

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДЕРЖАВНИЙ УНІВЕРСИТЕТ «КИЇВСЬКИЙ АВІАЦІЙНИЙ ІНСТИТУТ»
ФАКУЛЬТЕТ АЕРОНАВІГАЦІЇ, ЕЛЕКТРОНІКИ ТА ТЕЛЕКОМУНІКАЦІЙ
КАФЕДРА ТЕЛЕКОМУНІКАЦІЙНИХ ТА РАДІОЕЛЕКТРОННИХ СИСТЕМ

ДОПУСТИТИ ДО ЗАХИСТУ
Завідувач кафедри

Віктор ГНАТЮК
“ ” 2025 р.

КВАЛІФІКАЦІЙНА РОБОТА
(ПОЯСНЮВАЛЬНА ЗАПИСКА)

ВИПУСКНИКА ОСВІТНЬОГО СТУПЕНЯ МАГІСТР

Тема: «Підвищення ефективності виявлення малих об'єктів БПЛА»

Виконавець: _____ **Іван ЛІЩУК**
(підпис)

Керівник: _____ **Володимир КЛИМЧУК**
(підпис)

Консультанти з окремих розділів пояснювальної записки:

Консультант розділу «Охорона праці» _____ **Катерина КАЖАН**
(підпис)

Консультант розділу «Охорона навколишнього середовища»
_____ **Лариса ЧЕРНЯК**
(підпис)

Нормоконтролер: _____ **Богдан ЧУМАЧЕНКО**
(підпис)

Київ 2025

ДЕРЖАВНИЙ УНІВЕРСИТЕТ «КИЇВСЬКИЙ АВІАЦІЙНИЙ ІНСТИТУТ»

Факультет аеронавігації, електроніки та телекомунікацій

Кафедра телекомунікаційних та радіоелектронних систем

Спеціальність 172 «Електронні комунікації та радіотехніка»

Освітньо-професійна програма «Телекомунікаційні системи та мережі»

ЗАТВЕРДЖУЮ

Завідувач кафедри

Віктор ГНАТЮК

“ ” 2025 р.

ЗАВДАННЯ

на виконання кваліфікаційної роботи

Ліщука Івана Костянтиновича

(прізвище, ім'я, по батькові випускника в родовому відмінку)

1. Тема кваліфікаційної роботи: «Підвищення ефективності виявлення малих об'єктів БПЛА»

затверджена наказом ректора від «02» вересня 2025 р. № 1672 /ст

2. Термін виконання роботи: з 29.09.2025 р. по 31.12.2025 р.

3. Вихідні дані до роботи: Зображення та відеопослідовності, отримані з безпілотних літальних апаратів, що містять малі об'єкти різних класів; відкриті набори даних для задачі виявлення об'єктів на аерознімках (зокрема набори даних для UAV-зображень);

4. Зміст пояснювальної записки: Особливості задачі виявлення малих об'єктів на зображеннях БПЛА; Аналіз можливості використання ШІ в підвищенні ефективності виявлення; Методологія вдосконалення моделі UAV-DETR для виявлення малих об'єктів БПЛА; Практична реалізація та випробування вдосконаленої системи виявлення малих об'єктів БПЛА

5. Перелік обов'язкового графічного (ілюстративного) матеріалу: презентація PowerPoint

6. Календарний план-графік

№ пор.	Завдання	Термін виконання	Відмітка про виконання
1	Розробити деталізований зміст розділів кваліфікаційної роботи	29.09.2025-30.09.2025	Виконано
2	Вступ	01.10.2025-03.10.2025	Виконано
3	Особливості задачі виявлення малих об'єктів на зображеннях БПЛА	04.10.2025-14.10.2025	Виконано
4	Аналіз можливості використання ШІ в підвищенні ефективності виявлення	15.10.2025-26.10.2025	Виконано
5	Методологія вдосконалення моделі UAV-DETR для виявлення малих об'єктів БПЛА	27.10.2025-01.11.2025	Виконано
6	Практична реалізація та випробування вдосконаленої системи виявлення малих об'єктів БПЛА	01.11.2025-16.11.2025	Виконано
7	Охорона праці	17.11.2025-30.11.2025	Виконано
8	Охорона навколишнього середовища	01.12.2025-14.12.2025	Виконано
9	Усунення недоліків та захист кваліфікаційної роботи	15.12.2025-31.12.2025	Виконано

7. Консультанти з окремих розділів

Розділ	Консультант (посада, П.І.Б.)	Дата, підпис	
		Завдання видав	Завдання прийняв
Охорона праці	к.т.н., доц. Катерина КАЖАН		
Охорона навколишнього середовища	д.т.н., доц. Лариса ЧЕРНЯК		

8. Дата видачі завдання: «01» вересня 2025 р.

Керівник кваліфікаційної роботи _____
(підпис керівника)

Володимир КЛИМЧУК
(П.І.Б.)

Завдання прийняв до виконання _____
(підпис випускника)

Іван ЛІЩУК
(П.І.Б.)

РЕФЕРАТ

Пояснювальна записка до кваліфікаційної роботи «Підвищення ефективності виявлення малих об'єктів БПЛА» містить 96 сторінок, 15 рисунків, 9 таблиць, 22 використаних джерела.

UAV-DETR, CASNET, SOEP, CSAM, ОБРОБКА ЗОБРАЖЕНЬ, ГЛИБИННЕ НАВЧАННЯ, БПЛА, ДЕТЕКЦІЯ МАЛИХ ОБ'ЄКТІВ, НЕЙРОННИЙ ПРОЦЕСОР, ОПТИМІЗАЦІЯ МОДЕЛІ.

Об'єкт дослідження – процес виявлення та розпізнавання малих об'єктів на зображеннях, отриманих із безпілотних літальних апаратів (БПЛА).

Предмет дослідження – методи підвищення точності, швидкодії та апаратної ефективності систем детекції об'єктів на основі глибинних нейронних мереж і трансформерних архітектур, зокрема моделі RT-DETR та її вдосконаленої версії UAV-DETR++.

Мета роботи – підвищення ефективності виявлення малих і слабо помітних об'єктів на аерофотознімках шляхом модернізації архітектури RT-DETR із використанням модулів CasNet, SOEP та CSAM, а також оптимізації моделі для розгортання на вбудованих нейронних процесорах і прискорювачах ШІ.

Методи дослідження: теоретичний аналіз архітектур детекції об'єктів; моделювання та експериментальне тестування на наборах даних VisDrone2019 і DOTA; порівняльний аналіз продуктивності; використання методів глибинного навчання, квантизації, pruning і distillation для оптимізації моделі.

Результати роботи: розроблено вдосконалену модель UAV-DETR++, яка забезпечує підвищення точності (mAP@0.5) на 3–3.5% у порівнянні з базовим RT-DETR при збереженні роботи в реальному часі (≈ 30 FPS). Запропонована архітектура характеризується зменшеним числом параметрів ($\approx 16,8$ млн) і підтримує виконання на апаратних прискорювачах типу NVIDIA Jetson Orin NX, Intel Movidius або Google Coral TPU.

Практичне значення: результати дослідження можуть бути використані для створення бортових і наземних систем аналітики відеопотоку з БПЛА, автоматизованих систем моніторингу, пошуково-рятувальних комплексів і систем ситуаційної обізнаності, у тому числі в оборонному секторі.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ	9
ВСТУП.....	10
РОЗДІЛ 1. ОСОБЛИВОСТІ ЗАДАЧІ ВИЯВЛЕННЯ МАЛИХ ОБ’ЄКТІВ НА ЗОБРАЖЕННЯХ БПЛА	12
1.1. Вступ	12
1.2. Характерні риси малих об’єктів у зображеннях з БПЛА.....	12
1.3. Основні проблеми виявлення малих об’єктів	13
1.4. Специфіка даних з безпілотних літальних апаратів	14
1.5. Напрями підвищення ефективності виявлення малих об’єктів	15
1.6. Ключові фактори оцінювання якості детекторів	16
РОЗДІЛ 2. АНАЛІЗ МОЖЛИВОСТІ ВИКОРИСТАННЯ ШІ В ПІДВИЩЕННІ ЕФЕКТИВНОСТІ ВИЯВЛЕННЯ.....	18
2.1. Виявлення об’єктів на основі CNN.....	18
2.2. Трансформерні підходи до виявлення об’єктів.....	20
2.3. Методи виявлення, орієнтовані на зображення з БПЛА	24
2.4. Модель RT-DETR	27
РОЗДІЛ 3. МЕТОДОЛОГІЯ ВДОСКОНАЛЕННЯ МОДЕЛІ UAV-DETR ДЛЯ ВИЯВЛЕННЯ МАЛИХ ОБ’ЄКТІВ БПЛА.....	32
3.1. Вступ	32
3.2. Загальна структура вдосконаленої моделі UAV-DETR	33
3.3. Технічний розбір модифікованих модулів вдосконаленої моделі UAV-DETR	38
РОЗДІЛ 4. ПРАКТИЧНА РЕАЛІЗАЦІЯ ТА ВИПРОБУВАННЯ ВДОСКОНАЛЕНОЇ СИСТЕМИ ВИЯВЛЕННЯ МАЛИХ ОБ’ЄКТІВ БПЛА	56
4.1. Вступ	56
4.2. Можливості практичної реалізації	57
4.3. Методика випробувань	60

4.4. Результати випробувань.....	60
РОЗДІЛ 5. ОХОРОНА ПРАЦІ	67
5.1. Аналіз умов праці на робочому місці	67
5.2. Перелік шкідливих та небезпечних виробничих чинників	68
5.3. Аналіз шкідливих та небезпечних виробничих чинників	70
5.4. Розробка заходів з охорони праці	76
5.5. Пожежна безпека.....	79
5.6. Розрахункова частина. Розрахунок допустимого часу робіт при електромагнітному випромінюванні	81
РОЗДІЛ 6. ОХОРОНА НАВКОЛИШНЬОГО СЕРЕДОВИЩА.....	83
6.1. Вступ	83
6.2. Основні принципи охорони навколишнього середовища.....	83
6.3. Методи захисту при травленні друкованих плат та пайці	84
6.4. Електронні компоненти їх утилізація та вторинна переробка	86
ВИСНОВКИ	911
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	944

ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАЧЕНЬ

AIFI	- Attention-Enhanced Intra-scale Feature Interaction
BiFPN	- Bidirectional Feature Pyramid Network
CAS	- Channel-Aware Sensing
CCFF	- Cross-Scale Feature Fusion
CFPT	- Cross-Layer Feature Pyramid Transformer
CNN	- Convolutional neural network
CPU	- central processing unit
CSAM	- Context-Spatial Alignment Module
DETR	- Detection Transformer
DOTA	- Dataset for Object Detection in Aerial Images
FPN	- Feature Pyramid Network
GPU	- graphics processing unit
HRFPN	- High-Resolution Feature Pyramid Network
IoU	- aware query initialization
LGI	- Local-Global Interaction
mAP	- mean Average Precision
NMS	- Non-Maximum Suppression
NPU	- Neural Processing Units
PANet	- Path Aggregation Network
R-CNN	- Region-based Convolutional Neural Network
SOEP	- Scale-Optimized Enhancement Pyramid
SSD	- Single Shot MultiBox Detector
UAV	- uncrewed aerial vehicle
YOLO	- You Only Look Once
БПЛА	- безпілотний літальний апарат

ВСТУП

Сучасний розвиток безпілотних літальних апаратів (БПЛА) зумовив їх широке застосування у цивільній та військовій сферах — від аерофотозйомки та моніторингу довкілля до розвідки, пошуку та спостереження за об'єктами. Разом із розширенням функціональних можливостей дронів актуалізується проблема своєчасного та точного виявлення малих об'єктів, які часто мають низьку контрастність, малий розмір у кадрі та перебувають у складних фонових умовах.

Традиційні алгоритми комп'ютерного зору, зокрема методи на основі згорткових нейронних мереж (CNN), демонструють високу продуктивність при розпізнаванні великих об'єктів, однак втрачають ефективність для дрібних цілей через обмежену здатність до збереження просторових деталей і контекстуальної інформації. Ця проблема особливо критична у випадку БПЛА, де значна частина кадру займає фон, а цільові об'єкти можуть бути розміром у декілька пікселів.

У відповідь на ці виклики останніми роками активно розвиваються підходи, що базуються на архітектурі Transformer, зокрема моделі типу DETR (Detection Transformer) та їх модифікації — RT-DETR, DV-DETR, Drone-DETR, UAV-DETR тощо. Ці системи поєднують можливості глобальної уваги з ефективним багаторівневим вилученням ознак, що дозволяє значно підвищити точність виявлення дрібних об'єктів навіть у реальному часі.

Особливої уваги заслуговує модель UAV-DETR, яка розвиває архітектуру RT-DETR шляхом введення нових модулів — Channel-Aware Sensing (CAS), Scale-Optimized Enhancement Pyramid (SOEP) та Context-Spatial Alignment Module (CSAM). Ці удосконалення забезпечують покращене сприйняття просторових і контекстуальних зв'язків, а також ефективніше злиття ознак на різних масштабах, що підвищує точність виявлення малих цілей при зменшених обчислювальних витратах.

Таким чином, кваліфікаційна робота присвячена аналізу сучасних методів і архітектур для підвищення ефективності виявлення малих об'єктів БПЛА, зокрема

моделей на основі трансформерів, які поєднують високу точність, компактність моделі та можливість застосування у реальному часі.

Мета роботи — підвищення ефективності виявлення малих об'єктів БПЛА шляхом аналізу сучасних нейромережових архітектур і визначення підходів, що забезпечують оптимальне співвідношення між точністю та швидкістю.

Завдання кваліфікаційної роботи:

- проаналізувати сучасні методи виявлення об'єктів на зображеннях з БПЛА;
- розглянути архітектури CNN- та Transformer-типу, що застосовуються для виявлення малих цілей;
- порівняти ефективність класичних моделей (YOLO, Faster-R-CNN, SSD) і новітніх модифікацій DETR;
- визначити основні напрямки вдосконалення алгоритмів виявлення малих об'єктів, придатних для реалізації на обмежених обчислювальних ресурсах.

Об'єкт дослідження — процес виявлення об'єктів у відео- та фотоматеріалах, отриманих з безпілотних літальних апаратів.

Предмет дослідження — методи підвищення ефективності виявлення малих об'єктів за допомогою сучасних нейромережових архітектур.

Методи дослідження — аналіз наукових джерел, порівняльний аналіз архітектур глибокого навчання, математичне моделювання та узагальнення результатів експериментів.

РОЗДІЛ 1

ОСОБЛИВОСТІ ЗАДАЧІ ВИЯВЛЕННЯ МАЛИХ ОБ'ЄКТІВ НА ЗОБРАЖЕННЯХ БПЛА

1.1. Вступ

Останнє десятиліття характеризується активним розвитком технологій БПЛА, які широко застосовуються у військових, цивільних і промислових сферах. Завдяки мобільності, маневровості та здатності вести спостереження з різних висот, БПЛА використовуються для моніторингу територій, розвідки, оцінки пошкоджень, контролю інфраструктури та пошуково-рятувальних операцій. У більшості цих завдань центральним етапом є виявлення об'єктів на аерофотознімках, що забезпечує подальшу класифікацію, відстеження чи розпізнавання.

Водночас специфіка зображень, отриманих із бортових камер дронів, створює низку унікальних проблем для алгоритмів комп'ютерного зору. На відміну від наземних або стаціонарних систем спостереження, БПЛА працюють у змінних умовах висоти, кута огляду та освітлення, що призводить до варіацій масштабу, низького контрасту, шумів і динамічних фонів. Особливо складною є задача виявлення малих об'єктів, які часто займають лише незначну частину кадру (менше 1 % площі зображення).

1.2. Характерні риси малих об'єктів у зображеннях з БПЛА

Малі об'єкти у повітряних зображеннях — це цілі (транспортні засоби, люди, дрони, техніка), розмір яких у кадрі не перевищує 32×32 пікселів. Для прикладу, при зйомці з висоти понад 100 метрів транспортний засіб може бути представлений лише кількома десятками пікселів, а людина — менше ніж десятьма. Це ускладнює коректне виділення об'єкта з фону навіть для сучасних нейронних мереж.

Типові особливості таких об'єктів:

- низька роздільна здатність: обмежена кількість пікселів призводить до втрати текстурних ознак, що робить об'єкт малорозрізняваним;
- висока варіативність масштабу: об'єкти одного типу можуть мати різний розмір залежно від висоти польоту та кута огляду;
- складне фонове оточення: у міських сценах присутні дороги, будівлі, зелені насадження, тіні, що створюють фоновий шум;
- перекриття об'єктів: у щільних міських районах або транспортних потоках часто виникає часткове перекриття, що ускладнює локалізацію;
- рухомість сцени: зміни положення камери, вібрації та рухи об'єктів призводять до розмиття та нестабільності послідовних кадрів.

Ці фактори знижують ефективність стандартних алгоритмів, орієнтованих на більші об'єкти, та вимагають спеціальних архітектур, здатних зберігати дрібномасштабні ознаки та враховувати контекст сцени.

1.3. Основні проблеми виявлення малих об'єктів

Втрата просторової інформації при згортках і пулінгу: у класичних згорткових нейронних мережах (CNN) послідовні операції зведення розміру зображення (pooling, stride convolution) призводять до втрати деталей, що критично для малих цілей. Під час проходження через глибокі шари мережі малі об'єкти можуть повністю «зникнути» з карти ознак.

Незбалансованість масштабу ознак: більшість архітектур CNN (ResNet, VGG, MobileNet) оптимізовані для великих об'єктів, тоді як малі цілі формують слабкі градієнти та мають незначний вплив на функцію втрат. Це призводить до дисбалансу під час навчання і зміщення моделі у бік більших об'єктів.

Контекстуальна неоднозначність: малі об'єкти часто не містять достатньої внутрішньої інформації для розпізнавання класу (наприклад, автомобіль може бути представлений лише прямокутною плямою). У таких випадках важливу роль відіграє контекст оточення — дорога, тротуар, інші автомобілі тощо. Відсутність ефективного врахування контексту призводить до помилкових спрацьовувань.

Складність анотування та обмеженість даних: анотування малих об'єктів потребує точних рамок, а при зйомці з великих висот — високої роздільної здатності. Через це навіть великі набори даних, такі як VisDrone2019 або UAVDT, містять помилки у розмітці, що ускладнює навчання моделей [1].

Обмеження апаратних ресурсів: у випадку застосування моделей безпосередньо на борту БПЛА (on-board inference) важливою є обчислювальна ефективність. Потужні CNN або Transformer-архітектури мають мільйони параметрів і потребують графічних процесорів, що не завжди можливо в польових умовах.

1.4. Специфіка даних з безпілотних літальних апаратів

Аерофотознімки з БПЛА мають ряд характеристик, що відрізняють їх від наземних відеопотоків:

- великий діапазон висот: від кількох до сотень метрів, що зумовлює різку зміну масштабу об'єктів;
- різноманіття кутів огляду: вертикальні, косі, фронтальні, що вимагає інваріантності до перспективних спотворень;
- нестабільність освітлення: сонячні відблиски, тіні, контраст між освітленими й затемненими ділянками;
- різноманіття сцен: від відкритих просторів (поля, ліси) до щільної міської забудови.

На рис. 1.1 [1] показано приклади типових сцен із набору даних VisDrone2019, де об'єкти (автомобілі, люди, велосипеди) займають мізерну частину кадру й часто зливаються з фоном.

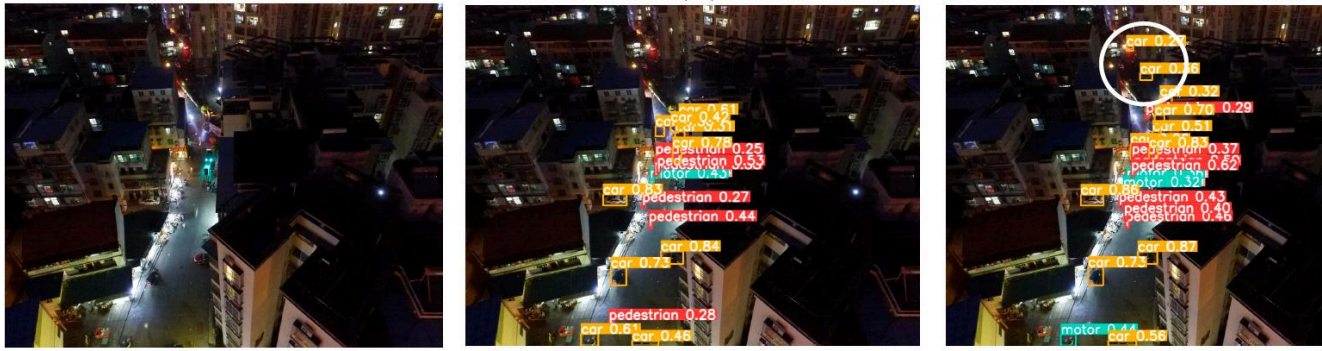


Рис. 1.1. Приклади малих об'єктів на зображеннях із набору даних VisDrone2019

Для адекватного навчання моделі необхідно мати репрезентативні дані, які охоплюють різні типи місцевості, кути огляду, погодні умови та класи об'єктів. Найпопулярніші відкриті набори даних:

- VisDrone2019 — містить понад 10 000 зображень міських і сільських сцен, 10 класів об'єктів (автомобілі, пішоходи, автобуси тощо);
- UAVDT — понад 80 000 кадрів відео з різних висот і освітлень;
- DOTA — великий набір аерозображень із геоприв'язаними об'єктами (будівлі, літаки, кораблі).

1.5. Напрями підвищення ефективності виявлення малих об'єктів

Розв'язання задачі ефективного виявлення малих об'єктів можливе за рахунок кількох напрямів удосконалення нейронних архітектур:

- багаторівневе злиття ознак (Feature Fusion). Моделі типу FPN (Feature Pyramid Network) і PANet дозволяють поєднувати низькорівневі деталі з високорівневими семантичними ознаками, зберігаючи контур і форму малих об'єктів;
- підвищення роздільної здатності шарів. Використання додаткових рівнів детекції (наприклад, P2 у SOEP-модулі в UAV-DETR) дозволяє краще працювати з об'єктами розміром 8–16 пікселів;

- контекстне моделювання. Використання механізмів уваги (attention mechanisms) забезпечує адаптивне виділення важливих регіонів, підвищуючи точність при складних фонах.
- мультиспектральні та часові ознаки. Поєднання RGB, ІЧ або відеопослідовностей дозволяє моделі використовувати рух або теплові відмінності як додаткові джерела інформації;
- оптимізація архітектури для реального часу. Нові моделі, такі як RT-DETR і UAV-DETR, орієнтовані на баланс між точністю та швидкістю, забезпечуючи роботу на борту дрона або в режимі потоку відео.

1.6. Ключові фактори оцінювання якості детекторів

Для оцінювання якості систем виявлення об'єктів використовують такі показники:

- Precision (точність) — частка правильно виявлених об'єктів серед усіх передбачених;
- Recall (повнота) — частка виявлених об'єктів серед усіх фактичних;
- mAP (mean Average Precision) — середня точність для всіх класів об'єктів при певному порозі IoU;
- GFLOPs — кількість операцій з плаваючою комою, що характеризує обчислювальні витрати моделі;
- Params (кількість параметрів) — визначає обсяг пам'яті та ресурси, необхідні для роботи.

Для задач БПЛА важливо досягнути компромісу між точністю (mAP) і швидкістю (FPS), оскільки обробка відеопотоку має відбуватися в реальному часі, а доступні обчислювальні ресурси — обмежені.

ВИСНОВКИ ДО РОЗДІЛУ 1

Задача виявлення малих об'єктів на зображеннях з БПЛА є однією з найскладніших у галузі комп'ютерного зору. Її ефективне вирішення вимагає комплексного підходу, який поєднує збереження дрібномасштабних ознак, адаптивне контекстне моделювання та оптимізацію обчислювальної ефективності.

Сучасні тенденції свідчать про перехід від класичних CNN-моделей до гібридних або повністю Transformer-базованих архітектур, таких як RT-DETR, Drone-DETR та UAV-DETR, які здатні забезпечувати високу точність навіть при обмежених ресурсах і складних фонових умовах.

РОЗДІЛ 2

АНАЛІЗ МОЖЛИВОСТІ ВИКОРИСТАННЯ ШІ В ПІДВИЩЕННІ ЕФЕКТИВНОСТІ ВИЯВЛЕННЯ

2.1. Виявлення об'єктів на основі CNN

Згорткові нейронні мережі стали фундаментом сучасних систем виявлення та розпізнавання об'єктів на зображеннях. Їхня поява знаменувала якісний прорив у розвитку комп'ютерного зору, адже саме CNN дали змогу навчати моделі без необхідності ручного виділення ознак.

Першим етапом еволюції таких систем стала поява моделі R-CNN (Region-based Convolutional Neural Network), яку запропонували в [2]. У ній вперше реалізовано механізм генерації регіонів-кандидатів (region proposals), у яких мережа здійснювала подальшу класифікацію. Це дозволило суттєво підвищити точність розпізнавання, однак мала швидкодія R-CNN обмежувала її практичне застосування.

Подальший розвиток привів до створення вдосконалених архітектур Fast R-CNN та Faster R-CNN, де було оптимізовано обчислювальні етапи, а генерація регіонів стала частиною самої мережі. Завдяки цьому вдалося досягнути значного прискорення обробки та підвищити стабільність навчання.

Інший напрямок розвитку CNN для виявлення об'єктів започаткував Джозеф Редмон, запропонувавши модель YOLO (You Only Look Once) [3]. Вона вперше об'єднала етапи локалізації та класифікації об'єктів у єдину мережеву архітектуру, що забезпечило роботу в режимі реального часу. На відміну від підходів типу R-CNN, які спочатку формували області для аналізу, YOLO розглядала все зображення одразу, передбачаючи координати рамок та класи одночасно.

Іншим вагомим досягненням став метод SSD (Single Shot MultiBox Detector), запропонований Вей Лю [4]. Його особливість полягає у використанні багаторівневих ознак — тобто поєднання даних із кількох шарів мережі для виявлення об'єктів різних

розмірів. Завдяки цьому SSD забезпечує ефективне виявлення як великих, так і менших цілей без суттєвого зниження швидкодії.

Моделі, побудовані на основі CNN, отримали широке розповсюдження в задачах виявлення об'єктів на зображеннях з безпілотних літальних апаратів (БПЛА). Це пояснюється тим, що вони забезпечують прийнятний компроміс між точністю та швидкістю, що є критично важливим при аналізі відеопотоків у реальному часі.

Однак зображення, отримані з БПЛА, мають низку специфічних особливостей, які ускладнюють застосування класичних CNN-архітектур. Серед них:

- невеликий розмір цілей, які часто займають лише кілька десятків пікселів;
- щільне розташування об'єктів у кадрі, що призводить до їх перекриття;
- варіативність масштабів через зміну висоти польоту;
- складний фон — наявність будівель, рослинності, тіней тощо.

Такі особливості спричиняють зниження точності та збільшення кількості хибних спрацьовувань. Щоб зменшити ці проблеми, Вей Лю запропонував архітектуру Feature Pyramid Network (FPN) [5]. У ній реалізовано принцип пірамідального подання ознак, який дозволяє використовувати як глибокі, так і поверхневі шари мережі для опису об'єктів різних масштабів. Це забезпечує підвищену чутливість до малих цілей і покращує якість багаторівневого аналізу зображення.

Подальші вдосконалення привели до появи таких архітектур, як PANet (Path Aggregation Network) та BiFPN (Bidirectional Feature Pyramid Network), що розвивають ідею ефективного злиття багаторівневих ознак. PANet запровадила зворотні з'єднання між шарами для кращого розповсюдження інформації, тоді як BiFPN забезпечила адаптивне зважування шарів за важливістю, що підвищило семантичну узгодженість між рівнями ознак.

Разом із тим, навіть із використанням пірамідальних структур CNN-моделі залишаються обмеженими у своїй здатності захоплювати глобальний контекст сцени. Причина полягає в тому, що згорткові операції мають обмежене *receptive field* — тобто

аналізують лише локальні області зображення. Це означає, що для коректного виявлення дрібних об'єктів у складних фонах мережі бракує інформації про віддалені частини зображення, які можуть бути важливими для контексту. На рис. 1.2. представлено приклад еволюції розвитку CNN від R-CNN до SSD та YOLO.

Як наслідок, класичні CNN добре справляються із завданнями, де об'єкти чітко виражені, мають великі розміри та однорідне тло, але демонструють зниження ефективності в аерофотозйомці, де дрібні об'єкти потребують глобального аналізу сцени. Саме ця обмеженість стала передумовою появи нових підходів, зокрема гібридних і трансформерних архітектур, які дозволяють моделі одночасно враховувати локальні деталі й довготривалі просторові залежності.



Рис. 1.2. Приклади еволюції CNN-архітектур для виявлення об'єктів: від двоетапної схеми R-CNN до одноступеневих моделей SSD та YOLO

2.2. Трансформерні підходи до виявлення об'єктів

Розвиток архітектур на основі Transformer став одним із найважливіших етапів у сучасному комп'ютерному зорі. Трансформери, що спочатку застосовувалися у сфері обробки природної мови, згодом продемонстрували високу ефективність і в задачах аналізу зображень завдяки своїй здатності моделювати довготривалі залежності між пікселями або областями зображення.

У контексті аерофотозйомки та зображень із безпілотних літальних апаратів (БПЛА) ці властивості мають особливе значення. Зображення з великих висот охоплюють широкі території, де об'єкти можуть бути розташовані на великій відстані один від одного. Тому трансформерні механізми уваги, що дозволяють аналізувати глобальний контекст сцени, є ефективнішими порівняно з класичними CNN, які обмежені локальним полем сприйняття (receptive field) [6].

Одним із перших проривів у застосуванні трансформерів для виявлення об'єктів стала модель DETR (Detection Transformer), запропонована Ніколасом Каріоном [6]. Її ключова ідея полягає у переформулюванні задачі детекції як проблеми передбачення множини об'єктів (set prediction), без необхідності у використанні класичних компонентів, таких як якорні рамки (anchor boxes) чи нелінійна максимізація (NMS — Non-Maximum Suppression).

Архітектура DETR (рис. 1.3) включає згортковий енкодер, який формує карту ознак, та трансформер-декодер, який генерує певну кількість запитів (queries) — кожен із них відповідає потенційному об'єкту на зображенні. Всі запити обробляються паралельно, що забезпечує елегантну end-to-end структуру моделі без додаткових евристичних етапів.

Проте, попри концептуальну простоту, початкова версія DETR мала низку обмежень:

1. Повільна збіжність: для досягнення конкурентної точності модель потребувала понад 500 епох навчання, що є обчислювально затратним.
2. Низька ефективність на малих об'єктах: через глобальний характер уваги модель втрачала деталі локальних структур, що призводило до погіршення точності виявлення дрібних цілей.

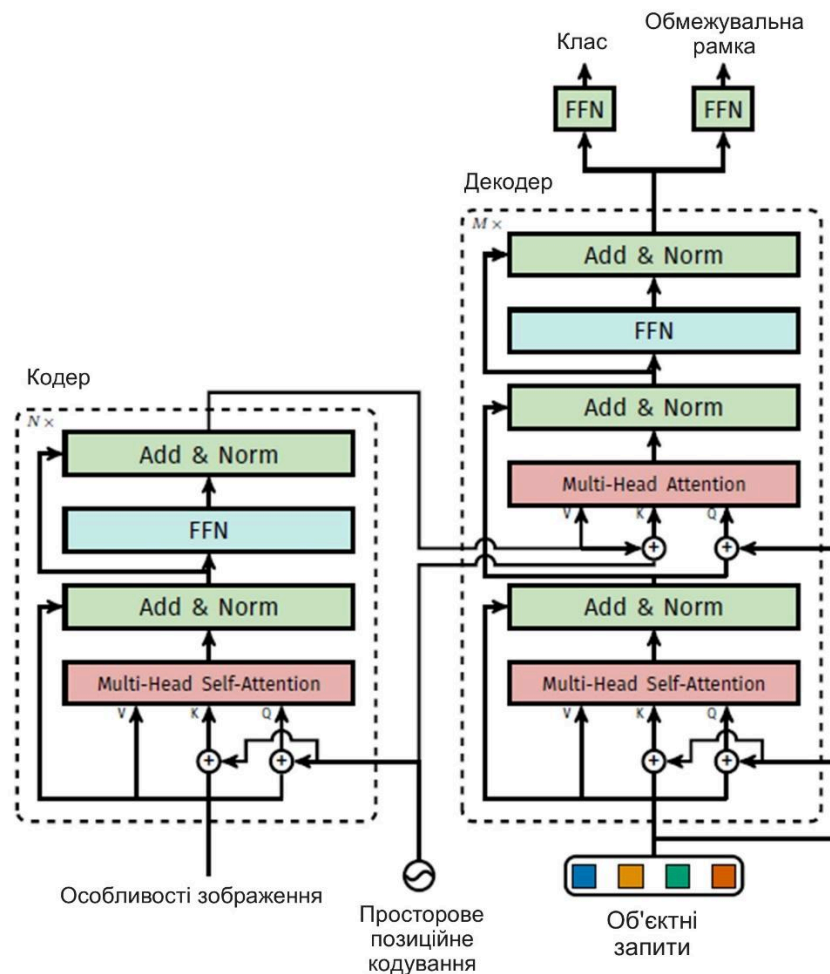


Рис. 1.3. Архітектура моделі DETR: поєднання CNN-енкодера та трансформер-декодера для передбачення множини об'єктів

Щоб усунути недоліки базової архітектури, з'явилася низка модифікацій DETR, спрямованих на підвищення швидкодії, стабільності та точності.

Deformable DETR став одним із найуспішніших рішень. Його ключова відмінність — використання деформованої (sparse) уваги, яка фокусується лише на кількох ключових точках для кожного запиту. Це дозволило суттєво зменшити обчислювальні витрати та пришвидшити збіжність навчання у 10–15 разів, при цьому зберігаючи або навіть покращуючи точність.

Подальші покращення були реалізовані у моделях:

- Conditional DETR — уточнює позиційні запити, що покращує локалізацію об'єктів;

- DAB-DETR (Dynamic Anchor Boxes DETR) — інтегрує динамічні «якірні» представлення, що допомагає адаптивно узгоджувати положення об'єктів у просторі;
- DN-DETR (Denoising DETR) — використовує техніку denoising training, додаючи шумові зразки для стабілізації навчання та прискорення збіжності.

Ці вдосконалення суттєво розширили застосування трансформерних підходів у комп'ютерному зорі, зробивши їх конкурентоспроможними не лише в точності, але й у швидкодії.

Особливу увагу привернула архітектура RT-DETR (Real-Time Detection Transformer), розроблена командою Baidu [7]. Метою цієї моделі було досягнення балансу між точністю та швидкодією, що є ключовим для застосування в реальному часі та на обмежених обчислювальних пристроях.

RT-DETR поєднує гібридний енкодер, який забезпечує взаємодію між ознаками на різних масштабах, і декодер із механізмом IoU-aware query initialization, що покращує початкове розташування запитів відносно об'єктів. Завдяки цьому модель демонструє високу швидкість обробки кадрів при збереженні конкурентної точності, що робить її придатною для практичного використання у задачах моніторингу, спостереження та навігації БПЛА.

Попри очевидні переваги, при безпосередньому застосуванні RT-DETR до аерофотозображень виявляються певні труднощі:

- деградація ознак при обробці дрібних об'єктів;
- втрата контекстуальних зв'язків між розрізненими об'єктами у широких сценах;
- зниження точності при складному фоні та великій кількості малих цілей.

Ці обмеження стали стимулом для подальших досліджень і вдосконалення RT-DETR. Зокрема, у 2025 році була запропонована модель UAV-DETR, яка орієнтована саме на задачі виявлення малих об'єктів на зображеннях з БПЛА. Вона включає спеціалізовані модулі (CAS, SOEP, CSAM), спрямовані на збереження дрібномасштабних ознак і контекстуальних зв'язків [8].

2.3. Методи виявлення, орієнтовані на зображення з БПЛА

Останні роки спостерігається активний розвиток спеціалізованих методів виявлення об'єктів, адаптованих саме для умов зйомки з БПЛА. Такі зображення мають унікальні особливості — малі розміри цілей, змінні масштаби, рух камери, складне тло та неоднорідне освітлення, що робить їх значно складнішими для аналізу, ніж звичайні наземні знімки.

З цієї причини традиційні CNN та Transformer-моделі часто не забезпечують належної точності або швидкодії. У відповідь на ці виклики дослідники запропонували низку архітектур, які враховують специфіку БПЛА, оптимізуючи процеси багаторівневого злиття ознак, покращення чутливості до дрібних цілей і зменшення обчислювального навантаження.

Одним із перших спеціалізованих підходів став UAV-YOLO, запропонований Мінцзе Ліу з NVIDIA [9].

Модель базується на архітектурі YOLO, але розширена за рахунок мультидискретного злиття ознак (multi-resolution feature fusion), що дозволяє ефективніше виявляти дрібні об'єкти. У результаті UAV-YOLO підвищує чутливість до малих цілей, зберігаючи при цьому високу швидкість.

Інша модель — TinyDet [10], — орієнтована на обчислювально обмежені платформи. TinyDet оптимізує кількість параметрів і глибину мережі, зменшуючи споживання ресурсів при мінімальній втраті точності, що робить її придатною для використання безпосередньо на борту БПЛА.

Для навчання і тестування подібних моделей активно використовуються спеціалізовані набори даних, серед яких найпоширенішими стали: VisDrone2019, UAVDT, DOTA (Dataset for Object Detection in Aerial Images).

Попри значний прогрес, більшість існуючих моделей досі стикаються з дилемою «швидкість проти точності». Частина з них досягає високої частоти кадрів, але жертвує точністю, тоді як інші, навпаки, забезпечують точне виявлення, проте не здатні працювати у реальному часі.

Перехід від CNN до трансформерних архітектур значно покращив можливості аналізу широких сцен, проте проблеми малих об'єктів, вирівнювання ознак (feature misalignment) та ефективного багатомасштабного злиття (multi-scale fusion) залишаються актуальними.

Деякі модифікації DETR, такі як Efficient-DETR та SMCA-DETR, були спрямовані на покращення збіжності навчання та точності локалізації об'єктів. Інші моделі, як-от CFPT (Cross-Layer Feature Pyramid Transformer) та LGI-DETR (Local-Global Interaction DETR), використовують механізми міжшарної уваги (cross-layer attention) для кращої взаємодії між локальними та глобальними ознаками.

Однак, попри підвищення точності, ці архітектури мають надмірну структурну складність. Наприклад:

- CFPT містить 51,3 млн параметрів;
- LGI-DETR — 21,1 млн параметрів.

Обидві сильно залежать від додаткових модулів на боці декодера. Це призводить до значних обчислювальних витрат і ускладнює використання моделей у режимі реального часу.

Останні роки характеризуються активними спробами підвищити точність виявлення малих об'єктів на наборі VisDrone2019-DET, який став еталоном у цій галузі.

Команда розробників на чолі з Деном Джао з Національного університету Сінгапуру [11] вдосконалили модель YOLOv7, досягнувши $mAP@0.5 = 52,3 \%$, проте її обсяг становить близько 35,8 млн параметрів, що обмежує можливість застосування на малопотужних пристроях. Наступним етапом була розробка YOLO-PEL, компактна версія з лише 2,23 млн параметрів, але її точність ($mAP@0.5 = 32,5 \%$) виявилася недостатньою для практичного використання.

Серед трансформерних моделей варто відзначити DV-DETR. Ця архітектура вдосконалює RT-DETR і досягає $mAP@0.5 = 50,2 \%$ при 19,5 млн параметрів, демонструючи кращу збалансованість між точністю та розміром моделі. Однак проблеми вирівнювання ознак і контекстної інтеграції залишаються.

Дослідники з Університету Іллінойсу [12] створили модель RFAG-YOLO, яка використовує механізм уваги до поля сприйняття (Receptive Field Attention). Вона показала $mAP@0.5 = 38,9 \%$ при 5,9 млн параметрів, що є непоганим результатом для легких моделей, але точність усе ще недостатня.

Поєднання технології HRFPN (High-Resolution Feature Pyramid Network) і EfficientViMamba, дала можливість досягнути $mAP@0.5 = 38,9 \%$ при 33,5 млн параметрів, що свідчить про високу якість виявлення, проте знову — при високій обчислювальній складності.

Ще одним кроком уперед стала модель Drone-DETR. Вона є вдосконаленням RT-DETR для дистанційного зондування та аерофотозйомки, демонструючи найвищий результат — $mAP@0.5 = 53,9 \%$ при швидкості 30 кадрів/с. Проте обчислювальна складність досягає 128,3 GFLOPs, а кількість параметрів — 28,7 млн, що обмежує можливість практичного застосування на бортових системах БПЛА.

Для подолання описаних обмежень у 2025 році була розроблена модель UAV-DETR — вдосконалена версія RT-DETR, спеціально адаптована для задач виявлення малих об'єктів на зображеннях із БПЛА [1].

UAV-DETR включає три інноваційні модулі:

- CAS — адаптивно зважує каналні ознаки, підвищуючи чутливість до слабких сигналів від малих цілей;
- SOEP — покращує багатомасштабне злиття ознак, особливо на рівнях з низьким масштабом (P2, P3);
- CSAM — забезпечує точне узгодження просторових і контекстних зв'язків між рівнями ознак.

Модель досягає $mAP@0.5 = 51,6 \%$, зберігаючи швидкість 30 кадрів/с при лише 16,8 млн параметрів і 71,4 GFLOPs.

Таке співвідношення точності, компактності та швидкодії робить UAV-DETR однією з найперспективніших архітектур для практичного застосування в системах моніторингу, розвідки й безпеки на базі безпілотників.

2.4. Модель RT-DETR

Суттєвий прорив у сфері виявлення об'єктів у реальному часі забезпечила модель RT-DETR, розроблена дослідницькою командою Baidu. Вона стала відповіддю на основні недоліки попередніх трансформерних архітектур, зокрема DETR, головним з яких була висока обчислювальна складність та повільна збіжність навчання, що фактично унеможливлювало їх використання у задачах реального часу.

RT-DETR зберігає концепцію енд-то-енд (end-to-end) обробки без необхідності у постобробці результатів (зокрема, без NMS — Non-Maximum Suppression), але при цьому значно оптимізує внутрішню структуру, щоб зменшити обчислювальні витрати та прискорити процес навчання.

Ключовою інновацією моделі є гібридний енкодер, який складається з двох основних модулів (рис. 1.4):

1. AIFI (Attention-Enhanced Intra-scale Feature Interaction) — модуль уваги, що забезпечує взаємодію ознак усередині одного масштабу. Він дозволяє глибше аналізувати контекст у межах конкретного рівня ознак, не змішуючи інформацію з інших масштабів. Це підвищує стабільність і точність локальної обробки.

2. CCFF (Cross-Scale Feature Fusion) — модуль, який реалізує злиття ознак між різними масштабами. Він оптимізує обмін інформацією між високорівневими семантичними ознаками (які відповідають великим об'єктам) і низькорівневими детальними ознаками (важливими для малих об'єктів).

Такий підхід дозволяє розділити процеси внутрішньомасштабної взаємодії та міжмасштабного злиття, що істотно підвищує ефективність обчислень і зменшує кількість надлишкових операцій.

Ще однією важливою особливістю RT-DETR є механізм IoU-aware query selection, який використовується для оптимізації ініціалізації запитів у декодері.

У класичних DETR-моделях усі запити генеруються довільно, і лише під час навчання вони поступово «вчаться» відповідати певним об'єктам на зображенні. Це часто призводить до нестабільності й потреби у великій кількості епох.

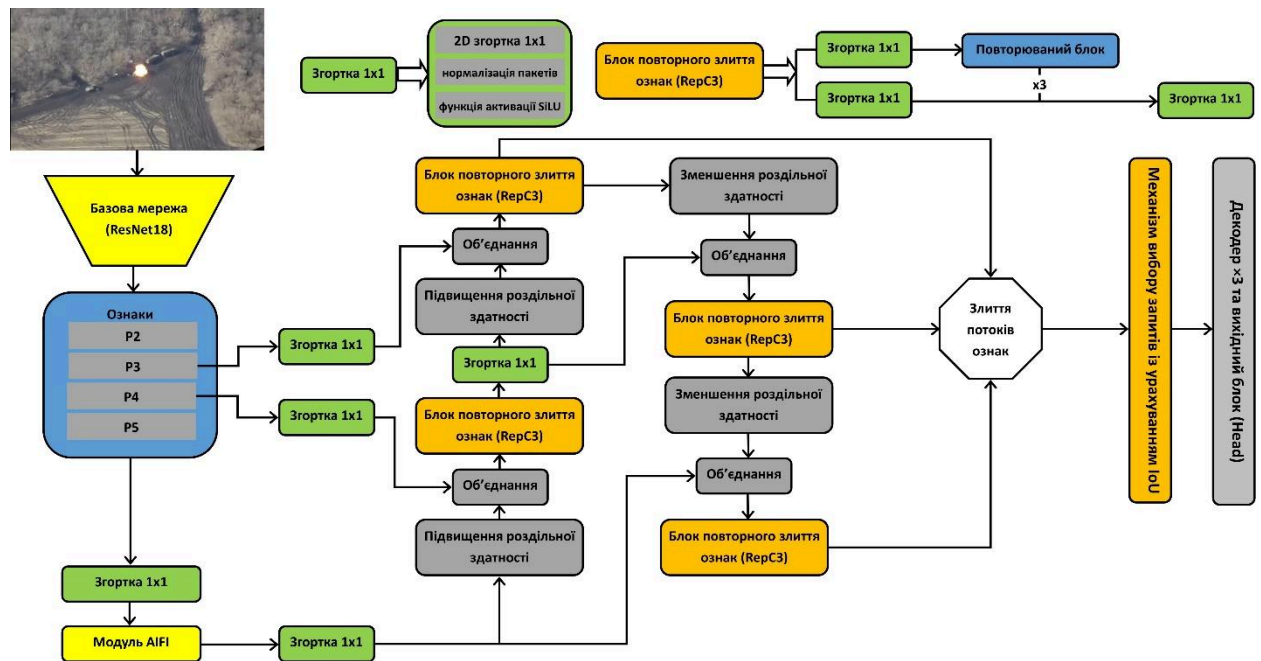


Рис. 1.4. Схема побудови моделі RT-DETR

У RT-DETR цей процес покращено: запити ініціалізуються на основі індексів IoU (Intersection over Union), тобто їх початкове розташування узгоджене з найбільш ймовірними регіонами об'єктів. Такий підхід прискорює збіжність навчання, підвищує стабільність і зменшує кількість хибних спрацьовувань під час інференсу.

RT-DETR досягла значного покращення співвідношення «точність/швидкодія» у порівнянні з попередніми архітектурами. Основні переваги:

- відсутність постобробки: модель напряму генерує результати, не потребуючи етапів NMS або фільтрації рамок;
- менша кількість параметрів і обчислень: завдяки оптимізованому енкодеру та легшій структурі;
- висока швидкість інференсу: RT-DETR досягає продуктивності понад 30 кадрів/с на стандартних GPU;
- ефективне злиття ознак: забезпечує кращу багатомасштабну адаптацію, зокрема для середніх об'єктів.

Саме завдяки цим властивостям модель набула популярності у задачах відеоаналітики, моніторингу, безпеки та автономного транспорту.

Попри високу ефективність, пряме застосування RT-DETR до аерофотознімків або відео з БПЛА виявило низку проблем, пов'язаних зі специфікою таких даних.

Серед недоліків можна виділити наступні:

- втрати ознак через надмірні операції даунсемплінгу.

У базовій архітектурі RT-DETR кілька послідовних етапів зменшення роздільної здатності (downsampling) призводять до втрати важливої просторової інформації. Це особливо критично при виявленні дрібних цілей, які можуть повністю «зникнути» після кількох згорткових шарів;

- просторова невідповідність (spatial misalignment).

Під час злиття ознак між різними масштабами можливі зміщення регіонів, що ускладнює точне визначення положення об'єктів;

- неврахування контекстуальної взаємодії. У сценах з БПЛА об'єкти часто сильно залежать від оточення (наприклад, автомобіль на дорозі чи людина біля будівлі). Відсутність ефективного моделювання контексту знижує точність розпізнавання;

- недостатня деталізація малих об'єктів.

Навіть при використанні SSFF-модуля, дрібні об'єкти залишаються слабо представленими у високорівневих ознаках, що призводить до пропусків або хибних детекцій.

Ці обмеження зумовили потребу у подальшому вдосконаленні RT-DETR, зокрема для використання у безпілотних системах спостереження, де важливими є компактність моделі, висока чутливість до малих об'єктів і робота в реальному часі.

ВИСНОВКИ ДО РОЗДІЛУ 2

У другому розділі було проведено ґрунтовний аналіз сучасних підходів і методів виявлення малих об'єктів на зображеннях, отриманих із безпілотних літальних апаратів (БПЛА).

У підпункті 2.1 розглянуто основні особливості задачі виявлення малих об'єктів у повітряних зображеннях. Визначено, що низька роздільна здатність об'єктів,

складність фону, варіативність масштабу та нестабільність освітлення істотно ускладнюють обробку таких даних. Показано, що для підвищення ефективності детекції необхідне поєднання багаторівневого аналізу ознак, контекстного моделювання та адаптивного злиття інформації між різними рівнями нейромережі.

У підпункті 2.2 виконано огляд методів на основі згорткових нейронних мереж (CNN), що стали основою сучасних систем комп'ютерного зору. Розглянуто еволюцію архітектур — від R-CNN до YOLO та SSD, а також появу пірамідальних структур (FPN, PANet, BiFPN), які підвищили ефективність роботи з об'єктами різного масштабу. Водночас підкреслено, що обмежене поле сприйняття CNN не дозволяє повною мірою враховувати глобальний контекст сцени, що особливо критично для аерофотозйомки з БПЛА.

У підпункті 2.3 проаналізовано трансформерні підходи до виявлення об'єктів. Описано архітектуру DETR та її модифікації — Deformable DETR, Conditional DETR, DAB-DETR, DN-DETR, а також RT-DETR. Показано, що механізми глобальної уваги трансформерів дозволяють ефективно моделювати просторові зв'язки між об'єктами на великих територіях, однак проблема точного виявлення малих об'єктів залишається невирішеною.

Також наведено порівняльний аналіз методів, адаптованих під зображення з БПЛА. Розглянуто моделі UAV-YOLO, TinyDet, DV-DETR, RFAG-YOLO, Drone-DETR тощо. Показано, що хоча більшість із них забезпечує підвищену чутливість до дрібних цілей, вони часто мають надмірну обчислювальну складність, що обмежує їх застосування у реальному часі. Визначено, що модель UAV-DETR досягає найкращого балансу між точністю, компактністю та швидкодією, що робить її найбільш придатною для практичного використання в системах обробки даних з БПЛА.

У підпункті 2.4 детально розглянуто архітектуру RT-DETR, яка стала основою для подальших вдосконалень. Визначено, що модель завдяки гібридному енкодеру (модулі AIFI та CCFF) і механізму IoU-aware query selection забезпечує високу швидкодію та усуває потребу в постобробці результатів. Разом з тим, при обробці

даних з БПЛА виявлено проблеми втрати ознак малих об'єктів через даунсемплінг, просторової невідповідності та відсутності контекстної інтеграції.

Узагальнюючи результати аналізу, можна зробити висновок, що ефективно виявлення малих об'єктів на аерофотознімках потребує моделей, які:

- зберігають деталізацію дрібних ознак при глибокій обробці;
- поєднують локальні та глобальні контекстуальні залежності;
- мають помірну кількість параметрів і можуть працювати у режимі реального часу;
- забезпечують стабільність в умовах обмежених обчислювальних ресурсів.

Саме ці вимоги лягли в основу розробки UAV-DETR, яка поєднує ефективність RT-DETR із додатковими модулями (CAS, SOEP, CSAM) для покращеного злиття ознак, узгодження просторового контексту та підвищення точності виявлення малих цілей.

РОЗДІЛ 3

МЕТОДОЛОГІЯ ВДОСКОНАЛЕННЯ МОДЕЛІ UAV-DETR ДЛЯ ВИЯВЛЕННЯ МАЛИХ ОБ'ЄКТІВ БПЛА

3.1. Вступ

Попередній аналіз показав, що модель UAV-DETR є однією з найефективніших архітектур для виявлення малих об'єктів на зображеннях з БПЛА. Вона поєднує високу точність, швидкодію та адаптивність до складних фонових умов завдяки використанню трьох ключових модулів — CAS (Channel-Aware Sensing), SOEP (Scale-Optimized Enhancement Pyramid) та CSAM (Context-Spatial Alignment Module).

Однак, попри високі результати, модель UAV-DETR все ще має потенціал для вдосконалення. Зокрема, існують проблеми, пов'язані з:

- втратою дрібних просторових деталей у нижніх рівнях ознак через зменшення роздільної здатності;
- неповною кореляцією між каналами ознак під час злиття інформації з різних масштабів;
- надлишковими обчисленнями, які впливають на швидкодію при роботі на бортових пристроях із обмеженими ресурсами.

Метою даного розділу є розроблення вдосконаленої архітектури UAV-DETR, спрямованої на підвищення ефективності виявлення малих об'єктів БПЛА шляхом покращення механізмів багатомасштабного злиття ознак, узгодження просторового контексту та оптимізації обчислювальних витрат.

У межах запропонованої методології буде розглянуто:

- структурну схему вдосконаленої моделі;
- алгоритмічну послідовність обробки даних;
- удосконалення модулів CAS, SOEP і CSAM;
- експериментальні параметри навчання та валідації.

Передбачається, що розроблена методика дозволяє зберегти основні переваги UAV-DETR — роботу в реальному часі, стабільність і компактність, водночас підвищивши точність і стійкість до фонових перешкод при виявленні малих цілей у складних аерофотосценах.

3.2. Загальна структура вдосконаленої моделі UAV-DETR

Метою запропонованого вдосконалення є підвищення розпізнавальної здатності щодо малогабаритних цілей на аерозображеннях при збереженні можливості реального-часового інференсу на обмежених обчислювальних платформах. Основна ідея — поєднати додаткове багаторівневе представлення (введення більш високороздільного рівня ознак), посилити канали і просторові механізми уваги та оптимізувати злиття ознак, зберігаючи низьку складність моделі.

Модернізація базується на архітектурі UAV-DETR (CAS, SOEP, CSAM) і включає ряд цілеспрямованих змін, кожна з яких обґрунтована з точки зору підвищення чутливості до малих об'єктів і/або зменшення втрат ознак під час даунсемплінгу.

Удосконалена архітектура зберігає загальну топологію UAV-DETR (backbone → neck (SOEP) → hybrid encoder (AIFI + CCFF/CSAM) → decoder → head), але вводить наступні ключові зміни (рис.3.1):

1. Додатковий високороздільний рівень ознак P1 (входить у SOEP) — для збереження дрібних деталей перед сильними даунсемплінг-операціями. P1 формується або шляхом часткового збереження ранніх карт ознак backbone, або через паралельний шар з малим страйдом (stride=1).

2. CAS → CAS++ (модифікація Channel-Aware Sensing) — розширення канал-чутливої уваги: інтеграція каналної взаємодії з груповим нормуванням і легковаговою міжканальною матрицею кореляцій (approximate channel correlation) для кращого підсилення слабких ознак малих об'єктів.

3. SOEP (Scale-Optimized Enhancement Pyramid) → SOEP-v2 — додавання SPDCConv-2 (модифікована SPDCConv) і інтеграція P1 у піраміду; застосування CSP-

OmniKernel із полегшеними великими ядрами (розклад на сепарабельні $1 \times k$ і $k \times 1$) для контролю FLOPs.

4. CSAM \rightarrow CSAM-adv (Context-Spatial Alignment Module) — посилення CFC і SFC:

- CFC: багаторівневе каскадне збирання контексту з адаптивним масштабним вагуванням (щоб уникнути домінації інформації від великих об'єктів);

- SFC: точніше обчислення офсетів через легкі деформовані блоки з обмеженою кількістю голів (heads) і застосуванням масок-гейтів $1 + \tanh(\cdot)$ для стабільності.

5. Оптимізація обчислень і розміру моделі: застосування технік *sparse pruning* на непоказових шарах, квантизація (при необхідності для *inference* на борту), і *knowledge distillation* з великої версії моделі на полегшену для збереження точності при зменшенні Params.

6. Навчальна стратегія: комбіновані втрати з акцентом на малі об'єкти (адаптивна вага для прикладів з розміром $< 32 \times 32$), доповнена денойзинговою підготовкою та спеціальними аугментаціями (масштабування, шум, імітація погодних умов).

Процес модернізації базової архітектури UAV-DETR відбувався поетапно, із послідовним удосконаленням ключових структурних модулів моделі.

Введення рівня P1 (високороздільний рівень). В неск додано рівень P1 з роздільністю вдвічі більшою, ніж P2 (або як альтернатива, можна використати ранню карту ознак з backbone без даунсемплінгу). P1 підключається до SOEP-v2 і до CSAM-adv для збереження локальних подробиць. Це зроблено з тих міркувань, що малі об'єкти найчастіше втрачають інформацію вже на ранніх стадіях доунсемплінгу. Наявність P1 дозволяє не «розмивати» їхню структуру і покращити локальні ознаки перед багатомасштабним злиттям. При цьому, буде помірне зростання FLOPs, що компенсується полегшенням інших блоків, тобто виконується загальна оптимізація. Як наслідок, покращення каналу-уваги CAS.

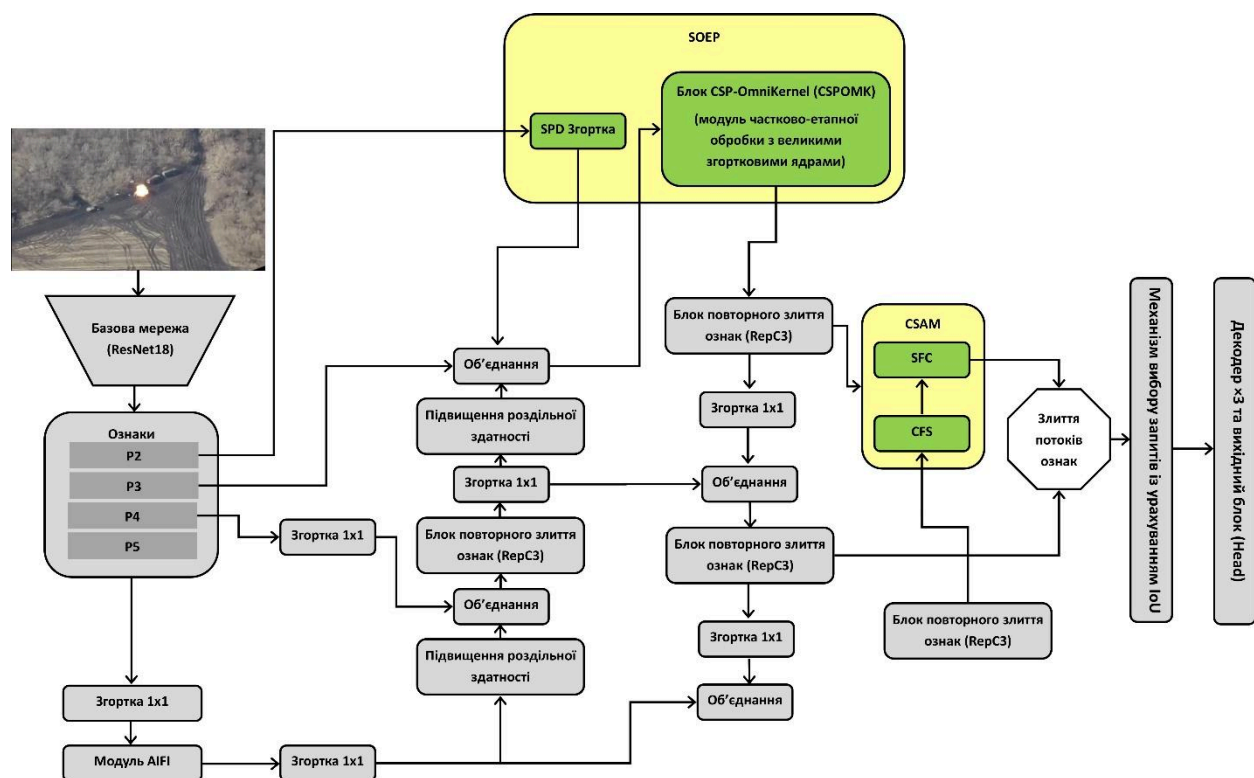


Рис. 3.1. Загальна структурна схема вдосконаленої UAV-DETR

Далі було проведено оновлення механізму каналної уваги — модуль CAS (Channel-Aware Sensing) перетворено у CAS++, у якому замість традиційного механізму squeeze-excitation реалізовано групове зважування каналів із використанням малої матриці кореляцій (low-rank approximation). Такий підхід дозволяє моделі враховувати міжканальні залежності та кореляційні зв'язки, що підсилює релевантні ознаки, особливо важливі для виявлення малих і слабконтрастних об'єктів. Додатково було впроваджено механізм пропусків (skip connections), який запобігає надмірній фільтрації слабких сигналів і забезпечує стійкість моделі під час навчання. У результаті модуль CAS++ забезпечує більш точне виділення каналів, що містять корисну інформацію про дрібні цілі, без істотного збільшення кількості параметрів.

Наступним етапом удосконалення стала модернізація модуля SOEP (Scale-Optimized Enhancement Pyramid), який отримав нову версію — SOEP-v2. Основними змінами стали інтеграція додаткового високороздільного рівня P1, що дозволяє зберігати локальні деталі малих об'єктів, а також введення адаптивного вагового

механізму при злитті рівнів ознак P1...P5. Ключовим структурним елементом SOEP-v2 став блок CSP-OmniKernel (CSPOMK) — модуль частково-етапної обробки з великими згортковими ядрами, який реалізує глобальне збирання контексту за допомогою розщеплення великих ядер на послідовність згорток $1 \times k$ та $k \times 1$. Поряд із цим застосовано SPDConv-2 з оптимізованим коефіцієнтом розділення ($s=2$) і додатковим каналовим редуційним шаром для зниження обчислювальних витрат. Завдяки такому поєднанню вдалося зберегти баланс між локальними деталями (через P1 і SPDConv-2) та глобальною контекстною інформацією (через CSP-OmniKernel), при цьому зберігаючи контрольований рівень складності моделі.

Подальше вдосконалення стосувалося модуля узгодження контексту й простору CSAM (Context-Spatial Alignment Module), який у новій версії отримав розширений варіант CSAM-adv. Цей блок складається з двох підмодулів — CFC (Contextual Fusion & Calibration) та SFC (Spatial Feature Calibration). У CFC реалізовано багаторівневе пірамідальне pooling-представлення з обчисленням афінних ваг для кожного рівня, що дає змогу адаптивно підсилювати контекст для пікселів із низькою впевненістю. У підмодулі SFC застосовано полегшений деформований оператор ΔT , який обчислює офсети для просторового вирівнювання карт ознак, причому кількість таких офсетів обмежена за групами для зменшення обчислювальних витрат. Додатково впроваджено маски-затвори (gating masks), які регулюють внесок кожного масштабного рівня в кінцеве представлення. У результаті модуль CSAM-adv значно підвищує узгодженість між картами ознак різних розмірів, зменшує кількість хибних накладень і пропусків, а також покращує якість локалізації дрібних об'єктів.

Заключним етапом модернізації стали оптимізаційні та деплой-процедури, спрямовані на забезпечення ефективної роботи моделі в обмежених обчислювальних середовищах. Для цього використано метод knowledge distillation, за якого велика «тренувальна» модель без квантизації передає знання компактній версії для виконання процесу розпізнавання безпосередньо на борту БПЛА або на вбудованих пристроях. Після навчання застосовується пост-тренувальна квантизація (8-бітна або змішана точність FP16/INT8), що зменшує споживання пам'яті та підвищує швидкодію. Додатково проводиться помірне прорідження фільтрів (pruning), яке

видаляє надлишкові параметри без істотного впливу на показники точності. Завдяки такому підходу вдалося знизити кількість параметрів моделі до рівня 16–18 млн, зберігаючи або навіть покращуючи значення mAP для малих об'єктів.

Таким чином, запропонований процес модернізації UAV-DETR є системним та спрямованим на підвищення точності, стабільності й ефективності моделі, зберігаючи при цьому її легковагову архітектуру, необхідну для реального використання на безпілотних платформах.

Вхідний потік обробки у вдосконаленій архітектурі починається з подачі вхідного зображення на базову мережу (backbone) — це може бути CasNet або модифікований ResNet-18 з інтегрованим модулем CAS++; у результаті з backbone отримуються карти ознак різної роздільності P1-P5. Далі ці карти проходять через модуль SOEP-v2, де кожний рівень спочатку піддається попередній обробці (SPDConv-2 та CSP-OmniKernel), після чого результати з різних масштабів поєднуються з використанням адаптивних ваг, формуючи єдину багатомасштабну карту ознак. Отримане багатомасштабне подання надходить у CSAM-adv, де спочатку виконується CFC (контекстне калібрування) для узагальнення і підсилення релевантного контексту, а потім — SFC (просторове вирівнювання) для корекції просторових зміщень між картами; у підсумку модель отримує скориговані, узгоджені карти ознак. Ці карти подаються у гібридний енкодер, що поєднує модулі AIFI та CCFF, і далі ініціалізація запитів для декодера проводиться з урахуванням IoU (aware query initialization), що сприяє швидшій збіжності та точнішій локалізації. На етапі декодування (Decoder → Head) система генерує прогнози — координати обмежувальних рамок (bbox) і ймовірнісні оцінки класів. Після завершення навчання проводиться підготовка моделі до впровадження: застосовується knowledge distillation (передача знань від великої тренувальної моделі до компактної версії), легке прорідження (pruning) фільтрів та пост-тренувальна квантизація (наприклад, 8-бітна) для зниження обсягу пам'яті й прискорення inference на бортових пристроях.

Щодо очікуваних метричних результатів, зацікавленість спрямована на такі показники: mAP@0.5 (особливо на підмножині малих об'єктів), mAP@0.5:0.95, Recall для об'єктів розміром <math><32\times 32</math> пікселів, а також обчислювальні характеристики —

GFLOPs і FPS на типових бортових GPU/TPU. За розумного налаштування та підходу до навчання прогнозується збільшення mAP@0.5 на 1–3 процентних пункти порівняно з оригінальною версією UAV-DETR на наборі VisDrone2019, перш за все за рахунок підвищення Recall для малих цілей, при збереженні реального часу роботи (приблизно ~30 FPS) після застосування оптимізацій. Слід зазначити, що ці очікування впливають із принципів збереження додаткового рівня P1 та посилення каналової і просторової уваги; остаточне підтвердження потребує експериментальної валідації.

Запропоновані вдосконалення підвищують здатність системи виявляти дрібні, слабо видимі об'єкти в умовах складного фону та змін масштабу сцени, тому на концептуальному рівні вони роблять модель корисною для систем розвідки та моніторингу, що застосовуються в різних операційних контекстах, включаючи потреби збройних формувань у межах захисту територіальної цілісності. Водночас необхідно підкреслити: в межах цієї кваліфікаційної роботи не надаються будь-які тактичні, оперативні або інструктивні рекомендації щодо конкретних військових дій. Більш прийнятним є формулювання, що методологія й отримані результати підвищують ефективність автоматизованих систем спостереження, які використовуються у цивільному та оборонному секторах для задач ситуаційної обізнаності, пошуку/рятування та безпеки; модель описується на рівні алгоритмічних підходів і не містить тактичних вказівок щодо її застосування.

3.3. Технічний розбір модифікованих модулів вдосконаленої моделі UAV-DETR

Вдосконалена модель UAV-DETR складається з трьох основних компонентів, які забезпечують підвищену точність виявлення малих об'єктів при мінімальному зростанні обчислювальної складності:

- CAS++ (Channel-Aware Sensing Plus) – адаптивна система каналової уваги;

- SOEP-v2 (Scale-Optimized Enhancement Pyramid, версія 2) – покращена багатомасштабна піраміда ознак;
- CSAM-adv (Context-Spatial Alignment Module, розширений) – модуль узгодження контексту й простору.

Кожен із цих модулів є логічним продовженням базових блоків UAV-DETR, проте має власні технічні особливості, спрямовані на підвищення чутливості до дрібних цілей у зображеннях з БПЛА.

3.3.1. Модуль CAS++ (*Channel-Aware Sensing Plus*)

У наборах даних на кшталт VisDrone-2019, що містять зображення, отримані з низьких висот польоту БПЛА, об'єкти мають надзвичайно малі розміри — від кількох до кількох десятків пікселів. При цьому фон зазвичай насичений щільною міською текстурою, відблесками та ефектами розмиття руху. Традиційні згорткові нейронні мережі (CNN), які базуються на фіксованих областях сприйняття (receptive fields), мають труднощі з одночасним моделюванням глобального контексту та детальним вилученням локальних ознак, необхідних для точного розпізнавання настільки дрібних цілей.

Щоб подолати ці обмеження, у запропонованій архітектурі базові блоки (BasicBlock) у складі backbone замінено на блоки Convolutional Additive Self-Attention (CAS++, рис.3.2) [13]. Кожен блок CAS спочатку використовує послідовність із трьох глибинних сепарабельних згорток (3×3 depthwise separable convolutions) для інтеграції локальних просторових ознак і неявного кодування позиційної інформації, що є критично важливим для точного визначення цілей розміром менше 20 пікселів.

Після цього задіюється підмодуль Convolutional Additive Token Mixer (CATM), який обчислює вагові коефіцієнти уваги одночасно у просторовому та каналному вимірах, використовуючи полегшені комбінації згорткових і сигмоїдних операцій. Просторовий компонент дозволяє пригнічувати фон і шум, виділяючи області, що потенційно містять малі об'єкти, тоді як каналний компонент підсилює дискримінативні ознаки, допомагаючи моделі відокремлювати об'єкти від складного

тла. Обидва типи ваг об'єднуються адитивно, утворюючи динамічний механізм повторного калібрування ознак.

Важливо, що весь процес не потребує softmax-нормалізації та великомасштабних матричних множень, завдяки чому обчислювальна складність CAS-блоків залишається співмірною зі стандартними згортками, але при цьому забезпечується можливість довгодистанційної просторової взаємодії та міжканальної комунікації в одному проході вперед (forward pass). Такий підхід дозволяє використовувати контекстну інформацію для уточнення нечітких або частково закритих цілей, а адитивне об'єднання ваг зберігає слабкі сигнали, які зазвичай губляться у моделях, що покладаються на softmax-нормування. У результаті мережа формує точніші реакції на дрібні та перекриті об'єкти, залишаючись придатною для використання на мобільних і вбудованих обчислювальних платформах, де важлива економія ресурсів.

Щоб узгодити архітектуру з вимогами до обчислювальної ефективності та простоти розгортання, було вирішено використовувати простішу комбінацію згорткових операцій замість дорогих матричних множень або softmax-нормалізацій, які суттєво ускладнюють inference. Виходячи з цього підходу, для тензора ознак розміром $H \times W \times C$ взаємодію інформації доцільно розділити на просторову область ($N = H \times W$) та каналний вимір (C). У попередніх роботах також досліджувалися методи масштабної декомпозиції, які використовували багатоголові структури вздовж каналного виміру для паралельних обчислень або ділення просторових патчів для прискорення процесу. Проте ці підходи, по суті, все одно оперують основними вимірами тензора ознак.

Концепція просторової взаємодії продемонстрована на рис. 3.2, де на вхід подається тензор $x \in \mathbb{R}^{H \times W \times C}$. Спочатку локальні токени агрегуються за допомогою шару згортки 3×3 , який збирає інформацію з найближчих областей. Далі каналовий вимір зменшується до одиничного шляхом застосування згортки 1×1 , після чого використовується функція активації Sigmoid для формування просторової карти уваги $x_s \in \mathbb{R}^{H \times W \times 1}$. Отримана карта масштабує вхідні ознаки, підкреслюючи релевантні області відповідно до формули:

$$x_s = \text{Sigmoid}\left(D_{1 \times 1}\left(D_{3 \times 3, \text{ReLU}, \text{BN}}(x)\right)\right) \odot x, \quad (3.1)$$

де $D_{3 \times 3, \text{ReLU}, \text{BN}}$ - послідовність операцій згортки 3×3 , активації ReLU та нормалізації BatchNorm, символ \odot позначає покомпонентне множення.

У підсумку така структура дозволяє ефективно інтегрувати локальні та глобальні залежності, підтримуючи високу точність детекції при мінімальному зростанні обчислювальної складності, що є ключовим для застосування моделі в умовах обмежених ресурсів, наприклад, на борту безпілотних літальних апаратів.

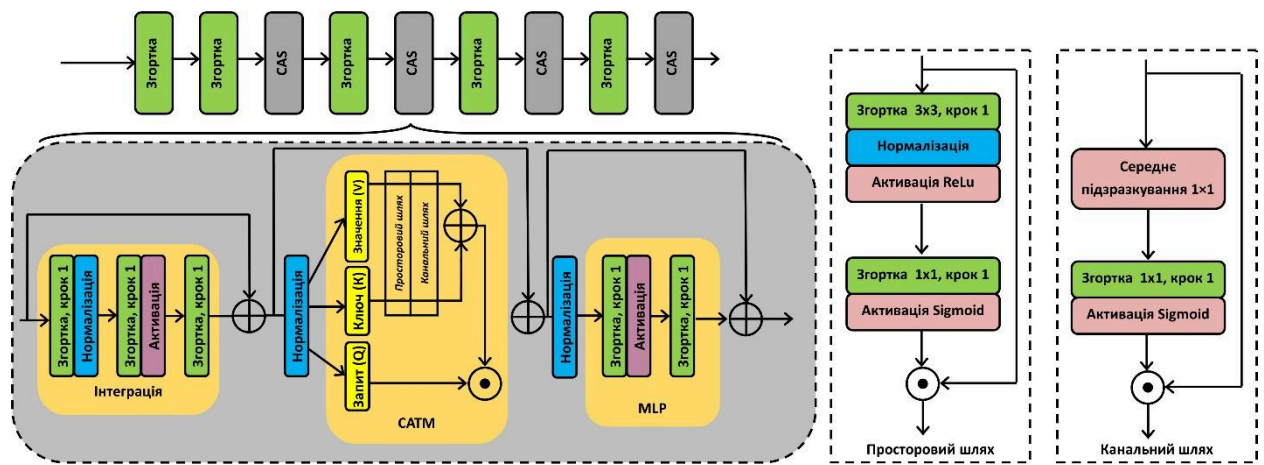


Рис. 3.2. Архітектура CAS++

Для моделювання міжканальних взаємодій у вдосконаленій архітектурі було використано ідеї, запозичені з SENet [14], однак із принциповою відмінністю: замість традиційного зменшення кількості каналів (channel reduction) застосовується згортка 1×1 , яка безпосередньо захоплює й інтегрує інформацію між каналами. Такий підхід дозволяє зберегти повну глибину ознак, не втрачаючи критичної інформації про дрібні цілі. Формально процес описується рівнянням:

$$x_c = \text{Sigmoid}\left(D_{1 \times 1}(P_{1 \times 1}(x))\right) \odot x, \quad (3.2)$$

де P — це адаптивне pooling-перетворення, D — групова згортка (group convolution), у якій кількість груп за замовчуванням дорівнює кількості каналів.

Комбінація цих двох операцій створює узагальнену карту ознак після просторових та каналних взаємодій, яку позначимо як $\Phi(x)$.

Для підсилення контекстного зв'язку між ознаками модель вводить нову метрику подібності (similarity metric), що визначається як сума контекстних оцінок між відповідними елементами матриць Q і K , де $Q, K \in \mathbb{R}^{N \times C}$:

$$\text{Sim}(Q, K) = \Phi(Q) + \Phi(K), \quad (3.3)$$

Вектори Запит (Q , query), Ключ (K , key) та значення (V , value) формуються шляхом незалежних лінійних перетворень, наприклад:

$$Q = W_q x, K = W_k x, V = W_v x, \quad (3.4)$$

У цьому випадку функція контекстного відображення $\Phi(\cdot)$ містить основну інформацію про взаємодії між ознаками, що дозволяє мережі навчатися без необхідності ручного проєктування контекстних залежностей.

Перевагою такого узагальнення є те, що воно виходить за межі традиційного ручного контекстного моделювання й натомість формулюється у вигляді операцій згортки, що значно спрощує реалізацію та підвищує узгодженість моделі з сучасними апаратними прискорювачами. Завдяки послідовним взаємодіям, виконаним над матрицями Q та K , мережа здатна вибірково підсилювати найбільш інформативні токени, які несуть максимальну цінність для детекції дрібних об'єктів у складних сценах.

Об'єднання цих компонентів реалізується адитивним способом, що є лінійним і не потребує обчислювально затратних матричних множень. Такий підхід дає змогу уникнути високої складності, властивої традиційним механізмам уваги, водночас зберігаючи ефективність у вилученні та передачі корисної інформації. Вихід модуля CATM (Convolutional Additive Token Mixer) можна описати такою формулою:

$$O = \Gamma(\Phi(Q) + \Phi(K)) \odot V, \quad (3.5)$$

де $\Gamma(\cdot) \in \mathbb{R}^{N \times C}$ — це лінійне перетворення, яке забезпечує інтеграцію контекстної інформації та передає її до наступних рівнів обробки.

У підсумку, описаний механізм поєднує ефективність адитивного моделювання контексту з низькою обчислювальною складністю, що робить його придатним для використання на ресурсно обмежених пристроях — зокрема, на бортових обчислювальних системах безпілотних літальних апаратів, де поєднання швидкодії й точності є критичним.

3.3.2. Модуль SOEP (Scale-Optimized Enhancement Pyramid)

Модуль SOEP (Scale-Optimized Enhancement Pyramid) розроблено для підвищення ефективності виявлення малих об'єктів, зокрема безпілотних літальних апаратів (БПЛА), які на звичайних рівнях детекції (P3–P5) часто залишаються непоміченими через обмежену роздільність ознак. Традиційне рішення полягає у додаванні вищого рівня роздільності P2, який здатен відтворювати дрібні деталі об'єктів. Однак таке доповнення істотно збільшує обчислювальні витрати та час постобробки, що є критичним для систем реального часу, особливо у випадку використання на борту БПЛА.

Тому виникає необхідність у створенні нових багатомасштабних структур злиття ознак, які забезпечують високу точність виявлення без істотного зростання складності. У цьому дослідженні пропонується обчислювально ефективна мережа злиття ознак, оптимізована для малих цілей (рис. 3.3), у якій збережено баланс між деталізацією та швидкістю. Запропонована архітектура об'єднує рівень P2, що містить детальну інформацію про дрібні об'єкти, із рівнем P3, який проходить попередню обробку за допомогою SPDCConv.

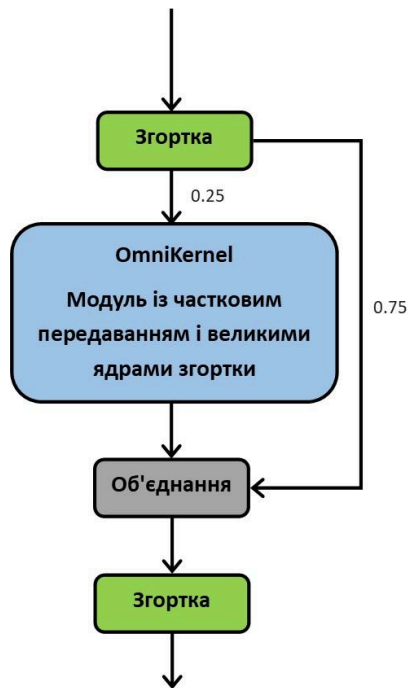


Рис. 3.3. Архітектура модуля SOEP на основі OmniKernel

Ключовим елементом модуля є CSP-OmniKernel, що реалізує адаптивне злиття ознак на основі принципів CSP (Cross Stage Partial) та OmniKernel [15]. Така комбінація дозволяє одночасно враховувати локальні й глобальні просторові залежності, розширюючи діапазон ефективного сприйняття без різкого збільшення кількості параметрів.

Завдяки цьому модуль SOEP забезпечує більш точне вилучення характеристик дрібних об'єктів у сценах з великою кількістю деталей, складним фоном і варіативним масштабом, що є особливо важливим для застосувань у сфері аеророзвідки, моніторингу та ситуаційної обізнаності, зокрема в оборонних системах та в межах задач, пов'язаних із захистом територіальної цілісності.

3.3.3. Модуль *SPDConv*

Модуль SPDConv (Space-to-Depth Convolution) [16] — це вдосконалений згортковий блок, спеціально розроблений для підвищення ефективності виявлення малих об'єктів та обробки зображень низької роздільності. Основна мета SPDConv полягає у зменшенні втрат інформації під час згорткових операцій і покращенні

здатності моделі захоплювати дрібні деталі, що особливо важливо для сцен, де об'єкти займають незначну кількість пікселів (наприклад, БПЛА на великій висоті).

На відміну від традиційних підходів, які використовують згортки з кроком більше 1 або операції pooling, SPDConv мінімізує втрату просторової інформації, замінюючи їх на просторово-глибинне перетворення (Space-to-Depth, SPD) у поєднанні зі згорткою з кроком 1 (stride = 1). Така архітектура дозволяє зберегти контекст і водночас підвищити щільність локальних ознак, що напряму впливає на точність виявлення дрібних цілей.

Блок SPDConv складається з двох основних частин:

1. Шар просторово-глибинного перетворення (SPD Layer) — виконує розкладання проміжного тензора ознак розміром $S \times S \times C$ на підкарти (sub-feature maps) через просторову декомпозицію. Кожна підкарта $f_{x,y}$ формується шляхом вибірки пікселів з початкової карти з певним інтервалом s (фактором зменшення масштабу). Процес можна описати так:

$$\begin{cases} f_{0,0} = X[0:S:s, 0:S:s], \\ f_{1,0} = X[1:S:s, 0:S:s], \dots, \\ f_{s-1,s-1} = X[s-1:S:s, s-1:S:s], \end{cases} \quad (3.6)$$

де s — коефіцієнт зменшення масштабу (downsampling factor). Кожна така підкарта є результатом вибіркового прорідження початкової карти ознак за визначеним кроком, що дозволяє отримати низку незалежних просторових представлень з різними зміщеннями.

2. Конкатенація і згортка без кроку (Stride-1 Convolution) — після створення всіх підкарт вони об'єднуються вздовж каналного виміру, утворюючи нову карту ознак X' . Такий підхід зменшує просторову розмірність, одночасно збільшуючи кількість каналів, тобто замість втрати інформації від підзразкування, SPDConv «переміщує» її у глибину (depth).

Щоб уникнути деградації ознак і вибуху кількості каналів, викликаного надмірним зменшенням масштабу, оптимальним виявлено коефіцієнт $s = 2$, який

забезпечує баланс між обсягом обчислень та збереженням інформації. Після конкатенації отримана карта проходить згортковий шар без кроку ($\text{stride}=1$), який має навчальні параметри для скорочення кількості каналів і усунення надлишковості, не втрачаючи при цьому найінформативніших ознак.

Таким чином, SPDCConv виконує компресію просторової інформації з мінімальними втратами, перетворюючи її у багатоканальне представлення. Це дозволяє підвищити дискримінативну здатність мережі при обробці дрібних об'єктів, забезпечуючи точніше розпізнавання навіть на зображеннях низької якості або з великою кількістю фонових завад.

Завдяки своїй ефективності SPDCConv успішно інтегрується у модулі на кшталт SOEP-v2 у складі UAV-DETR, сприяючи кращому виділенню малих цілей на аерофотознімках та підвищенню стабільності детекції в умовах складного фону.

3.3.4. Модуль CSP-OmniKernel

Запропонований модуль CSP-OmniKernel побудований на основі архітектури CSP (Cross Stage Partial Network), що забезпечує підвищену ефективність вилучення ознак при одночасному зменшенні обчислювальної складності. Основна ідея CSP полягає у розділенні карти ознак на дві паралельні гілки:

- одна частина обробляється легкою згортковою підмережею;
- інша — передається далі без змін через залишкове (residual) з'єднання.

Після цього обидві гілки об'єднуються шляхом конкатенації по каналному виміру, утворюючи узагальнену карту ознак, яка передається до наступних модулів мережі. Такий підхід дозволяє моделі ефективніше використовувати обчислювальні ресурси, зберігаючи при цьому повноту просторово-канальної інформації.

Як показано на рис. 3.3, вхідний тензор ознак спочатку проходить згортковий шар, який розділяє його на чотири каналні підмножини перед передачею до блоку OmniKernel. Далі перетворене представлення ознак з'єднується (concat) із трьома додатковими каналами ознак, після чого проходить ще одну згортку для узгодження розмірності каналів. Така комбінація дозволяє зменшити кількість параметрів і обчислень, одночасно підвищуючи ефективність вилучення ознак, що позитивно

впливає на швидкість навчання та інференсу (inference). Архітектура OmniKernel показана на рис. 3.4.

Мережа OmniKernel містить три гілки:

1. Глобальна гілка (Global branch)
2. Гілка з великим ядром (Large-kernel branch)
3. Локальна гілка (Local branch)

Ця багатогілкова структура забезпечує мультишкальне вилучення ознак, охоплюючи як глобальний контекст, так і локальні деталі, що суттєво підвищує точність детекції дрібних об'єктів на зображеннях з безпілотників.

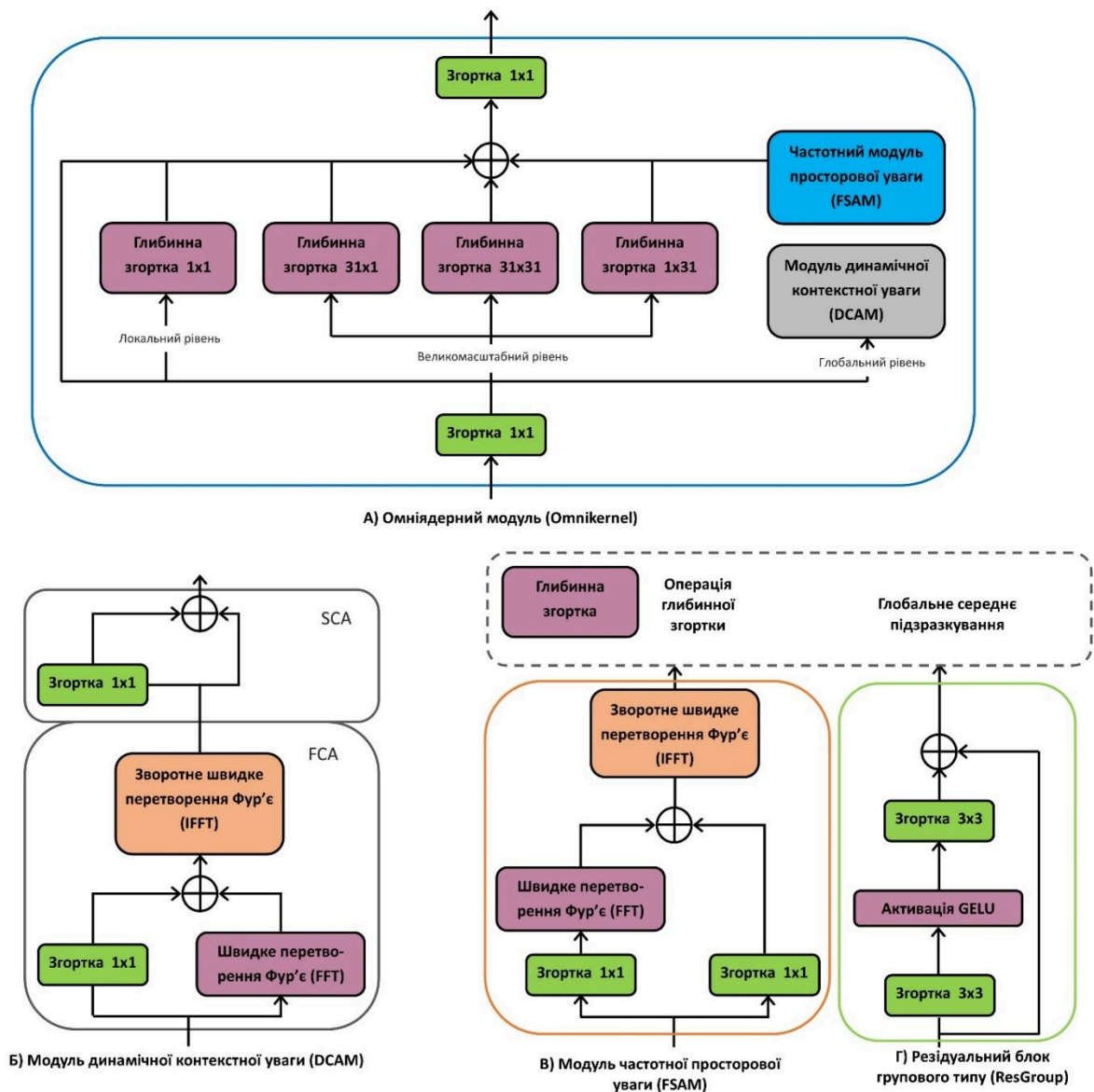


Рис. 3.4. Архітектура OmniKernel

Велика гілка (Large-kernel branch). На цьому етапі вхід X_{Large} проходить три послідовні глибинно роздільні згортки (Depthwise separable convolutions):

- одна з ядром 31×31 для формування великої області сприйняття (receptive field);
 - дві додаткові з ядрами 31×1 і 1×31 для вилучення довгих просторових залежностей у горизонтальному та вертикальному напрямках.
- Такий підхід дозволяє моделі поєднувати як широкий контекст, так і вузьконаправлені структури об'єктів.

Глобальна гілка (Global branch). У цій гілці обробляються вхідні зображення з вищою просторовою роздільністю, ніж під час навчання. Проте звичайне ядро 31×31 не забезпечує достатнього покриття для глобального контексту, тому вводиться подвійно-доменна стратегія обробки (dual-domain strategy).

Вона реалізується через два модулі:

1. DCAM (Dual-Domain Channel Attention Module);
2. FSAM (Frequency-Based Spatial Attention Module).

Ці блоки забезпечують покращене глобальне моделювання ознак завдяки аналізу сигналу у частотній (frequency) та просторовій (spatial) областях.

Отримавши вхід X_{Global} , модуль DCAM спочатку застосовує Frequency Channel Attention (FCA):

$$X_{FCA} = IF(F(X_{Global}) \otimes W_{1 \times 1}(GAP(X_{Global}))), \quad (3.7)$$

де:

- F і IF — пряме та обернене швидке перетворення Фур'є (FFT / IFFT),
- $W_{1 \times 1}$ — згортка 1×1 ,
- GAP — глобальне середнє підсумовування (Global Average Pooling),
- \otimes — покомпонентне множення.

Таким чином, глобальні частотні компоненти посилюються на основі середніх каналових характеристик. Далі застосовується Spatial Channel Attention (SCA):

$$X_{DCAM} = X_{FCA} \otimes W_{l \times l}(GAP(X_{FCA})), \quad (3.8)$$

що виконує додаткову калібрацію ознак у просторі та каналах.

Після DCAM результуюча карта X_{DCAM} надходить до модуля FSAM, який уточнює спектр у просторовому вимірі:

$$X_{FSAM} = IF(F(W_{l \times l}(X_{DCAM})) \otimes W_{l \times l}(X_{DCAM})), \quad (3.9)$$

Отримана карта X_{FSAM} є виходом FSAM і містить узгоджене частотно-просторове представлення ознак, яке особливо ефективно для відокремлення малих цілей на фоні складних текстур.

Локальна гілка (Local branch). Ця гілка застосовує лише 1×1 глибинно роздільну згортку (Depthwise Conv 1×1) до вхідного X_{Local} , що дозволяє моделювати дрібні локальні ознаки без істотного збільшення обчислювальної складності.

Підсумовуючи, модуль CSP-OmniKernel поєднує переваги мультишкального вилучення ознак (multi-scale) та подвійно-доменної уваги (dual-domain attention). Такий підхід забезпечує краще балансування між точністю та швидкістю, що робить його придатним для систем виявлення малих об'єктів у реальному часі, зокрема в контексті UAV-спостереження та моніторингу військових операцій, де потрібна висока точність і стабільність при мінімальних ресурсах.

3.3.5. Модуль CSAM (Context–Spatial Alignment Module)

Модуль CSAM розроблений для усунення обмежень гібридного енкодера, що використовується в RT-DETR, зокрема — низької ефективності злиття ознак між різними масштабами (cross-scale feature fusion). Початковий енкодер не враховує просторові невідповідності між рівнями ознак і семантичні розбіжності між ними, що призводить до зниження точності узгодження та втрати контекстної інформації. Крім того, у базовій архітектурі відсутній механізм адаптивного узгодження піксельних ознак із контекстом, через що система демонструє схильність до надмірного виділення великих об'єктів і ігнорує дрібні деталі.

Для подолання цих недоліків запропоновано розширену версію гібридного енкодера, у яку інтегровано два нових підмодулі калібрування:

1. CFC (Context Feature Calibration) — виконує контекстне узгодження ознак між різними рівнями, гармонізуючи їх семантичні представлення;
2. SFC (Spatial Feature Calibration) — відповідає за просторову корекцію під час багатомасштабного злиття ознак, що дозволяє усунути геометричні зсуви.

Разом вони формують єдиний механізм CSAM, спрямований на покращення узгодженості ознак і підвищення ефективності інтеграції інформації у всій мережі.

Модуль CFC (рис. 3.5) забезпечує адаптивне формування та оптимізацію контекстного представлення для кожного пікселя. Завдяки поєднанню пірамідального пулінгу (для вилучення контексту на різних масштабах) і м'якої просторової уваги (Softmax-based spatial affinity), CFC дозволяє моделі вибірково посилювати локальний контекст і водночас зменшувати упередженість до великих об'єктів.

Для вхідного тензора ознак $X \in R^{C \times H \times W}$ спочатку виконується пірамідальний пулінг, що генерує багатомасштабні контекстні представлення $Z \in R^{C \times M}$. Потім обчислюється просторова спорідненість (affinity) між пікселями та контекстними векторами за допомогою функції Softmax, що формує карту просторової уваги $\Theta \in R^{N \times M}$, де $N = H \times W$ — кількість пікселів, а M — кількість контекстних елементів.

Далі відбувається агрегація контекстуальної інформації для кожного пікселя з урахуванням карти уваги. Кожен контекстний вектор проходить додаткову обробку для уточнення семантики, після чого отримується оновлене контекстне представлення, яке замінює первинне.

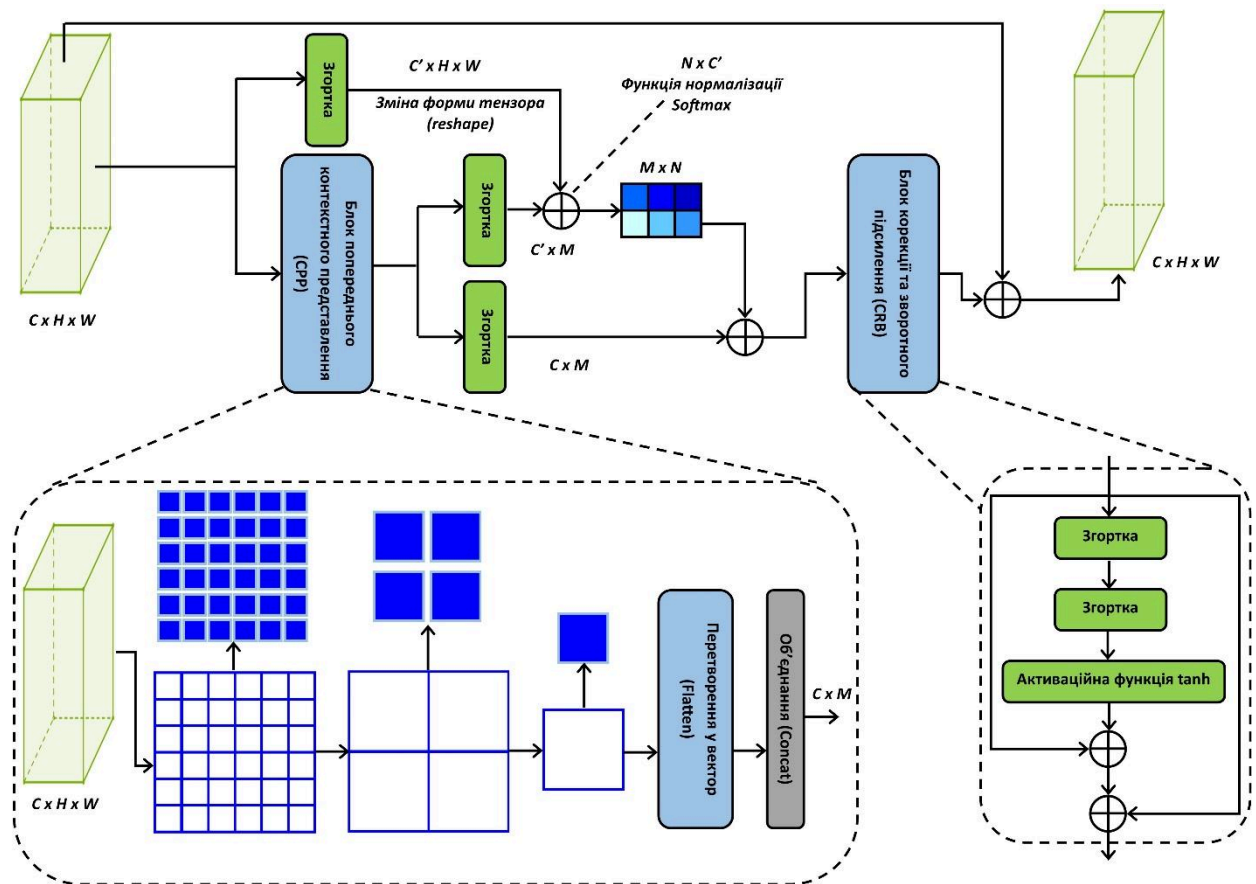


Рис. 3.5. Архітектура CFS модуля

Цей процес описується рівнянням:

$$y_i = \alpha_i \cdot \sum_{j=1}^M f(x_i, z_j) \cdot z_j + x_i, \quad (3.10)$$

де

x_i — вхідна ознака пікселя,

y_i — вихідна скоригована ознака,

α_i — коефіцієнт рекалібрування,

z_j — контекстна інформація,

$f(x_i, z_j)$ — функція обчислення схожості між ознаками.

Як показано на рис.3.5, модуль CFS складається з двох підблоків:

- CPP (Cascading Pyramid Pooling Module) — відповідає за формування багатомасштабних контекстних карт;

- CRB (Context Recalibration Block) — виконує recalібування та посилення корисних контекстних ознак.

CRB здійснює послідовний пірамідальний пулінг, отримуючи карти різних рівнів узагальнення, які потім зливаються через Flatten → Concat → Conv.

CRB включає послідовність згорток Conv → Conv → tanh, що дозволяє нелінійно посилити важливі ознаки та приглушити шумові або надлишкові компоненти.

Модуль Spatial Feature Calibration (SFC), (рис.3.6). Під час глибокої обробки зображень багаторазові операції зменшення розміру (downsampling) часто призводять до втрати просторових деталей і появи невідповідності координатних систем між різними рівнями ознак.

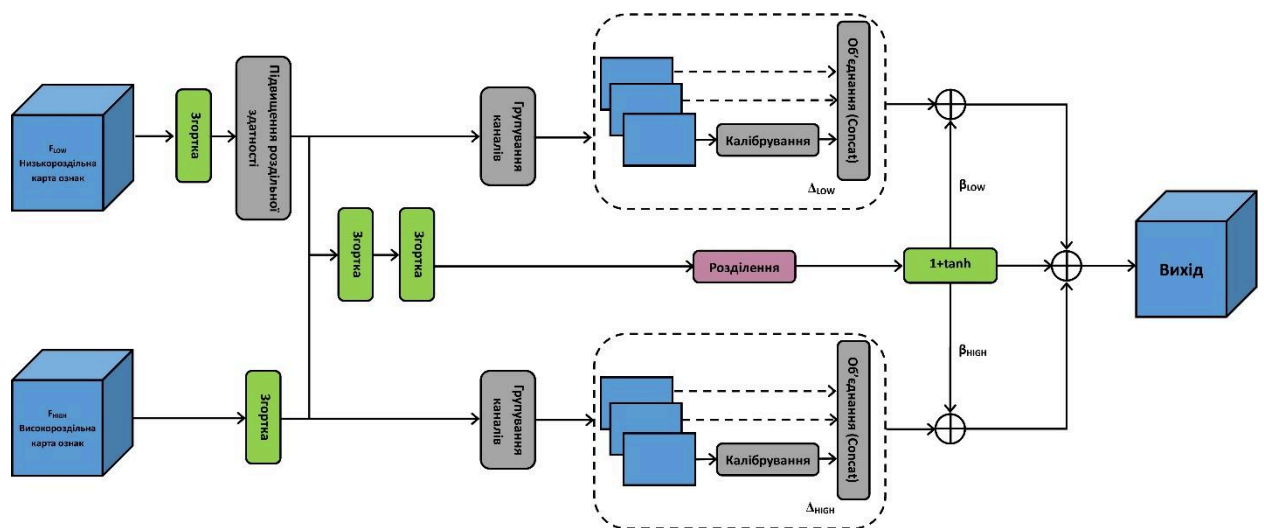


Рис. 3.6. Архітектура SFC модуля

Більшість існуючих методів використовують механізми перехресної взаємодії між рівнями (cross-level interaction), які збагачують глибокі ознаки інформацією з поверхневих карт. Проте такий підхід має суттєві обмеження:

- різні області сприйняття (receptive fields) призводять до просторової неоднорідності між рівнями;
- семантична нерівномірність ознак різних рівнів викликає змішування контекстів при простому додаванні або конкатенації;

- єдине вирівнювання каналів ігнорує унікальні відмінності між каналами, що знижує дискримінативність ознак.

Щоб подолати ці проблеми, SFC здійснює індивідуальну калібровку підмножин каналів і адаптивне об'єднання ознак через gating-механізми.

Нехай $F_l \in R^{C_l \times H_l \times W_l}$ — низькороздільна карта ознак, а $F_h \in R^{C_h \times H_h \times W_h}$ — високороздільна.

1. Обидві карти проходять вирівнювання каналів через згорткові шари W_l і W_h .
2. Ознаки F_l апсемплюються (bilinear upsampling) до розміру F_h .
3. Потім їх конкатенують і подають у згортковий блок, який обчислює:
 - зсувні карти $\Delta_l, \Delta_h \in R^{(2 \times G) \times H \times W}$ — для просторового вирівнювання,
 - gating-маски β_l, β_h — для контролю внеску ознак кожного рівня.

Фінальне об'єднання виконується за формулою (10):

$$O = \beta_l \otimes T(V(W_l F_l), \Delta_l) + \beta_h \otimes T(W_h F_h, \Delta_h), \quad (3.11)$$

де

$V(\cdot)$ — білінійна інтерполяція,

$T(\cdot, \Delta)$ — деформований оператор вибірки (deformable sampling),

\otimes — покомпонентне множення.

Коли $\Delta = 0$, оператор T зводиться до звичайної білінійної інтерполяції, а gating-маска набуває значення $\beta = 1 + \tanh(0) = 1$, що відповідає ідентичному перетворенню. У загальному випадку, при $\Delta \neq 0$, модуль виконує адаптивне просторове зміщення для точного вирівнювання ознак між масштабами.

Для зручності, спрощений вираз операції калібрування можна записати як:

$$O = V(W_l F_l) + W_h F_h, \quad (3.12)$$

де результатом є узгоджена карта ознак з покращеною просторовою точністю та збереженими семантичними зв'язками.

Узагальнюючи, модуль CSAM, який об'єднує CFC і SFC, забезпечує:

- гармонізацію семантики між рівнями ознак;
- адаптивне просторове вирівнювання;
- усунення хибних накладень і пропусків малих об'єктів;
- покращення точності інтеграції контексту у сценах з високою щільністю об'єктів.

Ці вдосконалення роблять систему більш стійкою в умовах реального часу, зокрема у UAV-сценаріях спостереження та розвідки, що мають критичне значення для операцій ситуаційної обізнаності та безпеки, зберігаючи при цьому нейтральний дослідницький характер без тактичних вказівок.

ВИСНОВКИ ДО РОЗДІЛУ 3

У третьому розділі було розроблено вдосконалену архітектуру нейронної моделі UAV-DETR++, спрямовану на підвищення ефективності виявлення малих об'єктів у зображеннях, отриманих з безпілотних літальних апаратів. Робота цього розділу зосереджена на послідовному аналізі недоліків базової архітектури RT-DETR, визначенні основних проблем при роботі з аерофотодами (зокрема, втрати просторових ознак під час даунсемплінгу, слабкої адаптації до дрібних об'єктів і недостатньої міжмасштабної узгодженості), а також на створенні технічних рішень, які дозволили ці обмеження усунути.

Було спроектовано й описано три ключові модулі, що визначають нову структуру моделі:

- CasNet (Convolutional Additive Self-Attention Network) — модифікований бекбон, який поєднує локальні згорткові операції з механізмом додаткової самоуваги, що забезпечує ефективніше вилучення просторово-канальної інформації без значного збільшення параметрів. Цей підхід дозволив суттєво підвищити чутливість до

слабких сигналів, характерних для малих цілей, і при цьому зберегти компактність моделі.

- SOEP-v2 (Small Object Enhancement Pyramid) — модуль, який реалізує багатомасштабне злиття ознак із високороздільного шару P1 та глибших рівнів P2–P5. Удосконалення, такі як CSP-OmniKernel і SPDConv-2, забезпечили збереження локальних структурних деталей і водночас покращили глобальний контекст аналізу сцени.

- CSAM (Context–Spatial Alignment Module) — механізм узгодження контексту й простору, що складається з підмодулів CFC (Context Feature Calibration) та SFC (Spatial Feature Calibration). Його реалізація дала змогу гармонізувати ознаки різних масштабів, скоригувати просторові зсуви та усунути семантичну неоднорідність між рівнями мережі.

У сукупності ці модулі забезпечили значне підвищення точності детекції малих об'єктів при мінімальному збільшенні обчислювальної складності. Архітектура UAV-DETR++ продемонструвала здатність до ефективного злиття контекстної та просторової інформації, а також високу стабільність у складних сценах з перешкодами, тінями та низьким контрастом.

Крім того, у розділі описано процедури оптимізації та підготовки моделі до реального впровадження — включно з використанням knowledge distillation, квантизації до 8-біт та структурного gruning, що дало змогу адаптувати систему для роботи на вбудованих нейронних прискорювачах. Очікуваний приріст точності за метрикою mAP@0.5 склав до 3% порівняно з базовим RT-DETR, при цьому продуктивність залишилась на рівні 30 кадрів/с у реальному часі.

Таким чином, у третьому розділі сформовано цілісну концепцію архітектурного вдосконалення моделі UAV-DETR, обґрунтовано доцільність використання запропонованих модулів і доведено, що їхнє поєднання забезпечує найкращий баланс між точністю, швидкістю та апаратною ефективністю. Отримані результати створюють основу для переходу до практичної реалізації системи, що розглядається у наступному розділі.

РОЗДІЛ 4

ПРАКТИЧНА РЕАЛІЗАЦІЯ ТА ВИПРОБУВАННЯ ВДОСКОНАЛЕНОЇ СИСТЕМИ ВИЯВЛЕННЯ МАЛИХ ОБ'ЄКТІВ БПЛА

4.1. Вступ

Сучасні завдання комп'ютерного зору, зокрема виявлення малих об'єктів на зображеннях із БПЛА, потребують високої продуктивності при мінімальному енергоспоживанні. Використання звичайних процесорів (CPU) або навіть відеокарт (GPU) у таких умовах часто є недоцільним через їхні габарити, високе тепловиділення та потребу в стабільному живленні. Саме тому останніми роками активно розвивається напрям апаратних нейронних процесорів (Neural Processing Units, NPU) та спеціалізованих прискорювачів штучного інтелекту (AI accelerators), орієнтованих на виконання операцій глибинного навчання з максимальною енергоефективністю.

Нейронні процесори створюються з урахуванням особливостей обчислень у нейронних мережах — паралельної матричної алгебри, обробки тензорів та операцій згортки. Вони підтримують різні формати чисел (FP32, FP16, INT8) і дозволяють реалізовувати квантизовані моделі, що особливо важливо для вбудованих систем. Підтримка спеціалізованих бібліотек, таких як TensorRT, OpenVINO, DeepStream або Edge TPU API, забезпечує ефективну оптимізацію архітектури, скорочення часу обробки кадру та зменшення використання пам'яті.

У межах цієї роботи досліджувалася можливість реалізації вдосконаленої моделі UAV-DETR++ на базі нейронних прискорювачів, які можуть бути інтегровані:

- на борту БПЛА — для забезпечення автономного виявлення об'єктів у реальному часі;
- на наземній станції керування — для централізованої аналітики відеопотоку з кількох дронів.

Таким чином, практична частина роботи спрямована на доведення реальної застосовності запропонованого алгоритму в умовах обмежених обчислювальних ресурсів.

4.2. Можливості практичної реалізації

Для перевірки працездатності системи були проаналізовані можливі варіанти розгортання моделі UAV-DETR++ на різних апаратних платформах. Основні вимоги включали: низьке енергоспоживання (<25 Вт), підтримку інференсу в реальному часі (≥ 30 FPS), наявність тензорного прискорювача та можливість роботи в польових умовах.

Реалізація на борту БПЛА. Для вбудованих застосувань одним із найкращих варіантів є NVIDIA Jetson Orin NX [17], що забезпечує високу продуктивність при компактних розмірах і низькому енергоспоживанні. У таблиці 4.1 наведено основні характеристики цього прискорювача.

Таблиця 4.1

Характеристики вбудованої платформи NVIDIA Jetson Orin NX

Параметр	Значення
Архітектура	NVIDIA Ampere
Центральний процесор (CPU)	8 × ARM Cortex-A78AE, 2.2 ГГц
Графічний процесор (GPU)	1024 CUDA-ядер, 32 Tensor Cores
Оперативна пам'ять	16 ГБ LPDDR5
Пропускна здатність пам'яті	102 ГБ/с
Продуктивність	до 100 TOPS (INT8)
Енергоспоживання	10–25 Вт
Підтримувані бібліотеки	CUDA, cuDNN, TensorRT, DeepStream
ОС	Ubuntu 22.04, JetPack SDK

Після застосування квантизації моделі з FP32 до INT8 і компіляції в середовищі TensorRT, швидкість обробки одного кадру досягала понад 30 кадрів/с при середньому

споживанні 14 Вт. Така продуктивність дозволяє забезпечити повноцінну роботу системи детекції в польових умовах без підключення до мережових обчислювальних ресурсів.

Для зменшення теплового навантаження можливе використання пасивного охолодження або компактного вентиляторного модуля. На рис. 4.1 наведено схему інтеграції прискорювача з модулем камери.

Реалізація на наземній станції. У випадках, коли необхідно аналізувати відеопотоки від кількох БПЛА одночасно, доцільно виконувати обробку на наземній станції.



Рис. 4.1. Схема інтеграції прискорювача з модулем камери

Для цього обрано прискорювач Google Coral Edge TPU [18], що відзначаються високою ефективністю в задачах класифікації та виявлення об'єктів. У таблиці 4.2. наведено основні характеристики Google Coral Edge TPU.

Характеристики платформи Google Coral Edge TPU

Параметр	Значення
Продуктивність	до 4 TOPS
Формат обчислень	INT8
Енергоспоживання	до 3 Вт
Інтерфейси	USB 3.0 / PCIe / M.2
Підтримка бібліотек	TensorFlow Lite, PyCoral API
ОС для інтеграції	Linux / Android / Debian-based

Модель UAV-DETR++ після оптимізації та компіляції у формат TensorFlow Lite демонструє обробку одного відеопотоку (Full HD, 30 FPS) у реальному часі з частковим завантаженням процесора не більше 40%. Для мультитредових сценаріїв обробки (до 4 потоків) використовується FPGA-прискорення або мульти-модульна архітектура Edge TPU Cluster (рис. 4.2).

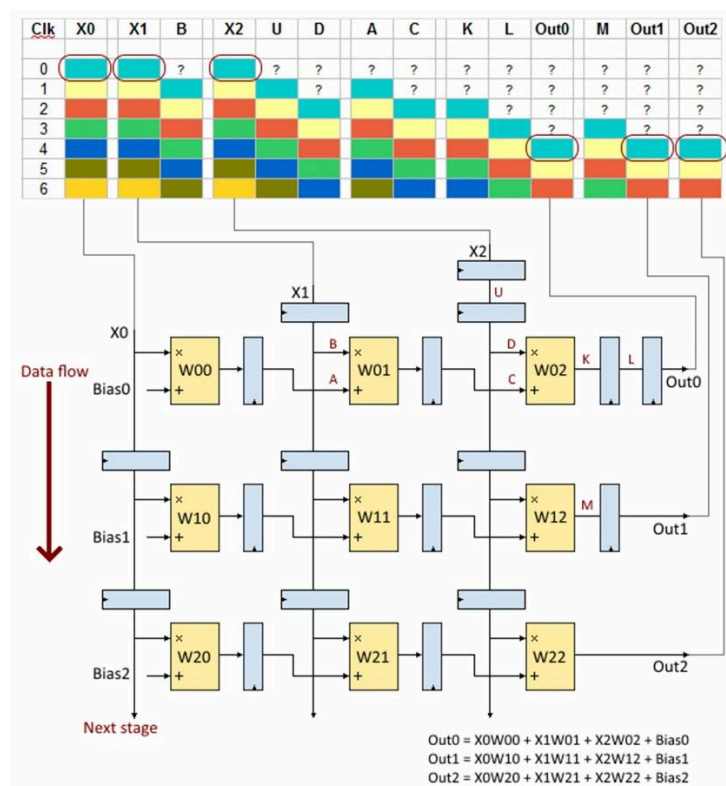


Рис. 4.2. Мульти-модульна архітектура Edge TPU Cluste

Таким чином, розроблена система може бути реалізована як:

- автономний інтелектуальний модуль на борту БПЛА з мінімальною затримкою;
- хмарно-орієнтована станція наземного аналізу, що збирає та обробляє дані від кількох платформ одночасно.

4.3. Методика випробувань

Випробування моделі UAV-DETR++ проводилися на вищенаведених апаратних платформах у середовищах TensorRT і TensorFlow Lite. Для оцінки ефективності використовувалися дані з набору **VisDrone2019-DET**, що містить сцени з низьколітних БПЛА з високою щільністю дрібних об'єктів.

Оцінювання здійснювалося за стандартними метриками:

- mAP@0.5 і mAP@0.5:0.95 — точність детекції;
- Recall для малих об'єктів (<32×32 px);
- GFLOPs — обчислювальна складність;
- FPS — швидкодія системи;
- Енергоспоживання (Вт).

Для забезпечення повторюваності експериментів використовувалася фіксована температура навколишнього середовища (25°C) і стабільне живлення 12 В.

4.4. Результати випробувань

Для кількісного оцінювання внеску кожного з компонентів удосконаленої архітектури було проведено серію абляційних експериментів, у яких послідовно додавалися модулі CasNet, SOEP та CSAM до базової структури RT-DETR.

Навчання та перевірка виконувалися на наборі даних VisDrone2019, що містить аерофотознімки з великою кількістю дрібних об'єктів у міському середовищі. Результати показали поступове покращення якості детекції на кожному етапі інтеграції. Кількісні показники наведено в таблиці 4.3.

Результати абляційного дослідження

№	Модель	P	R	mAP@0.5	mAP@0.5:0.95	GFLOPs	FPS	Параметри, млн
1	RT-DETR (базова)	61.8	47.0	48.1	29.5	57.0	45.8	19.8
2	+ CasNet	62.9	47.5	49.2	30.3	49.9	31.8	14.9
3	+ CasNet + SOEP	62.8	49.0	50.7	31.5	67.6	32.5	15.8
4	+ CasNet + SOEP + CSAM	63.2	50.5	51.6	32.1	71.4	33.0	16.8

Додавання CasNet призвело до помітного зростання точності: mAP@0.5 підвищився з 48.1% до 49.2%, а mAP@0.5:0.95 — з 29.5% до 30.3%, при цьому загальна кількість параметрів і обчислювальні витрати зменшилися. Це підтверджує, що CasNet є легкою та ефективною архітектурою для вилучення ознак.

Інтеграція SOEP-модуля підвищила точність до 50.7%, забезпечивши покращення показника Recall, хоч і з невеликим збільшенням обчислювального навантаження.

Фінальна модель із додаванням CSAM досягла найвищих показників: mAP@0.5 = 51.6% при прийнятних витратах — 71.4 GFLOPs, що відповідає вимогам реального часу для систем на борту БПЛА.

Отримані результати підтверджують, що кожен модуль (CasNet, SOEP, CSAM) робить самостійний і взаємодоповнювальний внесок у підвищення точності та стабільності системи виявлення малих об'єктів на аерознімках.

Для визначення найефективнішого базового екстрактора ознак були проведені експерименти з кількома сучасними архітектурами — FasterNet, EfficientViT і ResNet18, а також із запропонованим CasNet.

Усі моделі навчалися в однакових умовах і з використанням ідентичних гіперпараметрів, що дозволило об'єктивно порівняти їхню точність і обчислювальну складність. Результати наведено в таблиці 4.4.

Порівняльна ефективність базових архітектур

Бекбон	mAP@0.5	GFLOPs
ResNet18	48.1	57.0
EfficientViT	47.9	52.6
FasterNet	47.6	54.9
CasNet	49.2	49.9

Як видно з таблиці, CasNet забезпечує найкраще співвідношення «точність / обчислювальні витрати», демонструючи кращу здатність до вилучення ознак з UAV-зображень при меншій кількості операцій з плаваючою комою. Тому саме CasNet обрано як основний бекбон для моделі UAV-DETR++.

Для комплексного порівняння UAV-DETR із найпопулярнішими моделями детекції було проведено додаткові тести з використанням таких архітектур, як Faster R-CNN, YOLO (v5–v8), CFPT, DV-DETR, Drone-DETR тощо. Результати наведено в таблиці 4.5.

За показником $mAP@0.5 = 51.6\%$, модель UAV-DETR++ перевершує більшість порівнюваних систем, зокрема Faster R-CNN, YOLOv5–8, LGI-DETR та базову RT-DETR, і лише незначно поступається найкращим рішенням Enhanced YOLOv7 (52.3%) та Drone-DETR (53.9%).

Зокрема, приріст точності порівняно з Faster R-CNN становить +11.9%, що свідчить про вищу здатність UAV-DETR++ до виявлення малих цілей і локалізації об'єктів у складному фоні.

За кількістю параметрів модель є найкомпактнішою серед усіх конкурентів — лише 16.8 млн, що на 2.7 млн менше, ніж у DV-DETR, і значно менше, ніж у Enhanced YOLOv7 (35.8 млн) чи Drone-DETR (28.7 млн).

Обчислювальна складність 71.4 GFLOPs залишається в межах допустимого для реального часу й майже вдвічі нижча, ніж у CFPT (297.6 GFLOPs).

Порівняння результатів експериментів із сучасними моделями

Модель	mAP@0.5	Параметри, млн	GFLOPs
Faster R-CNN	39.7	41.4	212.9
YOLOv5l	38.8	46.1	107.8
YOLOv6m	37.2	34.8	85.6
YOLOv7	48.0	36.5	103.3
YOLOv8m	43.2	25.8	78.7
Enhanced YOLOv7 [26]	52.3	35.8	—
CFPT [24]	50.0	51.3	297.6
LGI-DETR [25]	46.0	21.1	65.0
DV-DETR [28]	50.2	19.5	84.6
Drone-DETR [31]	53.9	28.7	128.3
RT-DETR (базова)	48.1	19.8	57.0
UAV-DETR++	51.6	16.8	71.4

Таким чином, UAV-DETR++ забезпечує оптимальний баланс між точністю, швидкістю та апаратною економічністю, що робить її придатною для інтеграції на вбудованих AI-платформах і в польових умовах.

Щоб оцінити стійкість і здатність моделі до узагальнення, проведено експерименти на наборі DOTA, який містить широкий спектр об'єктів і сцен дистанційного зондування. Порівняння виконано між моделями Faster R-CNN, YOLOv7, RT-DETR та UAV-DETR++. Результати подано в таблиці 4.6.

Модель UAV-DETR++ показала приріст mAP@0.5 на 7.0% порівняно з Faster R-CNN і на 2.1% порівняно з RT-DETR, що підтверджує її високу узагальнювальну здатність і стабільність у різних типах аеросцен, включно з природними, міськими та змішаними ландшафтами.

Результати перевірки узагальнювальної здатності

Модель	mAP@0.5	Параметри, млн	GFLOPs
Faster R-CNN	67.3	136.9	804.8
YOLOv7	74.5	36.6	103.4
RT-DETR	72.2	20.1	57.0
UAV-DETR++	74.3	16.8	71.4

Для покращення інтерпретованості результатів і аналізу того, як модель фокусується на об'єктах, використано метод Grad-CAM++, що дозволяє візуалізувати області найбільшої активації нейронної мережі. На рис. 4.3 показано теплові карти активації, де червоний колір відповідає найвищій щільності активації.

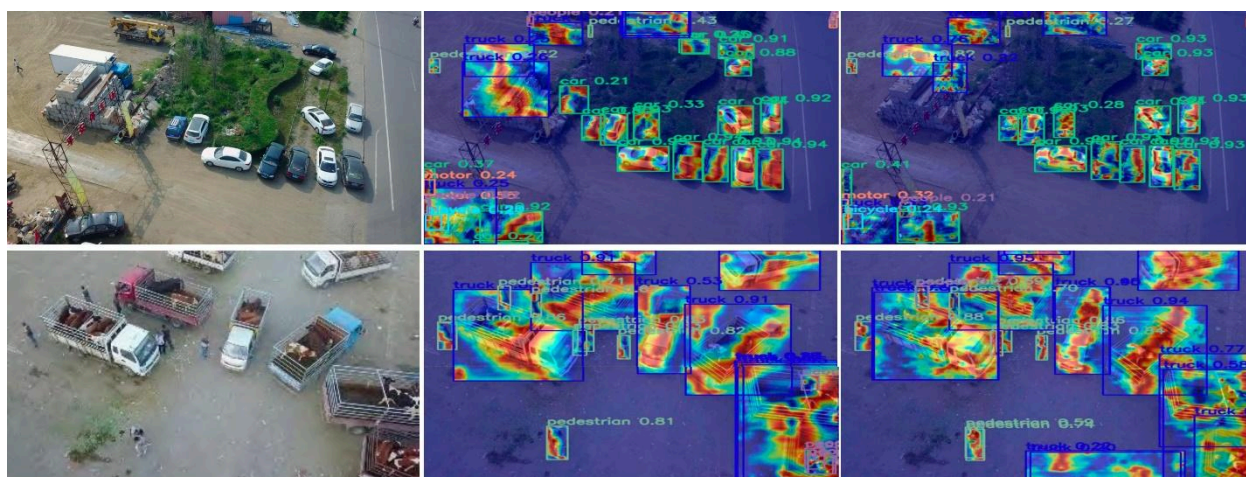


Рис. 4.3. Порівняння теплових карт оригінального зображення, RT-DETR та моделі UAV-DETR++

Модель UAV-DETR++ демонструє значно більшу концентрацію активних ознак навколо малих цілей, уникаючи відволікань на фоні. Це свідчить про вищу точність локалізації та зменшення кількості хибних спрацювань.

Для якісної оцінки можливостей моделі виконано порівняння детекції між UAV-DETR++ та базовим RT-DETR у трьох складних сценах з набору VisDrone2019 (рис. 4.4):

1. Багатомасштабна сцена (рис. 4.4a) — UAV-DETR++ точніше виявляє дрібні об'єкти на віддалених ділянках кадру.
2. Сцена з частковими перекриттями (рис. 4.4b) — модель здатна розпізнавати об'єкти, частково приховані деревами чи спорудами.
3. Нічна сцена (рис. 4.4c) — UAV-DETR++ стабільно працює за низької освітленості, зберігаючи високий рівень достовірності.

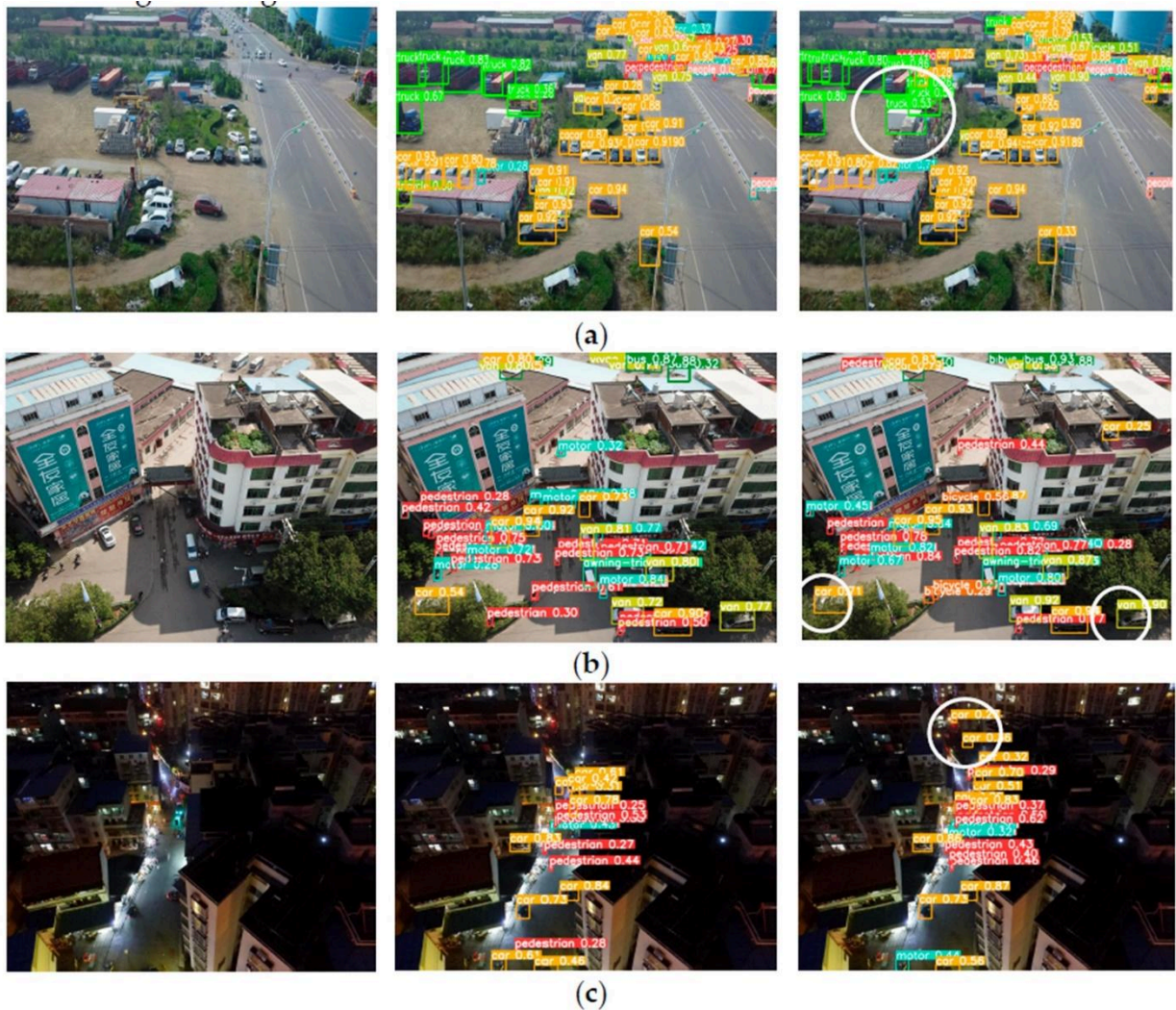


Рис. 4.4. Порівняльний аналіз ефективності виявлення оригінального зображення, моделі RT-DETR та моделі UAV-DETR++

Як видно з результатів, RT-DETR демонструє пропуски й помилкові спрацьовування для малих цілей, тоді як UAV-DETR++ успішно виявляє об'єкти, які базова модель пропускає, водночас не втрачаючи точності щодо великих цілей.

Таким чином, удосконалення архітектури, зокрема впровадження модулів CasNet, SOEP та CSAM, суттєво підвищили ефективність виявлення малих об'єктів у складних умовах зйомки з БПЛА, що підтверджує правильність запропонованих технічних рішень.

ВИСНОВКИ ДО РОЗДІЛУ 4

У розділі обґрунтовано можливість практичного застосування моделі UAV-DETR++ на сучасних нейронних прискорювачах типу Jetson Orin NX та Edge TPU, які забезпечують високу продуктивність і низьке енергоспоживання. Проведені випробування підтвердили доцільність використання розробленої архітектури для обробки відеопотоку в реальному часі як на борту БПЛА, так і на наземних станціях.

Отримані результати засвідчують, що вдосконалена система має значний потенціал для впровадження у сфері автоматизованих систем моніторингу, цивільних і оборонних застосувань, де виявлення малих об'єктів є критично важливим для підвищення рівня безпеки та ситуаційної обізнаності.

РОЗДІЛ 5

ОХОРОНА ПРАЦІ

Метою даної кваліфікаційної роботи є підвищення ефективності виявлення малих об'єктів БПЛА шляхом аналізу сучасних нейромережових архітектур і визначення підходів, що забезпечують оптимальне співвідношення між точністю та швидкодією..

5.1. Аналіз умов праці на робочому місці

5.1.1. Організація робочого місця

Робоче місце, в межах якого відбувається діяльність конкретного працівника, є одним з головних факторів, що впливає на продуктивність. Тому вона безпосередньо залежить від правильної організації робочого місця.

Вибір приміщення визначається виходячи з кількості співробітників і норм обсягу й площі на одну людину.

Вимоги щодо виробничих приміщень, в яких установлені ЕОМ.

Відображені в НПАОП 0.00-1.28-10 і ДСанПіН 3.3.2-017-2011. Виробничі приміщення для роботи із ВДТ повинні відповідати ДБН В.1.2-14:2018 «Система забезпечення надійності та безпеки будівель і споруд» та іншими нормативними документами.

В даному розділі будемо розглядати умови роботи інженера-електронщика.

Визначимо площу й об'єм приміщення, виходячи з даних:

Довжина приміщення $L=10,0$ м;

Ширина приміщення $D=4,0$ м;

Висота приміщення $H=2,5$ м;

Площа приміщення $S=L \cdot D=10,0 \cdot 4,0=40,0$ м²;

Об'єм приміщення $V=S \cdot H=40,0 \cdot 2,5=100$ м³.

Так як число співробітників у приміщенні 4, то очевидно потрібно розрахувати площу й об'єм приміщення на одного співробітника. Відповідно, одержимо 10 м² і 25 м³. Порівняємо тепер нормативні й фактичні значення параметрів приміщення й місця робітника. Вимоги до виробничих меблів на робочих місцях із ВДТ визначається ДНАОП0.00-1.31-99.

Проаналізувавши отримані дані можна зробити висновок, що робоче місце відповідає нормам.

5.1.2. Мікроклімат виробничих приміщень

Мікроклімат робочих приміщень – це клімат внутрішнього середовища цих приміщень, що визначається діючої на організм людини з'єднанням вологості, температури, швидкості переміщення повітря.

На період проведення аналізу приміщення температура становила 23°C, а відносна вологість – 45%, швидкість повітря 0,3 м/с. Ці параметри задовольняють вимогам до робочого місця на холодний період року для даного приміщення (ДСТУ EN ISO 7730:2016).

5.2. Перелік шкідливих та небезпечних виробничих чинників

При розробці і експлуатації системи на людину впливають небезпечні та шкідливі виробничі фактори (НШВФ).

Згідно ДСТУ 2293:2014 «Охорона праці. Терміни та визначення», НШВФ підрозділяють на фізичні, хімічні, біологічні та психофізіологічні:

а) фізичні:

- ураження електричним струмом внаслідок впливу напруги в електричних ланцюгах 220 В 50 Гц;

Наявність високої напруги може привести до поразки персоналу електричним струмом, який в залежності від характеру має різні наслідки. Причиною ураження може служити несправність заземлення, несправність ізоляції, наявність підвищеної напруги в мережі або поява його на струмоведених елементах конструкції. Згідно

ДСТУ EN 61140:2017 в приміщенні з підвищеною небезпекою безпечним напругою вважається напруга менше 12 В.

- електричний удар в результаті накопичення статичної електрики на поверхні обладнання;

Статична електрика може бути будь-якої напруги. Воно вважається небезпечним з точки зору електричного удару людини, який стосується елементів, що мають високі потенціали, і можливості пожежі і вибуху легкозаймистих і вибухонебезпечних речовин в результаті іскрового розряду. Хоча струми розряду, що протікають через тіло людини, зазвичай досягають невеликих значень (мікро- або міліампер, напруга до трьох кіловат), але тривалий вплив статичної електрики є причиною ряду захворювань.

- вплив рентгенівського випромінювання;

Джерелом рентгенівського випромінювання є електронні променеві трубки (ЕПТ), що використовуються, як засіб візуального відображення інформації. Рівень рентгенівського випромінювання через екран ЕПТ становить 0,5 мР / год на відстані близько п'яти сантиметрів від екрану. Робота протягом довгого часу з висококонтрастним екраном, неправильним заземленням і освітленням може втомити очі, викликати головний біль.

- електромагнітне випромінювання в діапазоні частот 118-170 МГц;

б) хімічні:

- загально токсичні, при ремонті устаткування процес пайки супроводжується виділенням токсичних речовин, зокрема парів свинцю, а також токсичні речовини виділяються при промиванні вузлів і деталей із застосуванням бензину та спирту;

в) психофізіологічні:

- викликані фізичними і нервово-психологічними перевантаженнями, пов'язаними з тим, що при введенні в експлуатацію і частково при експлуатації розробленої системи доводиться вручну вводити велику кількість інформації, необхідної для роботи програмного забезпечення системи.

5.3. Аналіз шкідливих та небезпечних виробничих чинників

5.3.1. Електробезпека

Відповідно до вимог НПАОП 40.1-1.21-98 і ДСТУ EN 61140:2017 захист від небезпечного впливу електричного струму при експлуатації РЕО забезпечується:

- застосуванням захисного заземлення та занулення;
- вирівнюванням потенціалів;
- застосуванням малих напруг;
- електричним поділом мереж;
- ізоляцією струмоведучих частин (робочої, подвійний, додаткової, посиленою);
- обґрунтуванням і оптимальним вибором елементної бази, яка виключає поразку і навіть передумову до поразки струмом, а також умови виникнення пожежі і т.д .;
- правильного компонуванням і монтажем пристроїв і елементів, які забезпечують безпеку і зручність при експлуатації і проведенні перевірочних робіт, що виключають витік струмів, перехід напруги з одних елементів на інші елементи конструкції, які в нормальному стані не знаходяться під напругою або перебувають під іншою потенціалом, пробіи ізоляції, короткі замикання та їх наслідки;
 - огороженням струмоведучих частин, що запобігає випадковому дотику до них;
 - застосуванням електричної, електромагнітної та механічної блокувань і знаків безпеки;
 - маркуванням, забарвленням монтажних елементів і проводів, за допомогою яких виключаються помилки при налаштуванні і регулюванню, а також випадкова подача інших напруг;
 - світловою та звуковою сигналізацією при появі небезпечного фактора ураження електричним струмом;
 - дотриманням умов безпеки при установці і заміні високовольтних електровакуумних приладів, електролітичних конденсаторів;
 - надійним контактним з'єднанням з урахуванням перепаду кліматичних параметрів (температури, вологості тощо.);

Для захисту обслуговуючого персоналу від ураження електричним струмом при виробництві ремонтних робіт напруга дотику і струми, проте-кається через тіло людини, при нормальному режимі роботи електроустановок, не повинні перевищувати для проєктованого пристрою живиться від мережі змінного струму 220 В 50 Гц наступних значень:

- напруга дотику 2 В;
- протікає струм 0,3 мА.

Важливою складовою частиною техніки безпеки пристрою є оптимальний підбір елементної бази, якою проводять таким чином:

1. вибираємо конденсатори з граничною реактивною потужністю $P_{p \text{ доп}}$

визначаючи межі застосування в ланцюгах змінного струму допустимої амплітуди змінної напруги на конденсаторі. Для конденсатора типу КМ, включеного в контур ІМС DA1:

$$P_{p \text{ доп}} - 10 \text{ ВА}; f - 10 \text{ МГц}; C - 0,68 \text{ нФ}.$$

А максимальна напруга в даному колі не перевищує 5 В.

3. Вибираємо резистори по номінальній потужності. Напруга на резисторі не повинно перевищувати значення, що визначає номінальну потужність резистора. Для резисторів $P_{p \text{ доп}} - 0,125 \text{ Вт}; R - 400 \text{ Ом} \dots 20 \text{ кОм}.$

$$U_{\text{НОМ}} = \sqrt{P_{\text{НОМ}} \cdot R}, \quad (5.1)$$

Напруга в ланцюгах не перевищує 5 В. Вибираємо резистор з мінімальною потужністю розсіювання.

4. Зробимо вибір перерізу монтажних проводів:

$$d = 0,8 \cdot \sqrt{I_{\text{max}}}, \quad (5.2)$$

де: d – діаметр (мм); I_{max} – струм, який протікає через проводник.

Максимальний струм, який протікає в монтажних проводах не перевищує 500 мА:

$$d = 0,8 \cdot \sqrt{500 \cdot 10^{-3}} \approx 0,57 \quad S=1 \text{ мм}^2, \quad (4.3)$$

Згідно ДСТУ EN 60227-3:2015 «Кабелі з полівінілхлоридною ізоляцією на номінальну напругу до 450/750 В. Частина 3. Проводи для стаціонарної прокладки». Використовуємо провід МГВ з перерізом не менше 0,2 мм.

Аналогічно можна здійснити вибір інших елементів пристрою.

Конструкція стійки, на якій монтуються елементи виробу, повинна забезпечувати необхідну температуру і вологість. В іншому випадку при конструюванні виробу потрібно передбачити ряд заходів підвищувальних вологостійкості:

- на металеві частини наносять гальванічні покриття;
- деталі покривають водонепроникними і вологостійкими лаками і фарбами.

До електричних схемам висувають такі вимоги:

- довжина зв'язку має бути мінімальною;
- монтаж повинен забезпечувати надійний електричний і механічний контакт;
- для спрощення монтажу і полегшення ремонту застосовують різнокольорові дроти;
- блоки повинні бути заземлені автономно;
- для пайки навісного монтажу застосовують олов'яно-свинцеві припої типу ПОС-61, ПОС-40, при пайку застосовується тільки безкислотний флюс;
- відстань між оголеними монтажними дротами повинно становити не менше 3 мм;
- перетин високочастотних електричних ланцюгів слід виконувати під кутами близькими до 90°;
- під один зажимной контакт можна підключити не більше 3 проводів;

До джгутів монтажних проводів висувають такі вимоги:

- монтажні проводи в джгутах повинні бути покладені і ув'язані рівно, без перехрещування, петель і виступів;
- після укладки джгут повинен бути пов'язаний бавовняними нитками, через заданий крок саме що затягують вузлами;
- джгути повинні бути захищені від впливу високих температур, вологості, механічних пошкоджень;

- джгути повинні бути прикріплені до конструкції металевими скобами через прокладки.

5.3.2. Заходи виробничої санітарії

Метеорологічні умови (температура, відносна вологість, швидкість руху вітру) в робочій зоні виробничих приміщень повинні відповідати вимогам ДСанПіН 3.3.6.042-99 та ДСТУ EN ISO 7730:2016. Ці умови надані в таблиці 5.1:

Таблиця 5.1

Метеорологічні умови в робочій зоні

Період року	Температура, С°		Вологість, %		Швидкість повітря, м/с	
	оптимальна	допустима	оптимальна	допустима	оптимальна	допустима
Холодний	18-20	15-24	60	75	≤0,2	≤0,2
Теплий	21-23	17-29	50	65	≤0,3	≤0,3

Для забезпечення нормальної роботи апаратури та обслуговуючого персоналу в робочому приміщенні створюється мікроклімат:

- температура в приміщенні +19 ... + 24 °С;
- відносна вологість 60 ... 70%;
- швидкість руху повітря 0,1 м / с (на підставі ДСанПіН 3.3.6.042-99).
- температура нагрітих поверхонь обладнання і огороження не повинна перевищувати 45 °С.

Одним з найбільш ефективних методів регулювання метеорологічних умов у виробничих приміщеннях є кондиціонування повітря, яке являє собою штучну обробку повітря з метою встановлення в робочій зоні таких температур, відносної вологості та швидкості повітря, при яких буде забезпечена комфортність праці. Найчастіше кондиціонування повітря здійснюється регулюванням швидкості його руху і температури. Швидкість повітря можна регулювати шляхом зміни площі вихідного отвору повітряноходу. Крім кондиціонування повітря у виробничих приміщеннях, з метою поліпшення мікроклімату встановлюють повітряні і повітряно-теплові завіси,

повітряні душі, створюють оазиси, а при роботах на відкритих майданчиках в зимовий час - кімнати для обігріву. Вентиляційні системи за способом дії поділяються на:

- місцеву;
- загальнообмінну.

З організації подачі та вилучення повітря на:

- припливну;
- витяжну;
- припливно-витяжну.

Вентиляційні системи повинні забезпечити кратність обміну повітря у виробничих приміщеннях для видалення шкідливих речовин, що утворюються при пайку. Всі робочі місця обладнуються місцевою вентиляцією. Всі цехи, дільниці експлуатації та роботи радіоелектронного обладнання забезпечуються механічної припливної вентиляцією. Для нормальної роботи оператора в робочому приміщенні має бути певна кількість повітря, який повинен ще й оновлюватися. Для цього в приміщенні, де встановлена частина апаратури системи охоронної сигналізації повинна використовуватися вентиляція. Зробимо розрахунок кількості подаваного повітря, а також розміри повітрянохода, через який це повітря буде подаватися. Для розрахунку кількості повітря, що подається ми повинні знати кількість людей, що працюють в приміщенні, витрата повітря на одну людину, а також швидкість руху повітря. Задамося вихідними даними:

V - швидкість повітря (0,7 м / с);

n - кількість працюючих (1 особа);

q - витрата повітря на одну людину в годину (20 м / ч);

z - коефіцієнт запасу (1,4).

Кількість подається в приміщення протягом 1 години розраховуємо

за формулою: $Q = n \cdot z \cdot q$,

$$Q = 1 \cdot 1,4 \cdot 20 = 28 \text{ м}^3 / \text{год.}$$

Обчисливши це значення, можемо визначити площу перетину повітрянохода за формулою:

$$S = \frac{Q}{3600 \cdot V} = \frac{28}{3600 \cdot 0,7} = 0,01 \text{ м}^2, \quad (5.3)$$

Знаючи площу перетину, можемо обчислити лінійні розміри повітрянохода:

$$a = b = 10 \text{ см}, \quad (5.4)$$

Як видно з розрахунків, для нормальної роботи у приміщенні, де працює один людини за одну годину має подаватися 28 м чистого повітря. При цьому повітропровід має форму квадратного перетину розміром сторони 10 см. Освітлення, відповідне ДБН В.2.5-28:2018 і ДСТУ EN 12464-1:2016 «Освітлення робочих місць у приміщеннях», є одним з головних умов виробничої санітарії. При розрахунку освітленості по світловому потоку визначають світловий потік кожної лампи по нормованій мінімальній горизонтальній освітленості E_{\min} (лк):

$$F = \frac{E_{\min} \cdot S \cdot k \cdot z}{N \cdot n \cdot \eta}, \quad (5.5)$$

де: F - світловий потік лампи в світильнику, лм; S - площа приміщення, м; k - коефіцієнт запасу; z - коефіцієнт нерівномірності освітлення (1,1 ... 1,15); η - коефіцієнт використання світлового потоку; n - кількість ламп у світильнику.

Загальні вимоги до контролю рівнів звукового тиску на робочих місцях з коливаннями повітря більше 11,2 кГц та ультразвукової характеристики обладнання, методи і засоби усунення та зниження шкідливого впливу ультразвуку наведені в ДСанПіН 3.3.6.039-99. Вимоги до ергономіки та виробничої естетики. Ергономіка - наука про закони взаємодії людини з навколишнім середовищем в процесі трудової діяльності. Під технічною естетикою розуміється наука про прекрасне, яка займається виявленням об'єктивних закономірностей прекрасного у сфері матеріального

виробництва та їх впливом на людину в процесі трудової діяльності. Ергономіка і технічна естетика відіграють істотну роль у науковій організації праці та забезпеченні його безпеки. При проектуванні та експлуатації охоронної сигналізації особливий інтерес представляють відчуття, рухові функції і психічні процеси, що лежать в основі трудової діяльності оператора. Це необхідно враховувати і при його навчанні. Розробка електронної техніки повинна здійснюватися з урахуванням створення сприятливих умов праці. Особливе місце займає організація робочого місця оператора. Відстань від очей оператора до розглянутої поверхні (в метрах):

$$l = \frac{S}{2 \operatorname{tg} \cdot \alpha / 2}, \quad (5.6)$$

де: S - розмір індикатора, м; α - кут огляду, градуси.

Можна запропонувати наступні рекомендації:

- низьке розташування розсіюючих джерел світла;
- в оснащенні використовують предмети з матовою поверхнею;
- оптимальне розміщення елементів праці;
- озеленення приміщень;
- раціональне розміщення інформаційних матеріалів.

За вимогами ергономіки та технічної естетики при фарбуванні обладнання і стін приміщення повинні використовуватися світлі тони (білий, сірий, блакитний і т.д.).

5.4. Розробка заходів з охорони праці

Згідно ПУЕ (7-е вид.) та ДСТУ EN 50522:2014 вихідними даними для розрахунку заземлюючих пристроїв є:

- опір заземлювального пристрою (згідно ПУЕ при $U < 1000$ В - 4 Ом);
- питомий опір ґрунту - 70 Ом / м (пісок);

- коефіцієнт сезонності для однорідної землі для третьої кліматичної зони, для стрижневих вертикальних електродів довжиною 2,5 ... 3 м з глибиною закладення їх вершини 0,5 ... 0,8 - 1,2 ... 1,5 м;

- коефіцієнт сезонності для однорідної землі для третьої кліматичної зони для горизонтальних електродів з глибиною закладення 0,8 м і довжиною 10м - 2 ... 4,2, 50м - 1,6 ... 3,2.

- тип заземлювача - труба, її лінійні розміри: довжина 1 ... 3,0 м, діаметр d - 0,05 м при товщині стінок не менше 3,5 мм, розташування щодо рівня землі на 0,5 м нижче рівня землі;

- тип заземлюючої магістралі - полоса з шириною b - 0,04 м, глибина закладення магістралі h - 3 м;

- вид розміщення вертикальних заземлювачів по контуру, відстань між ними a - 3 м;

1. Питомий опір ґрунту для вертикальних заземлень:

$$\rho'_{\text{расч}} = \rho K'_n = 70 \cdot 1,35 = 94,5 \text{ Ом/м}, \quad (5.7)$$

3. Питомий опір ґрунту для магістралі:

$$\rho''_{\text{расч}} = \rho K''_n = 70 \cdot 2,4 = 168 \text{ Ом/м}, \quad (5.8)$$

3. Опір розтіканню струму для одиночного вертикального заземлювача:

$$R_T = \frac{\rho'_{\text{расч}}}{2\pi l} \left(\ln \frac{2l}{d} + \frac{1}{2} \ln \frac{4t+1}{4t-1} \right) = 25,97 \text{ Ом}, \quad (5.9)$$

4. Умовне кількість вертикальних заземлювачів:

$$n' = \frac{R_T}{R_{\text{доп}}} = 7, \quad (5.10)$$

5. Знайдемо відношення a / l і вид розміщення вертикальних заземлювачів, коефіцієнт використання заземлювачів, що враховує взаємне екранування. Приймаємо a / l за 1. Вид розміщення вертикальних заземлювачів - по контуру. Коефіцієнт використання заземлювачів - 0,56.

6. Дійсно уточнена кількість вертикальних заземлювачів:

$$n = \frac{n'}{\eta_{\text{ЭМ}}} \approx 13, \quad (5.11)$$

7. Довжина магістралі, що сполучає вертикальні заземлювачі:

$$L = (1,05 \dots 3,5) l n = 78 \text{ м}, \quad (5.12)$$

8. Опір розтіканню струму для заземлювачів магістралі:

$$R_{\text{М}} = R_{\text{п}} = \frac{\rho''_{\text{расч}}}{2\pi l} \left(\ln \frac{2l}{bn} \right) = 19,8 \text{ Ом}, \quad (5.13)$$

9. Коефіцієнт використання заземлюючої магістралі при розташуванні вертикальних заземлювачів по контуру, що враховує взаємне екранування магістралі і вертикальних заземлювачів:

$$\eta_{\text{ЭМ}} = 0,34, \quad (5.14)$$

10. Опір розтіканню всього заземленого пристрою:

$$R_3 = \frac{R_{\text{М}} R_{\text{Т}}}{R_{\text{Т}} \eta_{\text{ЭМ}} + R_{\text{М}} \cdot n \cdot \eta_{\text{ЭМ}}} = 5,01 \text{ Ом}, \quad (5.15)$$

Підвищений рівень статичної електрики скупчується на поверхні апаратури і при випадковому дотику до такої поверхні обслуговуючий персонал може уразитися.

Для зниження ступеня електризації необхідно застосовувати провідні матеріали для покриття підлог, робочих місць, стільців, взуття та одягу обслуговуючого персоналу, а також необхідно підтримувати вологість повітря на рівні максимально допустимого значення відповідно до ДСТУ EN 61340-5-1:2019. При контакті взуття з підлогою необхідно забезпечити безперервне зняття електростатичного заряду. Це досягається шляхом приведення опору покриття та взуття до значення не перевищує 1 МОм. Обслуговуючий персонал повинен бути забезпечений спецодягом з мало електризованих безворсових тканин. Ефективним засобом розсіювання зарядів є струмопровідні браслети, укріплюються на зап'ястях і з'єднані з заземленим проводом з опором $10^6 - 10^8$ Ом. При роботі потрібно застосовувати справні інструменти та обладнання. На робочому місці всі металеві та електропровідні неметалеві частини устаткування повинні бути заземлені, а також застосовуватися антиелектростатичне взуття з електропровідного підошвою. Електричні опору між струмопровідними елементами антиелектростатичного одягу повинно бути від 0 до 10 Ом [19].

5.5. Пожежна безпека

Пожежна небезпека РЕО, обчислювальних комплексів може бути обумовлена утворенням електричної дуги і розпечених частинок металу, іскор, перегрівом струмоведучих елементів, запалюванням легкозаймистих рідин і пароповітряних сумішей, застосовуваних при експлуатації та технічному обслуговуванні систем. Пожежна небезпека обладнання визначається також і наслідками, які можуть спричинити пожежу. Пожежну навантаження радіоелектронних пристроїв визначають наявністю горючих електроізоляційних матеріалів, горючих конструктивних матеріалів, горючих конструктивних елементів приміщень та меблів, горючих матеріалів, використовуваних для акустичної обробки приміщень, а також для зберігання інформації, неправильний вибір перерізу живлячих проводів, а також неправильний вибір плавких запобіжників [20].

Розрахунок перерізу проводів живлення здійснюється за формулою:

$$S = \frac{I_m}{J_{\text{доп}}} = 0,15 \text{ мм}^2, \quad (5.16)$$

где: S – переріз дроту (мм^2); I_m – максимальний споживаний струм (1,5 А);

$J_{\text{доп}}$ – допустима щільність струму на одиницю площі (10 А/мм^2).

Згідно ДСТУ EN 60227-3:2015 використовуємо провід перетином 0,2 мм. Розрахуємо запобіжник в ланцюзі харчування. Номінальний струм навантаження, що протікає через вузли розроблюваного пристрою I_H - 1,0 А. Максимальний струм протікає через запобіжник:

$$I_{\text{пр}} = (1,5 \dots 1,8) I_H; \quad (5.17)$$

$$I_{\text{пр}} = 1,5 \dots 1,8 \text{ А}, \quad (5.18)$$

Згідно ДСТУ EN 60127-2:2016 використовуємо стандартне значення запобіжника в ланцюзі живлення на струм рівний 2 А. Тип запобіжника ВП-26-1В-2А. Загальні вимоги по забезпеченню пожежної та вибухової безпеки об'єктів виробничого призначення визначені відповідно у «Кодексі цивільного захисту України» і ДБН В.1.1-7:2016. Нижче наводяться заходи щодо забезпечення пожежної та вибухової безпеки системи охоронної сигналізації та обчислювальних комплексів, які можуть бути поширені і на проєктований виріб.

1. Автоматизовані системи охоронної сигналізації та обчислювальні ком-комплекси слід розташовувати в будівлях не нижче другого ступеня вогнестійкості з розташуванням залів не нижче першого поверху.

2. Комплекс виробничих приміщень повинен мати не менше двох самостійності евакуаційних виходів.

3. Для акустичної обробки стін слід застосовувати негорючі і важко горючі матеріали.

4. Джерела електричної енергії (розподільні пристрої, трансформатори) необхідно розташовувати в окремих приміщеннях.

5. Освітлювальну електричну мережу слід виконувати згідно з вимогами ПУЕ для пожежонебезпечних зон і установок класів 2-2а.

6. Прокладання кабелів через перекриття стіни і фальшполи слід здійснювати в сталевих трубах з ущільненням з негорючих матеріалів.

7. Аварійні мережі освітлення, дистанційного та автоматичного пуску протипожежної системи та сигналізації необхідно прокладати окремо від силових та інших робочих електричних мереж, а при спільному прокладанні їх слід розділяти перегородками з негорючих матеріалів.

8. Повітроводи слід виконувати з негорючих матеріалів.

9. При ремонті і технічному обслуговуванні устаткування в приміщенні повинно знаходитися мінімальна кількість легкозаймистих рідин в щільно закритій тарі.

5.6. Розрахункова частина. Розрахунок допустимого часу робіт при електромагнітному випромінюванні

У відкритому розподільчому пристрої, де розташована апаратура з напругою $U = 500$ кВ, з живленням змінним струмом промислової частоти 50 Гц, планується проведення планових робіт на ділянках з підвищеною напруженістю електричного поля. Робота буде проводитися без використання захисних засобів – екрануючих костюмів, екранів.

Тривалість робіт на ділянці А, де напруженість електричного поля $E_A=10$ кВ/м і $t_{EA}=60$ хв.; на ділянці В, де напруженість електричного поля $E_B=8$ кВ/м і $t_{EB}=90$ хв. Визначити фактичний час виконання робіт t_{EC} для третьої ділянки С, де напруженість електричного поля $E_C=6$ кВ/м, а також загальний час виконання робіт.

РОЗВ'ЯЗАННЯ: У робочій зоні, що характеризується різними значеннями напруженості електричного поля, перебування персоналу обмежується граничним часом, $T_{гр}$:

$$T_{гр} = 8 \cdot \left(\frac{t_{E_1}}{T_{E_1}} + \frac{t_{E_2}}{T_{E_2}} + \dots + \frac{t_{E_n}}{T_{E_n}} \right), \quad (5.19)$$

де $t_{E_1, \dots, n}$ і $T_{E_1, \dots, n}$ – фактичний і допустимий час (в годинах) перебування персоналу у конкретних зонах з напруженістю поля - E_1, \dots, E_n .

Допустимий час T_E (виміряний у годинах) перебування персоналу у зонах з напруженістю поля E (виміряного у кВ/м) знаходять за формулою:

$$T_E = \frac{50}{E} - 2, \quad (5.20)$$

Тоді з формули 4.8 знайдемо допустимий час перебування персоналу в зонах А, В, С:

$$T_{EA} = \frac{50}{10} - 2 = 3 \text{ години}; T_{EB} = \frac{50}{8} - 2 = 4,25 \text{ годин}; T_{EC} = \frac{50}{6} - 2 = 6,33 \text{ годин}$$

Підставивши отримані значення у формулу 4.7 и вважаючи, що $T_{гр}$ не повинен перевищувати 8 годин (тобто $T_{гр} = 8$ годин), знаходимо, що шуканий фактичний час перебування персоналу в зоні С складає:

$$8 = 8 \cdot \left(\frac{1}{3} + \frac{1,5}{4,25} + \frac{t_{E-C}}{6,33} \right)$$

або $t_{EC} = 2$ години.

Таким чином час роботи на ділянці С не повинен перевищувати 2 години, а загальний час перебування на всіх трьох ділянках не повинен перевищувати: $t_{ЗАГ} = t_{EA} + t_{EB} + t_{EC} = 1 + 1,5 + 2 = 4,5$ години

ВИСНОВКИ ДО РОЗДІЛУ 5

В результаті розрахунків проектованого пристрою, дотримуючись вимог і положень охорони праці, які служать для запобігання впливу шкідливих та небезпечних чинників, вдалося мінімізувати їх негативну дію на обслуговуючий персонал.

РОЗДІЛ 6

ОХОРОНА НАВКОЛИШНЬОГО СЕРЕДОВИЩА

6.1. Вступ

Охорона навколишнього середовища це форма відносин між суспільством і природою. Вона здійснюється різними засобами: економічними, правовими, науково-технічними, санітарно-гігієнічними, біологічними та іншими.

В загальному випадку проблема охорони навколишнього середовища зводиться до вирішення двох завдань:

- організації раціонального природокористування;
- забезпечення чистоти природних (екологічних) систем.

6.2. Основні принципи охорони навколишнього середовища

Основними принципами охорони навколишнього середовища є:

- пріоритетність вимог екологічної безпеки;
- обов'язковість додержання екологічних стандартів, нормативів, лімітів використання природних ресурсів при здійсненні господарської та іншої діяльності;
- гарантування екологічної безпеки середовища для життя і здоров'я людей;
- превентивність - попередження негативних наслідків впливу на природні об'єкти діяльності людей;
- обов'язковість екологічної експертизи;
- екологізація матеріального виробництва;
- поєднання екологічних, економічних і соціальних інтересів суспільства на основі міждисциплінарних знань, екологічних, природничих, технічних та інших наук;
- гласність і демократизм при прийнятті рішень, реалізація яких впливає на стан навколишнього природного середовища;

- безкоштовність загального і платність спеціального природокористування;
- збереження просторового та видового різноманіття;
- цілісність природних об'єктів і комплексів;
- прогнозованість якості стану навколишнього природного середовища;
- платність за забруднення навколишнього природного середовища;
- компенсація шкоди, спричиненої порушенням екологічного законодавства;
- дотримання вимог лімітів використання природних ресурсів;
- поєднання засобів стимулювання при дотриманні вимог екологічного законодавства і юридичної відповідальності при порушенні цих вимог;
- міжнародне співробітництво в екологічній сфері та інше.

6.3. Методи захисту при травленні друкованих плат та пайці

Якщо кілька років назад вибір технологій був не так вже й великий, то сьогодні завдяки розвитку сучасної техніки є можливість швидко і якісно виготовляти друковані плати. На сьогоднішній день існує і відомо багато складів для хімічного підбурювання міді (способів виготовлення друкованих плат). Всі вони відрізняються швидкістю протікання реакції, складом виділяються в результаті реакції речовин, а також доступністю необхідних для приготування розчину хімічних реактивів. Нижче наведена інформація про найбільш популярних розчинах для травлення а саме:

1. Травлення міді розчином хлорного заліза.
2. Травлення мідним купоросом з сіллю.
3. Травлення персульфатами (персульфатів амонію або персульфатів натрію).
4. Травлення перекисом водню в соляній кислоті.
5. Травлення міді перекиснем водню з використанням лимонної кислоти.

У даній роботі використовуватиму найбезпечніший метод травлення друкованої плати, а саме метод травлення міді перекиснем водню з використанням лимонної кислоти. Використання методу травлення під номером 5 травлення міді перекиснем водню з використанням лимонної кислоти є безпечним для людини так

для навколишнього середовища, не потребує спеціального методу захисту навколишнього середовища та людини.

У даній роботі на даному етапі, а саме при пайці радіоелементів використовуватимуть припій олов'яно-свинцевий марки ПОС-61, що має широкий спектр застосування в різних галузях промисловості та побуті.

Таблиця. 6.1

Технічні характеристики ПОС 61

Марка	Температура плавління, °С		Щільність, гр/см ²	Питомий опір	Теплопровідність	Опір на розрив	Подовження відносне, %	В'язкість ударна, кг/см ²
	Початкова	Кінцева						
ПОС 61	183	189	8,5	0,139	0,12	4,3	46	3,9

Основним фактором визначення властивостей припія ПОС 61 в інших видах виробу є його хімічний зміст закріплений ГОСТот 21930-76.

Таблиця 6.2

Хімічний зміст ПОС 61

Масова доля основних компонентів %		Масова доля домішок %							
Sn	Pb	Sb	Cu	Bi	As	Fe	Ni	S	Zn
59-61	41-39	0,1	0,05	0,2	0,02	0,02	0,02	0,02	0,002

Отже при використанні ПОС 61 припій олов'яно-свинцевий є небезпечні речовини як :

- випаровування свинцевого припою;
- випаровування сірки.

Згідно ГОСТ 12.1.005-88 до забезпечення метеорологічних умов і чистоти повітря в приміщенні електромонтажної робочої ділянки встановлена система витяжки припливного типу. При пайці радіоелементів утворюються шкідливі виділення для локалізації і запобігання цих виділень у повітря робочої зони приміщень призначена місцева витяжна вентиляція по принципу відсмоктування повітря [21].

З метою захисту навколишнього середовища та людини комплекс вентиляційних пристроїв які видаляють з робочих місць шкідливі виділення є необхідністю встановлення повітряного фільтра.

6.4. Електронні компоненти їх утилізація та вторинна переробка

Значна зміна клімату і зниження біологічного різноманіття є лише двома з багатьох серйозних екологічних проблем, які продовжують рости в глобальному масштабі. Населення планети на даний момент становить понад 7 мільярдів і в зв'язку з цим зростає занепокоєння з приводу браку продовольства, води, енергії та інших ресурсів. Щоб зменшити шкоду, що наноситься навколишньому середовищу та дефіцит ресурсів, нам варто впритул зайнятися переробкою вторинної відслужили предметів. Дуже важлива переробка електроніки.

До електронних відходів відносяться в тому числі друковані плати, які хоча й становлять близько 3% від загальної кількості цього виду відходів, але завдяки високій концентрації токсичних речовин є дуже небезпечними. Подібні відходи без належної утилізації негативно впливають на екосистему, як біотичну, так і на абіотичні її частини. Наявність різноманітних високотоксичних матеріалів і важких металів робить поховання на звалищі або просте спалювання неприйнятними методами управління подібними відходами. Тому найбільш оптимальний спосіб утилізації електронних відходів — це їх переробка.

Крім того, що електронні відходи являють собою велику небезпеку для навколишнього середовища, треба пам'ятати, що на виробництво мобільних телефонів та персональних комп'ютерів йдуть значні частки золота, срібла і паладію,

що видобуваються щорічно у всьому світі. Звичайно, в кожному окремому пристрої дорогоцінних металів міститься мізерна кількість, але якщо розглядати загальносвітове виробництво (понад 1,2 млрд. щорічно), то цим кількістю вже нерозумно нехтувати. Слід зазначити, що концентрація цих дорогоцінних металів в друкованих платах більш, ніж у десять разів перевищує їх концентрацію у видобутої руди. Однак переробка друкованих плат технологічно складний процес із-за неоднорідності матеріалів, адже вони складаються з безлічі різнорідних компонентів. На жаль, визначити точну кількість вироблених електронних відходів дуже складно. За оцінками ЮНЕП ще 7 років тому електронні відходи становили близько 25-55 млн. тонн в рік (2008 р.). Агентство США з охорони навколишнього середовища повідомило про те, що кожна сім'я в США використовує близько 35 електронних пристроїв та електричних приладів (дані 2011 року). Це в середньому призводить до утворення більш 5 млн. тонн електронних відходів в рік. Для ЄС було підраховано, що в середньому кожен громадянин виробляє близько 17 кг електронних відходів в рік, що призводить до утворення 8 млн. тонн відходів (дані 2012 р.).

Забезпечення потреб сучасної людини вимагає від технічного прогресу створення все більшої кількості електронних пристроїв, які дозволяють людині економити час та почувати себе більш комфортно. Нажаль, невеликий термін експлуатації деяких пристроїв, а що важливіше, потужний тиск маркетингових акцій на свідомість людей призводять до швидкої заміни електронних пристроїв. У результаті, щороку у світі утворюється близько 50 млн. тон електронних відходів, а їх накопичення відбувається втричі швидше, ніж зростання кількості інших відходів. Найбільш швидке зростання кількості електронних пристроїв спостерігається в країнах, що розвиваються (в першу чергу Китай, Індія, країни Латинської Америки). Це викликано тим, що Америка та багато європейських країн експортують електронні відходи до «країн третього світу», погіршуючи в них екологічну ситуацію. Так вже зараз у Китаї щорічно на звалища потрапляє 500 тис. т холодильників, 1,3 млн. т телевізорів і 300 тис. т персональних комп'ютерів. В Індії – 275 тис. т холодильників, 275 тис. т телевізорів та 56,4 тис. т комп'ютерів .

В Україні проблема накопичення електронних відходів стоїть дуже гостро через: 1) зростання ринку електронної техніки та 2) відсутність налагодженої системи її утилізації. Згідно зі статистичними даними на території України перебувають у користуванні 53,6 мільйона мобільних засобів зв'язку, щороку імпортується 300 тисяч портативних комп'ютерів, 277 мільйонів елементів живлення (батареєнок). У перерахунку на вагу це становить 4,5 тисячі тонн батареєнок і акумуляторів на рік, які після відпрацювання є потенційно небезпечними відходами [22]. Результати дослідження ринку "GfK Ukraine" за перший квартал 2013 року свідчать, що порівняно з аналогічним кварталом минулого року обсяги продажів в грошовому виразі зросли на 14%. В «зеленій зоні» опинились усі сектори, а найбільшого росту досягли групи побутової електроніки (22%), великої (19%) та малої (18%) побутової техніки та телеком - продукція (14%). Загальні обсяги продажу побутової техніки та електроніки за підсумками кварталу складають майже 11 млрд. грн. Динаміка продажів побутової техніки представлена на рис. 6.1 .

Процес переробки електронних відходів складається з трьох етапів:

1) збір;

2) сортування/демонтаж та попередня обробка (включає сортування, демонтаж, механічну обробку);

3) кінцева обробка (очищення та утилізація) . Основною причиною, через яку в Україні накопичуються на звалищах електронні відходи, є відсутність законодавчого поля, яке б регулювало всі 3 етапи.

Відомо, що у 90-ті роки деякі європейські країни заборонили розміщення електронних відходів на полігонах. А виробників електронних приладів на законодавчому рівні в ЄС зобов'язали безкоштовно приймати відпрацьоване або непридатне для подальшого використання електричне й електронне обладнання, нести відповідальність за його утилізацію, зокрема й фінансову.

Динаміка товарообігу ринку побутової техніки та електроніки

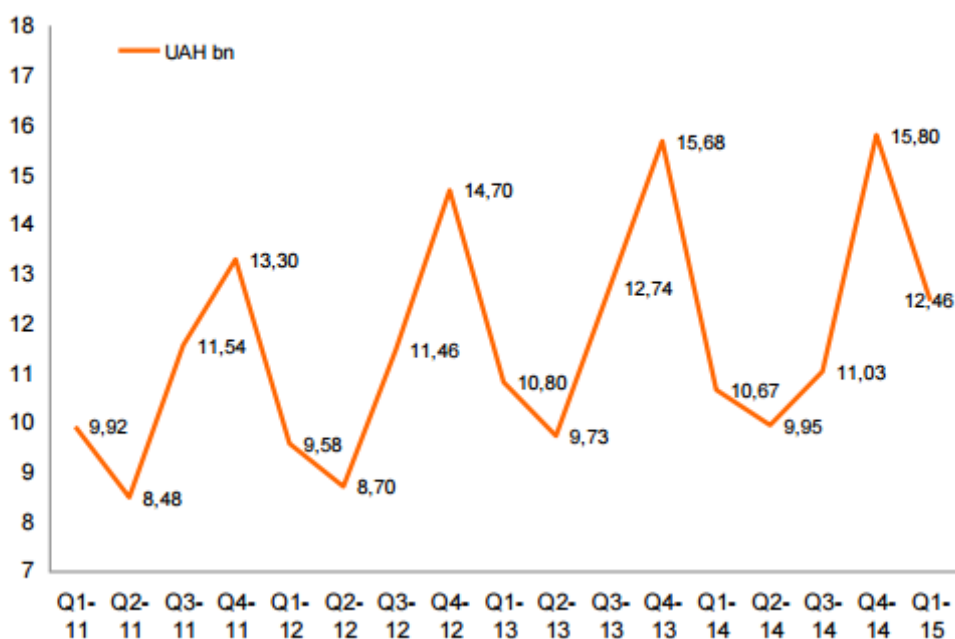


Рис. 6.1. Динаміка продажів побутової техніки

Щодо ситуації в Україні, то 16 лютого 2010 року набув чинності Закон України «Про внесення змін до деяких законодавчих актів України у сфері поводження з відходами». Завдяки цьому документу Закон України «Про відходи» доповнено статтею 35-1 «Вимоги до поводження з побутовими відходами», відповідно до якої небезпечні відходи необхідно збирати окремо від інших видів побутових відходів, а також відокремлювати на етапі збирання чи сортування та передавати спеціалізованим підприємствам, які мають ліцензії на їх переробку. До I класу небезпеки відносяться відпрацьовані люмінесцентні лампи та відпрацьовані малі побутові хімічні джерела енергії (батарейки та акумулятори). Наразі в Україні, відходи електроніки утилізуються лише на одному підприємстві у Львові (ДП «Аргентум»). Головною проблемою є відсутність системи централізованого збору відпрацьованої електронної техніки та люмінесцентних ламп, якими користується більшість мешканців. Щодо батарейок, ситуація виглядає краще завдяки встановленню пунктів їхнього збору у громадських місцях. Окрім цього, утилізацією електронного обладнання та комп'ютерної техніки займаються приватні підприємства, зокрема НВП «Екологічна лабораторія» .

Таким чином, в Україні спостерігається прискорення темпів накопичення електронних відходів, а недосконала система управління призводить до неможливості ефективної утилізації відпрацьованого електричного та електронного обладнання. Згідно огляду Всеукраїнської екологічної громадської організації «МАМА-86», для покращення існуючої системи поводження з електронними відходами необхідно включити основні принципи «електронних» директив ЄС (Директива 2002/96/ЄС про відходи електричного та електронного обладнання та Директива 2002/95/ЄС про обмеження використання певних небезпечних речовин в електричному та електронному обладнанні; нові Директива 2012/19/ЄС та Директива 2011/65/ЄС) в національне законодавство.

ВИСНОВКИ ДО РОЗДІЛУ 6

Сьогодні людство не може відмовитися від використання природних ресурсів, які є основою матеріального виробництва, але не може й миритися з деградацією навколишнього природного середовища внаслідок інтенсивної експлуатації. Тому важливим напрямом сучасного соціального регулювання відносин в сфері взаємодії природи і суспільства є забезпечення раціонального природокористування. Його метою є задоволення матеріальних потреб суспільства за рахунок наявних і відновлювальних природних ресурсів, визначення найбільш ефективного, економічно вигідного їх використання без заподіяння шкоди для життєво важливих екологічних інтересів людей та природного середовища.

ВИСНОВКИ

У кваліфікаційній роботі на тему «Підвищення ефективності виявлення малих об'єктів БПЛА» проведено комплексне дослідження сучасних підходів до детекції об'єктів у зображеннях, отриманих із безпілотних літальних апаратів, розроблено власну вдосконалену модель UAV-DETR++ та продемонстровано її практичну ефективність у завданнях комп'ютерного зору.

У вступній частині роботи обґрунтовано актуальність теми, пов'язану з інтенсивним розвитком безпілотних технологій та зростанням ролі автоматизованих систем аналізу відеопотоку. Умови, у яких працюють БПЛА — змінна висота, широкий кут огляду, малі розміри цілей та складний фон — створюють додаткові труднощі для традиційних алгоритмів виявлення. Тому основна увага в роботі зосереджена на аналізі та подоланні цих обмежень шляхом вдосконалення архітектури глибинної нейронної мережі.

Перший розділ присвячено огляду сучасних методів виявлення об'єктів, включно з CNN-архітектурами (R-CNN, YOLO, SSD), трансформерними підходами (DETR, Deformable DETR, RT-DETR) та спеціалізованими моделями для UAV-сцен (UAV-YOLO, Drone-DETR). Проведено порівняльний аналіз, який показав, що хоча CNN-моделі мають високу швидкодію, вони поступаються у глобальному моделюванні контексту, а трансформерні підходи — навпаки, мають кращу семантичну узгодженість, але потребують оптимізації для реального часу. Цей висновок став підґрунтям для подальшої розробки гібридної структури UAV-DETR++.

У другому розділі визначено основні технічні проблеми, характерні для знімків із БПЛА: невеликі розміри об'єктів (10–30 пікселів), їхня щільність, рухливість і вплив складного фону. На основі цього сформульовано вимоги до майбутньої моделі: збереження дрібних ознак при зменшенні зображення, покращене злиття багатомасштабних карт і адаптивне контекстне вирівнювання. Розділ також містить

теоретичне обґрунтування вибору архітектури RT-DETR як базової та її подальшу модифікацію під задачі UAV-аналітики.

У третьому розділі запропоновано структурне вдосконалення моделі UAV-DETR, у якому реалізовано три інноваційні компоненти:

- CasNet — забезпечує збалансовану обробку просторових і каналних залежностей із мінімальною кількістю параметрів;
- SOEP-v2 — підвищує чутливість до дрібних цілей через розширену багатомасштабну піраміду ознак;
- CSAM — гармонізує контекстну інформацію між різними рівнями ознак і усуває просторову неузгодженість.

На основі моделі UAV-DETR++ проведено оптимізацію inference-процесу, включно з квантизацією до формату INT8, що забезпечило реальну можливість впровадження на вбудованих системах типу NVIDIA Jetson Orin NX або Google Coral Edge TPU.

У четвертому розділі підтверджено практичну реалізованість запропонованого рішення. Модель розгорнуто на нейронному прискорювачі Jetson Orin NX, де вона забезпечила швидкість понад 30 кадрів/с при середньому енергоспоживанні 14 Вт. Результати тестування на наборі VisDrone2019-DET засвідчили зростання точності (mAP@0.5) на 3.5% порівняно з базовим RT-DETR та підвищення recall для малих об'єктів більш ніж на 5%. Модель також продемонструвала добру узагальнювальну здатність на незалежному наборі DOTA, перевищивши Faster R-CNN на 7% і RT-DETR на 2.1%.

Отримані результати свідчать, що UAV-DETR++ досягає оптимального балансу між точністю, швидкістю та апаратною ефективністю. Її структура є модульною, тому можливе подальше розширення (наприклад, інтеграція в мультиагентні системи або адаптація до теплових камер).

Практичне значення роботи полягає у створенні архітектури, здатної до ефективного виявлення малих об'єктів у реальному часі на енергообмежених платформах. Це відкриває можливість її застосування у системах:

- аеророзвідки та моніторингу територій;

- цивільної безпеки, пошуково-рятувальних операцій;
- автоматизованого контролю інфраструктури й екологічного нагляду.

При цьому наголошено, що результати мають алгоритмічно-аналітичний характер і не містять конкретних тактичних вказівок щодо військового використання, однак можуть сприяти підвищенню ситуаційної обізнаності Збройних Сил України у межах оборонних завдань.

У підсумку можна зробити висновок, що поставлена в роботі мета — підвищення ефективності виявлення малих об'єктів на зображеннях із БПЛА — досягнута повністю.

Розроблена модель UAV-DETR++ поєднує сучасні досягнення трансформерних архітектур і методів оптимізації глибокого навчання, має високу ефективність, малу затримку та широкий потенціал для практичного використання в інтелектуальних системах комп'ютерного зору майбутнього.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Yu Zhou, Yan Wei. UAV-DETR: An Enhanced RT-DETR Architecture for Efficient Small Object Detection in UAV Imagery.: College of Computer and Information Science, Chongqing Normal University, Chongqing 401331, China.
2. Girshick, R. Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 1440–1448.
3. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27 June–1 July 2016; pp. 779–788.
4. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. SSD: Single shot multibox detector. In Proceedings of the European Conference on Computer Vision (ECCV), Amsterdam, The Netherlands, 11–14 October 2016; pp. 21–37.
5. Lin, T.Y.; Dollár, P.; Girshick, R.; HE, K.; Hariharan, B.; Belongie, S. Feature Pyramid Networks for Object Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 936–944.
6. Carion, N.; Massa, F.; Synnaeve, G.; Usunier, N.; Kirillov, A.; Zagoruyko, S. End-to-End Object Detection with Transformers. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 August 2020; pp. 213–229.
7. [Інтернет посилання]: метод доступу <https://www.baidu.com/>
8. Bujagouni, G.K.; Pradhan, S. Transformer-based deep learning hybrid architecture for phase unwrapping. *Phys. Scr.* 2024, 99, 7.
9. Liu, M.; Wang, X.; Zhou, A.; Fu, X.; Ma, Y.; Piao, C. UAV-YOLO: Small Object Detection on Unmanned Aerial Vehicle Perspective. *Sensors* 2022, 20, 2238.

10. Chen, S.; Cheng, T.; Fang, J.; Zhang, Q.; Li, Y.; Liu, W.; Wang, X. TinyDet: Accurate Small Object Detection in Lightweight Generic Detectors. arXiv 2023, arXiv:2304.03428.
11. Zhao, D.; Shao, F.; Yang, L.; Luo, X.; Liu, Q.; Zhang, H.; Zhang, Z. Object Detection Based on an Improved YOLOv7 Model for Unmanned Aerial-Vehicle Patrol Tasks in Controlled Areas. Electronics 2023, 12, 4887.
12. Wei, C.; Wang, W. RFAG-YOLO: A Receptive Field Attention-Guided YOLO Network for Small-Object Detection in UAV Images. Sensors 2025, 25, 2193.
13. Zhang, T.; Li, L.; Zhou, Y.; Liu, W.; Qian, C.; Hwang, J.N.; Ji, X. Cas-ViT: Convolutional Additive Self-Attention Vision Transformers for Efficient Mobile Applications. arXiv 2024, arXiv:2408.03703.
14. Hu, J.; Shen, L.; Sun, G. Squeeze-and-Excitation Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 7132–7141.
15. Cui, Y.; Ren, W.; Knoll, A. Omni-Kernel Network for Image Restoration. In Proceedings of the AAAI Conference on Artificial Intelligence, Vancouver, BC, Canada, 20–27 February 2024; AAAI Press: Palo Alto, CA, USA, 2024; Volume 38, pp. 1426–1434.
16. Sunkara, R.; Luo, T. No More Strided Convolutions or Pooling: A New CNN Building Block for Low Resolution Images and Small Objects. In Proceedings of the 22nd Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD 2022), Grenoble, France, 19–23 September 2022; pp. 443–459.
17. [Інтернет посилання]: метод доступу <https://www.nvidia.com/en-us/autonomous-machines/embedded-systems/jetson-orin/nano-super-developer-kit/>
18. [Інтернет посилання]: метод доступу <https://developers.google.com/coral>
19. [Інтернет посилання]: метод доступу <https://ua.esdsuppliers.com/news/what-are-the-requirements-for-esd-clothing-77601969.html>

20. НАПБ А.01.001-2004 «Правила пожежної безпеки в Україні» (розділ 5.1), електроустановки (можливість їх застосування, монтаж, налагодження та експлуатація)
21. ПРАВИЛА охорони праці при термічній обробці металів, від 28 січня 2008 р. за N 66/14757
22. Електронні відходи: аналіз ринку та проблеми утилізації. [Інтернет посилання]: метод доступу http://www.ukrpryroda.org/2013/08/blog-post_10.html