

Технічні науки

УДК 621.391.822:623.746.7

Хоменко Євген Валентинович

аспірант, начальник

Центр досліджень

Khomenko Yevhen

PhD Student, Head

Research Center

ORCID: 0009-0006-7006-3439

Краснов Роман Володимирович

кандидат технічних наук, доцент,

офіцер відділу проблем розвитку озброєння та військової техніки,

застосування безпілотних комплексів та систем

радіоелектронної боротьби

Центр досліджень

Krasnov Roman

PhD in Technical Sciences, Associate Professor,

Officer of the Department for the Development of Armaments and

Military Equipment, Use of Unmanned Systems and Electronic Warfare Systems

Research Center

ORCID: 0009-0005-8889-1868

Буданов Олег Валентинович

офіцер відділу проблем розвитку озброєння та військової техніки,

застосування безпілотних комплексів та систем

радіоелектронної боротьби

Центр досліджень

Budanov Oleg

*Officer of the Department for the Development of Armaments and
Military Equipment, Use of Unmanned Systems and Electronic Warfare Systems*

Research Center

ORCID: 0000-0002-1105-4245

Ільїн Сергій Олександрович

*інженер відділу проблем розвитку озброєння та військової техніки,
застосування безпілотних комплексів та систем*

радіоелектронної боротьби

Центр досліджень

Ilyin Serhiy

*Engineer of the Department for the Development of Armaments and
Military Equipment, Use of Unmanned Systems and Electronic Warfare Systems*

Research Center

ORCID: 0009-0009-1066-4891

Бондар Віктор Юрійович

*начальник відділу досліджень проблем у сфері технічного прикриття
єдиної транспортної системи України та критичної інфраструктури*

Центр досліджень

Bondar Viktor

*Head of the Department for Research on Technical Security Issues
of Ukraine's Unified Transport System and Critical Infrastructure*

Research Center

ORCID: 0009-0004-8286-4060

**АНАЛІЗ ІСНУЮЧИХ ПІДХОДІВ ТА МЕЖ ЗАСТОСУВАННЯ
СУЧАСНИХ АЛГОРИТМІВ КОМП'ЮТЕРНОГО ЗОРУ І
МАШИННОГО НАВЧАННЯ В ЗАДАЧАХ ДЕТЕКТУВАННЯ
БЕЗПІЛОТНИХ ЛІТАЛЬНИХ АПАРАТІВ
ANALYSIS OF EXISTING APPROACHES AND LIMITATIONS OF
MODERN COMPUTER VISION AND MACHINE LEARNING
ALGORITHMS IN THE DETECTION OF UNMANNED AERIAL
VEHICLES**

***Анотація.** Вступ. Досліджено сучасний стан розвитку та застосування безпілотних літальних апаратів (БПЛА) у цивільній та військовій сферах, що зумовило зростання потреби в ефективних засобах їх виявлення та протидії, особливо навколо об'єктів критичної інфраструктури. Встановлено, що одним із перспективних напрямів є візуальне виявлення БПЛА із застосуванням алгоритмів комп'ютерного зору та машинного навчання. Показано, що цей підхід має низку обмежень, пов'язаних із зовнішніми умовами (освітлення, погода), параметрами сенсорів (роздільна здатність, оптика), а також обмеженнями обчислювальних ресурсів і точністю алгоритмів. Водночас визначено ключову перевагу — пасивність візуального методу, що ускладнює його виявлення засобами радіоелектронної розвідки, та є критично важливим у військовому контексті.*

У статті здійснено огляд сучасних комерційних і військових систем протидії БПЛА, визначено їхні технічні обмеження та відповідність актуальним викликам. Особливу увагу приділено методам виявлення, що ґрунтуються виключно на обробці візуальних даних. Проаналізовано межі ефективності таких методів з огляду на сучасні алгоритми глибинного навчання та доступні апаратні засоби.

Ключові слова: безпілотний літальний апарат, комп’ютерний зір, машинне навчання, детектування об’єктів.

Summary. *Introduction. The current state of development and application of unmanned aerial vehicles (UAVs) in both civilian and military domains is examined, highlighting the growing need for effective means of detection and counteraction, particularly around critical infrastructure sites. It is established that one of the promising directions is visual UAV detection using computer vision and machine learning algorithms. This approach, however, has several limitations related to external conditions (lighting, weather), sensor parameters (resolution, optics), as well as computational resource constraints and algorithm accuracy. At the same time, a key advantage is identified—the passive nature of the visual method, which complicates its detection by electronic intelligence systems, making it critically important in a military context.*

The article reviews modern commercial and military counter-UAV systems, identifying their technical limitations and how well they address current challenges. Special attention is given to detection methods based solely on visual data processing. The limits of effectiveness of such methods are analysed considering current deep learning algorithms and available hardware.

Key words: *unmanned aerial vehicle, computer vision, machine learning, object detection, counteraction.*

Аналіз важливості візуального детектування для СП БПЛА

Сучасні СП БПЛА є складними програмно-апаратними комплексами, які поєднують обробку даних із різноманітних сенсорів для точної ідентифікації загроз та інтеграції з існуючими системами безпеки. У переважній більшості сучасні комплексні СП БПЛА сконцентровані на відстеженні та нейтралізації загроз, пов’язаних із використанням цивільних БПЛА I класу [1, с. 3] або їх незначних модифікацій. Військові системи

також призначені для захисту від безпілотників II та III класів. Основними засобами виявлення (ЗВ) БПЛА є активні радары, радіоелектронне детектування протоколів та засобів зв'язку, візуальне виявлення оптичними системами та акустичне сканування. Усі ці канали інформації поєднуються завдяки алгоритмам обробки даних, що дозволяє покращувати точність кожного з них окремо, нівелювати їх недоліки, для того щоб заздалегідь реагувати на потенційні загрози й мінімізувати ризики для об'єктів, що перебувають під захистом.

Усі ЗВ БПЛА можна умовно поділити за радіусом на дальні (сотні метрів і більше) і ближні (кілька сотень метрів і менше) (див. рис. 1).

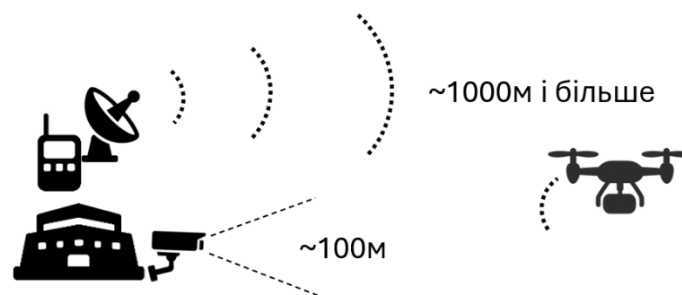


Рис. 1. Зони дії різних типів сенсорів

Джерело: розробка авторів

Критерієм потрапляння в категорію є здатність певного типу сенсора незалежно від інших каналів отримання інформації виявляти БПЛА в певному діапазоні відстаней [1, с. 9]. До дальніх ЗВ можемо віднести:

- радары (Ка-, Ку-, Х-, С-діапазонів), які виявляють цілі на значних відстанях, хоча можуть мати складнощі з низькою ефективною площею розсіювання, яку мають більшість БПЛА, і малою висотою польоту;
- радіочастотні сканери, що виявляють типові частоти сигналів керування і зворотного зв'язку чи телеметрії.

До ближніх ЗВ БПЛА можна віднести:

- оптичні та тепловізійні камери з можливістю розпізнання та класифікації БПЛА;
- акустичні сенсори.

Після детектування, локалізації, та визначення ступеня загрози від БПЛА, вживаються наявні засоби протидії. Серед них можна виокремити:

- Засоби радіоелектронної боротьби (РЕБ) – для перешкоджання або придушення каналів керування й навігації БПЛА [1, с. 16];
- Кінетичні та механічні методи – від фізичного знищення засобами вогневого ураження, до спеціалізованих БПЛА-перехоплювачів.
- Засоби високоенергетичної спрямованої дії (лазерна зброя, мікрохвильове випромінювання) [1, с. 23], які поки обмежені високою вартістю та непідтвердженими функціональними характеристиками. У більшості існують в якості дослідницьких прототипів.

Найпоширеніша конфігурація сучасних СП БПЛА – поєднання радарів і оптичних засобів (камер видимого та ІР діапазонів) зі спрямованими системами радіоелектронної боротьби (РЕБ) і рідше, у військових застосуваннях, кінетичних засобів ураження. Комбінування оптичних сенсорів з радарами обумовлене тим, що навіть у найкращих зразків активних радарів похибка в азимуті сягає близько 1° , а у відстані до виявленого об’єкта – кількох метрів, що навіть на відстанях у кілька сотень метрів значно погіршує точність наведення та ідентифікацію цілей. Це не є критичним для застосування засобів РЕБ, але використання кінетичних засобів без оптичного наведення майже неможливе.

Швидкий розвиток та здешевлення відповідних технологій привів до масштабного виготовлення і широкого використання БПЛА-камікадзе різного типу, які виконують функції розвідки та високоточних засобів ураження особового складу та техніки під час російсько-української війни [2]. Саме через масштаби використання БПЛА та колосальну різницю

вартості сучасних СП та об'єктів, які вони покликані захищати, виникає потреба в ефективних, легко масштабованих і економічно обґрунтованих рішеннях, які можна застосовувати на чисельних об'єктах вздовж усієї лінії бойового зіткнення.

Незважаючи на широкий набір ЗВ, існуючі СП БПЛА мають обмеження у боротьбі зі швидко еволюціонуючими технологічними рішеннями, орієнтованими на нейтралізацію або прихований обхід відомих засобів. Новітні БПЛА, що використовують нестандартні протоколи зв'язку, алгоритми автоматичного наведення, стають дедалі автономнішими і менш помітними для стандартних ЗВ. Наприклад, значні складнощі створює використання оптоволоконного кабелю для передачі сигналів керування БПЛА, що робить практично неможливим виявлення загрози радіочастотними сканерами (через повну відсутність випромінювання), а низька висота польоту при такому типі керування значно ускладнює виявлення за допомогою радарів. У результаті нівелюються зони виявлення дальнього радіусу дії. У найскладніших випадках СП БПЛА залишається покладатися лише на візуальні чи акустичні системи виявлення ближнього радіусу дії.

Мета роботи. У цій роботі автори оцінюють можливість створення повністю пасивного етапу детектування ближнього радіусу дії засобами візуальної ідентифікації з використанням простих оптичних систем та існуючих алгоритмів обробки візуальної інформації. Проведено порівняльний аналіз алгоритмів візуального детектування об'єктів і прикладів їх практичного застосування або тестування на реальних даних. Оскільки для задач захисту від БПЛА критичним є максимізація часу для використання наявних засобів протидії, аналізуються підходи, які дозволяють виявляти БПЛА на максимально можливій відстані згідно з оптичними характеристиками поширених сенсорів.

Виклад основного матеріалу. Підходи до оцінки якості детектування. У літературі по детектуванню об'єктів на зображеннях часто для оцінки успішності виявлення використовують класичний критерій Джонсона [3]. Він сформулював співвідношення між роздільною здатністю видимих лінійних пар відносно розміру об'єкту, та здатністю спостерігача виконувати завдання, такі як детектування, орієнтацію, розпізнавання та ідентифікацію з імовірністю 50%. На практиці це зручно, оскільки, знаючи оптичні параметри використовуваної системи, можна просто розраховувати відстань до виявленого об'єкта. Проте оригінальний критерій Джонсона має цілий ряд обмежень, якщо намагатись його застосовувати безпосередньо до результатів роботи сучасних алгоритмів КЗ та цифрових форматів представлення зображень. У спробах адаптувати критерій Джонсона під сучасні вимоги, запропоновано цілий ряд оновлених підходів до оцінки якості задач детектування об'єктів. В роботах, які стосуються перегляду критерія Джонсона, [4] оригінальні межі детектування, як правило, збільшуються, тобто показано, що для стійкості практичного використання результатів розпізнавання потрібно більше інформації.

У широкому спектрі задач розпізнавання складних за формою об'єктів, прив'язка оцінки результатів роботи алгоритмів візуального розпізнавання безпосередньо до розмірів об'єктів у пікселях на вхідному зображенні є складною і не завжди досяжною. На практиці значно частіше використовуються метрики, прив'язані до якості результатів розпізнавання [5]. Найпоширеніші з них:

- *Recall (R)* – повнота, фактично це відсоток виявлених алгоритмом об'єктів певного класу (або кількох класів);
- *Average Precision (AP)* – інтегральна характеристика точності та повноти розпізнавання, вона характеризує здатність алгоритму визначати клас об'єктів, розташування, та одночасно не пропускати об'єкти у вхідних зображеннях;

- *Mean Average Precision (mAP)* – середнє значення *AP* для кількох класів об'єктів;
- *Intersection over Union (IoU)* – міра точності детектування окремого об'єкта, характеризує наскільки алгоритм якісно визначив саме розташування об'єкта на вхідному зображенні.

Залежно від задач або меж застосування алгоритмів, для наочності, також часто виокремлюють метрики, специфічні для деяких підкласів, mAP_{small} – середня точність для розпізнавання об'єктів, визначених у рамках експерименту або тестового набору даних як маленькі

Ці та інші критерії мають свої межі та умови застосування, і їх інтерпретація може зміщуватись відповідно до специфіки задач, класу об'єктів розпізнавання, тощо. У контексті даної роботи вони нас цікавлять саме як найбільш поширені критерії для того, щоб мати змогу узагальнювати та інтерпретувати результати численних досліджень, фокусуючись на детектуванні дрібних об'єктів.

Архітектурні особливості в розпізнаванні дрібних об'єктів алгоритмами CNN

Останніми роками відбувся відчутний прогрес у сфері розпізнавання та класифікації об'єктів на цифрових зображеннях. У значній мірі це відбулось завдяки розвитку алгоритмів сімейства CNN, починаючи від фундаментальних архітектур, таких як LeNet та AlexNet [6], до сучасних складних багаторівневих нейронних мереж. Успіх CNN обумовлений не тільки численними інноваціями в архітектурі, а також вдосконаленням методів навчання, зростаючою доступністю великих обсягів даних та високопродуктивних обчислювальних платформ. Ці алгоритми не лише перевершують класичні методи КЗ в традиційних для останніх сферах застосування, але й ефективно застосовуються у задачах детектування об'єктів, семантичної сегментації та навіть генерації зображень.

Алгоритм розпізнавання на основі CNN обробляє зображення шляхом послідовного сканування для визначення ділянок, де ймовірно можуть бути присутні відомі об’єкти, та класифікації цих ділянок за типами [7]. У цьому процесі, серед інших, нас цікавлять два з чотирьох основних етапів (див. Рис. 2):

- *Згортка* – у CNN так називають спеціальну математичну операцію, що використовується для обробки зображень або інших типів даних. Параметри цих операцій визначаються автоматично під час процедури навчання;
- *Пулінг* – це операція зменшення розмірності мап ознак, отриманих після згортки, з метою узагальнення інформації та оптимізації обчислень.

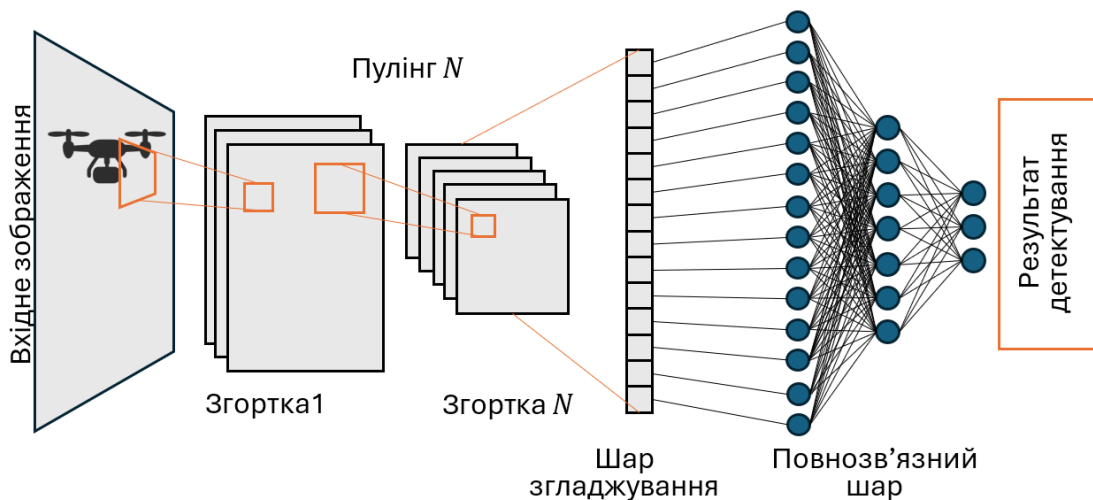


Рис. 2. Типова схема CNN мережі

Джерело: розробка авторів на основі [18]

Спочатку алгоритм CNN отримує змістовні візуальні ознаки із вхідного зображення, використовуючи згорткові шари. У результаті такої операції отримується карта ознак, у якій в концентрованому вигляді виділені певні характеристики вхідного зображення, наприклад, краї,

контури, текстури. Після цього CNN локалізує об'єкти в певних областях зображення, класифікує їх серед відомих категорій (транспортні засоби, люди, будівлі, тощо) та присвоює кожному детектуванню чисельну міру вірогідності.

Наведені особливості роботи CNN алгоритмів на жаль, є джерелом обмежень в задачах детектування дрібних об'єктів. Однією з головних проблем є те, що у процесі роботи CNN зображення послідовно зменшується, проходячи через шари мережі, саме через застосування згорток та пулінгу. Таке зменшення роздільної здатності призводить до втрати просторових деталей, що особливо критично для точного розпізнавання та локалізації малих об'єктів. Крім того, дрібні об'єкти містять меншу кількість пікселів, через що отримані ознаки часто є недостатньо чіткими й можуть бути легко переплутані з фоном чи шумами. Також у CNN алгоритмах іноді використовуються заздалегідь визначені опорні рамки, призначені для виділення можливих областей розташування об'єктів, або просто крок сканування може бути не оптимальним з тих чи інших причин для задачі детектування дрібних об'єктів, що значно погіршує міру точності *IoU* і відчутно знижує оцінку кінцевих результатів детектування.

Через згадані вище особливості архітектури алгоритмів на базі CNN, їх ефективність у детектуванні дрібних об'єктів зазвичай поступається результатам при виявленні відносно великих і виразних цілей. Зрозуміло, що така специфіка роботи і контекст задач вимагають адаптації загальних підходів. Виходячи з оглядових досліджень та порівняльного аналізу результатів тестів різних варіантів реалізації CNN-алгоритмів, наведених у працях [8; 9; 10; 11] можна виокремити кілька найперспективніших модифікацій базової архітектури, які оптимізують її для детектування дрібних об'єктів на зображеннях:

- *Мережа пірамід ознак (англ. Feature Pyramid Network)* – значно покращують здатність вилучення ознак на високих згорткових рівнях, що є ключовим для малих об'єктів, оскільки саме там відбуваються основні втрати інформації про цільовий об'єкт через зменшення роздільної здатності;
- *Механізми уваги* – модулі уваги допомагають CNN фокусуватись на регіонах, де імовірно знаходиться об'єкт детектування [12], підсилюючи інформативні ознаки (наприклад, дрібних об'єктів) та пригнічуючи зайву інформацію (наприклад, ділянки однорідного фону тощо);
- *Контекстне посилення ознак* - використання ширшої контекстної інформації суттєво покращує виявлення малих об'єктів завдяки інтеграції інформації про оточення. Так, для задач виявлення БПЛА цей підхід може покращити детектування на фоні хмарного неба;
- *Інтеграція Transformer-CNN* – це ще один механізм врахування контексту цілого вхідного зображення, Transformer-модулі додають до CNN можливість враховувати взаємозв'язки між далекими частинами зображення. Наприклад, Detection Transformer (DETR) використовує механізм самоуваги для одночасної обробки усієї сцени [19];
- *Розширені згортки (англ. Dilated Convolutions)* – це механізм штучного збільшення рецептивного поля згортки без збільшення кількості параметрів. Фактично, на вхід в операцію згортки подається зображення з пропусками окремих рядків та стовпчиків пікселів, що розширює область навколо об'єкта, яку аналізує мережа CNN. Це дозволяє зменшити ефект втрати просторової роздільної здатності і покращує точність виявлення малих об'єктів. Підбираючи коефіцієнт розширення, можна адаптувати детектування під потрібний розмір та масштаб вхідних даних;

- *Спеціалізовані функції втрат* – користувацькі функції втрат, орієнтовані на підвищення ефективності виявлення малих об'єктів. В таких задачах CNN-алгоритм здебільшого виконує передбачення для фону або інших неінформативних областей, що призводить до дисбалансу під час навчання (малі об'єкти статистично мають незначний вплив на загальні показники помилки). Через це мережа навчається поступово їх ігнорувати, ускладнюючи інтерпретацію результатів;
- *Стратегії розширення даних (англ. Data Augmentation)* – набір технік, які дозволяють штучно збільшити обсяг навчальних даних [13]. Якість та достатній об'єм даних, на якому навчається та або інша модель CNN, має критичне значення для отримання Бажаних результатів. На практиці побудова навчальних чи тестових наборів даних є дуже складною задачею, яка вимагає значних ресурсів. Тому стратегії розширення і набули своєї популярності, адже вони дозволяють підвищити якість навчання мережі та точність детектування, навіть за умови обмежених навчальних вибірок для об'єктів будь-яких розмірів;
- *Безанкерні методи (англ. Anchor-Free Methods)*– окремий підхід до розпізнавання об'єктів, який не використовує спеціалізовану вхідну розмітку, так звані анкери. І хоча безанкерні методи використовують аналогічні до інших CNN ановані дані для навчання, вони значно менше покладаються на розмітку для визначення меж об'єктів, що зменшує залежність результатів від якості розмітки;
- *Гіпер-виведення (англ. Slicing Aided Hyper Inference (SAHI))* – полягає в розбитті зображення на фрагменти з перекриттям [14], незалежному детектуванні об'єктів у кожному фрагменті та об'єднанні результатів. Метод SAHI може застосовуватися до різних моделей детектування на базі CNN без необхідності зміни їхньої архітектури або

перенавчання, оскільки він працює як зовнішній механізм і фактично адаптує зображення більшої роздільної здатності до оптимального розміру на якому проводилось навчання. Практичні тести показують, що SAHI покращує точність та повноту виявлення дрібних об'єктів. Значним недоліком методу є зростання обчислювального навантаження через паралельне виконання кількох розпізнавань.

Ці архітектурні та методологічні вдосконалення подекуди частково або повністю усувають специфічні обмеження CNN, які впливають на якість розпізнавання невеликих об'єктів. Багато з цих підходів можна використовувати одночасно модифікуючи базові варіанти реалізації різних алгоритмів. Це на практиці породжує десятки, якщо не сотні варіацій реалізації алгоритмів CNN для розпізнавання зображень. Це породжує окрему задачу визначення, яка з варіацій є найбільш вдалою в рамках поставлених задач, оскільки суто теоретично передбачити це неможливо.

Публічні набори даних для перевірки алгоритмів розпізнавання

Для навчання та оцінки результатів роботи CNN-алгоритмів доступні численні публічні набори даних, однак більшість із них орієнтовані на досить загальні задачі детектування об'єктів, де в пріоритеті надається кількості розпізнаваних класів, а не розмір у кадрі. У таких наборах як PASCAL VOC, VisDrone2019-DET чи MS COCO [9], хоч і присутні об'єкти різного масштабу, виявлення справді дрібних цілей не є основним завданням. Натомість спеціалізовані набори, як наприклад UA-DETRAC, DOTA-v2.0 та SODA-D [10], у значно більшій мірі фокусуються на малорозмірних об'єктах, у складних і реалістичних умовах, релевантних для цільового контексту.

На практиці існують десятки або навіть сотні різних вузько спеціалізованих наборів оптимізованих для розпізнавання тих чи інших класів об'єктів. Вони були вилучені з аналізу, оскільки орієнтовані на

розпізнавання об’єктів занадто специфічної геометричної форми (наприклад дорожні знаки), або на однорідному фоні (судна, або інші об’єкти які плавають на воді, люди на фоні одноманітних пейзажів, як-от пляж). Це зроблено свідомо, оскільки подібні фактори значно впливають на результати детектування. Крім того, так як нас значною мірою цікавить порівняння якості та межі розпізнавання різних алгоритмів між собою, логічним буде обмежитись популярними та широко використовуваними наборами даних у яких порівняння результатів є більш надійним.

На жаль, у публічних наборах тестових даних немає чіткого визначення того, що можна вважати “невеликим об’єктом”. Умовно їх розмір в пікселях на зображенні можна визначити відштовхуючись від одного з найбільших і найпопулярніших наборів MS COCO. В дослідженнях, які його використовують, до метрики результатів тестів AP_{small} , відносять усі об’єкти, які мають розміри на зображенні менші за 32×32 пікселі. Хоча це й важко назвати стандартом, але на практиці використовується як умовний орієнтир. Цієї межі переважно дотримуються і спеціалізовані набори даних.

Найбільший інтерес в рамках даної роботи звісно представляють набори даних, які безпосередньо містять фото та відео БПЛА: TIB-Net, Det-Fly, DUT Anti-UAV, ARD-MAV, UAVfly [15]. У зазначених публічних наборах можна знайти зображення лише комерційних моделей квадрокоптерів, що, безумовно, є найближчою апроксимацією до оцінюваних можливостей оптичного детектування БПЛА. До того ж, характер фону типового фону представленого в цих наборах, оточення, та форми самих об’єктів є найбільш релевантними для аналізу детектування.

Для порівняння, ключові характеристики згаданих вище тестових наборів представлено в таблиці 1.

Таблиця 1

Порівняння публічних наборів тестових даних

Характеристика	Розмір набору	Кількість категорій	Складність сцени	Відсоток малих об'єктів	Розміри малих об'єктів в пікселях
PASCAL-VOC	~2.6 млн	20	Середня	~10%	~200×200 – 50×50
VisDrone2019-DET	~0.5 млн	13	Висока	~20%	~80×80 – 10×10
MS COCO	~2.6 млн	80	Висока	~10%	~250×250 – 30×30
UA-DETRAC	~1.2 млн	5	Середня	~10%	~30×30 – 20×20
DOTA-v2.0,	~1.8 млн	18	Середня	~70%	~50×50 – 10×10
SODA-D	~280000	9	Висока	~70%	~32×32 – 16×16
TIB-Net	~3000	2	Висока	~80%	~45×45 – 10×10
Det-Fly	~13300	1	Висока	~50%	~60×60 – 20×20
DUT Anti-UAV	~10000	1	Висока	~90%	~30×30 – 10×10
ARD-MAV	~110000	1	Висока	~90%	~30×30 – 8×8
UAfly	~10000	1	Висока	~80%	~50×50 – 20×20

Джерело: розробка авторів на основі [34; 35; 36]

З огляду на відмінність підходів до тестування та розробки наборів даних, дуже складно визначити єдиний, однозначно інтерпретований критерій оцінки точності детектування. Через велику кількість варіацій реалізації кінцевої архітектури CNN-алгоритмів майже неможливо провести порівняльний тест в абсолютно однакових умовах. Крім того, далеко не всі версії доступні публічно для розгортання і виконання тестів незалежно. Тому так чи інакше доводиться покладатися на найбільш

поширені метрики AP_{small} , та групувати чисельні результати за розмірами, у тих випадках де дослідники їх явно диференціюють.

Найбільш вдалі алгоритми CNN для розпізнавання дрібних об'єктів

Ціла низка порівняльних досліджень містить набір результатів детектування дрібних об'єктів на зображеннях різними реалізаціями алгоритмів CNN [11; 16, с. 22; 17, с. 19; 15 с. 22]. На основі даних з оглядів було відібрано реалізації, які досягають найвищих показників AP_{small} . Ці моделі також демонструють високі результати для об'єктів середнього та великого розміру. Також численні тести показують, що моделі, які добре працюють на спеціалізованих даних для БПЛА, мають високі показники для інших класів об'єктів із загальних наборів даних і, відповідно, алгоритми розпізнавання значно більш чутливі до загальної архітектури і підбору даних для навчання, ніж до конкретного типу об'єктів. Тому до списку рекомендованих результатів можна включати алгоритми CNN, які мають підтверджені результати на малих об'єктах загального призначення. Це дозволяє мати ширший вибір для узагальнення особливостей CNN, які корисні в задачах розпізнавання дрібних об'єктів.

Отже, найвищої точності в різних дослідженнях на різних типах даних, де явно виокремлюються результати детектування дрібних об'єктів досягають наступні реалізації CNN: Faster R-CNN, Grid-RCNN [11, с. 23], STDnet-ST++ [18, с. 9], DETR [19], ESOD [20] та FreeAnchor [21]. Вони поділяють кілька вищезгаданих підходів до детектування, дрібних об'єктів, які підтверджені експериментально, як мінімум в кількох незалежних роботах:

- *Мережа пірамід ознак* – мабуть найпоширеніше і одне з найрезультативніших вдосконалень, яке значною мірою покращує врахування візуальних ознак дрібних об'єктів;

- *Механізми уваги* – у тій чи іншій формі більшість згаданих архітектур CNN реалізують механізми уваги і підсилюють окремі зони детектування;
- *Уточнення локалізації* – часто зниження загальної якості розпізнавання відбувається через низькі показники *IoU*, наявність різноманітних механізмів компенсації цієї проблеми для дрібних об'єктів також дає відчутний позитивний результат.

Також важливою спільною рисою усіх цих алгоритмів є те, що вони являють собою так звані двоступеневі детектори, через це досить вимогливі до обчислювальних потужностей. Це є суттєвим аргументом для практичного застосування у вбудованих або автономних системах. Як з точки зору споживання енергії, так і з огляду необхідності застосування відповідного обладнання.

Від цих алгоритмів зовсім незначно відстає найпопулярніший одноступеневий алгоритм сімейства YOLO (скорочення від You Only Look Once). Цей алгоритм надзвичайно популярний і на сьогодні існує аж одинадцять поколінь [22]. Останнє з яких з покоління в покоління інтегрувала варіації майже усіх згаданих вище вдосконалень CNN для детектування дрібних об'єктів, при цьому зберігши фокус на швидкість відпрацювання, загальний розмір даних навчених моделей, тощо. Тому, за умови виконання детектування автономними системами в режимі реального часу, цей алгоритм є однозначним лідером з практичної точки зору.

Згадані в цьому розділі реалізації CNN мають практичні результати які засвідчують їх здатність детектувати об'єкти розміром від 32 до 16 пікселів на зображеннях. Після цієї межі, якість розпізнавання різко зменшується, особливо відчутно падає повнота виявлення, тобто алгоритми перестають бачити значний відсоток об'єктів, що є неприпустимо для СП БПЛА. Розмір об'єкта в 16 пікселів на цифровому зображенні можна на

сьогодні оцінити як межу розпізнавання сучасних алгоритмів CNN при якому більшість об'єктів у кадрі будуть виявлені і локалізовані. З огляду на це, можна оцінити максимальну відстань виявлення об'єктів типового розміру для БПЛА квадрокоптерного типу (порядку 30 см по горизонталі). У таблиці 2 наведено результати розрахунку відстані для найпоширеніших стандартів відеоданих та камер, які мають типові кути огляду і які найчастіше використовуються у вбудованих обчислювальних платформах, робототехніці, та подібних галузях.

Таблиця 2

**Максимальна відстань детектування типового БПЛА
квадрокоптерного типу**

	Роздільна здатність	Кут 60°, відстань м	Кут 90°, відстань м .
VGA	640×480	~16	~10
HD	1280×720	~32	~20
FullHD	1920×1080	~47	~31
2K	2560×1440	~62	~41
4k	3870×2160	~93	~61
8k	7680×4320	~187	~122

Джерело: розробка авторів

Детектування і відстеження БПЛА в режимі обробки відео

В типових задачах СП БПЛА, корім первинного детектування потрібно здійснювати супровід цілі під час її руху для ідентифікації загрози та активації засобів протидії. Для цього широко використовуються спеціальні алгоритми відстеження. Якщо перейти від обробки зображень окремих кадрів до відео, то у потік вхідної інформації додається окремий часовий вимір і зростає обсяг даних для алгоритмів обробки. Це дає додаткові дані, які можуть впливати на кінцевий результат детектування.

В найбільш поширеному вигляді алгоритми стеження можуть базуватись на КЗ (статистичне диференціювання кадрів чи віднімання

фону) в комбінації з класичними фільтрами Калмана чи фільтра частинок [23, с. 3], для відсіювання хибних результатів. Для покращення результатів також широко використовуються різні алгоритми CNN [23]. Така комбінація свого часу стала досить поширеним підходом з наступними типовими кроками:

- виявлення руху використовується для звуження пошуку (привертання уваги);
- детектування об'єкта в цій області за допомогою CNN-алгоритму підтверджує результат;
- результат застосовується для стеження для зіставлення окремих детектувань між собою та відкидання хибних з цього ряду.

Найбільш відомими та вдалим реалізаціями цього підходу можна назвати алгоритми SORT, Deep SORT, Byte Track, [24]. Вони підвищують стабільність роботи алгоритмів відстеження, так як справжні цілі імовірно залишатимуться стійкими між кадрами, шум буде зазнавати більше випадкових змін, а використання CNN надає додаткові можливості для фільтрації. Проте, це не впливає значною мірою на здатність детектування менших об'єктів у кадрі і межа використання залишається близькою до обмежень використаного детектора, оскільки він незалежно виконує свою функцію на кожному окремому кадрі.

У процесі розвитку алгоритмів нейронних мереж, для обробки інформації з послідовною структурою, виник цілий підклас так званих Recurrent Neural Networks [24]. Ці мережі мають перевагу для обробки аудіо чи відео, так як не розглядають введення даних незалежно, а мають приховану пам'ять, яка враховує попередні стани вхідних даних. Для задач відстеження об'єкту в кадрі, природною є ідея інтегрувати часовий розподіл даних, який дає відео потік безпосередньо з алгоритмом розпізнавання, об'єднавши детектування і відстеження в одну мережу. Відомі спроби реалізувати цей підхід було зроблено в реалізаціях алгоритмів D&T, RCN,

FairMOT [24]. Усі вони поділяють загальну ідею: замість розподілу задач детектування та відслідковування як послідовних і незалежних, ці методи намагаються поєднати часові закономірності руху у вигляді траєкторій та просторові ознаки об’єктів виявлених на окремих зображеннях, таким чином, щоб вони підсилювали одні одних. Головний принцип полягає в тому, що часова зв’язаність (темпоральна узгодженість) певною мірою компенсує недостатні візуальні дані для детектування.

Слід зауважити, що паралельне виконання двох задач (незалежних чи зв’язаних), на перший погляд, підвищує вимоги до загальної обчислювальної потужності. Але для обох груп алгоритмів, описаних вище, детектування може оптимізуватись лише частинами зображень, де спостерігають рухомі об’єкти, тому однозначної відповіді дати не можна – вимоги можуть як збільшуватись, так і зменшуватись залежно від комбінації реалізації стеження/детектування. Це параметр на практиці завжди слід враховувати, оскільки неможливість обробляти кожен кадр в режимі реального часу може значно впливати на отримані результати.

Підсумовуючи вплив алгоритмів відстеження на загальну якість або ефективність детектування малих об’єктів, можна однозначно стверджувати, що поєднання обох задач має позитивний вплив на повноту виявлення R . Тобто додавання розподілу кадрів в часі з відео потоку, дозволяє виявляти більший відсоток малих об’єктів, ніж з використанням окремих кадрів. Але їх абсолютний розмір у пікселях, а отже, критичний для нас параметр, максимальна відстань детектування – залишається майже незмінним.

Обговорення і висновки. Проведений аналіз демонструє, що, незважаючи на значний прогрес у розвитку алгоритмів глибокого навчання та комп’ютерного зору, детектування малих об’єктів, зокрема БПЛА, за допомогою оптичних систем залишається складним завданням із низкою об’єктивних обмежень. Сучасні CNN-архітектури, включаючи

найуспішніші реалізації, починають суттєво втрачати точність розпізнавання на об’єктах, розмір яких у кадрі менше 16×16 пікселів. Це обмежує максимальну відстань ефективного виявлення БПЛА навіть за умови використання камер з високою роздільною здатністю та вузьким кутом огляду.

Разом з тим, певні архітектурні рішення (мережі пірамід ознак, механізми уваги, SANI, тощо), дозволяють частково компенсувати втрати інформативності, підвищуючи якість виявлення в складних умовах. Об’єднання алгоритмів детектування та стеження також дозволяє покращити повноту виявлення, особливо в умовах обробки відео даних.

У контексті побудови ефективних СП БПЛА ближнього радіусу дії, навіть обмежені можливості сучасних CNN-алгоритмів можуть бути використані в економічно ефективних рішеннях, особливо в ситуаціях, коли інші канали виявлення (радар, радіочастотний моніторинг) є неефективними.

Перспективні напрямки подальших досліджень включають:

- створення спеціалізованих навчальних вибірок із фокусом на реальні умови детектування БПЛА (фон, освітлення, масштаб);
- оптимізацію обчислювальних моделей для вбудованих платформ з обмеженими ресурсами;
- тестування в умовах, наближених до бойових або польових, із фіксацією реальних метрик якості виявлення для різних конфігурацій систем.

Таким чином, попри об’єктивні технологічні межі, сучасні алгоритми CNN залишаються придатними для використання в оптичному детектуванні БПЛА на обмежених відстанях, за умови правильного підбору архітектури, сенсорів та контексту застосування.

Література

1. Castrillo V. et al., A Review of Counter-UAS Technologies for Cooperative Defensive Teams of Drones. *Drones*, 6(3), 65. doi: 10.3390/drones6030065
2. Kunertova D., (2023). The war in Ukraine shows the game-changing effect of drones depends on the game *Bulletin of the Atomic Scientists*, 79, 95 - 102.
3. Sjaardema T., Smith C.S., & Birch, G.C. (2015). History and Evolution of the Johnson Criteria., doi: 10.2172/1222446.
4. DesAutels G., (2021). A Modern Review of the Johnson Image Resolution Criterion. *Optik*. doi: 10.1016/j.ijleo.2021.168246.
5. Padilla R., Netto S.L., & da Silva E.A., (2020). A Survey on Performance Metrics for Object-Detection Algorithms. 2020 International Conference on Systems, Signals and Image Processing (IWSSIP), 237-242.
6. Krizhevsky A., Sutskever I., Hinton GE., (2017). ImageNet classification with deep convolutional neural networks. *Commun ACM* 60(6):84-90. doi: 10.1145/3065386
7. Krichen, Moez. (2023). Convolutional Neural Networks: A Survey. *Computers*, 12(8), 151; doi: 10.3390/computers12080151.
8. Chen G. et al., A Survey of the Four Pillars for Small Object Detection: Multiscale Representation, Contextual Information, Super-Resolution, and Region Proposal, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 52, no. 2, pp. 936-953, Feb. 2022, doi: 10.1109/TSMC.2020.3005231
9. Kaur R., Singh S., A comprehensive review of object detection with deep learning. *Digit Signal Process*. 2022; doi: 10.1016/j.dsp.2022.103812

10. Xu C. et al., (2024). Oriented Tiny Object Detection: A Dataset, Benchmark, and Dynamic Unbiased Learning. ArXiv, abs/2412.11582.
11. Feng Q., Xu X., Wang Z., Deep learning-based small object detection: A survey. *Math Biosci Eng.* 2023;20(4):6551–6590. doi: 10.3934/mbe.2023282
12. Hu J., Shen L. and Sun G., "Squeeze-and-Excitation Networks," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018, pp. 7132-7141, doi: 10.1109/CVPR.2018.00745., doi: 10.1109/CVPR.2018.00745
13. Shorten C., Khoshgoftaar T., A survey on image data augmentation for deep learning. *J Big Data.* 2019;6(1):60. doi: 10.1186/s40537-019-0197-0
14. Zhang H. et al., Adaptive Slicing-Aided Hyper Inference for Small Object Detection in High-Resolution Remote Sensing Images. *Remote Sensing.* 2023; 15(5):1249. doi: 10.3390/rs15051249
15. Yasmeeen A.; Daescu O., Recent Research Progress on Ground-to-Air Vision-Based Anti-UAV Detection and Tracking Methodologies: A Review. *Drones* 2025, 9, 58. doi: 10.3390/drones9010058
16. Yu X. et al., Scale Match for Tiny Person Detection, 2020 IEEE Winter Conference on Applications of Computer Vision, pp. 1246-1254, doi: 10.1109/WACV45572.2020.9093394.
17. Wang X. et al., Small Object Detection Based on Deep Learning for Remote Sensing: A Comprehensive Review. *Remote Sensing.* 2023; 15(13):3265. doi: 10.3390/rs15133265
18. Bosquet B., Mucientes M, Brea V. M., STDnet-ST: Spatio-temporal ConvNet for small object detection, 2021,107929, <https://doi.org/10.1016/j.patcog.2021.107929>
19. Zhu X., et al. "Deformable detr: Deformable transformers for end-to-end object detection." arXiv preprint arXiv:2010.04159 (2020)

20. Liu K., et al., (2024). ESOD: Efficient Small Object Detection on High-Resolution Images. IEEE Transactions on Image Processing.
21. Zhang Q., Zhang H., Lu X. and Han X., Anchor-Free Small Object Detection Algorithm Based on Multi-scale Feature Fusion, 2022 5th International Conference on Pattern Recognition and Artificial Intelligence, 2022, pp. 370-374, doi: 10.1109/PRAI55851.2022.9904251.
22. Jegham et al., (2025). YOLO Evolution: A Comprehensive Benchmark and Architectural Review of YOLOv12, YOLO11, and Their Previous Versions. doi: 10.13140/RG.2.2.15952.83201.
23. Mirzaei B et al., Small Object Detection and Tracking: A Comprehensive Review. Sensors (Basel). 2023 Aug 3;23(15):6887. doi: 10.3390/s23156887.
24. Sajid, H. (2025, January 4). Top 10 video object tracking algorithms in 2025. URL: <https://encord.com/blog/video-object-tracking-algorithms/> (дата звернення: 03.04.2025).