

Метод нейромережевого виявлення та прогнозування траєкторій БПЛА у відеопотоці

Шифр: Прогнозування БПЛА

Зміст

Вступ.....	3
Розділ 1 Аналіз існуючих моделей, методів та реалізацій.....	5
1.1 Аналіз предметної області виявлення та відстеження об'єктів	5
1.2 Аналіз наукових праць за темою дослідження	7
1.3 Мета та завдання дослідження	8
Розділ 2 Метод нейромережевого виявлення та прогнозування траєкторій БПЛА	10
2.1 Пайплайн нейромережевого виявлення та прогнозування траєкторій БПЛА	10
2.2 Удосконалений підхід до прогнозування траєкторій БПЛА на основі позиційно-кодованого трансформера	11
2.3 Удосконалена архітектура на основі позиційно-кодованого трансформера ..	13
2.4 Набір даних дослідження	15
Розділ 3 Дослідження ефективності методу нейромережевого виявлення та прогнозування траєкторій БПЛА.....	17
3.1 Структура та функціональне призначення програмних складових системи ..	17
3.2 Результати досліджень	19
Висновки	26
Література	28

Вступ

Стрімке зростання використання БПЛА у цивільних, промислових та військових сферах створює як нові можливості, так і суттєві виклики для систем безпеки та моніторингу повітряного простору. БПЛА застосовуються для екологічного моніторингу, пошуку і рятування, доставки малогабаритних вантажів, сільськогосподарського нагляду та енергоефективного картографування територій, що прямо корелює з цілями сталого розвитку ООН – зокрема SDG 9 «Індустріалізація, інновації та інфраструктура», SDG 11 «Сталий розвиток міст та громад» і SDG 13 «Боротьба зі зміною клімату». У той же час їхнє масове впровадження породжує ризики несанкціонованого використання, що впливають на SDG 16 «Мир, справедливість та сильні інститути», адже малогабаритні некеровані дрони можуть становити загрозу для критичної інфраструктури, цивільного населення або військових об'єктів.

Традиційні методи контролю повітряного простору, такі як радіолокаційні системи або акустичні сенсори, є недостатньо ефективними щодо малої авіації, зокрема дронів з малим радіолокаційним відбитком, низькою швидкістю або складною траєкторією польоту. У цьому контексті комп'ютерний зір у поєднанні з нейромережевими моделями відкриває нові можливості для розроблення інтелектуальних систем виявлення та прогнозування руху БПЛА за даними відеоспостереження. На відміну від класичних підходів, які обмежуються лише фіксацією положення об'єкта в даний момент часу, сучасні глибинні моделі дозволяють прогнозувати траєкторію руху у майбутньому, що є критично важливим для автоматизованого ухвалення рішень у системах протидії або супроводу.

Попри значний прогрес у задачах об'єктної детекції та трекінгу, проблема прогнозування траєкторій у реальному часі з урахуванням маневреності БПЛА та шумів відеосигналу залишається відкритою. Більшість наявних рішень не враховують раптові зміни швидкості або напрямку, не забезпечують адаптивного перемикання між різними моделями руху, або демонструють значну похибку за

умов часткової втрати об'єкта з кадру. Це зумовлює потребу у комбінованому підході, який би поєднував швидкість статистичних методів фільтрації та адаптивність нейромережевого прогнозування.

Таким чином, актуальність даного дослідження полягає у розробленні методу нейромережевого виявлення та прогнозування траєкторій БПЛА у відеопотоці, який забезпечує стійке розпізнавання малорозмірних об'єктів, точне відстеження їхнього руху та формування короткострокового прогнозу з мінімальною похибкою. Такий підхід може бути інтегрований у системи моніторингу повітряного простору, охорони критичних об'єктів, супроводу гуманітарних вантажів або екологічного спостереження, сприяючи досягненню цілей сталого розвитку через підвищення безпеки, технологічної стійкості та відповідального використання автономних технологій.

Розділ 1 Аналіз існуючих моделей, методів та реалізацій

1.1 Аналіз предметної області виявлення та відстеження об'єктів

Задача виявлення об'єктів у відео- та візуальних даних є фундаментальною у сучасних системах комп'ютерного зору та відіграє ключову роль у таких прикладних областях, як транспортний моніторинг, робототехніка, медична діагностика та безпілотна авіація [1]. Її суть полягає у визначенні просторових координат одного або кількох об'єктів на кадрі (локалізація) та встановленні їх належності до певного класу (класифікація). Ці два процеси зазвичай реалізуються спільно: модель не лише формує обмежувальну рамку навколо об'єкта, але й визначає його тип із відповідною ймовірністю. Ефективність цього процесу значною мірою залежить від коректного виділення ознак, що описують структуру та візуальну форму об'єкта.

Попри значний прогрес у цій галузі, задача виявлення залишається складною через низку факторів. По-перше, необхідність одночасного розпізнавання великої кількості класів суттєво впливає на обчислювальну складність і може зменшувати точність моделей. По-друге, детектори часто демонструють залежність від ракурсу та умов освітлення: варіації положення, тіні або спотворення можуть призводити до хибних передбачень. По-третє, високоточні моделі потребують значних обчислювальних ресурсів, що обмежує їх використання на вбудованих системах. Нарешті, алгоритми, навчені лише на певному наборі об'єктів, здатні помилково інтерпретувати невідомі елементи через поверхневу подібність [2].

У сучасній практиці основним підходом до вирішення задачі виявлення об'єктів є застосування згорткових нейронних мереж. Вони дозволяють виділяти ієрархічні ознаки, переходячи від простих контурів до складних структурних патернів. За архітектурною організацією такі моделі поділяють на два типи – одноетапні (one-stage) та двоетапні (two-stage) [3]. У першому випадку вся обробка відбувається за один прохід, що забезпечує високу швидкість і робить цей тип придатним для роботи в реальному часі, як у

моделях родини YOLO або SSD [4]. Натомість двоетапні архітектури, зокрема Faster R-CNN, спочатку виділяють регіони потенційної присутності об'єктів, а потім класифікують їх окремо, що підвищує точність, але зменшує швидкодію. У контексті виявлення БПЛА, де ціль часто є малою і рухається на складному фоні, обидва типи можуть бути доцільними залежно від вимог до продуктивності та точності.

На відміну від виявлення, яке виконується незалежно в кожному кадрі, відстеження об'єктів передбачає встановлення їхньої часової послідовності та підтримання сталої ідентичності впродовж руху. Трекер формує внутрішню модель динаміки на основі попередніх спостережень, використовуючи інформацію про положення, швидкість та орієнтацію об'єкта. Спочатку він приймає початкові координати (як правило, з детектора), після чого прогнозує наступне положення на основі математичної моделі руху. Коли надходить новий кадр, прогноз зіставляється з реальними спостереженнями, і у випадку збігу модель оновлює траєкторію. Якщо об'єкт тимчасово зникає з поля зору, трекер продовжує екстраполювати його рух, доки не з'являться нові підтвердження [5].

Для реалізації такого підходу використовуються різні класи трекерів. Традиційні системи ґрунтуються на фільтрі Калмана, який добре працює для лінійних рухів з передбачуваною динамікою [6]. Для складніших сценаріїв із різкими змінами швидкості застосовуються фільтри частинок, які оперують набором гіпотез про можливі положення об'єкта, однак потребують значно більше ресурсів. Альтернативою є кореляційні трекери, що порівнюють фрагменти зображення та підтримують високу швидкість роботи, але чутливі до змін зовнішнього вигляду. Найбільш гнучкими є нейромережеві трекери, які здатні враховувати візуальні ознаки та адаптуватися до зміни форми чи освітлення, проте їхнє навчання та використання вимагає значних обчислювальних потужностей [7].

Таким чином, якісне відстеження об'єктів потребує балансу між точністю, швидкодією та стійкістю до часткової втрати об'єкта. Особливо актуально це для БПЛА, де рух є маневровим, об'єкти мають малі розміри, а фон часто містить перешкоди. Тому оптимальним підходом є поєднання швидких схем детекції з адаптивними моделями предиктивного трекінгу, що забезпечує безперервність траєкторії навіть за нестабільних умов спостереження.

1.2 Аналіз наукових праць за темою дослідження

За останні три роки питання виявлення та відстеження малогабаритних повітряних об'єктів у відеопотоці, зокрема БПЛА, стало одним із провідних напрямів досліджень у галузі комп'ютерного зору. Науковці відзначають, що основні труднощі цієї задачі пов'язані з малими розмірами цілей, значною варіативністю ракурсів, перекриттями та деградацією відеосигналу в реальних умовах [9]. У відповідь на це у сучасній літературі спостерігається тенденція до адаптації згорткових архітектур, насамперед родини YOLO, під специфіку дронів даних датасетів (VisDrone, UAVDT тощо), що підтверджується стабільними покращеннями детекції малих об'єктів у варіативних сценах [10]. У роботі модифіковано архітектуру YOLOv8 шляхом посиленого злиття ознак на різних масштабах, що дозволило покращити точність на наборі VisDrone2019. Автори підкреслюють, що найбільший приріст mAP спостерігається саме для об'єктів малої площі, завдяки оптимізації FPN-блоків.

У праці [11] запропоновано альтернативне рішення – SOD-YOLO, де увага спрямована на адаптивну вагову компенсацію масштабів. На відміну від [10], ця модель орієнтована не лише на покращення передачі дрібних ознак, але й на приглушення фонового шуму, що зменшило кількість хибних спрацьовувань на рухомих текстурах.

Ще один підхід представлено авторами [12], де застосовано локальне підсилення контурів на ранніх шарах нейронної мережі. Автори доводять, що попереднє виділення високочастотних компонентів дає змогу уникнути

«розмивання» об'єктів при згортках великого ядра, що є типовою проблемою у відео, знятому з великих відстаней.

Попри істотний прогрес у детекції, для практичного застосування критичним є збереження ідентичності та безперервності траєкторій у часі. Дослідження показують, що асоціативні багатооб'єктні трекери, такі як ByteTrack та DeepSORT, у поєднанні з детекторами останнього покоління забезпечують помітне зменшення ID-switch навіть за умов часткових перекриттів і зміни ракурсу [13, 14], а також у доменних роботах під БПЛА, де поєднання детекції й Re-ID покращує стійкість до нерівномірного руху та деградації відеосигналу [15].

Окремий пласт публікацій фокусується на прогнозуванні траєкторій БПЛА як наступному кроці після трекінгу. Поряд із класичними статистичними підходами (Калман/IMM), дедалі ширше застосовують послідовні неймережеві моделі – LSTM та Transformer [16]. Показано, що врахування історії траєкторії, контекстних ознак сцени та поведінкових намірів цілі знижує помилки прогнозу та підвищує стабільність оцінювання майбутніх положень у короткому горизонті передбачення [17], а для суміжних транспортних сценаріїв – і трансформерні предиктори траєкторій із урахуванням складної міської динаміки [18]. Водночас чимало робіт підкреслюють недостатню увагу до оцінювання невизначеності прогнозів та робастності до втрат спостережень – ці аспекти залишаються відкритими й визначають перспективні напрями подальших досліджень.

1.3 Мета та завдання дослідження

Метою роботи є розробка методу неймережевого виявлення та прогнозування траєкторій БПЛА у відеопотоці, який забезпечує стійке розпізнавання малорозмірних повітряних об'єктів на складному фоні, формування безперервних треків і короткостроковий прогноз руху з

мінімальною похибкою в умовах шумів, часткових перекриттів і втрат кадрів, придатний до інтеграції у системи моніторингу в реальному часі.

Для досягнення мети ставляться до виконання такі завдання:

- Проаналізувати сучасні підходи до виявлення, відстеження та прогнозування траєкторій БПЛА у відеопотоці, визначивши їх обмеження у контексті роботи в умовах шумів, перекриттів та малих розмірів об'єкта.
- Розробити пайплайн обробки відеоданих, що включає попередню обробку, детекцію об'єктів, асоціативний трекінг та прогнозування руху.
- Програмно реалізувати пайплайн.
- Провести експериментальне оцінювання точності прогнозування за метриками RMSE, ADE та FDE, а також зіставити статистичні й нейромережеві підходи при різних типах траєкторій.
- Реалізувати модуль візуалізації, що поєднує фактичні та прогнозовані траєкторії у вигляді карти руху та забезпечує відтворення результатів у режимі наближеному до реального часу.
- Виконати порівняльний аналіз із відомими методами на основі опублікованих реперних значень, враховуючи відмінність протоколів оцінювання.

Розділ 2 Метод неймережевого виявлення та прогнозування траєкторій БПЛА

2.1 Пайплайн неймережевого виявлення та прогнозування траєкторій БПЛА

Пайплайн на основі методу неймережевого виявлення та прогнозування траєкторій БПЛА реалізовано у вигляді послідовного конвеєра, що об'єднує модулі детекції, трекінгу, прогнозування та візуалізації результатів (рис. 2.1). На вхід системи може подаватися відеофайл, посилання на онлайн-потік або трансляція з камери у реальному часі. Кожен кадр, що надходить із джерела, передається до модуля YOLOv12 [19], який виконує виявлення дронів, повертаючи координати обмежувальних рамок і ймовірності належності до цільового класу.

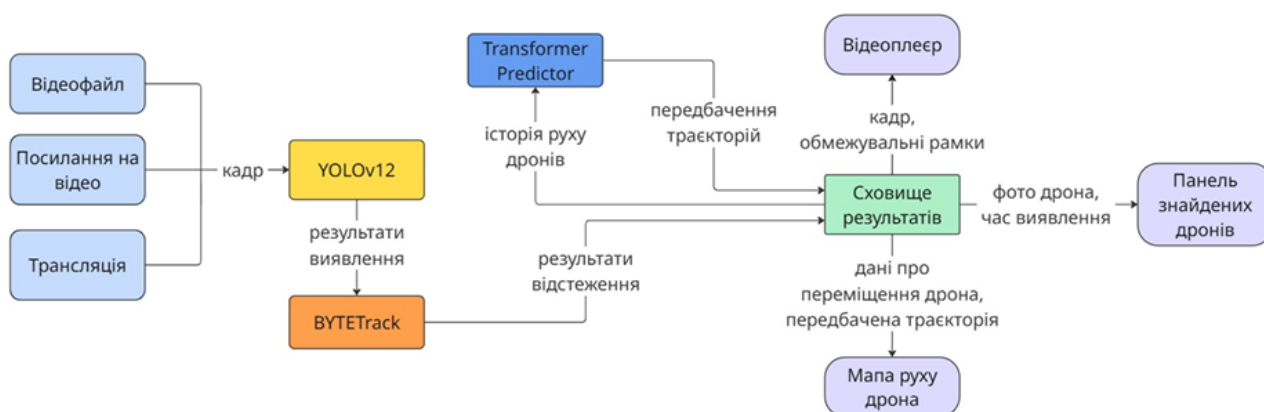


Рисунок 2.1 – Схема етапів навчання неймережевої моделі

Отримані детекції передаються до BYTETrack, який виконує асоціативне відстеження об'єктів у часовій послідовності кадрів, формуючи сталі ідентифікатори треків та історію руху кожного БПЛА. Для кожного сформованого треку накопичується послідовність координат, що використовується як вхід для модуля прогнозування Transformer Predictor. Трансформерна модель генерує передбачену траєкторію руху, враховуючи як поточне положення дрона, так і його історичний рух.

Як фактичні координати, так і прогнозні точки траєкторії передаються до сховища результатів, яке виконує функцію центрального буфера даних. У ньому накопичується фото дрона, час виявлення, геометрія траєкторії та ідентифікаційна інформація для подальшого відтворення.

Дані зі сховища паралельно передаються до двох вихідних компонентів: відеоплеєра, який виводить оригінальний відеокادر із накладеними обмежувальними рамками та карти руху дрона та фактичною/передбаченою траєкторією, де візуалізуються просторові переміщення об'єкта на координатній сітці або картографічній підкладці.

Додатково передбачено панель знайдених дронів, яка акумулює перелік активних об'єктів, їх характеристики та час останнього спостереження, що дозволяє оператору здійснювати моніторинг у режимі реального часу або аналізувати історію подій.

Таким чином, наведений пайплайн забезпечує повний цикл обробки відеопотоку від отримання сирих даних до формування прогнозованих траєкторій і їх наочної візуалізації. Інтеграція модулів детекції, трекінгу та трансформерного прогнозування у спільну архітектуру дозволяє не лише фіксувати присутність БПЛА, але й передбачати його подальший рух, що підвищує прикладну цінність інформаційної системи у задачах моніторингу та оперативного реагування.

2.2 Удосконалений підхід до прогнозування траєкторій БПЛА на основі позиційно-кодованого трансформера

Узагальнена схема підходу до прогнозування траєкторій БПЛА на основі позиційно-кодованого трансформера наведена на рисунку 2.2.

Запропонований підхід прогнозування руху БПЛА виконує передбачення наступних 60 положень об'єкта на основі попередніх 30 координат, що відповідає горизонту прогнозування у 2 секунди при частоті 30 кадрів за секунду.

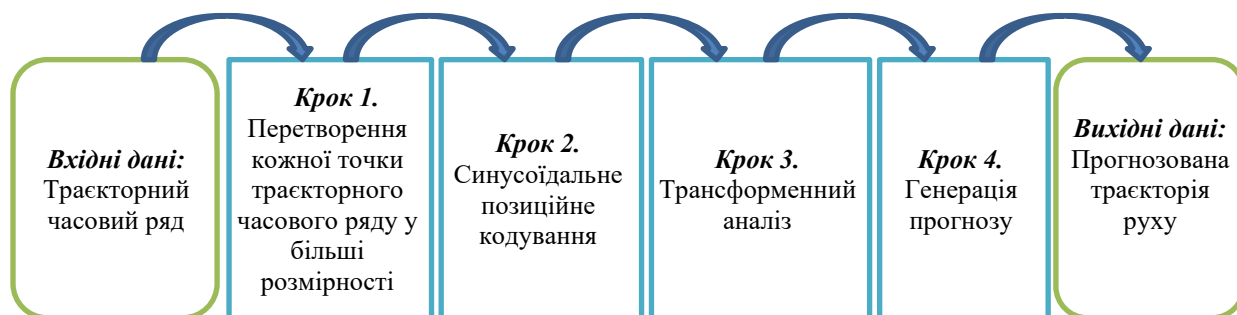


Рисунок 2.2 – Узагальнена схема підходу до прогнозування траєкторій БПЛА

Такий формат дозволяє використовувати лише одну секунду історії руху для отримання тактичного випереджувального прогнозу, придатного для систем моніторингу в реальному часі. На відміну від класичних регресійних або статистичних моделей, розроблена архітектура працює не з похідними величинами, такими як швидкість чи прискорення, а безпосередньо з координатною траєкторією, інтерпретуючи її як часову послідовність просторових спостережень. Для того щоб забезпечити здатність до узагальнення нелінійних патернів руху, кожна точка перетворюється з двовимірного представлення у вектор прихованих ознак більшої розмірності (Крок 1). Оскільки модель трансформерного типу не оперує поняттям порядку за замовчуванням, до векторних подань координат додається синусоїдальне позиційне кодування (Крок 2), що надає часову прив'язку кожному елементу послідовності. Центральним елементом системи є блок самоуваги (self-attention), який аналізує взаємозв'язки між усіма точками траєкторії одночасно (Крок 3), виявляючи як локальні зміни напрямку, так і більш довготривалі закономірності руху. Після обробки кодером усі сформовані ознаки агрегуються в єдиний контекстний вектор, який використовується як компактне відображення поточного стану об'єкта. На основі цього вектора багатошаровий перцептрон (Крок 4) генерує прогноз наступних координат, забезпечуючи плавне продовження траєкторії з урахуванням попереднього руху.

Отже, запропонований підхід дозволяє виконувати короткострокове прогнозування траєкторій БПЛА у реальному часі, забезпечуючи плавність

передбачення навіть за наявності нелінійних змін руху. Використання позиційного кодування та механізму самоуваги підвищує здатність моделі до узагальнення та робить її придатною для інтеграції у системи тактичного моніторингу та реагування.

2.3 Удосконалена архітектура на основі позиційно-кодованого трансформера

Розроблений підхід прогнозування траєкторій побудований на основі модифікованої послідовної трансформерної архітектури, адаптованої для роботи з координатними даними (рисунок 2.3). На відміну від типових NLP-трансформерів, які працюють із токенами дискретної природи, дана модель оперує числовими точками (x, y) , що описують траєкторію руху об'єкта у попередніх кадрах.

Вхідним шаром моделі є послідовність із 30 двовимірних координат, які представляють історію руху БПЛА. Оскільки сирі координати не здатні повноцінно відобразити латентні закономірності руху, кожна точка перетворюється за допомогою лінійного шару Linear ($2 \rightarrow 128$), що переводить її у високовимірний простір ознак розмірністю 128. Такий підхід дозволяє моделі виявляти складніші траєкторні патерни, недоступні при прямій обробці координат.

Для того щоб трансформер міг розпізнавати порядок елементів у послідовності, до ознак додається синусоїдальне позиційне кодування, яке задає часову прив'язку кожної точки. Основну обчислювальну частину становить триступеневий блок Transformer Encoder, кожен шар якого містить 4 голови самоуваги (multi-head attention), feedforward-проекцію з розмірністю 256 та dropout 0.2 для регуляризації. Завдяки механізму самоуваги модель одночасно аналізує як локальні зміни швидкості, так і довготривалі тренди руху, що забезпечує робастність до маневрів та різких змін напрямку.

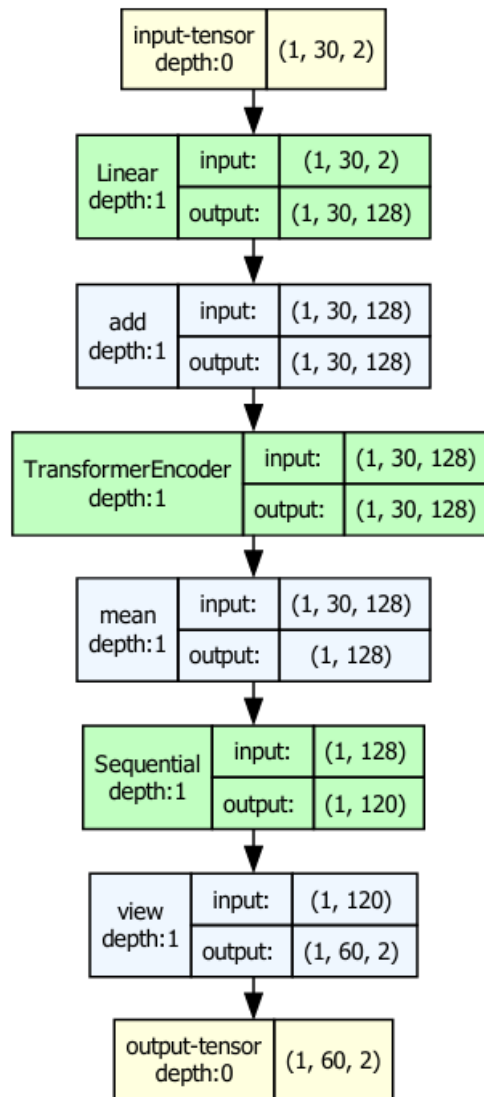


Рисунок 2.3 – Узагальнена архітектура нейромережі

На виході кодера формується тривимірний тензор memory, який агрегується у контекстний вектор шляхом усереднення ознак по часовій осі. Такий варіант агрегації значно спрощує структуру моделі порівняно з повноцінним трансформерним декодером і забезпечує низьку обчислювальну затримку, що є критично важливим для роботи у реальному часі.

Фінальний етап – декодер на основі багатошарового перцептрона. Він складається з LayerNorm, двох блоків $\text{Linear} \rightarrow \text{ReLU} \rightarrow \text{Dropout}$, після чого завершується вихідним Linear, який генерує 120 чисел, що представляють 60 прогнозованих координат майбутньої траєкторії.

Отже, удосконалена архітектура на основі позиційно-кодованого трансформера забезпечує повноцінну обробку координатних послідовностей без потреби у проміжному обчисленні похідних параметрів руху. Агрегація контексту через спрощений енкодер дозволяє зберегти високу прогностичну здатність при мінімальній обчислювальній затримці, що робить підхід сумісним із вимогами систем реального часу.

2.4 Набір даних дослідження

Набір даних «Anti-Unmanned Aerial Vehicle» [20] призначений для дослідження задач виявлення та відстеження безпілотних літальних апаратів у відеопотоці та містить записи, отримані в реальних умовах спостереження. Він охоплює широкий спектр сцен із різноманітними типами фонів, рівнями освітлення та кутами зйомки, що забезпечує репрезентативність для практичних сценаріїв моніторингу повітряного простору. Кожен кадр у наборі має точну анотацію у вигляді обмежувальних рамок, які вказують положення дронів, що дозволяє використовувати цей датасет як для детекції, так і для трекінгу. Особливістю набору є наявність записів як у видимому, так і в інфрачервоному спектрах, однак у межах даної роботи використовується лише підмножина відео у видимому діапазоні.

Структурно набір поділено на тренувальну, валідаційну та тестову частини, що містять відповідно 160, 67 та 91 відеозапис – загалом 318 відеофрагментів, які разом налічують 593 802 кадри. Така кількість спостережень забезпечує можливість повноцінного навчання глибоких моделей без додаткового штучного розширення вибірки.

Важливою властивістю датасета є цілеспрямована наявність у ньому складних сценаріїв спостереження, зокрема випадків тимчасового зникнення об'єкта з поля зору, змін масштабу та освітлення, часткових або повних перекриттів, низької роздільної здатності цілі та швидких переміщень, що призводять до значних міжкадрових зсувів. Це створює умови, наближені до

реального розгортання систем моніторингу, та дає змогу оцінити стабільність алгоритмів не лише за середніми показниками точності, а й у критичних ситуаціях, де типовий трекінг втрачає ефективність. Таким чином, «Anti-UAV Dataset» може розглядатися як повноцінна база для навчання та об'єктивного тестування методів виявлення, супроводу та прогнозування траєкторій БПЛА.

Розділ 3 Дослідження ефективності методу неймережевого виявлення та прогнозування траєкторій БПЛА

3.1 Структура та функціональне призначення програмних складових системи

UML-діаграма класів (рисунки 3.1–3.2) відображає структуру інформаційної системи, створеної на основі методу для реалізації у середовищі Python з використанням фреймворка Kivy [21]. Центральним елементом є клас DroneTrackingApp, який наслідує App із Kivy та виконує роль координатора роботи всіх компонентів, керує інтерфейсом користувача, обробкою відео та викликає послідовність етапів виявлення і трекінгу.

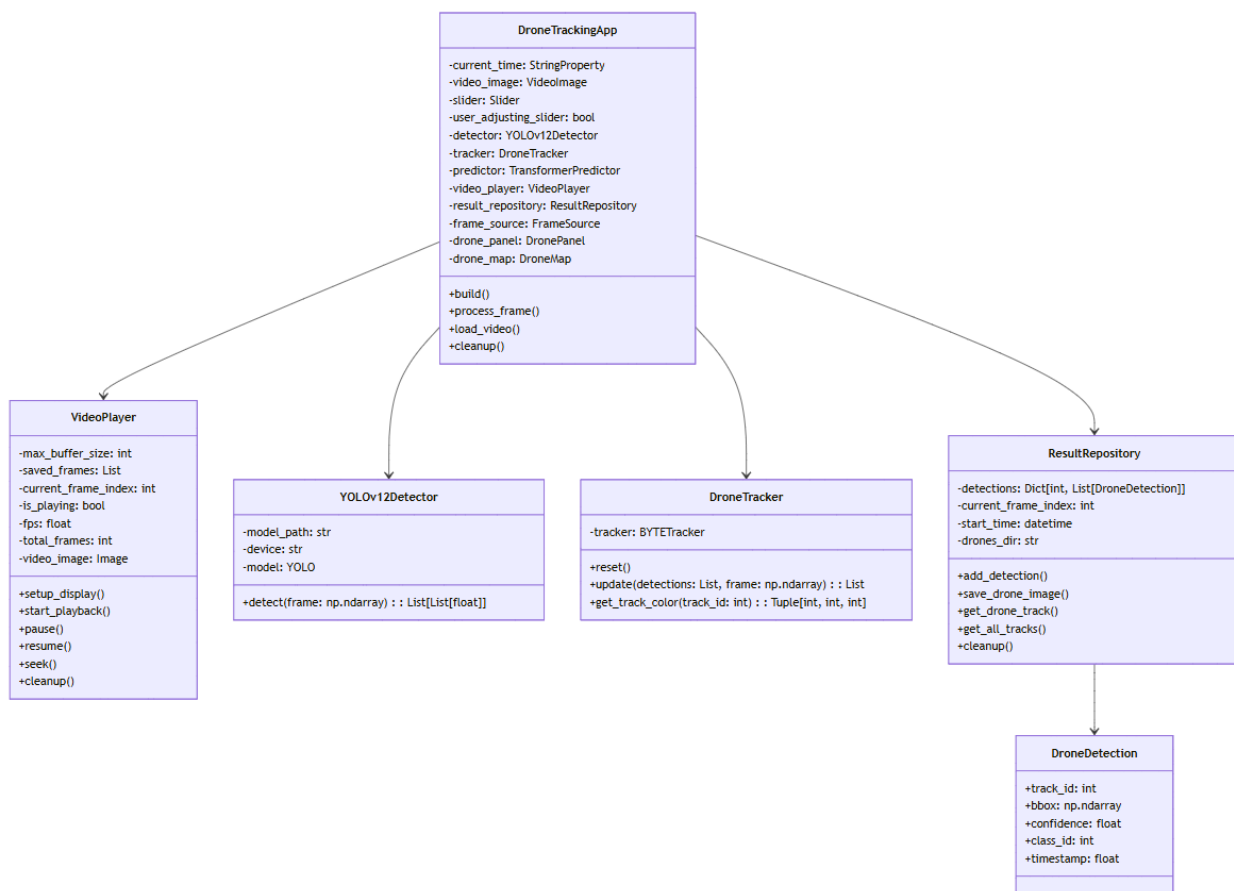


Рисунок 3.1 – Діаграма класів системи. Частина 1

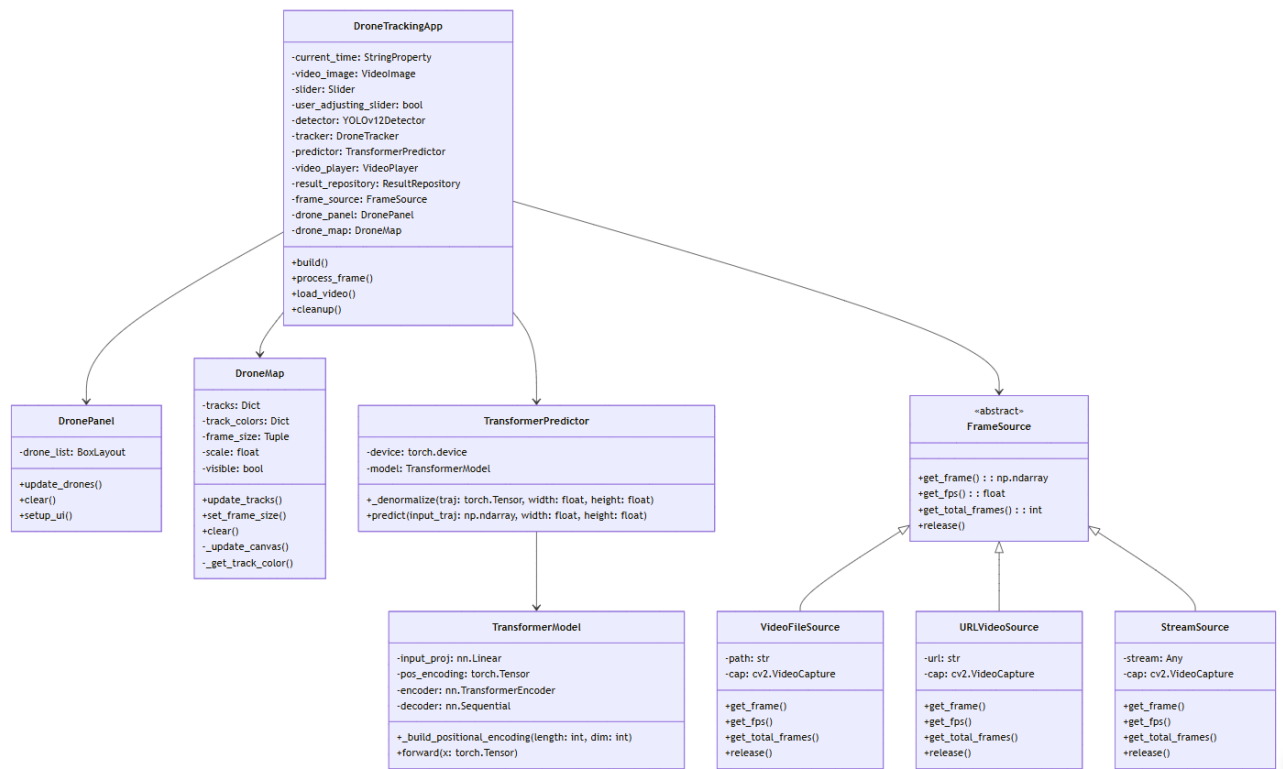


Рисунок 3.2 – Діаграма класів системи. Частина 2

За відтворення відео відповідає VideoPlayer, який забезпечує завантаження, перемикання кадрів і синхронізацію зі станом обробки. Для детекції використовується клас YOLOv12Detector, що інкапсулює фінальну модель YOLOv12, керує вибором пристрою (CPU / GPU) та повертає результати у стандартизованому форматі. Відстеження об'єктів виконує DroneTracker, що реалізує алгоритм BYTETrack, присвоює ідентифікатори, підтримує історію руху та містить службові методи на кшталт скидання стану чи вибору кольорів треків.

Збереження та управління результатами забезпечує ResultRepository, де накопичуються кадри, зображення дронів і їхня історія виявлень. Окремі екземпляри детекцій представлені класом DroneDetection, який уніфіковано зберігає ідентифікатор треку, координати рамки, час і впевненість.

Візуальні компоненти інтерфейсу реалізовано через DronePanel, який формує динамічний список знайдених дронів, та DroneMap, що відображає траєкторії руху на координатній сітці з урахуванням масштабування між відео

та простором відображення.

Джерела відео абстраговано через базовий клас `FrameSource`, а конкретні реалізації, `VideoFileSource`, `URLVideoSource` та `StreamSource`, забезпечують роботу відповідно з локальними файлами, потоковим відео за посиланням або трансляцією.

Отже, представлена UML-архітектура демонструє чіткий поділ системи на модулі обробки відео, виявлення, трекінгу, збереження результатів та візуалізації, що забезпечує гнучкість розширення і можливість подальшої інтеграції додаткових функціональних блоків без зміни базової логіки роботи застосунку.

3.2 Результати досліджень

Оцінювання продуктивності модуля детекції на основі YOLOv12 показало, що модель здатна стабільно розпізнавати малогабаритні повітряні об'єкти навіть у присутності шуму, зміни масштабу та варіативного фону. Високе значення mAP50 на тестовій вибірці свідчить про точне позиціонування обмежувальних рамок, тоді як співвідношення mAP50-95 демонструє, що модель зберігає адекватну чутливість навіть при посиленні критеріїв відповідності (таблиця 3.1). Робастність до дрібних цілей підтверджується тим, що кількість хибнонегативних спрацьовувань суттєво не зростає у випадках, коли площа обмежувальної рамки не перевищує кілька сотень пікселів.

Таблиця 3.1 – Метрики виявлення дронів моделю YOLOv12

mAP50 (%)			mAP50-95 (%)	FPS (CPU)	FPS (GPU)	GFLOPs
трен	валід	тест	62.76	13	55	6.3
99.05	94.65	91.04				

Матриця результатів класифікації (рисунок 3.3) відображає типову ситуацію для задач виявлення об'єктів, у яких фон не розглядається як окремий

клас. Через те, що потенційна варіативність фонових структур на одному зображенні є надто високою для детального маркування, несуттєві ділянки сцени традиційно не ануються, а лише присутні в наборі як контекст. Відповідно, випадки, що формально виглядають як помилкове віднесення фону до класу «дрон», не обов’язково свідчать про хибне спрацювання моделі. Значна частина таких спостережень відповідає об’єктам зі схожими структурними або силуетними ознаками, зокрема літальним апаратам або дроноподібним конструкціям, які в реальних умовах також можуть становити інтерес для системи моніторингу. За таких умов поведінка моделі радше демонструє її чутливість до потенційно релевантних цілей, ніж випадкову генерацію помилок. Подальше уточнення меж між цільовими та другорядними об’єктами може бути досягнуте шляхом включення до навчального набору окремої групи «суміжних» класів, однак навіть у поточному вигляді модель проявляє збалансовану здатність до узагальнення без тенденції до систематичного перенавчання на конкретні фонові патерни.

		Істинне	
		дрон	фон
Передбачене	дрон	5689	993
	фон	653	-

Рисунок 3.3 – Матриця помилок моделі YOLOv12 на тестовому наборі

Показник mAP50 (рисунок 3.4) демонструє стабільне зростання протягом перших десятків епох навчання. На початковому етапі спостерігається різке підвищення точності, після чого крива поступово вирівнюється. Близько 30–40

епох значення метрики виходить на плато й надалі змінюється лише незначними коливаннями. Це свідчить про те, що модель досягає стійкої якості детекції без ознак перенавчання, а подальше тренування практично не впливає на результат.

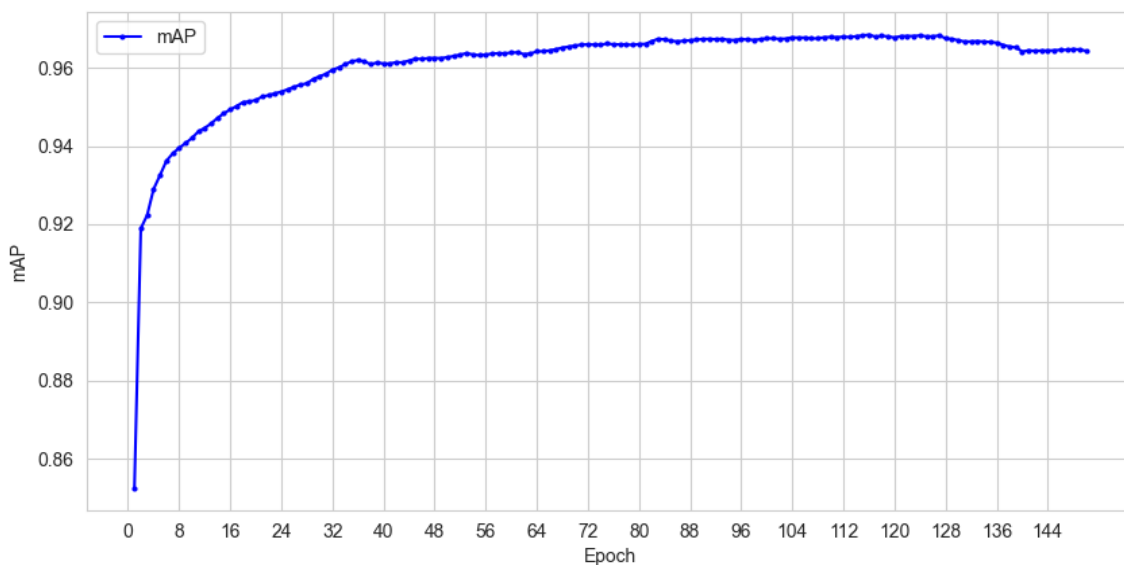


Рисунок 3.4 – Графік зміни метрики mAP50 на валідаційній вибірці за епохами

Графік зміни значення функції втрат наводить стрімке зниження похибки на початкових етапах навчання (рисунок 3.5).

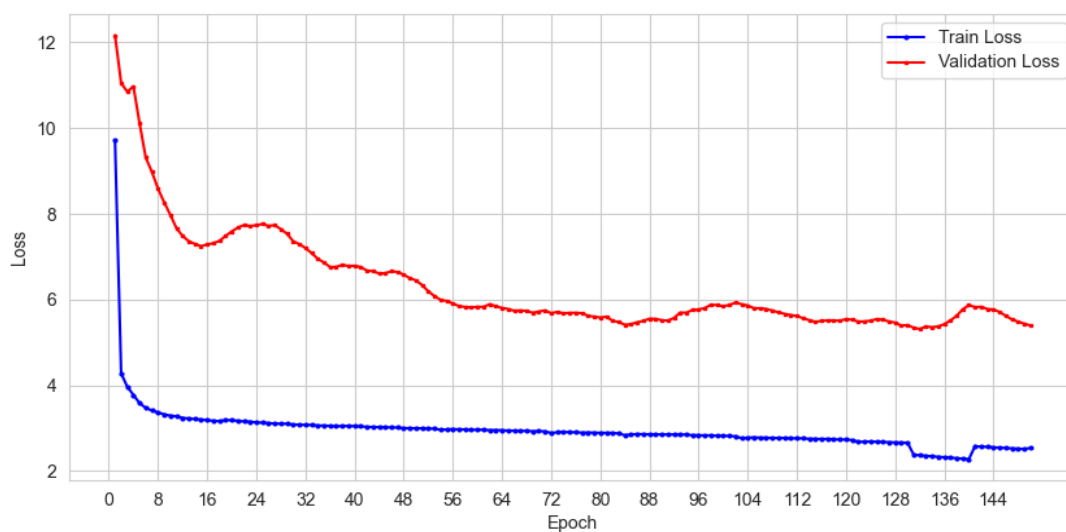


Рисунок 3.5 – Графік зміни значення функції втрат на тренувальній та валідаційній вибірках за епохами

Після стабілізації кривої на тренувальній вибірці спостерігається поступове згладжування коливань на валідаційній, що свідчить про досягнення збалансованого режиму роботи моделі без явних ознак перенавчання. Незначне розходження між кривими loss-вибірок вказує на притаманну задачі варіативність візуальних умов, але не впливає суттєво на загальну якість трекінгу.

Наступним етапом є дослідження можливостей відстеження здетектованих об'єктів. Результати оцінювання продуктивності трекера BYTETrack на тестовій вибірці демонструють достатньо високу узгодженість між виявленнями та сформованими траєкторіями. Значення MOTA = 0.79 свідчить про те, що більшість об'єктів зберігають свою ідентичність упродовж спостереження навіть за наявності часткових перекриттів або змін масштабу (таблиця 3.2). Показник IDF1 = 0.64 підтверджує помірний рівень точності зіставлення між детекціями та вже існуючими треками, що є характерним для сценаріїв з нерівномірною швидкістю руху. Висока швидкодія (332 кадри за секунду на CPU) дає підстави вважати, що обрана конфігурація алгоритму може використовуватися у системах реального часу без додаткової оптимізації.

Таблиця 3.2 – Метрики відстеження дронів тестової вибірки алгоритмом BYTETrack

MOTA	IDF1	FPS (CPU)
0.79	0.64	332

На рисунку 3.6 представлено криву тренувальних та валідаційних втрат протягом 200 епох навчання. Спостерігається стрімке зниження значення функції втрат на початкових етапах тренування (0–20 епох), після чого динаміка поступово стабілізується. Навчальна вибірка демонструє монотонне зменшення похибки аж до значення ≈ 0.05 , що свідчить про ефективне засвоєння закономірностей у траєкторіях. Валідаційна крива після первинного зниження залишається стабільною в межах 0.18–0.22, без значних коливань чи

різких стрибків, що підтверджує відсутність перенавчання. Механізм раннього зупинення не активувався передчасно, що означає стабільну динаміку покращень аж до фінальних епох. Слід зазначити, що тенденція збіжності втрат інтерпретується в нормованому просторі координат: усі позиційні значення були попередньо масштабовані в інтервал $[0,1]$ відносно ширини та висоти зображення, що забезпечує уніфіковану систему вимірювання та дозволяє порівнювати динаміку навчання незалежно від реальних масштабів сцени. Саме тому значення похибок у діапазоні $0.05\text{--}0.20$ на даному етапі відповідають малим відхиленням у нормалізованих координатах, тоді як у наступних розділах, де результати подано в пікселях, абсолютні значення можуть виглядати суттєво більшими, хоча фактично описують той самий рівень точності.



Рисунок 3.6 – Графік тренування трансформера

Підсумкові метрики, наведені в таблиці 3.3, демонструють середню кореневу середньоквадратичну помилку (RMSE) 12.66 пікселів, що відповідає точності рівня $\sim 1\text{--}2\%$ від типової ширини кадру. Метрики $ADE = 14.1$ та $FDE = 17.4$ пікселів вказують на те, що модель здатна не лише загалом правильно

оцінювати напрямок руху, але й досить точно прогнозувати кінцеву позицію у 2-секундному горизонті. Обчислювальна складність моделі становить лише 0.03 GFLOPs, що робить її придатною для розгортання на пристроях з обмеженими ресурсами та інтеграції у системи реального часу.

Таблиця 3.3 – Метрики трансформера в пікселях на тестовому наборі

RMSE	ADE	FDE	GFLOPs
12.66	14.1	17.4	0.03

Таким чином, отримані результати підтверджують, що модифікована трансформерна архітектура є здатною до стійкого короткострокового прогнозування траєкторій БПЛА, зберігаючи баланс між точністю, стабільністю та продуктивністю. У порівнянні з результатами класичних статистичних фільтрів (IMM/Kalman), що наведені в таблиці 3.4, модель демонструє суттєво меншу помилку при маневрових траєкторіях, що робить її перспективною для практичного використання у системах ситуаційної обізнаності.

Таблиця 3.4 – Метрики IMM/Kalman в пікселях на тестовому наборі

RMSE	ADE	FDE
126.27	161.1	250.95

У інформаційну систему інтегровано модуль передбачення траєкторій, що використовуючи трансформер, передбачає траєкторію дрона на 60 кадрів вперед. Результат виводиться на мапу руху дрона (приклади на рисунку 3.7).

Приклади роботи підходу до прогнозування траєкторій руху БПЛА наведено на рисунку 3.8.

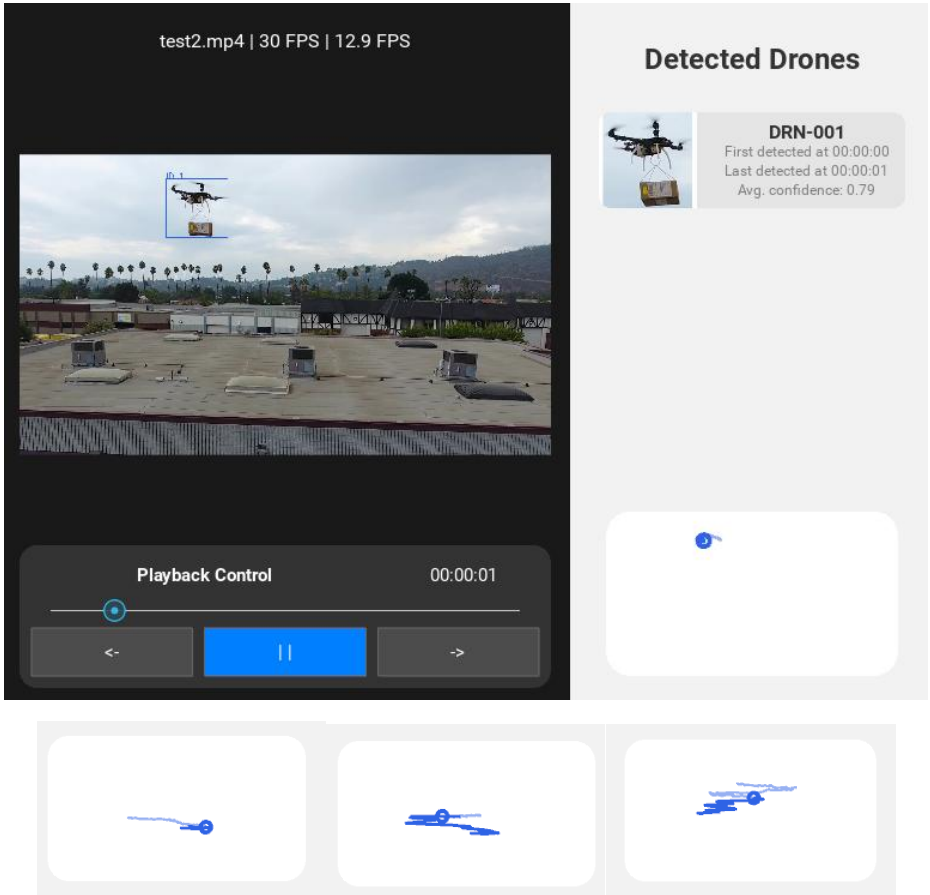


Рисунок 3.7 – Результат роботи програми

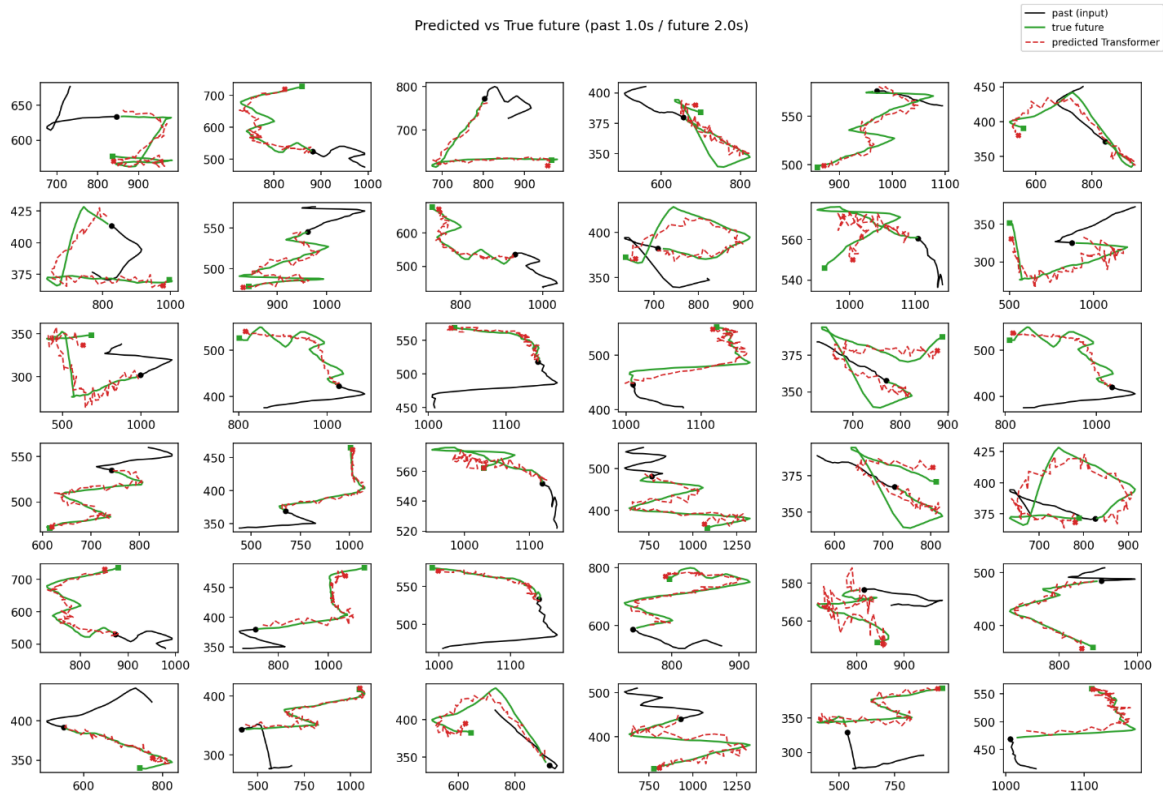


Рисунок 3.8 – Приклади роботи підходу до прогнозування траєкторій руху БПЛА

Візуально простежується, що модель здатна відтворювати як поступальні, так і маневрові типи руху з незначним відхиленням від реальної траєкторії, зберігаючи як загальний напрямок, так і локальну форму кривої. У більшості випадків розбіжність між передбаченим та фактичним шляхом зростає плавно, що свідчить про стабільну поведінку моделі навіть за умов різких змін напрямку.

За підсумком порівняння з відомими аналогами, запропонований метод демонструє вищу точність за метрикою MOTA, ніж більшість сучасних рішень для відстеження БПЛА. Наприклад, у проведеному дослідженні отримано значення метрики MOTA = 0.79, тоді як у роботі [22] повідомляється MOTA = 0.461.

Для коректного співставлення з роботами, у яких похибка прогнозування подається у метрах, слід врахувати відмінність форматів представлення координат. У поданому експерименті траєкторії оцінювалися в пікселях з нормалізацією відносно ширини кадру, тоді як у роботах на кшталт [23] похибка вимірюється у фізичному просторі. За типовим масштабом сцени у наборі Anti-UAV один метр у площині спостереження відповідає приблизно 40-60 пікселям (залежно від висоти польоту та кута огляду). За такого припущення значення RMSE = 12.66 px інтерпретується як приблизно 0.25-0.32 м, що є співмірним із похибкою 0.214-0.41 м, наведеною в [23]. Це вказує на те, що навіть за різних форматів оцінювання (зображення проти сенсорних координат), точність прогнозу перебуває у діапазоні сучасних моделей, адаптованих до фізичного простору.

Додатковою перевагою є висока швидкодія 332 FPS на CPU, що значно перевищує типові показники у відкритих реалізаціях мультиоб'єктного трекінгу, де продуктивність рідко перевищує 30-50 FPS. Це дозволяє віднести метод до категорії придатних для розгортання в умовах реального часу, зокрема в системах повітряного моніторингу або протидії БПЛА.

Висновки

У роботі розроблено метод нейромережевого виявлення та прогнозування траєкторій БПЛА у відеопотоці, який поєднує детекцію, трекінг та трансформерне прогнозування в єдиному конвеєрі обробки. Побудована система здатна формувати безперервні траєкторії руху малорозмірних об'єктів у складних умовах спостереження та генерувати короткостроковий прогноз з похибкою RMSE 12.66 пікселів (близько діапазону 0.25-0.32 м у фізичному просторі). Для підвищення точності системи передбачено використання глибинних архітектур прогнозування типу Transformer, що забезпечує адаптивність до різних типів траєкторій.

Проведені експерименти показали, що запропонований підхід забезпечує значення метрик MOTA = 0.79 та IDF1 = 0.64, що перевищує опубліковані показники більшості сучасних рішень на основі статистичних моделей та наближає систему до рівня, придатного для розгортання у реальному часі. Швидкодія 332 FPS на CPU підтверджує можливість інтеграції методу у практичні системи моніторингу.

Попри досягнуті результати, запропонований підхід має певні експериментальні обмеження, притаманні більшості робіт цього класу. Прогнозування здійснюється у двовимірних координатах зображення, що відповідає умовам стаціонарного спостереження з фіксованої камери, тому сценарії з активним переміщенням платформи потребують додаткової адаптації. У межах роботи розглядається лише короткостроковий горизонт передбачення, такий підхід є виправданим для систем оперативного моніторингу, а збільшення інтервалу прогнозу вимагатиме поєднання з поведінковими моделями або ймовірнісними оцінками невизначеності. Окремим напрямом розвитку є адаптація моделі до різних типів динаміки від поступального руху до складних маневрових режимів, що може бути реалізовано через попередню класифікацію траєкторій або багатомодельне прогнозування. Дані обмеження будуть опрацьовані в майбутніх дослідженнях.

Література

1. A Review of Video Object Detection: Datasets, Metrics and Methods / H. Zhu et al. *Applied Sciences*. 2020. Vol. 10, no. 21. P. 7834. URL: <https://doi.org/10.3390/app10217834> (date of access: 07.10.2025).
2. Hybrid multi-attention transformer for robust video object detection / S. Moorthy et al. *Engineering Applications of Artificial Intelligence*. 2025. Vol. 139. P. 109606. URL: <https://doi.org/10.1016/j.engappai.2024.109606> (date of access: 07.10.2025).
3. Synergistic Integration of Transfer Learning and Deep Learning for Enhanced Object Detection in Digital Images / S. R. Waheed et al. *IEEE Access*. 2024. P. 1. URL: <https://doi.org/10.1109/access.2024.3354706> (date of access: 07.10.2025).
4. Vijayakumar A., Vairavasundaram S. YOLO-based Object Detection Models: A Review and its Applications. *Multimedia Tools and Applications*. 2024. URL: <https://doi.org/10.1007/s11042-024-18872-y> (date of access: 07.10.2025).
5. Multiple object detection and tracking from drone videos based on GM-YOLO and multi-tracker / Y. Yuan et al. *Image and Vision Computing*. 2024. P. 104951. URL: <https://doi.org/10.1016/j.imavis.2024.104951> (date of access: 07.10.2025).
6. Jiang L., Wu L. Enhanced Yolov8 network with Extended Kalman Filter for wildlife detection and tracking in complex environments. *Ecological Informatics*. 2024. P. 102856. URL: <https://doi.org/10.1016/j.ecoinf.2024.102856> (date of access: 07.10.2025).
7. Correlation Filter based Single Object Tracking: A Review / A. Kumar et al. *Information Fusion*. 2024. P. 102562. URL: <https://doi.org/10.1016/j.inffus.2024.102562> (date of access: 07.10.2025).
8. CorrFormer: Context-aware tracking with cross-correlation and transformer / J. Zhang et al. *Computers and Electrical Engineering*. 2024. Vol. 114. P. 109075. URL: <https://doi.org/10.1016/j.compeleceng.2024.109075> (date of access: 07.10.2025).

9. Xu L., Luo Z. Anti-UAV detection and identification technology: Fundamentals, methods and challenges. *Physical Communication*. 2025. Vol. 71. P. 102676. URL: <https://doi.org/10.1016/j.phycom.2025.102676> (date of access: 07.10.2025).
10. A Small-Object Detection Model Based on Improved YOLOv8s for UAV Image Scenarios / J. Ni et al. *Remote Sensing*. 2024. Vol. 16, no. 13. P. 2465. URL: <https://doi.org/10.3390/rs16132465> (date of access: 07.10.2025).
11. SOD-YOLO: Small-Object-Detection Algorithm Based on Improved YOLOv8 for UAV Images / Y. Li et al. *Remote Sensing*. 2024. Vol. 16, no. 16. P. 3057. URL: <https://doi.org/10.3390/rs16163057> (date of access: 07.10.2025).
12. SRM-YOLO for Small Object Detection in Remote Sensing Images / B. Yao et al. *Remote Sensing*. 2025. Vol. 17, no. 12. P. 2099. URL: <https://doi.org/10.3390/rs17122099> (date of access: 07.10.2025).
13. ByteTrack: Multi-object Tracking by Associating Every Detection Box / Y. Zhang et al. *Lecture Notes in Computer Science*. Cham, 2022. P. 1–21. URL: https://doi.org/10.1007/978-3-031-20047-2_1 (date of access: 07.10.2025).
14. Wojke N., Bewley A., Paulus D. Simple online and realtime tracking with a deep association metric. *2017 IEEE International Conference on Image Processing (ICIP)*, Beijing, 17–20 September 2017. 2017. URL: <https://doi.org/10.1109/icip.2017.8296962> (date of access: 07.10.2025).
15. High-Performance Detection-Based Tracker for Multiple Object Tracking in UAVs / X. Li et al. *Drones*. 2023. Vol. 7, no. 11. P. 681. URL: <https://doi.org/10.3390/drones7110681> (date of access: 07.10.2025).
16. Lee J. K. Transformer-Based Vehicle-Trajectory Prediction at Urban Low-Speed T-Intersection. *Sensors*. 2025. Vol. 25, no. 14. P. 4256. URL: <https://doi.org/10.3390/s25144256> (date of access: 07.10.2025).
17. Multiple Object Tracking in Drone Aerial Videos by a Holistic Transformer and Multiple Feature Trajectory Matching Pattern / Y. Yuan et al. *Drones*. 2024. Vol. 8, no. 8. P. 349. URL: <https://doi.org/10.3390/drones8080349> (date of access: 07.10.2025).

18. Trajectory Planning of Unmanned Aerial Vehicles in Complex Environments Based on Intelligent Algorithm / Z. Cheng et al. *Drones*. 2025. Vol. 9, no. 7. P. 468. URL: <https://doi.org/10.3390/drones9070468> (date of access: 07.10.2025).

19. YOLOv12: State-of-the-Art Object Detection Model. *YOLOv12: State-of-the-Art Object Detection Model*. URL: <https://yolov12.com/> (date of access: 07.10.2025).

20. Anti-UAV: A Large-Scale Benchmark for Vision-Based UAV Tracking / N. Jiang et al. *IEEE Transactions on Multimedia*. 2021. P. 1. URL: <https://doi.org/10.1109/tmm.2021.3128047> (date of access: 07.10.2025).

21. ...

...

... *Наука і техніка сьогодні*. – 2025. ...

... DOI: <https://doi.org/>...

...

22. A Multi-Scale Feature-Fusion Multi-Object Tracking Algorithm for Scale-Variant Vehicle Tracking in UAV Videos / S. Liu et al. *Remote Sensing*. 2025. Vol. 17, no. 6. P. 1014. URL: <https://doi.org/10.3390/rs17061014> (date of access: 07.10.2025).

23. UAV Trajectory Prediction Based on Flight State Recognition / J.-d. Zhang et al. *IEEE Transactions on Aerospace and Electronic Systems*. 2023. P. 1–16. URL: <https://doi.org/10.1109/taes.2023.3303854> (date of access: 07.10.2025).