

ТЕЛІШЕВСЬКИЙ ПЕТРО

Національний університет «Львівська Політехніка»

<https://orcid.org/0009-0008-8328-0373>email: petro.a.telishevskiy@lpnu.ua**БОЙКО НАТАЛІЯ**

Національний університет «Львівська Політехніка»

<https://orcid.org/0000-0002-6962-9363>email: nataliya.i.boyko@lpnu.ua

ПОРІВНЯННЯ ОНДНОЕТАПНОЇ ТА ДВОЕТАПНОЇ СЕГМЕНТАЦІЇ ДЛЯ ДЕТЕКЦІЇ ПОШКОДЖЕННЯ АВТОТРАСПОРТУ НА ЗОБРАЖЕННІ

Стрімке зростання автомобільної промисловості та зростаючий попит на ефективну оцінку пошкоджень після аварій призвели до необхідності створення точних та автоматизованих систем виявлення пошкоджень транспортних засобів. У цій статті представлено порівняльний аналіз одноетапного та двоетапного методів сегментації для виявлення різних типів пошкоджень транспортних засобів з використанням комп'ютерного зору та глибокого навчання. Основна мета - визначити, який підхід до сегментації дає більш точні та надійні результати при виявленні таких типів пошкоджень, як вм'ятини, подряпини, тріщини, розбите скло, розбиті лампи та спушені шини.

Для навчання та оцінки використовувався набір даних CarDD, що складається з 4 000 зображень із випадками пошкоджень. Для порівняння було обрано дві найсучасніші моделі: YOLOv8, що представляє одноступеневий підхід до сегментації, та Mask R-CNN, як двоступеневий метод сегментації. YOLOv8 пропонує швидший висновок і простішу архітектуру, тоді як Mask R-CNN забезпечує вищу точність сегментації завдяки вдосконаленому механізму пропозиції областей і техніці RoIAlign.

Експериментальні результати показують, що Mask R-CNN перевершує YOLOv8 за точністю сегментації, особливо в складних сценаріях пошкодження. Для оцінки продуктивності для різних класів пошкоджень використовувалися такі показники, як середня точність (mAP50 і mAP50-95) і матриці плутанини. У той час як YOLOv8 продемонстрував високу ефективність у виявленні простих типів пошкоджень, таких як вм'ятини і подряпини, Mask R-CNN стабільно показував кращі результати у всіх категоріях, особливо в мінімізації плутанини з фоном.

Ключові слова: виявлення пошкоджень транспортних засобів, сегментація зображень, YOLOv8, Mask R-CNN, комп'ютерний зір.

TELISHEVSKIY PETRO, BOYKO NATALIYA

Lviv Polytechnic National University

COMPARE ONE-STAGE AND TWO-STAGE SEGMENTATION FOR DETECTING VEHICLE DAMAGE IN AN IMAGE

The rapid growth of the automotive industry and the increasing demand for efficient accident damage assessment have led to a need for accurate and automated vehicle damage detection systems. This paper presents a comparative analysis of one-stage and two-stage segmentation methods for detecting various types of damage in vehicles using computer vision and deep learning. The main goal is to determine which approach to segmentation brings more value and better results when detecting types of damage such as dents, scratches, cracks, broken glass, broken lights, and tires flat.

The Car Damage Detection dataset (CarDD), consisting of 4,000 images and over 9,000 annotated damage instances, was used for training and evaluation. Two models were chosen for comparison: YOLOv8, representing the one-stage segmentation approach, and Mask R-CNN, as the two-stage segmentation method. YOLOv8 offers faster inference and simpler architecture, while Mask R-CNN provides higher segmentation precision due to its refined region proposal mechanism and RoIAlign technique.

Experiments show that Mask R-CNN outperforms YOLOv8 in terms of segmentation accuracy, especially in complex damage scenarios. Metrics such as mean Average Precision and confusion matrices were used to analyze productivity for different types of damage. While YOLOv8 was highly effective in detecting simple types of damage such as dents and scratches, Mask R-CNN consistently performed better in all categories, especially in minimizing background blur.

This study concludes that while YOLOv8 is suitable for real-time applications due to its speed, Mask R-CNN is better suited for scenarios where detection accuracy is critical, such as insurance assessments or automated vehicle inspection systems.

Keywords: vehicle damage detection, image segmentation, YOLOv8, Mask R-CNN, computer vision.

Стаття надійшла до редакції / Received 01.05.2025

Прийнята до друку / Accepted 06.06.2025

Постановка проблеми

Раніше у [1] розглядали проблему як при великому зростанні автомобілів, швидко розглядати дорожньо-транспортні ситуації, щоб цей процес був швидким та об'єктивним. У даній статті формується концепція для вирішення такого завдання. Автоматизація наведеного в [1] процесу є доволі важливим та цінним етапом, де важливим рішенням буде розпізнати та виділити пошкодження на автотранспорті. Для цього слід застосувати методи машинного навчання і комп'ютерного зору, а саме – дослідити максимально точно область пошкодження на автомобілі. У даному дослідженні проаналізуємо методи сегментації, та вибір найкращого за результатами експериментів та методи одноетапної та двоетапної сегментації.

Аналіз досліджень та публікацій

Актуальність тематики гостро обговорюється в наукових колах дослідниками. Автори [2] розробили для задачі визначення пошкодження автомобіля метод на основі Mask-RCNN. У принцип пошуку полягає в тому, щоб пройти три кроки для розпізнавання пошкодження на автомобілі, а саме, розпізнати, локалізувати, класифікувати. Для цього автори на основі Mask R-CNN використали три притреновані моделі, а саме:

Inception ResNetV2, VGG-16 та VGG-19. З даних трьох методів найкращі результати показав метод Inception ResNetV2, з точністю 95%, для локалізації – точність складає 85% і для класифікації – точність складає 80%.

У дослідженні [3] було розроблено систему, що використовує різні види нейронних мереж, щоб максимізувати визначення точності пошкодження автомобіля. Авторами було запропоновано для кожного виду. Отож, було використано такі методи: YOLOv4, щоб перевірити чи присутній транспортний засіб, VGG-16, щоб визначити чи є у нас інтер'єр автомобіля, чи екстер'єр, ResNet-18 – під яким кутом зроблено зображення, ResNet-34 – для визначення чи відстріляні подушки безпеки у авто, Mask R-CNN – для сегментування частини автотранспорту і локалізація пошкодження.

У дослідженні [4] авторами було запропоновано використати Mask R-CNN щоб локалізувати деталі транспортного засобу. Було досягнуто високих показників оцінки методів, було використано mAP, що становило 90,27%.

Хоча у дослідженні [5] розглядається визначення пошкодження дорожнього полотна, однак воно сконцентровано на визначенні тріщини. Тому даний підхід можна використати у визначенні подряпин на автомобілі та демонструє високу точність визначення тріщин, а саме – 92,46%.

Аналізуючи наукові джерела, у більшості випадків для виконання задачі по визначенні пошкоджень на транспортних засобах використовували Mask R-CNN.

Метою роботи є провести аналіз методів сегментації пошкоджень автомобілів, використавши нейронні мережі, а саме, YOLOv8 та Mask R-CNN.

Виклад основного матеріалу

Перед тим, як почати аналіз досліджуваних методів, слід ознайомитись з набором даних, який використовуватиметься в експериментах. Даним набором даних виступає CarDD – Car Damage Detection. Автори цього набору даних дали його характеристику в дослідженні [6]. У даній роботі крім того що описано про інформацію та класи, що наповнюють його, дослідники CarDD визначили, в яких задачах його можна застосувати: класифікації, сегментації, object detection. Даний набір даних складається із 4000 зображень автомобілів, де на них представлено більше 9000 різних видів пошкоджень наземного транспорту. Дані види пошкоджень поділені на 6 класів:

- dent – вм'ятина;
- scratch – подряпина;
- crack – тріщина;
- glass shatter – розбите скло;
- lamp broken – розбиті ліхтарі;
- tire flat – спущене колесо.

В роботі в подальшому, для зручності, будуть використовуватися оригінальні назви класів, а не їх переклад. На рис. 1 можна побачити приклад набору даних.

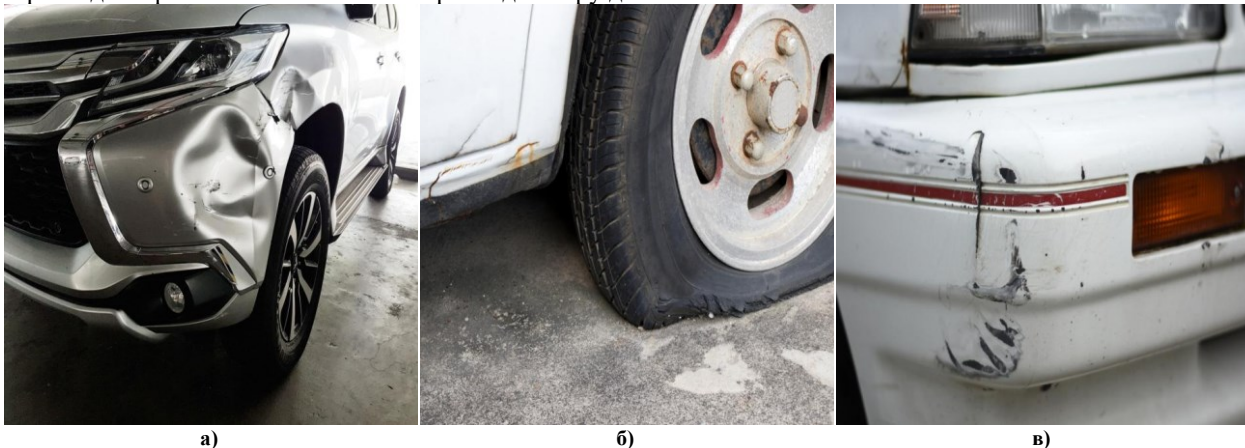


Рис. 1. Приклади зображень автомобілів із набору даних CarDD: а) – dent; б) – tire flat і crack; в) – crack, scratch

З наведених зображень можна побачити що не одну область пошкодження, їх може бути декілька на одному транспортному засобі (рис.1 б) – в).

Отож, для проведення експериментів було використано два види сегментації, а саме, одноетапна (One-stage segmentation) і двоетапна (Two-stage segmentation). One-stage segmentation полягає у тому, що визначення і процес сегментації об'єкта відбувається в один етап роботи нейронної мережі. До даного виду можна віднести YOLOv8, YOLOv5, SOLO та інші. Даний вид сегментації має ряд переваг, а саме, він простіший в архітектурі та на відміну two-stage segmentation є швидшим в обробці. Дані переваги корисні, коли нам потрібно швидко реагувати і використовувати менше ресурсів.

Однак, маючи такі переваги, які впливають на точність роботи моделі та мають нищу якість масок сегментації.

Two-stage segmentation на відміну від one-stage segmentation робить сегментацію у дві дії. Спершу відбувається пошук та виявлення певних частин пошкоджень, їх ще називають Region proposals. Це області,

що можуть містити об'єкти для сегментування. Отож, покращується якість класифікації об'єкта та визначення маски сегментації. Поширеними моделями для вирішення такого роду завдань є: Mask R-CNN, Cascade Mask R-CNN, Hybrid Task Cascade. Ці моделі дозволяють визначати точніше та швидше область пошкодження транспортного засобу. Це є перевагою при виборі інших методів для здійснення експериментів, для отримання високої оцінки щодо визначення пошкодження.

Однак, щоб отримати переваги у точності, потрібно використовувати більше ресурсів збільшити час на експерименти на відміну від one-stage segmentation. Для цього потрібно проаналізувати роботу моделей YOLOv8 та Mask R-CNN для визначення пошкоджень автомобілів.

Модель нейронної мережі YOLO використовується для задач розпізнавання об'єктів та сегментації, а саме для швидкого і точного результату. Оскільки за один прохід нейронна мережа намагається пройти усе та визначити клас об'єкта і його розташування.

YOLOv8 – це архітектура YOLO, що була запропонована компанією Ultralytics у січні 2023 року [7]. Для покращення моделі компанія ввела кілька покращень. Одне із них це введення нових модулів у архітектуру моделі, а саме C2A та SPPF. Також, одним із нововведень було перехід на модульну архітектуру, а саме можна гнучко налаштувати Backbone і Head для виконання різних задач. YOLOv8 може виконувати різні задачі від класифікації до сегментації.

На рис. 2 наведено приклад модуля Backbone для YOLOv8.

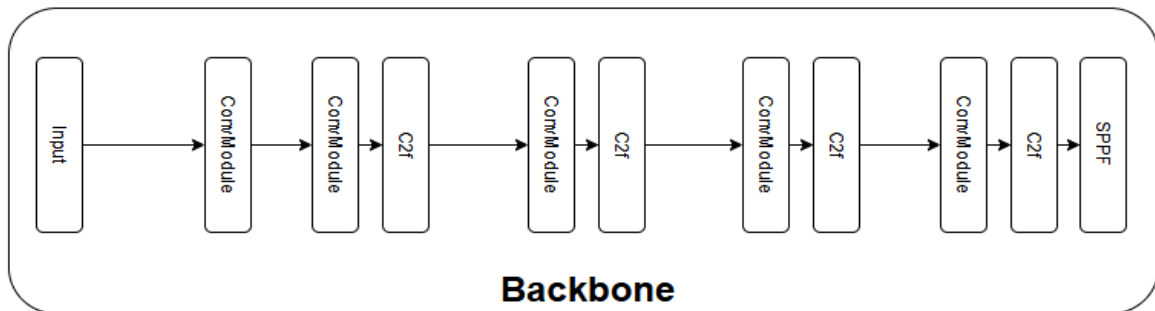


Рис. 2. Backbone модуль для YOLOv8

Mask R-CNN є розвитком Faster R-CNN, так як окрім класифікації та побудови обмежувальної рамки, в ньому також можна будувати маски для сегментації об'єктів. Перевагою Mask R-CNN є те що вона включає у своїй архітектурі гілку FCN(Fully Convolutional Neural Network) [8]. Дана мережа, що входить до архітектури Mask R-CNN зберігає просторову інформацію про об'єкт. Таким чином точність сегментації стає точнішою.

Також однією з особливостей Mask R-CNN є те що у нейронній мережі використовується RoIAlign замість шару RoIPool. І це є вагомою причиною для того, що Mask R-CNN має кращу точність у задачах сегментації. RoIPool працює подекуди некоректно на відмінну від RoIAlign. Причина полягає у тому, що RoIPool округлює координати областей до цілочисельних значень, а це в свою чергу призводить до не точних результатів. Також RoIAlign виконує обчислення координат для областей з плаваючою комою.

Результати дослідження

Ознайомившись із набором даних та методами, що використовується у даному дослідженні, було проведено експерименти. Для цього було використано confusion matrix, щоб визначити який з методів краще визначає класи пошкоджень автомобіля. На рис. 3 наведена confusion matrix для YOLOv8.

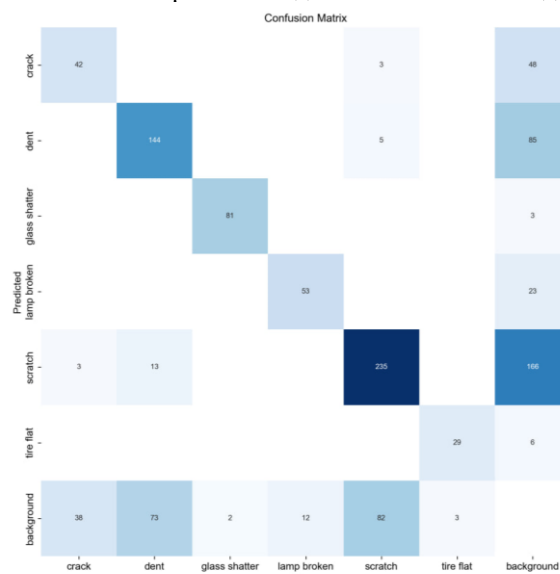


Рис. 3. Confusion matrix для моделі YOLOv8

Аналізуючи рис.3 можна побачити, що YOLOv8 добре справилася із розпізнаванням класів scratch та dent. Проте із решта класами у неї проблеми. Як можна побачити на Рис. 3, вона часто плутає різні класи між собою. А з класом tire flat модель розпізнала мало об'єктів. Також можна зазначити scratch найбільше плутався із фоном.

На рис. 4 наведена confusion matrix для Mask R-CNN

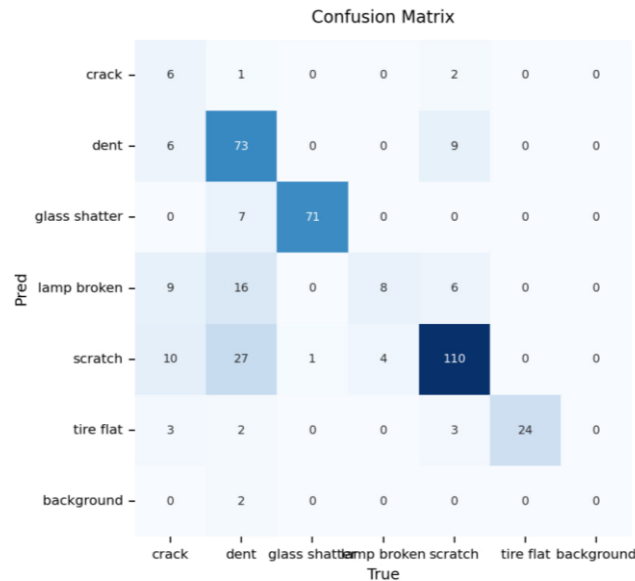


Рис. 4. Confusion matrix для моделі Mask R-CNN

Можна зазначити, що Mask R-CNN показує хороші результати, як і YOLOv8. Вона майже не плутає об'єкт (класи пошкодження) із фоном на відмінно від YOLOv8.

Також для оцінювання якості сегментації Mask R-CNN та YOLOv8 було використано метрики: mAP50 і mAP50-95. mAP (mean Average Precision) – це середня точність для всі класів. Вона обраховується за формулою:

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i,$$

де N – кількість класів; AP_i – середня точність для кожного класу.

У випадку mAP50 і mAP50-95 різниця між ними зводиться до того, що для першого показник IoU буде становити більшим за 0,5, а для mAP50-95 середній показник розраховується на основі 10 порогових значень IoU: від 0,50 до 0,95 з кроком в 0,05. У таблиці 1 наведені показники mAP50 і mAP50-95 для Mask R-CNN та YOLOv8 по класам.

Таблиця 1

Оцінка моделей Mask R-CNN та YOLOv8

Клас пошкодження	YOLOv8		Mask R-CNN	
	mAP50	mAP50-95	mAP50	mAP50-95
Crack	0.2142	0.0773	0.449	0.27
Dent	0.4483	0.209	0.61	0.369
Glass shatter	0.9788	0.8013	0.984	0.912
Lamp broken	0.6672	0.4878	0.851	0.704
Scratch	0.5335	0.2654	0.679	0.424
Tire flat	0.8364	0.72.44	0.942	0.93
	0.6131	0.4275	0.752	0.602

Провівши аналіз можна зазначити що Mask R-CNN справився краще, ніж YOLOv8. Оскільки Mask R-CNN перевершує по кожному класі пошкодження автомобіля. Однак, можна відміти, що показники в обох випадках мають високі показники в однакових класах, а також слабкі показники в однакових класах. Ще можна відмітити, що обом моделям важко справлятися із crack, а найточніше – це в glass shatter.

Висновок

В проведеному дослідженні використовувались дві різні моделі для сегментації для автоматичного виявлення пошкоджень автотранспорту на зображеннях. Проаналізувавши результати експериментів, можна визнати, що для виконання даної задачі краще використовувати Mask R-CNN. Хоча дана модель демонструє

перевагу над YOLOv8, її не варто використовувати в даній задачі, оскільки вона вимагає подальшого доопрацювання для покращення результатів. Адже експерименти проводились лише на окремих класах. Проте в загальному можна стверджувати, що Mask R-CNN краще справляється при виконанні типових задач.

Література

1. Телішевський, П., & Бойко, Н. (2024). Концепція автоматизованої системи оцінки пошкодження автотранспорту. *Herald of Khmelnytskyi National University. Technical sciences*, 335(3(1)), 251–254. <https://doi.org/10.31891/2307-5732-2024-335-3-34>
2. Qaddour, J., & Siddiq, S. A. (2023). Automatic damaged vehicle estimator using enhanced deep learning algorithm. *Intelligent Systems with Applications*, 18, 200192. <https://doi.org/10.1016/j.iswa.2023.200192>
3. Mallios, D., Xiaofei, L., McLaughlin, N., Rincon, J. M. D., Galbraith, C., & Garland, R. (2023). Vehicle Damage Severity Estimation for Insurance Operations Using In-The-Wild Mobile Images. *IEEE Access*, 11, 78644–78655. <https://doi.org/10.1109/ACCESS.2023.3299223>
4. Ojha, A., Sahu, S. P., & Dewangan, D. K. (2021). Vehicle Detection through Instance Segmentation using Mask R-CNN for Intelligent Vehicle System. *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 954–959. <https://doi.org/10.1109/ICICCS51141.2021.9432374>
5. Shim, S., & Cho, G.-C. (2020). Lightweight Semantic Segmentation for Road-Surface Damage Recognition Based on Multiscale Learning. *IEEE Access*, 8, 102680–102690. <https://doi.org/10.1109/ACCESS.2020.2998427>
6. Wang, X., Li, W., & Wu, Z. (2023). CarDD: A New Dataset for Vision-Based Car Damage Detection. *IEEE Transactions on Intelligent Transportation Systems*, 24(7), 7202–7214. <https://doi.org/10.1109/TITS.2023.3258480>
7. Sohan, M., Sai Ram, T., & Rami Reddy, Ch. V. (2024). A Review on YOLOv8 and Its Advancements. B. I. J. Jacob, S. Piramuthu, & P. Falkowski-Gilski (Ред.), *Data Intelligence and Cognitive Informatics* (с. 529–545). Springer Nature Singapore. https://doi.org/10.1007/978-981-99-7962-2_39
8. He, K., Gkioxari, G., Dollar, P., & Girshick, R. (2017). *Mask R-CNN*. 2961–2969. https://openaccess.thecvf.com/content_iccv_2017/html/He_Mask_R-CNN_ICCV_2017_paper.html

References

1. Telishevskiy, P., & Boiko, N. (2024). Kontseptsiiia avtomatyzovanoi systemy otsinky poshkodzhennia avtotransportu. *Herald of Khmelnytskyi National University. Technical sciences*, 335(3(1)), 251–254. <https://doi.org/10.31891/2307-5732-2024-335-3-34>
2. Qaddour, J., & Siddiq, S. A. (2023). Automatic damaged vehicle estimator using enhanced deep learning algorithm. *Intelligent Systems with Applications*, 18, 200192. <https://doi.org/10.1016/j.iswa.2023.200192>
3. Mallios, D., Xiaofei, L., McLaughlin, N., Rincon, J. M. D., Galbraith, C., & Garland, R. (2023). Vehicle Damage Severity Estimation for Insurance Operations Using In-The-Wild Mobile Images. *IEEE Access*, 11, 78644–78655. <https://doi.org/10.1109/ACCESS.2023.3299223>
4. Ojha, A., Sahu, S. P., & Dewangan, D. K. (2021). Vehicle Detection through Instance Segmentation using Mask R-CNN for Intelligent Vehicle System. *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 954–959. <https://doi.org/10.1109/ICICCS51141.2021.9432374>
5. Shim, S., & Cho, G.-C. (2020). Lightweight Semantic Segmentation for Road-Surface Damage Recognition Based on Multiscale Learning. *IEEE Access*, 8, 102680–102690. <https://doi.org/10.1109/ACCESS.2020.2998427>
6. Wang, X., Li, W., & Wu, Z. (2023). CarDD: A New Dataset for Vision-Based Car Damage Detection. *IEEE Transactions on Intelligent Transportation Systems*, 24(7), 7202–7214. <https://doi.org/10.1109/TITS.2023.3258480>
7. Sohan, M., Sai Ram, T., & Rami Reddy, Ch. V. (2024). A Review on YOLOv8 and Its Advancements. B. I. J. Jacob, S. Piramuthu, & P. Falkowski-Gilski (Ред.), *Data Intelligence and Cognitive Informatics* (с. 529–545). Springer Nature Singapore. https://doi.org/10.1007/978-981-99-7962-2_39
8. He, K., Gkioxari, G., Dollar, P., & Girshick, R. (2017). *Mask R-CNN*. 2961–2969. https://openaccess.thecvf.com/content_iccv_2017/html/He_Mask_R-CNN_ICCV_2017_paper.html