

<https://doi.org/10.62405/osiconf.2026.02>

Артиух В. О.,

Державний університет «Київський авіаційний інститут», Київ, Україна,
ORCID: 0009-0007-5847-526X, e-mail: Artiukh_Vladyslav@protonmail.com

ДОСЛІДНИЦЬКИЙ ПІДХІД ДО ВИЯВЛЕННЯ ТА СУПРОВОДЖЕННЯ РУХОМИХ ОБ'ЄКТІВ У ВІДЕОПОСЛІДОВНОСТЯХ

АНОТАЦІЯ. Метою роботи є дослідження впливу параметрів виявлення рухомих об'єктів на точність і стабільність їх подальшого супроводження у відеопослідовностях. Запропоновано експериментальний підхід, що передбачає послідовне виконання машинного виявлення об'єктів, міжкадрової асоціації та формування їх траєкторій із варіюванням порога достовірності детекції. Оцінювання здійснено за показниками точності та повноти виявлення, середньої точності детекції, якості й стабільності супроводження, кількості змін ідентифікаторів та швидкості обробки відео. Встановлено, що надмірне зниження порога достовірності збільшує кількість хибних детекцій і помилкових асоціацій, тоді як його надмірне підвищення призводить до втрати короткочасно перекритих та малоконтрастних об'єктів. Найкраща стабільність траєкторій досягається за проміжних значень параметра. Результати підтверджують доцільність спільного налаштування параметрів детекції та супроводження з урахуванням компромісу між точністю, стабільністю ідентифікації та обчислювальною ефективністю.

КЛЮЧОВІ СЛОВА: машинне навчання; відеопослідовність; виявлення об'єктів; супроводження об'єктів; траєкторія; комп'ютерний зір.

ВСТУП. Автоматичне виявлення та супроводження рухомих об'єктів є одним із базових завдань сучасного комп'ютерного зору та використовується у системах відеоспостереження, інтелектуального транспорту, робототехніки, аналізу поведінки та автоматизованого контролю середовища. На відміну від оброблення окремого зображення, аналіз відеопослідовності потребує не лише визначення положення об'єкта у поточному кадрі, але й встановлення його відповідності об'єкту, виявленому в попередніх кадрах, зі збереженням постійного ідентифікатора та формуванням безперервної траєкторії руху.

Одним із поширених способів розв'язання цього завдання є підхід tracking-by-detection, відповідно до якого спочатку для кожного кадру виконується виявлення об'єктів засобами машинного навчання, після чого отримані детекції асоціюються з уже сформованими траєкторіями. Така послідовність забезпечує можливість незалежного вдосконалення компонентів системи, однак водночас створює залежність кінцевої якості супроводження від характеристик детектора. Результати порівняльного дослідження різних комбінацій засобів виявлення та супроводження, наведені в, свідчать, що зміна детектора істотно впливає не лише на показники локалізації об'єктів, але й на кінцеві характеристики їх супроводження (Gragnaniello et al., 2023). Тому оцінювання лише точності детекції є недостатнім для визначення ефективності системи аналізу відеопослідовностей.

Додатковою проблемою є коректна міжкадрова асоціація об'єктів за умов часткового або повного перекриття, різкої зміни напрямку руху, низької достовірності окремих детекцій та тимчасового зникнення об'єкта з поля зору. Помилка асоціації може призвести до фрагментації траєкторії або присвоєння тому самому фізичному об'єкту нового ідентифікатора. У іншій роботі показано, що вдосконалення процедури асоціації детекцій із траєкторіями здатне суттєво підвищити стабільність багатооб'єктного супроводження (Stadler & Beyerer, 2023). Stanojevic i Todorovic (Stanojevic & Todorovic, 2024) у запропонованому алгоритмі BoostTrack показали, що врахування достовірності детекцій і сформованих траєкторій під час обчислення міри подібності дає змогу підвищити коректність їх міжкадрової асоціації та зменшити кількість помилкових змін ідентифікаторів.

Таким чином, невирішеним практичним питанням залишається вибір таких параметрів етапів виявлення та асоціації, за яких забезпечується баланс між кількістю коректно виявлених об'єктів, стабільністю їх ідентифікаторів і швидкістю оброблення відеопотоку.

Метою роботи є експериментальне дослідження впливу параметрів виявлення та супроводження рухомих об'єктів на точність і стабільність формування їх траєкторій у відеопослідовностях.

Методи. Дослідження виконано відповідно до схеми *tracking-by-detection*, що включає чотири послідовні етапи: отримання кадру відеопослідовності, машинне виявлення об'єктів, асоціацію отриманих детекцій із траєкторіями попередніх кадрів та оновлення або створення траєкторій. Програмну реалізацію експерименту виконано мовою Python 3.11 із використанням бібліотек Ultralytics 8.3, OpenCV 4.10 та PyTorch 2.4. Для виявлення об'єктів застосовано попередньо навчену на наборі даних COCO модель YOLOv8n із вагами yolov8n.pt.

На першому етапі засобами OpenCV відеопотік послідовно розбивався на окремі кадри. Перед поданням до детектора кадри масштабувалися до вхідного розміру 640×640 пікселів зі збереженням співвідношення сторін. Для кожного кадру детектор формував множину кандидатів, кожен з яких характеризувався координатами обмежувальної рамки, класом об'єкта та значенням достовірності виявлення. У дослідженні враховувалися об'єкти класу «person». Для усунення дубльованих обмежувальних рамок використовувалося немаксимальне пригнічення (Non-Maximum Suppression, NMS) із порогом IoU 0,50. Детекції, значення достовірності яких було нижчим за досліджуваний поріг T_{conf} , виключалися з подальшого аналізу.

Формально множину детекцій, що передаються до етапу супроводження, визначено як

$$D^* = \{d_i \in D \mid c_i \geq T_{conf}\}, \quad (1)$$

де D — множина сформованих детекцій, d_i — окрема детекція, c_i — її значення достовірності, а T_{conf} — встановлений поріг достовірності.

На другому етапі виконувалося супроводження об'єктів із використанням алгоритму SORT. Прогноз положення кожного активного об'єкта в наступному кадрі здійснювався за допомогою фільтра Калмана, а міжкадрова асоціація детекцій із прогнозованими траєкторіями виконувалася угорським алгоритмом на основі міри Intersection over Union (IoU) між обмежувальними рамками. Мінімальний IoU для встановлення відповідності становив 0,30. Параметр максимально допустимої тривалості відсутності детекції для активної траєкторії max_age було встановлено рівним 30 кадрам, а мінімальну кількість послідовних спостережень для підтвердження нової траєкторії min_hits — 3 кадрам. Якщо для детекції не знаходилося відповідної активної траєкторії, їй призначався новий ідентифікатор. Якщо об'єкт не виявлявся протягом понад 30 послідовних кадрів, відповідна траєкторія завершувалася.

Експериментальний набір містив 15 відеофрагментів загальною кількістю 18 750 кадрів. Роздільна здатність відео становила 1280×720 пікселів, частота — 25 кадрів за секунду. Для врахування різних умов супроводження до набору включено сцени з низькою, середньою та високою щільністю рухомих об'єктів, а також епізоди з частковими перекриттями та короточасним зникненням об'єктів.

Основним змінюваним параметром обрано поріг достовірності детекції T_{conf} . Досліджено три конфігурації: А — 0,25, В — 0,40 та С — 0,55. В усіх трьох конфігураціях використовувалися однакові ваги моделі YOLOv8n, вхідна роздільна здатність 640×640 пікселів, поріг NMS IoU 0,50 та параметри трекара: $max_age = 30$, $min_hits = 3$ та $IoU\ threshold = 0,30$. Таким чином, між експериментальними конфігураціями змінювався лише поріг T_{conf} , що дозволило безпосередньо оцінити його вплив на точність виявлення та стабільність подальшого формування траєкторій.

Експерименти виконувалися в середовищі Windows 10 на комп'ютері з процесором Intel Core i7, 16 ГБ оперативної пам'яті та графічним адаптером NVIDIA GeForce RTX 3060 6 ГБ. Для забезпе-

чення порівнюваності результатів усі конфігурації тестувалися на однаковому наборі відеофрагментів і за незмінного апаратного та програмного середовища.

Якість виявлення оцінювалася за показниками Precision, Recall та mAP. Детекція вважалася правильною за значення IoU з еталонною обмежувальною рамкою не менше 0,50.

Показники точності та повноти виявлення розраховувалися за співвідношеннями

$$\text{Precision} = \frac{TP}{TP + FP}, \text{ Recall} = \frac{TP}{TP + FN}, \quad (2)$$

де TP — кількість правильно виявлених об'єктів, FP — кількість хибних спрацьовувань, FN — кількість пропущених об'єктів.

Узагальнену якість роботи детектора оцінено за показником mAP@0.5:0.95.

Для оцінювання супроводження використано показники HOTA, IDF1, MOTA та IDSW, розраховані засобами TrackEval. HOTA застосовано як комплексний показник, що одночасно враховує якість виявлення та асоціації об'єктів. IDF1 характеризує коректність збереження ідентичності об'єкта протягом траєкторії, MOTA відображає загальну точність супроводження з урахуванням пропущених і хибних детекцій та помилок ідентифікації, а IDSW визначає кількість випадків неправильної зміни ідентифікатора об'єкта.

Окремо визначалася продуктивність оброблення за кількістю кадрів за секунду — FPS. Значення FPS розраховувалося як відношення загальної кількості оброблених кадрів до тривалості виконання повного циклу детекції та супроводження, без урахування часу завантаження відеофайлу та збереження результатів на диск. Необхідність одночасного врахування точності та швидкодії пов'язана з практичним використанням технологій супроводження в системах реального часу. Як показано в іншому дослідженні, підвищення точності алгоритму без урахування його обчислювальної складності не гарантує придатності рішення для практичного опрацювання відеопотоку (Medaglini et al., 2025).

РЕЗУЛЬТАТИ ТА ОБГОВОРЕННЯ. За результатами експерименту встановлено, що зміна порога достовірності впливає не лише на співвідношення правильних і хибних детекцій, але й на стабільність подальшої асоціації об'єктів між кадрами.

Для конфігурації А з порогом 0,25 отримано Precision 84,2% та Recall 91,1%. Низьке значення порога забезпечило максимальну повноту виявлення, однак одночасно збільшило кількість низькодостовірних детекцій. Частина таких детекцій помилково включалась у процес асоціації, що спричинило 143 зміни ідентифікаторів. Значення HOTA становило 65,8%, IDF1 — 70,9%, а MOTA — 67,6%.

За використання конфігурації В з порогом 0,40 показник Precision збільшився до 89,6% при Recall 87,3%. Водночас кількість змін ідентифікаторів зменшилася до 94. Для цієї конфігурації отримано максимальні значення показників супроводження: HOTA — 71,2%, IDF1 — 77,4% та MOTA — 73,1%. Це свідчить про те, що відсікання частини низькодостовірних детекцій зменшує кількість помилкових відповідностей без критичної втрати об'єктів.

Подальше підвищення порога до 0,55 у конфігурації С збільшило Precision до 92,4%, однак Recall знизився до 79,8%. Особливо це проявлялося для малих, частково перекритих та короткочасно видимих об'єктів. Пропуски детекцій призводили до розривів траєкторій і повторного створення ідентифікаторів після відновлення видимості об'єкта. Унаслідок цього IDSW зріс до 119, HOTA знизився до 67,5%, IDF1 — до 72,6%, а MOTA — до 69,3%.

Середнє значення mAP детектора на тестовому наборі становило 56,8%. Отриманий результат свідчить, що висока якість детекції сама по собі не визначає найкращу конфігурацію системи супроводження, оскільки різні правила відсікання результатів детектора змінюють набір спостережень, доступних алгоритму асоціації.

Порівняння основних метрик супроводження для трьох досліджених конфігурацій наведено на рис. 1.

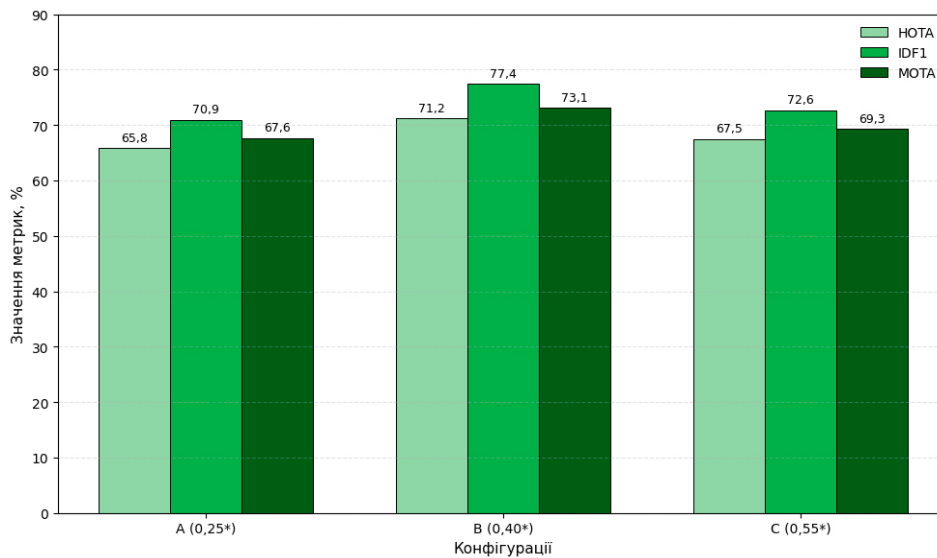


Рис. 1. Порівняння показників NOTA, IDF1 та MOTA для досліджених конфігурацій виявлення і супроводження

Джерело: спобудовано автором.

За даними рис. 1 найбільш збалансованою є конфігурація В, для якої одночасно отримано найбільші значення NOTA, IDF1 та MOTA. Порівняно з конфігурацією А значення IDF1 збільшилося на 6,5 відсоткового пункта, а кількість змін ідентифікаторів скоротилася приблизно на 34,3%. Це підтверджує, що надлишкова кількість низькодостовірних детекцій може негативно впливати на супроводження навіть за високого Recall.

Швидкість оброблення для конфігурацій А, В і С становила відповідно 31,8, 33,7 та 35,1 кадрів за секунду. Незначне зростання FPS зі збільшенням порога пояснюється зменшенням кількості детекцій, що передаються до процедури асоціації. Таким чином, збільшення порога частково знижує обчислювальне навантаження, однак після певного значення супроводжується помітним погіршенням стабільності траєкторій.

Характер зміни IDF1 та FPS залежно від порога достовірності наведено на рис. 2.

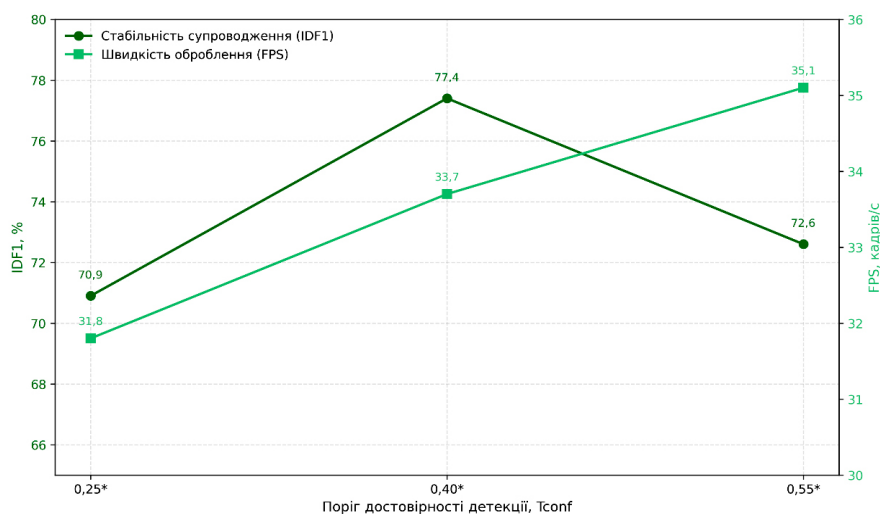


Рис. 2. Залежність стабільності супроводження та швидкості оброблення від порога достовірності детекції

Джерело: спобудовано автором.

Отримана залежність має нелінійний характер. Зі збільшенням T_{conf} від 0,25 до 0,40 якість супроводження підвищується внаслідок зменшення кількості хибних детекцій та помилкових асоціацій. Після підвищення порога до 0,55 тенденція змінюється: швидкість оброблення продовжує дещо зростати, але IDF1 та НОТА знижуються через збільшення кількості пропущених об'єктів і розривів траєкторій.

Додатковий аналіз відеосцен показав, що вплив параметра є найбільш вираженим за високої щільності об'єктів та наявності перекриттів. У сценах із невеликою кількістю ізольованих об'єктів різниця IDF1 між конфігураціями становила близько 3,1 відсоткового пункта, тоді як у складних сценах із частими перекриттями вона досягала 9,4 відсоткового пункта. Це пояснюється тим, що в умовах неоднозначності просторова близькість кількох об'єктів збільшує ймовірність неправильної міжкадрової асоціації.

Отже, результати експерименту підтверджують взаємозалежність етапів детекції та супроводження. Зменшення кількості хибних спрацьовувань детектора позитивно впливає на стабільність траєкторій лише до моменту, коли втрати правильних детекцій починають спричиняти їх фрагментацію. Відповідно, оптимальне значення порога не повинно визначатися виключно за Precision або Recall, а має обиратися за сукупністю характеристик виявлення, асоціації та продуктивності.

ВИСНОВКИ. Проведене дослідження дозволило встановити закономірність між параметрами виявлення рухомих об'єктів та стабільністю їх супроводження у відеопослідовностях. Зниження порога достовірності підвищує повноту виявлення, однак збільшує кількість низькодостовірних спостережень, що підвищує ймовірність помилкової міжкадрової асоціації та зміни ідентифікаторів. Надмірне підвищення порога, навпаки, покращує точність окремих детекцій і частково збільшує швидкість, але призводить до пропуску об'єктів, фрагментації траєкторій і повторної ідентифікації.

Експериментально найбільш ефективним виявився проміжний режим налаштування з порогом 0,40, за якого досягнуто максимального значення IDF1 77,4% при швидкості оброблення 33,7 кадрів за секунду та найменшій кількості змін ідентифікаторів — 94.

Практично це означає, що параметри детектора в інформаційній технології виявлення рухомих об'єктів доцільно оптимізувати не ізольовано, а з урахуванням їх впливу на наступний етап супроводження. Перспективним напрямом подальших досліджень є автоматичний адаптивний вибір параметрів детекції та асоціації залежно від щільності об'єктів, інтенсивності їх руху та складності поточної відеосцени.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

- Gragnaniello, D., Greco, A., Saggese, A., Vento, M., & Vicinanza, A. (2023). Benchmarking 2D multi-object detection and tracking algorithms in autonomous vehicle driving scenarios. *Sensors*, 23(8), 4024. <https://doi.org/10.3390/s23084024>
- Medaglini, A., Peccerillo, B., & Bartolini, S. (2025). High-performance multi-object tracking for autonomous driving in urban scenarios with heterogeneous embedded boards. *IEEE Access*, 13, 8649–8663. <https://doi.org/10.1109/ACCESS.2024.3484129>
- Stadler, D., & Beyerer, J. (2023). An improved association pipeline for multi-person tracking. In 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (pp. 3170–3179). IEEE. <https://doi.org/10.1109/CVPRW59228.2023.00319>
- Stanojevic, V. D. & Todorovic, B. T. (2024). BoostTrack: Boosting the similarity measure and detection confidence for improved multiple object tracking. *Machine Vision and Applications*, 35, Article 53. <https://doi.org/10.1007/s00138-024-01531-5>

Artiukh V. O.,

State University «Kyiv Aviation Institute», Kyiv, Ukraine,
ORCID: 0009-0007-5847-526X, e-mail: Artiukh_Vladyslav@protonmail.com

A RESEARCH APPROACH TO THE DETECTION AND TRACKING OF MOVING OBJECTS IN VIDEO SEQUENCES

ABSTRACT. *The aim of the study is to investigate the influence of moving object detection parameters on the accuracy and stability of their subsequent tracking in video sequences. An experimental approach based on sequential machine-learning detection, inter-frame association, and trajectory formation with variation of the detection confidence threshold is proposed. The evaluation is performed using detection precision and recall, average detection accuracy, tracking quality and stability, the number of identity changes, and video processing speed. The results demonstrate that an excessively low confidence threshold increases false detections and incorrect associations, whereas an excessively high threshold causes the loss of temporarily occluded and low-contrast objects. The most stable trajectories are obtained at intermediate parameter values. The findings confirm the need for joint adjustment of detection and tracking parameters while considering the trade-off between detection accuracy, identity stability, and computational efficiency.*

KEYWORDS: *machine learning; video sequence; object detection; object tracking; trajectory; computer vision.*