



МЕТОД ВИОКРЕМЛЕННЯ ВІЙСЬКОВИХ ОБ'ЄКТІВ З ВІДЕОПОТОКУ ДАНИХ FPV-ДРОНА З ВИКОРИСТАННЯМ МОДЕЛЕЙ МАШИННОГО НАВЧАННЯ

Проаналізовано наявні методи комп'ютерного зору та засоби машинного навчання для вирішення завдання автоматизованого виявлення та класифікації військових цілей безпілотними літальними апаратами (БПЛА). Обґрунтовано актуальність та вагомість застосування безпілотних засобів, зокрема – FPV-дронів (англ. *First Person View*), для виконання різноманітних бойових завдань за умов ведення сучасних бойових дій, зокрема – в контексті повномасштабного вторгнення Російської Федерації в Україну. Акцентовано увагу на проблемі вразливості БПЛА до засобів радіоелектронної боротьби, що виникла внаслідок необхідності постійного зв'язку з оператором. Окреслено одне з перспективних рішень, що базується на впровадженні засобів штучного інтелекту та комп'ютерного зору для забезпечення автономності безпілотного комплексу на завершальній стадії бойової місії, де ключовим елементом є програмне забезпечення (ПЗ) автоматизованого виокремлення цілей із відеопотоку даних з бортових сенсорів дрона. Охарактеризовано вимоги до програмного компонента бортового комплексу: оптимізація під обмежені обчислювальні ресурси периферійних бортових пристроїв, високі параметри точності виокремлення цілей, підвищені вимоги до надійності, а також мінімізація енергоспоживання обчислювального модуля як критична умова збереження автономності та польотних характеристик. Реалізовано метод, що базується на моделі машинного навчання YOLO26s та алгоритмі міжкадрового підтвердження виокремлень військових об'єктів, який формує результат на основі стійких сигналів послідовних кадрів відеопотоку даних з бортових сенсорів та виконує фільтрування хибнопозитивних результатів без істотного впливу на продуктивність. Сформовано навчальний набір даних на основі реальних фото- та відеоматеріалів, що охоплює шість класів військових об'єктів: танк, БТР-БМП, вантажівка, РСЗВ, ЗРК та комбатант. Описано використання різних методів аугментації для імітації роботи бюджетних сенсорів, умов навколишнього середовища та маскування шуканих цілей. За допомогою імплементованого методу досягнуто високих показників точності (середній mAP@50 становить $0,803^{+0,05}$, mAP@50-95 становить $0,624^{+0,06}$) за середньої тривалості інференсу 22^{+2} мс або 40^{+5} кадрів на секунду для відеопотоку даних з бортових сенсорів із роздільною здатністю 1920×1080 пікселів для апаратного комплексу Tesla T4. Апаратне забезпечення серії NVIDIA Jetson Orin запропоновано для розгортання розробленого програмного методу на борту FPV-дрона.

Ключові слова: машинне навчання моделей; виокремлення об'єктів; комп'ютерний зір; модель YOLO; міжкадрове підтвердження виокремлень; FPV-дрон; інференс моделі; військова техніка; бортове програмне забезпечення.

Вступ / Introduction

Повномасштабне вторгнення Російської Федерації в Україну докорінно змінило характер ведення сучасних бойових дій, загостривши потребу в актуальних розвідувальних даних, нанесенні високоточних ударів на значні відстані та масовому застосуванні економічно ефективних засобів ураження за умови мінімізації ризиків для особового складу. Сукупність цих чинників зумовила стрімке зростання ролі безпілотних літальних апаратів (БПЛА), які дають можливість виконувати небезпечні завдання дистанційно.

Серед широкого спектра безпілотних систем особливу ефективність продемонстрували FPV-дрони (англ. *First Person View*) – малогабаритні безпілотні апарати quadro- або гексакоптерної схеми, що керуються оператором від першої особи через відеопотік даних з борто-

вої камери в режимі реального часу. Завдяки високій маневреності, мінімальній радіолокаційній помітності та низькій вартості, FPV-дрони застосовують для виконання широкого спектра бойових завдань: ударних операцій, тактичної розвідки, коригування вогню та логістичного забезпечення передових позицій.

Вагомим недоліком використання FPV-дронів для виконання бойових завдань є необхідність у постійному зв'язку із оператором, що породжує істотні проблеми: вразливість до засобів радіоелектронної боротьби, підвищення ризику для операторів через потребу фізичного перебування у зоні ворожої протидії та обмеженість кількості одночасно задіяних апаратів через потребу в індивідуальному ручному управлінні. Одним із перспективних рішень є впровадження технологій штучного інтелекту та засобів комп'ютерного зору для автономного виокремлення цілей безпосередньо на борту

© Copyright 2026 by the author(s)

Здебський Степан Володимирович, студент, кафедра програмного забезпечення.

Email: stepan.zdebskyi.mnpzm.2024@lpnu.ua; <https://orcid.org/0009-0000-4454-2539>

Сенів Максим Михайлович, канд. техн. наук, доцент, кафедра програмного забезпечення.

Email: maksym.m.seniv@lpnu.ua; <https://orcid.org/0000-0003-1044-4628>

Цитування за ДСТУ: Здебський С. В., Сенів М. М. Метод виокремлення військових об'єктів з відеопотоку даних FPV-дрона з використанням моделей машинного навчання. Науковий вісник НЛТУ України. 2026, т. 36, № 3. С. 26–35.

Citation APA: Zdebskyi, S. V., & Seniv, M. M. (2026). Method for extracting military objects from an FPV drone video stream using machine learning models. *Scientific Bulletin of UNFU*, 36(3), 26–35. <https://doi.org/10.36930/40360303>

апарата, забезпечуючи виконання остаточної стадії бойової місії без участі оператора та незалежно від стану каналів зв'язку. Програмні засоби із заданим набором функціоналу мають дві основні групи вимог. Перша група відповідає за вимоги точності виокремлення об'єктів. Наявне ПЗ має забезпечувати високу достовірність виокремлення військових об'єктів із мінімальним рівнем хибнопозитивних результатів, стійке супроводження цілей за умов їх замаскованості, часткового перекриття, різних параметрів погоди та ландшафту. Друга група містить нефункціональні вимоги продуктивності, що окреслені фізичними обмеженнями: мінімальне енергоспоживання, маса та габарити апаратного модуля. Це зумовлює низький рівень обчислювальної потужності бортового апаратного забезпечення, доступної для бортового ПЗ.

Сучасні архітектури згорткових нейронних мереж, зокрема – сімейство моделей YOLO, демонструють значну ефективність у вирішенні завдань виокремлення та класифікації літальних об'єктів у режимі реального часу завдяки оптимізованій архітектурі з єдиним обходом мережі нейронів, що забезпечує низьку обчислювальну складність інференсу зі збереженням достатньої точності виокремлення об'єктів. Поєднання подібних моделей із алгоритмом міжкадрового підтвердження виокремлень літальних об'єктів, що накопичує та верифікує результати виокремлення об'єктів впродовж послідовних кадрів відеопотоку даних з бортових сенсорів, дає змогу додатково підвищити стійкість до хибнопозитивних спрацювань без істотного приросту обчислювального навантаження.

Актуальним завданням є розроблення методів виявлення та класифікації військових об'єктів на відеопотоці даних на основі моделей машинного навчання та алгоритму міжкадрового підтвердження виокремлень у контексті сукупності вимог щодо удосконалення ПЗ під бортове апаратне забезпечення, параметрів точності і стійкості відстеження цілей.

Об'єкт дослідження – виокремлення військових об'єктів з відеопотоку даних FPV-дрона.

Предмет дослідження – методи і засоби виокремлення військових об'єктів з відеопотоку даних FPV-дрона, які дадуть змогу формалізувати вибір конфігурації програмно-апаратних засобів їх реалізації як задачу багатокритеріальної оптимізації.

Мета роботи – запропонувати метод виокремлення військових об'єктів з відеопотоку даних FPV-дрона з використанням моделей машинного навчання, що дасть змогу підвищити стійкість ПЗ до хибнопозитивних спрацювань у відеопотоці даних з бортових сенсорів без істотних втрат їх продуктивності за умов обмежених обчислювальних ресурсів.

Для досягнення зазначеної мети визначено такі основні завдання дослідження:

- виконати аналіз сучасних методів комп'ютерного зору та моделей машинного навчання для завдання виокремлення літальних об'єктів у відеопотоці даних з бортових сенсорів у контексті визначених програмно-апаратних вимог, що дасть змогу обрати оптимальну базову архітектуру програмного засобу;
- обґрунтувати доцільність використання моделі машинного навчання YOLO26s як базового програмного засобу для виокремлення військових об'єктів на основі онтологічної моделі предметної області, що дасть змогу формалізувати вибір конфігурації програмно-апаратних засобів як задачу багатокритеріальної її оптимізації;

- сформувати навчальний набір даних на основі фото- та відеоматеріалів із застосуванням програмних засобів аугментації для імітації вихідних даних бюджетних сенсорів, погодних умов та камуфляжного забарвлення цілей, що уможливить навчання моделі на репрезентативній вибірці бойових сценаріїв;
- виконати навчання моделі YOLO26s та програмно реалізувати метод із використанням алгоритму міжкадрового підтвердження виокремлень з дворівневою валідацією результатів інференсу на основі високого та низького рівнів впевненості моделі, що дасть змогу підвищити стійкість ПЗ до хибнопозитивних спрацювань у відеопотоці даних з бортових сенсорів без істотних втрат продуктивності інференсу за умов обмежених обчислювальних ресурсів;
- провести тестування програмної реалізації методу на відеоматеріалах знімання із безпілотних комплексів, оцінити відповідність досягнутих показників продуктивності й точності вимогам до розгортання на периферійних бортових обчислювальних пристроях, що дасть змогу підтвердити практичну придатність розробленого програмного засобу до застосування за умов реальних бойових завдань.

Аналіз останніх досліджень та публікацій. У роботі [17] подано комплексну класифікацію кіберзагроз для БПЛА, що охоплює три основні категорії: атаки на канал зв'язку між апаратом та командним центром, атаки на взаємодію між БПЛА, а також зловмисні маніпуляції із функціональністю апарата. Встановлено, що глушіння радіоканалу та GPS-спуфінг є найпоширенішими типами атак у військовому контексті, при цьому більшість моделей FPV-дронів є особливо вразливими через відсутність захищених апаратних архітектур, адже підходи до виготовлення цих літальних апаратів походять із цивільного та споживчого секторів. Автори також окремо виділяють досягнення автономності літальних апаратів як одне із перспективних рішень для значної частини вказаних проблем. Робота [9] надає результати кількісного оцінювання критичності зазначених загроз: спуфінг збільшує похибку позиціонування об'єктів до 20,45 м, тоді як глушіння знижує успішність виконання місій на 40 %.

У роботі [6] детально розглянуто методи досягнення повної автономності БПЛА за умов відсутності зовнішніх каналів зв'язку. Встановлено, що бортові системи комп'ютерного зору є одним із небагатьох класів навігаційних сенсорів, що забезпечують визначення положення апарата без використання бездротового зв'язку, що робить їх принципово стійкими до РЕБ-атак. Водночас, зазначено, що точність візуальної навігації може знижуватись за умов малотекстурованого середовища або за значних висот польоту.

У роботі [15] досліджено завдання виявлення та класифікації бронетехніки у відеопотоці даних з бортових сенсорів дрона за допомогою моделі машинного навчання YOLOv5, розгорнутої на апаратній платформі FPGA. Авторам вдалося встановити придатність архітектури моделі YOLO для вирішення завдань виокремлення військових об'єктів у реальному часі на спеціалізованих вбудованих пристроях. У роботі [8] обґрунтовано, що оброблення відеоданих на наземних пристроях є принципово непридатним для завдань виокремлення цілей у реальному часі. Запропоновано систему на основі вбудованого обчислювального модуля, що виконує повний цикл оброблення відеопотоку даних з бортової камери безпосередньо на борту апарата.

У роботах [3, 4] підтверджено ефективність розгортання алгоритму на базі моделі YOLO на апаратних платформах із обмеженою продуктивністю, зокрема – NVIDIA Jetson, однак встановлено, що стандартні архітектури мають обмежену точність виокремлення мало-розмірних об'єктів, що є критичною проблемою для аерофотознімання з великих висот.

З огляду на результати проаналізованих досліджень, завдання виокремлення військових об'єктів залишається актуальним, оскільки більшість наявних програмних засобів не забезпечують оптимального балансу між високою точністю виокремлення військових об'єктів та стабільною роботою за умов жорстких обмежень бортового апаратного забезпечення.

Матеріали та методи дослідження. Основними методами дослідження є онтологічне моделювання предметної області та математична формалізація. Онтологічне моделювання призначене для формування структурованого подання знань у вигляді системи класів, атрибутів та семантичних відношень, що дає змогу об'єднати різноманітні поняття військової галузі та комп'ютерних наук в єдину формалізовану базу знань. На основі побудованої онтологічної моделі завдання розроблення методу виокремлення військових об'єктів формалізується як задача багатокритеріальної оптимізації конфігурації програмно-апаратних засобів БПЛА, де критеріями виступають точність виокремлення військових об'єктів та швидкодії оброблення відеопотоку даних з бортових сенсорів за умов фізичних обмежень бортової платформи. Експериментальна частина дослідження охоплює формування навчального набору даних на основі реальних матеріалів аерофотознімання, навчання моделі YOLO26s та кількісне оцінювання її характеристик за стандартними метриками комп'ютерного зору, а саме: mAP@50, mAP@50-95, Precision (точність) та Recall (повнота).

Навчальний набір даних сформовано на основі відкритих наборів даних платформи Roboflow. Зібрані матеріали загальним обсягом близько 12 тис. зображень було очищено від низькоякісних зразків, проведено ручну перевірку анотацій, виділено та згруповано основні класи шуканих об'єктів: танк, БТР-БМП, вантажівка, РСЗВ, ЗРК та комбатант. Попереднє оброблення да-

них містило приведення зображень до вхідного формату моделі шляхом масштабування до роздільної здатності 640×640 пікселів алгоритмом рівномірного розтягування. У межах навчання моделі YOLO бібліотеки Ultralytics використовували стандартний набір ваг для версії моделі YOLO26s. Розроблений алгоритм міжкадрового підтвердження виокремлень інтегровано з алгоритмом трекінгу ByteTrack, реалізація якого також входить до складу бібліотеки Ultralytics.

Також виконано тестування продуктивності програмної реалізації алгоритму міжкадрового підтвердження виокремлень у комбінації з моделлю YOLO26s на відеофайлах формату.mp4 та роздільних здатностей 1280×720 та 1920×1080. Навчання моделі та тестування ПЗ здійснено у середовищі Google Colab у межах наданих безкоштовних ресурсів для віддаленого середовища виконання Tesla T4.

Результати дослідження та їх обговорення / Research results and their discussion

Метод виокремлення військових об'єктів – це технологія, яка автоматично знаходить військову техніку та споруди на відео з дрона за допомогою штучного інтелекту. Програма аналізує зображення під час польоту та миттєво підказує оператору, де саме ховається ціль. Онтологічна модель для виокремлення військових об'єктів – це формалізована система знань про воєнні технології та тактичне середовище. Вона структурує військові об'єкти (техніку, позиції, інфраструктуру) за їхніми класами, просторовими зв'язками, станами та характеристиками для інтеграції з алгоритмами розпізнавання.

Для формалізованого подання знань про предметну область у середовищі Protégé побудовано онтологічну модель, що описується кортежем

$$O = \langle C, R, F, RI, Ind, Prop \rangle,$$

де: $C = \{c_j, j = \overline{1, n_C}\}$ – множина класів; $R = \{r_j, j = \overline{1, n_R}\}$ – множина відношень; $F = \{f_j, j = \overline{1, n_F}\}$ – множина функцій інтерпретації; $RI = \{rl_j, j = \overline{1, n_{RI}}\}$ – множина семантичних правил; $Ind = \{ind_j, j = \overline{1, n_{Ind}}\}$ – множина екземплярів онтологічних класів; $Prop = \{prop_j, j = \overline{1, n_{Prop}}\}$ – множина властивостей.

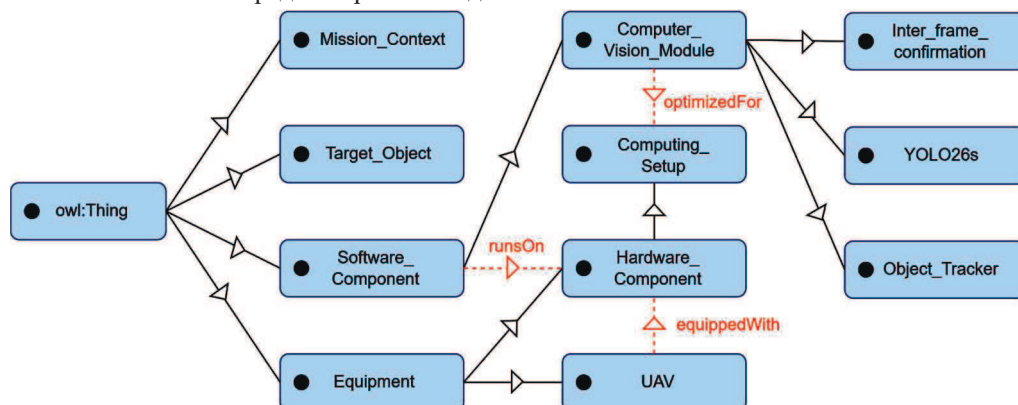


Рис. 1. Частина схеми онтології апаратно-програмного комплексу / Part of the ontology diagram for the hardware-software complex

Модель охоплює чотири групи сутностей (рис. 1): безпілотний комплекс з бортовим апаратним забезпеченням (*Equipment*), ПЗ для виокремлення цілей (*Software_Component*), умови зовнішнього середовища (*Mission_Context*) та шукані військові об'єкти (*Target_Object*). Клас *Equipment* позначає множину усіх фі-

зичних складових апаратно-програмного комплексу, зокрема – обчислювальне обладнання (*Computing_Setup*), до якого можуть входити різні конфігурації обчислювальних модулів із процесором (CPU), відеокартою (GPU), ОЗП та накопичувачем. Також у межах класу *Equipment* розглядаються параметри бортових сенсорів

(відеокамери видимого та інфрачервоного спектрів), польотних контролерів та інших складових, що мають бути адаптовані до умов експлуатації на борту дрона.

Окрім цього, важливо враховувати фізичні параметри самого дрона, окреслені у класі *UAV*, – масу корисного навантаження, потужність блоку живлення та ємність акумуляторів, що безпосередньо впливають на вимоги до бортового обчислювального обладнання. Наприклад, сучасні високопродуктивні відеокарти серії Nvidia GeForce RTX здатні ефективно виконувати інференс моделей комп'ютерного зору, однак їх габарити, маса та енергоспоживання унеможливають застосування як компонента бортового апаратно-програмного комплексу FPV-дронів. Залежності між фізичними параметрами дрона та апаратним забезпеченням позначає відношення *equippedWith*.

Клас *Mission_Context* позначає різні умови роботи комплексу БПЛА, зокрема – ворожа протидія засобами радіоелектронної боротьби, погода та особливості місцевості. Клас *Target_Object* описує характеристики шуканих військових цілей: габарити та пропорції об'єктів, засоби маскування та додаткового захисту. Вказані властивості обох класів безпосередньо впливають на те, які об'єкти буде виокремлювати модель машинного навчання, а також які дані та їх джерела будуть використовуватися для навчання цієї моделі.

Клас *Software_Component* позначає програмні компоненти для виокремлення військових об'єктів з відеопотоку даних бортових сенсорів дрона. Вимоги до цього класу та його характеристики залежать від параметрів усіх інших базових класів. До прикладу, формальне відношення *runsOn* пов'язує клас *Software_Component* із класом апаратного забезпечення *Hardware_Component* та позначає параметри сумісності інтеграції між ними у рамках єдиного апаратно-програмного комплексу. Відношення *optimizedFor* характеризує обмеження параметрів продуктивності обчислювального обладнання (тактова частота та кількість ядер процесора, обчислювальна потужність графічного процесора у TFLOPS, обсяг і пропускна здатність ОЗП) класу *Computing_Setup*, що впливають на параметри засобів комп'ютерного зору та моделей машинного навчання (*Computer_Vision_Module*).

На основі створеної онтологічної моделі завдання вибору придатної конфігурації програмно-апаратних засобів виокремлення військових цілей формалізовано як задачу багатокритеріальної оптимізації конфігурації програмно-апаратних засобів.

Нехай система дрона S описується таким кортежем:

$$S = \langle H, M, E, O \rangle,$$

де: $H = \{h_j, j = \overline{1, n_H}\}$ – множина доступних апаратних платформ і засобів; $M = \{m_j, j = \overline{1, n_M}\}$ – множина наявних програмних засобів штучного інтелекту та моделей машинного навчання; $E = \{e_j, j = \overline{1, n_E}\}$ – множина умов зовнішнього середовища та варіацій властивостей шуканих цілей; O – онтологічна база знань, що задає семантичні зв'язки між елементами H, M, E .

Тоді завдання зводиться до знаходження програмного засобу m , що максимізує зважену суму нормованих критеріїв точності та швидкодії для поточної апаратної платформи та умов середовища, за дотримання сукупності обмежень: мінімально допустимої швидкодії для

коректного оброблення відеопотоку даних з бортових сенсорів, граничного енергоспоживання бортової платформи та сумісності програмного та апаратного забезпечення.

$$F(m) = \omega_1 \cdot \frac{Acc(m, e_{cur})}{Acc_{max}} + \omega_2 \cdot \frac{FPS(h_{cur}, m)}{FPS_{max}} \rightarrow \max,$$

де: ω_1, ω_2 – вагові коефіцієнти (що важливіше – висока швидкодія чи висока точність для виконання певного завдання), $\omega_1 + \omega_2 = 1$, $\omega_1, \omega_2 \geq 0$; $Acc(\cdot)$ – показник точності програмного засобу за поточних умов середовища; $FPS(\cdot)$ – частота оброблення кадрів відеопотоку даних з бортових сенсорів програмного засобу за поточних умов апаратного забезпечення; Acc_{max} – максимальний теоретичний показник точності (100 % або 1,0); FPS_{max} – максимальна частота бортової камери.

Сукупність обмежень має такий вигляд:

$$\begin{cases} FPS(h_{cur}, m) \geq FPS_{min}; \\ Pow(h_{cur}, m) \leq Pow_{limit}; \\ Compatibility(h_{cur}, m) = true, \end{cases}$$

де: FPS_{min} – мінімальне значення швидкодії для коректної роботи із відеопотоком даних; $Pow(\cdot)$ – поточне енергоспоживання апаратної платформи під час роботи програмного рішення; Pow_{min} – обмеження живлення; $Compatibility(\cdot)$ – сумісність апаратного та програмного забезпечення.

Зважаючи на значну кількість параметрів для кожного елементу кортежу системи дрона S , завдання детальної формалізації складових кортежу може бути підгрупам для подальших етапів дослідження.

У межах описаної задачі багатокритеріальної оптимізації конфігурації програмно-апаратних засобів, формалізованої на основі онтологічної моделі, одним із обґрунтованих програмних рішень є використання моделі машинного навчання YOLO26s сімейства YOLO, що демонструє найкращий баланс між точністю виокремлення військових об'єктів та обчислювальною ефективністю на периферійних пристроях серед розглянутих підходів. Варіант "small" обрано як компроміс між точністю та швидкістю, що відповідає критеріям цільової функції оптимізації конфігурації програмно-апаратних засобів за умов апаратних обмежень периферійних бортових пристроїв.

Наступним етапом є формування навчального набору даних. Відповідно до специфіки завдання визначено шість цільових класів об'єктів: танк, БТР/БМП, вантажівка, РСЗВ, ЗРК та комбатант. Вибір класів зумовлений пріоритетністю цілей на полі бою в контексті застосування FPV-дронів, а також необхідністю забезпечення достатньої репрезентативності кожного класу для ефективного навчання моделі. Техніку типу БМП та БТР об'єднано в єдиний клас, зважаючи на їх високу візуальну схожість на фотознімках, характерних для FPV-дронів. Обидва типи машин належать до категорії гусеничних або колісних бойових броньованих машин і мають подібні пропорції, розміри та компонування.

Основу навчального набору даних становлять публічні набори даних платформи Roboflow Universe, що містять зображення із широким охопленням різноманітних умов знімання, та попередньо сформовані файли розмітки об'єктів військової техніки. Другу категорію становлять вручну відібрані фотоматеріали та кадри із відео бойових дій з відкритих джерел мережі Інтернет,

що підвищує відповідність навчальних даних реальним умовам застосування. Незначну частку вибірки становлять матеріали з інших джерел, зокрема – 3D-моделювання, що використовувались для часткового поповнення недостатньо поданих класів. Зібрані матеріали за-

гальним обсягом 12616 зображень пройшли ручну верифікацію анотацій та здійснено очищення від низькоякісних зразків, після чого ці матеріали розподілені на навчальну, валідаційну та тестову вибірки у співвідношенні 80/16/8 (рис. 2).

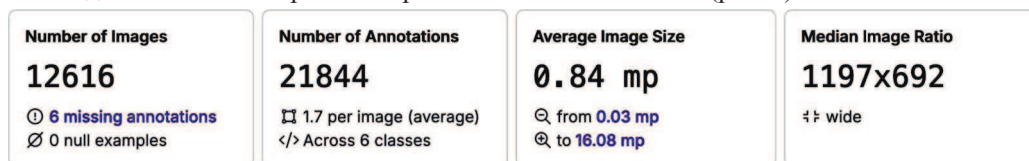


Рис. 2. Метадані навчального набору даних / Training dataset metadata

Перед навчанням усі зображення приведено до вхідного формату моделі 640×640 пікселів. Оскільки медіанне співвідношення сторін датасету становить 1197×692 пікселів, виникає невідповідність між прямокутним форматом вхідних даних та квадратним вхідним шаром мережі. Для вирішення цієї проблеми обрано алгоритм рівномірного розтягування замість стандартного letterbox-масштабування, оскільки внаслідок його застосування близько 42 % вхідного тензора займають смуги заповнення, які не несуть корисної інформації про сцену. Алгоритм розтягування, попри незначне геометричне спотворення пропорцій об'єктів, забезпечує повне використання вхідного простору мережі за умови консистентного застосування як під час навчання, так і під час інференсу. Нюанс відповідності вхідних даних для навчання та інференсу є критичним, адже модель навчається розпізнавати об'єкти на основі статистичних закономірностей, притаманних саме тому формату зображень, на якому вона тренувалась. Будь-яке відхилення у попередньому обробленні зображень на етапі інференсу фактично змінює розподіл вхідних даних, що призводить до деградації точності виокремлення військових об'єктів за незмінних ваг моделі.

Для підвищення узагальнювальної здатності моделі та імітації реальних умов аерофотознімання застосовано комплекс онлайн-аугментацій бібліотеки Ultralytics та Albumentations, що охоплює такі групи перетворень:

- колірна аугментація у просторі HSV – зсув відтінку до 0,015, насиченості до 0,3, яскравості до 0,2;
- геометрична аугментація – обертання до $\pm 10^\circ$, масштабування від 0,4х до 1,6х, переміщення до 5 %;
- мозаїчна аугментація – компонування 4-х зображень, копіювання та вставлення об'єктів в інші зображення для 15 % вибірки;
- просторові віддзеркалення – горизонтальне та вертикальне, до 50 % зображень;
- імітація характеристик оптичних сенсорів – випадкове розмиття Гауса до 30 %, зміна контрасту до 30 %, перетворення у відтінки сірого для 18 % вибірки.

Усі аугментації застосовують безпосередньо під час завантаження зображень у процесі навчання, що забезпечує практично необмежену варіативність вхідних даних. Це означає, що одне й те саме зображення на кожній епісоді надходить до моделі в новій аугментованій формі. Мозаїчна аугментація навчає модель виокремлювати об'єкти в різноманітному контексті та за різних масштабів одночасно, що є особливо важливим для класу комбатантів, які часто опиняються поряд із технікою. Колірна аугментація у просторі HSV моделює природну варіативність освітлення та атмосферних умов. Масштабування є ключовим геометричним перетворенням для аерофотознімання, оскільки безпосередньо симулює зміну відстані до об'єкта та висоти

польоту дрона. Вертикальне віддзеркалення, на відміну від наземних наборів даних, є обґрунтованим для аерофотознімання, оскільки за погляду зверху вниз сцена не має фіксованої орієнтації. Для імітації характеристик бюджетних оптичних сенсорів програмними засобами бібліотеки Albumentations реалізовано два додаткових перетворення: розмиття Гауса із розміром ядра від 1 до 3 пікселів та стандартним відхиленням від 0,5 до 1,8, що моделює оптичну нечіткість дешевих камер та розфокусування на великих відстанях, тоді як випадкова зміна контрасту в межах $\pm 15\%$ без зміни яскравості відтворює варіативність налаштувань матриці сенсора. Перетворення у відтінки сірого для 18 % вибірки підвищує стійкість моделі до монохромних або інфрачервоних відеопотоків.

Навчання моделі YOLO26s виконано протягом 220 епох із використанням оптимізатора AdamW із початковою швидкістю навчання 0,001 та косинусним розкладом із остаточним коефіцієнтом 0,01. Розмір пакету становив 16 зображень, розмір вхідного зображення – 640×640 пікселів. Для стабілізації початкового етапу навчання застосовано три епохи прогріву; регуляризація забезпечувалась коефіцієнтом weight decay 0,0005. Збереження контрольних точок виконувалось кожні 5 епох для забезпечення можливості відновлення навчання.

Таблиця. Показники точності навченої моделі на основі моделі YOLO26s / Accuracy metrics for the trained model based on YOLO26s

Клас об'єктів	Метрика точності виокремлення військових об'єктів			
	Precision	Recall	mAP@50	mAP@50-95
Всі	0,819	0,74	0,803	0,624
БТР/БМП	0,84	0,76	0,855	0,653
Комбатант	0,676	0,418	0,489	0,273
РЗЗВ	0,921	0,756	0,852	0,725
Танк	0,829	0,795	0,853	0,574
Вантажівка	0,859	0,816	0,867	0,714
ЗРК	0,792	0,896	0,903	0,804

За результатами проведення перевірок на тестовій вибірці з 1009 зображень модель досягла середнього показника mAP@50 рівня $0,803^{\pm 0,05}$ та mAP@50 рівня $0,624^{\pm 0,06}$ за точності $0,819^{\pm 0,09}$ та повноти $0,740^{\pm 0,08}$. Як наведено в таблиці, найвищі результати продемонстрував клас ЗРК із mAP@50 рівня 0,903, тоді як клас комбатант виявився найбільш проблемним – 0,489 mAP@50 та повнота тільки 0,418, що зумовлено недостатньою поданістю зразків малого і середнього розмірів в наборі даних, а також малими розмірами шуканих об'єктів на вхідних відеопотоках під час інференсу.

Вхідними даними методу у процесі інференсу є відеопотік даних з бортової камери дрона з типовою роздільною здатністю 1280×720 або 1920×1080 пікселів та частотою 24-30 кадрів/секунду. Кожен кадр попередньо

масштабується за потреби до формату 640×640 пікселів методом рівномірного розтягування та подається на вхід моделі YOLO26s. Паралельно з інференсом виконується трекінг виокремлень алгоритмом ByteTrack, який супроводжує об'єкти з рівнем впевненості, що вищий за основний поріг (0,3), при цьому поріг створення нового треку встановлено на 10 % пунктів вище, ніж основний, для запобігання появи нестабільних треків від слабких виокремлень.

Ключовою відмінністю аналізу відеопотоку даних з бортових сенсорів від оброблення окремих зображень є наявність послідовності кадрів, що несе не тільки інформацію про розміщення об'єктів у певному моменті часу, але і дані про динаміку сцени. Основним викликом при цьому є необхідність стійкого утримання виокремленого об'єкта між кадрами за умов його переміщення, часткового перекриття та інших параметрів середовища. Водночас додатковий вимір даних відкриває можливість оцінювання стабільності виокремлень у часі, що є недоступною за аналізу окремих зображень. Реальні об'єкти породжують стійкі повторювані сигнали виокремлення військових об'єктів впродовж послідовних кадрів, тоді як хибнопозитивні спрацювання, зазвичай, є короточасними. Це припущення є базовим для реалізованого алгоритму міжкадрового підтвердження виокремлень.

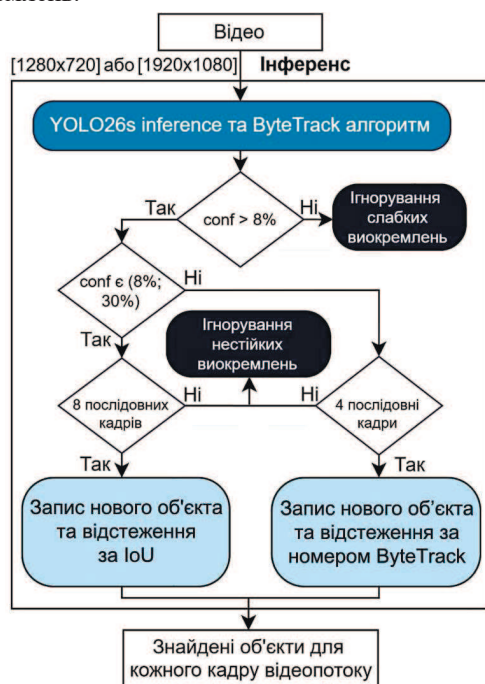


Рис. 3. Блок-схема алгоритму міжкадрового підтвердження виокремлень військових об'єктів / Block diagram of the inter-frame confirmation algorithm for the military objects extraction

Кроки алгоритму міжкадрового підтвердження виокремлень (рис. 3):

- 1) Виконати виявлення та класифікацію об'єктів за допомогою моделі машинного навчання YOLO26s на основі вхідного кадру відеопотоку даних з бортових сенсорів із роздільною здатністю 1280×720 або 1920×1080 пікселів. Вихідними даними є масив виокремлених об'єктів із параметрами: назва класу, координати прямокутника обмеження об'єкта та значення впевненості виокремлення об'єкта. Для нових виокремлень із впевненістю, яка більша або становить 40 %, алгоритм ByteTrack призначає унікальний ідентифікатор. Усі виокремлення об'єктів із впевненістю, що менша за 8 %, відкинути.

- 2) Поділити виокремлення об'єктів на дві групи низького та високого порогів залежно від значення впевненості моделі YOLO. Перша група міститиме об'єкти із впевненістю від 8 до 30 % не включно, друга – виокремлення об'єктів від 30 до 100 % включно.
- 3) Для кожного виокремлення об'єкта обох масивів присвоїти значення кількості послідовних кадрів, на яких він успішно виокремлюється. Об'єкти групи низького порогів порівняти між кадрами за допомогою метрики Intersection over union із часткою схожості 60 %. Групи впевненості високого порогів відстежувати між кадрами за унікальними ідентифікаторами алгоритму ByteTrack із мінімальним порогом знайдених виокремлень 30 % включно.
- 4) Якщо об'єкт групи низького порогів успішно виокремлюється протягом 8 послідовних кадрів відеопотоку даних з бортових сенсорів – додати у масив підтверджених виокремлень.
- 5) Якщо об'єкт групи високого порогів успішно виокремлюється протягом 4 послідовних кадрів відеопотоку даних з бортових сенсорів – додати у масив підтверджених виокремлень.
- 6) Перевірити всі об'єкти масиву підтверджених виокремлень: якщо підтверджене на попередніх кадрах виокремлення об'єкта не з'явилося хоча б на 1 кадрі із 4 наступних – видалити об'єкт із масиву підтверджених виокремлень. Якщо об'єкт групи А на певному кадрі отримав значення впевненості моделі групи В – залишити його у попередній групі А, поки не буде виконана умова послідовного виокремлення групи впевненості В.
- 7) Відобразити на поточному кадрі відеопотоку даних з бортових сенсорів кожен об'єкт із масиву знайдених виокремлень залежно від групи впевненості.

Непідтверджені об'єкти групи низького порогів візуально не відображаються на кадрах відеопотоку даних з бортових сенсорів дрона. Підтверджені об'єкти групи низького порогів відображаються у вигляді зафарбованого напівпрозорого прямокутника із міткою класу та символом "?" замість значення впевненості. Непідтверджені об'єкти групи високого порогів відображаються як підтверджені об'єкти групи низького порогів, але із числовим значенням кількості кадрів до підтвердження виокремлення об'єкта (рис. 4). Підтверджені об'єкти групи високого порогів відображаються у вигляді прямокутника із суцільними краями та прозорою серединою із міткою класу та значення впевненості.

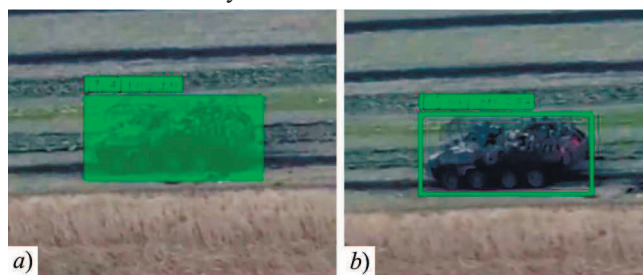


Рис. 4. Результат роботи алгоритму підтвердження виокремлень для високого порогів на прикладі двох послідовних кадрів / Result of the inter-frame confirmation algorithm for a high threshold detection on a two consecutive frames

Кількість кадрів, потрібних для підтвердження виокремлення об'єкта як для низького, так і для високого порогів, є емпірично визначеними метапараметрами методу, що можуть бути скориговані залежно від частоти кадрів відеопотоку даних з бортових сенсорів та оперативних вимог до затримки підтвердження цілі. Чим більшим є число послідовних кадрів для реєстрації нового об'єкта, тим ПЗ буде більш консервативним, реагу-

ючи тільки на стійкі достовірні сигнали, однак при цьому збільшуючи затримку між моментом появи об'єкта у кадрах та підтвердженням його виокремлення.

Розроблений алгоритм також демонструє підвищену стійкість до несприятливих умов зовнішнього середовища, зокрема – часткового задимлення, атмосферних опадів та перекриття об'єктів рослинністю або елементами рельєфу. Короткочасне погіршення якості відеопотоку даних з бортової камери дрона призводить тільки до збільшення часу підтвердження виокремлення об'єкта, але не до втрати вже накопиченого стану треку (рис. 5).

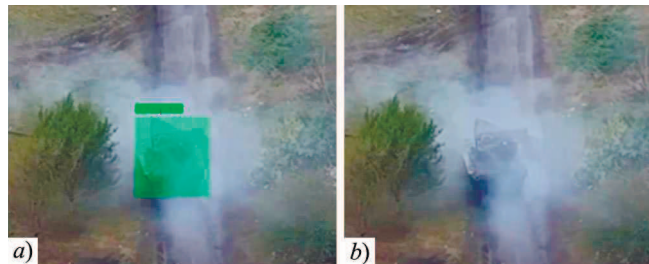


Рис. 5. Порівняння результату роботи з алгоритмом міжкадрового підтвердження виокремлень та без / Comparison of the results using the inter-frame confirmation algorithm versus no confirmation

За результатами тестування на апаратному комплексі Tesla T4 середній час інференсу становить $22^{\pm 2}$ мс, що відповідає максимальній швидкодії $40^{\pm 5}$ кадрів/с для відеопотоку даних з бортових сенсорів дрона з роздільною здатністю 1920×1080 пікселів. Вплив алгоритму міжкадрового підтвердження виокремлень об'єктів на загальну продуктивність програмної реалізації методу є незначним та знаходиться в межах зазначеної похибки вимірювань.

Для розгортання на цільовому периферійному апаратному забезпеченні, як один із розв'язків задачі багатокритеріальної оптимізації конфігурації програмно-апаратних засобів, доцільним є використання платформи NVIDIA Jetson Orin NX 16GB. Завдяки підтримці апаратного пришвидшення TensorRT та оптимізації моделі до формату FP16 очікуваний час інференсу становитиме орієнтовно $35^{\pm 5}$ мс, що відповідає швидкодії $26^{\pm 4}$ кадрів/с та є достатнім для надійної роботи реалізованого методу в режимі реального часу. Платформа Jetson AGX Orin 64GB забезпечила б вищу продуктивність, наближену до показників Tesla T4, однак її істотно вища вартість та більше енергоспоживання роблять Jetson Orin NX 16GB пріоритетним вибором з огляду на вимоги до масштабованості та економічної доцільності бойового застосування.

Подальше вдосконалення розробленого методу пов'язане з низкою викликів. Застосований метод рівномірного розтягування зображень до формату 640×640 пікселів забезпечує збереження інформаційного контексту сцени, однак вносить геометричне спотворення пропорцій об'єктів. Альтернативний метод на основі letterbox-масштабування зі збереженням пропорцій потребує повного перенавчання моделі з узгодженою попередньою обробкою на обох етапах – навчанні та інференсі, що є предметом окремого порівняльного дослідження.

Архітектура моделі YOLO демонструє високу ефективність для об'єктів середнього та великого розмірів, однак є менш пристосованою до виокремлення малорозмірних об'єктів. З огляду на те, що реальні бойові сценарії передбачають значні відстані між дроном та

ціллю, де об'єкти займають тільки декілька десятків пікселів у кадрі, точність виокремлення військових об'єктів може істотно знижуватись. Це безпосередньо проявляється у результатах тестування класу комбатант, що продемонстрував найнижчі показники точності серед усіх класів. Попри відносно збалансований розподіл екземплярів цього класу за розміром у наборі даних, фізичні розміри людини на аерофотографіях з реальних висот польоту дрона є істотно меншими порівняно з об'єктами техніки. Одним із можливих підходів до вирішення цієї проблеми є алгоритм SAHI [11], що реалізує інференс із застосуванням перекриття фрагментів, однак його використання може призвести до значних втрат продуктивності, що є неприйнятним за умов обмежених ресурсів бортового обчислювального комплексу.

Обговорення результатів дослідження. У роботі [10] аналізують різні методи створення програмного засобу для розпізнавання військової техніки на зображеннях. Виконано порівняння ефективності моделей керованого навчання ResNet50 та Xception, аналіз переваг та недоліків їхньої архітектури. Згідно із експериментальними результатами дослідження, архітектура ResNet50 дає змогу досягти точності близько 88 %, на відміну від моделі Xception, яка володіє високими показниками втрат. Поза увагою авторів залишилися питання обчислювальної ефективності та швидкості інференсу, недостатньо розкрито можливості нейронних мереж обробляти зображення із низькою роздільною здатністю, а також замасковані об'єкти.

У роботі [14] автори фокусуються на виявленні та класифікації військової техніки за різних умов навколишнього середовища. Вони пропонують архітектуру нейронної мережі на основі ResNet із стратегією ієрархічної репрезентації ознак та навчання з підкріпленням. Проведено порівняння точності запропонованого методу та популярних моделей: R-FCN, SSD, YOLOv3, YOLOv4. Авторам вдалося досягти високої точності у сегментації та класифікації військової техніки. У статті не акцентовано уваги на обробленні відеопотоку даних, а також використовується обмежений набір навчальних даних та класів, що не повністю репрезентує різні типи військової техніки.

У роботах [1, 13, 18] пропонують використання моделі YOLOv8 різних версій для вирішення завдання розпізнавання та класифікації. У роботі [1] додатково пропонують застосування методів попереднього оброблення вхідних даних, алгоритм виокремлення ознак SIFT та мережу класифікації Deep Belief Network, що дає змогу виокремлювати транспорт від елементів урбаністичного середовища із високою точністю. Недоліком вказаного підходу є надлишкова комплексність, що призводить до низької продуктивності у межах завдання оброблення відеопотоку даних з бортових сенсорів дрона. Стаття [18] демонструє результати розроблення ПЗ для виокремлення малих морських суден та людей, що потребують порятунку внаслідок надзвичайних ситуацій на відкритих водоймах. Для підвищення точності за незначної втрати швидкодії використовується згортковий модуль уваги (англ. *Convolutional Block Attention Module*). Авторам [13] не вдається досягти високої точності виявлення та класифікації, також вони не розглядають завдання аналізу відеопотоку даних з бортових сенсорів дрона. У роботах [1, 18] зацентовано на виокремленні цивільного транспорту та морських су-

ден, проте результати точності, досягнуті за допомогою моделі YOLOv8, вказують на перспективність використання цього підходу для вирішення завдання виокремлення військових об'єктів, хоча в контексті продуктивності використання додаткових модулів призводить до підвищення часу інференсу, що може негативно вплинути на швидкість ПЗ.

У роботі [2] подано комплексне дослідження застосування БПЛА для військового спостереження та відстеження цілей у режимі реального часу. Автори сформували власний навчальний набір даних на основі відеоматеріалів із безпілотників та провели порівняльний аналіз моделей YOLOv5, YOLOv8 та YOLOv11 у варіантах nano та small задля досягнення балансу між точністю виокремлення об'єктів та обчислювальною ефективністю. За результатами експериментів найвищі показники продемонструвала модель YOLOv11s у поєднанні з алгоритмом ByteTrack, що забезпечило надійне виокремлення об'єктів та супроводження цілей із мінімальним залученням оператора. У роботі не розглядають проблему фільтрації хибнопозитивних спрацювань, а також не аналізують впливу маскування цілей на стабільність відстеження.

Для вирішення проблеми наявності великої кількості класів військових об'єктів, у роботі [7] запропоновано нову модель, яка поєднує методи виокремлення деяких функціонально важливих компонентів (зброї, сенсорів, ходової частини) замість розпізнавання техніки як єдиного цілісного об'єкта. Цей підхід дає змогу розпізнавати як базові моделі, так і модифіковану техніку, яка широко використовується в сучасних конфліктах. Також вказаний метод дає змогу проводити класифікацію об'єктів, які не були безпосередньо включені в тренувальний набір, що зменшує потребу у широкій репрезентативності для великої кількості класів. Окрім цього, використання атрибутів відкритого світу дає змогу моделі динамічно адаптуватися до нових класів або конфігурацій техніки без необхідності повного перенавчання. Недоліком цього підходу є значне підвищення складності побудови навчального набору даних.

Підходи до вирішення завдання виокремлення замаскованих об'єктів розглядають у роботах [5, 12, 16]. Наприклад, у роботі [12] автори зосереджують увагу на виокремленні замаскованих військових цілей за допомогою автономних безпілотних засобів. Основою реалізованої системи є пірамідална модель-трансформер із багатовимірною мапою ознак із високою роздільною здатністю. Окрім цього, автори пропонують окремі модулі для виокремлення і комбінації меж, а також злиття ознак. У роботі [16] автори використовують згорткову нейронну мережу, що деформується, та алгоритми виокремлення країв (DCEGNet). У цій статті також проведено порівняння запропонованого методу та наявних моделей COD (англ. *Camouflaged Object Detection*): ZomNet, SegMaR, COD10K та інші. У статті [5] автори пропонують використання моделі YOLOv8 та алгоритм покращення ознак для виокремлення замаскованих об'єктів. Також спроектовано багатогілковий згортковий механізм виділення розмитих сегментів, а також алгоритм постпроцесингу E-Soft-NMS для зменшення кількості хибних виокремлень. Незважаючи на те, що програмний засіб у роботі [12] пристосований до виокремлення цілей, досягнута точність, що становить близько 49 %, не є достатньою, оскільки військова галузь

передбачає високі ризики. Оскільки у роботах [5, 16] для навчання здебільшого використовується набір даних із зображеннями тварин у природному середовищі, питання використання запропонованої моделі для розпізнавання замаскованої військової техніки залишається відкритим.

Отже, обговорення результатів дослідження підтверджує перспективність використання архітектури моделі YOLO як базового програмного засобу для виокремлення військових об'єктів з відеопотоку даних з бортових сенсорів FPV-дрона, проте жоден із розглянутих програмних засобів не вирішує проблему ефективності їх виокремлення повною мірою. Архітектури відеозасобів на основі CNN та трансформерів забезпечують прийнятну точність виокремлення військових об'єктів, проте їх обчислювальна складність унеможливорює оброблення відеопотоку даних з бортових сенсорів дрона в режимі реального часу на периферійних бортових пристроях. Більшість розглянутих підходів орієнтована на аналіз окремих зображень і не вирішує проблеми фільтрації хибнопозитивних спрацювань у відеопотоці даних з бортових сенсорів дрона. Програмна реалізація методу на основі моделі YOLO26s у поєднанні з алгоритмом міжкадрового підтвердження виокремлень усуває зазначені недоліки: середній час інференсу 22^{+2} мс забезпечує оброблення відеопотоку даних з бортових сенсорів дрона зі швидкістю 40^{+5} кадрів/с, досягнутий середній показник точності mAP@50, що становить 0,803, перевищує результати робіт із виокремлення замаскованих військових цілей [12], а алгоритм міжкадрового підтвердження виокремлень забезпечує стійке супроводження цілей.

Отже, внаслідок виконаної роботи можна сформулювати такі наукову новизну та практичну значущість результатів дослідження.

Наукова новизна отриманих результатів дослідження – розроблено метод виокремлення військових цілей у відеопотоці даних з бортових сенсорів FPV-дрона, що дає змогу формалізувати вибір придатної конфігурації програмно-апаратних засобів як задачу їх багатокритеріальної оптимізації та забезпечує дворівневу валідацію високо- та низькодостовірних сигналів без істотних втрат продуктивності за умов обмежених обчислювальних ресурсів бортового апаратного забезпечення.

Практична значущість результатів дослідження – можна використати як програмний компонент бортового комплексу FPV-дрона або безпілотника іншої комплектації для забезпечення автономного виокремлення військових об'єктів та супроводження військових цілей на завершальній стадії бойової місії за умов активної радіоелектронної протидії, що дає змогу істотно знизити залежність від каналу зв'язку з оператором, покращити рівень успішності виконання бойових завдань та збільшити кількість одночасно задіяних апаратів за обмеженого особового складу.

Висновки / Conclusions

Розроблено метод виокремлення військових об'єктів з відеопотоку даних FPV-дрона з використанням моделей машинного навчання, що дало змогу підвищити стійкість ПЗ до хибнопозитивних спрацювань у відеопотоці даних з бортових сенсорів без істотних втрат їх продуктивності за умов обмежених обчислювальних ресурсів. За результатами проведеного дослідження можна зробити такі основні висновки.

1. Виконано аналіз сучасних методів комп'ютерного зору та моделей машинного навчання для завдання виокремлення об'єктів у відеопотоці даних з бортових сенсорів дрона у контексті визначених програмно-апаратних вимог.
2. Обґрунтовано вибір моделі машинного навчання YOLO26s як базового програмного засобу для виокремлення військових об'єктів на основі онтологічної моделі предметної області та задачі багатокритеріальної оптимізації конфігурації програмно-апаратних засобів за критеріями точності, швидкодії та енергоспоживання.
3. Сформовано навчальний набір даних обсягом 12616 зображень із шести класів військових об'єктів на основі відкритих джерел та реальних аерофотоматеріалів; застосовано комплекс аугментацій для імітації умов бюджетних оптичних сенсорів, камуфляжу та різноманітних умов зовнішнього середовища.
4. Виконано навчання моделі YOLO26s, реалізовано програмний метод з використанням алгоритму міжкадрового підтвердження виокремлень з дворівневою валідацією, що забезпечує стійке супроводження цілей та фільтрацію хибнопозитивних спрацювань без істотного впливу на продуктивність реалізованого методу.
5. Виконано верифікацію методу виокремлення військових цілей у відеопотоці даних з бортових сенсорів FPV-дрона та тестування його програмної реалізації. Унаслідок проведеного тестування методу досягнуто середнього показника точності $mAP@50$ 0,803^{±0,05} за тривалості інференсу 22^{±2} мс, що відповідає швидкодії 40^{±5} кадрів/с для відеопотоку даних з бортових сенсорів дрона з роздільною здатністю 1920×1080 пікселів на апаратному комплексі Tesla T4; як цільову апаратну платформу для розгортання на борту дрона запропоновано сімейство NVIDIA Jetson Orin.

References

1. Al Mudawi, N., Qureshi, A. M., Abdelhaq, M., Alshahrani, A., Alazeb, A., Alonazi, M., & Algarni, A. (2023). Vehicle detection and classification via YOLOv8 and deep belief network over aerial image sequences. *Sustainability*, 15(19), 14597–14597. <https://doi.org/10.3390/su151914597>
2. Chakrabarty, S., Chatterjee, R., Chakraborty, S., Roy Shuvo, S., & Chowdhury, R. (2025). Drones in defense: Real-time vision-based military target surveillance and tracking. *2025 3rd International Conference on Intelligent Systems, Advanced Computing and Communication (ISACC)*, 508–513. <https://doi.org/10.1109/ISACC65211.2025.10969335>
3. Chen, Z., Zhang, Y., & Xing, S. (2025). YOLO-LE: A lightweight and efficient UAV aerial image target detection model. *Computers, Materials & Continua*, 84(1), 1787–1803. <https://doi.org/10.32604/cmc.2025.065238>
4. Cheng, Q., Wang, H., Zhu, B., Shi, Y., & Xie, B. (2023). A Real-Time UAV target detection algorithm based on edge computing. *Drones*, 7(2), article number 95. <https://doi.org/10.3390/drones7020095>
5. Han, T., Cao, T., Zheng, Y., Chen, L., Wang, Y., & Fu, B. (2023). Improving the detection and positioning of camouflaged objects in YOLOv8. *Electronics*, 12(20), 4213–4213. <https://doi.org/10.3390/electronics12204213>
6. Hashim, H. A. (2024). Advances in UAV avionics systems architecture, classification and integration: A comprehensive review and future perspectives. *Results in Engineering*, 25, 103786–103786. <https://doi.org/10.1016/j.rineng.2024.103786>
7. Henkel, M., Keilbach, S., Burkart, N., & Schumann, A. (2024). Part-aided military vehicle detection. *Artificial Intelligence for Security and Defence Applications II*, 13206, 321–333. <https://doi.org/10.1117/12.3034615>
8. Huang, F., Chen, S., Wang, Q., Chen, Y., & Zhang, D. (2023). Using deep learning in an embedded system for real-time target detection based on images from an unmanned aerial vehicle: vehicle detection as a case study. *International Journal of Digital Earth*, 16(1), 910–936. <https://doi.org/10.1080/17538947.2023.2187465>
9. Joaneke, P. C., Obioha-Val, O., Olaniji, O. O., Ogungbemi, O. S., Olisa, A. O., & Akinola, O. I. (2024). Protecting autonomous UAVs from GPS spoofing and jamming: a comparative analysis of detection and mitigation techniques. *Journal of Engineering Research and Reports*, 26(10), 71–92. <https://doi.org/10.9734/jerr/2024/v26i101291>
10. Legendre, D., & Vankka, J. (2020). Military vehicle recognition with different image machine learning techniques. *Communications in Computer and Information Science*, 220–242. https://doi.org/10.1007/978-3-030-59506-7_19
11. Liang, X., Xiang, J., Qin, S., Xiao, Y., Chen, L., Zou, D., Ma, H., Huang, D., Huang, Y., & Wei, W. (2025). Small target detection algorithm based on SAHI-Improved-YOLOv8 for UAV imagery: A case study of tree pit detection. *Smart Agricultural Technology*, 12, article ID 101181. <https://doi.org/10.1016/j.atech.2025.101181>
12. Liu, K., Li, A., Yang, S., Wang, C., & Zhang, Y. (2025). Multi-scale attention and boundary-aware network for military camouflaged object detection using unmanned aerial vehicles. *Signal, Image and Video Processing*, 19(1), article number 184. <https://doi.org/10.1007/s11760-024-03766-1>
13. Nazarkevych, M., & Oleksiv, N. (2024). Method of identification of combat vehicles based on YOLO. *Information Systems and Networks*, 15, 87–105. <https://doi.org/10.23939/sisn2024.15.087>
14. Ouyang, Y., Wang, X., Hu, R., Xu, H., & Shao, F. (2022). Military vehicle object detection based on hierarchical feature representation and refined localization. *IEEE ACCESS*, 10, 99897–99908. <https://doi.org/10.1109/ACCESS.2022.3207153>
15. Soumya, D. Sree, Aishwarya, Ch., & Vasavi, S. (2023). FPGA-based military vehicles detection and classification from drone videos using YOLOv5. *Lecture Notes in Electrical Engineering*, 1057, 265–276. https://doi.org/10.1007/978-981-99-3691-5_22
16. Wang, X., Chen, C., Ruan, K., et al. (2024). Camouflaged object detection based on deformable convolution and edge guidance [Preprint]. *Research Square*. <https://doi.org/10.21203/rs.3.rs-5259126/v1>
17. Yu, A., Kolotylo, I., Hashim, H. A., & Eltoukhy, A. E. E. (2025). Electronic warfare cyberattacks, countermeasures, and modern defensive strategies of UAV avionics: A survey. *IEEE ACCESS*, 13, 68660–68681. <https://doi.org/10.1109/ACCESS.2025.3561068>
18. Zhao, B., Zhou, Y., Song, R., et al. (2024). Modular YOLOv8 optimization for real-time UAV maritime rescue object detection. *Scientific Reports*, 14(1), article number 24492. <https://doi.org/10.1038/s41598-024-75807-1>

S. V. Zdebskyi, M. M. Seniv

Lviv Polytechnic National University, Lviv, Ukraine

METHOD FOR EXTRACTING MILITARY OBJECTS FROM AN FPV DRONE VIDEO STREAM USING MACHINE LEARNING MODELS

The study investigates existing computer vision methods and machine learning approaches for the automated detection and classification of military targets by unmanned aerial vehicles. The significance and growing operational relevance of UAVs, particularly FPV (First Person View) drones, across a broad spectrum of combat tasks is substantiated in the context of modern warfare, with particular emphasis on the full-scale invasion of Ukraine by the Russian Federation. The vulnerability of unmanned aerial systems to electronic warfare countermeasures is identified as a critical limitation, arising from their dependence on continuous communication

with the operator. One of the promising solutions is the implementation of artificial intelligence and computer vision technologies to achieve partial or full autonomy of UAVs. The key element of such a system is a software component for automated target extraction from the onboard sensor video stream, allowing the drone to complete the attack without operator participation at the terminal stage of a military operation. Strict requirements for integration into onboard systems are taken into account: optimization for limited power supply and computing resources (CPU, memory), high precision metrics, algorithm reliability and fault tolerance. To address this problem, a method combining the YOLO26s model with an interframe confirmation algorithm has been implemented, generating the final output based on signal consistency across consecutive video frames to filter out false positives with minimal impact on performance. A complex training dataset has been formed, featuring six classes of military objects: tank, IFV/APC, truck, MLRS, SAM, and combatant, sourced from real aerial photo and video materials, with augmentation strategies applied to simulate low-cost sensor characteristics, varying environmental conditions, and target camouflage. The developed system achieves high accuracy metrics, with mean mAP@50 of $0.803^{\pm 0.05}$ and mAP@50-95 of $0.624^{\pm 0.06}$, at an average inference time of $22^{\pm 2}$ milliseconds, corresponding to $40^{\pm 5}$ frames per second for a 1920×1080 -pixel video stream on Tesla T4 hardware. The NVIDIA Jetson Orin hardware series is proposed for deploying the developed software solution onboard an FPV drone.

Keywords: machine learning models; object extraction; computer vision; YOLO model; interframe confirmation of extractions; FPV drone; model inference; military vehicles; onboard software.