping between context features and similarity of object appearance and the background, and the fuzziness of object boundaries is also considered.

Identification of military equipment has not been sufficiently investigated in previous research. All of the above studies are oriented towards performing tasks in the civilian sector, and do not include the ability to identify and classify military equipment in the image. Therefore, the purpose of this study was to create a deep learning model for identifying military equipment, taking into account camouflage and counteraction technologies obtained from unmanned aerial vehicles.

MATERIALS AND METHODS

Modern technologies of machine learning and computer vision open up new opportunities for effective monitoring and classification of objects in the image. This article is devoted to a comparative analysis of various deep learning models, including ResNet, CNN, EfficientDet, YOLOv8 (nano, small, medium). The efficiency and possibilities of integration into automated monitoring systems have been investigated.

To work with machine learning methods, the primary task is to collect and label the appropriate dataset. In the case of equipment identification, this is the main means of imaging, since the modern Ukrainian-Russian war, which has been ongoing with varying intensity since 2014, is the largest source of the required data. So, to search for images, it is worth focusing on various social networks, namely Oryx, GeoConfirm and Twitter, where military analysts, journalists, fighters, as well as the brigade show the results of their units' work. All collected data have been summarised and published in the open access on the platform of Figshare (2024). A generalised order of image processing consists of several main stages, each of which plays an important role in ensuring precise and efficient image processing. The study considers four key steps (Fig. 1).

Preprocessing → Object detection → Object classification → Result

Figure 1. Generalised order of the image processing pipeline by computer vision systems of the proposed system

Source: authors' development

Preprocessing is the initial stage of image processing, which includes image preparation for further analysis. Such image preparation helps to improve its quality and make it more suitable for further processing stages. The main methods used for image preprocessing are: image alignment and normalisation, noise filtering, lighting correction. Object detection is the process of determining the location of objects in the image, that is, in other words, it is the segmentation of the image and the localisation of objects in it. Effective object detection is critical for successful classification and further analysis. Object classification includes identification and determination of the type of object that has been detected in the image at previous steps of the processing pipeline. The classification allows for identification of objects and specific labelling of each of detected objects in the image, according to the classes provided for the search using the designed system. The classification adds a more situational understanding and will also make it possible to better prioritise the tasks of defeating various enemy targets (Benmhahe & Chentoufi, 2021). Results are a process that is responsible for the preparation and correct presentation of the obtained results, which enables the data to be used later by users or other systems for further processing or decision-making.

It is also necessary to determine what types of videos/photos to choose for analysis. In general, first of all, the choice is between night and day videos, as well as between their different types. Optical cameras (standard cameras) are the most common ones and capture optical image (Panwar *et al.*, 2023). Infrared cameras (thermal cameras) capture thermal radiation of objects and build a heat map of the battlefield (Ippalapally *et al.*, 2020). There are many types of images, but they are not practically used in UAVs (unmanned aerial vehicles) and are irrelevant in this study. Infrared cameras capture images in a completely different way than optical cameras (Park & Lee, 2021), in them it is much easier to identify equipment and weapons, so for these cameras the research itself is not very relevant. Images from different types of cameras are shown in Figure 2.

Figure 2. An example of an image from different types of cameras

Source: authors' development

The described leaves only optical cameras. These cameras are mounted on almost all UAVs (Chen *et al.*, 2024), and are the main eyes of the front. And it is for them that this study is the most relevant, since it is quite difficult for the human eye to see the camouflaged equipment, it requires great skills and knowledge. For image processing, this study has used the YOLO model, as it provides a high speed of object real-time detection and classification. Due to its architecture, YOLO processes the image only once, which significantly reduces the computation time.

RESULTS

There are many methods for image processing and object detection in the model used in the study. Convolutional Neural Networks (CNNs) are the basis for many computer vision tasks. They work by training a network on a large number of images, automatically finding and highlighting important features for object classification and detection. Histograms of Oriented Gradients (HOG) are a method for detection of image features to identify objects, which is used to describe the textures and shapes of objects in an image (Zhou *et al.*, 2020). It analyses pixel brightness gradients and creates histograms that indicate the direction and intensity of changes in the image. You Only Look Once (YOLO) is a modern method for simultaneous detection and classification of objects in the image (Hussain, 2023). Speed is the main advantage of YOLO, as it processes the image only once, unlike other methods that may require multiple passes. For a better understanding, a comparative characteristics and identification precision of various researched methods are given (Table 1).

Table 1. Comparative table of results

Method name	Precision (%)	F1 measure	Model size (MB)
ResNet	66.4	0.86	46
CNN	56.8	0.64	50
EfficientDet	63.4	0.82	32
YOLOv8m	71.8	0.91	54

Notes: the results obtained during the study may differ in other systems and vehicles
Source: authors' development

Additionally, a system deployment diagram that shows how various components interact in a real environment, including hardware and software modules is provided. The proposed system consists of the following components: communication networks, Backend and Database (Fig. 3).

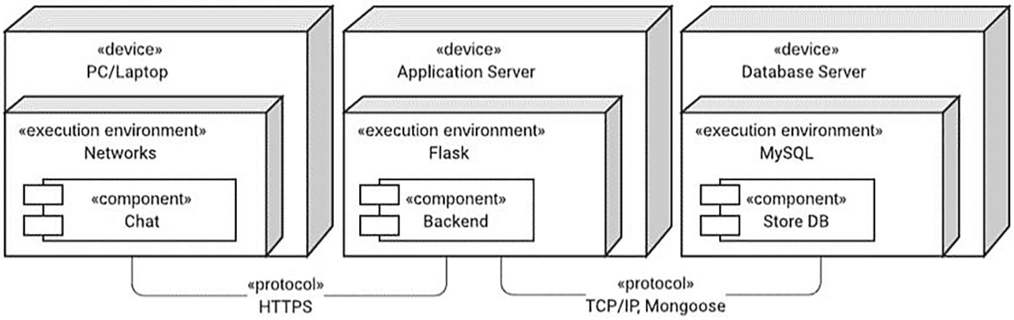

Figure 3. Diagram of the proposed system deployment
Source: authors' development

During the experiments, it has been found that adding of some context to the classification network significantly increases the classification precision and almost eliminates the localisation problem. Thus, the main attention has been focused on the precision of the detection of artillery and military vehicles (Fig. 4).

Class	Images	Instances	Box(P	R	mAP50
all	203	203	0.718	0.61	0.637
artillery	203	30	0.706	0.4	0.43
vehicle	203	173	0.73	0.821	0.845

Figure 4. Precision of the model for two classes of equipment
Source: authors' development

As shown, the precision of artillery detection is significantly lower (more than twice) compared to the precision of military vehicle detection. This can be explained by the quality of images of towed artillery: most images of artillery have low resolution and are fuzzy due to the specificity of its use, namely from closed firing positions and nests. In Figure 5, gun muzzle can be identified on the right, while only a silhouette of the gun is visible on the left, making it difficult to identify the artillery. There is also a problem of sufficient saturation of the dataset with different types of images, in particular with the use of various fortifications and caponiers for equipment. As for military equipment, the results are much better. When testing the neural network in typical images, high results have been obtained. Stroke precision is slightly lower, but this is not critical, since detection precision is the main parameter (Fig. 6).

a

b

Figure 5. Examples of fortification nests of towed artillery
Source: authors' development

Figure 6. Examples of equipment localization in a test dataset
Source: authors' development

The obtained model shows high results of identification precision, but there is a problem with sufficient amount of data. There is a big problem with the variety of equipment forms in the images and with the variety of environments. At first glance, the network appears to find most equipment, but in fact it only detects it in certain image formats. Manual testing has shown that the network works well with images where equipment is not too dark and where certain structural elements are visible. This is due to insufficient number of similar images in the dataset. The problem has been partially solved by increasing the number of such images in the dataset. Currently, the model finds about 60% of small objects. Other similar problems have been corrected by analysing the images and supplementing the dataset. For example, an image of tanks from the side projection has been added (Fig. 7).

Figure 7. Results of equipment classification from the side projection
Source: authors' development

During the study, it has been found that smoke is an important artifact that significantly impairs the performance of the identification model. Not only does smoke reduce the precision of equipment identification in images where it is present, but often thick clouds of smoke are classified as equipment. This creates significant obstacles for precise identification of military objects, which is especially critical in the conditions of hostilities. Examples of identification of a military equipment convoy in the video of the movement of a convoy of vehicles in a smoke screen demonstrate these difficulties.

There are differences between civilian and military vehicles that affect the process of their identification. Military equipment is often camouflaged in every possible way to minimise its detection by reconnaissance devices, especially drones. This is done with the help of various camouflage nets, forest plantations or by hiding equipment in dense terrain. Civilian vehicles are characterised by distinctive features such as visible wheels and bodywork, which make them easy to identify, even when they are partially covered. In the case of military equipment, there are no clear patterns for identification; often only individual pieces of equipment are visible, such as a barrel that can belong to either a tank gun or towed artillery, or a turret that can belong to either an infantry fighting vehicle (IFV) or a tank. This greatly complicates the identification of the type of equipment.

Additional difficulties are created by active countermeasures from the enemy, who tries to damage the operation of reconnaissance vehicles and shoot them down, in particular, UAVs. For this, means of Electronic Warfare (EW) and Air Defence (AD) are used, which significantly limits the efficiency of UAVs. EW tools are used to create electronic interference that can affect the operation of navigation systems, communication systems and UAV control (Zhang *et al.*, 2024). This can result in loss of communication with the vehicle, its disorientation or even complete loss of control over it.

In view of this, the study has been aimed at developing two models: classification and localisation ones, which have been considered separately. The localisation model has two classes to search for: vehicle and artillery. The YOLO8s neural model, which has demonstrated optimal identification results, has been chosen. The results of this model are presented in detail in Figure 8.

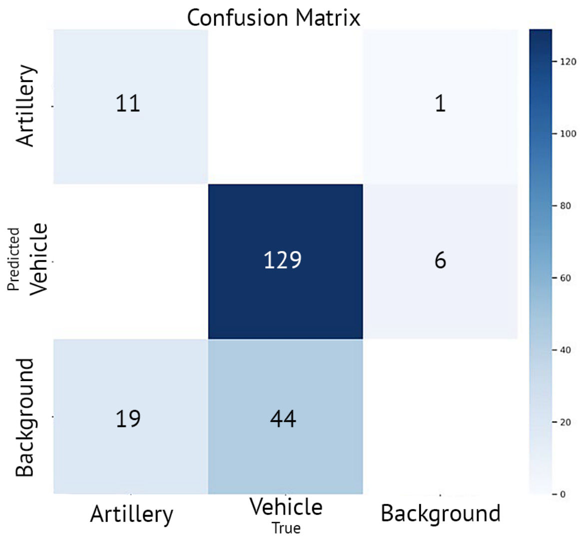

Figure 8. Matrix of uncertainties of the proposed YOLO model
Source: authors' development

The obtained results show that the proposed model has achieved a precision of 71.8%, which exceeds the indicators of other considered models. The value of F1-measure for the proposed model has been 0.91, which indicates a high balance between the precision and completeness of identification (Fig. 9).

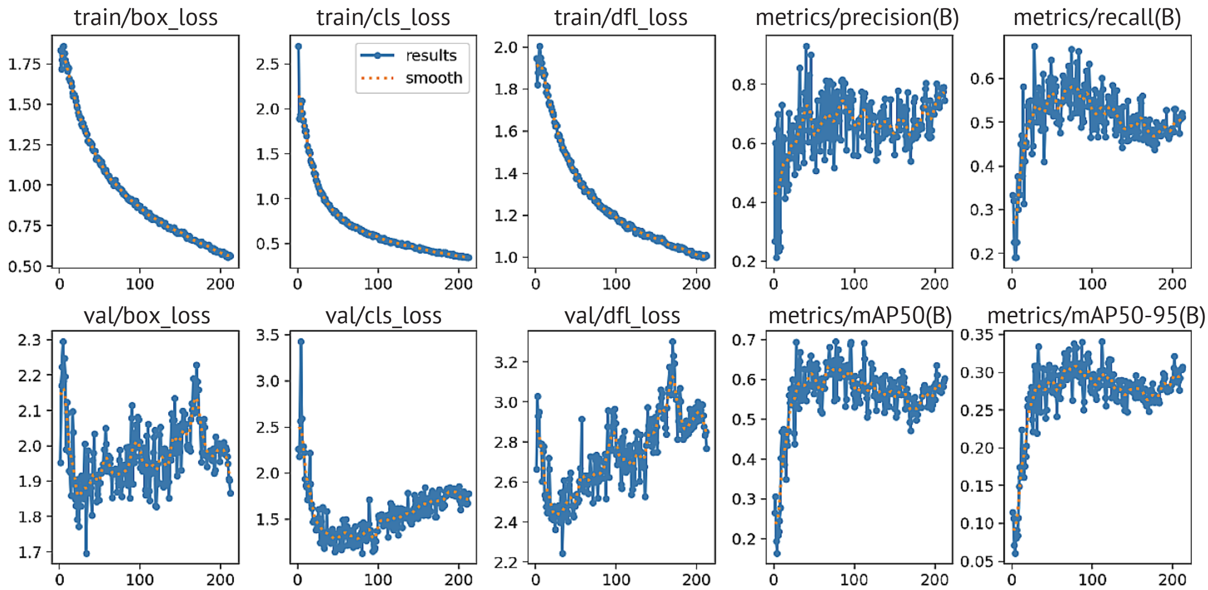

Figure 9. Results of the proposed YOLO model training

Source: authors' development

Precision is based on two indicators: precision of vehicle or artillery identification, as well as precision of equipment boundaries. For this work, the precision of classification is much more important than that of localisation. The results of the study confirm that the model is able to identify fragments of military equipment, such as artillery barrels and turrets of combat vehicles, which is important for precise identification of the type of equipment.

The proposed model demonstrates a high level of precision and efficiency in the tasks of identifying military equipment under camouflage conditions. The results of the study show that the model can be used to improve the effectiveness of intelligence operations and provide more precise detection of threats, which is critical for ensuring national security and defence.

DISCUSSION

The model proposed in this study will combine modern deep learning methods with new algorithms to increase the precision and reliability of military objects detection and classification. This is especially true in conditions where the enemy uses various methods to complicate the detection of its equipment, including EW and AD equipment.

It is worth noting that the images obtained from a multispectral camera capture images in different ranges (visible light, ultraviolet radiation, etc.). A comprehensive study on the topic of multispectral imaging has been conducted by V. Mohammadi *et al.* (2022). The authors consider a camera sensor consisting of a multispectral image sensor, a spectral filter matrix, a sensor board, and a control board. A hybrid system that operates using eight bands in the visible range, in the interval of 400-700 nm is proposed. Based on a genetic algorithm, a program has been developed to find the best combination of filters. The program selects Gaussian filters using the genetic algorithm based on the Wiener filter estimation method. For band selection, a minimum RMS of 0.0016 has been obtained to select bands in the visible range. The developed camera provides eight high-resolution spectral images.

The article by L. Alzubaidi *et al.* (2021) describes the sequence of CNNs development, starting with the AlexNet network and ending with the High-Resolution Network (HR.Net). Also a list of problems and standardised solutions that have helped researchers to better understand existing research gaps by applying pre-defined techniques and approaches is provided. In addition, a list of the main areas of application of the proposed approaches is provided, and hardware requirements for each of the proposed models are given. Considering the further development of convolutional neural networks, namely the EfficientDet approach, M. Tan *et al.* (2020) propose a weighted bidirectional feature pyramid network (BiFPN) that allows easy and fast fusion of multi-scale features, and a combined scaling method that uniformly scales resolution, depth, and width for all base networks, feature networks, and class prediction networks simultaneously. Collectively, all these improvements have resulted in the development of a new family of object detectors, called EfficientDet, which consistently achieve much better efficiency. In particular, when using the EfficientDet-D7 model to compare with convolutional networks, the former achieves 52.2 AP on the COCO test device with 52M and 325B FLOPs1, being 4-9 times smaller and using 13-42 times less FLOPs than that of CNN (Kishore & Balamurugan, 2024).

G. Csurka *et al*. (2022) summarise two decades of image segmentation research. Since current segmentation models require large amounts of annotated data, which are expensive, the authors highlight the success of Unsupervised Domain Adaptation (UDA), which allows for a significant reduction in data labelling costs by using domain-based image clustering. A review of five years of experience with Domain Adaptation for Semantic Image Segmentation (DASiS), which highlights the importance of adapting segmentation models to new conditions, is the second key contribution. Finally, commonly used datasets and benchmarks are described, and supporting machine learning techniques, namely the latest trend in the use of transformers, are discussed. Another method involves the use of the Gray Level Co-occurrence Matrix (GLCM) method, which has been analysed in detail in the work of N. Iqbal *et al*. (2021). The GLCM method, analysing spatial relationships between pixels, allows for extraction of valuable information for such tasks as texture classification and segmentation. Experimenting with various parameters and functions, it is possible to fine-tune the analysis to meet specific application requirements, which makes this method a universal tool in the field of image processing and computer vision.

R. Gavrilescu *et al*. (2018) consider the Faster Regional based Convolutional Neural Network (Faster R-CNN) algorithm. Faster R-CNN is the result of merging of Region Proposal Network (RPN) and Fast-RCNN algorithms into one network. To increase video processing power, a Graphics Processing Unit (GPU) has been used to train and test real-time object identification at 15 frames per second in a dataset containing 3000 images for 4 classes. The dataset consists of images containing three traffic light phases. The use of the You Only Look Once (YOLO) approach, which has been also used in the above study, and its advantages and disadvantages have been described by M. Sohan *et al*. (2024). The authors have reviewed the YOLOv8 model, which is an improved version for real-time object detection. Of all popular methods for object identification and machine learning models, such as Faster RCNN, SSD, and RetinaNet, YOLO is the most well-known method in terms of precision, speed, and efficiency. This study presents an analysis of YOLO v8, highlighting its innovative features, improvements, applicability in various environments, and a detailed comparison of its performance indicators with other versions and models. Additionally, the authors A. Vijayakumar & S. Vairavasundaram (2024) describe the stages of development of deep learning methods over the decades, emphasise the relevance of the topic, and state that most researchers are working on improving the detection, segmentation and classification of objects.

In the conducted research, the application of technology for military equipment identification, which has not been proposed in any of the reviewed studies, is a unique feature of the final development. This technology is based on the integration of modern approaches to image analysis, which allows for effective identification of military equipment in conditions of limited computing resources. The approach uses advanced deep learning techniques, including neural networks optimised for use in unmanned aerial vehicles. During the research, an effective selection strategy to ensure high image quality has been found. This strategy takes into account the structural similarity between images before and after transformation, based on gradient-induced transition. This approach provides a soft grouping between context features and similarity of object appearance and the background. Taking into account of the fuzziness of object boundaries within the framework of the developed system allows for increase in the precision of identification even in difficult conditions.

The use of an image paradigm to detect camouflaged objects using RGB-D images is one of key innovations in the conducted research. This approach allows for a significant improvement in the precision of camouflaged equipment detection, as it includes additional information about the depth of the scene. The use of RGB-D images in combination with developed data processing algorithms allows for effective separation of camouflaged objects from the background, which is important in the context of military applications. The proposed system has been tested in an optical camera with 24 frames per second, with the possibility of increasing to 60, and a response time within 400 milliseconds. These characteristics provide a significant advantage compared to other considered systems, especially in complex conditions of battlefield images and when detecting camouflaged equipment. The high frame rate and fast response time allow the system to work in real time, which is critical for prompt detection and identification of military objects.

CONCLUSIONS

Within this study, a wide range of architectural solutions based on neural networks for searching for objects in images has been analysed. In particular, various approaches to solving the task of finding equipment in the image have been considered, ultimately choosing localisation and identification. Various approaches for object identification, including one-stage and two-stage methods, have been investigated. A dataset from identified sources has been collected and models for object localisation and classification using YOLOv8 have been trained.

Despite significant limitations in the amount of data, the models show improved results compared to known and publicly available solutions. For image formats well represented in the dataset, the network demonstrates 80% of localisation precision and 70% of classification precision. Overall, this approach has proven to be effective and has the potential to be improved with more data that will make it possible to more precisely locate, segment, and classify the equipment in images. This approach to solving the problem has proven itself well, even despite significant limitations in the dataset. High precision of identification and classification of most types of images has been achieved.

The main conclusion is that if there is a sufficiently large dataset, it is possible to create a neural network that will effectively and reliably identify equipment at the battlefield. Further research and improvement of the machine learning model will take place precisely in the direction of supplementing and retraining the model on larger (newer) datasets, as well as in the way of integrating additional sources of information with various, in particular thermal, multispectral images, as well as various, in particular audio, tracks with the sound of equipment and battlefields.

ACKNOWLEDGEMENTS

None.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Alzubaidi, L., *et al*. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8, article number 53. doi: 10.1186/s40537-021-00444-8.

[2] Benmhahe, B., & Chentoufi, J.A. (2021). Automated pavement distress detection, classification and measurement: A review. *International Journal of Advanced Computer Science and Applications*, 12(8), 708-718.

[3] Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y.M. (2020). YOLOv4: Optimal speed and accuracy of object detection. *arXiv*, 2004, article number 10934. doi: 10.48550/arXiv.2004.10934.

[4] Chen, X., Liu, C., Chen, L., Zhu, X., Zhang, Y., & Wang, C. (2024). A pavement crack detection and evaluation framework for a UAV inspection system based on deep learning. *Applied Sciences*, 14(3), article number 1157. doi: 10.3390/app14031157.

[5] Csurka, G., Volpi, R., & Chidlovskii, B. (2022). Semantic image segmentation: Two decades of research. *Foundations and Trends® in Computer Graphics and Vision*, 14(1-2), 1-162. doi: 10.1561/0600000095.

[6] Figshare. (2024). *Vehicles and fortification image recognition dataset*. doi: 10.6084/m9.figshare.25726470.

[7] Gavrilescu, R., Zet, C., Foșalău, C., Skoczylas, M., & Cotovanu, D. (2018). Faster R-CNN: An approach to real-time object detection. In *International conference and exposition on electrical and power engineering (EPE)* (pp. 165-168). Iasi: IEEE. doi: 10.1109/ICEPE.2018.8559776.

[8] Hussain, M. (2023). YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection. *Machines*, 11(7), article number 677. doi: 10.3390/machines11070677.

[9] Ippalapally, R., Mudumba, S.H., Adkay, M., & Vardhan, H.R.N. (2020). Object detection using thermal imaging. In *2020 IEEE 17th India Council international conference (INDICON)* (pp. 1-6). New Delhi: IEEE India Council. doi: 10.1109/INDICON49873.2020.9342179.

[10] Iqbal, N., Mumtaz, R., Shafi, U., & Zaidi, S.M.H. (2021). Gray level co-occurrence matrix (GLCM) texture based crop classification using low altitude remote sensing platforms. *PeerJ Computer Science*, 7, article number e536. doi: 10.7717/peerj-cs.536.

[11] Ji, G.-P., Fan, D.-P., Chou, Y.-C., Dai, D., Liniger, A., & Gool, L.C. (2023). Deep gradient learning for efficient camouflaged object detection. *Machine Intelligence Research*, 20, 92-108. doi: 10.1007/s11633-022-1365-9.

[12] Kishore, A.S.K., & Balamurugan, G. (2024). Faster region based convolution neural network with context iterative refinement for object detection. *Measurement: Sensors*, 31, article number 101025. doi: 10.1016/j.measen.2024.101025.

[13] Liang, J. (2020). Image classification based on RESNET. *Journal of Physics: Conference Series*, 1634, article number 012110. doi: 10.1088/1742-6596/1634/1/012110.

[14] Mohammadi, V., Sodjinou, S.G., Katakpe, K.K., Rosse, M., & Gouton, P. (2022). Development of a multi-spectral camera for computer vision applications. In *30. jubilee international conference in central Europe on computer graphics, visualization and computer vision* (pp. 24-27). Plzen: Vaclav Skala University of West Bohemia. doi: 10.24132/CSRN.3201.4.

[15] Panwar, K., Singh, A., Kukreja, S., Singh, K.K., Shakhovska, N., & Boichuk, A. (2023). Encipher GAN: An end-to-end color image encryption system using a deep generative model. *Systems*, 11(1), article number 36. doi: 10.3390/systems11010036.

[16] Park, S., & Lee, H. (2021). Deep learning approach to optical camera communication receiver design. In *2021 IEEE region 10 symposium (TENSYMP)* (pp. 1-5). Jeju: IEEE. doi: 10.1109/TENSYMP52854.2021.9550896.

[17] Sohan, M., Sai Ram, T., & Rami Reddy, C.V. (2024). A review on YOLOv8 and its advancements. In I.J. Jacob, S. Piramuthu & P. Falkowski-Gilski (Eds.), *Data intelligence and cognitive informatics* (pp. 529-545). Singapore: Springer. doi: 10.1007/978-981-99-7962-2_39.

[18] Tan, M., Pang, R., & Le, Q.V. (2020). EfficientDet: Scalable and efficient object detection. In *IEEE/CVF conference on computer vision and pattern recognition (CVPR)* (pp. 10778-10787). Seattle: IEEE. doi: 10.1109/CVPR42600.2020.01079.

[19] Vijayakumar, A., & Vairavasundaram, S. (2024). YOLO-based object detection models: A review and its applications. *Multimedia Tools and Applications*, 83, 83535-83574. doi: 10.1007/s11042-024-18872-y.

[20] Yang, S., Jiang, L., Liu, Z., & Loy, C.C. (2023). GP-UNIT: Generative prior for versatile unsupervised image-to-image translation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(10), 11869-11883. doi: 10.1109/TPAMI.2023.3284003.

[21] Zhang, D., Wang, C., & Fu, Q. (2024). A new benchmark for camouflaged object detection: RGB-D camouflaged object detection dataset. *Open Physics*, 22(1), article number 20240060. doi: 10.1515/phys-2024-0060.

[22] Zhang, Z., Xie, X., Guo, Q., & Xu, J. (2024). Improved YOLOv7-Tiny for object detection based on UAV aerial images. *Electronics*, 13(15), article number 2969. doi: 10.3390/electronics13152969.

[23] Zhao, Z.-Q., Zheng, P., Xu, S.-T., & Wu, X. (2019). Object detection with deep learning: A review. *IEEE Transactions on Neural Networks and Learning Systems*, 30(11), 3212-3232. doi: 10.1109/TNNLS.2018.2876865.

[24] Zhou, W., Gao, S., Zhang, L., & Lou, X. (2020). Histogram of oriented gradients feature extraction from raw bayer pattern images. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 67(5), 946-950. doi: 10.1109/TCSII.2020.2980557.

[25] Zou, Z., Chen, K., Shi, Z., Guo, Y., & Ye, J. (2023). Object detection in 20 years: A survey. *Proceedings of the IEEE*, 111(3), 257-276. doi: 10.1109/JPROC.2023.3238524.

Класифікація військової техніки на основі методів комп'ютерного зору

Олег Басистюк
Асистент
Національний університет «Львівська політехніка»
79000, вул. Степана Бандери, 12, м. Львів, Україна
https://orcid.org/0000-0003-0064-6584
Зоряна Рибчак
Кандидат технічних наук, доцент
Національний університет «Львівська політехніка»
79000, вул. Степана Бандери, 12, м. Львів, Україна
https://orcid.org/0000-0002-5986-4618
Дмитро Беца
Студент
Національний університет «Львівська політехніка»
79000, вул. Степана Бандери, 12, м. Львів, Україна
https://orcid.org/0009-0006-6582-8900

Анотація. Найбільш поширеними засобами для вирішення завдань пошуку та виявлення цілей є засоби висотної розвідки, зокрема супутники, розвідувальні дрони та авіаційні комплекси. Дана робота фокусується на покращенні процесу знаходження та ідентифікації цілей шляхом імплементації системи автоматичного пошуку із використанням штучного інтелекту, з особливим акцентом на використання цієї технології на дронах, в умовах обмежених обчислювальних ресурсів. Метою роботи було створення моделі машинного навчання, яка дозволить локалізувати та класифікувати військову техніку за допомогою зображень, отриманих з безпілотних літальних апаратів. Методами дослідження є моделі машинного навчання, що використовуються для локалізації об'єктів на зображеннях, що базуються на підходах CNN, ResNet, Fast CNN, EfficientDet та YOLO. Було досліджено різні підходи комп'ютерного зору, на основі згорткових мереж для локалізації та класифікації військової техніки на зображеннях отриманих з безпілотних літальних апаратів. Найефективніше себе показав підхід, який базується на методі YOLO8. Узагальнена точність пропонованої моделі сегментації техніки на зображення становить 70 %, а точність класифікації наближається до 90 %, inference time пропонованої моделі становить менше 400 мілісекунд. Система приймає як вхідні дані зображення і повертає вхідне зображення із знайденою військовою технікою. Додатково проведено апробацію методів YOLO8 (nano, small, medium) у проблематиці розпізнавання та класифікації техніки на зображеннях з безпілотних літальних апаратів. Підхід виявився ефективним і має потенціал для подальшого застосування, а також покращення при наявності більших наборів. Система може використовуватися на практиці для оптимізації пошуку цілей, спрощуючи таким чином завдання для оператора безпілотних літальних апаратів. Також, у випадку подальшого доопрацювання та оптимізації під конкретні апаратні ресурси, має потенціал до впровадження у реальному секторі оборони. Потенційно дане вирішення може стати важливим інструментом для військової розвідки та інших суміжних галузей, де важлива точна ідентифікація об'єктів на зображеннях в реальному часі. Впровадження таких систем може значно підвищити ефективність і швидкість реагування в різних сценаріях використання безпілотних літальних апаратів

Ключові слова: локалізація; розпізнавання; маркування; згорткова нейронна мережа; аналіз зображень