

DOI: <https://doi.org/10.36910/6775-2524-0560-2026-62-02>

УДК: 004.8:614.8

Кравець Ілля Вячеславович¹, аспірант

<https://orcid.org/0009-0001-8467-391X>

Угрин Дмитро Ілліч¹, професор, д.т.н.

<https://orcid.org/0000-0003-4858-4511>

Терлецький Тарас Володимирович², доцент, к.т.н.

<https://orcid.org/0000-0002-4114-0734>

¹Чернівецький національний університет ім. Ю. Федьковича, м. Чернівці, Україна

²Луцький національний технічний університет, м. Луцьк, Україна

ДОСЛІДЖЕННЯ ГЛИБИННОГО НАВЧАННЯ ДЛЯ РОЗПІЗНАВАННЯ УРАЖЕНЬ ОБ'ЄКТІВ У НАДЗВИЧАЙНИХ СИТУАЦІЯХ

Кравець І.В., Угрин Д.І., Терлецький Т.В. Дослідження глибокого навчання для розпізнавання уражень об'єктів у надзвичайних ситуаціях. Розглянуто задачу автоматичного розпізнавання та оцінювання уражень об'єктів у надзвичайних ситуаціях на основі сучасних методів глибокого навчання. Проведено аналіз сучасних методів глибокого навчання для задач автоматичного розпізнавання та оцінювання уражень об'єктів у надзвичайних ситуаціях. Також було розглянуто особливості застосування моделей глибокого навчання в умовах візуальних шумів, часткової видимості об'єктів та складного фону. На основі розглянутих підходів виявлення та сегментації було виявлено, що однорівневі архітектури мають обмеження при обробці даних з БПЛА. На основі цього запропоновано підхід до побудови комбінованої архітектури, що складається з швидкого виявлення за допомогою моделі YOLOv8, сегментацію використовуючи U-Net та класифікацію пошкоджень з EfficientNet. Запропонований підхід орієнтований на використання в умовах обмежених обчислювальних ресурсів та нестабільного зв'язку, що є критично важливим для зон надзвичайних ситуацій. В результаті було запропоновано модель обробки даних, яка враховує особливості різних типів вхідних даних, а також вплив зовнішніх факторів (дим, низьке освітлення). Отримані результати підтверджують доцільність використання комбінованих архітектур для підвищення надійності автоматизованої оцінки пошкоджень. Було продемонстровано результати тестового використання окремих складових запропонованої моделі, що показує можливість створення систем реагування на надзвичайні ситуації. Практична цінність дослідження полягає у формуванні концептуальних засад для створення автономних інтелектуальних систем реагування на надзвичайні ситуації.

Ключові слова: глибоке навчання, сегментація уражень, детекція пошкоджень, мультиспектральні зображення, функція втрат, класифікація ступеня ураження.

Kravets I., Ugrin D., Terletskyi T. Research on deep learning for recognizing damage to objects in emergency situations. The problem of automatic recognition and assessment of damage to objects in emergency situations based on modern deep learning methods is considered. An analysis of modern deep learning methods for automatic recognition and assessment of damage to objects in emergency situations is carried out. The peculiarities of applying deep learning models in conditions of visual noise, partial visibility of objects, and complex backgrounds were also considered. Based on the considered approaches to detection and segmentation, it was found that single-level architectures have limitations when processing data from UAVs. Based on this, an approach to building a combined architecture was proposed, consisting of rapid detection using the YOLOv8 model, segmentation using U-Net, and damage classification using EfficientNet. The proposed approach is designed for use in conditions of limited computing resources and unstable communication, which is critical for emergency areas. As a result, a data processing model was proposed that takes into account the peculiarities of different types of input data, as well as the influence of external factors (smoke, low lighting). The results confirm the feasibility of using combined architectures to improve the reliability of automated damage assessment. The results of testing individual components of the proposed model were demonstrated, showing the possibility of creating emergency response systems. The practical value of the research lies in the formation of conceptual foundations for the creation of autonomous intelligent emergency response systems.

Keywords: deep learning, damage segmentation, damage detection, multispectral images, loss function, damage classification.

Постановка наукової проблеми. Сьогодні ми часто спостерігаємо надзвичайні ситуації, такі як природні катастрофи, техногенні аварії, пожежі, вибухи та військові дії, які призводять до великих руйнувань будівель та об'єктів інфраструктури. За таких умов, для планування рятувальних операцій, розподілу ресурсів та швидкого реагування просто необхідна швидка попередня оцінка пошкоджень. Однак, звичні методи аналізу, що базуються на візуальній оцінці рятувальниками фотографій, або безпосередньо на місці, є повільними, небезпечними та часто суб'єктивними, що знижує ефективність реагування в критичний час після події.

Наразі вже існують рішення які працюють здебільшого на основі аналізу супутникових знімків, однак вони не враховують, що типи ураження можуть відрізнятися між собою. Також вони не враховують специфічні умови, у яких ці дані отримуються. Крім того, для отримання супутникових знімків потрібен значний час та кошти, що спричиняє затримку обробки, а також обмежує точність в умовах диму, пилу чи неповної руйнації. Також, ці методи не призначені для роботи на edge-пристроях (дрони, портативні комплекси, польові сервери), а це важливо у зонах

бойових дій, де відсутній стабільний зв'язок. Проблемними місцями також є виявлення зон пошкодження, визначення меж цих пошкоджень, визначення типу ураження (руйнування, обвал, пробоїна тощо), та оцінка ступеня пошкоджень.

Багато моделей глибинного навчання часто показують високу точність в лабораторних умовах. Однак, застосування цих моделей на реальних відеозаписах із камер чи тепловізорів ця точність знижується. Часто відсутні потрібні тестові датасети для тренування моделей, які можуть працювати в умовах надзвичайних ситуацій, а це створює труднощі для побудови ефективних систем. Через це виникає потреба у розробці системи автоматичного аналізу уражень, що поєднує сучасні архітектури глибинного навчання, таких як CNN, U-Net, EfficientNet, Vision Transformer, YOLO, та яка здатна працювати у режимі реального часу з різними типами даних. Така система має забезпечувати стійкість до шумів, неповної видимості об'єктів, змін освітлення, а також буде придатною для автономної роботи на пристроях з обмеженими обчислювальними ресурсами.

Аналіз останніх досліджень і публікацій. Для необхідного налізу джерел було проведено пошук наукових джерел у міжнародних базах Scopus, IEEE Xplore та Web of Science. Розглядалися публікації датовані 2024-2025 роками, оскільки в них використовуються останні наявні моделі комп'ютерного зору. Для пошуку використовувались такі ключові слова, як «deep learning damage assessment», «UAV real-time detection», «semantic segmentation for disaster relief», «edge computing AI». Основними критеріями вибору робіт на огляд були зосередженість на оцінці руйнувань/пошкоджень будинків чи інфраструктури, використання сучасних моделей комп'ютерного зору та наявність інформації про ефективність описаних моделей.

Загалом було опрацьовано близько 30 джерел, з яких було відібрано 17 найактуальніших робіт, що лягли в основу цього дослідження.

Провівши огляд та аналіз наукових робіт можна виділити три основні підходи до оцінки руйнувань і пошкоджень будинків та інфраструктури, а саме аналіз супутникових знімків, проведення моніторингу за допомогою БПЛА та порівняння сучасних архітектур нейромереж між собою.

Переважна частина підходів створена навколо аналізу та обробки супутникових знімків. Зокрема була створена CNN-архітектура для класифікації пошкоджень. Ця архітектура показала високу точність на датасетах із знімками з великим масштабом [1]. Основою цього методу був датасет xBD та моделі, що порівнювали знімки «до» та «після» [2]. В іншому дослідженні застосовувались злиття даних з супутників Sentinel-1 та Sentinel-2 за допомогою варіаційних автокодувальників [6]. Хоча ці дослідження й показали високу точність, вони сильно обмежені в плані часу, оскільки для отримання даних з супутника потрібен час (Рис. 1, Рис. 2).



Рис. 1, 2. Приклад роботи існуючої системи для аналізу пошкоджень будинків базуючись на супутникових знімках [2].

Залежність від прольоту супутника, низька роздільна здатність (що унеможливує детекцію дрібних тріщин чи пробоїн) та неможливість роботи в умовах хмарності роблять їх допоміжним, а не основним інструментом для тактичного рівня [8].

Моделі самі по собі споживають значну кількість ресурсів комп'ютера, а в деяких випадках виникала необхідність використання комп'ютерного зору у зв'язці з дронами. Це стимулювало розвиток легких моделей. Зокрема було показано ефективність моделі YOLOv8 для розпізнавання об'єктів у реальному часі. Вона досягнула mAP 0.91 при обробці відеопотоків, а ця швидкість є достатньою для використання міні-комп'ютера, який використовує відеопотік камери дрона [4]. Також такий підхід був застосований для детекції пошкоджень дахів будівель, що показує здатність роботи YOLO з різними ракурсами [7].

Хоча детектори (YOLO) й є швидкими, вони дають інформацію лише у вигляді шматочків зображення, що мають назву «bounding boxes». Це дає лише загальну інформацію про те, де саме знаходиться враження, а цього недостатньо для повної оцінки. Також такі детектори можуть некоректно спрацьовувати на специфічних ландшафтах чи поверхнях, що потребує додаткової перевірки [13].

Для вирішення цієї проблеми застосовуються методи сегментації, які безпосередньо створюють маску пошкодження. Зокрема, для цього була використана модель TDBOSegNet, яка робить сегментацію пошкоджених об'єктів в зонах стихійних лих [6]. Аналогічно використовується U-Net, який було адаптовано для створення детальних масок пошкоджень [7]. Головним недоліком використання чисто моделей сегментації високе споживання ресурсів. Сегментація повних зображень, навіть якщо обробляти не кожен, а n-ий кадр відеопотоку у реальному часі на edge-пристроях (типу Raspberry Pi або Jetson Nano) за допомогою U-Net є технічно неможливим без попереднього виокремлення пошкоджених зон [12].

Окремим питанням є вибір архітектури. Порівняльний аналіз показав, що Vision Transformers (ViT) перевершують CNN (MobileNetV3) у точності класифікації руйнувань (точність до 97% на тестових датасетах) [3, 5]. Однак обчислювальні витрати ViT залишаються занадто високими для використання мобільних платформах. За умов обмежених ресурсів edge-пристроїв оптимізовані CNN-моделі будуть найкраще справлятися обробкою зображень в режимі реального часу.

Проведений аналіз показує, що між швидкими, але грубими методами детекції (YOLO) та точними, але повільними методами сегментації (U-Net/Transformers) існує певний розрив. Жоден із розглянутих монолітних підходів не забезпечує одночасно і деталізацію оцінки пошкоджень, і можливість автономної роботи на edge-пристроях.

Систематизація методів та обґрунтування запропонованого підходу. Огляд та порівняльний аналіз існуючих підходів. У сучасних задачах розпізнавання уражень об'єктів у зоні надзвичайних ситуацій найбільш важливим та актуальним є забезпечення виявлення, локалізації, оцінки уражень та оперативного прийняття рішень у режимі реального часу з мінімальною затримкою. Для впровадження подібного рішення необхідно забезпечити роботу з різними типами даних, зокрема відеопотоки, медичні або мультиспектральні зображення, тепловізійні потоки та сенсорні дані. Особливість таких даних полягає у високій варіативності: руйнування можуть бути точковими або масштабними, проявлятися як руйнація будівлі або часткова пожежа, локальні обвали або комплексні структурні дефекти.

Тому класичні комп'ютерні методи виявилися недостатніми, а сучасні існуючі системи покладаються переважно на супутникові знімки, аналіз яких відбувається зі значною затримкою, оскільки необхідні зображення ще повинні бути зроблені самим супутником.

Аналіз сучасних наукових джерел та існуючих систем моніторингу, що широко використовуються та робота яких пов'язана з обробкою супутникових знімків та аерозйомки з дронів, показує, що існуючі підходи можна розділити за рівнем обробки інформації та особливостями архітектури. Зокрема вони діляться на три основні групи: методи детекції, методи семантичної сегментації та гібридні (каскадні) підходи.

Методи детекції (Detection-based methods). Ці методи орієнтовані на виявлення та локалізацію певного типу об'єктів, які подаються у вигляді рамок (bounding boxes). Типовим прикладом є моделі YOLO (You Only Look Once), SSD та R-CNN.

Вони розбивають зображення на вході на сітку, де кожна комірка перевіряється на наявність об'єкта, а після цього визначаються координати самої рамки. Наприклад, модель YOLOv8 чудово

працює для детекції транспортних засобів та людей на зображеннях чи відеопотоках. Швидкість детекції на таких кейсах досягає близько 238 мс, а точність - 0.91% (Рис. 3).

Таблиця 1. Порівняльна таблиця характеристик архітектур YOLOv8 та Faster R-CNN [4].

Модель	Роздільна здатність кадру	Точність	Швидкість виконання (мс)
Faster R-CNN	900x900	0.64 %	2766.94 мс
YOLOv8	900x900	0.72 %	238.41 мс

Також натреновану YOLOv4 використали на супутникових знімках для виявлення будівель після землетрусів. Недоліком є те, що метод детекції надає лише загальну інформацію про те, в якому саме місці знаходиться руйнування. При цьому незрозуміло, яка саме форма руйнування. Крім того, такі детектори можуть не працювати з дрібними об'єктами, які містяться на супутникових знімках, якщо їх попередньо не оптимізувати.

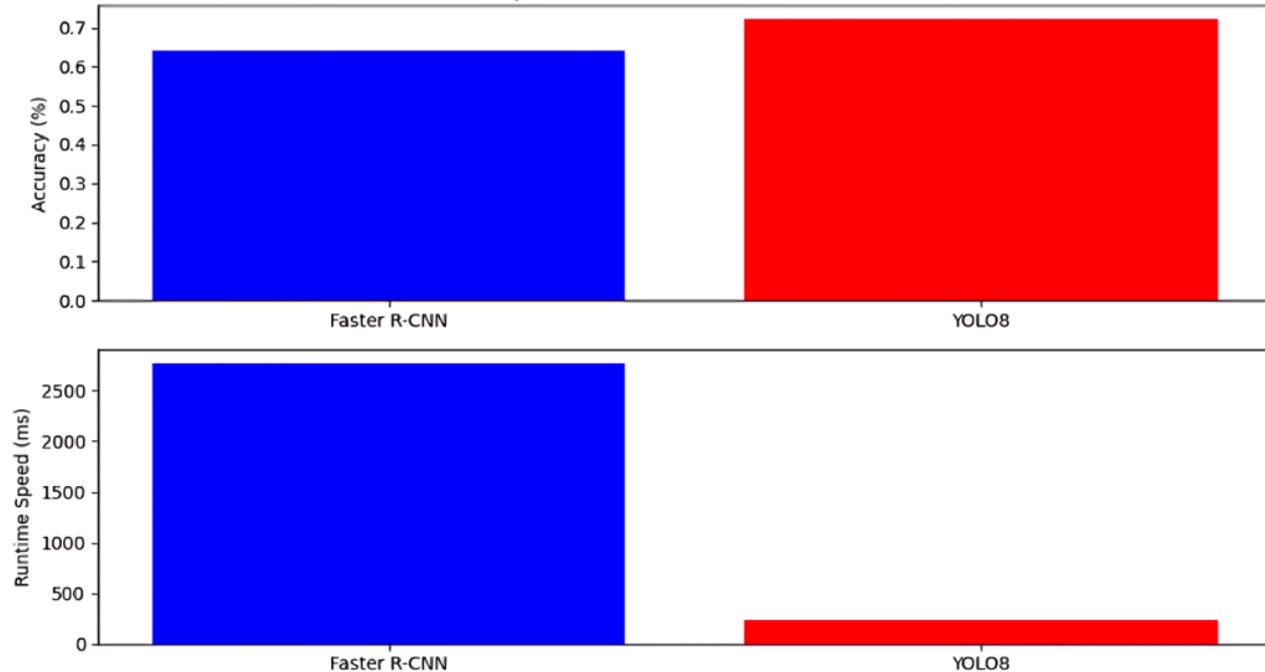


Рис. 3. Графік порівняння архітектур YOLOv8 та Faster R-CNN [4].

Методи семантичної сегментації (Segmentation-based methods). Цей підхід передбачає класифікацію кожного пікселя зображення, що дозволяє отримати точні контури пошкоджених зон. Для цього широко використовуються архітектури U-Net, SegNet, PSPNet. Наприклад, модель TDBOSegNet (створена на базі U-Net) ефективно використовується для сегментації об'єктів на знімках з БПЛА, що зазнали пошкоджень внаслідок стихійного лиха. До недоліків підходу відносять високу залежність від якості розмічених даних та проблему втрати інформації про простір. Також традиційні методи можуть генерувати надлишкові дані для передачі з борту БПЛА.

Гібридні та каскадні методи (Hybrid/Cascade methods). З огляду на досвід використання попередніх підходів, найбільш ефективним рішенням є поєднання декількох моделей. Такі методи використовують вихідні дані однієї мережі як вхідні для іншої. Прикладом є двоетапний метод, де спочатку детектор (YOLO) локалізує будівлю, а потім класифікатор (SVM або CNN) визначає ступінь пошкодження.

Порівняльний аналіз архітектур: CNN vs Vision Transformers (ViT)

В даний час для задач комп'ютерного зору активно досліджуються як згорткові нейромережі (CNN), так і трансформери (ViT). У кожній архітектурі є свої переваги і недоліки. Дослідження 2024-2025 років показують, що архітектури ViT демонструють кращу здатність до узагальнення та вищу точність у завданнях класифікації складних форм руйнувань (точність 0.97 проти 0.95 у MobileNetV3) завдяки механізмам самоуваги (self-attention) (Рис. 4, Рис. 5). Ці підходи ефективно працюють із глобальним контекстом, на відміну від CNN, які орієнтуються на локальні області зображення.

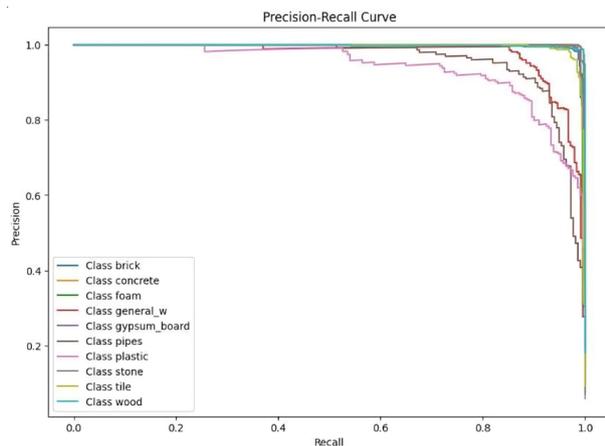


Рис. 4. Криві Precision-Recall для моделі ViT [3]

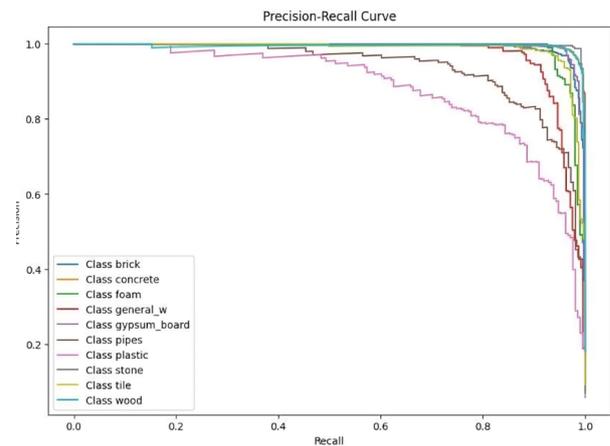


Рис. 5. Криві Precision-Recall для моделі MobileNetV3 [3]

Обґрунтування переваги CNN для задач реального часу

Оскільки кінцевою метою є інтеграція моделі на edge-пристроях (бортових системах БПЛА), для цих цілей найчастіше застосовуються саме CNN-моделі. Незважаючи на вищу точність трансформерів, CNN залишаються кращим вибором з наступних причин оскільки трансформери мають високі обчислювальні витрати, що збільшує складність при розгортанні на пристроях з обмеженими обчислювальними ресурсами. В свою чергу архітектури типу MobileNetV3 та YOLOv8 спеціально оптимізовані для швидкої детекції. Також для задач реального часу критичною є низька затримка. Оптимізовані моделі YOLOv8 досягають швидкості обробки близько 238 мс, тоді як важчі моделі потребують значно більше часу (до 2766 мс). Плюсом також є те, що квантовані моделі CNN (наприклад, YOLOv8-INT8) забезпечують достатній рівень точності при значному зменшенні розміру моделі (до 70%) та споживання пам'яті.

Узагальнюючи проведений аналіз, можна зробити висновок, що хоча CNN є оптимальним вибором для edge-пристроїв, жоден із розглянутих класів моделей самостійно не вирішує задачу комплексної оцінки руйнувань. Монолітні детектори (YOLO) забезпечують необхідну швидкість, але не дають точної інформації про площу пошкоджень. Семантична сегментація (U-Net) гарантує високу деталізацію, але є надто ресурсоємною для обробки повних кадрів у реальному часі. Виходячи з цього, у даній роботі пропонується методичне рішення, що полягає у каскадному поєднанні цих архітектур. Така стратегія дозволяє використовувати легкі детектори для фільтрації даних, а глибокі моделі - виключно для точкового аналізу, що забезпечує швидкість та точність.

2. Методологічні засади побудови каскадної архітектури

Наразі існує багато рішень глибокого навчання для роботи з різними типами задач, однак, як зазначалось вище неможливо об'єктивно провести розпізнавання уражень за допомогою лише однієї моделі глибокого навчання, оскільки жодна окрема модель не може водночас забезпечити і швидку детекцію, і високоточну сегментацію, і достовірну оцінку ступеня ураження. Тому в цьому дослідженні запропоновано комбінований підхід, який дозволить поєднати декілька архітектур в одну (Рис. 6.). Кожна підзадача розв'язує окрему задачу аналізу. Цей підхід дозволить компенсувати обмеження та недоліки однієї архітектури іншою.

Основою запропонованого підходу є модель, яка складається з декількох етапів: від початкового виокремлення вражених зон до визначення ступеня пошкодження. Основна ідея тому, що кожен етап працює виключно з валідними даними, які були попередньо відібрані та оброблені на попередніх етапах. Така архітектура дозволяє з'єднати і високу швидкість сучасних детекторів, і точність методів сегментації, і аналітику моделей класифікації. Ці етапи є критичними для задач моніторингу в реальному часі.



Рис. 6. Схема роботи комбінованого методу глибокого навчання

Крок 1. Першим етапом запропонованого підходу є автоматичне виявлення зон ураження на зображеннях та відеокадрах, отриманих, зокрема, з безпілотних літальних апаратів. Зокрема тут відбувається швидке сканування кадрів з метою визначення регіонів інтересу (ROI - Region of Interest), що можуть містити такі уражені об'єкти, як будівлі або елементи інфраструктури. Фактично, на цьому етапі модель має не просто побачити пошкоджені об'єкти, а швидко і точно локалізувати їх у кадрі у вигляді прямокутників які мають назву «bounding boxes». Для виконання цього завдання буде використано модель YOLOv8 модифікації Nano. Було обрано модифікацію Nano, оскільки існує необхідність оптимізації процесу під зйомку з дрона, де об'єкти часто мають малі кутові розміри.

Архітектуру YOLO було обрано, оскільки вона забезпечує ключові вимоги до системи. Модель здатна працювати в режимі реального часу, зокрема аналізуючи кожен п'ятий кадр відеопотоку. Оскільки YOLO належить до одноетапних детекторів, у яких детекція та класифікація виконуються за один прохід, забезпечується висока швидкість роботи та аналіз відео в реальному часі.

Також YOLO має офіційні та неофіційні реалізації з підтримкою експорту в ONNX, TensorRT, OpenVINO та інші фреймворки оптимізації. Це дозволяє розгорнути модель на Raspberry Pi, чи міні-комп'ютерах, що дозволить розгорнути її для роботи в польових умовах, без доступу до потужних серверів, наприклад на дроні.

YOLOv8 є новим детектором, в якому впроваджено модулі C2F (Cross-Stage Partial connections with Fusion) та SPPF (Spatial Pyramid Pooling Fast). Впровадження цих архітектурних рішень дозволяє моделі значно ефективніше захоплювати мультимасштабну інформацію та виділяти ознаки дрібних об'єктів. Дослідження підтверджують, що YOLOv8 здатна досягати показників точності на рівні 0.91% при детекції людей та транспорту, при цьому забезпечуючи час інференсу близько 238 мс для зображень розміром 900x900 пікселів. Для порівняння, двокрокові детектори, такі як Faster R-CNN, потребують для аналогічної задачі близько 2766 мс, що унеможливує їх використання в умовах аналізу відеопотоку, де можлива велика кількість динамічних зображень.

Модифікація Nano (YOLOv8n) має найменшу кількість параметрів, що дозволяє досягти оптимального балансу між швидкістю та точністю на пристроях з обмеженими обчислювальними ресурсами. Взагалі YOLO розглядає зображення як цілісне поле, на якому модель розбиває простір зображення на сітку (grid), а потім для кожної комірки прогнозує: наявність об'єкта (objectness), координати обмежувального прямокутника (bounding box), а також визначає клас пошкодження цього об'єкта (наприклад – «пожежа», «руйнування»).

Крок 2. Наступним ключовим етапом запропонованої системи є точне визначення контурів та форми пошкоджень на зонах, які були виділені на першому кроці моделлю YOLO. Локалізовані фрагменти зображення (patches), що містять знайдені об'єкти, передаються на вхід нейронної мережі U-Net. Метою цього кроку є формування точної бінарної маски пошкоджень та має назву сегментація. Перший етап забезпечує швидке виявлення приблизного місця ураження, в той час як задачею сегментації є детальний опис пошкодженої ділянки, який необхідний для подальшої класифікації ступеня руйнування.

Необхідність введення етапу сегментації пояснюється обмеженнями методу детекції. Стандартна обмежувальна рамка (bounding box) дає лише приблизне розташування об'єкта, а також включає різні типи шумів, такий як земля, рослинність, дим, туман або сусідні об'єкти. Для оцінки специфічних пошкоджень, наприклад руйнування стін, покрівлі, обвалені стіни, пробоїни чи згорілі ділянки, критично важливо знати точні межі пошкодженої ділянки. Використання архітектури U-Net дозволяє вирішити цю проблему завдяки наявності зв'язків типу «skip connections», які передають просторову інформацію з шарів енкодера безпосередньо в декодер. Це забезпечує розпізнавання деталей руйнувань, які зазвичай втрачаються в глибоких згорткових мережах. Такий підхід дозволяє

ефективно вирішувати проблему розмитих меж (blurred boundaries), типову для знімків з дрону, та точно відділяти об'єкт від складного фону. Крім того, U-Net дозволяє відновлювати високоточні маски, а також не потребує великих датасетів для навчання, що критично важливо, оскільки доступ до зображень обмежений, а дані отримуються переважно з відеопотоку дрону. Крім того U-Net легко модифікується під різні типи вхідних даних. В нашому випадку це RGB, тепловізійні ІЧ-канали, а також сенсорні дані.

U-Net складається з двох симетричних етапів. Першим етапом є Encoder, в якому зображення проходить через згорткові шари та пулінг, де виявляються особливості пошкоджень, фільтрується шум, а також формується компактне представлення зони уражень. Фактично, Encoder «бачить» структуру об'єкта та форму його пошкодження. Наступним етапом є Decoder, завдяки якому є відновлення деталізації після стиснення decoder-ом, використовуючи upsampling та злиття з шарами encoder-а. Це дозволяє зберегти дрібні деталі пошкоджень, виділити точні контури та виділити маску з високою роздільністю.

В результаті даного кроку ми отримуємо точну форму пошкодження, реальну площу руйнування, що дозволить нам підняти точність на наступному етапі класифікації, який буде використовувати готові маски, замість зашумлених зображень, які виділяє Yolo.

Крок 3. На заключному етапі, система застосовує класифікатор на базі архітектури EfficientNet для визначення категорії пошкодження, що включає класи від «непошкоджений» до «повністю зруйнований». Після сегментації пошкоджених зон за допомогою U-Net та отримавши точну маску об'єкта, у системі необхідно визначити ступінь ураження, тобто оцінити рівень пошкодження об'єкта (будівлі, інфраструктури, матеріальних конструкцій). Аналіз проводиться відповідно до стандартизованої шкали (наприклад, xBD) та дозволяє віднести ураження до одного з визначених класів, зокрема легкі пошкодження, помірні пошкодження, суттєві пошкодження, критичні руйнування.

Модель EfficientNetV2 використовується в якості потужного екстрактора ознак. Це компактна модель з оптимізованим масштабуванням глибини, ширини та роздільної здатності мережі ($\text{depth} \times \text{width} \times \text{resolution}$), що забезпечує високу точність класифікації при меншій кількості параметрів порівняно з аналогами типу ResNet або DenseNet. У гібридних архітектурах EfficientNet демонструє найкращі показники (F1-score до 84%) саме при класифікації складних шаблонів руйнувань (Рис. 7.). Важливою особливістю запропонованого методу є те, що класифікація здійснюється з урахуванням маски, отриманої на попередньому етапі. Це дозволяє зосередити модель на аналізі виключно в межах об'єкту, ігноруючи при цьому навколишній візуальний шум, що суттєво підвищує надійність оцінки. За рахунок цього вона підійде для edge-пристроїв, оскільки дає достатню точність без використання великих обчислювальних ресурсів.

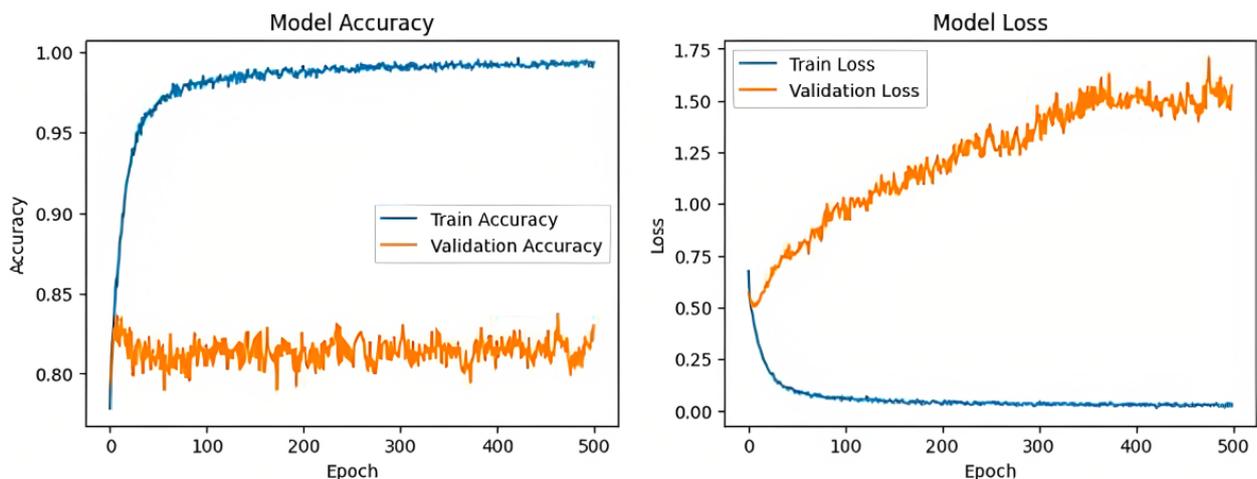


Рис. 7. Навчальні точність та втрати для моделі EfficientNet [1]

На вхід класифікатора іде маска пошкодження, яка раніше була виділена за допомогою U-Net. Вона проходить через такі етапи як «Feature extraction» в якому мережа витягує ознаки

(пошкодження, деформації); «Global representation», де модель формує узагальнене представлення патернів пошкоджень; «Fully-connected classification head», який визначає ступінь ураження.

Аналіз ефективності використання ресурсів. Реалізація описаного каскадного поділу дозволяє досягти суттєвої економії обчислювальних ресурсів пристрою, що базується на двох факторах. По-перше, легка модель YOLOv8n сканує кадр і відкидає непотрібні частини, залишаючи тільки необхідну ділянку з ураженнями. По-друге, важкі моделі сегментації та класифікації запускаються не для всього зображення, а лише для невеликих фрагментів, які раніше були вирізані YOLO (наприклад, 128x128 пікселів). Такий підхід зменшить обсяг даних, а таким чином і ресурсів, що дозволить продити швидкий аналіз територій з мінімальними затримками.

Дослідження математичних моделей детекції, сегментації та класифікації

Математична постановка задачі. Для навчання та оцінки запропонованого комбінованого методу необхідно визначити цільові функції втрат (Loss Functions), які мінімізуються під час тренування неймереж, а також метрики якості, що дозволяють кількісно оцінити ефективність детекції та сегментації руйнувань.

Нехай маємо вихідне зображення або відеокадр:

$$x \in \mathbb{R}^{H \times W \times C},$$

де H, W – висота і ширина зображення, C – кількість каналів (RGB, IR, NIR тощо).

Метою системи є обчислення функції:

$$f(x) = k,$$

де $k \in \{1, 2, 3\}$ – ступінь ураження (легке, середнє, важке).

Оскільки повна модель складається з кількох етапів обробки, це відображення формується як композиція:

$$f(x) = f_{cls}(f_{seg}(f_{det}(x))).$$

Математична модель детекції уражених зон за допомогою моделі глибокого навчання Yolo. Для навчання моделі детекції об'єктів (YOLO) критично важливою є точність регресії обмежувальної рамки (bounding box). У сучасних версіях YOLO, включаючи YOLOv4 та YOLOv8, стандартна функція втрат IoU (Intersection over Union) була вдосконалена для врахування відстані між центрами рамок та їх співвідношення сторін. Модель детекції формалізується як функція:

$$f_{det}(x) = \{(b_i, c_i)\}_{i=1}^N,$$

де $b_i = (x_1, y_1, x_2, y_2)$ – координати рамки, c_i – клас пошкодження або тип об'єкта, N – кількість знайдених зон [7].

CIoU Loss (Complete IoU Loss) – це функція яка враховує три геометричні фактори: площу перекриття, відстань між центральними точками та співвідношення сторін (aspect ratio). Вона визначається формулою:

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v,$$

де IoU – коефіцієнт перекриття між передбаченою рамкою B та істинною рамкою B^{gt} ; $\rho^2(b, b^{gt})$ – евклідова відстань між центральними точками передбаченої рамки (b) та істинної рамки (b^{gt}); c – довжина діагоналі найменшого прямокутника, що охоплює обидві рамки; α – ваговий коефіцієнт; v – параметр, що оцінює узгодженість співвідношення сторін [12].

Використання L_{CLoU} дозволяє прискорити збіжність моделі та покращити точність локалізації об'єктів, що є критичним для щільної забудови. Для новіших версій, таких як YOLOv8, також використовується **Distribution Focal Loss (DFL)**, яка дозволяє моделі адаптуватися до складних розподілів ознак, покращуючи локалізацію в умовах невизначеності. Навчання Yolo базується на мінімізації сумарної функції втрат:

$$L = L_{cls} + L_{box} + L_{obj},$$

де L_{cls} – помилка класифікації (cross entropy), L_{box} – помилка регресії координат (CIoU / DIoU), L_{obj} – помилка виявлення об'єкта (objectness) [4]. Мета оптимізації:

$$\theta_{det}^* = \arg \min_{\theta} L(\theta).$$

Математична модель сегментації уражених ділянок за допомогою моделі глибокого навчання U-Net. Сегментація є задачею відображення:

$$f_{seg}: \mathbb{R}^{H \times W \times C} \rightarrow [0,1]^{H \times W},$$

де кожен піксель отримує оцінку приналежності до ураженої області [6].

Позначимо P як прогнозовану маску, та G – як істинну маску. Для задачі сегментації пошкоджених зон (наприклад, затоплених територій або зруйнованих дахів) використовується функція втрат **Dice Loss**. Вона базується на коефіцієнті Дайса (Dice coefficient), який вимірює схожість двох множин.

$$L_{dice} = 1 - \frac{2 * | P \cap G |}{| P | + | G |}$$

де P (Prediction) – множина пікселів, передбачених моделлю як об'єкт; G (Ground Truth) – множина пікселів істинної маски; $| P \cap G |$ – кількість пікселів у перетині (спільна область) [6].

Ефективність при дисбалансі класів: У задачах оцінки руйнувань часто спостерігається значний дисбаланс класів: кількість пікселів, що відповідають «зруйнованим» ділянкам, значно менша за кількість пікселів «фону» (непошкоджена територія, вода, ліс). Стандартна функція втрат (наприклад, Cross-Entropy) може давати високу точність, просто класифікуючи все як «фон». Dice Loss вирішує цю проблему, оскільки вона фокусується безпосередньо на області перекриття цільового об'єкта, ігноруючи переважну кількість фонових пікселів. Це дозволяє моделі ефективніше навчатися на складних прикладах з малою площею руйнувань.

Математичний запис комбінованої функції втрат:

$$L = \lambda_1 L_{dice} + \lambda_2 L_{bce},$$

при

$$L_{bce} = -[G \log P + (1 - G) \log(1 - P)],$$

де L – результуюче значення комбінованої функції втрат, L_{dice} – складова функції втрат на основі коефіцієнта Дайса (відповідає за точність форми об'єкта), L_{bce} – складова бінарної перехресної ентропії (відповідає за попиксельну класифікацію), λ_1, λ_2 – вагові коефіцієнти, що визначають пріоритетність кожної складової (наприклад, для компенсації дисбалансу класів), G (Ground Truth) – істинне значення класу для конкретного пікселя (1 – пошкодження, 0 – фон), P (Prediction) – передбачена моделлю ймовірність приналежності пікселя до класу пошкоджень (значення в діапазоні $[0, 1]$) [1, 13].

Метою оптимізації є:

$$\theta_{seg}^* = \arg \min_{\theta} L(\theta).$$

Математична модель класифікації ступеня ураження за допомогою моделей EfficientNet/ResNet. На вхід класифікатор отримує зону ураження та її маску:

$$R' = x \odot M,$$

де M – маска сегментації, а \odot це поканальне множення (елементне маскування). Формалізація виглядає наступним чином:

$$f_{cls}(R') = k, k \in \{1, 2, 3\}.$$

Навчання класифікатора виконується за стандартною функцією перехресної ентропії:

$$L_{ce} = - \sum_{i=1}^c y_i \log p_i,$$

де y_i – істинна мітка (one-hot), а p_i є передбаченою ймовірністю класу [1]. Оптимізація:

$$\theta_{cls}^* = \arg \min_{\theta} L_{ce}(\theta).$$

Метрики оцінки якості. Для кількіснооцінки ефективності запропонованого методу використовуються метрики, що базуються на поняттях True Positive (TP), False Positive (FP) та False Negative (FN).

Precision (Точність) та Recall (Повнота):

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

де TP – правильно виявлені об'єкти (руйнування), FP – хибні спрацьовування (непошкоджений об'єкт визначено як зруйнований), FN – пропущені об'єкти (зруйнований об'єкт не виявлено) [4, 11].

F1-Score: Ця метрика є гармонічним середнім між точністю та повнотою. Вона є критично важливою для задач оцінки збитків, оскільки забезпечує баланс між мінімізацією хибних тривог та максимальним виявленням реальних загроз.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

mAP (mean Average Precision): mAP є стандартною метрикою для оцінки детекторів об'єктів. Вона обчислюється як середнє значення Average Precision (AP) для всіх класів (k). AP – це площа під кривою Precision-Recall (PR curve):

$$mAP = \frac{1}{k} \sum_{i=1}^k AP_i = \frac{1}{k} \sum_{i=1}^k \int_0^1 P(R) dR$$

де k – загальна кількість класів об'єктів, що розпізнаються системою, AP_i (Average Precision) – середня точність для окремого i -го класу, $P(R)$ – функція залежності точності (Precision) від повноти (Recall), що описує форму кривої Precision-Recall [4].

Фізичний зміст для оцінки руйнувань. mAP показує загальну стабільність моделі при різних порогах впевненості. Високий mAP (наприклад, > 0.9) означає, що модель стабільно знаходить більшість пошкоджених будівель з високою точністю локалізації. **F1-Score** є індикатором надійності в реальних умовах. Якщо F1 низький через низький Recall, це означає, що система пропускає багато руйнувань (ризик для рятувальників). Якщо F1 низький через низьку Precision, система генерує багато хибних тривог, витрачаючи ресурси на перевірку цілих будівель.

Висновки. У роботі здійснено узагальнення теоретичних та практичних аспектів автоматизованої оцінки руйнувань з використанням БПЛА. На основі проведеного аналізу запропоновано підхід, що дозволить зумістити важкі моделі для оцінки збитків з обмеженими обчислювальними ресурсами.

У ході дослідження було виявлено, що наразі існуючі дослідження зосереджені переважно на використанні тільки одного методу (тільки детекція або тільки сегментація) для аналізу уражених ділянок. Крім того жоден метод не може забезпечити достатньої швидкості оцінки за обмежених обчислювальних ресурсів edge-пристроїв. Також були розглянуті дослідження, які на противагу попереднім використовували гібридну архітектуру для коректної роботи пристроїв з обмеженими ресурсами. В результаті була запропонована концепція комбінованої архітектури обробки даних, яка поєднує детекцію уражених ділянок за допомогою YOLOv8, та сегментацію з U-Net. Архітектура запропонована внаслідок обмеження існуючих робіт та пропонує спосіб оптимізації під легкі пристрої.

Надалі дослідження будуть зосереджені на підготовці наборів даних для кожного з кроків, програмній реалізації та тестуванні кожного з етапів. Головними проблемами залишаються адаптація моделей під типи візуальних шумів та інтеграція різноманітних типів вхідних даних моделей, а також оптимізація моделей.

Потенційна практична цінність роботи полягає у визначенні архітектурних вимог до автономних систем швидкого реагування. Очікується, що впровадження комбінованого підходу дозволить у майбутньому створити ефективний інструмент для дистанційної оцінки збитків на територіях із високим рівнем небезпеки (зони бойових дій, заміновані ділянки). Крім того, запропонована методологія зможе бути використана як база для розробки інтелектуальної системи для моніторингу за постраждалими ділянками внаслідок різних типів надзвичайних ситуацій. Це дозволить зменшити ризики для рятувальників, а також аналізувати об'єкти, які потенційно можуть бути небезпечні для людей.

Список бібліографічного опису

1. Сонаг С., Юхандрі Ю., Таджуддін М. Гібридний підхід CNN для класифікації пошкоджень будівель після катастроф за допомогою супутникових знімків. *Journal of Applied Data Sciences*. 2025. Т. 6, № 4. С. 2824–2837. DOI: <https://doi.org/10.47738/jads.v6i4.931>.
2. Гупта Р., Хосфелт Р., Саджесв С. та ін. xBD: набір даних для оцінки пошкоджень будівель за супутниковими знімками. *arXiv preprint*. 2019. arXiv:1911.09296. DOI: <https://doi.org/10.48550/arXiv.1911.09296>.
3. Дидо Р., Собко О., Молчанова М., Мазурець О. Аналіз точності виявлення зруйнованих залишків будівель на фотографіях із використанням нейронних мереж MobileNetV3 та ViT. *Science and Technology: New Horizons of Development: Collection of Scientific Papers of the 1st International Scientific and Practical Conference (May 14–16, 2025, Prague, Czech Republic)*. Прага: International Scientific Unity, 2025. С. 208–214. DOI: <https://doi.org/10.70286/ISU-14.05.2025>.
4. Мааруф М. К. А., Бухлель М. С. Виявлення об'єктів у реальному часі за допомогою моделі YOLO-8: підхід на основі дронів. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*. 2025. Т. 16, № 1. С. 190–204. DOI: <https://doi.org/10.58346/JOWUA.2025.I1.011>.

5. Янкович Б., Джангірова С., Уллах В., Хан Л. У., Гуїзані М. Виявлення катастроф у реальному часі за допомогою БПЛА з використанням оптимізованої моделі Transformer. *arXiv preprint*. 2025. arXiv:2501.12087. DOI: <https://doi.org/10.48550/arXiv.2501.12087>.
6. Суй Ю., Шань В., Ван М., Ян Ю., Ван М. TDBOSegNet: модель семантичної сегментації типових об'єктів, що постраждали від катастроф, на зображеннях з БПЛА під час повеней. *Canadian Journal of Remote Sensing*. 2025. Т. 51, № 1. Арт. 2589555. DOI: <https://doi.org/10.1080/07038992.2025.2589555>.
7. Сілва А. С., Азеведо А. Р., Нето Ф. Х. А. М., Феррейра П. Х. Модель на основі YOLOv8 для автоматичного виявлення пошкоджень дахів житлових будинків. *Revista ALCONPAT*. 2025. Т. 15, № 1. С. 50–63. DOI: <https://doi.org/10.21041/ra.v15i1.783>.
8. Каштан В. Ю., Гнатушенко В. В. Автоматизоване виявлення пошкоджень будівель на цифрових зображеннях із використанням методів машинного навчання. *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*. 2023. № 6. С. 134–140. DOI: <https://doi.org/10.33271/nvngu/2023-6/134>.
9. Чжу Р., Лань С. Мультиmodalний метод оцінювання пошкоджень будівель із поєднанням адаптивного механізму уваги та моделей стан-простору. *Sensors*. 2026. Т. 26. Арт. 638. DOI: <https://doi.org/10.3390/s26020638>.
10. Сентгабтін К., Лілатат Н., Аумнойсомбат Р. та ін. Використання генеративного штучного інтелекту для покращення управління катастрофами: систематичний огляд. *Big Earth Data*. 2025. DOI: <https://doi.org/10.1080/20964471.2025.2521157>.
11. Салису М. Л., Гамбо Ф. Л., Муса А., Абдуллахі А. А. Удосконалена модель виявлення дронів для периферійних пристроїв із використанням дистилляції знань та байєсівської оптимізації. *Engineering Proceedings*. 2025. Т. 87. Арт. 71. DOI: <https://doi.org/10.3390/engproc2025087071>.
12. Лінь С., Мен Ю., Сунь Л., Ян С., Лен С., Лі Ю., Нью Ц., Гун В., Сяо С. Виявлення дефектів поверхні будівель на основі покращеної моделі YOLOv8. *Buildings*. 2025. Т. 15. Арт. 1865. DOI: <https://doi.org/10.3390/buildings15111865>.
13. Ван Ц., Фань С., Ці Ю., Не В., Ань Ц., Чжан С., У Х. Виявлення пошкоджень будівель шляхом злиття мультизональних знімків Sentinel-1/2 із використанням варіаційного автокодувальника та мережі MLP-Mixer: на прикладі землетрусу в Цзішишані. *Big Earth Data*. 2025. DOI: <https://doi.org/10.1080/20964471.2025.2597590>.

References

1. Sonang S., Yuhandri Y., Tajuddin M. Hybrid CNN Approach for Post-Disaster Building Damage Classification Using Satellite Imagery. *Journal of Applied Data Sciences*. 2025. Vol. 6, No. 4. P. 2824–2837. DOI: <https://doi.org/10.47738/jads.v6i4.931>.
2. Gupta R., Hosfelt R., Sajeev S. et al. xBD: A Dataset for Assessing Building Damage from Satellite Imagery. *arXiv preprint*. 2019. arXiv:1911.09296. DOI: <https://doi.org/10.48550/arXiv.1911.09296>.
3. Dydo R., Sobko O., Molchanova M., Mazurets O. Analysis of Precision of Finding the Destroyed Remains Buildings on Photos using MobileNetV3 and ViT Neural Networks. *Science and Technology: New Horizons of Development: Collection of Scientific Papers of the 1st International Scientific and Practical Conference (May 14–16, 2025, Prague, Czech Republic)*. Prague: International Scientific Unity, 2025. P. 208–214. DOI: <https://doi.org/10.70286/ISU-14.05.2025>.
4. Maarooof M. K. A., Bouhleh M. S. Real-Time Object Detection Using YOLO-8 Model: A Drone-Based Approach. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*. 2025. Vol. 16, No. 1. P. 190–204. DOI: <https://doi.org/10.58346/JOWUA.2025.II.011>.
5. Jankovic B., Jangirova S., Ullah W., Khan L. U., Guizani M. UAV-Assisted Real-Time Disaster Detection Using Optimized Transformer Model. *arXiv preprint*. 2025. arXiv:2501.12087. DOI: <https://doi.org/10.48550/arXiv.2501.12087>.
6. Sui Y., Shan W., Wang M., Yang Y., Wang M. TDBOSegNet: A Semantic Segmentation Model for Typical Disaster-Bearing Objects in UAV Images of Flood Disasters. *Canadian Journal of Remote Sensing*. 2025. Vol. 51, No. 1. Art. 2589555. DOI: <https://doi.org/10.1080/07038992.2025.2589555>.
7. Silva A. S., Azevedo A. R., Neto F. H. A. M., Ferreira P. H. YOLOv8-based model for automatic detection of residential roof damage. *Revista ALCONPAT*. 2025. Vol. 15, No. 1. P. 50–63. DOI: <https://doi.org/10.21041/ra.v15i1.783>.
8. Kashtan V. Yu., Hnatushenko V. V. Automated building damage detection on digital imagery using machine learning. *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*. 2023. No. 6. P. 134–140. DOI: <https://doi.org/10.33271/nvngu/2023-6/134>.
9. Zhu R., Lan X. Multimodal Building Damage Assessment Method Fusing Adaptive Attention Mechanism and State-Space Modeling. *Sensors*. 2026. Vol. 26. Art. 638. DOI: <https://doi.org/10.3390/s26020638>.
10. Saengtattim K., Leelawat N., Aumnoysombat R., Adelfar M., Saengwongwattana N., Suktavornprasit G., Suppasri A., Imamura F., Tang J. Harnessing generative AI for enhanced disaster management: a systematic review. *Big Earth Data*. 2025. DOI: <https://doi.org/10.1080/20964471.2025.2521157>.
11. Salisu M. L., Gambo F. L., Musa A., Abdullahi A. A. Enhanced drone detection model for edge devices using knowledge distillation and Bayesian optimization. *Engineering Proceedings*. 2025. Vol. 87. Art. 71. DOI: <https://doi.org/10.3390/engproc2025087071>.
12. Lin X., Meng Y., Sun L., Yang X., Leng C., Li Y., Niu Z., Gong W., Xiao X. Building Surface Defect Detection Based on Improved YOLOv8. *Buildings*. 2025. Vol. 15. Art. 1865. DOI: <https://doi.org/10.3390/buildings15111865>.
13. Wang J., Fan X., Qi Y., Nie W., An J., Zhang X., Wu H. Building damage detection from multi-feature fusion of Sentinel-1/2 imagery using variational autoencoder and MLP-Mixer network: insights from the Jishishan earthquake. *Big Earth Data*. 2025. DOI: <https://doi.org/10.1080/20964471.2025.2597590>.

Історія статті:

Отримано: 19.02.2026 Доопрацьовано: 26.02.2026 Прийнято до друку: 23.03.2026 Опубліковано: 29.03.2026