ISSN 2707-1898 (print)

Український журнал інформаційних технологій
Ukrainian Journal of Information Technology
http://science.lpnu.ua/uk/ujit
https://doi.org/10.23939/ujit2024.01.041
Article received 27.09.2024 р.
Article accepted 19.11.2024 р.
UDC 004.[93+8+4]

Correspondence author
B. P. Borkivskyi
bohdan.p.borkivskyi@lpnu.ua

***B. P. Borkivskyi, V. M. Teslyuk***

*Lviv Polytechnic National University, Lviv, Ukraine*

# OPTIMIZATION OF OBJECT DETECTION IN CLOSED SPACE USING MOBILE ROBOTIC SYSTEMS WITH OBSTACLE AVOIDANCE

Introducing neural network training process modification that uses combination of several datasets to optimize search of objects and obstacles using mobile robotic systems in a closed space. The study includes an analysis of papers and existing approaches aiming to solve the problem of object boundary detection and discovers the key features of several neural network architectures. During research, it was discovered that there is an insufficient amount of data about the effectiveness of using obstacle detection approaches by mobile robotics systems in a closed space. The presented method is a combination of a deep neural network-based approach for object boundary recognition and visual data for obstacles in a closed space. The base for the object detection neural network is a Deeplab model architecture, trained on the NYU Depth Dataset V2 and ADE20K datasets with extensive data for varying scene types and object categories along with corresponding annotations. The first one consists exclusively of indoor images with precise masks even for little objects, while the latter contains both indoor and outdoor images. The paper provides the results of several conducted experiments aiming to estimate the performance and feasibility of using the developed system for various tasks with the UNet and Deeplab architectures on different datasets. The experiments determined that the developed system built utilizing the Deeplab neural network architecture trained using the combination of ADE20K scene parsing dataset and NYU Depth indoor data dataset reached an accuracy of 86.9 % in multi-object segmentation. The visual results were presented to demonstrate the obstacle detection of various objects in a closed space and to compare object detection accuracy of several approaches in different situations, like an empty apartment with random obstacles on the way or a crowded study space at a university. The results show an excellent distinction between room sides like ceiling and floor, walls and doors, as well as detecting people and pieces of furniture. The benefits of this study constitute that the proposed neural network training process modification facilitates precise obstacle detection in a closed space by mobile robotic systems, provides more optimal solution to be used in navigation component of such systems by gathering information about precise objects outlines and constructing more optimal path to destination, and is an effective approach to solving the assigned tasks in various environments.

***Keywords:*** image processing, convolutional network, pattern recognition, image segmentation.

## Introduction / Вступ

In recent years, the introduction of mobile robotic systems has been observed in various fields, ranging from industrial automation to service robotics and autonomous vehicles. Central to the functioning of such systems is their ability to perceive and interact with the environment, which is often accompanied by the need to detect objects. Traditional methods of object detection typically involve defining bounding boxes around objects of interest in photo or video frames. While these methods are effective in many scenarios, they often have trouble accurately defining object boundaries and handling failures, which are common problems in real-world environments. As a result, there is a growing interest in using object segmentation methods to improve the efficiency of object detection by mobile robots.

The importance of optimizing object detection for mobile robots is that it directly affects the efficiency, reliability, and safety of robotic systems operating in dynamic and unstructured environments. In industrial environments, accurate object detection is crucial for tasks such as automated material handling, assembly, and quality control, where errors or inconsistencies can lead to significant disruptions and economic losses. In service robotics, including home care, healthcare, and retail, reliable object detection enables robots to efficiently assist humans and confidently navigate complex environments. In addition, in autonomous vehicles and unmanned aerial vehicles (UAVs), accurate object detection is essential for collision avoidance, route planning, and situational awareness, ensuring safe and efficient operation in a variety of scenarios.

Despite advances in object detection methods, challenges remain in optimizing the performance of mobile robots, especially in terms of accuracy, speed, and resource efficiency. Traditional object detection methods often have trouble accurately detecting small or enclosed objects, distinguishing between similar objects, and adapting to changes in lighting conditions or environmental factors. In addition, the computational and memory requirements of modern object detection models pose challenges for deployment on resource-constrained mobile platforms such as robotic drones or autonomous vehicles. By using object segmentation methods and optimizing their integration with object detection pipelines, researchers and practitioners are striving to solve these problems and unlock the full potential of mobile robots in various fields. Accordingly, optimization of obstacle recognition in mobile systems is an urgent task of today.

*Object of research* – the process of object recognition in a closed environment for autonomous mobile systems.

*Subject of research* – methods and means of object recognition with finding the contours of objects in an enclosed environment for autonomous mobile systems.

*The purpose of the research* – to increase the efficiency of object recognition in an enclosed environment for autonomous mobile systems with obstacle avoidance by finding clear boundaries of obstacles.

To achieve this purpose, the following main research objectives are identified:

Conduct a literature analysis of methods and means of object recognition.

Choose the optimal architecture of a neural network for object detection.

Select a data set for training the neural network and develop a model.

Conduct training experiments with different combinations of architectures and datasets.

Evaluate the accuracy of the selected algorithms.

***Analysis of recent research and publications.*** Researchers from various fields are constantly improving approaches to solving the problems of controlling and navigating robotic systems. In particular, researchers in the field of computer networks are working to improve communication with robotic systems [1][2], and specialists in the field of artificial intelligence, namely vision, are constantly improving approaches to solving object recognition problems [3], [4], and there are several that have been widely used around the world due to their effectiveness in solving applied problems. However, all algorithms and models have both strengths and weaknesses that need to be investigated to determine which solution is best suited for the task at hand.

DeepLab v3 is a powerful convolutional neural network (CNN) model designed for semantic image segmentation, which involves labeling each pixel of an image with a corresponding class. Developed by Chen et al. in 2017 [5], DeepLab v3 leverages the work of its predecessors, including advanced convolution and atropy spatial pyramid pooling (ASPP) modules, to efficiently capture contextual information at different scales. The DeepLab v3 architecture consists of a backbone network, typically a pre-trained convolutional network such as ResNet [6], [7] or Xception [8], followed by multiple atros convolutional layers and ASPP modules. These components allow the model to effectively understand both local and global context, making it very accurate in segmenting objects of different sizes in an image.

In terms of performance, DeepLab v3 demonstrated high accuracy and efficiency on a variety of sample datasets as well as in practical applications. It achieved high results in semantic segmentation tasks, significantly outperforming previous methods. By using advanced convolution and ASPP modules, DeepLab v3 can capture fine details and contextual information, which is crucial for accurate segmentation even in large-scale images. In addition, DeepLab v3 is capable of efficiently processing high-resolution images, making it suitable for applications such as autonomous driving, satellite image analysis, and medical imaging.

UNet is a widely used convolutional neural network (CNN) architecture primarily designed for biomedical image segmentation tasks. Proposed by Ronneberger et al. in 2015 [9], the UNet architecture consists of a gradually narrowing path followed by a gradually expanding path. The narrowing path, reminiscent of a traditional convolutional network architecture, captures contextual information through a sequence of layers, while the expanding path facilitates precise localization through scaling operations and result pooling. This design enables UNet to efficiently segment objects in images with high accuracy, making it particularly suited for medical imaging applications where precise delineation of structures such as organs, tumors, or cells is critical.

UNet has demonstrated impressive results in a variety of biomedical image and dataset segmentation tasks [10]. High accuracy is facilitated by its ability to capture both local and global contextual information by skipping tapering and expanding connections between parts. In addition, the UNet architecture allows for efficient learning even with a limited amount of annotated data, making it appropriate for situations where there is a lack of datasets with a high number of labels. In addition, the UNet architecture is flexible and can be customized or modified to meet specific segmentation tasks or to accommodate different input data sizes and resolutions.

PSPNet (Pyramid Scene Parsing Network) is a deep learning model designed for semantic image segmentation, introduced by Zhao et al. in 2017 [11]. One of the key innovations of PSPNet is the use of pyramidal pooling modules that capture contextual information at different scales. The architecture of PSPNet consists of a backbone convolutional neural network, such as ResNet or VGG, followed by a pyramidal pooling module. This module divides object maps into regions of different sizes and aggregates global contextual information from each region independently. By incorporating contextual information of different scales, PSPNet is able to accurately segment objects of different sizes in an image, making it particularly effective for scene parsing tasks in computer vision.

PSPNet has achieved strong results on various benchmark datasets, outperforming previous methods in terms of semantic segmentation accuracy [12]. The pyramid fusion module allows PSPNet to efficiently capture both local and global context, leading to more accurate segmentation results even for objects with complex shapes or different scales. In addition, PSPNet achieves competitive performance with relatively fewer parameters compared to other deep learning models, making it more memory efficient and suitable for deployment on resource-constrained devices or real-time applications.

HRNet (High-Resolution Network) is a deep learning architecture designed for tasks such as human pose estimation and semantic segmentation, introduced by Wang et al. in 2019 [13]. Unlike traditional convolutional neural networks (CNNs), which reduce the sampling of feature maps to reduce computational costs, HRNet maintains high resolution throughout the network. The HRNet architecture consists of parallel branches with different resolutions that simultaneously capture information at different scales. These branches are interconnected by high–resolution connections, enabling the model to capture fine detail as well as global context. By supporting high resolution, HRNet achieves superior performance in tasks requiring precise localization and detailed feature extraction.

HRNet has demonstrated strong results in a variety of computer vision tasks, including human pose estimation [14] and semantic segmentation. The ability to support high-resolution images allows HRNet to accurately capture fine details and subtle features in images, resulting in more accurate predictions. In addition, HRNet achieves competitive performance with relatively fewer parameters compared to other deep learning architectures, making it more memory efficient and suitable for deployment on resource-constrained devices or in real-time applications. In addition, the HRNet architecture is highly flexible and can be adapted or extended to accommodate different input sizes, resolutions, and task requirements.

Despite demonstrating good results in solving segmentation problems in the medical field or in the field of self-driving cars operating on the street, there is a lack of data on the effectiveness of the reviewed models in mobile robotic systems with obstacle avoidance in enclosed spaces.

## Research results and their discussion / Результати дослідження та їх обговорення

**Choosing a neural network architecture.** To evaluate the accuracy of object recognition, we propose to compare some of the reviewed architectures, namely UNet and Deeplab v3.

Although the UNet model was one of the first solutions in image segmentation, its variants are still used by the scientific community [15] and demonstrate good results in the field of semantic recognition, particularly in medicine. UNet uses pass-through connections between short and long paths, facilitating direct communication between low- and high-level functions. The collection of information from all levels looks like this:

$$z = \left[ x_1, \quad x_2, \quad \ldots, \quad x_n \right],$$

where $x_i$ – information from the $i$-th level.

This allows the model (Fig. 1) to simultaneously capture fine details and context, which leads to more accurate segmentation results, especially in scenarios with objects of different sizes and shapes.

The more modern Deeplab v3 model was introduced several years ago, and it still demonstrates impressive results [16] when performing recognition in various conditions, including in the medical field. Thanks to the use of atropic convolution modules [17] (Fig. 2) and bilinear interpolation, as well as atropic spatial pyramid pooling (ASPP), DeepLab v3 allows for efficient collection of context information at different scales. Atros convolutions are calculated by the formula:

$$y[i] = \sum_{k=1}^{K} x[i + r \cdot k] \cdot w[k],$$

where $y$ – input signal; $x$ – output signal; $w$ – filter weights; $r$ – expansion rate coefficient.

This allows the model to take into account both local and global context when predicting segmentation. DeepLab v3 is capable of creating high-resolution segmentation masks, which allows you to accurately delineate object boundaries and fine details in segmented areas. This is especially useful for tasks where precise localization and detailed segmentation are essential, such as medical image analysis and satellite image interpretation.

**Dataset selection.** The study uses two datasets for segmentation: NYU Depth Dataset V2 and ADE20K. These datasets offer a large number of annotations for a variety of scenes and object categories, and are valuable resources for further research in semantic segmentation and related fields.

The NYU Depth Dataset V2 is a widely used dataset for understanding indoor scenes and depth estimation. It was created by Nathan Silberman et al. at New York University [18]. This dataset consists of RGB-D images acquired from various indoor scenes including offices, living rooms, kitchens, bedrooms, and bathrooms. The ADE20K dataset is a large–scale scene parsing dataset designed for semantic segmentation tasks. Developed by B. Zhou et al. at the Massachusetts Institute of Technology [19], ADE20K contains a variety of scenes from both indoor and outdoor environments, including city streets, natural landscapes, and indoor spaces. The dataset consists of more than 20 thousand high-resolution images with pixel annotations for more than 150 categories of objects, such as people, animals, vehicles, buildings, and natural elements.

In general, the choice between these datasets depends on the specific requirements of the segmentation task, such as the application, scene complexity, and availability of depth information. Therefore, in this study, we conducted experiments using both datasets.

**Description of the performed experiments.** The aim of the experiments was to achieve high recognition accuracy after training neural network with certain data. After training UNet architecture with data from NYU Depth dataset with different configurations of training parameters, the accuracy of 71 % was achieved. To further improve the accuracy of system, we used another architecture Deeplab v3 and used a combination of two datasets. Firstly we trained the model using ADE20K dataset that has a big variety of images of different themes, both indoor scenes and outdoor landscapes, therefore teaching the model how to distinguish between different objects. After that we proceeded with tuning the model with domain-specific data of detailed indoor interiors provided by NYU Depth dataset, therefore improving model's understanding of the indoor objects and obstacles.

The recognition accuracy was tested on different images that reproduce the use of mobile robotic systems in different conditions – both in the absence of obstacles ahead and with obstacles of different types.

During the experiments, we compared the accuracy (Tab. 1) of three approaches: determining inaccurate boundaries (bounding boxes) of image objects using the SSD classifier, and determining accurate boundaries (segmented object masks) using the UNet and Deeplab models.

When used in a residential environment, the robotic system must be able to find and avoid objects that are atypical for this environment and that were left behind by residents (Fig. 3).

In this case, the SSD classifier accurately detected the position of two boxes in the foreground, identifying them as suitcases, and the guitar behind them. Segmentation models also detected one of the boxes, but had some difficulty with the second one.

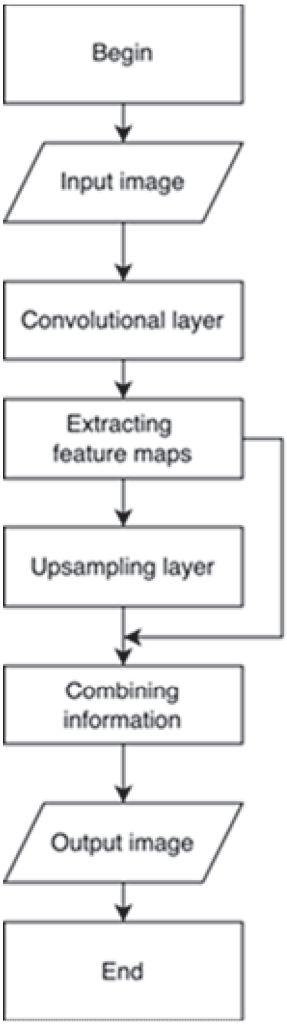


**Fig. 1.** Block diagram of the UNet model operation algorithm / Блок-схема алгоритму роботи моделі Unet

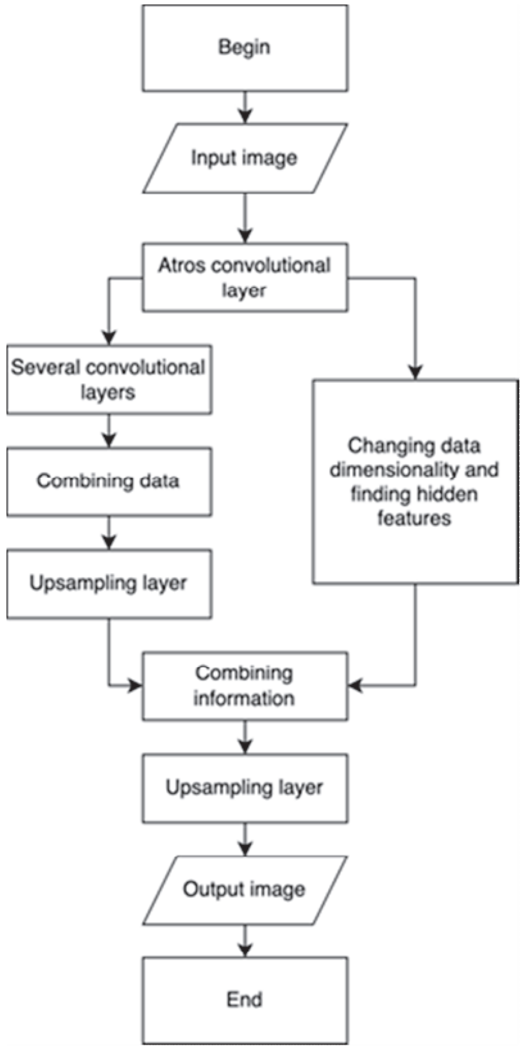


**Fig. 2.** Block diagram of the Deeplab model operation algorithm / Блок-схема алгоритму роботи моделі Deeplab

Developing a mobile robotic system for use in a single type of enclosed environment (a residential dwelling) is impractical, so it is worth checking the correctness of the developed systems in other types of environments with a large number of obstacles, such as a university classroom (Fig. 4).

The corridor may not always be completely empty, so it is necessary to check how the developed systems react to various obstacles, such as a suddenly opened door (Fig. 5).

A frequent case is the presence of a person who wants to interact with the robotic system in some way or accidentally gets into the robot's field of vision (Fig. 6).

Another typical obstacle in the classroom is student furniture, so the robotic system must not only be able to find it in front of it, but also understand whether it is obstructing further movement (Fig. 7).

***Discussion of the research results.*** Comparing our results with those of [2, 5, 10, 13], we can conclude that the developed object recognition method improves the efficiency of object search.

Despite the high accuracy of image contour recognition at 82.9 % due to the use of the SSD classification model [4], the stated solution is not able to determine the exact boundaries of the found objects, and also does not provide information about the main obstacles of the closed environment (e.g., walls), and therefore is not optimal for the task.

In [7], the authors applied label smoothing and layer weight attenuation to improve model generalization, but the accuracy of 86.6 % is inferior to our result.

**Tab. 1.** Results of models training / Результати тренування моделей

| Acrhitecture | Dataset | Accuracy (mIoU), % |
|---|---|---|
| UNet | NYU Depth | 71.4 |
| Deeplab v3 | ADE20K + NYU Depth | 86.9 |

**Fig. 3.** Recognizing obstacles in the corridor using: a – SSD classifier; b – UNET segmentation; c – Deeplab segmentation / Розпізнавання перешкод у коридорі за допомогою: а – класифікатора SSD; б – сегментації UNET; в – сегментації Deeplab

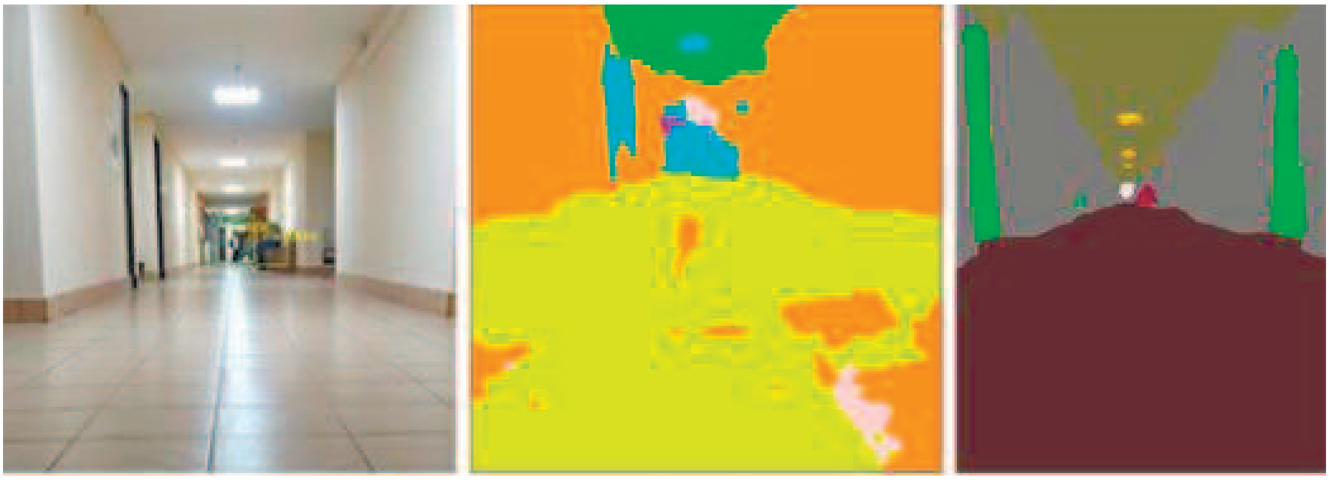

**Fig. 4.** Recognition in the empty corridor of ACS department using: a – SSD classifier; b – UNET segmentation; c – Deeplab segmentation / Розпізнавання пустого коридору кафедри АСУ за допомогою: а – класифікатора SSD; б – сегментації UNET; в – сегментації Deeplab

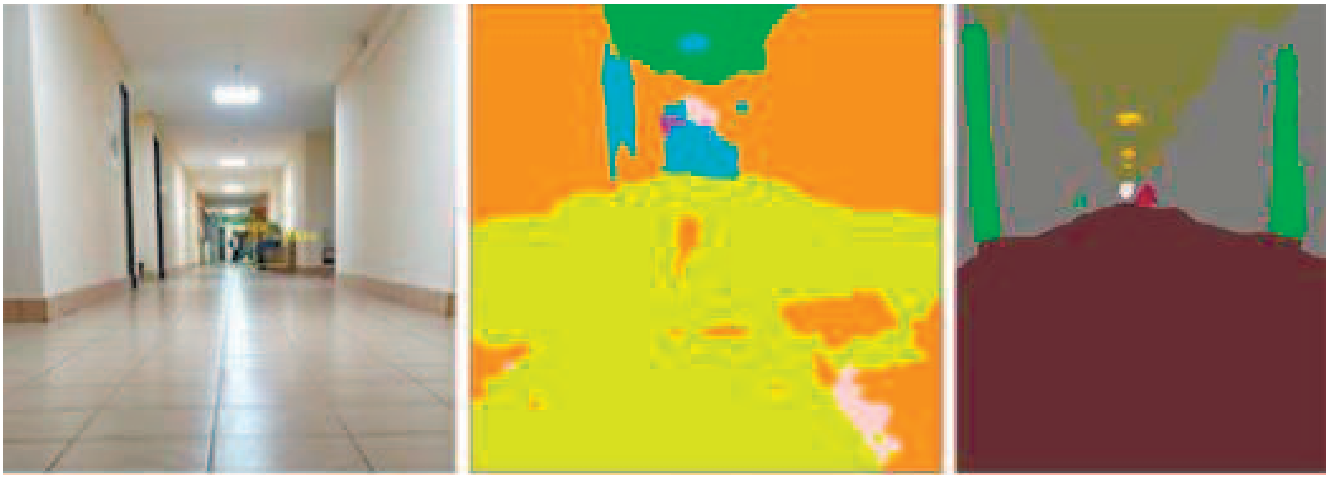

**Fig. 5.** Recognizing obstacles in the corridor of ACS department using: a – SSD classifier; b – UNET segmentation; c – Deeplab segmentation / Розпізнавання перешкоди в коридорі кафедри АСУ за допомогою: а – класифікатора SSD; б – сегментації UNET; в – сегментації Deeplab

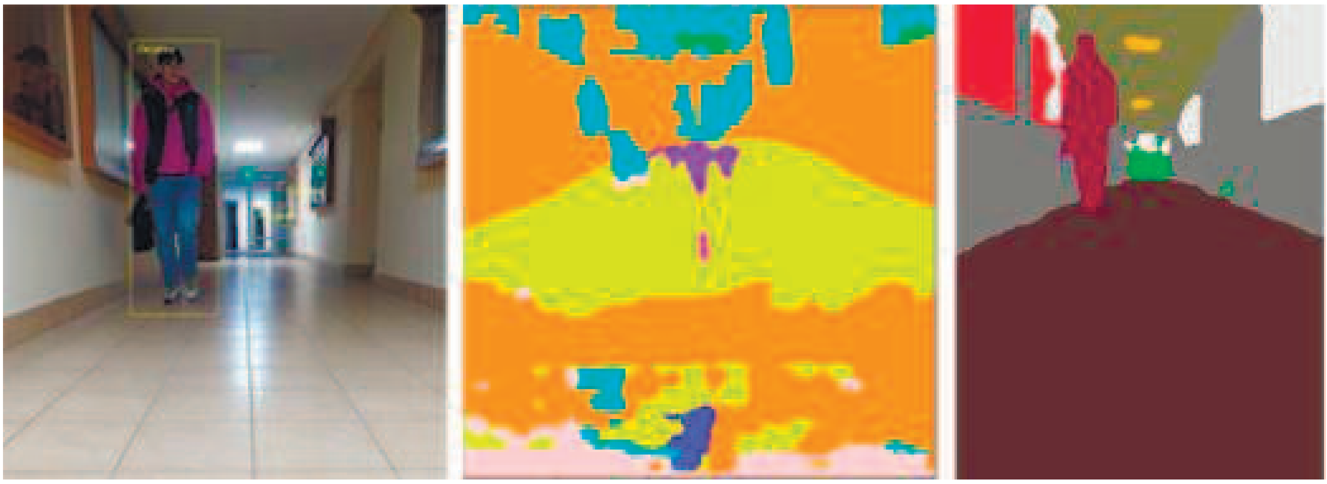

**Fig. 6.** Recognizing people in the corridor of ACS department using: a – SSD classifier; b – UNET segmentation; c – Deeplab segmentation / Розпізнавання людини в коридорі кафедри АСУ за допомогою: а – класифікатора SSD; б – сегментації UNET; в – сегментації Deeplab

**Fig 7.** Recognizing furniture in the ACS department using: a – SSD classifier; b – UNET segmentation; c – Deeplab segmentation / Розпізнавання меблів на кафедрі АСУ за допомогою: а – класифікатора SSD; б – сегментації UNET; в – сегментації Deeplab

The authors of [12] used new approaches to train recognition models for autonomous vehicles on several datasets simultaneously, but due to the large amount of data for training, they obtained low recognition accuracy.

The use of tripling the internal architecture of the UNet model presented in [15], although it complicated the internal structure of the model, gave a significant increase in accuracy, which allowed the accuracy rate of 86.5 % to almost equal our result.

Thus, based on the results of the work performed, we can formulate the following scientific novelty and practical significance of the research results.

*Scientific novelty of the obtained research results* – that the training process of indoor object recognition using neural networks has been modified by introducing two-phased training involving pre-training using more general data from one dataset and post-training with more specific task-related data from another dataset that results in improving object detection accuracy in indoor space.

*Practical significance of the research results* – that the improved method of object recognition can be used in robotic assistants systems in various indoor conditions for social purposes like assisting people in need with finding or moving things, as well as in auxiliary robots in manufacturing processes.

## Conclusions / Висновки

As a result of the work performed, the existing software solutions for solving the problem of finding the exact boundaries of an object in an image were investigated, and existing approaches to the design and construction of neural networks for object boundary recognition were analyzed. A comparative characterization of the known models for object boundary recognition, namely UNet, PSPNet, HRNet, and Deeplab, is carried out, and experiments are conducted to determine the qualitative characteristics of the proposed models.

To ensure the quality of the model, the NYU Depth Dataset and ADE20K datasets, which contain a large number of images of the closed environment, were chosen for training. The method of object recognition using neural networks has been improved by introducing two-phased training of the neural network to obtain accurate boundaries of indoor objects in the image, namely by pre-training using more general data from one dataset and post-training with more specific task-related indoor scenes data from another dataset that results in improving object detection accuracy.

To determine the efficiency of the system, experiments were conducted on the NYU Depth Dataset and ADE20K datasets, and the object recognition accuracy was achieved at 87 %. The system is able to recognize typical objects and objects found in an indoor environment, such as walls, ceilings, floors, furniture, foreign objects, etc.

An example of a situation where the solution can be useful is in the field of auxiliary robots. Such a system can be an assistant, helping a person navigate the building and help them find the things they need. The developed system can also perform tasks and work autonomously, move independently in an enclosed environment, and automatically avoid obstacles ahead.

## References


1. [1] Tsmots, I. G., Teslyuk, V. M., Opotiak, Yu. V., & Oliinyk, O. O. (2023). Development of the scheme and improvement of the motion control method of a group of mobile robotic platforms. Ukrainian Journal of Information Technology, 5(2), 97–104. https://doi.org/10.23939/ujit2023.02.097
2. [2] Tsmots I. G., Opotiak Yu. V., Obelovska K. M., & Tesliuk S. V. (2024). Methods and means of conflict-free data exchange in the group of mobile robotic platforms. Ukrainian Journal of Information Tecnology, 6(1), 65–75. https://doi.org/10.23939/ujit2024.01.065
3. [3] Huan Wang, Can Qin, Yue Bai, Yulun Zhang, & Yun Fu. (2022). Recent Advances on Neural Network Pruning at Initialization. https://doi.org/10.48550/arXiv.2103.06460
4. [4] Borkivskyi, B. P., Teslyuk, V. M. (2023). Application of neural network tools for object recognition in mobile systems with obstacle avoidance. Scientific Bulletin of UNFU, 33(4), 84–89. https://doi.org/10.36930/40330412
5. [5] Liang–Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, & Alan L. Yuille. (2017). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. https://doi.org/10.48550/arXiv.1606.00915
6. [6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, & Jian Sun. (2015). Deep Residual Learning for Image Recognition. https://doi.org/10.48550/arXiv.1512.03385
7. [7] Omar Mohamed Awad, Habib Hajimolahoseini, Michael Lim, Gurpreet Gosal, Walid Ahmed, Yang Liu, & Gordon Deng. (2023). Improving Resnet-9 Generalization Trained on Small Datasets. https://doi.org/10.48550/arXiv.2309.03965
8. [8] François Chollet. (2017). Xception: Deep Learning with Depthwise Separable Convolutions. https://doi.org/10.48550/arXiv.1610.02357
9. [9] Olaf Ronneberger, Philipp Fischer, & Thomas Brox. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. https://doi.org/10.48550/arXiv.1505.04597
10. [10] Gaurav Prasanna, John Rohit Ernest, Lalitha G, & Sathiya Narayanan (2023). Squeeze Excitation Embedded Attention UNet for Brain Tumor Segmentation. https://doi.org/10.48550/arXiv.2305.07850
11. [11] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, & Jiaya Jia (2017). Pyramid Scene Parsing Network. https://doi.org/10.48550/arXiv.1612.01105
12. [12] Li Wang, Dong Li, Han Liu, Jinzhang Peng, Lu Tian, & Yi Shan (2021). Cross-Dataset Collaborative Learning for Semantic Segmentation in Autonomous Driving. https://doi.org/10.48550/arXiv.2103.11351
13. [13] Jingdong Wang, Ke Sun, Tianheng Cheng, Borui Jiang, Chaorui Deng, Yang Zhao, Dong Liu, Yadong Mu, Mingkui Tan, Xinggang Wang, Wenyu Liu, & Bin Xiao (2020). Deep High-Resolution Representation Learning for Visual Recognition. https://doi.org/10.48550/arXiv.1908.07919
14. [14] Yixuan Zhou, Xuanhan Wang, Xing Xu, Lei Zhao, & Jingkuan Song (2023). X-HRNet: Towards Lightweight Human Pose Estimation with Spatially Unidimensional Self-Attention. https://doi.org/10.48550/arXiv.2310.08042
15. [15] Liu, G., Guo, Y., Jin, Q., Chen, G., Saheya, B., & Wu, C. (2024). A region of interest focused Triple UNet architecture for skin lesion segmentation. International Journal of Imaging Systems and Technology, 34(3). https://doi.org/10.48550/arXiv.2311.12581
16. [16] Vinh Quoc Luu, Duy Khanh Le, Huy Thanh Nguyen, Minh Thanh Nguyen, Thinh Tien Nguyen, & Vinh Quang Dinh. (2024). Semi-Supervised Semantic Segmentation using Redesigned Self-Training for White Blood Cells. https://doi.org/10.48550/arXiv.2401.07278
17. [17] Liang-Chieh Chen, George Papandreou, Florian Schroff, & Hartwig Adam. (2017). Rethinking Atrous Convolution for Semantic Image Segmentation. https://doi.org/10.48550/ arXiv.1706.05587
18. [18] Silberman, N., Hoiem, D., Kohli, P., & Fergus, R. (2012). Indoor Segmentation and Support Inference from RGBD Images. In Computer Vision – ECCV 2012 (pp. 746–760). Springer Berlin Heidelberg. https://doi.org/10.1007/978–3–642–33715–4_54
19. [19] Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., & Torralba, A. (2017). Scene Parsing through ADE20K Dataset. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 5122–5130). https://doi.org/10.1109/CVPR.2017.544



***Б. П. Борківський, В. М. Теслюк***

*Національний університет "Львівська політехніка", м. Львів, Україна*


## МЕТОД ОПТИМІЗАЦІЇ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ У ЗАКРИТОМУ СЕРЕДОВИЩІ ДЛЯ МОБІЛЬНИХ РОБОТИЗОВАНИХ СИСТЕМ З ОБХОДОМ ПЕРЕШКОД


Запропоновано використання модифікованого процесу навчання нейронної мережі з почерговим використанням декількох наборів даних для розпізнавання об'єктів мобільними роботизованими системами у закритому просторі. Проаналізовано останні дослідження та підходи до вирішення проблеми розпізнавання точних контурів об'єктів, виявлено ключові особливості декількох архітектур нейронних мереж. Виявлено нестачу даних про ефективність використання систем розпізнавання об'єктів у мобільних роботизованих системах із обходом перешкод у закритому просторі. Розроблено метод, що ґрунтується на поєднанні нейронної мережі глибинного навчання для отримання точних контурів об'єктів із візуальними даними перешкод у закритому середовищі. Застосовано архітектуру Deeplab для пошуку точних меж об'єктів та використано для тренування моделей набори даних, які пропонують велику кількість анотацій для різноманітних сцен і категорій об'єктів, а саме NYU Depth Dataset V2 і ADE20K. Здійснено експерименти для порівняння точності моделей UNet та Deeplab на різних наборах даних та їх комбінаціях для визначення доцільності використання запропонованого підходу для розв'язання поставлених завдань. Визначено точність розпізнавання розробленого методу на рівні 86,9 % із використанням архітектури Deeplab та комбінації наборів даних ADE20K та NYU Depth. Подано результати у графічному вигляді розпізнавання об'єктів у різних типах закритого простору із порівнянням точності розпізнавання різних підходів. Встановлено, що запропонована модифікація процесу навчання сприяє підвищенню точності розпізнавання перешкод у закритому просторі, надає оптимальніше рішення для навігаційної компоненти роботизованих систем за допомогою отримання інформації про точні обриси перешкод і побудови оптимальнішого шляху, і є ефективною для виконання поставлених завдань.

***Ключові слова:*** оброблення зображень, згорткова мережа, розпізнавання образів, сегментація зображень.

---

**Інформація про авторів:**

**Борківський Богдан Петрович**, аспірант, кафедра автоматизованих систем управління.
**Email:** bohdan.p.borkivskyi@lpnu.ua; https://orcid.org/0000–0003–3301–476X

**Теслюк Василь Миколайович**, д-р техн. наук, професор, завідувач кафедри автоматизованих систем управління.
**Email:** vasyl.m.teslyuk@lpnu.ua; https://orcid.org/0000–0002–5974–9310

**Цитування за ДСТУ:** Борківський Б. П., Теслюк В. М. Метод оптимізації розпізнавання об'єктів у закритому середовищі для мобільних роботизованих систем з обходом перешкод. Український журнал інформаційних технологій. 2024, т. 6, № 2, С. 41–48.

**Citation APA:** Borkivskyi, B. P., & Teslyuk, V. M. (2024). Optimization of object detection in closed space using mobile robotic systems with obstacle avoidance. *Ukrainian Journal of Information Technology, 6*(2), 41–48. https://doi.org/10.23939/ujit2024.02.041