

О. С. ЛЮБИМОВ

аспірант кафедри програмних засобів і технологій
Херсонський національний технічний університет
ORCID: 0009-0004-7800-7322

В. Г. ШЕРСТЮК

доктор технічних наук, професор,
проректор з навчальної роботи
Херсонський національний технічний університет
ORCID: 0000-0002-9096-2582

ВДОСКОНАЛЕННЯ МОДЕЛІ YOLOV8 ДЛЯ ДИСТАНЦІЙНОГО ЗОНДУВАННЯ ШВИДКОПЛИННИХ ДЕСТРУКТИВНИХ ПРОЦЕСІВ

Своєчасне та точне виявлення осередків лісових пожеж як найпоширенішого класу швидкоплинних деструктивних процесів має вирішальне значення для припинення горіння та мінімізації наслідків. Для цього можуть бути застосовані технології дистанційного зондування з безпілотних літальних апаратів, машинного навчання та комп'ютерного зору. Однак, вплив зовнішнього середовища і низки факторів невизначеності, спотворень та динаміки руху створюють проблеми ідентифікації ознак пожежі, а обчислювальна складність перешкоджає роботі алгоритмів розпізнавання в режимі реального часу. Для вирішення цих проблем в статті запропоновано «полегшену» з погляду на обчислювальну складність модель розпізнавання осередків лісових пожеж при дистанційному зондуванні YOLOv8N, що вдосконалює базову модель YOLOv8n за рахунок застосування на магістральному рівні модуля GhostNetv2 разом з модулем уваги DFC замість традиційної операції згортки, що дозволяє суттєво зменшити кількість параметрів моделі, зберігаючи її продуктивність, та механізму звернення уваги MHSA в операціях C2f, що покращує здатність отримувати ознаки осередків горіння та підвищує точність виявлення невеликих осередків горіння, також на проміжному рівні моделі використано механізм самоуваги SegNeXt в операціях C2f, що дозволило підвищити точність виявлення ознак пожежі у складних умовах. Модель YOLOv8N збільшує показники влучності, повноти, середнього гармонійного і точності на 4,3 %, 7,5 %, 4,8 % і 5,9 % відповідно у порівнянні з базовою моделлю YOLOv8n, кількість параметрів зменшується на 33,3 %. Отже, запропонована модель забезпечує високий рівень точності виявлення ознак лісової пожежі, зберігаючи при цьому баланс між обчислювальною складністю і ефективністю моделі, що гарантує її спроможність працювати в системах дистанційного зондування з безпілотних літальних апаратів в режимі реального часу.

Ключові слова: дистанційне зондування, безпілотний літальний апарат, швидкоплинні деструктивні процеси, лісова пожежа, осередок горіння, механізм уваги, ідентифікація ознак, розпізнавання об'єктів, зображення.

O. S. LIUBIMOV

Postgraduate Student at the Department of Software Tools and Technologies
Kherson National Technical University
ORCID: 0009-0004-7800-7322

V. G. SHERSTJUK

Doctor of Technical Science, Professor,
Vice-Rector for Academic Affairs
Kherson National Technical University
ORCID: 0000-0002-9096-2582

IMPROVEMENT OF THE YOLOV8 MODEL FOR REMOTE SENSING OF RAPID DESTRUCTIVE PROCESSES

Timely and accurate detection of forest fires as the most common class of rapid destructive processes is crucial for stopping them and minimizing their consequences. Remote sensing technologies from unmanned aerial vehicles, machine learning, and computer vision can be used for this purpose. However, the influence of the external environment and several uncertainties, distortions, and motion dynamics create problems in identifying fire signs, as well as computational complexity hinder the operation of recognition algorithms in real time. To solve these problems, the paper proposes a “lightweight” model for recognizing forest fire sources during remote sensing, YOLOv8N, which improves the basic YOLOv8n model by using the GhostNetv2 module at the backbone level together with the DFC attention module instead of the traditional convolution operation, which allows significantly reducing the number of model parameters while maintaining its performance, and the MHSA attention mechanism in C2f operations, which improves the ability to obtain

signs of burning sources and increases the accuracy of detecting small burning areas. Also, at the intermediate level of the model, the SegNeXt self-attention mechanism is used in C2f operations, which allows to increase the accuracy of detecting fire signs in difficult conditions. The YOLOv8N model increases the accuracy, completeness, harmonic mean, and precision by 4.3 %, 7.5 %, 4.8 % and 5.9 % respectively compared to the base model YOLOv8n, the number of parameters is also reduced by 33.3 %. Therefore, the proposed model provides a high level of accuracy in detecting forest fire features, while maintaining a balance between computational complexity and model efficiency, which guarantees its ability to work in remote sensing systems from unmanned aerial vehicles in real time.

Key words: remote sensing, unmanned aerial vehicles, rapid destructive processes, forest fire, combustion area, attention mechanism, feature identification, object recognition, images.

Постановка проблеми

Внаслідок зміни клімату та зовнішніх впливів на просторово-розподілені територіальні системи утворюються швидкоплинні деструктивні процеси (ШДП), при цьому тригерні та каскадні явища часто призводять до природних катастроф. На сьогоднішній день не існує способів побудови математичних моделей ШДП для їх прогнозування та попередження, а їх слабка спостережуваність, значні масштаби, інтенсивність та швидкість розповсюдження суттєво ускладнюють умови вирішення задач ідентифікації, локалізації, реагування та ліквідації наслідків.

Найпоширенішим класом ШДП є лісові пожежі (ЛП) – мабуть, один із найбільш руйнівних типів природних катастроф, що мають високу частоту виникнення та суттєво впливають на життя людей, соціально-економічний стан та природні екосистеми у глобальному масштабі [1]. Зростання чисельності населення, урбанізація та певні антропогенні фактори, зокрема незаконна вирубка лісів і зміни ґрунтового покриву, значно підвищують ризик виникнення та інтенсивність ЛП [2].

Лісові пожежі, особливо верхові, дуже швидко поширюються в умовах високого вмісту кисню і при наявності сильних повітряних потоків [3]. Отже, проблема своєчасного виявлення осередків загоряння, локалізації, реагування та ліквідації наслідків ЛП є надзвичайно актуальною.

Ще десятиліття тому використовувались такі традиційні методи ідентифікації ЛП та вирішення задач локалізації, протидії, ліквідації наслідків, як візуальне спостереження. Хоча цей метод забезпечує можливість безпосереднього моніторингу ситуації, але спостереження є неповними і неточними, а сам метод – малоефективним та занадто небезпечним [4]. На сьогоднішній день з'явилися достатньо ефективні методи дистанційного зондування, що дають змогу ефективно й швидко виявляти осередки загоряння та спостерігати процес розповсюдження ЛП. Можливість дистанційного зондування «зверху» та «збоку» дозволяє значно ширше охоплювати просторово-розподілені події, досягаючи принципово іншого рівня ситуаційної обізнаності осіб, що приймають рішення щодо локалізації, реагування та ліквідації наслідків ЛП, на відміну від візуального спостереження [5].

Супутникові системи, такі як Himawari-9 [6] і MODIS [7], які забезпечують безперервний моніторинг ЛП на значних територіях, однак мають обмежену просторову роздільну здатність, що ускладнює спостереження, ідентифікацію та моніторинг осередків горіння, особливо малого масштабу. Нові можливості дистанційного зондування ЛП відкрилися з розвитком безпілотних літальних апаратів (БПЛА), технологій обробки зображень, машинного навчання, що дає можливість спостерігати за розвитком ЛП «зверху», та, з одного боку, дозволяє максимально наблизитися до спостережуваних подій, а з іншого боку – утримуватись на безпечній відстані від них [8], що, в свою чергу, надає можливість успішного спостереження в умовах, небезпечних для здоров'я і навіть життя спостерігача. Крім того, використання БПЛА для дистанційного зондування ЛП забезпечує отримання високоякісних знімків із більшою просторовою деталізацією, а також дає змогу ефективно функціонувати в складних ландшафтах [9].

Однак, навіть із застосуванням найсучасніших технологій дистанційного зондування з БПЛА залишається необхідною участь оператора для підтвердження результатів ідентифікації, що зумовлює потребу у використанні методів обробки зображень та комп'ютерного зору, інтегрованих з системами дистанційного зондування на БПЛА, що може забезпечити більш точне та ефективне виявлення ЛП у реальному часі [10], а отже, є перспективним напрямком досліджень.

Аналіз останніх досліджень і публікацій

На початкових етапах впровадження методів обробки зображень та машинного навчання в першу чергу визначали набори ознак, пов'язаних із ЛП, такі як полум'я, відблиски, задимлення. Ці ознаки оброблялись відповідними алгоритмами для ідентифікації осередків горіння. Однак, таких ознак виявилось дуже багато, і майже всі вони є динамічними, що потребувало вирішення задачі ідентифікації за наявності неповної та неточної інформації про наявність ознак. Так, в [11] запропоновано модель опорних векторів (PreVM), яка ефективно розв'язує проблему ідентифікації нетипових для ЛП зразків даних. В [12] автори використали додаткові (екологічні) параметри та зенітний кут сонця для підвищення точності класифікації ЛП за допомогою випадкових лісів. Однак, такі підходи не гарантують визначення оптимального набору ознак, оскільки залежать від вибору ознак вручну, що вимагає і значних експертних знань, і достатньо значного часу. Отже, на наступних етапах впровадження

методів обробки зображень та машинного навчання почали застосовувати алгоритми глибокого навчання, які мають суттєву перевагу, дозволяючи автоматично вилучати складні подання ознак із необроблених даних [13].

Методи глибокого навчання [14, 15] використовують штучні нейронні мережі (НМ), що дозволяє ефективно аналізувати дані, розпізнавати закономірності та приймати рішення [16]. У задачах ідентифікації і моніторингу ЛП використовують двоетапні [17] та одноетапні [18] алгоритми глибокого навчання. Двоетапні алгоритми спочатку визначають потенційні ділянки горіння, а потім уточнюють їхню класифікацію та коригують регресію просторових обмежувальних рамок [19]. Їх недоліком є значні обчислювальні витрати як на навчання, так і на отримання висновків, що унеможлиблює їхнє застосування в режимі реального часу.

Одноетапні алгоритми, зокрема YOLO [20], забезпечують швидке виявлення завдяки використанню анкерів для прогнозування положення та класифікації об'єктів. Механізми уваги [21] дозволяють НМ вибірково зосереджуватися на певній інформації при обробці величезної кількості вхідних даних і дають змогу ефективно обробляти великі масиви даних. Як наслідок, вони широко використовуються в поєднанні з НМ типу YOLO. Наприклад, в [22] автори інтегрували Swin Transformer і механізм уваги в мережу YOLOX, що підвищило точність виявлення до 92,26 %. Проте запропонований підхід все ж залишається обчислювально затратним, має значну кількість параметрів і вимагає ще більше обчислювальних ресурсів, що не сприяє ідентифікації ЛП у реальному часі. В [23] автори розглянули проблему низької точності розпізнавання та запропонували модель FCDM на основі YOLOv5. Ця модель покращила здатність виявлення різних типів ЛП шляхом модифікації функції втрат обмежувальної просторової рамки та запровадженням механізму згорткового блоку уваги. В роботі [24] автори представили модель LMDFS на основі YOLOv7, що вдосконалює можливості визначення функції диму, використавши механізм координації уваги та механізм підвищення дискретизації функції повторного складання з урахуванням вмісту. Точність було покращено на 5,9 % порівняно з базовою моделлю YOLOv7, однак, ця модель також має обмеження, коли стикається з нетиповими (нерегулярними) зображеннями ЛП або невеликими за розміром ЛП.

З проведеного аналізу публікацій можна зробити висновок, що незважаючи на інтенсивні дослідження і значний прогрес у розробці алгоритмів машинного навчання та комп'ютерного зору, використання БПЛА для дистанційного зондування ЛП продовжує стикатися з проблемами. По-перше, динаміка руху БПЛА, його вібрації посилюють заплутаність фону на зображеннях. По-друге, ця заплутаність суттєво ускладнюється такими факторами, як непередбачувані метеорологічні обставини, коливання освітленості, наявність хмар, туману та диму, що додає впливу факторів невизначеності. По-третє, існує певне протиріччя між складністю структури НМ, яка може допомогти подолати неточність і невизначеність спостереження, та обчислювальною складністю НМ, що перешкоджає роботі алгоритму в режимі реального часу. Подолання цих проблем потребує додаткових досліджень.

Формулювання мети дослідження

Зважаючи на необхідність балансування між ефективністю та обчислювальною складністю, щоб подолати вищезазначені проблеми, необхідно розробити полегшений алгоритм дистанційного зондування БПЛА для моніторингу ЛП, що дозволяє досягти певного компромісу між ефективністю, що забезпечується завдяки мінімізації кількості параметрів, і збереженням високої точності ідентифікації осередків горіння при забезпеченні можливості роботи алгоритму в режимі реального часу.

Цю статтю спрямовано на розробку одноетапного алгоритму глибокого навчання, заснованого на YOLOv8, для вирішення задач дистанційного зондування ЛП, який задовільнив би вимогам точності ідентифікації при збереженні продуктивності, достатньої для здійснення моніторингу в реальному часі.

Щоб подолати проблему обчислювальної складності, оптимізовано магістральну архітектуру НМ, що суттєво зменшує складність мережі та обчислювальні вимоги. Щоб подолати проблему великої кількості вхідних даних, застосовано механізм зосередження уваги для підвищення точності виявлення НМ дрібних об'єктів спостереження. Нарешті, щоб подолати вплив багатьох факторів невизначеності, використано механізм семантичної сегментації для покращення розпізнавання осередків горіння у складних умовах.

Метою дослідження є підвищення точності ідентифікації осередків горіння у складних умовах спостереження під час дистанційного зондування з БПЛА за рахунок використання полегшеного алгоритму дистанційного зондування, що дозволяє досягти компромісу між ефективністю та точністю, дозволяючи здійснення моніторингу ШДП в реальному часі.

В розділі 1 представлено принципи роботи та архітектуру моделі YOLOv8, розділ 2 присвячено вдосконаленню операцій згортки на магістральному рівні моделі, в розділі 3 розглядаються питання вдосконалення операцій C2f на магістральному рівні моделі, розділ 4 вдосконалює операції C2f на проміжному рівні, в розділі 5 розглядаються результати експерименту.

Викладення основного матеріалу дослідження

1. Принципи роботи та архітектура YOLOv8

В основу роботи покладено найновішу модель НМ для виявлення об'єктів, класифікації зображень і сегментації зразків YOLOv8 [25], яка використовує глибоке навчання для ідентифікації об'єктів на зображеннях, забезпечуючи баланс між швидкістю обробки та точністю прогнозування.

Популярність моделі YOLO пояснюється достатньо високою точністю при збереженні порівняно невеликого її розміру. Моделі YOLO можна навчати на одному GPU і розгортати з низькими витратами на периферійному обладнанні або в хмарі.

Залежно від глибини та ширини НМ її можна представити п'ятьма різними фреймворками: YOLOv8n, YOLOv8l, YOLOv8s, YOLOv8x і YOLOv8m. Враховуючи вимоги до дистанційного зондування ЛП у реальному часі, нами обрано полегшену модель YOLOv8n (нано) та оптимізовано саме її.

Архітектура YOLOv8n в основному поділяється на чотири рівня: вхідний, магістральний, проміжний та вихідний, як показано на рис. 1.

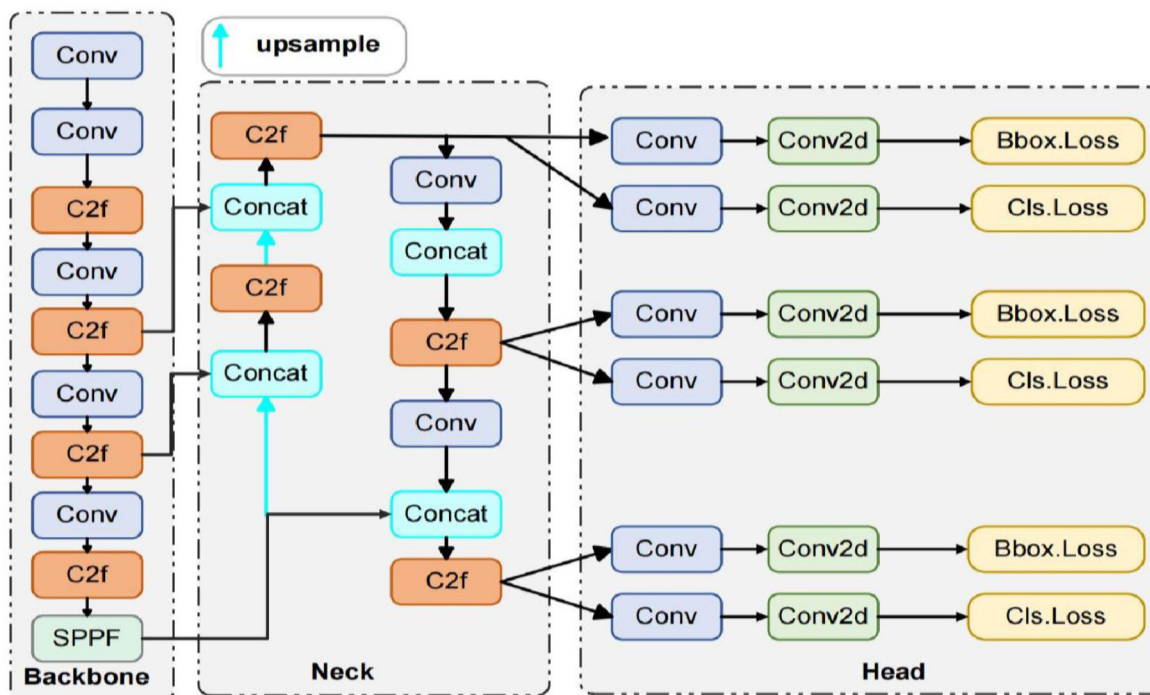


Рис. 1. Структура YOLOv8: Conv – операція згортки; C2f – операція CSPDarknet53 до 2-ступеневої пірамідної мережі; Conv2d – операція 2-вимірної згортки; SPPF – операція об'єднання просторових пірамід

Принцип роботи YOLOv8 базується на одноетапному підході до виявлення об'єктів. На відміну від двоетапних методів типу R-CNN, які спочатку генерують регіони-кандидати, а потім класифікують їх, YOLO обробляє зображення як єдине ціле.

Ключові принципи роботи моделі:

1. Розбиття зображення на рівномірну сітку, де кожна її комірка відповідає за прогнозування об'єктів, центр яких знаходиться в межах цієї області.
2. Єдине одноетапне проходження через мережу для одночасного визначення класу, координат обмежувальної рамки та ймовірності присутності об'єкта.
3. Застосування анкерних боксів для ефективного розпізнавання об'єктів різних розмірів і форм.
4. Застосування сучасних механізмів уваги, які покращують точність визначення об'єктів за рахунок фокусування на релевантних частинах зображення.

Модульна архітектура YOLOv8 складається з кількох основних компонентів:

- Backbone (основа мережі, магістральний рівень), що використовує вдосконалену архітектуру CSPDarknet, яка поєднує ефективність традиційних згорткових нейронних мереж (CNN) з механізмом залишкових зв'язків (ResNet), що зменшує кількість параметрів та обчислювальні витрати, забезпечуючи високу продуктивність.
- Neck (проміжний рівень), що використовує компоненти PAN (Path Aggregation Network) та FPN (Feature Pyramid Network), які покращують збереження просторової інформації та злиття ознак з різних рівнів абстракції, що підвищує якість розпізнавання дрібних об'єктів.
- Head (вихідний рівень), що відповідає за генерацію прогнозів, включаючи координати обмежувальних рамок, класифікацію об'єктів та впевненість.

Порівняно з попередніми версіями, в YOLOv8 застосовано низку вдосконалень а саме [25]:

- використано нову стратегію прогнозування для кращої адаптації до складних фонових сцен;

- використано конволюційну НМ GhostNet для оптимізації обчислень і зменшення числа параметрів;
- запроваджено додаткові механізми нормалізації для стабілізації навчання;
- покращено функції втрат для найкращого узгодження обмежувальних рамок із реальними об'єктами;
- оптимізовано структуру моделі, що дозволяє використовувати НМ на вбудованих пристроях та мобільних процесорах, що є дуже важливим для досягнення поставленої нами мети.

Вхідний рівень виконує задачу обробки зображень дистанційного зондування ЛП за допомогою операцій доповнення даних. Ці операції, серед іншого, забезпечують налаштування відтинку, масштабування та мозаїчне доповнення даних, що полягає у випадковому виборі регіонів із чотирьох різних зображень, а потім об'єднанні цих регіонів у нове зображення після випадкового обрізання та масштабування.

Магістральний рівень відповідає за вилучення ключових характеристик із зображення та складається із модулів згортки Conv, НМ CSPDarknet53, конволюційних 2-ступеневих блоків C2f, що формують піраміди, і модулів об'єднання просторових пірамід SPPF.

CSPDarknet53 – це згорткова НМ та магістраль для виявлення об'єктів, яка використовує модель DarkNet-53 (глибиною 53 шари) та працює за стратегією поділу та злиття CSPNet, спочатку розділяючи карту функцій базового рівня на дві частини, а потім об'єднуючи їх за допомогою проміжної ієрархії, що забезпечує більш градієнтний потік через мережу [26].

Модуль Conv обробляє дані за допомогою операцій згортки, пакетної нормалізації BN і функцій активації сигмоїдної лінійної одиниці SiLU.

Модуль C2f покращує розповсюдження градієнта та доповнює потік інформації в мережі вилучення ознак завдяки інтеграції міжрівневих з'єднань.

Модуль SPPF використовує три послідовні операції об'єднання для зниження обчислювальних вимог, зберігаючи при цьому багатомасштабну інформацію і таким чином розширюючи поле сприйняття.

Основна функція проміжного рівня полягає в досягненні міжвимірної інтеграції функціональних ознак, що може ефективно об'єднувати відображення функцій різних рівнів, забезпечуючи точне збереження просторової інформації завдяки використанню мережі піраміди ознак (FPN) і мережі агрегації шляхів (PAN). Цей рівень дозволяє моделі більше зосереджуватися на інформації про цільову функцію, значно покращуючи ефективність виявлення [27].

Вихідний рівень виконує задачу формування остаточних результатів виявлення об'єктів, використовуючи детальні відображення функцій, отримані з проміжного рівня, для визначення положення обмежувальної просторової рамки, ймовірності категорій та іншу необхідну інформацію для кожної з функцій. Вихідний рівень використовує механізм NMS [28] для видалення можливих дублікатів, зберігаючи при цьому точні результати прогнозування.

2. Вдосконалення операцій згортки на магістральному рівні

Враховуючи важливість якнайшвидшого виявлення ЛП для запобігання її швидкому поширенню, ефективність ідентифікації є критичним фактором, який слід неодмінно враховувати при оцінці алгоритму ідентифікації її ознак.

Магістральний рівень мережі моделі YOLOv8n використовує низку послідовних операцій згортки (рис. 1), збільшуючи тим самим кількість каналів і розширюючи поле сприйняття, однак, внаслідок цього збільшується кількість параметрів моделі та її обчислювальна складність, що зовсім не сприяє вирішенню задачі у реальному часі. Підвищити ефективність розпізнавання об'єктів та ідентифікації ознак горіння можна шляхом впровадження більш ефективних згорткових мереж, вдосконалюючи безпосередньо операції згортки.

Пропонується оптимізувати магістральний рівень мережі YOLOv8n, замінюючи стандартні операції згортки Conv модулями GhostNetV2 [29] і «полегшуючи» тим самим процес екстракції функцій. Звичайно, це може спровокувати певне зниження точності розпізнавання, але в той же час забезпечить суттєве зниження кількості параметрів моделі та вимог до обчислень.

Як відомо, GhostNetV2 поєднує в собі механізм уваги DFC із модулями Ghost (рис. 2), що дозволяє максимально підтримувати продуктивність, зберігаючи легку структуру.

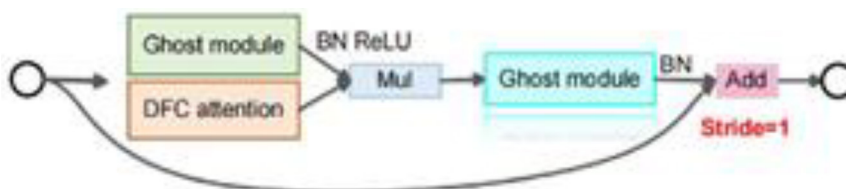


Рис. 2. Структура GhostNetV2

Увага (attention) в технологіях машинного навчання є механізмом, за допомогою якого обчислюються «м'які» ваги для певної функції, точніше, для її вкладення у контекстному вікні. У GhostNetV2 ці ваги обчислюються послідовно, як прийнято у рекурентних НМ. Однак, «м'які» ваги можливо обчислювати паралельно, що і пропонується

в роботі, завдяки чому ці ваги можуть змінюватися в процесі кожного наступного циклу виконання, замість того, щоб бути статичними після тренування НМ.

Застосування механізму DFC дозволяє подолати недолік прихованих виходів рекурентних НМ, які завжди віддають перевагу найсвіжішим кадрам зображення, пригнічуючи більш ранні кадри. Обчислення уваги дозволяє прихованому рівню мати рівний доступ до будь-якої частини зображення безпосередньо, а не через попередній прихований стан (рис. 3).

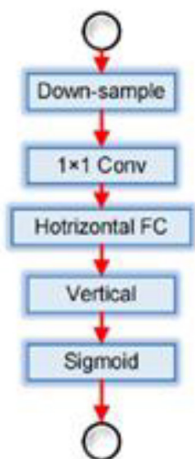


Рис. 3. Структура уваги DFC

Механізм уваги DFC використовує динамічні фільтри для зважування функцій, виділення з них найбільш важливих і придушення неважливих, а отже, цей механізм динамічно регулює фокус моделі на різних функціях, тим самим підвищуючи її точність.

Зі спрощеної точки зору, механізм уваги DFC розкладає карту уваги на два повністю пов'язані шари (рис. 4) та збирає функції вздовж горизонтального та вертикального напрямків окремо.

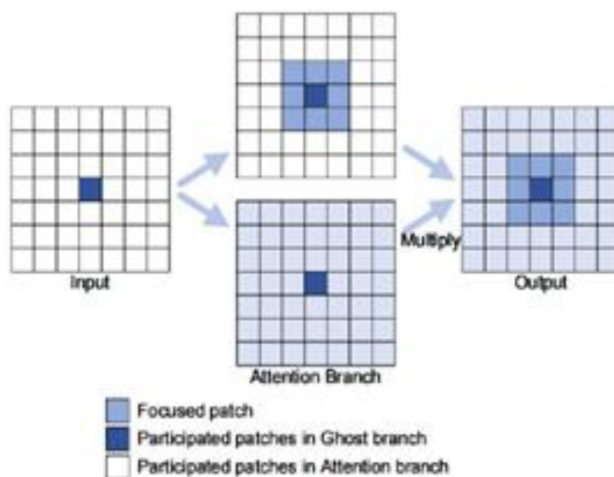


Рис. 4. Механізм уваги DFC

Механізм DFC працює наступним чином.

Нехай a'_{hw} і a_{hw} – карти уваги відповідно вертикального та горизонтального напрямку; $\odot$ – операція поелементного множення, W – ширина карти ознак; H – висота карти ознак; F – вагові коефіцієнти у шарі повного зв'язку (FC), що підлягають навчанню, F^W і F^H – трансформаційні ваги вертикального та горизонтального напрямку відповідно.

Тоді

$$a'_{hw} = \sum_{h'=1}^H F_{h,h'w}^H \odot z_{h'w}, \quad h=1,2,\dots,H, \quad w=1,2,\dots,W$$

$$a_{hw} = \sum_{w'=1}^W F_{w,hw'}^W \odot a'_{hw}, \quad h=1,2,\dots,H, \quad w=1,2,\dots,W.$$

За рахунок роз'єднання горизонтальних і вертикальних перетворень карти ознак обчислювальна складність модуля DFC знижується до $O(H^2W + HW^2)$, що суттєво покращує обчислювальну ефективність моделі.

Як показано на рис. 2, в HM GhostNetV2 модуль Ghost працює паралельно механізму уваги DFC, їх результати поелементно множаться, що і дозволяє реалізувати вдосконалення та розширення вихідних функцій, фіксуючи залежності між пікселями в різних просторових локаціях на дальній відстані. Так, після паралельного виконання модуля Ghost і механізму DFC вихідні функції зменшуються, оскільки є потреба зменшити просторовий розмір відображення, що відповідно зменшує обсяг обчислень та використаної пам'яті. З іншого боку, це збільшує поле сприйняття, надаючи можливість захоплення ширшого діапазону контекстних даних та мінімізуючи втрати градієнта. Роботу модуля GhostNetV2 можна додатково вдосконалити, підвищуючи обчислювальну ефективність, якщо ввести додаткову згортку по глибині (рис. 5) після операції поелементного множення, тоді розмір вихідної функції може бути відновлено модулем Ghost, що дозволить забезпечити узгодженість із вхідними даними. Назвемо таку структуру GhostNetV21.

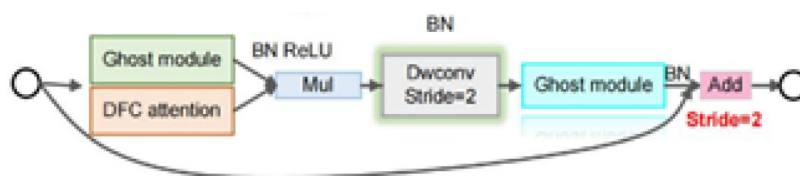


Рис. 5. Вдосконалена структура GhostNetV21

3. Вдосконалення операцій C2f на магістральному рівні

У контексті моніторингу ЛП наявність невеликих за розміром осередків займання є їх первинною ознакою, тож важливо їх швидко ідентифікувати. Однак, чим більше ми «полегшуємо» модель YOLOv8n обчислювально, спрощуючи її структуру, тим більше обмежуємо точність в задачі виявлення ознак невеликих ЛП. Зберегти достатню точність розпізнавання, не ускладнюючи модель, можливо, якщо додатково до замієнених на GhostNetV21 операцій згортки вдосконалити ще й модулі C2f.

C2f означає «Coordinates-To-Features», ці модулі призначено для ефективного перетворення інформації про координати в подання ознак [30].

C2f працює наступним чином (рис. 6):

- вхідна інформація проходить перший перехідний шар;
- спрацьовує функція розбиття уздовж розмірності каналів, завдяки чому визначаються значення функцій a і b , які подаються списком y ;
- останній елемент y в циклі передається n разів до послідовності вузькостей, де кожен результат додається до списку y , який після всіх кроків має $n + 1$ значень функцій для об'єднання перед переходом до другого перехідного рівня;
- y об'єднується та проходить через другий перехідний рівень.

Яким чином можна вдосконалити модуль C2f? По перше, ми можемо замінити всі операції конволюції, які містяться в цьому модулі, на GhostNetV21, як це було запропоновано в попередньому розділі, полегшуючи структуру модуля і підвищуючи його обчислювальну ефективність.

По друге, у структуру модулів C2f магістральної мережі доцільно додати свої механізми уваги, що дозволить вивчати ширший спектр вхідних функцій в різних репрезентативних підпросторах, повністю враховуючи контекст зображення, що може підвищити точність ідентифікації ознак для порівняно невеликих просторових об'єктів.

Мережі уваги розроблялися для встановлювання найвищих кореляцій між ознаками, виходячи з припущення, що вони навчаються цим закономірностям під час тренування. Така кореляція фіксується в нейронних вагах через зворотне поширення або з самокерованого попереднього тренування, або з керованого тонкого настроювання. Однак, традиційні механізми уваги безпосередньо обчислюють залежності між елементами всередині послідовності, ефективно фіксуючи глобальну інформацію про зображення, але вивчаючи ознаки лише з одного репрезентативного підпростору, що обмежує в цілому здатність моделі фіксувати різні розмірні характеристики вхідних даних [31].

В цій статті пропонується додати до модуля C2f механізм багатоголової самоуваги MHSA [32]. Механізм MHSA подібний до перегляду простору ознак кілька разів, щоразу зосереджуючись на іншому аспекті, як-от ознаки диму, або полум'я, або відблисків. Поєднуючи ці зосереджені перегляди, можливо отримати глибше розуміння ознак на зображенні. Цей механізм використовує кілька головок уваги для захоплення різних аспектів взаємозв'язків у даних, однак, враховуючи вимоги до ефективності обчислень, кількість головок уваги доцільно мінімізувати.

Механізм MHSA ефективно фіксує інформацію про ознаки з різних підпросторів, використовуючи паралельно декілька незалежних головок самоконтролю та передаючи їм одночасно для обробки різні ознаки вхідного зображення.

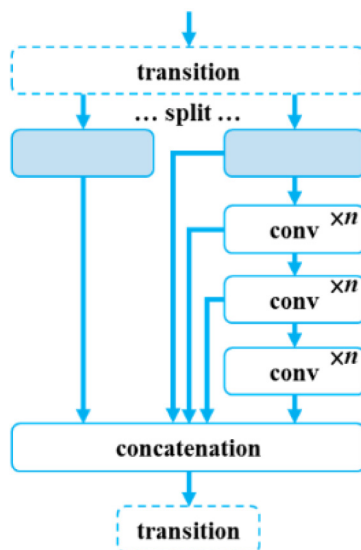


Рис. 6. Структура модуля C2f

Відомо, що повільну НМ можна навчити за допомогою градієнтного спуску програмувати швидкі ваги іншої НМ через тензорні добутки самопороджуваних шаблонів збудження, які називають «ключем» (key) та «значенням» (value) (за аналогією до «FROM» і «TO»), а відображення уваги швидкими м'якими вагами називають запитами (queries).

Отже, нехай у кожній голові модель обчислює перетворення запиту Q , ключа K і значення V вхідних даних X (рис. 7), а кожна голова розраховує ваги уваги на основі подібності між запитами та ключами, що реалізується за допомогою операції скалярного добутку, за якою слідує функція softmax.

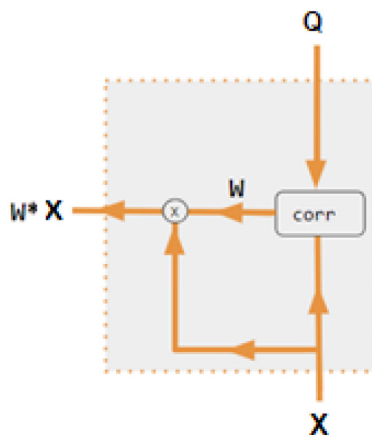


Рис. 7. Обчислення ваги уваги

Нехай, W_q , W_k і W_v відповідають навченим ваговим матрицям Q , K , V відповідно, d_k – розмірність ключа, або його масштабний коефіцієнт. Тоді стандартна масштабована скалярно-добуткова увага для кожної голови визначається як:

$$Q = W_q X; \quad K = W_k X; \quad V = W_v X$$

$$head_k = Attention_k(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) \cdot V.$$

Функція softmax використовується для перетворення показників уваги в розподіл ймовірностей, що гарантує, що ваги є позитивними та підсумовуються до 1. Отримані ваги уваги вказують на важливість кожного елемента для інших елементів у даних та використовуються для зважування відповідних значень.

Результати, вироблені всіма головами, об'єднуються та проходять через лінійне перетворення для формування остаточного результату:

$$MultiHead(Q, K, W) = \text{Concat}(head_1, \dots, head_n) \cdot W^T.$$

Вдосконалену структуру модуля C2f, яка використовує запропонований в роботі механізм самоуваги у вигляді шару НМ, представлено на рис. 8, де кожену голову механізму MHSA реалізовано окремим шаром, а останній шар традиційного модуля C2f замінено на шар самоуваги. Останній модуль (CBS) виконує згорткове перетворення GhostNetV21+BatchNorm2d+SiLU замість стандартного для YOLOv8 Conv2d+BatchNorm2d+SiLU [33].

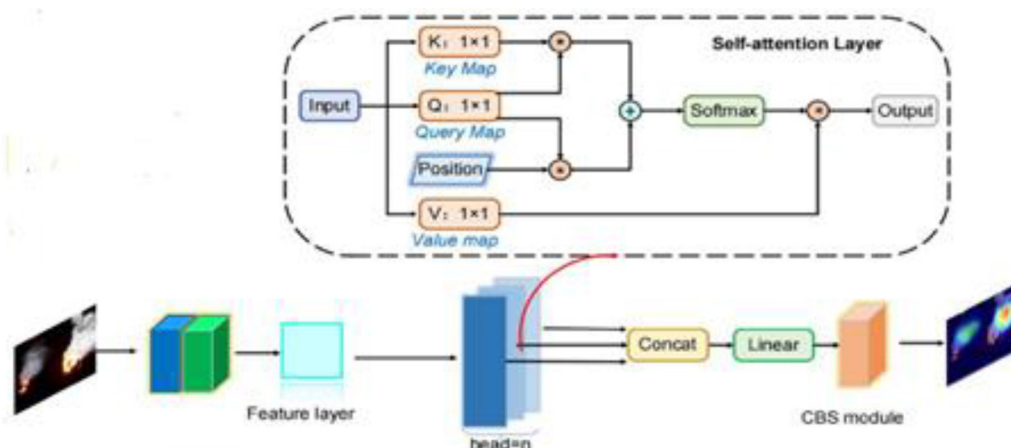


Рис. 8. Структура вдосконаленого модуля c2f з механізмом MHSA

4. Вдосконалення операцій C2f на проміжному рівні.

Під час дистанційного зондування умови освітлення можуть сильно відрізнятися внаслідок складності природного середовища, наприклад, десь може бути пряме сонячне світло, а десь тінь. Крім того, різноманітність ландшафту в лісі, в тому числі видів дерев, кущів і рельєфу місцевості, ускладнює розрізнення осередків ЛП від їх природного фону, що, безумовно, висуває більш суворі вимоги до точності моделі.

Хоча модель YOLOv8n використовує у проміжному шарі структури C2f для злиття семантичної інформації з різних шарів і масштабів, що до певної міри покращує ефективність виявлення за допомогою вилучення ознак і міжрівневих зв'язків та підвищує продуктивність розпізнавання ознак, таке рішення суттєво збільшує обчислювальну складність моделі та негативно впливає на кількість параметрів.

Для того, щоб ефективно вирізняти ознаки диму і полум'я на складному фоні та пригнічувати фоновий шум, пропонується вдосконалити модулі C2f проміжного рівня моделі YOLOv8n, доповнивши їх механізмом уваги до особливостей ознак полум'я. Враховуючи поставлену нами задачу, доцільно використати механізми самоуваги, що використовуються для семантичної сегментації при обробці просторової інформації в згорткових НМ.

Безумовно, мережа згорткової самоуваги, що використовує обчислювально легкі згорткові операції, має бути більш ефективною при обробці просторової контекстної інформації, ніж традиційні трансформаційні механізми [34]. Одним з найбільш ефективних механізмів даного класу є механізм SegNeXt Attention [35].

Механізм привернення уваги SegNeXt покращує здатність моделі обробляти просторові зв'язки та складний фон у зображеннях за рахунок використання декількох масштабів одночасно, а застосований кодер-декодер використовує прогресивний ієрархічний підхід для обробки зображень, як показано на рис. 9.

Під час фази кодування вхідне зображення піддається поступовому зниженню дискретизації, що досягається за допомогою згорткових і об'єднуючих шарів, які зменшують просторові розміри зображення, водночас роблячи акцент на інформації про ознаки, що дозволяє отримати більш глибокий результат семантичної сегментації.

В структурі кодера застосовано алгоритм мульти-масштабної згортки уваги MSCA (Multi-Scale Contextual Attention), який складається з поглибленої згортки, що агрегує локальну інформацію, розгалуженої поглибленої згортки, яка фіксує мульти-масштабний контекст, та згортки 1×1 , що моделює зв'язки між різними каналами.

Застосування механізму MSCA покращує сприйняття моделлю цільових об'єктів шляхом зважування карт об'єктів для виділення важливих просторових областей і придушення нерелевантного фону, що забезпечує більш точну локалізацію та розпізнавання осередків горіння (займання).

Нехай *Attention* і *Out* відповідно карта уваги та вихід, *F* представляє вхідну функцію, $\otimes$ – операція поелементного множення матриці, *DW-Conv* означає згортку по глибині, *Scale_i* – зв'язок ідентичності, $i \in \{0, 1, 2, 3\}$ позначає *i*-ту гілку.

Тоді

$$Attention = Conv_{1 \times 1} \left(\sum_{i=0}^3 Scale_i (DW - Conv(F)) \right)$$

$$Out = Attention \otimes F$$

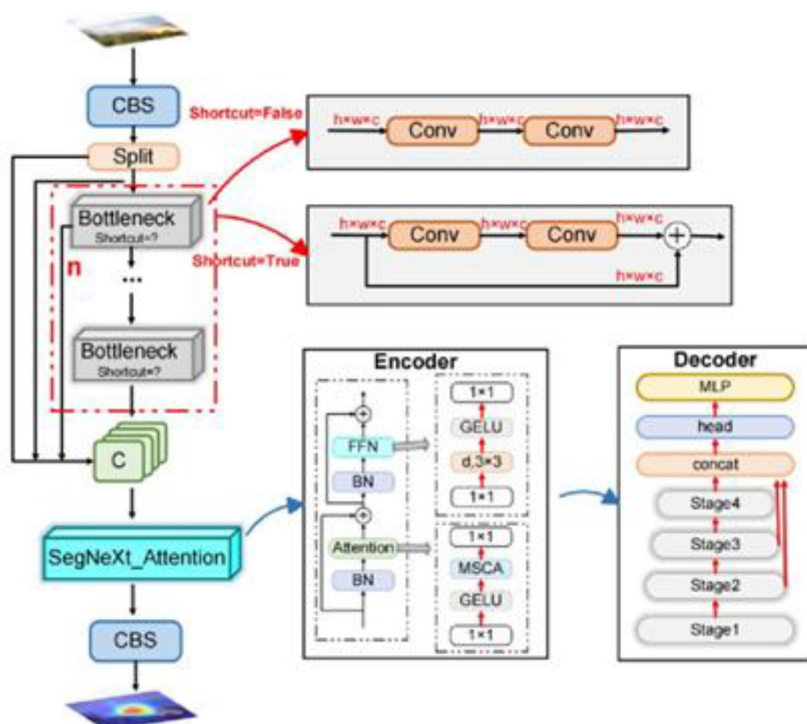


Рис. 9. Структура вдосконаленого модуля c2f з механізмом SegNeXt

На етапі декодування механізм SegNeXt поступово відновлює просторову роздільну здатність зображення за допомогою серії операцій підвищення дискретизації та згорткових шарів, використовуючи глибинні семантичні характеристики, отримані на етапі кодування. На цьому етапі використовуються методи об'єднання ознак, що об'єднують глибоку семантичну інформацію з фази кодування з функціями високої роздільної здатності з фази декодування, підвищуючи деталізацію реконструйованих карт функцій.

Під час цього процесу унікальні міжшарові з'єднання SegNeXt гарантують, що навіть у випадках насичених деталей зображення або складного фону модель може ефективно визначати цільові області. Використання карт високої роздільної здатності на виході декодера дозволяє механізму самоуваги генерувати більш точні результати, досягаючи ефективного виявлення та розпізнавання ознак ЛП на складному фоні спостереження. Для полегшення структури декодера на вході використовується багатошаровий перцептрон MLP, а функції з останніх трьох етапів агрегуються, зберігаючи високу обчислювальну ефективність моделювання глобального контексту.

Механізм SegNeXt може бути ефективно вбудовано у традиційний модуль C2f проміжного рівня моделі YOLOv8n, як показано на рис. 9, що дозволяє інтегрувати функції різних рівнів, гарантуючи, що, ефективно розпізнаючи глобальну форму ЛП, він при цьому може точно ідентифікувати краї та деталі осередків горіння.

5. Результати дослідження.

Модель YOLOv8 містить значну кількість згорткових операцій, завдяки чому схильна до неефективних обчислень. Відповідно, збільшення обчислювального навантаження не сприяє вирішенню задачі дистанційного зондування ЛП у реальному часі.

Запропонована в статті модель (назвемо її YOLOv8N) містить низку архітектурних вдосконалень, які мають зменшити її обчислювальну складність, забезпечуючи продуктивність, достатню для дистанційного зондування ЛП в реальному часі, при збереженні високої точності для ідентифікації осередків горіння малого розміру, які звичайно виникають на початковому етапі ЛП. Однак, з балансом точності і ефективності не все виглядає так просто: якщо для підвищення продуктивності моделі зменшувати кількість параметрів на магістральному рівні, модель втрачає значну кількість детальної інформації малого масштабу, при цьому втрачаючи можливість ідентифікувати осередки горіння малого розміру, і навпаки.

Отже, запроваджені в моделі YOLOv8N новачки і архітектурні вдосконалень потребують експериментального дослідження. Експеримент проводився на прототипі тактичної системи моніторингу лісових пожеж [36, 37], що заснована на мережі персональних комп'ютерів (процесор Intel Core i7-14700K 3.4GHz/33MB, 32ГБ оперативної пам'яті, відеокарта NVIDIA GeForce RTX4090, 24ГБ відеопам'яті). Програмне забезпечення реалізовано мовою Python з використанням бібліотеки PyTorch, операційна система 64-розрядна Windows 11.

Оцінка точності моделі визначається за чотирма традиційними метриками: влучність P , повнота R , середнє гармонійне $F1$ та точність $Accuracy$. Влучність P вимірює частку фактичних осередків ЛП, правильно

ідентифікованих моделлю, повнота R подає частку осередків ЛП, правильно ідентифікованих моделлю, серед усіх фактичних позитивних зразків, $F1$ є гармонійним середнім значенням влучності та повноти, що забезпечує усереднене уявлення про загальну продуктивність моделі, а точність $Accuracy$ є часткою всіх правильно класифікованих даних (позитивних і негативних результатів). Ці показники обчислюються наступним чином:

$$P = \frac{TP}{TP + FP}; \quad R = \frac{TP}{TP + FN}; \quad F1 = \frac{2 \cdot P \cdot R}{P + R};$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN},$$

де TP – істинно-позитивний результат: кількість справжніх спрацьовувань, тобто кількість осередків ЛП, правильно виявлених моделями; TN – істинно-негативний результат, коли певна ділянка, на якій немає ЛП, класифікована як така, що не містить ЛП; FP – хибно-позитивний результат: кількість помилкових спрацьовувань, що вказує на виявлення моделями осередків ЛП, які насправді не є пожежами; FN – кількість хибно негативних результатів, яка вказує на те, що модель не виявила осередки ЛП, які насправді існують.

Крім того, було досліджено кількість параметрів моделі N (включаючи вагові коефіцієнти та зміщення) та кількість операцій з плаваючою комою за секунду (FLOP) для подальшого вимірювання складності та обчислювальних вимог: показник FLOP безпосередньо визначає швидкість навчання та логічного висновку моделі глибокого навчання – чим він більше, тим вищі обчислювальні потреби моделі та суворіші вимоги до апаратного забезпечення.

Експеримент проводився для порівняння запропонованої моделі YOLOv8N з моделлю YOLOv8n.

Для аналізу було вибрано загальнодоступний великий набір даних дистанційного зондування із зображеннями лісових пожеж M4SFWD, знятих з точки зору БПЛА [38]. Він містить різні типи місцевості, метеорологічних умов, інтенсивності освітлення та різноманітні лісові пожежі, всього 3974 зображень з ознаками вогню і диму. Зображення мають різні розміри, щоб задовольнити різноманітні сценарії експерименту. Всі зображення були аналізовані в їх реальному розмірі. Набір даних ретельно аналізувався, і всі випадки вогню та диму інтерпретувалися досліджуваною моделлю.

В процесі експерименту набір даних було розділено на набори зображень для навчання, перевірки та тестування із співвідношенням розподілу відповідно 5:1:1. Експеримент дав наступні результати (табл. 1).

Таблиця 1

Результати експерименту

Модель	P , %	R , %	$F1$, %	$Accuracy$, %	N	FLOPs
YOLOv8n	85,3	83,2	87,3	91,4	5	8,7
YOLOv8N	89,6	90,7	92,1	97,3	3	7,1
	+4,3	+7,5	+4,8	+5,9	-2	-1,6

Порівняння отриманих результатів ілюстровано на рис. 10, рис. 11.

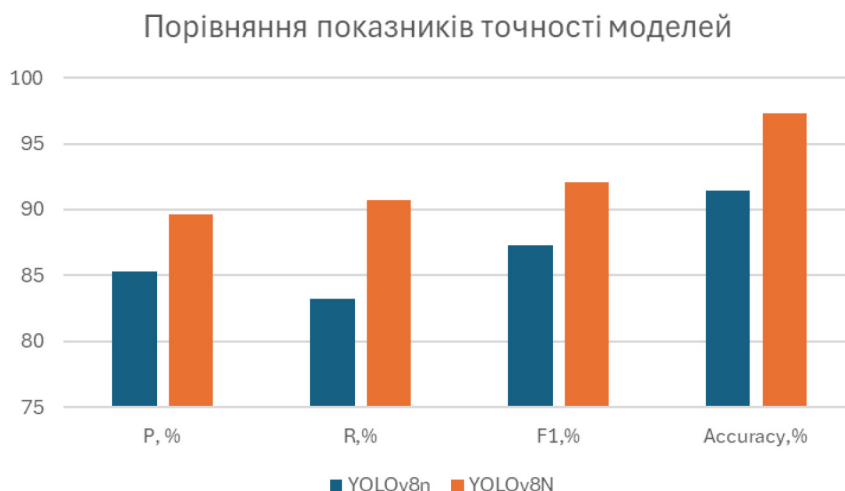


Рис. 10. Порівняння отриманих показників точності моделей YOLOv8n та YOLOv8N

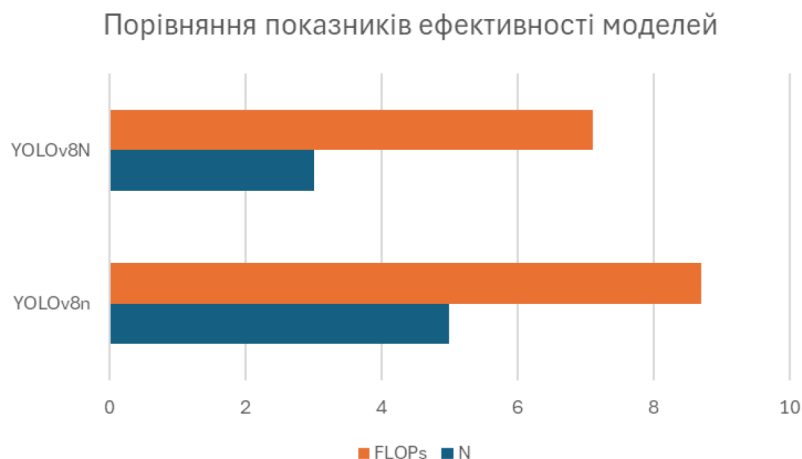


Рис. 11. Порівняння отриманих показників ефективності моделей YOLOv8n та YOLOv8N

Експериментальні результати, подані в табл. 1, показують, що модель YOLOv8N має кращі показники точності розпізнавання ознак ЛП (влучність, повнота, середнє гармонійне і точність), ніж базова модель YOLOv8n, відповідно на 4,3 %, 7,5 %, 4,8 % і 5,9 %. Це пояснюється, в першу чергу, вдосконаленням магістрального шару, що покращує здатність моделі виявляти невеликі за розмірами осередки ЛП. Крім того, вдосконалення проміжного шару покращує здатність моделі точно виявляти осередки ЛП на складному фоні. Заміна звичайних згорткових шарів на GhostNetV21 дозволила зменшити кількість параметрів, зменшуючи обчислювальну складність моделі, що, в свою чергу, дозволило знизити обчислювальні потреби моделі та забезпечити відповідність умовам дистанційного зондування ЛП в реальному часі.

Отже, запропонована модель YOLOv8N значно підвищує точність виявлення ЛП, зберігаючи при цьому достатню продуктивність, що має велике значення для своєчасного виявлення та моніторингу ЛП.

Висновки

1. В роботі запропоновано «полегшену» з погляду на обчислювальну складність модель розпізнавання осередків ЛП при дистанційному зондуванні з БПЛА YOLOv8N, яка покращує базову модель YOLOv8n за рахунок застосування:

- на магістральному рівні модуля GhostNetv2 разом з модулем уваги DFC замість традиційної операції згортки Conv, що дозволяє суттєво зменшити кількість параметрів моделі, зберігаючи її продуктивність;
- на магістральному рівні механізму звернення уваги MHSA в операціях C2f, що покращує здатність отримувати ознаки осередків горіння, а отже, підвищує точність виявлення порівняно невеликих за розміром ЛП на різних цільових масштабах зображень за рахунок більш глибокого дослідження різних репрезентативних підпросторів та врахування більш детальної контекстної інформації;
- на проміжному рівні механізму самоуваги SegNeXt в операціях C2f, що дозволяє підвищити точність виявлення ознак ЛП у складних умовах за рахунок акценту уваги на цільові об'єкти на складному фоні, пригнічення шумів і покращення здатності моделі розпізнавати ознаки ЛП на основі поглибленого семантичного контрасту.

2. Запропонована модель YOLOv8N збільшує показники влучності, повноти, середнього гармонійного і точності на 4,3 %, 7,5 %, 4,8 % і 5,9 % порівняно з моделлю YOLOv8n, при цьому кількість параметрів зменшується на 40 %. Отже, модель YOLOv8N забезпечує високий рівень точності виявлення ознак ЛП, зберігаючи при цьому баланс між обчислювальною складністю і ефективністю моделі, що гарантує її спроможність працювати в системах дистанційного зондування ЛП з БПЛА в режимі реального часу.

3. Впровадження моделі YOLOv8N для дистанційного зондування ЛП з БПЛА дає особі, яка приймає рішення, більш точне і надійне візуальне уявлення про динаміку ЛП, а отже, вірогідність прийняття рішення щодо реагування на ЛП підвищується, що робить доцільним практичне використання запропонованої моделі в тактичних системах моніторингу ЛП.

Список використаної літератури

1. Hillayová, M., Holécý, J., Koristeková, K., Bakšová, M., Ostrihoň, M., Škvarenina, J.: Ongoing Climatic Change Increases the Risk of Wildfires. Case Study: Carpathian Spruce Forests. *Journal of Environmental Management*, 2023, vol. 337, pp. 117620.
2. Stocks, B., Mason, J., Todd, J., Bosch, E.M., Wotton, B.M., Amiro, B.D., Flannigan, M.D., Hirsch, K.G., Logan, K.A., Martell, D.L.: Large Forest Fires in Canada, 1959–1997. *J. of Geographical Research*, 2002, vol. 107, 8149.

3. Howell, A.N., Belmont, E.L., McAllister, S.S., Finney, M.A.: An Investigation of Oxygen Availability in Spreading Fires. *Fire Technology*, 2023, vol. 59, pp. 2147–2176.
4. Sherstjuk, V., Zharikova, M., Dorovskaja, I., Sheketa, V.: Assessing Forest Fire Dynamics in UAV-Based Tactical Monitoring System. *Advances in Intelligent Systems and Computing*, 2021, vol. 1246, pp. 285–301.
5. Sherstjuk, V., Zharikova, M., Dorovskaja, I.: 3D Fire Front Reconstruction in UAV-Based Forest-Fire Monitoring System. *Data Stream Mining & Processing: Proc. of IEEE Third Int. Conf. DSMP'2020*, Lviv, 2020, pp. 243–248.
6. Bessho, K., Date, K., Hayashi, M., Ikeda, A., Imai, T., Inoue, H., Kumagai, Y., Miyakawa, T., Murata, H., Ohno, T.: An Introduction to Himawari-8/9 Japan's New-Generation Geostationary Meteorological Satellites. *Journal of the Meteorological Society of Japan*, Ser. II 2016, vol. 94, pp. 151–183.
7. Justice, C. O., Giglio, L., Korontzi, S., Owens, J., Morisette, J. T., Roy, D., Descloitres, J., Alleaume, S., Petitcolin, F., Kaufman, Y.: The MODIS Fire Products. *Remote sensing of Environment*, 2002, vol. 83(1–2), pp. 244–262.
8. Sherstjuk, V., Zharikova, M., Dorovskaja, I., Chorny, D., Gusev, V., Sokol, I.: 3D Fire Front Reconstruction Based on Multi-View Observation and Complex Uncertainty Model. *Advanced Computer Information Technologies: Proc. of 11th Int. Conf., Deggendorf, Germany, 2021*, pp. 761–765.
9. Osco, L.P., Junior, J.M., Ramos, A.P.M., de Castro Jorge, L.A., Fatholahi, S.N., de Andrade Silva, J., Matsubara, E. T., Pistori, H., Gonçalves, W.N., Li, J.: A Review on Deep Learning in UAV Remote Sensing. *International Journal of Applied Earth Observation and Geoinformation*, 2021, vol. 102, 102456.
10. Li, C., Li, G., Song, Y., He, Q., Tian, Z., Xu, H., Liu, X.: Fast Forest Fire Detection and Segmentation Application for UAV-Assisted Mobile Edge Computing System. *IEEE Internet of Things Journal*, 2023, vol. 11(16), pp. 26690–26699.
11. Yang, X., Hua, Z., Zhang, L., Fan, X., Zhang, F., Ye, Q., Fu, L.: Preferred Vector Machine for Forest Fire Detection. *Pattern Recognition*, 2023, vol. 143, 109722.
12. Maeda, N., Tonooka, H.: Early Stage Forest Fire Detection from Himawari-8 AHI Images Using a Modified MOD14 Algorithm Combined with Machine Learning. *Sensors*, 2022, vol. 23, 210.
13. Sathishkumar, V.E., Cho, J., Subramanian, M., Naren, O.S.: Forest Fire and Smoke Detection Using Deep Learning-Based Learning without Forgetting. *Fire Ecology*, 2023, vol. 19, p. 9.
14. Liu, F., Chen, R., Zhang, J., Xing, K., Liu, H., Qin, J.: R2YOLOX: A Lightweight Refined Anchor-Free Rotated Detector for Object Detection in Aerial Images. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, vol. 60, pp. 1–15.
15. Guan, R., Li, Z., Tu, W., Wang, J., Liu, Y., Li, X., Tang, C., Feng, R.: Contrastive Multi-View Subspace Clustering of Hyperspectral Images Based on Graph Convolutional Networks. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, vol. 62, pp. 1–14.
16. Lv, Q., Quan, Y., Sha, M., Feng, W., Xing, M.: Deep Neural Network-Based Interrupted Sampling Deceptive Jamming Countermeasure Method. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2022, vol. 15, pp. 9073–9085.
17. Li, Y., Zhang, S., Wang, W.-Q.: A Lightweight Faster R-CNN for Ship Detection in SAR Images. *IEEE Geoscience and Remote Sensing Letters*, 2020, vol. 19, pp. 1–5.
18. Zhang, W., Liu, Z., Zhou, S., Qi, W., Wu, X., Zhang, T., Han, L.: LS-YOLO: A Novel Model for Detecting Multi-Scale Landslides with Remote Sensing Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024, 1–14.
19. Zhang, L., Wang, M., Ding, Y., Bu, X.: MS-FRCNN: A Multi-Scale Faster RCNN Model for Small Target Forest Fire Detection. *Forests*, 2023, vol. 14(3), 616.
20. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You Only Look Once: Unified, Real-Time Object Detection. *Arxiv*, 2016, 1506.02640.
21. Niu, Z., Zhong, G., Yu, H.: A Review on the Attention Mechanism of Deep Learning. *Neurocomputing*, 2021, vol. 452, pp. 48–62.
22. Luo, M., Xu, L., Yang, Y., Cao, M., Yang, J.: Laboratory Flame Smoke Detection Based on an Improved YOLOX Algorithm. *Applied Sciences*, 2022, vol. 12(24), 12876.
23. Xue, Q., Lin, H., Wang, F.: FCDM: An Improved Forest Fire Classification and Detection Model Based on YOLOv5. *Forests*, 2022, vol. 13(12), 2129.
24. Chen, G., Cheng, R., Lin, X., Jiao, W., Bai, D., Lin, H.: LMDFS: A Lightweight Model for Detecting Forest Fire Smoke in UAV Images Based on YOLOv7. *Remote Sensing*, 2023, vol. 15(15), 3790.
25. Wang, Z., Hua, Z., Wen, Y., Zhang, S., Xu, X., Song, H.: E-YOLO: Recognition of Estrus Cow Based on Improved YOLOv8n Model. *Expert Systems with Applications*, 2024, vol. 238, 122212.
26. Mahasin, M., Dewi, I.A.: Comparison of CSPDarkNet53, CSPResNeXt-50, and EfficientNet-B0 Backbones on YOLO V4 as Object Detector. *International Journal of Engineering, Science and Information Technology*, 2022, vol 2(3), pp. 64–72.

27. Yi, H., Liu, B., Zhao, B., Liu, E.: Small Object Detection Algorithm Based on Improved YOLOv8 for Remote Sensing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023, vol. 17, pp. 1734–1747.
28. Neubeck, A., Van Gool, L.: Efficient Non-Maximum Suppression. *Pattern recognition: Proc. of the 18th Int. Conf. ICPR'06*, Hong Kong, China, 2006, vol. 3, pp. 850–855.
29. Tang, Y., Han, K., Guo, J., Xu, C., Xu, C., Wang, Y.: GhostNetv2: Enhance Cheap Operation with Long-Range Attention. *Advances in Neural Information Processing Systems*, 2022, vol. 35, pp. 9969–9982.
30. Zhang, Z.: Drone-YOLO: An Efficient Neural Network Method for Target Detection in Drone Images. *Drones*, 2023, 7(8), 526.
31. Zhao, F., Feng, F., Ye, S., Mao, Y., Chen, X., Li, Y., Ning, M., Zhang, M.: Multi-head self-attention mechanism-based global feature learning model for ASD diagnosis. *Biomedical Signal Processing and Control*, 2024, vol. 91, 106090.
32. Zhao, Z., Du, J., Li, C., Fang, X., Xiao, Y., Tang, J.: Dense Tiny Object Detection: A Scene Context Guided Approach and a Unified Benchmark. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, vol. 62, pp. 1–13.
33. Wang, Z., Meng, Q., Tang, F., Qi, Y., Li, B., Liu, X., Kong, S., Li, X.: MSG-YOLO: A Lightweight Detection Algorithm for Clubbing Finger Detection. *Electronics*, 2024, vol. 13(22), 4549.
34. Guo, M.-H., Lu, C.-Z., Hou, Q., Liu, Z.-N., Cheng, M.-M., Hu, S.-M.: SegNeXt: rethinking convolutional attention design for semantic segmentation. *Neural Information Processing Systems: Proc. of the 36th Int. Conf. NIPS'22*, Red Hook, NY, USA, 2022, vol. 84, pp. 1140–1156.
35. Guo, M., Lu, C., Hou, Q., Liu, Z., Cheng, M., Hu, S.: SegNeXt: Rethinking Convolutional Attention Design for Semantic Segmentation. *ArXiv*, 2022, abs/2209.08575.
36. Sherstjuk, V., Zharikova, M.: Fire-Front Recognition in UAV-Based Forest-Fire Monitoring System Using Fuzzy Rough Soft Sets. *Electrical and Computer Engineering: Proc. of IEEE 2nd Ukraine Conference UKRCON*, Lviv, Ukraine, 2019, pp. 1091–1096.
37. Sherstjuk, V., Zharikova, M.: Evaluation of Fire Intensity Based on Neural Networks in a Forest-Fire Monitoring System. *Electronics and Nanotechnology: Proc. of IEEE 39th Int. Conf. ELNANO'2019*, Kyiv, Ukraine, 2019, pp. 802–807.
38. Wang, G., Li, H., Li, P., Lang, X., Feng, Y., Ding, Z., Xie, S.: M4SFWD: A Multi-Faceted Synthetic Dataset for Remote Sensing Forest Wildfires Detection. *Expert Systems with Applications*, 2024, vol. 245, 123489.

References

1. Hillaýová, M., Holéý, J., Korísteková, K., Bakšová, M., Ostrihoň, M., Škvarenina, J. (2023) Ongoing Climatic Change Increases the Risk of Wildfires. Case Study: Carpathian Spruce Forests. *Journal of Environmental Management*, vol. 337, pp. 117620. DOI: <https://doi.org/10.1016/j.jenvman.2023.117620>
2. Stocks, B., Mason, J., Todd, J., Bosch, E.M., Wotton, B.M., Amiro, B.D., Flannigan, M.D., Hirsch, K.G., Logan, K.A., Martell, D.L. (2002) Large Forest Fires in Canada, 1959–1997. *J. of Geographical Research*, vol. 107. DOI: <https://doi.org/10.1029/2001JD000484>
3. Howell, A.N., Belmont, E.L., McAllister, S.S., Finney, M.A. (2023) An Investigation of Oxygen Availability in Spreading Fires. *Fire Technology*, vol. 59, pp. 2147–2176. DOI: <https://doi.org/10.1007/s10694-023-01396-6>
4. Sherstjuk, V., Zharikova, M., Dorovskaja, I., Sheketa, V. (2021). Assessing Forest Fire Dynamics in UAV-Based Tactical Monitoring System. *Advances in Intelligent Systems and Computing*, 2021, vol. 1246, pp. 285–301. DOI: https://doi.org/10.1007/978-3-030-54215-3_18
5. Sherstjuk, V., Zharikova, M., Dorovskaja, I. (2020) 3D Fire Front Reconstruction in UAV-Based Forest-Fire Monitoring System. *Data Stream Mining & Processing: Proc. of IEEE Third International Conference DSMP'2020*, Lviv, pp. 243–248. DOI: <https://doi.org/10.1109/DSMP47368.2020.9204196>
6. Bessho, K., Date, K., Hayashi, M., Ikeda, A., Imai, T., Inoue, H., Kumagai, Y., Miyakawa, T., Murata, H., Ohno, T. (2016) An Introduction to Himawari-8/9 Japan's New-Generation Geostationary Meteorological Satellites. *Journal of the Meteorological Society of Japan*, Ser. II, vol. 94, pp. 151–183. DOI: <https://doi.org/10.2151/jmsj.2016-009>
7. Justice, C.O., Giglio, L., Korontzi, S., Owens, J., Morisette, J.T., Roy, D., Descloitres, J., Alleaume, S., Petitcolin, F., Kaufman, Y. (2002) The MODIS Fire Products. *Remote sensing of Environment*, vol. 83(1–2), pp. 244–262. DOI: [https://doi.org/10.1016/S0034-4257\(02\)00076-7](https://doi.org/10.1016/S0034-4257(02)00076-7)
8. Sherstjuk, V., Zharikova, M., Dorovskaja, I., Chornyi, D., Gusev, V., Sokol, I. (2021) 3D Fire Front Reconstruction Based on Multi-View Observation and Complex Uncertainty Model. *Advanced Computer Information Technologies: Proc. of 2021 11th International Conference, Deggendorf, Germany*, pp. 761–765. DOI: <https://doi.org/10.1109/ACIT52158.2021.9548517>
9. Osco, L.P., Junior, J.M., Ramos, A.P.M., de Castro Jorge, L.A., Fatholahi, S.N., de Andrade Silva, J., Matsubara, E.T., Pistori, H., Gonçalves, W.N., Li, J. (2021) A Review on Deep Learning in UAV Remote Sensing. *International Journal of Applied Earth Observation and Geoinformation*, vol. 102, 102456. DOI: <https://doi.org/10.1016/j.jag.2021.102456>

10. Li, C., Li, G., Song, Y., He, Q., Tian, Z., Xu, H., Liu, X. (2023) Fast Forest Fire Detection and Segmentation Application for UAV-Assisted Mobile Edge Computing System. *IEEE Internet of Things Journal*, vol.11(16), pp. 26690–26699. DOI: <https://doi.org/10.1109/JIOT.2023.3311950>
11. Yang, X., Hua, Z., Zhang, L., Fan, X., Zhang, F., Ye, Q., Fu, L. (2023) Preferred Vector Machine for Forest Fire Detection. *Pattern Recognition*, vol. 143, 109722. DOI: <https://doi.org/10.1016/j.patcog.2023.109722>
12. Maeda, N., Tonooka, H. (2022) Early Stage Forest Fire Detection from Himawari-8 AHI Images Using a Modified MOD14 Algorithm Combined with Machine Learning. *Sensors*, vol. 23(1), 210. DOI: <https://doi.org/10.3390/s23010210>
13. 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*, vol. 19, p. 9. DOI: <https://doi.org/10.1186/s42408-022-00165-0>
14. Liu, F., Chen, R., Zhang, J., Xing, K., Liu, H., Qin, J. (2022) R2YOLOX: A Lightweight Refined Anchor-Free Rotated Detector for Object Detection in Aerial Images. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–15. DOI: <https://doi.org/10.1109/TGRS.2022.3215472>
15. Guan, R., Li, Z., Tu, W., Wang, J., Liu, Y., Li, X., Tang, C., Feng, R. (2024) Contrastive Multi-View Subspace Clustering of Hyperspectral Images Based on Graph Convolutional Networks. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–14. DOI: <https://doi.org/10.1109/TGRS.2024.3370633>
16. Lv, Q., Quan, Y., Sha, M., Feng, W., Xing, M. (2022) Deep Neural Network-Based Interrupted Sampling Deceptive Jamming Countermeasure Method. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 9073–9085. DOI: <https://doi.org/10.1109/JSTARS.2022.3214969>
17. Li, Y., Zhang, S., Wang, W.-Q. (2020) A Lightweight Faster R-CNN for Ship Detection in SAR Images. *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5. DOI: <https://doi.org/10.1109/LGRS.2020.3038901>
18. Zhang, W., Liu, Z., Zhou, S., Qi, W., Wu, X., Zhang, T., Han, L. (2024) LS-YOLO: A Novel Model for Detecting Multi-Scale Landslides with Remote Sensing Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1–14, DOI: <https://doi.org/10.1109/JSTARS.2024.3363160>
19. Zhang, L., Wang, M., Ding, Y., Bu, X. (2023) MS-FRCNN: A Multi-Scale Faster RCNN Model for Small Target Forest Fire Detection. *Forests*, vol. 14(3), 616. DOI: <https://doi.org/10.3390/f14030616>
20. Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. Arxiv, 1506.02640. DOI: <https://doi.org/10.48550/arXiv.1506.02640>
21. Niu, Z., Zhong, G., Yu, H. (2021) A Review on the Attention Mechanism of Deep Learning. *Neurocomputing*, vol. 452, pp. 48–62. DOI: <https://doi.org/10.1016/j.neucom.2021.03.091>
22. Luo, M., Xu, L., Yang, Y., Cao, M., Yang, J. (2022) Laboratory Flame Smoke Detection Based on an Improved YOLOX Algorithm. *Applied Sciences*, vol. 12(24), 12876. DOI: <https://doi.org/10.3390/app122412876>
23. Xue, Q., Lin, H., Wang, F. (2022) FCDM: An Improved Forest Fire Classification and Detection Model Based on YOLOv5. *Forests*, vol. 13(12), 2129. DOI: <https://doi.org/10.3390/f13122129>
24. Chen, G., Cheng, R., Lin, X., Jiao, W., Bai, D., Lin, H. (2023) LMDFS: A Lightweight Model for Detecting Forest Fire Smoke in UAV Images Based on YOLOv7. *Remote Sensing*, vol. 15(15), 3790. DOI: <https://doi.org/10.3390/rs15153790>
25. Wang, Z., Hua, Z., Wen, Y., Zhang, S., Xu, X., Song, H. (2024) E-YOLO: Recognition of Estrus Cow Based on Improved YOLOv8n Model. *Expert Systems with Applications*, vol. 238, 122212. DOI: <https://doi.org/10.1016/j.eswa.2023.122212>
26. Mahasin, M., Dewi, I.A. (2022) Comparison of CSPDarkNet53, CSPResNeXt-50, and EfficientNet-B0 Backbones on YOLO V4 as Object Detector. *International Journal of Engineering, Science and Information Technology*, vol 2(3), pp. 64–72. DOI: <https://doi.org/10.52088/ijesty.v2i3.291>
27. Yi, H., Liu, B., Zhao, B., Liu, E. (2023) Small Object Detection Algorithm Based on Improved YOLOv8 for Remote Sensing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 1734–1747. DOI: <https://doi.org/10.1109/JSTARS.2023.3339235>
28. Neubeck, A., Van Gool, L. (2006) Efficient Non-Maximum Suppression. Pattern recognition: Proc. of the 18th Int. Conf. ICPR'06, Hong Kong, China, vol. 3, pp. 850–855. DOI: <https://doi.org/10.1109/ICPR.2006.479>
29. Tang, Y., Han, K., Guo, J., Xu, C., Xu, C., Wang, Y. (2022) GhostNetv2: Enhance Cheap Operation with Long-Range Attention. *Advances in Neural Information Processing Systems*, vol. 35, pp. 9969–9982. DOI: <https://doi.org/10.48550/arXiv.2211.12905>
30. Zhang, Z. (2023) Drone-YOLO: An Efficient Neural Network Method for Target Detection in Drone Images. *Drones*, 7(8), 526. DOI: <https://doi.org/10.3390/drones7080526>
31. Zhao, F., Feng, F., Ye, S., Mao, Y., Chen, X., Li, Y., Ning, M., Zhang, M. (2024) Multi-head self-attention mechanism-based global feature learning model for ASD diagnosis. *Biomedical Signal Processing and Control*, vol. 91, 106090. DOI: <https://doi.org/10.1016/j.bspc.2024.106090>

32. Zhao, Z., Du, J., Li, C., Fang, X., Xiao, Y., Tang, J. (2024) Dense Tiny Object Detection: A Scene Context Guided Approach and a Unified Benchmark. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–13. DOI: <https://doi.org/10.1109/TGRS.2024.3357706>.
33. Wang, Z., Meng, Q., Tang, F., Qi, Y., Li, B., Liu, X., Kong, S., Li, X. (2024) MSG-YOLO: A Lightweight Detection Algorithm for Clubbing Finger Detection. *Electronics*, vol. 13(22), 4549. DOI: 10.3390/electronics13224549
34. Guo, M.-H., Lu, C.-Z., Hou, Q., Liu, Z.-N., Cheng, M.-M., Hu, S.-M. (2022) SegNeXt: rethinking convolutional attention design for semantic segmentation. *Neural Information Processing Systems: Proc. of the 36th Int. Conf. NIPS'22*, Red Hook, NY, USA, vol. 84, pp. 1140–1156. DOI: <https://doi.org/10.5555/3600270.3600354>
35. Guo, M., Lu, C., Hou, Q., Liu, Z., Cheng, M., Hu, S. (2022). SegNeXt: Rethinking Convolutional Attention Design for Semantic Segmentation. *ArXiv*, abs/2209.08575. DOI: <https://doi.org/10.48550/arXiv.2209.08575>
36. Sherstjuk, V., Zharikova, M. (2019) Fire-Front Recognition in UAV-Based Forest-Fire Monitoring System Using Fuzzy Rough Soft Sets. *Electrical and Computer Engineering: Proc. of IEEE 2nd Ukraine Conf. UKRCON*, Lviv, Ukraine, pp. 1091–1096. DOI: <https://doi.org/10.1109/UKRCON.2019.8879829>
37. Sherstjuk, V., Zharikova, M. (2019) Evaluation of Fire Intensity Based on Neural Networks in a Forest-Fire Monitoring System. *Electronics and Nanotechnology: Proc. of IEEE 39th Int. Conf. ELNANO'2019*, Kyiv, Ukraine, pp. 802–807. DOI: <https://doi.org/10.1109/ELNANO.2019.8783410>
38. Wang, G., Li, H., Li, P., Lang, X., Feng, Y., Ding, Z., Xie, S. (2024) M4SFWD: A Multi-Faceted Synthetic Dataset for Remote Sensing Forest Wildfires Detection. *Expert Systems with Applications*, vol. 245, 123489. DOI: <https://doi.org/10.1016/j.eswa.2024.123489>