

УДК 004.932.72'1:004.8]:621.397.43

ДОСЛІДЖЕННЯ ПИТАННЯ МОНІТОРИНГУ ДІЙ ЛЮДЕЙ У ПРИМІЩЕННІ НА ОСНОВІ АНАЛІЗУ ВІДЕОПОТОКУ З КАМЕР ВІДЕОСПОСТЕРЕЖЕННЯ

Ісаєв Є.А.

Науковий керівник – к.н.т., доц. Яковлева О.В.

Харківський національний університет радіоелектроніки, каф. ІНФ,
м. Харків, Україна

e-mail: yevhenii.isaiev@nure.ua

This work is devoted to the study of the question of monitoring the actions of people indoors based on the analysis of the video stream from surveillance cameras with the aim of further developing a plugin for the iSpy application. Special attention is paid to the study of trackers from BoT-SORT, ByteTrack from the modern Ultralytics framework, as well as experiments with different variants of the YOLOv8 model, which is designed for detection and classification. The results of the research showed that it is reasonable to use the ByteTrack tracker with the YOLOv8n model for implementation in the iSpy application as the fastest tracker and one with high enough accuracy to accomplish the task.

У сучасному світі, де технології невпинно вдосконалюються, застосування камер відеоспостереження стає не лише розповсюдженим, але й необхідним з метою забезпечення безпеки, контролю та ефективного управління різноманітними просторами. Існує велика кількість застосунків, яка використовує штучний інтелект для аналізу зображень. За допомогою штучного інтелекту можна розпізнавати обличчя людей, їх емоції, маски, натовпи, вік людей, та багато інших можливостей. Існують як безкоштовні версії застосунків, так і платні їх версії. Серед найпопулярніших визначають iSpy, Xeoma, Zoneminder та IP Camera Viewer. Було виявлено, що більшість із них мають можливість розпізнавати образи за допомогою штучного інтелекту, але на жаль, ці застосунки не є безкоштовними.

Варто зазначити, що за минулі роки, напрямок детектування та розпізнавання образів у штучному інтелекті мав дуже стрімкий розвиток [1, 2]. Трекінг об'єктів, що базується на детекції та розпізнаванні, також має суттєвий прогрес. Але, на жаль, зазначені вище застосунки ще використовують застарілі алгоритми аналізу образів, які мають велику похибку. Також, наведені вище застосунки не мають трекінгу об'єктів, що може призвести у майбутньому до багатьох проблем, наприклад, проблема оклюзії, через яку детектор не зможе виявити об'єкт. Застосунок iSpy має відкритий вихідний код, завдяки якому можливо створити плагін, або модифікувати застосунок, додавши туди можливість трекінгу об'єктів.

Робота присвячена дослідженню питання трекінгу об'єктів для розв'язання задачі відстеження дій людей в приміщенні з метою подальшого розроблення плагіну трекінгу об'єктів для застосунку iSpy.

Для дослідження було обрано трекери з фреймворку від Ultralytics YOLO [3]. Даний фреймворк було обрано через те, що він має значні переваги серед його аналогів, такі як: обробка відеопотоків в режимі реального часу, підтримка безлічі алгоритмів та конфігурації відстеження, простий Python API та опції CLI для швидкої інтеграції та розгортання, простий у використанні, завдяки спеціально навченим моделям YOLO.

Існує велика кількість моделей YOLO, які розроблені спеціально для певних задач, такі як: детекція об'єктів, сегментація, визначення поз, визначення орієнтації об'єктів, класифікація об'єктів. У нашому випадку, було використано лише моделі для детекції та класифікації об'єктів, такі як YOLOv8n, YOLOv8m, YOLOv8x. У ході дослідження взаємодії моделей та трекерів, було виявлено, що найлегша модель YOLOv8n справлялася із задачею детекції та класифікації об'єктів, але мала помилки. Також, було досліджено модель YOLOv8x, яка є найважчою серед усіх моделей детекції, вона є найповільнішою через те, що вона використовує найбільше ресурсів комп'ютера, але вона мала найкращі результати при виконанні задачі.

Трекінг об'єктів можна класифікувати на дві основні категорії, виходячи з типу та функціональності доступних трекерів – це SOT (Single Object Tracking), та MOT (Multiple Object Tracking). У класі SOT трекерів, перший кадр позначається прямокутником, який вказує на місцезнаходження об'єкта, який ми хочемо відстежувати. Згодом, цей об'єкт відстежується в наступних кадрах за допомогою алгоритму відстеження. У більшості реальних застосунках, ці трекери використовуються з детектором об'єктів. З розвитком швидкості трекерів, стало можливо застосовувати відстеження для кількох об'єктів в одному, або послідовних кадрах. Стандартний підхід в алгоритмах відстеження декількох об'єктів (MOT) - це відстеження за виявленням, де виявлення (обмежувальні рамки, що ідентифікують об'єкти у відеокадрах) керують процесом відстеження. Ці детектори пов'язані з підтримкою однакових ідентифікаторів для одних і тих самих цілей на різних кадрах. Фреймворк YOLOv8 від Ultralytics, має два MOT трекери: BoT-SORT, ByteTrack. BoT-SORT – надійний сучасний трекер, який поєднує в собі переваги інформації про рух та зовнішній вигляд, компенсацію руху камери та більш точний вектор стану фільтра Калмана. BoT-SORT посідає перше місце в наборах даних MOTChallenge на тестових наборах MOT17 і MOT20 за всіма основними метриками MOT: MOTA, IDF1 та HOTA (для MOT17 отримано 80,5 MOTA, 80,2 для IDF1 та 65,0 для HOTA) (рис. 1) [3]. ByteTrack – простий, ефективний і загальний метод асоціації, який відстежує, пов'язуючи кожне поле виявлення, а не тільки ті, що мають високу оцінку. Для блоків визначення з низькою оцінкою ми використовуємо їхню схожість з трекетами, щоб знайти валідні об'єкти та відфільтрувати фонові виявлення. Нижче наведено результати тестування трекера ByteTrack у MOTChallenge на тестових наборах (рис. 2) [3]:

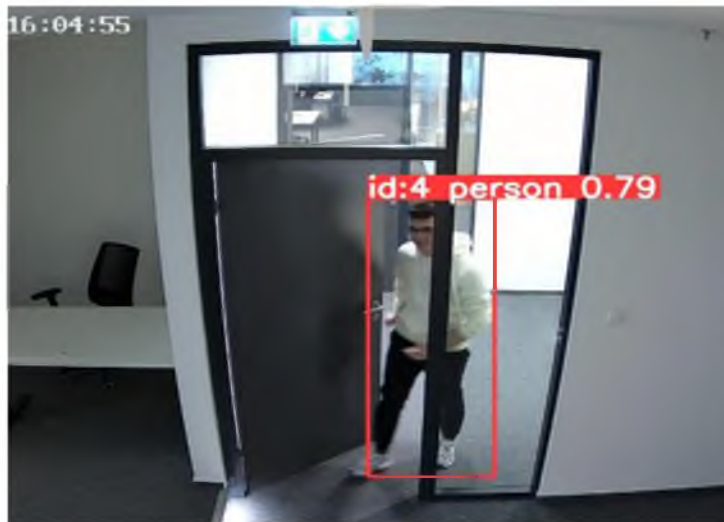


Рисунок 1 – Дослідження трекеру BoT-SORT

Dataset	MOTA	IDF1	HOTA	MT	ML	FP	FN	IDs	FPS
MOT17	80.3	77.3	63.1	53.2%	14.5%	25491	83721	2196	29.6
MOT20	77.8	75.2	61.3	69.2%	9.5%	26249	87594	1223	13.7

Рисунок 2 – Результати тестування трекера ByteTrack у MOTChallenge [3]

Дослідивши трекери ByteTrack та BoT-SORT для відео з камери відеоспостереження із застосування моделей YOLOv8n, та YOLOv8x для детекції постаті людини, можна зробити такі висновки: для детектування та класифікації достатньо можливостей YOLOv8n, яка має значно кращу швидкість; порівняння результатів трекерів ByteTrack та BoT-SORT на основі YOLOv8n показало, що трекери мають схожу точність, обидва трекери в складних випадках, наприклад, за наявності перекриття об'єктів, допускалися помилки, але у випадку BoT-SORT їх було менше ніж у ByteTrack, але швидкості BoT-SORT є недостатньо для обробки відео потоку у режимі реального часу. Таким чином, дослідження показали, що для впровадження у застосунок для відстеження дій людей в приміщенні доцільно використовувати трекер ByteTrack з моделлю YOLOv8n.

Список використаних джерел:

1. Yakovleva, O., Kovtunenکو, A., Liubchenko, V., Honcharenko, V., & Kobylin, O. (2023). Face Detection for Video Surveillance-based Security System. CEUR Workshop Proceedings Vol. 3403. pp. 69-86. ISSN 1613-0073
2. Yakovleva, O., Kovač, M., Ardasov, V. & Yeremenko, I. (2023). Study on adding functionality to the Zoom online conference system for monitoring the participant activities. Public Administration and Regional Development, 19(1), pp. 158–184.
3. Ultral Multi-Object Tracking with Ultralytics YOLO. (2024). Retrieved from <https://docs.ultralytics.com/modes/track/>