



Б. Б. Вдовін, В. С. Яковина, О. М. Томашевський

Національний університет "Львівська політехніка", м. Львів, Україна

ВІЯВЛЕННЯ ЛІСОВИХ ПОЖЕЖ НА ПОТОКОВИХ ЗОБРАЖЕННЯХ З БПЛА МЕТОДАМИ ГЛИБИННОГО НАВЧАННЯ

Досліджено проблему виявлення лісових пожеж у режимі реального часу із використанням безпілотних літальних апаратів і методів глибокого навчання. Проаналізовано сучасні методи виявлення та класифікації пожеж, визначено їхні обмеження та недоліки, зокрема – низьку ефективність під час роботи в реальному часі та залежність від специфіки місцевості. Встановлено, що традиційні методи виявлення лісових пожеж, засновані на супутникових знімках або сенсорних дачах, не забезпечують достатньої точності на ранніх стадіях займання, що ускладнює своєчасне реагування. Досліджено можливості застосування штучного інтелекту, зокрема – згорткових нейронних мереж, для виявлення полум'я та диму на знімках. Розроблено архітектуру моделі на підставі моделі YOLOv5, яка використовує механізм спільного навчання детектора та класифікатора з використанням алгоритму LBTW.2. Створено набір даних, що містить 4104 зображення із 3289 розміченими рамками класу "вогонь" та 2788 рамками класу "дим", отриманих із відеоматеріалів, знятих камерами дронів. Проведено порівняння розробленої моделі з наявними підходами, що підтвердило її переваги в точності класифікації та виявлення об'єктів. Порівняно з базовою моделлю YOLOv5s, навченою без спільного навчання моделей, запропонований підхід показав кращі результати, зокрема – метрика mAP@50 зросла на 11,6 % (0,917 проти 0,822). Класифікатор моделі продемонстрував високу ефективність, досягнувши Accuracy = 0,986, Recall = 0,977, Precision = 0,985 та F1-score = 0,981. Розроблений підхід забезпечує стабільну роботу моделі під час опрацювання зображень у реальному часі та може бути інтегрований у системи моніторингу, що використовують безпілотні літальні апарати. Результати дослідження свідчать про ефективність використання глибокого навчання для раннього виявлення лісових пожеж і відкривають перспективи подальших наукових розробок щодо підвищення енергоефективності роботи моделей, вдосконалення їхньої архітектури та створення нових наборів даних. Показано, що перспективним є напрям спрощення архітектури моделей сімейства YOLO задля зниження обчислювальної складності та підвищення енергоефективності роботи без істотного зниження показників точності, що є особливо важливим для використання таких моделей на безпілотних літальних апаратах із обмеженими ресурсами.

Ключові слова: штучний інтелект; згорткові нейронні мережі; модель YOLOv5; комп'ютерний зір; ідентифікація об'єктів; класифікація об'єктів.

Вступ / Introduction

Наразі лісові пожежі, що охоплюють значні площі, стали доволі поширеним явищем. Причиною цього є низка факторів, зокрема – зміни клімату – підвищення середньорічних температур, тривалі посухи, раптові хвилі спеки. Важливу роль відіграють також антропогенні чинники, пов'язані з відпочинком людей на природі та туризмом. Унаслідок таких пожеж було спалено тисячі гектарів землі, багато людей втратили домівки, дехто – навіть життя. Значної шкоди зазнала інфраструктура, а економічні втрати в окремих регіонах є надзвичайно великими.

Окрім цього, викиди вуглекислого газу від масштабних лісових пожеж можуть бути дуже значними. Наприклад, у 2023 р. в Канаді [5] сталося багато лісових пожеж, і загальні викиди CO₂ в атмосферу становили близько 647 TgC – у 4,3 раза більше, ніж усі промислові ви-

киди CO₂ у країні за той самий рік. За останні десятиліття проблема лісових пожеж посилилася і в Україні, що пов'язано зі змінами клімату, змінами в суспільстві та землекористуванні. Зокрема, катастрофічні пожежі у природних ландшафтах у 2014 р., 2015 р. та 2020 р. [40], дві великі пожежі в Чорнобильській зоні відчуження, які в 2015 р. охопили загальну площу близько 14,9 тис. га, а також катастрофічні пожежі, які трапились у 2016–2018 рр. в Австралії, Греції, Іспанії, Каліфорнії, Німеччині, Португалії, Швеції [39], свідчать, що Україна перебуває у нових, за рівнем пожежної небезпеки, кліматичних умовах [40]. При цьому ситуація ще більше ускладнилася з початком широкомасштабної агресії Російської Федерації та активними бойовими діями протягом останніх років.

Загалом літо – сезон, під час якого трапляється найбільше лісових пожеж. Це зумовлено високими темпе-

Інформація про авторів:

Вдовін Богдан Богданович, магістр, кафедра систем штучного інтелекту.

Email: bohdan.vdovin.mkns@lpnu.ua; <https://orcid.org/0009-0004-4698-4284>

Яковина Віталій Степанович, д-р техн. наук, професор, кафедра систем штучного інтелекту.

Email: vitality.s.yakovyna@lpnu.ua; <https://orcid.org/0000-0003-0133-8591>

Томашевський Олег Михайлович, канд. техн. наук, доцент, кафедра систем штучного інтелекту.

Email: Oleh.M.Tomashevskiy@lpnu.ua; <https://orcid.org/0009-0009-9134-7520>

Цитування за ДСТУ: Вдовін Б. Б., Яковина В. С., Томашевський О. М. Виявлення лісових пожеж на поточних зображеннях з БПЛА методами глибокого навчання. Науковий вісник НЛТУ України. 2025, т. 35, № 5. С. 97–107.

Citation APA: Vdovin, B. B., Yakovyna, V. S., & Tomashevskyy, O. M. (2025). Detection of forest fires in UAV streaming images using deep learning methods. *Scientific Bulletin of UNFU*, 35(5), 97–107. <https://doi.org/10.36930/40350511>

ратурами, тривалими періодами посухи та низькою кількістю опадів у деяких регіонах. Водночас, пожежі можуть виникати й восени, взимку або ранньою весною – особливо на відкритих територіях, які мають велику щільність дерев.

Початком виникнення пожежі можуть бути три чинники: матеріал, що може добре горіти; джерело тепла, наприклад іскра; кисень. Основна небезпека лісових пожеж полягає в тому, що з часом їх стає дедалі складніше ліквідувати. Тому навіть незначне займання у лісі може швидко перерости у масштабну пожежу, яку вже можна вважати екологічною катастрофою. Надзвичайно важливим є те, як швидко пожежа буде виявлена і відповідні дії з її припинення будуть застосовані [4, 12, 15].

Сучасні методи виявлення лісових пожеж часто передбачають використання давачів диму, супутникових систем спостереження та CCTV-камер. Останнім часом також зростає популярність застосування різних типів безпілотних літальних апаратів (БПЛА), що працюють на різних висотах. Безпілотники на основі крил зазвичай здійснюють моніторинг на висотах від 350 до 5000 м, тоді як апарати на пропелерах, які володіють значно більшою маневреністю, – на висотах від 10 до 350 метрів. Використання окремих методів не заперечує їх комбінування, оскільки комплексний підхід часто забезпечує навіть кращу ефективність.

Основними недоліками давачів диму є їх невеликий радіус дії, тому їхнє використання є дуже обмеженим. Інші підходи передбачають використання штучного інтелекту, зокрема – комп'ютерного зору, для виявлення вогню та диму часто на підставі RGB-зображень, інфрачервоних знімків, додаткових метеорологічних параметрів, як-от вологість повітря, температура і тиск.

Використання супутникових знімків, попри очевидні переваги, має також недоліки. Зокрема – відсутність можливості виявляти невеликі осередки займання чи диму, що не дає змоги виявляти лісові пожежі на ранніх стадіях її появи.

CCTV-камери зазвичай розміщують по периметрах територій на відносно низькій висоті, це водночас обмежує їхній огляд тільки навколишніми ділянками. І навіть, якщо є змога розмістити їх досить високо над землею [13], то це тільки часткове рішення, що не працюватиме для складних ландшафтів.

У цьому контексті БПЛА є одним із найперспективніших засобів не тільки для виявлення лісової пожежі, а й для моніторингу, локалізації та підтримки операцій із гасіння. Їхня ефективність зросла особливо за останні роки завдяки розвитку технологій штучного інтелекту та глибокого навчання моделей. Проте, також існують деякі труднощі їхнього використання: 1) складний ландшафт на зображеннях; 2) об'єкти часто мають велику варіативність у розмірах; 3) часті зміни зображень, зміна кута огляду, позиції, нахилу камери тощо; 4) обмежені обчислювальні ресурси, які впливають на тривалість роботи безпілотника на одному заряді батареї. Усі ці фактори мають істотне значення під час застосування БПЛА у завданнях виявлення та моніторингу лісових пожеж.

Об'єкт дослідження – виявлення лісових пожеж за даними, отриманими із камер безпілотних літальних апаратів.

Предмет дослідження – алгоритми глибокого навчання моделей для ідентифікації лісових пожеж у реальному часі із зображень, знятих безпілотними літаль-

ними апаратами, що дасть змогу ефективно натренувати детектор та класифікатор.

Мета роботи – розробити моделі на підставі їх глибокого навчання для підвищення ефективності виявлення лісових пожеж у реальному часі зі зображень, знятих безпілотними літальними апаратами, що дасть змогу встановити переваги, недоліки та область застосування запропонованого підходу.

Для досягнення зазначеної мети визначено такі основні завдання дослідження:

- дослідити сучасні методи та алгоритми глибокого навчання моделей, які використовують для виявлення лісових пожеж, зокрема – на їх ранніх стадіях появи, що дасть змогу обґрунтовано визначити базову модель для створення удосконаленої моделі виявлення лісових пожеж;
- розробити новий набір даних, який міститиме зображення лісових пожеж із класами "вогонь" та "дим". Зображення повинні бути зібрані із використанням БПЛА, або відповідати поставленій задачі. Також набір даних повинен містити достатню кількість зображень, що не містять міток, що дасть змогу ефективно натренувати детектор і класифікатор;
- розробити модель глибокого навчання для виявлення лісових пожеж та здійснити її програмну реалізацію, що дасть змогу підвищити точність виявлення лісових пожеж на підставі зображень із дронів;
- провести тестування, порівняння та аналіз моделей глибокого навчання, що дасть змогу встановити переваги, недоліки та область застосування запропонованого підходу.

Аналіз останніх досліджень та публікацій. У роботі [28] дослідники порівнювали три моделі завдання класифікації лісової пожежі (полум'я та дим), а саме VGG16, InceptionV3 та Xception. Для цього вони застосували навчання без забування (LwF) до уже навчених моделей для того, щоб вирішити проблему катастрофічного забування під час доналаштування моделей для завдання класифікації вогню та диму. Модель Xception продемонстрували найкращу точність у 96,89 %, що є найкращим результатом серед розглядуваних моделей.

Автори [16] розглянули завдання виявлення об'єктів у контексті завдання лісових пожеж. Щоб вилучити об'єкти, які часто спричиняють появу помилкових спрацювань (false-positives) під час виявлення вогню чи диму, дослідники внесли деякі додаткові класи (чорний дим, сірий дим, хмари, світло та вогонь). Основою для моделі стали YOLOv5, ResNeXt-50 та EfficientNet-b4. Запропонована вченими модель складається з ансамблю класифікаторів з використанням Late Average Fusion (LAF), а також моделі для виявлення об'єктів. Авторам вдалося отримати деякі покращення mAP@50 на 4,3 % у завдання виявлення об'єктів порівняно з моделлю YOLOv5s на їхньому наборі даних.

У роботі [13] автори вирішували проблему виявлення лісових пожеж у розрізі проєкту CICLOPE. Сам проєкт є використанням башт спостереження з високороздільними камерами, які покривають понад 2,7 млн га лісів Португалії. Метою було покращити точність системи і зробити її менш вразливою до помилкових спрацювань. Набір даних було зібрано на підставі відео та знімків із CICLOPE-башт. Для цього автори використали Dual-Channel CNN на підставі DenseNet та SCAM-SCNN з використанням CA (англ. *Channel Attention*), як механізму уваги. Це дало дослідникам змогу зменшити кількість негативних спрацювань до 0,20 % на їхньому наборі даних. Утім, сам механізм уваги не був повноцінно досліджений авторами, а отримані результати є

значущими тільки в контексті португальських лісів та проєкту CICLOPE. Вони можуть не відтворюватися на інших вибірках або загалом у завданнях виявлення лісових пожеж через обмежену варіативність місцевості.

Автори [2, 6, 7, 18] досліджували різні архітектури моделі YOLO в контексті завдання виявлення лісових пожеж у реальному часі, зокрема – версії моделей YOLOv5, YOLOv6, YOLOv7, YOLOv8, YOLO-NAS. Загалом найкращі результати продемонстрували версії моделей YOLOv5 та YOLOv8.

У роботі [17] вирішували проблему класифікації з використанням EfficientNetB7 та ACNet (англ. *Attention Guided Network*). Автори отримали приріст з F1-score 94,09 % на наборі даних FLAME [29] до 97,12 %. Загалом підхід дав змогу краще помічати мале полум'я порівняно з базовою моделлю.

У роботі [3] автори запропонували модель на підставі поєднання архітектури моделі YOLOv5s та ансамблю на підставі XceptionNet, MobileNetV2 та ResNet-50. Процес здійснюється у два етапи: спершу ансамблевий класифікатор голосуванням визначає наявність ознак диму/вогню в кадрі, після чого детектор YOLO локалізує відповідні області. Унаслідок навчання моделей класифікатор досяг F1-score 0,95 та точності 0,99, тоді як для виявлення отримано $mAP@0,5 = 0,76$ за поєднання двох моделей. Потрібно зазначити, що автори стикнулися з проблемою малої кількості хороших зображень та відео справжніх лісових пожеж, знятих за допомогою БпЛА. Також дослідники наголосили, що запропонований підхід, у контексті роботи в обмежених ресурсах БпЛА, не є ідеальним стосовно енергоефективності роботи.

Також автори роботи [21] розглянули використання системи виявлення на підставі відео з використанням Faster R-CNN для виявлення потенційних регіонів пожежі (SRoF) з подальшим зберіганням ознак в LSTM, де фінальне рішення обчислюється на підставі більшості голосів (англ. *majority voting*). У межах дослідження науковці зібрали набір даних, що містить 1309 відеокліпів. Faster R-CNN показав $mAP@50 = 88,3$ %. На рівні відео найкраща конфігурація LSTM (512 прихованих одиниць) дала 95,0 % точності; додавання короткого голосування більшістю на 10 с підняло точність до 97,92 % за рахунок зниження FP до 2,47 % та FN до 1,38 %. Загалом зі збільшенням вікна можна істотно зменшити кількість хибних спрацювань, але це не є end-to-end-рішення, оскільки дві моделі тренуються окремо, та і такий підхід не прийнятний для виявлення зображень в реальному часі і зовсім не є енергоефективним, щоправда, використання його як додаткового механізму для зменшення кількості негативних спрацювань і збільшення точності може бути.

У дослідженні [35] було розглянуто розробку полегшеної моделі для виявлення об'єктів із зображень у градаціях сірого для БпЛА на підставі моделі YOLOv8. Для цього дослідники застосували цілу низку покращень, спрямованих на зменшення моделі та виявлення малих об'єктів на різнопланових зображеннях із складним заднім фоном. Зокрема – було використано ShuffleNetV2 як основу для мережі та Multi Scale-Dilated Attention, що дало змогу зменшити кількість обчислень (GFLOPs) на 65,3 % порівняно із звичайною моделлю. Для покращення виявлення малих об'єктів було використано Wise-IoU-функцію витрат. Загалом авторам

вдалося отримати хороший компроміс для UAV-задач, істотно зменшивши складність підрахунків, та навіть дещо отримати кращі значення метрик для $mAP@50$ та F1-score. Проте, як зазначили самі автори, ця модель була тільки перевірена на зображеннях з одним відтінком, а тому може не демонструвати покращення на зображеннях з кількома каналами, як-от RGB.

У роботі [30] автори запропонували архітектуру на підставі YOLOv11n для системи виявлення лісових пожеж у реальному часі. Основні покращення стосувалися зміни блоків для отримання ознак на C3k2MBNV2 та C3K2WTDC, які поєднали раніше відомі ідеї з MobileNetV2 та вейвлет-згортки. Це дало змогу істотно зменшити кількість параметрів моделі та кількість обчислень. SCDwn блок було використано для того, щоб збільшити кількість збереженої інформації під час операції зменшення роздільності (downsample). Унаслідок вдалося досягти $mAP@50 = 91,6$, кількості параметрів 1,2 млн та GFLOPs = 4,5.

Можна стверджувати, що наразі архітектура YOLO є провідною у завданнях виявлення об'єктів, особливо в системах реального часу через високу швидкодію (FPS) та ефективність. У контексті завдання виявлення лісових пожеж чудові результати демонструють YOLOv5 та YOLOv8. Важливими є дослідження, націлені на зменшення кількості параметрів та кількості обчислень для виявлення та класифікації об'єктів (GFLOPs). Найбільше це стосується edge-пристроїв, враховуючи й безпілотні літальні апарати, де енергоефективність є однією із ключових характеристик. Іншою проблемою є кількість негативних спрацювань. Зокрема – у багатьох працях було помічено, що моделі часто класифікують хмари та туман як дим, або сонячне світло – як полум'я. Також однією з проблем є виявлення малих об'єктів на зображеннях, переважно це стосується зображень, знятих БпЛА на камеру, і є важливим для завдання раннього виявлення лісової пожежі.

Результати дослідження та їх обговорення / Research results and their discussion

Формування набору даних. Підготовка набору даних для завдання виявлення вогню/полум'я та диму з безпілотних літальних апаратів є доволі складним і клопітким процесом. Це пов'язано з великою варіативністю ландшафтів і погодних умов, а також з обмеженою кількістю реальних відео- чи фотоматеріалів пожеж, знятих із дронів.

Передусім, набір даних повинен містити зображення полум'я та диму, отримані з висоти до 150 м і з достатньо високою роздільною здатністю. На зображеннях мають бути представлені як невеликі осередки займання, так і пожежі середнього та великого масштабу. Окрім цього, важливо забезпечити різноманітність даних – зображення повинні охоплювати широкий спектр типів лісів і різні умови знімання.

Окрему увагу потрібно приділити наявності у наборі даних зображень, які можуть бути помилково сприйняті моделлю як вогонь чи дим – наприклад, відблиски сонця, світло ліхтарів, туман або хмари. Це допоможе підвищити точність роботи моделі та зменшити кількість хибно позитивних спрацювань. Тому для підготовки набору даних було використано кілька джерел.

Boreal Forest Fire: UAV-collected Videos of Fire Events in Finland [27]. Цей набір даних є обширним на-

бором відео та зображень керованих лісових пожеж у Фінляндії, знятих дроном DJI Phantom 4 із роздільною здатністю 4096×2160. Для нашої задачі було використано відео з локацій Eva, Karkkila та Ruokolahti. Отримані відео було конвертовано у набори зображень із частотою 1 кадр на секунду (1 FPS). Після цього зображення були попередньо оброблені: вилучено дублікати та надто схожі кадри, а також створено анотації для двох класів – полум'я та дим.

FLAME3: UAV-Collected Radiometric Thermal and Visual Spectrum Wildfire Imagery [14]. Цей набір даних містить зображення лісових пожеж, знятих дроном DJI M30T із роздільною здатністю 4000×3000. Хоча набір даних містить понад 600 зображень, значна їх частина повторюється або є дуже схожою між собою. Тому для використання у нашій роботі набір було додатково очищено від надлишкових зображень, після чого для нього створено власні анотації.

D-Fire: An Image Dataset for Fire and Smoke Detection [9]. Цей набір даних містить різноманітні зображення пожеж із відповідними анотаціями для класів вогонь та дим. Однак, тільки незначну частину зображень було знято безпілотними літальними апаратами. Більшість кадрів отримано з відеозаписів камер спостереження у міських умовах, на парковках та поблизу парків. З огляду на специфіку нашої задачі, ми здійснили додаткову фільтрацію та відібрали тільки ті зображення, що відповідають умовам знімання із дронів.

FASDD: An Open Flame and Smoke Detection Dataset for Deep Learning in Remote Sensing-Based Fire Detection [33]. Цей набір даних містить велику кількість зображень лісових пожеж, отриманих за допомогою різних пристроїв, зокрема – і безпілотних літальних апаратів. Проте, під час перегляду виявлено низку зображень, що потребували фільтрації та корекції (усунення дублікатів, видалення неякісних зображень тощо).

Унаслідок проведеної роботи було сформовано узагальнений набір даних, який містить 4104 зображення. Для них створено 3298 анотацій, що позначають клас полум'я, та 2788 анотацій для класу дим. Усі анотації мають формат YOLO. Після цього сформований набір даних було поділено на три окремих набори: тренуваль-

ний, валідаційний та тестувальний зі співвідношенням кількості зображень 7:2:1. При цьому було збережено розподіл зображень у кожному створеному наборі даних.

Архітектура моделі. Для використання в роботі як базову архітектуру було обрано модель YOLOv5 [20]. Це рішення було здійснено після попереднього аналізу літератури з теми, де ця модель демонструвала одні з найкращих результатів, а також деяких попередніх експериментальних тестів.

Модель YOLOv5 використовує трирівневу архітектуру (рис. 1), яка складається із Backbone-, Neck- та Head-компонент. Backbone мережі становить CSPDarknet53 – варіант базового екстрактора ознак, у якому використано принцип Cross Stage Partial connections.

Для агрегації контексту в YOLOv5 застосовують модифікований модуль Spatial Pyramid Pooling – Fast (SPPF), який забезпечує узагальнення просторової інформації з кількох масштабів. Це дає змогу моделі враховувати як локальні, так і глобальні контексти, що є важливим для виявлення об'єктів різного розміру або з нечіткими межами, наприклад, зон диму.

У компоненті Neck використовують Path Aggregation Network (PANet), який покращує передачу ознак між шарами різної глибини. Такий підхід забезпечує поєднання детальних просторових характеристик із високорівневою семантикою, що підвищує точність локалізації об'єктів.

У модулі Head реалізовано багаторівневе виявлення на трьох масштабах ознак, що дає змогу виявляти як великі, так і малі об'єкти на одному зображенні.

Окрім цього, YOLOv5 застосовує механізм автоматичного підбору розмірів для рамок "якорів", це здійснюється на підставі оптимізації параметрів під конкретний набір даних, ще перед початком тренування моделі. Що, водночас, зменшує потребу у ручному налаштуванні, як це було у попередніх версіях.

Навчання відбувається з використанням детектора YOLOv5s та додаткового класифікатора, який використовує карти ознак, отримані із модуля Backbone детектора, для здійснення багатокласової класифікації. Архітектуру класифікатора зображено на рис. 2.

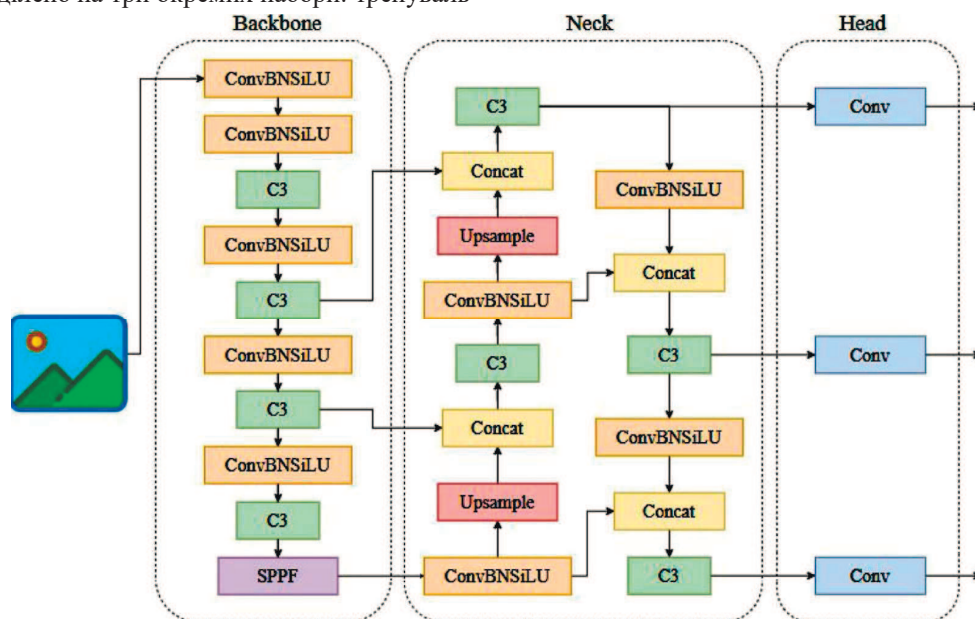


Рис. 1. Загальна архітектура нейронної мережі для виявлення об'єктів YOLOv5 / General architecture of the YOLOv5 neural network for object detection [20]

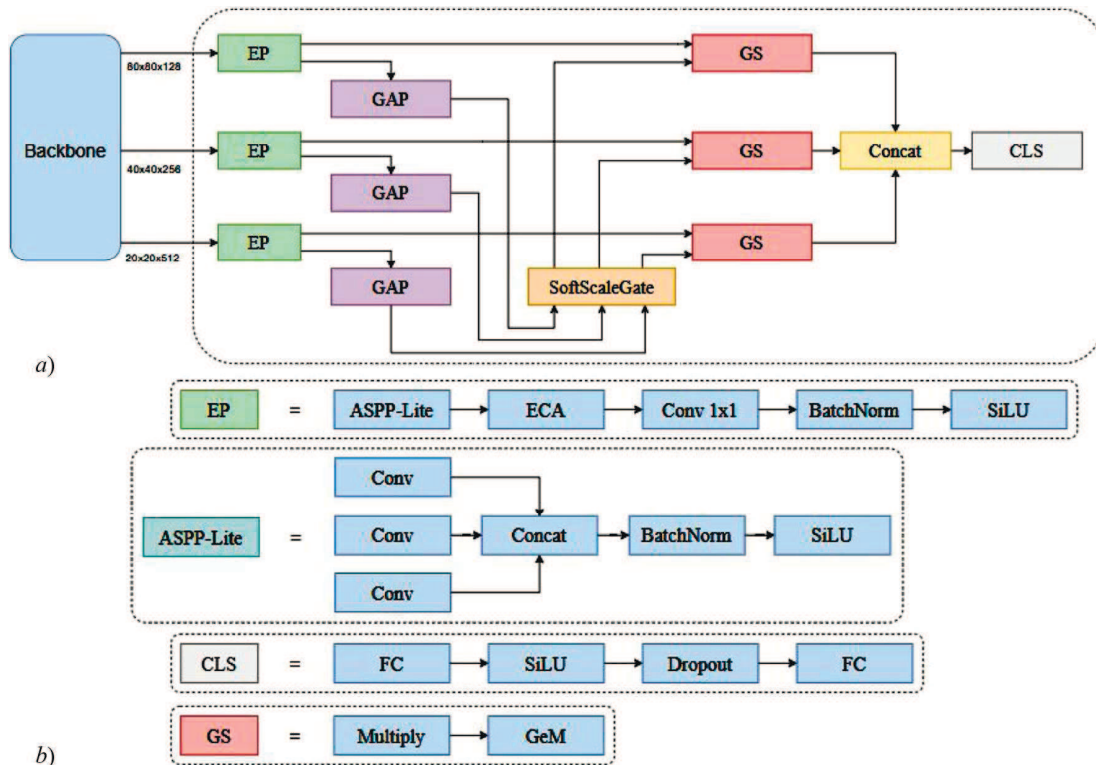


Рис. 2. Архітектура розробленого класифікатора ForFv3 / Architecture of the developed ForFv3 classifier

У цьому класифікаторі отримані карти ознак (P_3, P_4, P_5), що потрапляють на вхід моделі, окремо передаються до блоку EP. У середині цього блоку ознаки спочатку потрапляють у шар ASPP-Lite. Цей шар є спрощеним варіантом Atrous Spatial Pyramid Pooling (ASPP), який застосовують для покращення сприйняття контексту різних масштабів. У ньому застосовано кілька паралельних згорток із різними розмірами рецептивного поля, результати яких об'єднуються конкатенацією [8]:

$$G = \text{Concat}(\text{Conv}_1(X), \text{Conv}_2(X), \text{Conv}_3(X)). \quad (1)$$

Після нормалізації (BatchNorm) та активації SiLU отримуємо початково узагальнену карту ознак.

Далі використовують шар ECA (Efficient Channel Attention) [34], який підсилює найінформативніші канали (depth). Наступним етапом є згортка з ядром 1×1 , що зводить усі карти ознак до єдиної глибини. Після цього застосовують BatchNorm та SiLU. Отже, модуль EP завершує оброблення даних, створюючи карти ознак Z_p єдиної розмірності для всіх масштабів.

На наступному етапі до кожної карти Z_p застосовують Global Average Pooling (GAP) [23]:

$$d_{p,c} = \frac{1}{H_p W_p} \sum_{i,j} Z_{p,i,j,c}, \forall p \in P, \forall c \in C, \quad (2)$$

де: $d_{p,c}$ – значення каналу, для якого GAP обчислює середнє значення всіх значень на карті ознак цього каналу; H_p – висота обраної карти ознак; W_p – ширина обраної карти ознак; $Z_{p,i,j,c}$ – значення в карті ознак Z_p , де p – індекс, що вказує на обрану карту ознак, що приймає значення з множини $\{3, 4, 5\}$; i – координата першого виміру (ширини), j – координата другого виміру (висоти); c – координата третього виміру (глибини).

Після цього отримані вектори передаються у блок SoftScaleGate. Цей блок виконує каналне зважування між різними масштабами, обчислюючи ваги за допомогою функції softmax, а саме

$$w_{p,c} = \frac{\exp(d_{p,c} / \tau)}{\sum_k \exp(d_{k,c} / \tau)}, \forall p \in P, \forall c \in C, \quad (3)$$

де τ – параметр температури, який контролює ступінь розподілу ваг.

Після цього кожна масштабована карта $\hat{Z}_p = \omega_p \odot Z_p$ проходить через блок GS, у якому виконується GeM (англ. *Generalized Mean Pooling*) [36]:

$$p_{p,c} = \sqrt[r]{\frac{1}{H_p W_p} \sum_{i,j} (\hat{Z}_{p,i,j,c})^r}, \forall p \in P, \forall c \in C, \quad (4)$$

де: $\hat{Z}_{p,i,j,c}$ – значення в карті ознак Z_p , де p – індекс, що вказує на обрану карту ознак, що приймає значення з множини $\{3, 4, 5\}$; i – координата першого виміру (ширини), j – координата другого виміру (висоти), c – координата третього виміру (глибини), r – параметр GeM, що визначає ступінь узагальнення між середнім і максимальним пулінгом.

Після цього, отримані вектори ознак з усіх масштабів, об'єднуються операцією конкатенації і потрапляють у блок CLS. Цей блок складається із двох повністю зв'язаних шарів (fully connected layer) з функцією активації SiLU і шаром Dropout для регуляризації.

Далі обчислюється функція витрат, використовуючи бінарну крос-ентропію з логітами (англ. *Binary Cross-Entropy with Logits Loss*) для багатокласової класифікації [19, 26].

$$L = -\frac{1}{C} \sum_{c=1}^C [y_c \log \sigma(\hat{y}_c) + (1 - y_c) \log (1 - \sigma(\hat{y}_c))]. \quad (5)$$

Бінарна перехресна ентропія з втратою логітів – це комбінована сигмоподібна активаційна функція та бінарні перехресні ентропійні втрати, які використовують для бінарної класифікації даних. Вона чисельно стабільна, оскільки поєднує ці два кроки, і приймає необроблені "логіти" як вхідні дані, спрощуючи вивід моделі та код. Це означає, що користувачу не потрібен окремий сигмоподібний шар у розробленій моделі або не потрібно обчислювати функцію втрат з ймовірностей їхньої появи.

Для об'єднаного тренування детектора та класифікатора в цій роботі використано модифікований алгоритм LBTW [24], механізм дії якого зображено на рис. 3. Ідея підходу полягає у додатковому зважуванні втрат від завдання класифікатора, а коефіцієнт β використовують для надання пріоритету певній задачі, він приймає дійсні значення від 0 до 1. Надалі називатимемо таку модель ForFv3.

Algorithm 1: LBTW з використанням коефіцієнта впливу β

Input: Пакет даних B , ваги нейронної мережі W , T задач класифікації, K задач детекції та параметра β

Output: Оновлені ваги нейронної мережі W

Ініціалізувати ваги нейронної мережі W

for кожну епоху i **do**

for кожен пакет даних B **do**

 Обчислити втрати для кожної задачі класифікації $l_{(B,t)}$

 Зберегти втрати для першого пакета як $l_{(0,i)}$

 Обчислити втрати для кожної задачі детекції $l_{(B,k)}$

for кожну задачу t **do**

 Встановити вагу задачі $w_t = \left(\frac{l_{(B,t)}}{l_{(0,i)}}\right)^\alpha$

 Оновити зважену втрату $l_{(B,t)} = l_{(B,t)} \cdot w_t$

end

 Обчислити загальну втрату для пакета:

$$l_t = \frac{\beta}{|K|} \sum_{k=1}^K l_{(B,k)} + \frac{1-\beta}{|T|} \sum_{t=1}^T l_{(B,t)}$$

 Оновити ваги W відносно l_t

end

end

Рис. 3. Удосконалений алгоритм роботи LBTW з використанням коефіцієнта впливу β / Improved LBTW algorithm with the influence factor β

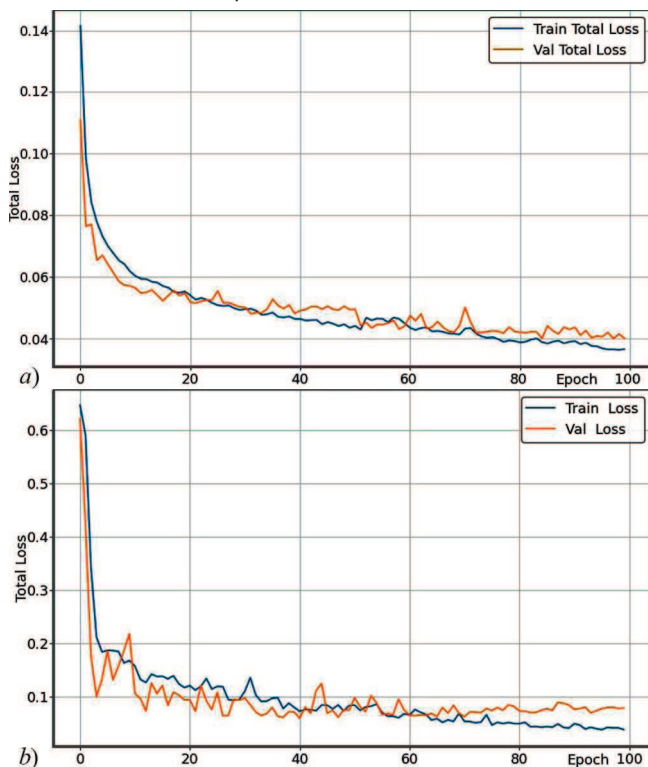


Рис. 4. Графіки втрати під час навчання для детектора та класифікатора ForFv3 / Training loss dependencies for the ForFv3 detector and classifier

Тренування моделі. Розроблення та навчання моделі відбувалося мовою Python 3.12.11 з використанням фреймворку PyTorch 2.8. Модель тренували у середовищі Google Collab з використанням графічного процесора Nvidia A100 (розмір пам'яті 40GB). Навчання здійснювалося упродовж 100 епох, розмір пакета (batch size) 16, розмір зображень $640 \times 640 \times 3$. Значення параметрів

впливу β становило 0,7. Графіки функцій втрат для класифікатора та детектора запропонованої моделі наведено на рис. 4 як для тренувального, так і для тестового наборів даних.

Табл. 1. Оцінка метрик детектора ForFv3 / Performance metrics of the ForFv3 detector

Модель ForFv3	Значення метрик				
	mAP@50	mAP@50-95	Recall	Precision	F1
Усі	0,917	0,619	0,875	0,911	0,892
Вогонь	0,871	0,498	0,815	0,871	0,840
Дим	0,963	0,741	0,935	0,955	0,944

З отриманих результатів виявлення об'єктів (табл. 1) помітно, що клас полум'я має значення метрик, які значно поступаються метрикам класу дим. Проте, для завдання розпізнавання лісових пожеж виявлення диму є більш пріоритетним, оскільки дим часто є помітнішим швидше, ніж полум'я, що є сигналом про можливу пожежу. Значення метрик класифікації виявлених об'єктів наведено в табл. 2.

Причиною гірших значень метрик передусім є те, що на більшості зображень, які вдалося отримати в наборі даних, полум'я має дещо складнішу природу, ніж дим. Дим часто виступає як один об'єкт, що спричинене самою його природою і тим, що він піднімається вгору. Полум'я ж часто на зображеннях виступає або малими запалюваннями – тоді модель здатна їх добре розпізнати – або, якщо пожежа вже переросла у середню, або малу, які об'єкти класу вогонь стає переважна більшість, і вони знаходяться дуже близько один до одного. У таких умовах модель часто бачить вогонь, але інколи плутає два вогнища як одне, або навпаки. Цю проблему доволі складно вирішити, якщо йдеться про виявлення об'єктів в контексті моделей YOLO, з використанням прямокутних рамок. Якщо виявлення чітких обрисів полум'я є важливим, то значно кращим підходом є використання моделей для сегментації зображень.

Табл. 2. Оцінка метрик класифікатора ForFv3 / Performance metrics of the ForFv3 classifier

Модель ForFv3	Значення метрик			
	Accuracy	Precision	Recall	F1
Вогонь	0,972	0,967	0,949	0,958
Дим	0,991	0,995	0,990	0,993
Усі	0,986	0,985	0,977	0,981

Для порівняння результатів також було здійснено навчання моделей YOLOv5s та YOLOv8s на 100 епохах, розмір пакета 16, розмір зображення $640 \times 640 \times 3$. Тренування здійснювалося на цьому ж наборі даних. На рис. 5-7 зображено приклади роботи всіх трьох моделей на зображеннях з досліджуваного набору даних, а в табл. 3 наведено значення метрик та потрібних обчислюваних ресурсів у GFLOPs.

Табл. 3. Порівняння результатів виявлення об'єктів / Comparing object detection results

Модель	Значення метрик					
	mAP@50	mAP@50-95	Recall	Precision	F1	GFLOPs
YOLO v5s	0,822	0,544	0,790	0,812	0,810	15,8
YOLO v8s	0,810	0,553	0,766	0,801	0,780	23,3
ForFv3	0,918	0,620	0,874	0,916	0,894	15,8

Згідно з даними табл. 3, модель YOLOv5s дає кращі результати, ніж YOLOv8s. А тому вибір моделі YOLOv5s, як основи для удосконаленої моделі ForFv3, виявився

виправданим. Понад це, оскільки модель YOLOv5s широко використовують в IoT, то багато її компонент доволі добре оптимізовані для безпосереднього застосування на пристроях з низькими обчислювальними можливостями, що полегшить її практичне застосування.

Також з табл. 3 видно, що запропонована модель ForFv3 дає кращі результати порівняно з базовою моделлю без використання LBTW.

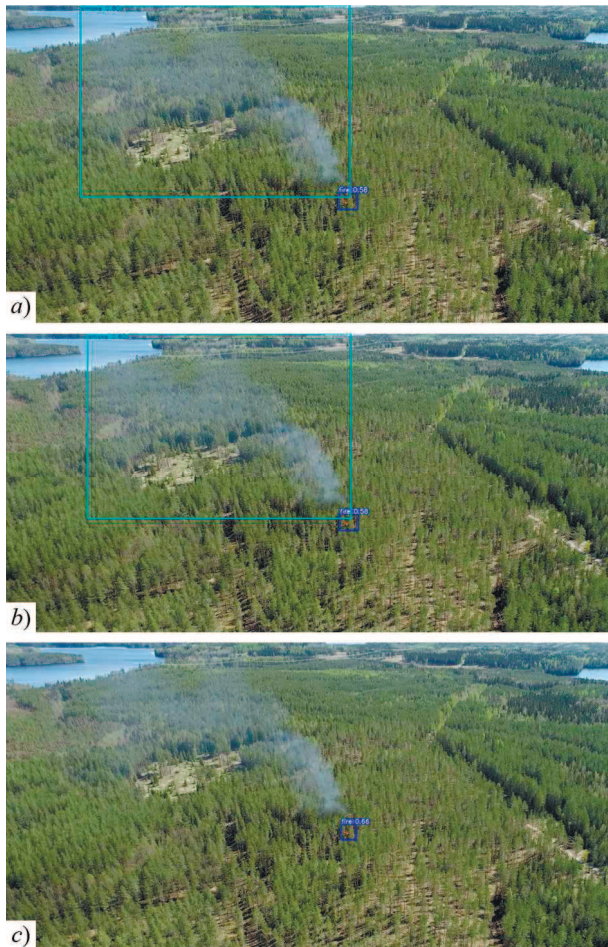


Рис. 5. Результати роботи моделей / Results of the models: a) модель ForFv3 / ForFv3 model; b) модель YOLOv5s / YOLOv5s model; c) модель YOLOv8s / YOLOv8s model

Отже, аналіз отриманих результатів показує, що детектор моделі ForFv3 продемонстрував покращені показники метрик, досягнувши $mAP@50 = 0,917$, $mAP@50-95 = 0,619$, $Recall = 0,875$, $Precision = 0,911$ та $F1-score = 0,892$. Порівняно з базовою моделлю YOLOv5s, навченою без спільного навчання моделей, запропонований підхід показав кращі результати, зокрема – метрика $mAP@50$ зросла на 11,6 % (0,917 проти 0,822). Класифікатор моделі продемонстрував високу ефективність, досягнувши $Accuracy = 0,986$, $Recall = 0,977$, $Precision = 0,985$ та $F1-score = 0,981$. Окрім цього, базова модель YOLOv5s є оптимізованою для безпосереднього застосування на пристроях з низькими обчислювальними можливостями, що дає можливість її практичного застосування для виявлення лісових пожеж на підставі зображень з безпілотних літальних апаратів.

Загалом можна зазначити, що застосований у цій роботі підхід, у поєднанні зі сформованим набором даних, дав змогу отримати кращі результати для завдання виявлення, особливо під час виявлення диму. Щодо класифікації, деякі публікації демонструють дещо вищі показники метрик, однак загалом отримані в цій роботі

результати є порівнюваними й конкурентними відносно наявних рішень.

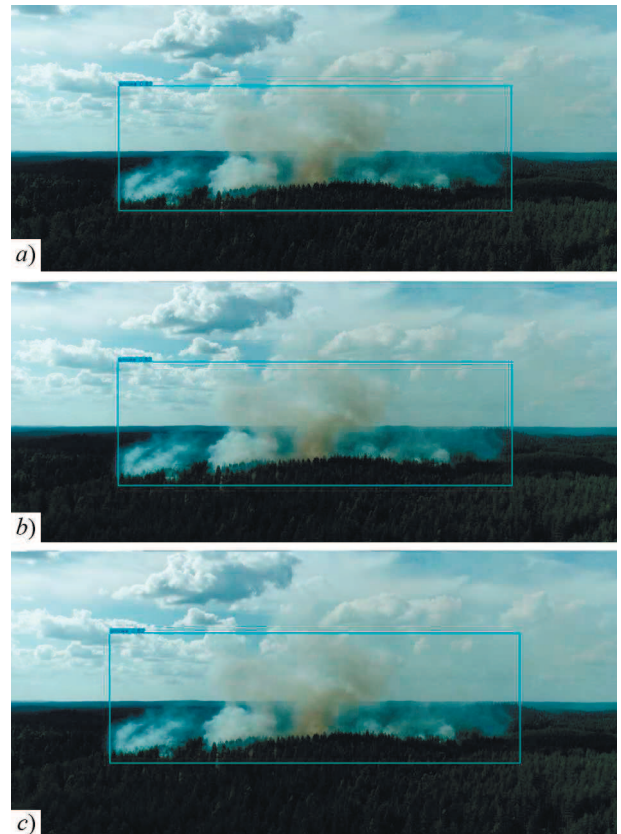


Рис. 6. Результати роботи моделей / Results of the models: a) модель ForFv3 / ForFv3 model; b) модель YOLOv5s / YOLOv5s model; c) модель YOLOv8s / YOLOv8s model

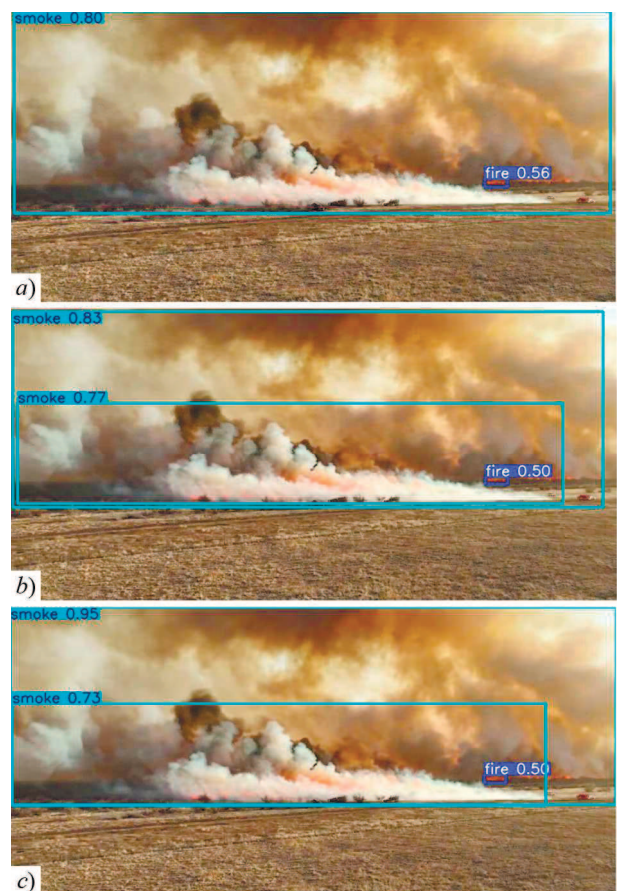


Рис. 7. Результати роботи моделей / Results of the models: a) модель ForFv3 / ForFv3 model; b) модель YOLOv5s / YOLOv5s model; c) модель YOLOv8s / YOLOv8s model

У подальших дослідженнях перспективним є напрям спрощення архітектури моделей сімейства YOLO задля зниження обчислювальної складності та підвищення енергоефективності роботи без істотного зниження показників точності. Це має особливе значення у контексті використання таких моделей на безпілотних літальних апаратах із обмеженими ресурсами.

Обговорення результатів дослідження. Важливість завдання виявлення лісових пожеж підтверджується численними дослідженнями у цій галузі. Значна частина робіт зосереджена на удосконаленні методів класифікації лісових пожеж, їх виявлення та сегментації, а також на спрощенні моделей для їх ефективного використання у системах реального часу, зокрема – в безпілотних системах спостереження або системах відеомоніторингу.

У роботі [1] автори розробили модель штучного інтелекту для виявлення пожежі за зображеннями лісового ландшафту. Ця стаття надала нового розуміння надійного та наукового моделювання для виявлення лісових пожеж шляхом аналізу карт ознак на підставі архітектур поєднаних трансформерів за допомогою глибоких нейронних мереж. Результати перевірки показують, що запропонована авторами гібридна модель досягає метрик точності 99,58 та 96,79 % для двох наборів даних з використанням 5-кратної кросвалідації. Окрім цього, висока чутливість та висока точність означають, що модель має мало хибнонегативних та хибнопозитивних результатів. Враховуючи отриману точність, запропоновану гібридну модель може бути використано для комплексного моделювання виявлення пожеж. Також автори роботи вважають, що завдяки використанню запропонованої моделі на БпЛА, оснащеному системою реального часу, виявлення пожеж надасть підтримку рішенням польових фахівців щодо зменшення збитків та управління лісовими пожежами.

Автори роботи [10] оцінили вплив трьох методів покращення зображення – вирівнювання гистограми (HE), адаптивного вирівнювання гистограми з обмеженим контрастом (CLAHE) та гібридного методу під назвою DBST-LCM CLAHE – на продуктивність моделі виявлення об'єктів YOLOv11 у виявленні пожеж та диму. Для навчання та оцінювання моделі використовували набір даних D-Fire, що складається з 21527 анованих зображень, отриманих за різних сценаріїв навколишнього середовища та рівнів освітлення. Кожен метод покращення було застосовано до набору даних перед навчанням. Продуктивність моделі оцінювали за допомогою кількох показників, враховуючи Precision, Recall, mean Average Precision at 50 % IoU (mAP50), F1-score та візуальний аналіз за результатами обмежувальної рамки. Експериментальні результати показують, що всі три методи покращення підвищили продуктивність виявлення. Метод HE дав найвищий показник mAP50 0,771, а також збалансовану Precision 0,784 та Recall 0,703, що демонструє високий ступінь узагальнення в різних умовах. DBST-LCM CLAHE досяг найвищого значення Precision – 79 %, що ефективно зменшує кількість хибнопозитивних результатів, особливо у сценах з розсіяним димом або складними текстурами. Метод CLAHE, з дещо нижчими загальними показниками, сприяв покращенню виявлення локальних ознак. Кожен метод продемонстрував чіткі переваги: HE покращив глобальний контраст; метод CLAHE покращив види-

мість локальної структури; а DBST-LCM CLAHE забезпечив оптимальний баланс завдяки динамічному визначенню розміру блоків та збереженню локального контрасту. Ці результати вказують на важливість вибору методів попереднього оброблення відповідно до пріоритетів виявлення, таких як мінімізація хибних спрацювань. Результати роботи підтверджують інтеграцію методів попереднього оброблення для підвищення точності виявлення, пропонуючи основу для систем виявлення лісових пожеж у реальному часі на периферійних пристроях або БпЛА, особливо в таких регіонах, як Індонезія.

Наявні методи виявлення лісових пожеж стикаються з такими проблемами, як відсутність інформації про краї, неефективне багатомасштабне об'єднання ознак та низька точність під час ідентифікації малих або віддалених цілей у складних лісових умовах. Для вирішення цих проблем автори роботи [32] запропонували нову систему виявлення під назвою Динамічна контекстна неглибока мережа (DCSNet). Ця система підвищує продуктивність виявлення та ефективність роботи в режимі реального часу. Запропонований у цій роботі метод містить три ключові компоненти: (1) контекстну часткову перехресну мережу ознак (CPCFNet), яка використовує механізм перепараметризованої нелокальної уваги та часткове розділення каналів для посилення контекстного представлення; (2) динамічну багатомасштабну пірамідальну мережу об'єднання (DMFPN), яка використовує динамічну вибірку та деформовану згортку для адаптивного об'єднання багатомасштабних ознак; (3) шар виявлення неглибоких ознак (SFDL), який уточнює неглибокі ознаки для покращення виявлення малих цілей диму та полум'я. Експериментальні оцінювання наборів даних дистанційного зондування видимого світла та інфрачервоного випромінювання, БпЛА, продемонстрували ефективність DCSNet. Зокрема, DCSNet досягає значення метрики mAP@50 76,0 %, частоти кадрів 283 кадри на секунду та F1-score 71,9 % для зображень у видимому світлі. На наборі даних інфрачервоного випромінювання фреймворк досягає mAP@50 77,5 %, mAP@50:95 45,5 % та F1-score 72,9 %. Ці результати свідчать про те, що DCSNet забезпечує високу точність та надійність у виявленні лісових пожеж у режимі реального часу за різних умов навколишнього середовища.

Автори роботи [31] запропонували CEVG-RTNet – архітектуру виявлення диму від лісових пожеж у реальному часі, розроблену для підвищення стійкості за таких складних умов. Модель CEVG-RTNet містить кілька нових компонентів. Модуль просторово-канальної апріорної перцептивної згортки (SCPP-Conv) покращує здатність моделі локалізувати дим та сприймати його морфологію навіть у середовищах з низькою контрастністю та високою прозорістю. Модуль ієрархічного вирівнювання залишкових ознак (HRFA) вирішує проблему багатомасштабного вилучення ознак шляхом вирівнювання локальних та великомасштабних ознак диму за допомогою стратегії вирівнювання на підставі залишків та агрегації на підставі багатопередового перцептрона. Для подальшого вдосконалення динамічного виявлення диму модуль динамічного рекурсивного покращення ознак (DRFE) застосовує рекурсивне адаптивне покращення каналу та стратегії міжканальної уваги. Окрім цього, введено нову функцію втрат Polygonal-Intersection over Union (PolyIoU) Loss для оброблення морфологічної складності областей диму. Архітектура

використовує механізм розрідженої уваги на графах для підвищення точності без надмірних обчислювальних витрат. Експериментальні результати, які отримали автори, демонструють ефективність моделі CEVG-RTNet, при цьому варіант моделі CEVG-RTNet-n досягає значення метрик Precision 89,1 %, Recall 82,9 %, mAP@0,5 на рівні 89 % та mAP@0,5:0,95 на рівні 58,9 %. Модель демонструє високий ступінь узагальнення та придатність для складного виявлення диму від лісових пожеж.

Для вирішення проблеми неоптимального виявлення лісових пожеж у складних середовищах на середніх та пізніх стадіях, автори [38] розробили новий алгоритм виявлення об'єктів – FF-net. Автори використовували набори даних про лісові пожежі та полум'я з повітря та, використовуючи 16 метрик, порівняли експерименти із класичними алгоритмами CNN, алгоритмами SOTA (State-of-the-Art), алгоритмами на підставі трансформерів та сучасними алгоритмами виявлення пожеж, виявивши, що загальна ефективність виявлення FF-net перевершує інші алгоритми. Зокрема, значення Precision перевищує продуктивність класичних алгоритмів CNN, алгоритмів SOTA та трансформерів на 6,46, 6,40 та 21,50 % відповідно. На наборі даних про полум'я середня Ассигасу є на 2,47 % вищою, ніж у інших методів.

Автори [22] розробили високоточну мережу виявлення диму від лісових пожеж. Для вирішення проблеми пропуску виявлення, вони розробили багатовимірний екстрактор вікон Swin (SMWE), який покращує обмін інформацією між вікнами як у горизонтальному, так і у вертикальному вимірах для вилучення глобальних текстурних ознак із зображень з димом. Потім було запропоновано мережу піраміди ознак гільйотини (GFPN) разом з новим методом згортки гільйотини для зменшення надлишкової інформації про ознаки з погляду об'єднання ознак, цим самим покращуючи стійкість моделі до перешкод. Окрім цього, враховуючи тонкість і нерівномірність диму поблизу меж, запропоновано адаптивну функцію втрат контурів для мінімізації розміття меж, спричиненого зменшенням дискретизації карти ознак у мережі. Експериментальні результати показали, що SMWE-GFPNNet досягає значень 80,92 % mAP, 90,01 % mAP50 та 83,38 % mAP75 на наборі даних комплексного фонових виявлення диму від лісових пожеж.

Робота [11] зосереджена на розробленні автоматизованої системи виявлення лісових пожеж за допомогою моделі глибокого навчання на підставі злиття (AFFD-FDL) для моніторингу навколишнього середовища. Метод AFFD-FDL містить розробку моделі злиття на підставі ентропії для вилучення ознак. Підхід поєднує створені вручну за допомогою гістограми градієнтів (HOG) ознаки з ознаками, отриманими за допомогою моделей глибокого навчання SqueezeNet та Inception v3. Окрім цього, для ідентифікації наявності пожежі використовували класифікатор на підставі оптимальної машини екстремального навчання (ELM). Для правильного налаштування параметрів моделі ELM застосовували алгоритм Oppositional Glowworm Swarm Optimization (OGSO), що покращує продуктивність виявлення лісових пожеж.

Автори [25] розробили систему виявлення лісових пожеж на підставі Інтернету речей з використанням інтеграції моделі котячого рою з моделлю LSTM. Через важливість виявлення об'єктів у дослідженні для реалізації було обрано модель EfficientDet. Запропоновано

виявляти спалахи пожеж за допомогою рекурентної нейронної мережі LSTM (RLSTM-NN). Автори використали алгоритм оптимізації дробового числення Cat Swarm (CSFCO) для глибокого навчання, який поєднує найкращі характеристики оптимізації Cat Swarm (CSO) з дробовим численням для оптимальних результатів навчання. Отримані результати показують, що запропонований підхід може ідентифікувати початок пожежі з точністю 98,6 % та коефіцієнтом помилок тільки 0,14 %.

Автори роботи [37] запропонували модифіковану архітектуру YOLOv5 для ефективного виявлення пожеж у розумних містах. Модифікована модель інтегрує модуль Stem у магістраль, замінює більші ядра меншими в SPP (шийці) та додає модуль P6 у голову архітектури. Запропонована модель має меншу складність та розмір, і здатна виявляти як малі, так і великі області пожежі на зображеннях. Окрім цього, в цій роботі створено набір даних про пожежі середнього масштабу, який складається з трьох класів (пожежа транспортних засобів, пожежа в будівлях та пожежа в приміщенні), з ручною анотацією. Для порівняння запропонованого підходу автори роботи використали 11 різних сучасних моделей виявлення об'єктів, які були навчені на створеному у роботі наборі даних. Автори зазначають, що запропонована модифікована архітектура YOLOv5 досягла кращої продуктивності виявлення та може бути застосована в реальних сценаріях.

Варто зазначити, що якісних наборів даних для завдання виявлення лісових пожеж існує небагато. Найкращих практичних результатів, як показують дослідження, вдається досягти під час використання даних, що відповідають конкретній місцевості та типу ландшафту, на відміну від універсальних наборів даних. Ймовірно, ця тенденція збережеться ще тривалий час через обмежену доступність репрезентативних та збалансованих даних. У нашій роботі ми акцентували увагу на помірних широтах у денну пору часу.

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

Наукова новизна отриманих результатів дослідження – удосконалено модель виявлення лісових пожеж на базі архітектури YOLOv5, яка, на відміну від наявної, використовує класифікатор ForFv3 та спільне їхнє навчання на підставі модифікованого алгоритму LBTW з використанням коефіцієнта впливу β , що дало змогу досягти високих результатів під час ідентифікації об'єктів та їх класифікації, зокрема – диму та вогню з безпілотних літальних апаратів.

Практична значущість результатів дослідження – натреновану модель глибокого навчання можна використовувати на безпілотних літальних апаратах для виявлення лісових пожеж у реальному часі, зокрема – диму та вогню, а також їх класифікації, що дає змогу своєчасно встановлювати осередки займання на ранніх стадіях і, відповідно, забезпечує можливість оперативного реагування на небезпеку та своєчасної її ліквідації.

Висновки / Conclusions

Розроблено модель на підставі її глибокого навчання для підвищення ефективності виявлення лісових пожеж в реальному часі зі зображень, знятих безпілотними літальними апаратами, що дало змогу встановити переваги, недоліки та область застосування запропоно-

ваного підходу. За результатами проведеного дослідження можна зробити такі висновки.

1. Здійснено аналіз сучасних підходів до виявлення лісових пожеж у режимі реального часу із використанням безпілотних літальних апаратів, а також методів глибокого навчання для класифікації та виявлення об'єктів на зображеннях.
2. Створено набір даних, що містить 4104 зображення із 3289 розміченими рамками класу "вогонь" та 2788 рамками класу "дим", отриманих із відеоматеріалів, знятих камерами дронів.
3. Розроблено та реалізовано модель ForFv3, що використовує механізм спільного навчання та алгоритм LBTW. Модель реалізовано мовою програмування Python із застосуванням фреймворку PyTorch.
4. Проведено навчання запропонованої моделі ForFv3 та здійснено порівняльний аналіз її результатів із моделями YOLOv5s і YOLOv8s, натренованих на тому самому наборі даних. За результатами експериментів детектор моделі ForFv3 продемонстрував покращені показники метрик, досягнувши $mAP@50 = 0,917$, $mAP@50-95 = 0,619$, $Recall = 0,875$, $Precision = 0,911$ та $F1-score = 0,892$. Порівняно з базовою моделлю YOLOv5s, навченою без спільного навчання моделей, запропонований підхід показав кращі результати, зокрема – метрика $mAP@50$ зростає на 11,6 % ($0,917$ проти $0,822$). Класифікатор моделі продемонстрував високу ефективність, досягнувши $Accuracy = 0,986$, $Recall = 0,977$, $Precision = 0,985$ та $F1-score = 0,981$.

References

1. Akyol, K. (2025). An innovative hybrid method utilizing fused transformer-based deep features and deep neural networks for detecting forest fires. *Advances in Space Research*, 75(12), 8583–8598. <https://doi.org/10.1016/j.asr.2025.04.020>
2. Al-Smadi, Y., Alauthman, M., Al-Qerem, A., Aldweesh, A., Quaddoura, R., Aburub, F., Mansour, K., & Alhmiedat, T. (2023). Early Wildfire Smoke Detection Using Different YOLO Models. *Machines*, 11(2), article ID 246. <https://doi.org/10.3390/machines11020246>
3. Bahhar, C., Ksibi, A., Ayadi, M., Jamjoom, M. M., Ullah, Z., Soufiene, B. O., & Sakli, H. (2023). Wildfire and Smoke Detection Using Staged YOLO Model and Ensemble CNN. *Electronics*, 12(1), article ID 228. <https://doi.org/10.3390/electronics12010228>
4. Bushnaq, O. M., Chaaban, A., & Al-Naffouri, T. Y. (2021). The Role of UAV-IoT Networks in Future Wildfire Detection. *IEEE Internet of Things Journal*, 8(23), 16984–16999. <https://doi.org/10.1109/JIOT.2021.3077593>
5. Byrne, B., Liu, J., Bowman, K. W., Pascolini-Campbell, M., Chatterjee, A., Pandey, S., Miyazaki, K., Van Der Werf, G. R., Wunch, D., Wennberg, P. O., Roehl, C. M., & Sinha, S. (2024). Carbon emissions from the 2023 Canadian wildfires. *Nature*, 633(8031), 835–839. <https://doi.org/10.1038/s41586-024-07878-z>
6. Casas, E., Ramos, L., Bendek, E., & Rivas-Echeverría, F. (2023). Assessing the Effectiveness of YOLO Architectures for Smoke and Wildfire Detection. *IEEE Access*, 11, 96554–96583. <https://doi.org/10.1109/ACCESS.2023.3312217>
7. Casas, E., Ramos, L., Bendek, E., & Rivas-Echeverría, F. (2024). YOLOv5 vs. YOLOv8: Performance Benchmarking in Wildfire and Smoke Detection Scenarios. *Journal of Image and Graphics*, 12(2), 127–136. <https://doi.org/10.18178/joig.12.2.127-136>
8. Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2016). *DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs* (Version 2). arXiv. <https://doi.org/10.48550/ARXIV.1606.00915>
9. De Venâncio, P. V. A. B., Lisboa, A. C., & Barbosa, A. V. (2022). An automatic fire detection system based on deep convolutional neural networks for low-power, resource-constrained devices. *Neural Computing and Applications*, 34(18), 15349–15368. <https://doi.org/10.1007/s00521-022-07467-z>
10. Dewi, C., Santoso, M. V. V., Chernovita, H. P., Mailoa, E., Philemon, S. A., & Chen, A. P. S. (2025). Integration of YOLOv11 and Histogram Equalization for Fire and Smoke-Based Detection of Forest and Land Fires. *Computers, Materials & Continua*, 84(3), 5361–5379. <https://doi.org/10.32604/cmc.2025.067381>
11. Duhayyim, M. A., Eltahir, M. M., Ali, O. A. O., Albraikan, A. A., Al-Wesabi, F. N., Hilal, A. M., Hamza, M. A., & Rizwanullah, M. (2023). Fusion-Based Deep Learning Model for Automated Forest Fire Detection. *Computers, Materials & Continua*, 77(1), 1355–1371. <https://doi.org/10.32604/cmc.2023.024198>
12. Edwards, R. B., Naylor, R. L., Higgins, M. M., & Falcon, W. P. (2020). Causes of Indonesian forest fires. *World Development*, 127, article ID 104717. <https://doi.org/10.1016/j.worlddev.2019.104717>
13. Gonçalves, A. M., Brandão, T., & Ferreira, J. C. (2024). Wildfire Detection With Deep Learning – A Case Study for the CICLOPE Project. *IEEE Access*, 12, 82095–82110. <https://doi.org/10.1109/ACCESS.2024.3406215>
14. Hopkins, B., O'Neill, L., Marinaccio, M., Rowell, E., Parsons, R., Flanary, S., Nazim, I., Seielstad, C., & Afghah, F. (2024). *FLAME 3 Dataset: Unleashing the Power of Radiometric Thermal UAV Imagery for Wildfire Management* (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2412.02831>
15. Hu, J., Niu, H., Carrasco, J., Lennox, B., & Arvin, F. (2022). Fault-tolerant cooperative navigation of networked UAV swarms for forest fire monitoring. *Aerospace Science and Technology*, 123, article ID 107494. <https://doi.org/10.1016/j.ast.2022.107494>
16. Hwan Shin, D., Mook Kang, J., & Cheong, T. (2025). FDN: A Real-Time Ensemble Fire Detection Network. *IEEE Access*, 13, 96723–96739. <https://doi.org/10.1109/ACCESS.2025.3575057>
17. Islam, A. M., Masud, F. B., Ahmed, Md. R., Jafar, A. I., Ullah, J. R., Islam, S., Shatabda, S., & Islam, A. K. M. M. (2023). An Attention-Guided Deep-Learning-Based Network with Bayesian Optimization for Forest Fire Classification and Localization. *Forests*, 14(10), article ID 2080. <https://doi.org/10.3390/f14102080>
18. Islam, A., & Habib, Md. I. (2023). *Fire Detection From Image and Video Using YOLOv5* (Ver. 1). arXiv. <https://doi.org/10.48550/ARXIV.2310.06351>
19. Keren, G., Sabato, S., & Schuller, B. (2017). *Fast Single-Class Classification and the Principle of Logit Separation* (Version 4). arXiv. <https://doi.org/10.48550/ARXIV.1705.10246>
20. Khanam, R., & Hussain, M. (2024). *What is YOLOv5: A deep look into the internal features of the popular object detector* (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2407.20892>
21. Kim, B., & Lee, J. (2019). A Video-Based Fire Detection Using Deep Learning Models. *Applied Sciences*, 9(14), article ID 2862. <https://doi.org/10.3390/app9142862>
22. Li, R., Hu, Y., Li, L., Guan, R., Yang, R., Zhan, J., Cai, W., Wang, Y., Xu, H., & Li, L. (2024). SMWE-GFPNet: A high-precision and robust method for forest fire smoke detection. *Knowledge-Based Systems*, 289, article ID 111528. <https://doi.org/10.1016/j.knsys.2024.111528>
23. Lin, M., Chen, Q., & Yan, S. (2013). *Network In Network* (Version 3). arXiv. <https://doi.org/10.48550/ARXIV.1312.4400>
24. Liu, S., Liang, Y., & Gitter, A. (2019). Loss-Balanced Task Weighting to Reduce Negative Transfer in Multi-Task Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01), 9977–9978. <https://doi.org/10.1609/aaai.v33i01.33019977>
25. Mahaveerakannan, R., Anitha, C., Aby K Thomas, Rajan, S., Muthukumar, T., & Govinda Rajulu, G. (2023). An IoT based forest fire detection system using integration of cat swarm with LSTM model. *Computer Communications*, 211, 37–45. <https://doi.org/10.1016/j.comcom.2023.08.020>
26. Mannor, S., Peleg, D., & Rubinstein, R. (2005). The cross entropy method for classification. *Proceedings of the 22nd International Conference on Machine Learning – ICML 05*, 561–568. <https://doi.org/10.1145/1102351.1102422>
27. Pesonen, J., Raita-Hakola, A.-M., Joutsalainen, J., Hakala, T., Akhtar, W., Koivumäki, N., Markelin, L., Suomalainen, J., Alves De Oliveira, R., Pölonen, I., & Honkavaara, E. (2025). Boreal Forest Fire: UAV-collected Wildfire Detection and Smoke Segmen-

- tation Dataset. *Scientific Data*, 12(1), article ID 1419. <https://doi.org/10.1038/s41597-025-05634-0>
28. Sathishkumar, V. E., Cho, J., Subramanian, M., & Naren, O. S. (2023). Forest fire and smoke detection using deep learning-based learning without forgetting. *Fire Ecology*, 19(1), 9. <https://doi.org/10.1186/s42408-022-00165-0>
 29. Shamsoshoara, A., Afghah, F., Razi, A., Zheng, L., Fulé, P. Z., & Blasch, E. (2021). Aerial imagery pile burn detection using deep learning: The FLAME dataset. *Computer Networks*, 193, article ID 108001. <https://doi.org/10.1016/j.comnet.2021.108001>
 30. Tao, Y., Li, B., Li, P., Qian, J., & Qi, L. (2025). Improved Lightweight YOLOv11 Algorithm for Real-Time Forest Fire Detection. *Electronics*, 14(8), article ID 1508. <https://doi.org/10.3390/electronics14081508>
 31. Wang, J., & Yan, C. (2026). CEVG-RTNet: A real-time architecture for robust forest fire smoke detection in complex environments. *Neural Networks*, 194, article ID 108187. <https://doi.org/10.1016/j.neunet.2025.108187>
 32. Wang, L., Guo, L., Li, H., He, B., Yang, J., & Huang, Y. (2025). Enhanced forest fire detection via dynamic multiscale fusion and contextual partial cross features. *Engineering Applications of Artificial Intelligence*, 162, article ID 112531. <https://doi.org/10.1016/j.engappai.2025.112531>
 33. Wang, M., Jiang, L., Yue, P., Yu, D., & Tuo, T. (2022). FASDD: An Open-access 100,000-level Flame and Smoke Detection Dataset for Deep Learning in Fire Detection. *Earth System Science Data Discussions* [preprint]. <https://doi.org/10.5194/essd-2022-394>
 34. Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., & Hu, Q. (2019). ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks (Version 4). arXiv. <https://doi.org/10.48550/ARXIV.1910.03151>
 35. Weng, S., Wang, H., Wang, J., Xu, C., & Zhang, E. (2025). YO-LO-SRMX: A Lightweight Model for Real-Time Object Detection on Unmanned Aerial Vehicles. *Remote Sensing*, 17(13), article ID 2313. <https://doi.org/10.3390/rs17132313>
 36. Wu, X., Irie, G., Hiramatsu, K., & Kashino, K. (2018). Weighted Generalized Mean Pooling for Deep Image Retrieval. *2018 25th IEEE International Conference on Image Processing (ICIP)*, 495–499. <https://doi.org/10.1109/ICIP.2018.8451317>
 37. Yar, H., Khan, Z. A., Ullah, F. U. M., Ullah, W., & Baik, S. W. (2023). A modified YOLOv5 architecture for efficient fire detection in smart cities. *Expert Systems with Applications*, 231, article ID 120465. <https://doi.org/10.1016/j.eswa.2023.120465>
 38. Yuan, J., Wang, H., Yang, T., Su, Y., Song, W., Li, S., & Gong, W. (2025). FF-net: A target detection method tailored for mid-to-late stages of forest fires in complex environments. *Case Studies in Thermal Engineering*, 65, article ID 105515. <https://doi.org/10.1016/j.csite.2024.105515>
 39. Zibtsev, S. V., Soshenskyi, O. M., Gumeniuk, V. V., & Koren, V. A. (2019). Long term dynamic of forest fires in Ukraine. *Ukrainian Journal of Forest and Wood Science*, 10(3), 27–40. <https://doi.org/10.31548/forest2019.03.027>
 40. Zibtsev, S. V., Soshenskyi, O. M., Myroniuk, V. V., & Gumeniuk, V. V. (2020). Wildfire in Ukraine: An overview of fires and fire management system. *Ukrainian Journal of Forest and Wood Science*, 11(2), 15–31. <https://doi.org/10.31548/forest2020.02.015>

B. B. Vdovin, V. S. Yakovyna, O. M. Tomashevskyy

Lviv Polytechnic National University, Lviv, Ukraine

DETECTION OF FOREST FIRES IN UAV STREAMING IMAGES USING DEEP LEARNING METHODS

The problem of real-time forest fire detection using unmanned aerial vehicles and deep learning methods was studied. Modern methods of fire detection and classification were analyzed, and their limitations and shortcomings were identified, in particular, their low efficiency during real-time operation and their dependence on terrain-specific factors. It was found that traditional methods of forest fire detection based on satellite images or sensor data do not provide sufficient accuracy in the early stages of ignition, which complicates timely response. The potential of artificial intelligence, particularly convolutional neural networks, for flame and smoke detection in images was studied. The model architecture was developed based on the YOLOv5 model, which uses the LBTW2 algorithm for joint learning of the detector and classifier. A dataset containing 4104 images, with 3289 labeled frames of the "fire" class and 2788 frames of the "smoke" class, was created from drone-captured video footage. The developed model was compared with existing approaches, confirming its advantages in classification and object detection accuracy. Compared with the base model YOLOv5s, trained without co-training, the proposed approach achieved better results, particularly in the mAP@50 metric, which increased by 11.6 % (0.917 versus 0.822). The model classifier demonstrated high efficiency, achieving Accuracy = 0.986, Recall = 0.977, Precision = 0.985 and F1-score = 0.981. The developed approach ensures stable model operation during real-time image processing and can be integrated into monitoring systems using unmanned aerial vehicles. The results of the study indicate the effectiveness of deep learning for early detection of forest fires and open up prospects for further scientific development, including increasing the energy efficiency of the models, improving their architecture, and creating new data sets. It is shown that simplifying the architecture of YOLO family models is promising for reducing computational complexity and increasing energy efficiency without significantly compromising accuracy indicators, which is especially important for using such models on unmanned aerial vehicles with limited resources.

Keywords: artificial intelligence; convolutional neural networks; YOLOv5 model; computer vision; object detection; object classification.