

АНОТАЦІЯ

Стець С.Ю. Підвищення точності та швидкодії детектування зображень автомобілів засобами згорткової нейронної мережі YOLO. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 121 – «Інженерія програмного забезпечення» – Чернівецький національний університет імені Юрія Федьковича, Чернівці, 2026.

Детектування зображень транспортних засобів є базовим етапом для автоматизованого дистанційного зондування та контролю стану об'єктів у різноманітних системах Інтернету речей (IoT). Автомобілі суттєво впливають на стан дорожньої обстановки, тому підвищення ефективності їх виявлення є ключовим фактором надійності транспортних систем. Для розпізнавання та детектування зображень автомобілів широко застосовуються штучні нейронні мережі (ШНМ), зокрема, згорткові нейронні мережі (ЗНМ). Одним з найбільш ефективних засобів детектування зображень об'єктів є ЗНМ з архітектурою YOLO (англ. You Only Look Once), які застосовують технології глибокого навчання.

Проте стандартні моделі YOLO, особливо моделі малого розміру, часто демонструють недостатню точність детектування у складних сценах, а моделі великого розміру вимагають значних обчислювальних ресурсів, що ускладнює їх використання у вбудованих системах. Тому підвищення точності та швидкодії детектування зображень автомобілів є **актуальним завданням**.

Метою дисертаційної роботи є підвищення точності та швидкодії детектування зображень автомобілів шляхом вибору архітектури та версії згорткових нейронних мереж, попередньої обробки зображень та донавчання нейронних мереж з архітектурою YOLO на основі автоматизовано створених наборів даних.

Для досягнення мети вирішувались такі **завдання**:

1. Провести аналіз існуючих нейромережових архітектур, методів та програмних засобів для детектування зображень об'єктів, і на основі такого аналізу виконати вибір архітектури та версії ЗНМ, які б задовольняли вимоги до швидкодії та точності детектування зображень автомобілів.

2. Розробити методику та програмні засоби для попередньої обробки зображень шляхом еквалізації їх гістограм та підвищення локального контрасту, які б забезпечували підвищення точності детектування зображень автомобілів.

3. Розробити методику та програмні засоби для навчання різних версій моделей YOLO та аналізу результатів навчання, які б давали змогу підвищити точність детектування за рахунок донавчання та вибирати версію YOLO з урахуванням вимог для конкретної задачі аналізу зображень.

4. Розробити методику та програмні засоби для автоматизованого створення набору даних із зображень автомобілів засобами моделі ЗНМ YOLO великого розміру, які б забезпечували селекцію унікальних кадрів відеопотоку.

5. Розробити методику та програмні засоби з архітектурою «Вчитель-Учень» для донавчання згорткової нейронної мережі YOLO малого розміру на основі датасету, створеного в автоматизованому режимі.

6. Виконати програмну реалізацію інтелектуальної системи для детектування зображень автомобілів із використанням навчених моделей YOLO.

7. Провести експериментальне дослідження точності та швидкодії розробленої системи на прикладі детектування об'єктів на тестових зображеннях.

Для досягнення мети застосовувалися такі **методи дослідження**: методи глибокого навчання для згорткових нейронних мереж архітектури YOLO; методи цифрової обробки зображень для еквалізації їх гістограм та підвищення локального контрасту; методи інтелектуального аналізу даних (регресійного аналізу, зокрема, лінійної регресії); перцептуальне хешування для селекції унікальних кадрів

відеопотоку; методи комбінаторної оптимізації (угорський алгоритм); методи класифікації даних та зображень; методи комп'ютерного експерименту з використанням програм на мові Python для емпіричного підтвердження адекватності отриманих результатів.

У вступі обґрунтовано актуальність теми дисертаційної роботи, сформульовано мету та завдання проведених досліджень, описано наукову новизну та практичне значення одержаних результатів, представлено методи, об'єкт і предмет досліджень, вказано особистий внесок здобувача, а також наведено дані щодо публікацій за темою дисертації.

У першому розділі здійснено комплексний аналіз сучасних методів і засобів обробки детектування зображень автомобілів. Детально розглянуто структуру та принципи роботи ЗНМ, проаналізовано вплив гіперпараметрів на процес навчання. Описано принципи побудови глибоких залишкових мереж (ResNet), архітектури MobileNet із застосуванням роздільних згорток, а також архітектури EfficientNet. Проведено огляд двостадійних (Faster R-CNN) та одностадійних (YOLO, SSD) детекторів зображень об'єктів. Обґрунтовано вибір архітектури YOLO як базової для подальших досліджень з огляду на її показники точності та швидкодії. Проаналізовано існуючі програмні інструменти (Google Cloud Vision, Amazon Rekognition) та бібліотеки (OpenCV), призначені для детектування зображень.

У другому розділі розроблено та програмно на мові Python реалізовано методику підвищення якості детектування зображень шляхом попередньої обробки вхідних даних. Розділ присвячено архітектурним особливостям обраної моделі YOLOv8 та інтеграції в її конвеєр модуля попередньої обробки. Детально описано структуру YOLOv8. Попередня обробка зображень полягає підвищенні їх контрасту трьома способами: 1) глобальне вирівнювання (еквалізація) гістограми; 2) підвищення локального контрасту методом адаптивної еквалізації гістограми (CLAHE); 3) вирівнювання та центрування гістограми. Точність детектування

зображень автомобілів оцінено за метриками Precision, Recall, F1-score та середньої точності (mAP). Встановлено, що запропонований спосіб попередньої обробки зображень шляхом вирівнювання та центрування гістограми забезпечує найвищу точність детектування автомобілів. Проведено інтелектуальний аналіз даних, а саме регресійний та кореляційний аналіз результатів детектування зображень для знаходження взаємозв'язків між метриками якості детектування.

У третьому розділі розроблено методику для донавчання різних версій ЗНМ YOLO, виконано програмну реалізацію методики, проведено машинне навчання моделей YOLO та проаналізовано його результати. Описано процес формування спеціалізованого набору зображень автомобілів, який включає збір, ручну анотацію зображень та балансування класів. Аугментація даних полягала у розширенні навчальної вибірки зображень шляхом застосування геометричних (повороти, відображення, зміна масштабу) та колористичних (зміна яскравості, насиченості, додавання шуму) перетворень. Це дозволило збільшити обсяг датасету з 1542 до 3578 зображень та запобігти перенавчанню моделі. Виконано донавчання моделей YOLO, попередньо натренованих на датасеті COCO. Шляхом детектування тестових зображень отримано, що донавчання дозволяє підвищити точність детекції за метрикою повноти Recall з 0.58 до 0.91, а за інтегральною метрикою mAP50-95 – з 0.37 до 0.72. Проведено порівняльний аналіз ефективності ЗНМ двох версій: YOLOv8m та новітньої YOLOv11m. Встановлено, що при співставній точності детектування (mAP50 \approx 0.96) модель YOLOv11m демонструє на 22% менший розмір файлу ваг. Показано, що модель YOLOv8m доцільно вибирати для забезпечення максимальної точності, а YOLOv11m – максимальної швидкодії.

У четвертому розділі розроблено методику для автоматизованого створення набору даних із зображень автомобілів засобами моделі згорткової нейронної мережі YOLO великого розміру. Методика передбачає автоматичну селекцію унікальних кадрів з відеопотоку з використанням методу перцептуального

хешування (pHash). Це дозволяє оцінювати візуальну схожість кадрів і відфільтровувати дублікати, зменшуючи надлишковість даних на 88% (з 993 до 121 кадра у тестовому експерименті). Виявлення схожих об'єктів (автомобілів) на зображеннях виконано угорським алгоритмом з урахуванням метрики детектування IoU та середньоквадратичної помилки (RMSE) між дескрипторами зображень.

Розроблено методику з архітектурою «Вчитель-Учень», яка полягає у донавчанні згорткової нейронної мережі YOLO малого розміру «Учень» (YOLOv8n) на основі автоматизовано створеного спеціалізованого датасету зображень мережею YOLO середнього розміру «Вчитель» (YOLOv8m). Експериментально підтверджено, що такий підхід дозволяє підвищити точність моделі «Учень» за метрикою mAP50 з 0.6864 до 0.9292.

Розроблені методики інтелектуальної системи реалізовано в програмі на мові Python, яка виконується на апаратній платформі одноплатного комп'ютера Raspberry Pi 5 з використанням технологій контейнеризації (Docker). Розроблено веб-інтерфейс програми. Результати обробки тестових зображень показали високу точність детектування автомобілів на зображеннях, отриманих засобами БПЛА.

Наукова новизна отриманих результатів полягає у наступному:

1. Вперше розроблено методику з архітектурою «Вчитель-Учень», особливістю якої є донавчання згорткової нейронної мережі YOLO малого розміру на основі автоматизовано створеного спеціалізованого датасету зображень мережею YOLO середнього або великого розміру, що забезпечує високу швидкодію детектування зображень автомобілів засобами моделі нейромережі малого розміру та підвищує точність детектування за метрикою mAP50 (до 24 %).

2. Вперше розроблено методику для попередньої обробки зображень, які подаються на входи згорткової нейронної мережі з архітектурою YOLO, особливістю якої є підвищення контрасту зображень запропонованим способом вирівнювання (еквалізації) та центрування гістограм зображень, що призводить до

підвищення точності детектування зображень автомобілів за метрикою перекривання рамок IoU (на 16 %).

3. Подальшого розвитку отримала методика для донавчання різних версій ЗНМ YOLO з використанням датасетів, створених у ручному режимі, особливістю якої є порівняльний аналіз результатів донавчання, що забезпечує цілеспрямований вибір версії YOLO відповідно вимогам до точності, швидкодії та обсягу використаних ресурсів при детектуванні зображень автомобілів.

4. Подальшого розвитку отримала методика для автоматизованого створення набору даних із зображень автомобілів засобами моделі згорткової нейронної мережі YOLO середнього або великого розміру, особливістю якої є селекція унікальних кадрів відеопотоку з використанням перцептивного хешування, що дає змогу суттєво скоротити час формування датасету і отримувати спеціалізовані набори без дублювання даних із мінімальною участю людини-експерта.

Практичне значення отриманих результатів полягає у тому, що розроблені у дисертаційній роботі методики та програмні засоби на мові Python можуть застосовуватися для автоматизованого донавчання нейронних мереж YOLO та для високоточного детектування з їх допомогою зображень автомобілів та інших учасників дорожнього руху в прикладних системах комп'ютерного зору, які застосовуються для оцінки дорожнього трафіку, аналізу зайнятості автомобільних парковок, контролю безпеки дорожнього руху, аналізу дорожньої обстановки та в технологіях автономного водіння. Розроблені програмні засоби штучного інтелекту можуть функціонувати на мобільних обчислювальних пристроях.

Ключові слова: аналіз даних (інтелектуальний аналіз даних), безпілотні літальні апарати (БПЛА), глибоке навчання, дистанційне зондування, інтелектуальна система, класифікація (класифікація даних), класифікація зображень, машинне навчання, регресійний аналіз, лінійна регресія, штучний інтелект, штучні нейронні мережі (ШНМ), YOLO, згорткові нейронні мережі (ЗНМ).

ABSTRACT

Serhii Stets. Improving the accuracy and speed of car image detection using the YOLO convolutional neural network. – Qualification scientific work as a manuscript.

Thesis for the degree of Doctor of Philosophy in the specialty 121 – “Software Engineering” – Yuriy Fedkovych Chernivtsi National University, Chernivtsi, 2026.

Detecting images of vehicles is a basic step for automated remote sensing and control of objects in various Internet of Things (IoT) systems. Cars have a significant impact on road conditions, so improving the efficiency of their detection is a key factor in the reliability of transport systems. Artificial neural networks (ANN), in particular convolutional neural networks (CNN), are widely used for recognizing and detecting images of cars. One of the most effective means of detecting images of objects is CNN with YOLO (You Only Look Once) architecture, which use deep learning technologies.

However, standard YOLO models, especially small ones, often demonstrate insufficient detection accuracy in complex scenes, while large models require significant computing resources, which complicates their use in embedded systems. Therefore, improving the accuracy and speed of vehicle image detection **is a pressing task**.

The goal of this dissertation is to improve the accuracy and speed of car image detection by selecting the architecture and version of convolutional neural networks, preprocessing images, and retraining neural networks with YOLO architecture based on automatically generated datasets.

To achieve this goal, the following **tasks** were solved:

1. Analyze existing neural network architectures, methods, and software tools for detecting images of objects, and based on this analysis, select the architecture and version of the CNN that would meet the requirements for speed and accuracy of detecting images of cars.

2. Develop methods and software tools for preprocessing images by equalizing their histograms and increasing local contrast, which would improve the accuracy of vehicle image detection.

3. Develop methods and software tools for training different versions of YOLO models and analyzing training results, which would make it possible to improve detection accuracy through retraining and select the YOLO version based on the requirements for a specific image analysis task.

4. Develop a methodology and software tools for the automated creation of a dataset of car images using large-scale YOLO models, which would ensure the selection of unique frames from the video stream.

5. Develop a methodology and software tools with a “Teacher-Student” architecture for retraining a small YOLO convolutional neural network based on a dataset created in an automated mode.

6. Implement software implementation of an intelligent system for detecting images of cars using trained YOLO models.

7. Conduct an experimental study of the accuracy and performance of the developed system using the example of object detection in test images.

To achieve this goal, the following **research methods** were used: remote learning methods for convolutional neural networks of YOLO architecture; digital image processing methods for equalizing their histograms and increasing local contrast; methods of data mining (regression analysis, in particular, linear regression); perceptual hashing for selecting unique frames from a video stream; combinatorial optimization methods (Hungarian algorithm); data and image classification methods; computer experiment methods using Python programs for empirical confirmation of the adequacy of the results obtained.

The introduction substantiates the relevance of the dissertation topic, formulates the purpose and objectives of the research, describes the scientific novelty and practical

significance of the results obtained, presents the methods, objects, and subject of the research, indicates the personal contribution of the applicant, and provides data on publications related to the dissertation topic.

The first chapter provides a comprehensive analysis of modern methods and means of processing and detecting images of vehicles. The structure and principles of operation of deep neural networks are examined in detail, and the influence of hyperparameters on the learning process is analyzed. The principles of constructing deep residual networks (ResNet), MobileNet architecture with the use of separate convolutions, and EfficientNet architecture are described. A review of two-stage (Faster R-CNN) and single-stage (YOLO, SSD) object image detectors are provided. The choice of YOLO architecture as the basis for further research is justified, given its accuracy and performance indicators. Existing software tools (Google Cloud Vision, Amazon Rekognition) and libraries (OpenCV) designed for image detection are analyzed.

The second chapter develops and implements in Python a methodology for improving image detection quality through preprocessing of input data. The chapter is devoted to the architectural features of the selected YOLOv8 model and the integration of a preprocessing module into its pipeline. The structure of YOLOv8 is described in detail. Preprocessing of images consists of increasing their contrast in three ways: 1) global histogram equalization; 2) increasing local contrast using the adaptive histogram equalization (CLAHE) method; 3) histogram equalization and centering. The accuracy of car image detection is evaluated using the metrics Precision, Recall, F1-score, and mean accuracy (mAP). It has been established that the proposed method of preprocessing images by equalizing and centering the histogram provides the highest accuracy of car detection. Intelligent data analysis was performed, namely regression and correlation analysis of image detection results to find correlations between detection quality metrics.

The third section develops a methodology for retraining different versions of the YOLO neural network, implements the methodology in software, machine learning of

YOLO models is performed, and its results are analyzed. It describes the process of forming a specialized set of car images, which includes collection, manual annotation of images, and class balancing. Data augmentation consisted of expanding the training sample of images by applying geometric (rotations, reflections, scaling) and coloristic (brightness, saturation, noise addition) transformations. This allowed to increase the size of the dataset from 1542 to 3578 images and prevent model overfitting. Performed retraining of YOLO models previously trained on the COCO dataset. By detecting test images, it was found that retraining allows increasing the detection accuracy according to the Recall completeness metric from 0.58 to 0.91, and according to the mAP50-95 integral metric – from 0.37 to 0.72. A comparative analysis of the effectiveness of two versions of CNNs was performed: YOLOv8m and the latest YOLOv11m. It was found that with comparable detection accuracy ($mAP50 \approx 0.96$), the YOLOv11m model has a 22 % smaller weight file size. It was shown that the YOLOv8m model should be chosen to ensure maximum accuracy, and YOLOv11m – for maximum performance.

The fourth section develops a methodology for the automated creation of a dataset of car images using a large YOLO convolutional neural network model. The methodology involves the automatic selection of unique frames from a video stream using the perceptual hashing (pHash) method. This allows evaluating the visual similarity of frames and filtering out duplicates, reducing data redundancy by 88% (from 993 to 121 frames in a test experiment). The detection of similar objects (cars) in images is performed by a Hungarian algorithm, taking into account the IoU detection metric and the root mean square error (RMSE) between image descriptors.

A methodology with a “Teacher-Student” architecture was developed, which consists in retraining a small convolutional neural network YOLO ‘Student’ (YOLOv8n) based on an automatically created specialized image dataset by a medium-sized network YOLO “Teacher” (YOLOv8m). It has been experimentally confirmed that this approach allows increasing the accuracy of the “Student” model according to the mAP50 metric

from 0.6864 to 0.9292. The developed methods of the intelligent system are implemented in a Python program that runs on a Raspberry Pi5 single-board computer hardware platform using containerization technologies (Docker). A web interface for the program has been developed. The results of processing test images showed high accuracy in detecting cars in images obtained by unmanned aerial vehicles (UAV).

The scientific novelty of the results obtained lies in the following:

1. For the first time, a methodology with a “Teacher-Student” architecture has been developed, the feature of which is the retraining of a small-sized YOLO convolutional neural network based on an automatically created specialized image dataset by a medium or large-sized YOLO network, which ensures high-speed detection of car images using a small neural network model and increases detection accuracy according to the mAP50 metric (up to 24 %).

2. For the first time, a method has been developed for preprocessing images that are fed to the inputs of a convolutional neural network with YOLO architecture, which features an increase in image contrast using the proposed method of equalization and centering of image histograms, resulting in an increase in the accuracy of vehicle image detection according to the IoU metric (by 16%).

3. A method for retraining different versions of the YOLO CNN using manually created datasets has been improved, featuring a comparative analysis of retraining results, which ensures a targeted selection of the YOLO version in accordance with the requirements for accuracy, speed, and the amount of resources used in detecting car images.

4. The methodology for automated creation of a dataset of car images using a medium or large YOLO convolutional neural network model has been improved. which features the selection of unique video stream frames using perceptual hashing, which significantly reduces the time required to form a dataset and obtain specialized sets without data duplication with minimal human expert involvement.

The practical significance of the results obtained lies in the fact that the methods and software tools developed in the dissertation in Python can be used for automated retraining of YOLO neural networks and for high-precision detection of images of cars and other road users in applied computer vision systems used for traffic assessment, analyzing car park occupancy, monitoring road safety, analyzing road conditions, and in autonomous driving technologies. The developed artificial intelligence software tools can run on mobile computing devices.

Keywords: data analysis (data mining), unmanned aerial vehicles (UAV), deep learning, remote sensing, intelligent system, classification (data classifier), image classification, machine learning, regression analysis, linear regression, artificial intelligence, artificial neural networks (ANN), YOLO, convolutional neural networks (CNN).

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Наукові праці, в яких опубліковані основні наукові результати дисертації:

публікації у фахових виданнях України:

1. **Стець С.** Аналіз точності та швидкодії детекції автомобілів за допомогою нейронних мереж YOLOV8 та YOLOV11. *Herald of Khmelnytskyi National University. Technical Sciences.* 2025. Т. 357 (5.2). С. 123-130. <https://doi.org/10.31891/2307-5732-2025-357-74>.
2. Balovsyak S., **Stets S.** Preprocessing of object images before their detection using YOLO neural network. *Security of Infocommunication Systems and Internet of Things.* 2025. Vol. 3, No. 2, Paper 02002. P. 1-5. ISSN 2786-8443. <https://doi.org/10.31861/sisiot2025.2.02002>. (Баловсяк С. – наукове керівництво і

редагування; **Стець С.** – розробка методики та програмного забезпечення для попередньої обробки зображень перед їх детектуванням).

3. Баловсяк С., **Стець С.** Автоматизоване створення спеціалізованого датасету для зображень автомобілів. *Herald of Khmelnytskyi National University. Technical Sciences*. 2025. Т. 359 (6.2). С. 278-285. <https://doi.org/10.31891/2307-5732-2025-359-111>. (Баловсяк С. – наукове керівництво і редагування; **Стець С.** – розробка методики та програмного забезпечення для автоматичного створення набору даних та донавчання згорткової нейронної мережі).

Наукові праці, які засвідчують апробацію матеріалів дисертації:

4. Баловсяк С.В., **Стець С.Ю.** Аналіз сучасних методів розпізнавання зображень автомобілів. *Проблеми інформатики та комп'ютерної техніки: праці XI Міжнародної науково-практичної конференції (ПШКТ – 2022)*, м. Чернівці, 10–13 листопада 2022 р. Чернівці: Черн. нац. ун-т, 2022. С. 77-80. (Баловсяк С.В. – наукове керівництво і редагування; **Стець С.Ю.** – аналіз сучасних методів розпізнавання зображень автомобілів).
5. Баловсяк С.В., **Стець С.Ю.** Використання модуля Inception для підвищення точності розпізнавання зображень у згорткових нейронних мережах. *Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення: матеріали міжнародної наукової інтернет-конференції*, 8-9 червня 2023 р. Вип. 78. Тернопіль, 2023. С. 30-32. (Баловсяк С.В. – наукове керівництво і редагування; **Стець С.Ю.** – розробка програмного забезпечення з використанням модуля Inception).
6. **Стець С.Ю.** Підвищення точності нейромережного розпізнавання зображень за допомогою модуля Inception. *Проблеми інформатики та комп'ютерної техніки:*

- праці XII Міжнародної науково-практичної конференції (ПКТ – 2023), м. Чернівці, 10-12 листопада 2023 р. Чернівці: Черн. нац. ун-т, 2023. С. 61-63.
7. Баловсяк С.В., **Стець С.Ю.** Детектування зображень пішохідних переходів за допомогою нейронних мереж YOLO. *Проблеми інформатики та комп'ютерної техніки*: праці XIII Міжнародної науково-практичної конференції (ПКТ – 2024), м. Чернівці, 1–3 листопада 2024 р. Чернівці: Черн. нац. ун-т, 2024. С. 49-51. (Баловсяк С.В. – наукове керівництво і редагування; **Стець С.Ю.** – розробка програмного забезпечення для детектування зображень пішохідних переходів).
 8. Баловсяк С.В., **Стець С.Ю.** Попередня обробка зображень об'єктів перед їх детектуванням засобами нейронної мережі YOLO. *Фізико-технологічні проблеми передачі, обробки та зберігання інформації в інфокомунікаційних системах*: матеріали X Міжнародної науково-практичної конференції (ПРЕДТ-2025). 15-17 травня 2025 р. Чернівці: Черн. нац. ун-т, 2025. С. 181-182. (Баловсяк С.В. – наукове керівництво і редагування; **Стець С.Ю.** – розробка програмних засобів для попередньої обробки зображень).
 9. **Стець С.Ю.** Донавчання нейронної мережі YOLO за допомогою автоматизовано створеного датасету зображень автомобілів. *Проблеми інформатики та комп'ютерної техніки*: праці XIV Міжнародної науково-практичної конференції (ПКТ – 2025), м. Чернівці, 13-15 листопада 2025 р. Чернівці: Черн. нац. ун-т, 2025. С. 64-66.

Публікації, які додатково відображають наукові результати дисертації:

10. Balovsyak S., Kroitor O., Odaiska Kh., Salem A.B.M., **Stets S.** Car image recognition using convolutional neural network with efficient met architecture. *CEUR Workshop Proceeding*. 2024. Vol. 3675: 5th International Workshop on Intelligent Information Technologies and Systems of Information Security, IntelITSIS 2024. P. 182-195

(Scopus). (Balovsyak S. – наукове керівництво, Kroitor O. – аналіз наборів даних, Odaiska Kh. – редагування, Salem A.B.M. – аналіз архітектур згорткових нейронних мереж, **Stets S.** – розробка моделі згорткової нейронної мережі з архітектурою EfficientNet, призначеної для розпізнавання зображень автомобілів).