

DOI: <https://doi.org/10.15276/ict.03.2026.11>

УДК 004.8:004.93:624.04(477)

## Метод багаторакурсного оцінювання пошкоджень інфраструктурних об'єктів на основі мультимодального аналізу зображень

Іванов Олексій Володимирович<sup>1)</sup>

ORCID: <https://orcid.org/0000-0002-8620-974X>; [lesha.ivanoff@gmail.com](mailto:lesha.ivanoff@gmail.com). Scopus Author ID: 59135127100

Беляєв Кирило Олександрович<sup>1)</sup>

ORCID: <https://orcid.org/0009-0001-7135-3562>; [kirillbelyaev2921@gmail.com](mailto:kirillbelyaev2921@gmail.com). Scopus Author ID: 59390187700

Олійник Вадим Миколайович<sup>1)</sup>

ORCID: <https://orcid.org/0009-0000-5527-7750>; [vadimol081@gmail.com](mailto:vadimol081@gmail.com). Scopus Author ID: 59390493400

Манікаєва Ольга Сергіївна<sup>1)</sup>

ORCID: <https://orcid.org/0000-0002-0631-8883>; [manikaeva@gmail.com](mailto:manikaeva@gmail.com). Scopus Author ID: 57190488371

<sup>1)</sup> Національний університет «Одеська політехніка», пр. Шевченка, 1. Одеса, 65044, Україна

### АНОТАЦІЯ

Масштабні руйнування цивільної інфраструктури, спричинені війною в Україні, зумовили нагальну потребу в автоматизованих, надійних і масштабованих методах оцінювання пошкоджень будівельних конструкцій. У цьому дослідженні представлено порівняльну оцінку сучасних моделей «бачення-мова» (Vision-Language Models, VLM) та гібридного підходу комп'ютерного зору для класифікації пошкоджень інфраструктури за даними з кількох ракурсів. Було сформовано спеціалізований набір даних на основі наземних зображень пошкоджених війною будівель, зібраних у різних регіонах України, де кожен об'єкт інфраструктури представлений від двох до восьми фотографій, включно із зображеннями до та після пошкодження, якщо такі були доступні. Ступінь пошкодження класифіковано за п'ятьма класами: від непошкодженого до повністю зруйнованого. Запропонована методологія поєднує підхід багаторакурсного аналізу з семантичною сегментацією на основі архітектури U-Net із енкодером ResNet, а також евристичною моделлю на основі правил, яка оцінює ступінь структурних пошкоджень за піксельними метриками. Інформація з різних ракурсів агрегується для формування єдиної оцінки на рівні об'єкта, що зменшує неоднозначність, спричинену перекриттями та обмеженими кутами огляду. Запропонований підхід експериментально порівняно з кількома сучасними мультимодальними VLM, зокрема ChatGPT версії 5.5 (High Intelligence), Claude Sonnet версії 5 (Medium), Gemini версії 3.1 Pro (High) та Grok версії 4 (Fast), із використанням таких метрик, як Accuracy, Precision, Recall, F1-score, а також коефіцієнтів переоцінювання та недооцінювання. Експериментальна оцінка, проведена на спеціалізованому наборі даних, що містить зображення пошкоджених об'єктів інфраструктури, продемонструвала, що сучасні VLM перевершують базову модель сегментації за загальною точністю класифікації, водночас забезпечуючи надійне міркування у складних сценаріях структурних пошкоджень. Водночас запропонований гібридний підхід на основі сегментації забезпечує інтерпретоване оцінювання пошкоджень на рівні пікселів і залишається придатним для використання в умовах обмежених обчислювальних ресурсів. Розроблена методологія багаторакурсного оцінювання є ключовим аналітичним компонентом інтелектуальної багатофункціональної ГІС, призначеної для підтримки інвентаризації будівельних і демонтажних відходів, оптимізації логістики їх перероблення та сприяння сталому повоєнному відновленню на засадах принципів циркулярної економіки.

**Ключові слова:** Оцінювання пошкоджень інфраструктури, моделі «бачення-мова», багаторакурсний аналіз зображень, семантична сегментація, будівельні та демонтажні відходи, штучний інтелект, машинне навчання.

*Для цитування:* Іванов О. В., Беляєв К. О., Олійник В. М., Манікаєва О. С. «Метод багаторакурсного оцінювання пошкоджень інфраструктурних об'єктів на основі мультимодального аналізу зображень». *Інформатика. Культура. Техніка*. 2026; Том 3 № 1 (3): 127–143. DOI: <https://doi.org/10.15276/ict.03.2026.11>

**Актуальність.** Масштаби спричинених війною руйнувань інфраструктури в Україні призвели до безпрецедентних обсягів будівельних і демонтажних відходів (БДВ), що зумовлює потребу в оперативних автоматизованих системах їх інвентаризації, здатних замінити повільні та небезпечні ручні обстеження. У цій статті представлено інтелектуальну систему аналізу зображень, розроблену як ключовий компонент інтелектуальної багатофункціональної геоінформаційної системи (ГІС) для інвентаризації БДВ.

Для навчання та валідації системи було сформовано рідкісний, високоспеціалізований набір даних, що містить багаторакурсні зображення пошкоджених війною об'єктів інфраструктури по всій Україні – від 2 до 8 ракурсів для кожного об'єкта, включно із зображеннями до та після пошкодження. Об'єкти класифіковано за п'ятьма рівнями

пошкодження (від 0 – «Непошкоджений» до 4 – «Зруйнований») на основі уніфікованих візуальних критеріїв. Запропоновано гібридний метод класифікації, що поєднує вилучення ознак за допомогою глибокого навчання на основі архітектури U-Net з бекбоном ResNet та евристичну модель, засновану на правилах. Ця евристична модель оцінює співвідношення пошкодженого об'єму в пікселях для трьох ключових зон інтересу: покрівлі, конструктивної та фасадної цілісності, а також скління.

Ефективність запропонованого гібридного підходу U-Net–ResNet–heuristic порівнюється із сучасними моделями «бачення-мова» (Vision-Language Models, VLM) із використанням стратегії багаторакурсної агрегації, яка визначає загальний стан об'єкта на основі максимального рівня видимих пошкоджень. Оцінювання проводиться за стандартними метриками: Accuracy, Precision, Recall, F1-score, а також коефіцієнтами переоцінювання та недооцінювання.

Результати демонструють, що поєднання сегментації на основі глибокого навчання з евристичними, заснованими на правилах, забезпечує високу інтерпретованість і точність класифікації. Запропонований автоматизований конвеєр оцінювання пошкоджень є важливою основою інтелектуальної багатофункціональної ГІС-платформи, що дає змогу оптимізувати логістику перероблення будівельних і демонтажних відходів та підтримувати принципи циркулярної економіки в процесі сталого відновлення України. Водночас вимірюваність ставить перед інженерами низку компромісів: продуктивність проти енергоощадності, зручність користувача проти «важких» мультимедійних елементів, точність модельних обчислень проти їхньої обчислювальної вартості [6].

**Метою дослідження** є розробка та експериментальна оцінка методу автоматизованої класифікації ступеня пошкодження інфраструктурних об'єктів на основі аналізу багаторакурсних зображень з використанням сучасних VLM, а також гібридного методу класифікації на основі архітектури U-Net з бекбоном ResNet та евристичної моделі, заснованої на правилах як ключового компоненту інтелектуальної багатофункціональної ГІС для інвентаризації БДВ.

Для досягнення цієї мети вирішуються такі завдання:

- запропонувати концептуальну архітектурну схему інтелектуальної багатофункціональної ГІС для інвентаризації БДВ з модулем III для аналізу зображень пошкоджених інфраструктурних об'єктів;
- створити та підготувати спеціалізований набір даних об'єктів інфраструктури України, представлених наборами багатовидових зображень (від 2 до 8 фотографій на об'єкт);
- розробити логічну основу для інтелектуальної технології багаторакурсної класифікації, що включає підготовку даних, контекстний аналіз зображень з використанням мультимодальної моделі, агрегацію результатів та валідацію отриманих прогнозів;
- реалізувати метод багаторакурсної оцінки пошкоджень, який поєднує аналіз зображень, виконаний за допомогою моделей Vision-Language, з піксельними характеристиками пошкоджених структурних компонентів та алгоритмом інтеграції результатів з кількох точок зору;
- провести порівняльну експериментальну оцінку сучасних моделей Vision-Language для класифікації пошкоджень інфраструктури з використанням метрик Accuracy, Precision, Recall та F1-score, а також оцінити схильність моделей до переоцінки або недооцінки ступеня пошкодження.

Авторський науковий внесок порівняно з попередніми роботами полягає у наступному:

- розроблено багаторакурсний метод оцінювання ступеня пошкодження інфраструктурних об'єктів, який забезпечує формування єдиної об'єктної оцінки на основі сукупного аналізу кількох зображень. Запропонований метод поєднує підхід U-Net із порівняльним оцінюванням сучасних мультимодальних VLM;
- розроблено концептуальну схему багатофункціональної ГІС для інвентаризації БДВ при відбудові пошкоджених війною інфраструктурних об'єктів;
- вперше сформовано спеціалізований набір багаторакурсних зображень 108 пошкоджених інфраструктурних об'єктів, включаючи зображення об'єкта до руйнування;

– проведення порівняльного аналізу запропонованого гібридного підходу із сучасними VLM за комплексом показників точності, а також переоцінювання і недооцінювання з урахуванням Класу 0 (непошкоджені об'єкти) та із виключенням цього класу.

**Огляд літератури.** Інтеграція штучного інтелекту (ШІ), комп'ютерного зору та геопросторових технологій останнім часом стала одним із ключових напрямів у сфері сталого будівництва та відновлення після катастроф. У цьому розділі розглянуто сучасні методології оцінювання пошкоджень будівель і управління будівельними та демонтажними відходами, структуровані за трьома основними напрямками.

Глибоке навчання для оцінювання пошкоджень після катастроф і класифікації БДВ. У роботі Samal et al. [1] здійснено критичний аналіз різних методів машинного навчання, що використовуються для оцінювання, класифікації та прогнозування обсягів БДВ, із наголосом на їхній ролі у забезпеченні сталого управління відходами. На рівні супутникового дистанційного зондування знаковою стала робота Gupta et al. [2], у якій було представлено набір даних xBD, що й досі залишається одним із еталонних наборів даних для виявлення змін та оцінювання пошкоджень будівель. У роботі [3] автори запропонували двоетапний конвеєр машинного навчання для автоматизованого оцінювання пошкоджень, який поєднує семантичну сегментацію для виявлення будівель із моделями класифікації та регресії для оцінювання ступеня пошкодження. Розвиваючи ці прогностичні моделі, комп'ютерний зір почали застосовувати для класифікації відходів безпосередньо за візуальними даними. Наприклад, Molchanova et al. [4] запропонували спеціалізований метод класифікації БДВ із використанням двофакторного нейромережевого аналізу зображень. Guo et al. [5] представили новий zero-shot підхід до сегментації тріщин для моніторингу стану цивільної інфраструктури, який використовує нагляд на рівні зображень і Segment Anything Model, що дає змогу усунути потребу у трудомісткому ручному створенні піксельних анотацій. Zheng et al. [6] представили ChangeOS – глибоку об'єктно-орієнтовану систему семантичного виявлення змін, яка поєднує локалізацію об'єктів із класифікацією пошкоджень для підвищення ефективності оцінювання стану будівель за бітемпоральними супутниковими зображеннями. Duarte et al. [7] запропонували три підходи до злиття ознак convolutional neural network (CNN) із різною просторовою роздільною здатністю для підвищення ефективності оцінювання пошкоджень будівель шляхом інтеграції зображень із різних платформ і просторових масштабів, зокрема супутникових та аерофотознімків.

Новітні VLM та генеративний ШІ для візуальної діагностики. Якщо традиційні CNN добре справляються з вузькоспеціалізованими завданнями класифікації та сегментації, то стрімкий розвиток моделей «бачення-мова» (VLM) відкрив нові парадигми. Генеративний ШІ та VLM демонструють перспективні можливості завдяки здатності розуміти контекст зображень за допомогою природномовних запитів і токенизації. Estêvão [8] оцінював ефективність моделей генеративного ШІ для оцінювання пошкоджень після землетрусів та показав, як такі сучасні системи можуть токенизувати структурні характеристики для визначення рівня пошкодження та класифікації типів матеріалів за мінімального спеціалізованого навчання під конкретне завдання. Ширший потенціал таких інтелектуальних систем досліджено Alimi et al. [9]. Їхній огляд літератури розглядає можливості поєднання ШІ, машинного навчання та оброблення природної мови для оптимізації робочих процесів управління БДВ, прогнозування обсягів утворення відходів та активного виявлення можливостей для циркулярного повторного використання й перероблення матеріалів.

Оцінювання руйнувань, спричинених війною, та виклики циркулярної економіки. У дослідженні [10] запропоновано комплексну стратегію управління будівельними відходами, що утворюються внаслідок бойових дій в Україні, з акцентом на впровадження принципів циркулярної економіки та формування ринку вторинних матеріалів. У контексті війни в Україні особливо актуальним є проєкт, описаний Aouf [11], який демонструє можливість використання ШІ на периферійних пристроях (edge AI) для аналізу відеоматеріалів із дронів, отриманих над обстріляними будівлями. Aïmaiti et al. [12] оцінюють ефективність супутникових зображень середньої роздільної здатності Sentinel-1 (SAR) та Sentinel-2 (оптичних) для оперативного

оцінювання пошкоджень будівель у Київському регіоні в умовах триваючої війни. Для виявлення критично важливих структурних деталей, які залишаються невидимими на супутникових зображеннях, значного поширення набуло оцінювання пошкоджень будівель за наземними зображеннями. Вагомий внесок у цей напрям зробили Lysanets et al. [13], які запропонували архітектуру Embedding-enhanced U-Net, що використовує підхід на основі патчів, та представили перший набір даних для сегментації будівель за наземними зображеннями з воєнними пошкодженнями.

**Розробка концептуальної схеми інтелектуальної ГІС для інвентаризації БДВ.** Це дослідження розвиває концептуальну основу, запропоновану в нашій попередній роботі [14], у якій було представлено інтелектуальну інформаційну технологію управління будівельними та демонтажними відходами. У зазначеному дослідженні було закладено фундаментальну ГІС-базу знань, що інтегрує просторові та атрибутивні дані для системи картографування «Sinks–Sources» («місця накопичення – джерела»), із використанням LLM для інтерпретації візуальних ознак та оптимізації маршрутів транспортування вторинних ресурсів.

Загальну концептуальну схему робочого процесу запропонованої інтелектуальної геоінформаційної системи для інвентаризації будівельних відходів при відбудові пошкоджених війною інфраструктурних об'єктів наведено на Рис. 1.

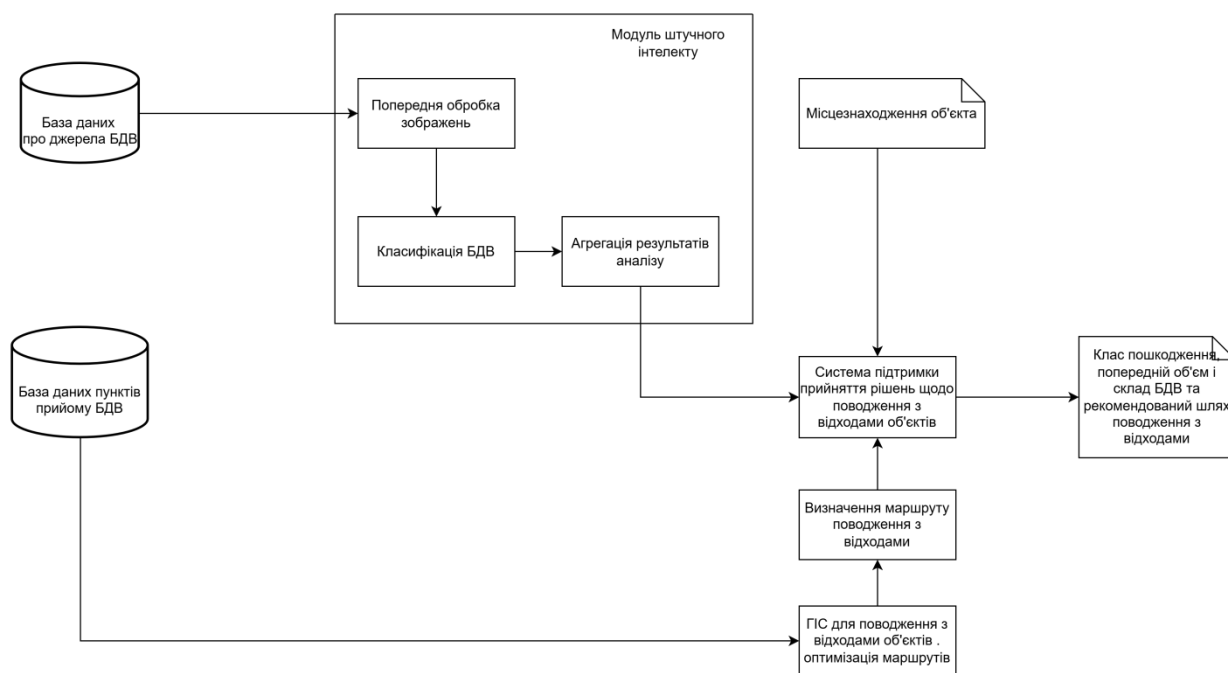


Рис. 1. Концептуальна схема інтелектуальної ГІС для інвентаризації БДВ

Джерело: розроблено авторами

Першим етапом є підготовка даних. Перший набір даних представляє собою базу даних про джерела відходів будівництва та зносу. Цей набір даних включає в себе фото інфраструктурного об'єкту до та після пошкодження, просторові координати, а також клас пошкодження об'єкта на основі попередньої оцінки. Інший набір даних представляє собою базу даних пунктів прийому відходів БДВ, яка включає в себе просторові координати об'єкта, його тип, та типи відходів, які може приймати цей пункт прийому.

Наступним компонентом схеми є модуль штучного інтелекту. На цьому етапі здійснюється автоматизований аналіз отриманих зображень для визначення змін, які відбулися з об'єктом внаслідок пошкодження або руйнування. Спочатку проводиться попередня обробка зображень, яка включає їх вирівнювання, нормалізацію та підготовку до подальшого аналізу. Після цього виконується порівняння зображень об'єкта до та після пошкодження з метою виявлення ділянок, на яких відбулися зміни. Наступним кроком є класифікація БДВ. За

допомогою методів комп'ютерного зору та мультимодальних моделей штучного інтелекту визначається склад відходів, що утворилися внаслідок руйнування. Зокрема, система повинна визначати основні типи матеріалів, такі як бетон, цегла, метал та інші будівельні матеріали. Для складних або змішаних відходів можуть використовуватися великі мультимодальні моделі, здатні аналізувати зображення та формувати їх текстовий опис. Після класифікації виконується агрегація результатів аналізу. На цьому етапі результати, отримані для окремих ділянок зруйнованого об'єкта, об'єднуються в єдину характеристику джерела БДВ. Визначається загальний склад відходів та їх орієнтовний обсяг.

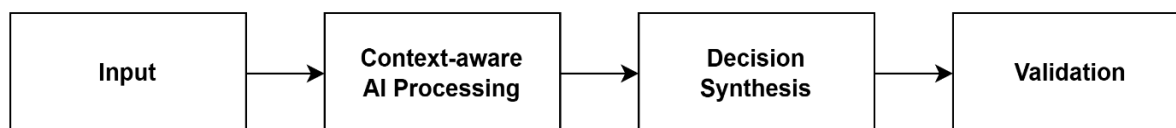
USC використовується для забезпечення просторового представлення інформації про джерела БДВ та пункти їх прийому. На карті відображаються місцезнаходження пошкоджених об'єктів, їх характеристики, склад та обсяг утворених відходів, а також розташування доступних пунктів прийому. Крім того, ГІС містить інформацію про дорожню мережу та просторові обмеження, які необхідно враховувати під час транспортування відходів. На основі просторових даних, отриманих у ГІС, визначається маршрут переміщення БДВ від джерела їх утворення до відповідного пункту прийому. При цьому враховуються відстань між об'єктами, характеристики дорожньої мережі, допустиме навантаження доріг, транспортні обмеження та початкове місце базування транспортного засобу.

Завершальним компонентом є система підтримки прийняття рішень (СППР) щодо поводження з відходами. СППР використовує результати аналізу БДВ, просторову інформацію з ГІС та розраховані маршрути для формування кінцевої рекомендації щодо поводження з відходами. Система здійснює багатокритеріальну оцінку можливих варіантів, враховуючи відповідність класу відходів можливостям пункту прийому, його потужність, відстань, час транспортування та логістичні обмеження. У результаті СППР визначає найбільш доцільний пункт прийому та рекомендований варіант транспортування БДВ.

Таким чином, сформована система забезпечує послідовний процес від отримання інформації про пошкоджений об'єкт, автоматичного визначення складу та обсягу БДВ до вибору пункту прийому і визначення оптимального маршруту їх транспортування.

### **Розробка методу баготоракурсної оцінки пошкоджень інфраструктурних об'єктів.**

Запропонований метод розроблений для автоматичного оцінювання пошкоджень об'єктів цивільної інфраструктури з використанням мультимодального штучного інтелекту. На відміну від традиційних підходів до класифікації зображень, де кожен знімок обробляється незалежно, запропонований робочий процес розглядає об'єкт інфраструктури як єдину аналітичну сутність, представлену множиною фотографій, отриманих із різних ракурсів. Така мультиракурсна стратегія знижує ймовірність некоректної оцінки, зумовленої частковою видимістю структурних руйнувань, і забезпечує більш надійну класифікацію загального технічного стану об'єкта. Загальна блок-схема робочого процесу запропонованої інтелектуальної інформаційної системи оцінювання пошкоджень будівель проілюстрована на Рис. 2.



**Рис. 2. Блок-схема робочого процесу оцінювання пошкоджень будівель на основі ШІ**

*Джерело: розроблено авторами*

Першим етапом робочого процесу є введення даних (Input). Вхідними даними для системи є пакет фотографій, що відповідають одному об'єкту інфраструктури. Кожен об'єкт описується набором із  $k$  зображень, зроблених із різних ракурсів. Пакет зображень має містити фотографії, зафіксовані до та після події руйнування, що дає змогу моделі зіставляти структурні зміни в часі.

Другим етапом є контекстно-залежна обробка засобами ШІ (Context-aware AI processing). Замість окремого аналізу кожного знімка повний пакет фотографій одночасно подається на вхід

мультимодальної моделі зору та мови (VLM) разом із проблемно-орієнтованим запитом (промптом). Промпт інструктує модель щодо необхідності сумісного аналізу всіх доступних ракурсів та визначення найвищого спостережуваного рівня структурних пошкоджень. Враховуючи просторову узгодженість між різними точками зйомки, модель здатна виявляти пошкодження, які можуть бути непомітними з одного ракурсу камери.

Третім етапом є синтез рішень (Decision synthesis). Мультимодальна модель агрегує візуальні свідчення, вилучені з усіх ракурсів, і генерує єдину мітку пошкодження, яка відображає загальний стан об'єкта інфраструктури. Цей процес відповідає принципам мультиракурсної класифікації, де остаточний прогноз формується на рівні об'єкта загалом, а не на рівні окремого зображення.

Фінальним етапом робочого процесу є валідація результатів (Validation). Спрогнозований клас пошкодження зіставляється з еталонною розміткою (ground truth), виконаною експертами-анотаторами. Ефективність моделі оцінюється за допомогою стандартних метрик класифікації, зокрема Accuracy, Precision, Recall та F1-score.

У запропонованому методі фундаментальною одиницею аналізу розглядається об'єкт інфраструктури, а не окреме зображення. Кожен об'єкт представлений сукупністю фотографій, зроблених із кількох ракурсів, що забезпечує всебічну оцінку структурних руйнувань. Таке представлення відповідає принципам мультиракурсної класифікації зображень, за якої підсумковий прогноз призначається об'єкту після спільного аналізу всіх наявних візуальних спостережень.

Формально кожен об'єкт інфраструктури описується як:

$$O_i = \{I_1, I_2, \dots, I_j, \dots, I_k\}, k \geq 2, \quad (1)$$

де  $O_i$  позначає  $i$ -й об'єкт інфраструктури;  $I_j$  – окреме зображення, зняте з відповідного ракурсу;  $k$  – кількість доступних ракурсів.

За наявності набір зображень може містити фотографії, зроблені як до, так і після події руйнування.

У цьому випадку представлення об'єкта набуває вигляду:

$$O_i = \{I_1^{before}, \dots, I_m^{before}, I_1^{after}, \dots, I_n^{after}\}, \quad (2)$$

де  $m + n = k$ . Зображення «до» надають еталонну інформацію щодо початкового вигляду споруди, тоді як зображення «після» містять свідчення можливих руйнувань. Якщо еталонні кадри недоступні, оцінювання виконується виключно на основі спостережень після події.

На відміну від звичайних датасетів для класифікації зображень, де кожне зображення є незалежним зразком, у запропонованому наборі даних знімки впорядковані у групи на рівні об'єктів. Відповідно, кожен прогноз стосується об'єкта інфраструктури в цілому, а не окремої фотографії. Така структура істотно зменшує невизначеність, зумовлену оклюзіями, обмеженими кутами огляду або локально невидимими дефектами несучих конструкцій, та забезпечує більш надійну оцінку фактичного ступеня руйнування.

**Схема класифікації даних.** Для забезпечення узгодженості анотування та оцінювання всі об'єкти інфраструктури відносяться до одного з п'яти взаємовиключних класів пошкодження. Запропонована класифікація побудована за прогресивною шкалою від неушкоджених споруд до повного знищення та ґрунтується виключно на видимих конструктивних характеристиках, що спостерігаються за всіма доступними ракурсами.

Під час розмітки остаточний клас визначається за найвищим рівнем пошкоджень, зафіксованим на будь-якому зі знімків об'єкта:

– Клас 0 – без пошкоджень (Undamaged): будівля візуально ціла, без видимих конструктивних дефектів. На фасаді відсутні тріщини, дах не пошкоджений, скління вікон ціле, а загальна геометрія споруди збережена повністю;

– Клас 1 – незначні пошкодження (Slight Damage): спостерігаються дрібні поверхневі дефекти, тоді як структурна цілісність залишається непорушеною. Типовими ознаками є розбите віконне скло при збережених рамах, незначні дефекти фасаду (наприклад, злущена штукатурка чи сліди від уламків) та обмежені пошкодження покрівлі даху без наскрізних отворів чи структурних деформацій;

– Клас 2 – помірні пошкодження (Moderate Damage): будівля має локальні структурні руйнування, проте зберігає загальну стійкість. До видимих ознак належать наскрізні отвори в зовнішніх стінах або покрівлі, повна втрата віконних чи балконних блоків, а також локальні пошкодження конструктивних елементів. Основний несучий каркас і загальна геометрія будівлі залишаються збереженими, обвали окремих кімнат чи поверхів відсутні;

– Клас 3 – важкі пошкодження (Severe Damage): наявні значні руйнування несучих конструкцій. Спостережувані характеристики включають частковий обвал стін, оголення внутрішніх приміщень, обвалення плит міжповерхового перекриття, виражену деформацію тримального каркаса та значну втрату секцій даху. Хоча окремі частини будівлі продовжують стояти, її конструктивна стійкість критично порушена;

– Клас 4 – зруйновано (Destroyed): початкова геометрія будівлі більше не розпізнається. Споруда обвалилася й перетворилася на завали; збереглися лише фрагменти цокольних стін, поодинокі конструктивні залишки або фундамент. Об'єкт вважається повністю знищеним і більше не придатний до експлуатації.

**Стратегія мультиракурсного промптингу.** На відміну від класичних класифікаторів глибокого навчання, що обробляють окремі зображення незалежно, VLM забезпечують можливість одночасного логічного міркування щодо множини зображень у межах спільного контексту. У запропонованому методі всі фотографії, що відповідають одному об'єкту інфраструктури, передаються моделі в межах єдиного промπτу. Це дозволяє моделі агрегувати візуальні ознаки з різних ракурсів і формувати єдину оцінку на рівні всього об'єкта.

Промпт інструктує модель виконувати роль незалежного експерта з будівельної інженерії та оцінювати виключно фактичні візуальні свідчення, наявні на наданих фотографіях. Оцінювання прямо виключає будь-які припущення щодо прихованих конструктивних елементів або причин руйнувань. Крім того, модель отримує вказівку аналізувати всі доступні зображення сумісно та присвоювати один остаточний клас пошкодження, що відповідає загальному стану споруди.

Текст промπτу, використаного в цьому дослідженні, наведено нижче:

*«Ти виступаєш як незалежний експерт з технічного обстеження будівель і споруд, спеціаліст з оцінки технічного стану будівельних конструкцій та аналізу наслідків руйнувань.*

*Твоє завдання – виконати об'єктивний візуальний аналіз одного об'єкта за наданими фотографіями до пошкодження та після пошкодження і визначити клас його пошкодження.*

*Для одного об'єкта може бути надано різну кількість фотографій (одна або декілька) як стану «до», так і стану «після». Використовуй усі надані фотографії одного об'єкта як єдине джерело інформації та сформулюй єдину підсумкову оцінку.*

*Оцінюй лише те, що можна безпосередньо побачити на фотографіях. Не роби припущень щодо конструктивних елементів, які не видно, та не враховуй можливу причину пошкодження. Якщо окремі елементи будівлі не проглядаються, не роби висновків про їхній стан.*

*При визначенні класу пошкодження оцінюй будівлю комплексно. Жодна окрема ознака сама по собі не визначає клас пошкодження. Наприклад, повністю вибиті вікна або значні пошкодження фасаду не означають автоматично найвищий клас пошкодження. Основним критерієм є стан несучих конструкцій, геометрії будівлі та загальної конструктивної цілісності.*

*Використовуй таку шкалу класифікації.*

*Клас 0 – Немає пошкоджень. Будівля візуально ціла. Відсутні видимі тріщини на фасаді. Вікна, покрівля та зовнішні конструкції не мають видимих пошкоджень. Геометрія будівлі повністю збережена.*

*Клас 1 – Легке пошкодження. Геометрія будівлі не порушена. Вибите лише скло у вікнах, але рами залишаються на місці. Можливі дрібні пошкодження фасаду (сліди уламків, локальне осипання штукатурки), пошкодження окремих елементів покрівлі. Наскрізні отвори та конструктивні руйнування відсутні.*

*Клас 2 – Середнє пошкодження. Наявні локальні конструктивні пошкодження. Можуть бути наскрізні пробоїни у стінах або покрівлі, вибиті цілі віконні чи балконні блоки. При цьому загальна геометрія будівлі, кути та основний каркас залишаються збереженими. Відсутній обвал поверхів або значних частин будівлі.*

*Клас 3 – Сильне пошкодження. Наявний частковий обвал конструкцій. Можуть бути обвалені окремі стіни, видно внутрішні приміщення, пошкоджені або обвалені міжповерхові перекриття, будівля має помітну деформацію або перекис, значні частини покрівлі відсутні. Однак частина конструкції все ще зберігається.*

*Клас 4 – Зруйновано. Будівля втратила свою конструктивну форму. Об'єкт являє собою переважно руїни або безформну купу будівельного сміття. Збереглися лише окремі фрагменти фундаменту або залишки нижніх частин стін.*

*Якщо ознаки відповідають двом сусіднім класам або наявної інформації недостатньо для однозначної оцінки, обері найбільш імовірний клас та відповідно змени рівень впевненості.*

*Формат відповіді:*

- 1. Візуальний аналіз.*
- 2. Основні ознаки, що вплинули на оцінку.*
- 3. Обґрунтування.*
- 4. Фінальний клас (0–4).*
- 5. Рівень впевненості (0–100 %) із коротким поясненням.»*

Структурований промпт стандартизує процес логічного висновування мультимодальної моделі та гарантує, що кожен об'єкт інфраструктури оцінюється за ідентичними критеріями. Завдяки явній інструкції моделі щодо агрегації даних із кількох ракурсів та фокусування на структурній цілісності, а не на ізольованих візуальних дефектах, запропонована стратегія промптингу знижує неоднозначність і підвищує узгодженість пооб'єктної класифікації пошкоджень. Крім того, структурований текстовий вихід спрощує подальше зіставлення результатів з експертно розміченими еталонними мітками (ground truth) та уможлиблює прямий розрахунок метрик оцінювання. Приклад поданого зображень об'єкту 722-СК-0001 та відповідей на промпт моделей генеративного штучного інтелекту ChatGPT версії 5.5 (High Intelligence), Claude Sonnet версії 5 (Medium), Gemini версії 3.1 Pro (High) та Grok версії 4 (Fast) наведений у додатку А.1 [16].

**Оцінка пошкоджень на основі аналізу сегментації за кількома ракурсами та піксельних показників.** Для семантичної сегментації пошкоджень будівель із наземних ракурсів використано глибоку згорткову нейронну мережу, що поєднує симетричний декодер U-Net із попередньо навченим кодувальником (енкодером) ResNet-34. Прямі наскрізні зв'язки (skip connections) передають високочастотні просторові ознаки, такі як конструктивні тріщини та уламки скла, безпосередньо з ранніх етапів кодування до декодера, запобігаючи втратам просторової дискретизації під час субдискретизації (downsampling).

Модель навчено на спеціалізованому наборі даних воєнних руйнувань із наземного ракурсу [15], що містить 290 високороздільних бічних зображень постраждалих будівель в Україні. Попіксельна розмітка охоплює шість семантичних класів: «Інше» (клас 0, 66,10 % пікселів), «Будівля» (клас 1, 19,42 %), «Дах» (клас 2), «Пошкодження» (клас 3), «Розбите вікно» (клас 4, 2,66 %) та «Пошкоджений дах» (клас 5, 1,92 %). Для збереження дрібних текстур зображення сегментуються методом ковзного вікна на фрагменти (патчі) розміром 640 × 640 пікселів із кроком 160. Логіти перекритих фрагментів усереднюються перед застосуванням функції Softmax для усунення крайових артефактів.

Критичним обмеженням цього набору даних є те, що розмітка обмежується виключно зображеннями стану «після руйнування». За відсутності еталонної геометрії «до руйнування»

стандартні наскрізні (end-to-end) моделі демонструють високу класифікаційну упередженість (bias) і схильні недооцінювати реальну деградацію несучих конструкцій.

Основним внеском цього дослідження є математичний фреймворк інтерпретації, розроблений для усунення цієї упередженості. Завдяки застосуванню оптимізованого мікропорогування та мультиракурсною зваженої агрегації оцінок фреймворк перетворює первинні маски сегментації на робастні класи структурної цілісності з урахуванням інженерних обмежень безпеки.

Шість базових класів агрегуються у три окремі області інтересу (Regions of Interest, ROI) для вилучення фізичних метрик площі:

– площа будівлі ( $P_{\text{building}}$ ). Сумарна кількість пікселів, що представляють наявні (активні) конструкції:

$$P_{\text{building}} = \sum \text{pixels} (\text{mask} \in [1,2,3,4,5]), \quad (3)$$

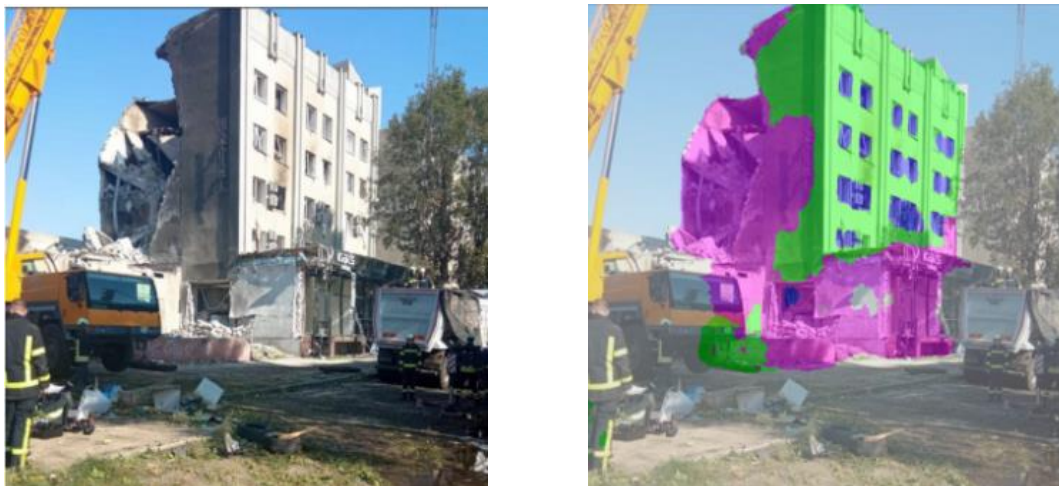
– площа структурних пошкоджень ( $P_{\text{structural\_damage}}$ ). Сукупна кількість пікселів пошкоджених стін і даху, які загрожують несній здатності:

$$P_{\text{structural\_damage}} = \sum \text{pixels} (\text{mask} \in [3,5]), \quad (4)$$

– площа пошкодження скління ( $P_{\text{window\_damage}}$ ). Кількість пікселів, що відповідають пошкодженим елементам скління:

$$P_{\text{window\_damage}} = \sum \text{pixels} (\text{mask} == 4). \quad (5)$$

Отримані сегментовані масиви пікселів формують вхідні дані для конвеєра прийняття рішень. Якісні результати роботи модифікованої моделі U-Net щодо виділення цільових меж конструкцій та дефектів наведено на Рис. 3.



**Рис. 3. Візуалізація роботи моделі сегментації: вхідне зображення фасаду в поєднанні з накладеними багатоколірними масками областей інтересу (ROI)**

*Джерело: розроблено авторами*

Для усунення масштабної варіативності, зумовленої відстанню зйомки, площі переводяться у нормовані коефіцієнти. Коефіцієнт структурних пошкоджень ( $R_{\text{str}}$ ) та коефіцієнт пошкодження скління ( $R_{\text{w}}$ ) розраховуються відносно загального видимого силуету будівлі:

$$R_{\text{str}} = \left( \frac{P_{\text{structural\_damage}}}{P_{\text{building}}} \right) \times 100 \%. \quad (6)$$

$$R_w = \left( \frac{P_{\text{window\_damage}}}{P_{\text{building}}} \right) \times 100 \%. \quad (7)$$

Для компенсації систематичного зміщення (bias) моделі, зумовленого обмеженням навчальної вибірки виключно знімками «після руйнування», впроваджено нелінійне експертне дерево рішень. Порогові значення відкалібровано відповідно до емпіричних інженерних стандартів за вузькими інтервалами деструкції конструкцій (6,0 % – 14,0 % – 35,0 % – 70,0 %):

$$\text{Class} = \begin{cases} 4, \text{ if } R_{str} > 70 \% \\ 3, \text{ if } 35 \% < R_{str} \leq 70 \% \\ 2, \text{ if } 14 \% < R_{str} \leq 35 \% \vee (R_{str} \leq 6 \% \wedge R_w > 20 \%) \\ 1, \text{ if } 6 \% < R_{str} \leq 14 \% \vee (R_{str} \leq 6 \% \wedge R_w \leq 20 \% \wedge (R_w > 6 \% \vee R_{str} > 2 \%)) \\ 0, \text{ else} \end{cases} \quad (8)$$

Оцінювання на основі одного наземного ракурсу є вразливим до помилок оклюзії (наприклад, неушкоджений передній фасад маскує обвалену задню стіну). Для мультиракурсних наборів ( $2 \leq N \leq 8$ ) реалізовано двофакторний конвеєр агрегації. Для кожного окремого ракурсу обчислюється лінійний первинний бал шляхом масштабування структурних компонентів та елементів скління з локальними інженерними вагами:

$$\text{Score}^{(i)} = 1,0 \times R_{str}^{(i)} + 0,2 \times R_w^{(i)}. \quad (9)$$

Для забезпечення балансу між локалізованими критичними руйнуваннями та загальним станом усього об'єкта підсумковий глобальний бал ( $S_{\text{object}}$ ), формулюється як опукла комбінація максимальної та середньої оцінок за ракурсами:

$$S_{\text{object}} = \alpha \times \max_{i=1 \dots N} (\text{Score}^{(i)}) + (1 - \alpha) \times \text{Mean}(\text{Score}^{(i)}), \quad (10)$$

де  $\alpha = 0,35$  математично оптимізований ваговий коефіцієнт.

Підсумковий конструктивний клас будівлі ( $\text{Class}_{\text{final}}$ ) визначається за допомогою відкаліброваних меж прийняття рішень ( $T_1 = 6,0$ ,  $T_2 = 17,0$ ,  $T_3 = 25,0$ ):

$$\text{Class}_{\text{final}} = \begin{cases} 4, \text{ if } S_{\text{object}} > 70,0 \\ 3, \text{ if } T_3 < S_{\text{object}} \leq 70,0 \\ 2, \text{ if } T_2 < S_{\text{object}} \leq T_3 \\ 1, \text{ if } T_1 < S_{\text{object}} \leq T_2 \\ 0, \text{ if } S_{\text{object}} \leq T_1 \end{cases} \quad (11)$$

Параметр  $\alpha$  та пороги прийняття рішень визначено за допомогою протоколу емпіричної оптимізації методом сіткового пошуку (Grid Search) у заданому просторі параметрів ( $\alpha \in [0,35; 0,50]$  із кроком 0,05;  $T_1, T_2, T_3 \in [1; 100]$  із кроками 1,0 та 2,0). Цільова функція була спрямована на максимізацію загальної точності багатокласової класифікації (Ассигасу) та метрики Macro-F1 на всьому наборі даних. Ця процедура калібрування гарантує оптимальний баланс між пріоритезацією локальних важких пошкоджень та придушенням ізольованого піксельного шуму сегментації.

Запропонована двофакторна модель накладає інженерне обмеження безпеки. Якщо критичне руйнування зафіксоване бодай на одному ракурсі, максимальний член через параметр  $\alpha$  спрямовує глобальний бал у відповідний клас небезпеки. Одночасно компонент середнього значення пригнічує шум від поодиноких аномалій сегментації, забезпечуючи надійність даних

перед їх передачею для ГІС-картографування та оптимізації маршрутів утилізації будівельних відходів у межах циркулярної економіки.

**Проведені експерименти та результати.** З метою оцінки ефективності запропонованої архітектури інтерпретації та агрегації даних з різних ракурсів ми зібрали та анотували спеціальний регіональний тестовий набір даних. Цей набір даних складається з об'єктів цивільної інфраструктури, що зазнали пошкоджень внаслідок військових ударів у різних адміністративних регіонах України.

Для кожної фізичної будівлі ми розробили структурований запис у базі даних, що містить як просторові атрибути, так і пари візуальних зображень. Схеми зібраних метаданих спрощено до трьох функціональних категорій:

- просторові та класифікаційні метадані: містять унікальний ідентифікатор об'єкта, геопросторові параметри (широта, довгота), адміністративне розташування (адреса, код регіону), функціональну категорію (тип об'єкта);

- пари часових візуальних ракурсів (масив багаторакурсних зображень): структуровані таким чином, щоб поєднати історичні зображення до пошкодження (Before\_Front, Before\_Back, Before\_Left, Before\_Right) з ракурсами після пошкодження (After\_Front, After\_Back, After\_Left, After\_Right, After\_Iso). Ізометрична проекція (After\_Iso) включена спеціально для забезпечення подальшої оцінки глибини фізичних куп уламків;

- адміністративні записи та записи «земної істини»: документують показники стану даних, умови ліцензування, збирачів даних та перевірений клас об'єкта, визначений експертами (Пошкодження).

Дотримуючись цієї схеми, ми зібрали загалом 108 унікальних тестових об'єктів. Розподіл міток цільової групи серед цих тестових зразків показано на Рис. 4. Розподіл зміщений у бік рівнів пошкодження від помірного до значного (на клас 3 припадає 42 зразки, або 38,9 % набору даних), що відображає реалістичні фізичні умови, за яких удари спричиняють значні пошкодження конструкції без повного руйнування. Класи незначного та повного руйнування (клас 0 та клас 4) зустрічаються рідше (відповідно 13,9 % та 11,1 %).

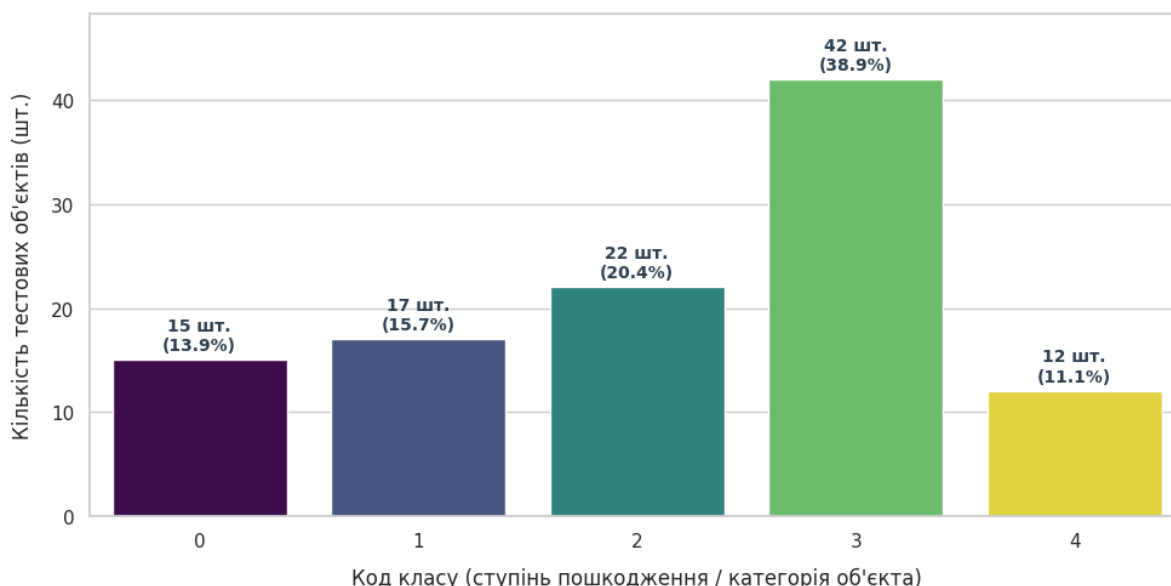


Рис. 4. Початковий розподіл експертних класів у тестовій вибірці

Джерело: розроблено авторами

Для визначення ефективності запропонованого методу оцінки пошкоджень на основі багаторакурсного аналізу було проведено експерименти з використанням однієї традиційної моделі глибокого навчання та чотирьох найсучасніших мультимодальних великих моделей (Large VLMs). Обрані моделі представляють різні архітектурні парадигми, що дозволяє провести порівняльний аналіз між підходом на основі комп'ютерного зору з навчанням під наглядом та

новітніми мультимодальними системами штучного інтелекту загального призначення. Короткий огляд обраних моделей наведено в Таблиці 1.

Для всебічного оцінювання класифікаційної спроможності моделей в умовах дисбалансу вибірки застосовано багаторівневий комплекс статистичних та інженерних метрик. Базову частку правильно визначених споруд серед усієї вибірки відображає загальна точність (Accuracy), для якої розраховано 95 % довірчий інтервал Вілсона з метою перевірки статистичної значущості відмінностей. Водночас головним референтним показником для якісного порівняння та ранжування систем визначено збалансовану метрику F1-Macro. Вона обчислюється як незважене середнє гармонійне покласових значень точності й повноти, забезпечуючи рівноцінний внесок усіх категорій технічного стану споруд та запобігаючи викривленню підсумкових оцінок через домінування мажоритарних класів над рідкісними.

**Таблиця 1. Моделі штучного інтелекту, що використовуються для класифікації пошкоджених інфраструктурних об'єктів**

| Модель                          | Категорія             | Короткий опис                                                                       |
|---------------------------------|-----------------------|-------------------------------------------------------------------------------------|
| Claude Sonnet 5 (Medium)        | Vision-Language Model | Мультимодальна модель Anthropic                                                     |
| Gemini 3.1 Pro (High)           | Vision-Language Model | Мультимодальна базова модель Google                                                 |
| Grok 4 (Fast)                   | Vision-Language Model | Мультимодальна модель xAI                                                           |
| ChatGPT-5.5 (High Intelligence) | Vision-Language Model | Мультимодальна велика мовна модель OpenAI                                           |
| Unet + ResNet                   | Deep Learning         | Спеціалізована модель комп'ютерного зору з симетричною архітектурою «кодер-декодер» |

*Джерело: розроблено авторами*

Додатково визначаються зважені показники точності (Precision (Weighted)) та повноти (Recall (Weighted)), причому зважена повнота за умов багатокласової схеми математично тотожна загальній точності. Порядкову природу п'ятирівневої шкали пошкоджень враховують метрики рангової узгодженості та фізичної відстані помилок. Квадратично зважений коефіцієнт каппа Коена (Quadratic Weighted Kappa, QWK) оцінює ступінь збігу прогнозів із експертними мітками, накладаючи квадратичний штраф за величину розбіжності між класами, через що грубі помилки класифікації караються суттєво суворіше за відхилення на сусідній рівень. Поряд із цим розраховується середня абсолютна похибка (Mean Absolute Error, MAE), яка кількісно визначає середній розмір відхилення прогнозів моделі від еталона безпосередньо в одиницях градацій шкали руйнувань.

Спрямованість і характер інженерних похибок контролюються за допомогою коефіцієнтів переоцінювання (Overestimation Ratio) та недооцінювання (Underestimation Ratio), які демонструють частку об'єктів із завищеним або заниженим ступенем деградації. Показник недооцінювання є критичним критерієм будівельної безпеки, оскільки він відображає ймовірність пропуску аварійного стану несучих конструкцій.

Загальну тенденцію алгоритмів до гіпердіагностики чи применшення масштабів руйнувань фіксує середній системний зсув (Mean Bias Error, MBE), чий знак і величина вказують на інтегральний напрямок відхилення прогнозів від фактичного стану об'єктів.

Порівняльні результати ефективності оцінених моделей наведено в Таблиці 2.

Для перевірки статистичної достовірності зафіксованої загальної точності було розраховано 95 % довірчі інтервали за методом Вілсона: для лідерів Claude та Gemini інтервал становить [0,545; 0,723], для моделі Grok – [0,452; 0,637], для ChatGPT [0,443; 0,628], а для базового пайплайну U-Net – [0,319; 0,501].

Співставлення довірчих інтервалів підтверджує, що моделі вищого ешелону (Claude та Gemini) демонструють статистично значущу перевагу над базовою U-Net, оскільки нижня межа інтервалу VLM (0,545) суворо перевищує верхню межу сегментаційної моделі (0,501). Проте

між моделями середнього рівня (Grok та ChatGPT) та базовою U-Net спостерігається виражене взаємне накладання довірчих діапазонів, що вказує на відсутність статистично ізольованої різниці між ними на поточному обсязі спостережень. Аналіз порядкових та інженерних метрик розкриває відмінності у внутрішній логіці систем: Claude продемонстрував найбільш узгоджену поведінку з найвищим показником рангової відповідності  $QWK = 0,819$ , найменшою похибкою  $MAE = 0,407$  та практично нульовим системним зсувом  $MBE = -0,037$ . Gemini показав безпекоорієнтований інженерний профіль, звівши критичні помилки недооцінювання до мінімуму (0,130). Моделі Grok та ChatGPT схильні до помірної гіперболізації ( $MBE > +0,21$ ), тоді як базова U-Net через брак навчальних прикладів непошкоджених споруд виявила високий коефіцієнт переоцінювання (0,398) та зсув  $MBE = +0,352$ , що призвело до просідання її початкового  $QWK$  до 0,456.

Таблиця 2. Підсумкова таблиця зведеної оцінки моделей

| Модель                           | Accuracy | Precision (Weighted) | Recall (Weighted) | F1-Macro | Over-estimation Ratio | Under-estimation Ratio | MBE    | MAE   | QWK   |
|----------------------------------|----------|----------------------|-------------------|----------|-----------------------|------------------------|--------|-------|-------|
| Claude Sonnet 5 (Medium)         | 0,639    | 0,674                | 0,639             | 0,639    | 0,176                 | 0,185                  | -0,037 | 0,407 | 0,819 |
| Gemini 3.1 Pro (High)            | 0,639    | 0,691                | 0,639             | 0,570    | 0,231                 | 0,130                  | +0,111 | 0,426 | 0,787 |
| Grok 4 (Fast)                    | 0,546    | 0,563                | 0,546             | 0,527    | 0,343                 | 0,111                  | +0,269 | 0,546 | 0,755 |
| ChatGPT -5.5 (High Intelligence) | 0,537    | 0,550                | 0,537             | 0,515    | 0,315                 | 0,148                  | +0,213 | 0,528 | 0,782 |
| Unet + ResNet                    | 0,407    | 0,424                | 0,407             | 0,348    | 0,398                 | 0,194                  | +0,352 | 0,852 | 0,456 |

Джерело: розроблено авторами

З метою поглибленого дослідницького аналізу та усунення методологічного викривлення (bias), зумовленого специфікою навчання базової моделі сегментації (яка тренувалася виключно на знімках категорії «після» та не містила референсних масок повністю неушкоджених споруд), було проведено повторний експериментальний інференс. У межах цього етапу з вибірки було виключено об'єкти Класу 0 (15 тестових екземплярів), що дозволило оцінити точність роботи моделей безпосередньо у сфері розпізнавання та градації реальних деструктивних модифікацій (Класи 1-4, загальна вибірка  $N = 93$ ). Отримані результати були систематизовані та розраховані за аналогічною методологією. Зведені порівняльні показники для вибірки без Класу 0 наведено в Таблиці 3.

Розрахунок довірчих інтервалів Вілсона для 4-класового експерименту показав такі значення: Gemini – [0,511; 0,706]; Claude – [0,500; 0,696]; Grok – [0,405; 0,604]; ChatGPT – [0,375; 0,574]; U-Net – [0,364; 0,563]. У 4-класовому експерименті довірчі інтервали перекриваються між усіма досліджуваними моделями, що свідчить про відсутність статистично достовірної різниці за точністю між U-Net та будь-якою з VLM.

Фільтрація тривіального для розпізнавання Класу 0 викликала закономірне зниження загальної точності всіх мовних моделей на 2,6-6,4 %. Це підтверджує, що диференціація суміжних ступенів деструкції (зокрема межі між помірними та аварійними деформаціями) є нетривіальним завданням і для складних VLM. Натомість для локального конвеєра U-Net усунення нецільового класу забезпечило суттєве підвищення збалансованого показника F1-Macro з 0,348 до 0,408. За цією метрикою U-Net зрівнялася з ChatGPT (0,406) і скоротила розрив

із Grok (0,447), а її довірчий інтервал майже повністю перекрив діапазони комерційних моделей середнього класу. До того ж коефіцієнт переоцінювання U-Net (0,333) виявився нижчим, ніж у ChatGPT (0,355) та Grok (0,366), що підтверджує стійкість її локальних просторових ознак.

**Таблиця 3. Підсумкова таблиця зведеної оцінки моделей без класу 0**

| Модель                           | Accuracy | Precision (Weighted) | Recall (Weighted) | F1-Macro | Over-estimation Ratio | Under-estimation Ratio | MBE    | MAE   | QWK   |
|----------------------------------|----------|----------------------|-------------------|----------|-----------------------|------------------------|--------|-------|-------|
| Claude Sonnet 5 (Medium)         | 0,602    | 0,631                | 0,602             | 0,573    | 0,183                 | 0,215                  | -0,075 | 0,441 | 0,666 |
| Gemini 3.1 Pro (High)            | 0,613    | 0,655                | 0,613             | 0,498    | 0,237                 | 0,151                  | +0,065 | 0,430 | 0,619 |
| Grok 4 (Fast)                    | 0,505    | 0,522                | 0,505             | 0,447    | 0,366                 | 0,129                  | +0,269 | 0,591 | 0,517 |
| ChatGPT -5.5 (High Intelligence) | 0,473    | 0,490                | 0,473             | 0,406    | 0,355                 | 0,172                  | +0,237 | 0,602 | 0,531 |
| Unet + ResNet                    | 0,462    | 0,470                | 0,462             | 0,408    | 0,333                 | 0,204                  | +0,204 | 0,720 | 0,337 |

*Джерело: розроблено авторами*

Повні матриці невідповідності використаних моделей штучного інтелекту, що використовуються для класифікації пошкоджень інфраструктури, наведені у додатку А.2 [16].

**Висновки та майбутня робота.** У межах проведеного дослідження було представлено та всебічно розглянуто концептуальну архітектуру інтелектуальної багатофункціональної геоінформаційної системи (ІБГІС), у якій модуль комп'ютерного зору виконує роль первинного аналітичного ядра для автоматизованого аудиту пошкодженої цивільної забудови.

Експериментальне співставлення мультимодальних візуально-мовних систем та спеціалізованого сегментаційного алгоритму на базі архітектури U-Net дозволило визначити їхні якісні відмінності та операційні компроміси. Передові закриті VLM виступають орієнтиром інтегральної візуальної інтерпретації, демонструючи високу якість узагальнення глобального контексту руйнувань. Разом із тим, їхній загальний рівень влучності засвідчує, що використання виключно текстових генеративних міркувань залишається недостатнім для цілковито автономного прийняття інженерних рішень без додаткової експертної верифікації. До того ж впровадження комерційних хмарних моделей у реальний контур моніторингу суттєво лімітується фінансовими витратами на регулярні API-запити, нестабільністю телекомунікаційного зв'язку в районах надзвичайних ситуацій та критичними вимогами щодо захисту просторової інформації.

З огляду на це, локалізований метод на базі модифікованої моделі U-Net становить ключову прикладну цінність для практичної реалізації системи. На відміну від мовних моделей, що повертають лише узагальнену текстову категорію, згорткова мережа продукує просторово детерміновані піксельні маски деструкцій окремих компонентів споруди (стін, покрівлі, скління). Ці маски виступають безпосередньою геометричною основою для математичного розрахунку фізичних площ, прогнозування об'ємів і маси конкретних фракцій будівельних відходів та автоматизованого планування транспортної логістики їх перероблення. Повна автономність функціонування на локальних терміналах, відсутність витрат на сторонні сервіси та гарантія конфіденційності роблять даний підхід пріоритетним для польового розгортання.

Напрями подальших досліджень передбачають масштабування бази еталонних даних за рахунок залучення збалансованого корпусу зображень непошкоджених будівель та застосування

генеративних методів синтезу архітектурних фасадів. Додаткове донавчання енкодера на розширеному масиві зразків дозволить нівелювати зафіксовану схильність до гіпердіагностики неушкоджених об'єктів і сформує високоточний, масштабований інструмент підтримки інвентаризації та відновлення інфраструктури на засадах циркулярної економіки.

**Конфлікт інтересів:** Автори декларують, що не мають конфлікту інтересів, зокрема фінансового, особистого, авторського чи будь-якого іншого характеру, який міг би вплинути на дослідження, а також на результати, опубліковані в цій статті.

**Фінансування:** Це дослідження було профінансовано Міністерством освіти і науки України в рамках науково-дослідної держбюджетної роботи «Розробка інтелектуальної багатофункціональної геоінформаційної системи для інвентаризації будівельних відходів при відбудові пошкоджених війною інфраструктурних об'єктів» (номер державної реєстрації 0126U001737), що виконується в Національному університеті «Одеська політехніка». Автори також висловлюють подяку Анастасії Барановій, Богдану Іпатову та Дмитру Паласу за їхню цінну допомогу у зборі набору зображень пошкоджень війною та проведенні експериментів з порівняльного аналізу VLM.

**Доступність даних:** Усі дані доступні в цифровій або графічній формі в основному тексті рукопису. Додаткові матеріали до статті (приклад виконання промпту у моделі ШІ та матриці еневідповідності) наведені у додатку за посиланням [16].

**Використання засобів штучного інтелекту:** Під час підготовки цієї роботи автор(и) використовував(и) інструменти генеративного штучного інтелекту ChatGPT версії 5.5 (High Intelligence), Claude Sonnet версії 5 (Medium), Gemini версії 3.1 Pro (High) та Grok версії 4 (Fast) для того, щоб: реалізувати метод багатовидової оцінки пошкоджень, який поєднує аналіз зображень, виконаний за допомогою моделей Vision-Language у розділах 4, 6 статті (дані в [16]). Крім того, автор(и) використовував(и) інструмент генеративного штучного інтелекту Gemini версії 3.5 Flash для того, щоб: покращити переклад тексту, перефразування та переформулювання. Після використання цих інструментів/сервісів автор(и) переглянув(и) та відредагував(и) контент за потреби та несе(ють) повну відповідальність за зміст публікації.

## СПИСОК ЛІТЕРАТУРИ

1. Samal, C. G., Biswal, D. R., Udgata, G. & Pradhan, S. K. “Estimation, classification, and prediction of construction and demolition waste using machine learning for sustainable waste management: A critical review”. *Construction Materials*. 2025; 5 (1): 10, <https://www.scopus.com/sourceid/21101240286>. DOI: <https://doi.org/10.3390/constrmater5010010>.
2. Gupta, R., Hosfelt, R., Sajeev, S., Patel, N., Goodman, B., Doshi, J., Heim, E., Choset, H. & Gaston, M. “Xbd: A dataset for assessing building damage from satellite imagery”. *arXiv*. 2019. DOI: <https://doi.org/10.48550/ARXIV.1911.09296>.
3. Klym, Y. “Assessment of damage to Ukrainian infrastructure using satellite imagery”. *Bachelor’s thesis, Ukrainian Catholic University, Lviv, Ukraine*. 2025. – Available from: <https://hdl.handle.net/20.500.14570/5516>. – [Accessed: 03.05.2026].
4. Molchanova, M., Didur, V., Mazurets, O., Sobko, O. & Zakharkevich, O. “Method for construction and demolition waste classification using two-factor neural network image analysis”. *Proceedings of the 2nd International Conference on Smart Automation & Robotics for Future Industry*. Lviv, Ukraine. 2025; 3970: 168–182, <https://www.scopus.com/sourceid/21100218356> – Available from: <https://ceur-ws.org/Vol-3970/PAPER14.pdf> – [Accessed: 06.05.2026].
5. Guo, S., Cockayne, M. J., Ortolani, M., Al-Hamd, R. & Al-Bander, B. “Zero-Shot crack segmentation for structural health monitoring using class activation map guided segment anything model”. *SSRN*. 2025. DOI: <https://doi.org/10.2139/ssrn.5197086>.
6. Zheng, Z., Zhong, Y., Wang, J., Ma, A. & Zhang, L. “Building damage assessment for rapid disaster response with a deep object-based semantic change detection framework: From natural disasters to man-made disasters”. *Remote Sensing of Environment*. 2021; 265. DOI: <https://doi.org/10.1016/j.rse.2021.112636>.
7. Duarte, D., Nex, F., Kerle, N. & Vosselman, G. “Multi-resolution feature fusion for image classification of building damages with convolutional neural networks”. *Remote Sensing*. 2018; 10, <https://www.scopus.com/sourceid/12503>. DOI: <https://doi.org/10.3390/rs10101636>.
8. Estêvão, J. M. C. “Effectiveness of Generative AI for post-earthquake damage assessment”. *Buildings*. 2024; 14 (10): 3255, <https://www.scopus.com/sourceid/26980>. DOI: <https://doi.org/10.3390/buildings14103255>.
9. Alimi, K., Jin, R., Nguyen, B. N., Nguyen, Q., Chen, W. & Hosking, L. “Exploring artificial intelligence applications in construction and demolition waste management: a review of existing

literature”. *Journal of Science and Transport Technology*. 2025; 5 (1): 104–136, <https://www.scopus.com/sourceid/21101274771>. DOI: <https://doi.org/10.58845/jstt.utt.2025.en.5.1.104-136>.

10. Dadiverina, L., Laukhin, D., Hryhoriev, O., Khalymendyk, D. & Bondarenko, V. “Management of construction waste generated by military actions: analysis and solutions”. *Collection of Research Papers of the National Mining University*. 2024; 79: 163–174. DOI: <https://doi.org/10.33271/crpnmu/79.163>.

11. Aouf, R. S. “Circularity on the Edge AI finds reusable materials in rubble of Ukrainian buildings”. *Dezeen*. 2025. – Available from: <https://www.dezeen.com/2025/07/07/circularity-on-the-edge-ai-ukraine>. – [Accessed: 10.05.2026].

12. Aimaiti, Y., Sanon, C., Koch, M., Baise, L. G. & Moaveni, B. “War related building damage assessment in Kyiv, Ukraine, using Sentinel-1 Radar and Sentinel-2 optical images”. *Remote Sensing*. 2022; 14 (24), <https://www.scopus.com/sourceid/86430> DOI: <https://doi.org/10.3390/rs14246239>.

13. Lysanets, A., Kosukha, O., Panchenko, T. & Tereshchenko, Y. “Embedding-Enhanced U-Net with a patch-based approach and a novel dataset for ground-level building damage segmentation”. *Applied Artificial Intelligence*. 2026; 40 (1), <https://www.scopus.com/sourceid/23673>. DOI: <https://doi.org/10.1080/08839514.2026.2685357>.

14. Arsirii, O., Ivanov, O., Bieliaiev, K. & Cudecka-Purina, N. “Intelligent information technology for inventory and utilization of construction and demolition waste from damaged infrastructure facilities”. *Proceedings of the 13-th International Conference on Information Control Systems & Technologies*. Odesa, Ukraine. 2025; 4048: 276–289, <https://www.scopus.com/sourceid/21100218356>. DOI: <https://doi.org/10.55056/ceur-ws.org/Vol-4048/paper22.pdf>.

15. Lysanets, A. “Ukrainian building damage assessment system: Ground-level Ukrainian building damage segmentation dataset”. *Kaggle*. 2025. – Available from: <https://www.kaggle.com/datasets/andrewlysanets/ukrainian-damaged-buildings/data>. – [Accessed: 22.06.2026].

16. Іванов, О., Беляєв, К., Олійник, В. & Манікаєва, О. “Додаток А: додаткові матеріали до статті Метод багаторакурсного оцінювання пошкоджень інфраструктурних об'єктів на основі мультимодального аналізу зображень”. *Zenodo*. 2026. DOI: <https://doi.org/10.5281/zenodo.22181084>.

Отримано 20.07.2026

Отримано після перегляду . 28.08.2026

Прийнято 03.09.2026

DOI: <https://doi.org/10.15276/ict.03.2026.11>

UDC 004.8:004.93:624.04(477)

## Multi-view infrastructure damage assessment method based on multimodal image analysis

**Oleksii V. Ivanov<sup>1)</sup>**

ORCID: <https://orcid.org/0000-0002-8620-974X>; [lesha.ivanoff@gmail.com](mailto:lesha.ivanoff@gmail.com). Scopus Author ID: 59135127100

**Kyrylo O. Bieliaiev<sup>1)</sup>**

ORCID: <https://orcid.org/0009-0001-7135-3562>; [kirillbelyaev2921@gmail.com](mailto:kirillbelyaev2921@gmail.com). Scopus Author ID: 59390187700

**Vadym M. Oliinyk<sup>1)</sup>**

ORCID: <https://orcid.org/0009-0000-5527-7750>; [vadimol081@gmail.com](mailto:vadimol081@gmail.com). Scopus Author ID: 59390493400

**Olga S. Manikaeva<sup>1)</sup>**

ORCID: <https://orcid.org/0000-0002-0631-8883>; [manikaeva@gmail.com](mailto:manikaeva@gmail.com). Scopus Author ID: 57190488371

<sup>1)</sup> Odesa Polytechnic National University, Shevchenko Ave. 1. Odesa, 65044, Ukraine

## ABSTRACT

Massive destruction of civil infrastructure caused by the war in Ukraine has created an urgent need for automated, reliable, and scalable methods for assessing structural damage to buildings. This study presents a comparative evaluation of state-of-the-art Vision-Language Models (VLM) and a hybrid computer vision approach for classifying infrastructure damage using multi-view data. A specialized dataset was curated based on ground-level imagery of war-damaged buildings collected across various regions of Ukraine, where each infrastructure object is represented by two to eight photographs, including pre- and post-damage images when available. The degree of damage is classified into five categories ranging from undamaged to completely destroyed. The proposed methodology combines a multi-view analysis approach with semantic segmentation based on a U-Net architecture with a ResNet encoder, alongside a rule-based heuristic model that estimates the degree of structural damage using pixel-level metrics. Information from different viewpoints is aggregated to form a unified object-level assessment, which reduces ambiguity caused by occlusions and limited viewing angles. The proposed approach is experimentally compared with several modern multimodal VLM, including ChatGPT version 5.5 (High Intelligence), Claude Sonnet version 5 (Medium), Gemini version 3.1 Pro (High), and Grok version 4 (Fast), using metrics such as Accuracy, Precision, Recall, F1-score, as well as overestimation and underestimation coefficients. The experimental evaluation, conducted on a specialized dataset containing images of damaged infrastructure objects, that modern VLMs outperform the baseline segmentation model in overall classification accuracy while providing robust reasoning in complex structural damage scenarios. Meanwhile, the proposed segmentation-based hybrid approach provides interpretable pixel-level damage assessment and remains suitable for deployment in resource-constrained environments. The developed multi-view assessment methodology serves as a key analytical component of an intelligent multi-functional GIS designed to support the inventory of construction and demolition waste, optimize its recycling logistics, and promote sustainable post-war recovery based on circular economy principles.

**Keywords:** Infrastructure damage assessment, Vision-Language Models, multi-view image analysis, semantic segmentation, construction and demolition waste, artificial intelligence, machine learning.

## ІНФОРМАЦІЯ ПРО АВТОРІВ



**Іванов Олексій Володимирович** - доктор філософії, доцент кафедри Інформаційних систем. Національний університет «Одеська політехніка». пр. Шевченка, 1. Одеса, 65044, Україна

**Галузь досліджень:** Геоінформаційні інформаційні системи; екологічний ризик; нечітка логіка; промислова безпека; системи підтримки прийняття рішень, системи управління базами даних

**Oleksii V. Ivanov** – PhD, Associate Professor, Information Systems Department. Odesa Polytechnic National University, 1, Shevchenko Ave. Odesa, 65044, Ukraine

ORCID: <https://orcid.org/0000-0002-8620-974X>; [lesha.ivanoff@gmail.com](mailto:lesha.ivanoff@gmail.com). Scopus Author ID: 59135127100

**Research field:** Geoinformation information systems; environmental risk; fuzzy logic; industrial safety; decision support systems, database management systems



**Бєляєв Кирило Олександрович** - аспірант кафедри Інформаційних систем. Національний університет «Одеська політехніка». пр. Шевченка, 1. Одеса, 65044, Україна

**Галузь досліджень:** Геоінформаційні інформаційні системи; системи підтримки прийняття рішень; машинне навчання; нейронні мережі

**Kyrylo O. Bieliaiev** - Postgraduate student of the Information Systems Department. Odesa Polytechnic National University, 1, Shevchenko Ave. Odesa, 65044, Ukraine

ORCID: <https://orcid.org/0009-0001-7135-3562>; [kirillbeliyaev2921@gmail.com](mailto:kirillbeliyaev2921@gmail.com). Scopus Author ID: 59390187700

**Research field:** Geographic information systems; decision support systems; machine learning; neural networks



**Олійник Вадим Миколайович** - аспірант кафедри Інформаційних систем. Національний університет «Одеська політехніка». пр. Шевченка, 1. Одеса, 65044, Україна

**Галузь досліджень:** Геоінформаційні інформаційні системи; системи підтримки прийняття рішень; машинне навчання; нейронні мережі

**Vadym M. Oliinyk** - Postgraduate student, Information Systems Department. Odesa Polytechnic National University, 1, Shevchenko Ave. Odesa, 65044, Ukraine

ORCID: <https://orcid.org/0009-0000-5527-7750>; [vadimol081@gmail.com](mailto:vadimol081@gmail.com). Scopus Author ID: 59390493400

**Research field:** Geographic information systems; decision support systems; machine learning; neural networks



**Манікаєва Ольга Сергіївна** - кандидат технічних наук, доцент кафедри Інформаційних систем. Національний університет «Одеська політехніка». пр. Шевченка, 1. Одеса, 65044, Україна

**Галузь досліджень:** Методи та системи штучного інтелекту; нейронні мережі

**Olga S. Manikaeva** - PhD, Associate Professor, Information Systems Department. Odesa Polytechnic National University, 1, Shevchenko Ave. Odesa, 65044, Ukraine

ORCID: <https://orcid.org/0000-0002-0631-8883>; [manikaeva@gmail.com](mailto:manikaeva@gmail.com). Scopus Author ID: 57190488371

**Research field:** Artificial intelligence methods and systems, neural networks