

Вікторія Смолій

доктор технічних наук, професор, професор кафедри інформаційних систем і технологій
Національний університет біоресурсів і природокористування України, Київ, Україна

ORCID ID 0000-0002-1268-7837

vmsmolij@nubip.edu.ua

Натан Смолій

магістрант кафедри інформаційних систем і технологій Національний технічний університет
України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна

ORCID ID: 0009-0002-3763-6726

hoibbitizukrainy@gmail.com

МОДУЛЬ МАШИННОГО ЗОРУ ДЛЯ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ

Анотація. В роботі для реалізації завдання сегментації зображень було сформовано спеціалізований датасет для об'єктів захисного спорядження. На етапі підготовки до тренування моделі до цього набору даних було застосовано процедури аугментації, зокрема віддзеркалення, обертання, масштабування та зміну яскравості, що дозволило суттєво підвищити різноманітність навчальних прикладів. Збільшення розміру датасету забезпечило більш збалансоване представлення даних і підвищило узагальнювальну здатність моделі. Отримані в роботі результати підтверджують доцільність і ефективність запропонованого авторами для даної проблеми окремого навчання моделей для сегментації та класифікації зображень.

Ключові слова: виявлення об'єктів, модель детекції, сегментація зображень, попереднє навчання моделі, процедури аугментації, донавчання, бінарні маски об'єктів.

Робота присвячена розробці та дослідженню методів виявлення об'єктів на зображеннях і у відеопотоці з використанням сучасних технологій глибокого навчання. Основна увага приділяється створенню, навчанню та практичному застосуванню моделей детектування об'єктів на основі нейронних мереж, що забезпечують високу точність і швидкість обробки візуальних даних [1]. У процесі виконання дослідження необхідно визначити об'єкти заданого типу на зображеннях, отриманих із відеокамери, а також розробити модель детекції, яка зможе ефективно ідентифікувати та локалізувати ці об'єкти в різних умовах освітлення, масштабу та перспективи. Для досягнення поставленої мети пропонується використати попередньо натренований детектор, провести його донавчання (fine-tuning) на спеціалізованій вибірці даних і адаптувати модель до конкретних умов задачі [2, 3]. Отримані результати моделювання та навчання підлягають оцінюванню на тестовій вибірці, що дозволяє проаналізувати точність, повноту та швидкодію моделі, а також визначити можливі напрямки її подальшої оптимізації та вдосконалення.

У межах дослідження було застосовано модель YOLOv11s для сегментації зображень, що містять об'єкти захисного спорядження. Проведено порівняльний аналіз якості результатів між підходами з попереднім навчанням моделі та zero-shot learning. Тестування моделей здійснювалося у середовищі Ultralytics із використанням мови програмування Python 3.11.

Для реалізації завдання сегментації зображень було сформовано спеціалізований датасет, що складався із 150 початкових зображень, на яких представлені об'єкти захисного спорядження. На етапі підготовки до тренування моделі до цього набору даних було застосовано процедури аугментації, зокрема віддзеркалення, обертання, масштабування та зміну яскравості, що дозволило суттєво підвищити різноманітність навчальних прикладів. У результаті цих дій обсяг датасету збільшився до 390 зображень, що забезпечило більш збалансоване представлення даних і підвищення узагальнювальної здатності моделі. Приклади зображень із сформованого датасету подано на рис. 1.



Рис. 1. Приклад зображення із сформованого датасету

Ненавчена модель виявилася малоприсадоною для прямого порівняння з донавченою, оскільки попереднє тренування було орієнтоване на розпізнавання базових категорій об'єктів. Для оцінювання ефективності роботи моделі в межах створеного датасету було проаналізовано графік зміни метрики IoU (Intersection over Union) протягом процесу навчання, що дало змогу відстежити динаміку покращення якості сегментації.

Оскільки донавчання виконувалося у віддаленому середовищі, застосовувалися рекомендовані параметри, визначені розробниками сервісу [4]. Серед них ключову роль для задачі сегментації відіграють такі гіперпараметри: оптимізатор Stochastic Gradient Descent, швидкість навчання (learning rate) — 0.01, коефіцієнт моменту (momentum) — 0.937 та регуляризаційний параметр weight decay — 0.0005, які забезпечують стабільність і збіжність процесу оптимізації моделі.

Функція втрат у моделі сегментації формується як поєднання трьох складових компонентів. Box loss реалізовано у вигляді CIOU Loss (Complete IoU), який відповідає за точність регресії обмежувальних прямокутників. Classification loss базується на Binary Cross-Entropy (BCE) Loss і використовується для класифікації об'єктів за їх типами. Mask loss також побудовано на BCE Loss, що забезпечує коректну бінаризацію масок об'єктів на зображенні. Додатково для підвищення стабільності та точності сегментації застосовувався Dice loss, який покращує збалансованість між передбаченими та реальними областями об'єктів.

Zero-shot модель була застосована для розв'язання задачі сегментації зображень з метою виявлення елементів захисного будівельного спорядження, таких як жилети, маски та рукавиці. Для цього використано SAM (Segment Anything Model) — універсальну модель сегментації, створену компанією Meta. Її робота ґрунтується на принципі promptable segmentation, тобто модель отримує підказку (prompt) від користувача і виконує точну сегментацію обраного об'єкта на зображенні.

Архітектурно SAM побудована на Vision Transformer (ViT), який відповідає за витягнення ознак, та масковому предикторі, що генерує бінарні маски об'єктів. Результати сегментації, отримані за допомогою SAM, характеризуються високою точністю відтворення контурів і форм об'єктів, а також універсальністю — модель ефективно працює з будь-якими типами об'єктів без потреби у додатковому донавчанні.

Результати роботи моделі zero-shot підходу за наперед заданими точками та порівняння з YOLO наведені на рис. 2, а та б відповідно.

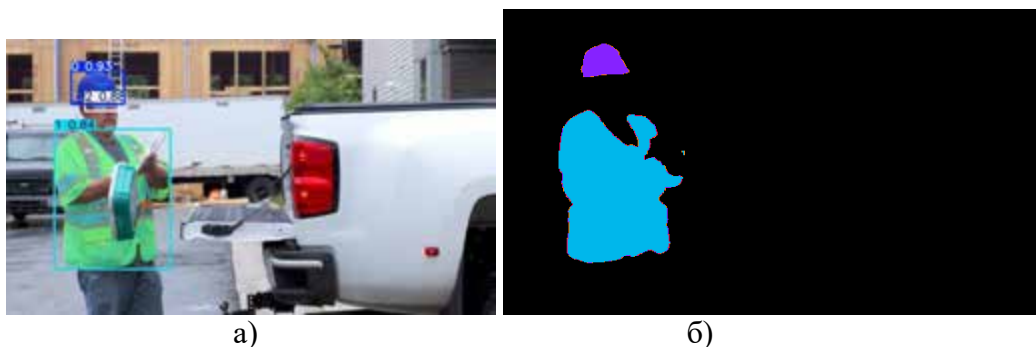


Рис. 2. Результати роботи моделі zero-shot підходу та порівняння з YOLO

Під час навчання моделі було застосовано спеціальний параметр, характерний для задач сегментації — $\text{mask_ratio} = 4$, який визначає ступінь наближення форми сегментованої області до графічних примітивів. У результаті значення метрики IoU (перетину з розміченими даними) досягло приблизно 40%, що свідчить про задовільну точність моделі для поставленого завдання. Додатково відзначено ефективність моделі за швидкістю — середній час виконання операції сегментації одного зображення становив близько 120 мс, що демонструє її придатність для використання в системах реального часу.

ВИСНОВКИ

У результаті проведених досліджень було сформовано спеціалізований датасет для сегментації зображень об'єктів захисного спорядження, що забезпечив якісну основу для навчання моделей комп'ютерного зору. Застосування процедур аугментації, таких як віддзеркалення, обертання, масштабування та зміна яскравості, дозволило суттєво розширити обсяг і різноманітність даних, підвищивши їх репрезентативність. Отримані результати свідчать, що збільшення обсягу датасету позитивно вплинуло на збалансованість даних і узагальнювальну здатність моделей. Проведене порівняння підтвердило ефективність використання підходу окремого навчання моделей для задач сегментації та класифікації. Таким чином, запропонований підхід може бути використаний як ефективне рішення для автоматизованої обробки зображень у завданнях технічного контролю та безпеки.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. V.M.Smolij, N.V.Smolij, S.P.Sayapin, Search and classification of objects in the zone of reservoirs and coastal zones, CEUR Workshop Proceedings 3666 (2024) 37–51. doi:ceur-ws.org/Vol-3666/ paper04.pdf10.1002/pssb.2220940160.
2. J.Tang, N.Xie, K.Li, Y.Liang, X.Shen, Trajectory tracking control for fixed-wing uav based on ddpg, Journal of Aerospace Engineering 37 (2024) 251–283. doi:10.1061/JAEEZ.ASENG-5286.
3. V.M.Smolij, N.V.Smolij, O.Y.Kovalenko, M.Z.Shvydenko, Channel extractor for uav ppm signals, CEUR Workshop Proceedings 3917 (2025) 226–236. doi:https://ceur-ws.org/Vol-3917/.
4. <https://docs.ultralytics.com/modes/train/#train-settings>

MINISTRY OF EDUCATION
AND SCIENCE OF UKRAINE

NATIONAL UNIVERSITY
OF LIFE AND ENVIRONMENTAL
SCIENCES OF UKRAINE

FACULTY OF INFORMATION
TECHNOLOGY

МІНІСТЕРСТВО ОСВІТИ
І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
БІОРЕСУРСІВ І
ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ

ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ
ТЕХНОЛОГІЙ

PROCEEDINGS

XIII International scientific
and practical conference

**GLOBAL AND
REGIONAL PROBLEMS OF
INFORMATIZATION IN
SOCIETY AND
NATURE USING
'2025**

13-14 November 2025

Kyiv, NULES of Ukraine

Kyiv 2025

МАТЕРІАЛИ

XIII Міжнародної науково-
практичної конференції

**ГЛОБАЛЬНІ ТА
РЕГІОНАЛЬНІ ПРОБЛЕМИ
ІНФОРМАТИЗАЦІЇ В
СУСПІЛЬСТВІ І
ПРИРОДОКОРИСТУВАННІ
'2025**

13-14 листопада 2025 року

Київ, НУБіП України

Київ 2025

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ
І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

МАТЕРІАЛИ

ХІІІ Міжнародної науково-практичної конференції

ГЛОБАЛЬНІ ТА РЕГІОНАЛЬНІ ПРОБЛЕМИ ІНФОРМАТИЗАЦІЇ В СУСПІЛЬСТВІ І ПРИРОДОКОРИСТУВАННІ '2025

13-14 листопада 2025 року

Київ, НУБіП України

Київ 2025

УДК 004

Рекомендовано до друку вченою радою факультету інформаційних технологій Національного університету біоресурсів і природокористування України (протокол № 4 від 18.12.2025).

Укладач: д.т.н., доцент Шкарупило В.В.

Збірник матеріалів XIII Міжнародної науково-практичної конференції "Глобальні та регіональні проблеми інформатизації в суспільстві і природокористуванні '2025", 13–14 листопада 2025 року, НУБіП України, Київ. – К.: НУБіП України, 2025. – 206 с.

Відповідальність за зміст публікацій несуть автори.

© Національний університет біоресурсів
і природокористування України, 2025