

Денис Віннічук

Аспірант кафедри комп'ютерних наук НУБіП України

0009-0001-9441-7696

Vinniboy2709@gmail.com

РОЗПІЗНАВАННЯ, КЛАСИФІКАЦІЯ ТА ВІДСТЕЖЕННЯ ОБ'ЄКТІВ В УМОВАХ ДИНАМІЧНОГО СПОСТЕРЕЖЕННЯ

Анотація. Було розглянуто принципи побудови системи спостереження для дронів, яка дозволяє розпізнавати, відстежувати та визначати положення об'єктів в реальному часі. Запропонована система враховує динаміку камери за допомогою інформації від датчиків висотоміру та дальноміру, а також значень кутів курсу, тангажу та крену.

Ключові слова: комп'ютерний зір, нейромережі, штучний інтелект, трекінг.

Задачі аналізу об'єктів на фото та відео в динамічних складних сценах сьогодні є надзвичайно актуальними в багатьох сферах, особливо в галузі оборони в період повномасштабного вторгнення. Завдання аналізу фото та відео умовно можна поділити на такі складові: Розпізнавання об'єкту; Класифікація з відомого набору класів; Відстеження, тобто присвоєння ідентифікатору між кадрами; Визначення фізичних характеристик для прийняття подальших рішень.

В ході роботи було розглянуто задачу де дрон з встановленою камерою веде спостереження, а саме оглядає плоску поверхню по якій рухаються деякі об'єкти, які необхідно розпізнати, класифікувати та визначити їх координати відносно камери. Важливо відмітити, що система повинна працювати в режимі реального часу. Дрон є відносно нерухомим (тобто виконує тільки задачу спостереження), однак камера на дроні може обертатися по трьом осям та збільшувати або зменшувати масштаб. Зображення передаються на сервер за допомогою RTMP потоку. Також дрон здатний передавати метадані, а саме значення кутів обертання камери та коефіцієнту зумування. Додатково дрон обладнаний дальноміром і висотоміром та передає дані від цих датчиків (рис. 1). Метадані є критично важливими при визначенні характеристик розпізнаних об'єктів.

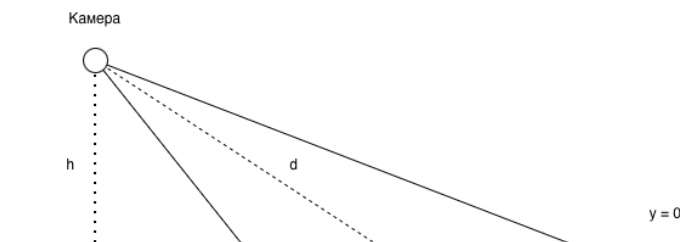


Рис. 1. Схема розташування поля зору камери. d – значення відстані від дальноміру, h – значення висоти

Завдання розпізнавання та класифікації об'єктів ефективно здатні виконувати сучасні нейромережеві підходи. Однією з найефективніших архітектур є одностадійна згортова нейронна мережа прямого поширення YOLO (You only look once) [1]. Така архітектура поступається в метриках точності двохстадійним підходам. Однак, головною перевагою YOLO є можливість розпізнавати та класифікувати об'єкти в режимі реального часу при відносно невеликих обчислювальних витратах. В задачі, що розглядається було використано YOLO v8m, яку було донавчено на різноманітних відкритих датасетах з військовою технікою.

При виявленні об'єкту часто потрібно також визначити його певні фізичні характеристики, такі як швидкість, траєкторію, розмір, тощо. Для цього потрібно

ідентифікувати координати розташування об'єкту. Це можливо зробити, якщо врахувати динамічну поведінку камери. Камера може змінювати кут курсу (yaw), тангажу (pitch) та крену (roll). Обертання здійснюється в порядку yaw-roll-pitch, що зумовлене особливістю конструкції шарнірів камери DJI Zenmuse H20. Також камера здатна виконувати зумування, тобто зміну масштабу камери. Додатково експериментально було визначено кути огляду камери в градусах (горизонтальний та вертикальний) FOV_{base} .

Нейромережа YOLO повертає локальні піксельні координати u, v обмежувальної рамки (які масштабуються в діапазон $[-1;1]$) та клас об'єкту на зображення. Щоб визначити глобальні координат спочатку потрібно знайти напрямляючий одиничний вектор від камери на об'єкт, який не буде змінюватись при обертанні камери чи зумуванні. Для цього отримане зображення проєктується на одиничну сферу навколо камери за допомогою відомих значень куту огляду і зуму. При чому впливу зуму на кут огляду визначається як:

$$\tan\left(\frac{FOV_{new}}{2}\right) = \frac{\tan\left(\frac{FOV_{base}}{2}\right)}{zoom_ratio} \quad (1)$$

Таким чином кожен об'єкт отримує координати $v_{local} = (x_{local}, y_{local}, z_{local})$ на одиничній сфері навколо камери. Наступним кроком є врахування обертання камери. Оскільки порядок обертання осей відомий, будується матриця обертання, яка є добутком матриць обертання по окремим осям в строго визначеному порядку, оскільки множення матриць є некомутативною операцією:

$$R = R_{yaw} R_{pitch} R_{roll}, \quad (2)$$

де R_{yaw} – матриця обертання навколо осі Y, R_{pitch} – матриця обертання навколо осі X, R_{roll} – матриця обертання навколо осі Z. Загальна матриця множиться на вектор координат:

$$v_{global} = R * v_{local} \quad (3)$$

Таким чином отримується одиничний напрямляючий на об'єкт вектор на одиничній сфері навколо камери. Для того щоб тепер отримати позицію об'єкту відносно камери потрібно врахувати дані від дальноміру та висотоміру. Знаючи висоту камери Y_{cam} та вектор напрямку v_{global} можна зайти точку перетину даного вектору з площиною землі:

$$\begin{pmatrix} X_{real} \\ Y_{real} \\ Z_{real} \end{pmatrix} = \begin{pmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \end{pmatrix} + t * v_{global} \quad (4)$$

Оскільки $Y_{real} = 0$ (об'єкт розташований на землі), то $t = -\frac{Y_{cam}}{v_y}$.

$X_{cam} = 0$; $Y_{cam} = h$; $Z_{cam} = 0$, тобто top-view координати камери беремо за початок відліку. Недоліком такого методу є неможливість визначення координат у випадку знаходження розпізнаного об'єкту вище рівня землі ($Y_{real} > 0$).

Іншим підходом є використання відстані до об'єкту, оскільки це є значення від дальноміру (рис. 2):

$$\begin{pmatrix} X_{real} \\ Y_{real} \\ Z_{real} \end{pmatrix} = \begin{pmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \end{pmatrix} + d * v_{global} \quad (5)$$

Однак такий підхід також має значний недолік, який полягає в тому, що дальномір працює точково, тобто повертає відстань від камери до об'єкту який розташований в центрі зображення. У випадку, якщо ж об'єкт знаходиться, наприклад на краю зображення то його реальна відстань може бути суттєво іншою в порівнянні з даними від датчику. Враховуючи вищезазначене для обчислення реальних координати було обрано метод з врахуванням висоти.

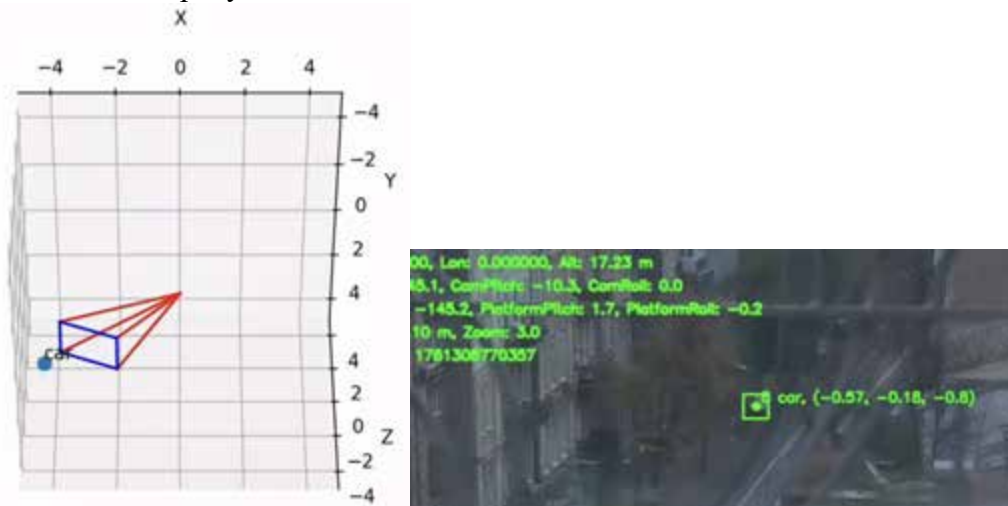


Рис. 2. Приклад роботи системи на дроні з побудовою 3D мапи

Трекінг об'єктів, тобто присвоєння ідентифікатору та відстеження руху між кадрами, дозволяє визначити швидкість та напрям руху об'єкту. Ефективним алгоритмом 2D трекінгу є ByteTrack [2], який дозволяє врахувати як висококонфідентні так і низькоконфідентні детекції. Однак, при використанні стандартного підходу використання піксельних координат, такий алгоритм не враховує динаміку камери. Тому було запропоновано використовувати для трекінгу координати об'єкту обраховані за допомогою вищепри описаного підходу з врахуванням зумування, обертання камери та даних від датчиків. Оскільки за моделлю, що розглядається завжди $Y_{real} = 0$, тому для трекінгу використано X_{real} та Z_{real} (розглядаємо об'єкти у форматі top – view, тобто дивимось згори).

В результаті було отримано систему, яка здатна розпізнавати та класифікувати об'єкти. Додатково за допомогою даних від датчиків вона здатна обраховувати координати відносно камери, а також здійснювати трекінг об'єктів незалежно від повороту камери та зумування, що дозволяє ефективно вести спостереження за полем бою в режимі реального часу. Систему протестовано на дроні DJI matrice з камерою DJI H20. В подальшому система може бути модифікована для роботи з рухомим дроном.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.
2. Zhang, Yifu & Sun, Peize & Jiang, Yi & Yu, Dongdong & Yuan, Zehuan & Luo, Ping & Liu, Wenyu & Wang, Xinggang. (2021). ByteTrack: Multi-Object Tracking by Associating Every Detection Box. URL: <https://doi.org/10.48550/arXiv.2110.06864>.

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

ХІІ 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