

АВТОМАТИЗОВАНА ІДЕНТИФІКАЦІЯ ДАХОВИХ СОНЯЧНИХ ЕЛЕКТРОСТАНЦІЙ НА ОСНОВІ ВЕЛИКИХ МУЛЬТИМОДАЛЬНИХ МОДЕЛЕЙ В УМОВАХ ОБМЕЖЕНОЇ ВІЗУАЛЬНОЇ ІНФОРМАЦІЇ

Отримано 22 лют. 2026 р.; рекомендовано до публікації 23 бер. 2026 р.
Доступно онлайн 31 бер. 2026 р.

Шаповалова С. І.¹, Матях С. В.², Головакін М. А.³

Автор для кореспонденції: Матях Сергій,
e-mail: krypto@ukr.net

¹ канд. техн. наук, доцент
<https://orcid.org/0000-0002-3431-5639>

² канд. техн. наук
<https://orcid.org/0000-0002-1707-3519>

³ аспірант
<https://orcid.org/0009-0006-1456-8926>

^{1,3} Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського», м. Київ, Україна

² Інститут відновлюваної енергетики НАН
України, м. Київ, Україна

Анотація. Запропоновано метод визначення дахових об'єктів сонячної генерації на супутникових та аерофотознімках за умов деградації їх просторової роздільної здатності. Виявлення об'єкта базується на розв'язанні задачі класифікації на зображеннях великою мовною моделлю для зменшення обчислювального навантаження та пришвидшення процесу ідентифікації. Запропонований метод призначено для оперативної інвентаризації об'єктів сонячної генерації з метою підвищення стійкості енергетичної інфраструктури України. На відміну від підходів семантичної сегментації, що потребують спеціалізованого навчання моделей і значних обчислювальних ресурсів, запропонований підхід з бінарною класифікацією локальних фрагментів зображення LLM не потребує додаткового навчання та дає змогу зменшити сумарні обчислювальні витрати за рахунок переходу від ресурсоемної задачі семантичної сегментації до задачі бінарної класифікації локальних фрагментів зображення. Запропоновано чотири промпт-стратегії виявлення заданого об'єкта на зображеннях. Розроблено алгоритм моделювання обмеженої просторової роздільної здатності знімків, що базується на контрольованому масштабуванні та інтерполяції локальних фрагментів зображення. Проведено обчислювальні експерименти на великій мовній моделі GPT-4o з порівняння стратегій формування критеріїв ідентифікації об'єктів сонячної генерації за різної просторової роздільної здатності. Встановлено закономірності впливу деградації просторової інформації в діапазоні 1–2 м/піксель на ефективність автоматизованої ідентифікації об'єктів. Підтверджено доцільність експериментального вибору логіки ідентифікації до умов просторової деградації зображення. Встановлено, що за високої деталізації, 0.1–0.3 м/піксель, найвищу точність забезпечує стратегія «бінарна класифікація на прикладах зображень» ($F1\text{-score} = 0.9523$), тоді як у разі збільшення масштабу пікселя, 1–2 м/піксель, стійкішими виявляються «класифікація на основі покрокового аналізу ознак» ($F1\text{-score} = 0.6801$) та «класифікація на основі гіпотез» ($F1\text{-score} = 0.6502$). Запропонований метод застосування мультимодальних мовних моделей надає можливість масштабованої автоматизованої інвентаризації об'єктів розподіленої сонячної генерації на великих територіях.

Ключові слова: аерофотознімок, супутниковий знімок, сонячні панелі, визначення об'єктів на зображеннях, LLM, GPT, просторове розділення аерофотознімків, енергетична інфраструктура.

Перелік використаних позначень та скорочень

ДЗЗ – дистанційне зондування Землі

СЕС – сонячна електростанція

CEPA (Center for European Policy Analysis) – Центр аналізу європейської політики

WMS (Web Map Service) – Вебсервіс карт

OGC (Open Geospatial Consortium) – Відкритий геопросторовий консорціум

CRS (Coordinate Reference System) – Система координат

EPSG – код системи координат (у роботі використано EPSG:25832)

GeoTIFF – геоприв'язаний растровий формат зображення
GetMap – тип запиту сервісу WMS для отримання растрового фрагмента карти

LLM (Large Language Model) – Велика мовна модель
API (Application Programming Interface) – прикладний програмний інтерфейс

Вступ

Актуальність автоматизованої ідентифікації дахових об'єктів сонячної генерації суттєво зросла внаслідок масштабних руйнувань енергетичної інфраструктури

України, спричинених воєнними діями. У таких умовах оперативна інвентаризація об'єктів генерації на великих територіях є критично важливою для оцінювання потенціалу відновлюваних джерел енергії та планування

екстрених дій із забезпечення стійкості енергетичної інфраструктури. За даними аналітичного центру Center for European Policy Analysis (CEPA) [1], унаслідок ракетних та дронних атак у 2022–2024 роках Україна зазнала суттєвих втрат електрогенерувальних потужностей. Зокрема, у 2024 році було втрачено близько 9 ГВт генерації, що відповідає приблизно одній третині довоєнного рівня споживання електроенергії, а частка теплової генерації скоротилась з 23.5 % до близько 5 %. У цих умовах особливого значення набувають об'єкти децентралізованої відновлюваної генерації, зокрема дахові об'єкти сонячної генерації, які здатні підвищувати енергетичну стійкість локальних споживачів та зменшувати навантаження на пошкоджену інфраструктуру. Проте відсутність актуальних реєстрів та оперативних інструментів інвентаризації таких об'єктів істотно ускладнює їх урахування в енергетичному плануванні. Традиційні підходи до аналізу аерофотознімків [2–4], зокрема методи семантичної сегментації [5], характеризуються високими обчислювальними витратами та залежать від спеціалізованих вибірок, що обмежує їх застосування в умовах різної якості зйомки та обмеженої просторової роздільної здатності. Тому актуальним є створення методів швидкої ідентифікації дахових об'єктів сонячної генерації, здатних працювати в умовах обмеженої візуальної інформації та масштабуватись на великі досліджувані площі без попереднього навчання.

Задача автоматизованої ідентифікації дахових сонячних панелей на аерофотознімках у більшості сучасних досліджень розв'язується із застосуванням згорткових нейронних мереж [6]. Основний акцент у таких роботах робиться на підвищенні точності розпізнавання шляхом удосконалення архітектур моделей та використання багаторівневих візуальних ознак.

У дослідженні [7] розглядається задача автоматичного виявлення сонячних панелей на основі семантичної сегментації аерофотознімків. Дослідники використовують згорткову глибинну нейронну мережу, навчену на спеціалізованому наборі розмічених даних, та демонструють високі показники точності ідентифікації сонячних панелей на зображенні високої якості. Досягнуті результати підтверджують ефективність використання згорткових глибинних нейронних мереж для цієї задачі, однак підхід потребує попереднього навчання моделі, значних обчислювальних ресурсів та високої якості зображення.

Інший приклад представлено в роботі [8], де автори зосереджуються на підвищенні точності сегментації сонячних панелей. Для цього використовується ускладнена архітектура згорткової нейронної мережі з багатомасштабним аналізом ознак і додатковими модулями просторової реконструкції. Цей підхід забезпечує покращення якості розпізнавання порівняно з базовими моделями, проте досягнення таких результатів супроводжується збільшенням складності моделі та обчислювальних витрат.

Сучасні дослідження з автоматизації розпізнавання сонячних панелей переважно спрямовані на підвищення точності шляхом ускладнення архітектури згорткових нейронних мереж, використання додаткових рівнів обробки ознак та спрямовані на обробку чітких знімків. Хоча такі підходи демонструють високі показники якості, вони залишаються ресурсоемними та потребують спеціалізованого навчання моделей, що ускладнює їх застосування для оперативної обробки великих масивів аерофотознімків або роботи в умовах обмеженої просторової роздільної здатності.

Проблема полягає в необхідності великої кількості обчислювальних ресурсів, насамперед часу, для точного визначення заданого об'єкта на великій множині знімків методами семантичної сегментації мережами глибокого навчання. Тому необхідно мати засіб швидкого нересурсоемного попереднього визначення підмножини знімків, на яких з великою ймовірністю зображений заданий об'єкт.

Запропонований у цій роботі підхід спрямовано на зменшення використання обчислювальних ресурсів за рахунок відмови від спеціалізованого навчання моделей і перенесення акценту з ускладнення архітектури на формалізацію критеріїв прийняття рішень при аналізі зображень. Для автоматизованої класифікації об'єктів на аерофотознімках застосовуються великі мультимодальні моделі [9–12]. На відміну від класичних методів комп'ютерного зору [3, 4], де логіка класифікації закладена в параметрах навченої моделі, у разі використання великих мовних моделей результат аналізу значною мірою залежить від способу задання критеріїв інтерпретації візуальної інформації. Швидкий та попередній аналіз без етапу спеціалізованого навчання на основі LLM множини аерофотознімків дасть змогу створити обмежену множину зображень, які з високою ймовірністю містять сонячні панелі, і суттєво скоротити ресурси для їх подальшої семантичної сегментації.

Метою роботи є створення методу визначення об'єктів сонячної генерації на супутникових та аерофотознімках за умов деградації їх просторової роздільної здатності.

Для досягнення поставленої мети у роботі необхідно виконати такі завдання:

1. Формалізувати задачу ідентифікації об'єктів сонячної генерації на аерофотознімках для використання мультимодальної моделі (LLM).
2. Розробити промпт-стратегії для LLM.
3. Підготувати набір зображень для обчислювальних експериментів.
4. Розробити алгоритм моделювання умов обмеженої просторової роздільної здатності, для дослідження впливу деградації візуальної інформації на результати ідентифікації.
4. Провести порівняльний аналіз результатів обчислювальних експериментів з визначення сонячних

панелей за різними промпт-стратегіями та умовами роздільної здатності.

Об'єктом дослідження є методи та засоби розв'язання задачі бінарної класифікації об'єктів на аерофотознімках.

Предметом дослідження є застосування мультимодальних мовних моделей для класифікації об'єктів на локальних фрагментах зображень за різної просторової роздільної здатності.

Методи дослідження: навчені великі мовні моделі, GPT-4o, стратегії промпт-орієнтованої ідентифікації на зображенні, моделювання різних рівнів деталізації аерофотознімку.

1. Задача визначення наявності сонячної панелі на аерофотознімку на основі LLM

Задача визначення об'єктів сонячної генерації на аерофотознімку ставиться як задача бінарної класифікації на зображенні.

Для коректного аналізу зображення, отримані в результаті аерофотозйомки, мають бути попередньо підготовані. Тому визначення сонячних панелей відбувається за два етапи:

1. Просторова декомпозиція.

Область інтересу на зображенні розбивається на множину локальних фрагментів фіксованого розміру, що забезпечує локалізацію аналізу та уніфікацію вхідних даних для подальшої обробки.

2. Класифікація сонячної панелі або її відсутності на локальних фрагментах аерофотознімка.

Для кожного фрагмента зображення на основі LLM виконується перевірка наявності об'єктів сонячної генерації.

Для безпосередньої класифікації перед представленням кожного зображення LLM пропонується промпт-контекст з описом характеристик об'єктів сонячної генерації та завданням. LLM надає оцінку наявності сонячної панелі за шкалою від 0 до 1. Вважається, що знімок містить сонячну панель, якщо оцінка в діапазоні від 0.6 до 1.

Задача визначення об'єкта на знімках ускладнюється тим, що вхідні зображення можуть суттєво відрізнятися:

1. Просторовою роздільною здатністю – розміром території, що припадає на 1 піксель.
2. Якістю – наявністю втрат фрагментів зображення або нечітких сегментів.
3. Схожістю з іншими об'єктами, оскільки характерні візуальні ознаки об'єктів сонячної генерації можуть бути частково / повністю втрачені або мати спільні риси з іншими об'єктами.

За таких обставин безпосередня класифікація зображення без явного опису ознак призводить до нестабільних результатів і зростання кількості хибних рішень.

Під час опису візуальних ознак сонячних панелей необхідно враховувати, що зображення об'єктів сонячної

генерації можуть мати спільні риси з такими об'єктами, як сільськогосподарські теплиці зі світлопрозорим покриттям, металеві або світловідбивні покрівлі, дахові вікна, інженерні надбудови й тіні від конструктивних елементів. Наявність подібних об'єктів у межах щільної міської забудови істотно ускладнює автоматизовану інтерпретацію знімків та підвищує ризик хибнопозитивної класифікації.

Використання LLM дає змогу уникнути повного аналізу всього зображення, обмежуючи обчислення лише тими локальними фрагментами, для яких бінарна класифікація моделі виявляє наявність об'єктів сонячної генерації з достатньою високою оцінкою. Завдяки цьому задачу з ресурсоємної семантичної сегментації, що потребує спеціалізованого навчання моделей, можна трансформувати в задачу локальної бінарної класифікації, що сприяє зменшенню сумарних обчислювальних витрат алгоритмічного циклу та забезпечує масштабованість алгоритму при обробці супутникових знімків великих територій.

2. Промпт-стратегії для LLM

Ефективність автоматизованої ідентифікації об'єктів сонячної генерації безпосередньо залежить від способу формалізації критеріїв аналізу візуальної інформації.

Для зов'язання задачі було використано чотири підходи, які відрізняються логікою виявлення заданого об'єкта:

1. Класифікація на основі базової інтерпретації ознак.
2. Бінарна класифікація на прикладах зображень.
3. Класифікація на основі покрокового аналізу ознак.
4. Класифікація на основі гіпотез.

Класифікація на основі базової інтерпретації ознак.

При цьому підході для LLM визначалась роль експерта. Для виявлення сонячної панелі надавався опис її характеристик ознак: темна прямокутна структура, регулярність форми, контраст з навколишнім фоном.

Такий підхід фактично імітує інтуїтивну інтерпретацію зображення без деталізованої перевірки окремих характеристик. За високої просторової роздільної здатності цей спосіб забезпечує прийнятну точність. Однак у разі зниження деталізації сцени зростає кількість хибнопозитивних рішень, зокрема через подібність з темними покрівлями або тінями.

Бінарна класифікація на прикладах зображень.

При цьому підході для LLM визначалась роль «You are a specialized AI for detecting solar arrays in satellite imagery». LLM надавалось два приклади зображень сонячних панелей, кожен з яких є типовим для класу наявності та відсутності. Промпт-текст містив визначення заключення щодо об'єкта для кожного прикладу у форматі спілкування з ШІ-експертом.

У цьому підході використано уточнення критеріїв шляхом використання репрезентативних прикладів об'єктів цільового класу та типових граничних випадків. Рішення

приймалося на основі оцінювання подібності аналізованого фрагмента до попередньо визначених зразків.

На зображеннях високої якості цей підхід дає змогу суттєво підвищити точність класифікації, оскільки структурні елементи сонячних панелей добре зберігаються. Проте зі зниженням просторової роздільної здатності ефективність методу зменшується, оскільки дрібні текстурні деталі, що визначають подібність до еталону, частково або повністю втрачаються.

Класифікація на основі покровового аналізу ознак. При цьому підході для LLM визначалась роль експерта-аналітика. Визначення об'єкта відбувається за три етапи:

1. Виокремлення найбільш значущих характеристик ознак об'єкта: Luminance, Uniformity, Edges.
2. Дедуктивне виведення класу на основі комбінацій значень ознак.
3. Перетворення вердикту на оцінку від 0 до 1.

Таким чином, на першому етапі оцінювалася відносна яскравість об'єкта та його геометрична форма. На другому етапі перевірялась наявність просторового групування елементів.

Такий підхід дає змогу зменшити залежність від дрібних текстурних ознак і перейти до оцінювання більш загальних характеристик, що зберігаються навіть за помірного зниження просторової роздільної здатності. Практичні результати показали підвищення стабільності класифікації для зображень середньої якості.

Класифікація на основі гіпотез. При цьому підході для LLM визначалась роль експерта-дослідника: «You are a Lead Scientist in Remote Sensing». На відміну від покровового аналізу ознак, процес класифікації організовано як паралельну перевірку кількох альтернативних гіпотез щодо природи об'єкта на даху. За цим підходом визначення відбувається в чотири етапи:

1. Формування та оцінювання гіпотези А (світлий відбивний об'єкт / теплиця). Аналізуються ознаки високої яскравості (High Luminance), наявність відблисків та внутрішньої структури.
2. Формування та оцінювання гіпотези В (ізолюваний малий об'єкт / мансардне вікно). Перевіряється ізолюваність об'єкта та його відносно малий розмір порівняно з площею даху.
3. Формування та оцінювання гіпотези С (сонячні панелі). Аналізуються ознаки нижчої яскравості відносно покрівлі (Low Luminance), а також наявність просторового групування модулів (Clustering) або регулярної структури.
4. Синтез результатів. Порівнюються відносні оцінки гіпотез. Клас «сонячні панелі» підтверджується лише за умови, що гіпотеза С є більш обґрунтованою, ніж А та В. У разі переваги гіпотез А або В рішення приймається на користь відсутності сонячних панелей. Такий підхід продемонстрував найвищу стійкість до зниження просторової роздільної здатності, ос-

кільки базується не на одній окремій ознаці, а на зіставленні альтернативних пояснень. Водночас у випадках темних однорідних покрівель можливе зростання хибнопозитивних результатів через переоцінювання гіпотези С.

Запропоновані промпт-тексти для виявлення сонячних панелей на аерофото знімках за кожною описаною стратегією представлено в [13].

3. Підготовка набору даних

Набір зображень створено на основі цифрових аерофото знімків високої просторової роздільної здатності, отриманих з відкритого картографічного сервісу державного рівня Web Map Service (WMS) [14] через API. Для завантаження зображень використовувався запит типу GetMap у проєкції EPSG:25832 [15], що забезпечує коректну метричну прив'язку та узгодженість розмірів області зображення.

Для кожної територіальної зони визначався обмежувальний прямокутник фіксованого розміру в метрах та відповідне йому аерофото у растровому форматі з заданими параметрами ширини та висоти зображення, що відповідають цільовій просторовій роздільній здатності 0.1 м/піксель. Завантажені зображення додатково перетворювалися у формат GeoTIFF із записом просторової прив'язки (CRS і геотрансформації), що забезпечує відтворюваність експериментів та можливість подальшої просторової обробки.

Для проведення експериментального дослідження сформовано тестову вибірку, що містить 1000 локальних фрагментів зображень, з яких 390 відповідають наявності об'єктів сонячної генерації та 610 – їх відсутності. Розподіл класів є незбалансованим, що враховано під час інтерпретації метрик точності. У реальних задачах дистанційного зондування Землі просторові характеристики зображень суттєво залежать від типу системи, що робить знімки, параметрів сенсора, висоти зйомки, атмосферних умов та етапів подальшої обробки зображення пристроєм. Різні супутникові й аерознімальні платформи формують зображення з неоднаковою просторовою роздільною здатністю, що призводить до варіативності рівня деталізації сцени, чіткості контурів об'єктів та ступеня просторового згладжування.

4. Моделювання умов обмеженої просторової роздільної здатності зображення

У реальних умовах вхідні дані не завжди містять чітко окреслені межі об'єктів, а дрібні структурні елементи можуть частково або повністю втрачатися. За таких умов алгоритми аналізу зображень повинні бути стійкими до змін просторової роздільної здатності та здатними коректно працювати з даними різної якості.

У зв'язку з цим виникає необхідність формування репрезентативного набору вхідних даних, який відтворює типові обмеження реальних систем спостереження. Задача моделювання обмеженої просторової роздільної здатності полягає в штучному відтворенні деградації

просторових характеристик високодеталізованого аерофотознімка з метою імітації даних, отриманих із різних знімальних платформ. Такий підхід дає змогу дослідити вплив зниження просторової роздільної здатності на результати подальшої класифікації об'єктів та забезпечує коректну валідацію методів аналізу зображень в умовах, наближених до реальних.

Моделювання умов обмеженої просторової роздільної здатності виконується у вигляді формального алгоритму, що складається з двох послідовних етапів:

1. Просторова декомпозиція аерофотознімка на локальні фрагменти.
2. Контрольоване зниження просторової роздільної здатності кожного фрагмента.

Вихідний аерофотознімок розглядається як багатоканальне растрове зображення. Для забезпечення локального аналізу та уніфікації вхідних даних зображення поділяється на локальні фрагменти однакового розміру, сформовані у вигляді сітки. Розбиття виконується в такий спосіб, щоб забезпечити повне просторове покриття області інтересу, включно з граничними ділянками зображення. Таким чином, зображення розбивається на фрагменти розміром 512×512 пікселів. Базова просторова роздільна здатність вихідних аерофотознімків становить $r_0 = 0.10$ м/піксель.

Для моделювання умов зйомки використовується скінченна множина значень просторових роздільних здатностей

$$\mathfrak{R} = \{0.10, 0.20, 0.30, 0.40, 0.50, 0.70, 1.00, 1.30, 1.50, 2.00\} \text{ м/піксель.}$$

Для кожного значення $r \in \mathfrak{R}$ визначається коефіцієнт відповідності пікселя

$$s_r = \frac{r}{r_0}, \quad (1)$$

де r – референтна просторова роздільна здатність фрагмента знімка (м/піксель),

r_0 – значення цільової просторової роздільної здатності (м/піксель).

Дискретне багатоканальне зображення має вигляд

$$I: \{0, \dots, H-1\} \times \{0, \dots, W-1\} \times \{1, \dots, C\} \rightarrow \mathbb{R}, \quad (2)$$

де H – висота зображення (кількість рядків),

W – ширина зображення (кількість стовпців),

C – кількість спектральних каналів,

$x \in \{0, \dots, W-1\}$ – горизонтальна координата пікселя,

$y \in \{0, \dots, H-1\}$ – вертикальна координата пікселя,

$c \in \{1, \dots, C\}$ – індекс спектрального каналу.

Зменшене зображення масштабу r визначається оператором

$$I^{(r)} = D_s(I), \quad (3)$$

де D_s – оператор усереднення з коефіцієнтом масштабу s , що перетворює зображення розміру $H \times W \times C$ на зображення розміру $H' \times W' \times C$.

Розміри зменшеного зображення визначаються як

$$H' = \left\lfloor \frac{H}{s} \right\rfloor, W' = \left\lfloor \frac{W}{s} \right\rfloor, \quad (4, 5)$$

де H – висота зображення в пікселях

W – ширина зображення в пікселях

s – коефіцієнт масштабу.

Для забезпечення єдиного формату вхідних даних виконується зворотне масштабування:

$$\tilde{I}^{(r)} = u_s(I^{(r)}), \quad (6)$$

де u_s – оператор білінійної інтерполяції, що масштабує зображення розміру $H' \times W' \times C$ до фіксованого розміру $H \times W \times C$.

Під час аналізу аерофотознімків міських територій особлива увага приділялася граничним випадкам, тобто об'єктам, морфологічні характеристики яких частково збігаються з ознаками елементів сонячної генерації.

Для визначення потенційно схожих об'єктів з ознаками реальних сонячних панелей було проведено порівняльний аналіз геометричних і текстурних характеристик об'єктів зображень міських територій. Оцінювалися регулярність структури, просторові пропорції, однорідність інтенсивності в межах об'єкта та контраст з навколишнім фоном. Додатково досліджувалося видозмінення таких об'єктів при контрольованому зменшенні просторової роздільної здатності, що дало змогу простежити, як деградація деталізації зображення впливає на збереження характерної структури.

5. Обчислювальні експерименти

Обчислювальні експерименти полягали в розв'язанні задачі класифікації моделлю GPT-4o [16] за усіма зображеннями створеного датасету з 1000 аерофотознімків. За кожним зображенням модель робила чотири висновки відповідно до стратегій, представлених у розділі 1. Висновки оцінювались за precision, recall та F1-score.

Результати обчислювальних експериментів представлено у таблиці.

Таблиця. Результати класифікації сонячних панелей на зображеннях

Просторова роздільна здатність, м/піксель	Стратегія ідентифікації	Точність	Повнота	F1-міра
0.1	Класифікація на основі покрокового аналізу ознак	0.8365	0.959	0.8936
0.1	Бінарна класифікація на прикладах зображень	0.957	0.9477	0.9523
0.1	Класифікація на основі гіпотез	0.6225	0.9431	0.75
0.1	Класифікація на основі базової інтерпретації ознак	0.7906	0.9573	0.866
0.5	Класифікація на основі покрокового аналізу ознак	0.69	0.828	0.7527
0.5	Бінарна класифікація на прикладах зображень	0.9404	0.7029	0.8045
0.5	Класифікація на основі гіпотез	0.5603	0.9566	0.7067
0.5	Класифікація на основі базової інтерпретації ознак	0.6494	0.8511	0.7367
1	Класифікація на основі покрокового аналізу ознак	0.5529	0.6575	0.6007
1	Бінарна класифікація на прикладах зображень	0.5853	0.4029	0.4773
1	Класифікація на основі гіпотез	0.5106	0.9153	0.6555
1	Класифікація на основі базової інтерпретації ознак	0.5205	0.8654	0.65
2	Класифікація на основі покрокового аналізу ознак	0.552	0.8857	0.6801
2	Бінарна класифікація на прикладах зображень	0.6653	0.2972	0.4109
2	Класифікація на основі гіпотез	0.6991	0.6077	0.6502
2	Класифікація на основі базової інтерпретації ознак	0.4573	0.5769	0.5102

Порівняльний аналіз результатів з різними промпт-стратегіями підтверджує суттєву залежність точності від просторової роздільної здатності аерофотознімків.

На рисунку, представлено залежності F1-міри від просторової роздільної здатності зображення, отримані за всіма промпт-стратегіями. За високої деталізації найвищі значення F1-міри отримано за стратегією «бінарна

класифікація на прикладах зображень», однак при 1 – 2 м/піксель спостерігається різке зниження цієї характеристики. Стратегії «класифікація на основі покрокового аналізу ознак» та «класифікація на основі гіпотез» характеризуються помірнішою деградацією показників, що свідчить про їх більшу стійкість до втрати просторової інформації.

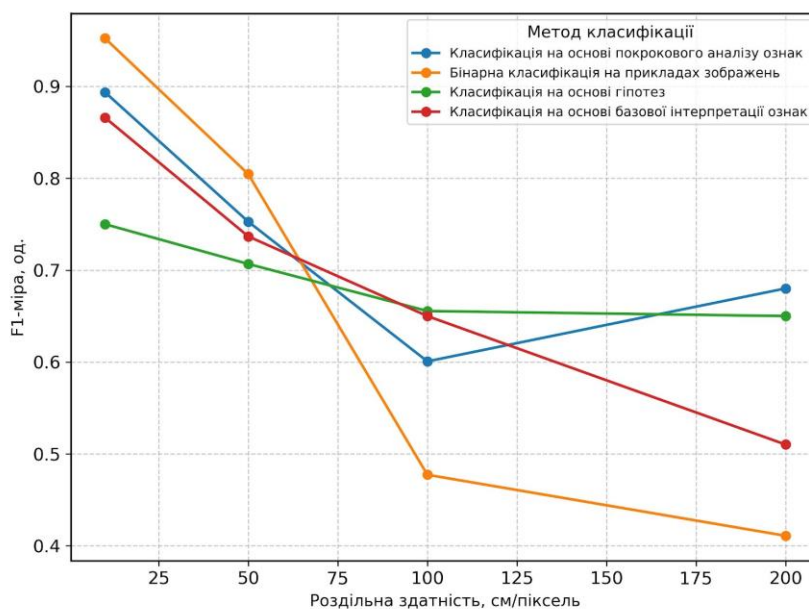


Рисунок. Залежність F1-міри від просторової роздільної здатності

Отримані результати засвідчують наявність різних режимів ефективності стратегій залежно від масштабу пікселя. За високої просторової роздільної здатності найефективнішим є підхід, орієнтований на подібність до еталонних прикладів, тоді як зі зниженням деталізації стійкішими виявляються стратегії, засновані на структурованому та ієрархічному аналізі ознак. Результати вказують на різну ефективність стратегій залежно від рівня просторової деталізації. Підходи, що спираються на локальні текстурні ознаки, є оптимальними при високій роздільній здатності, тоді як зі зростанням масштабу пікселя перевагу отримують методи, орієнтовані на інтегральні геометричні та фотометричні характеристики об'єкта. Це підтверджує необхідність адаптації логіки ідентифікації до умов просторової деградації зображення.

Висновки

1. Формалізовано задачу автоматизованої ідентифікації об'єктів сонячної генерації на аерофотознімках як задачу бінарної класифікації локальних фрагментів з використанням великої мультимодальної моделі без зміни її параметрів. Запропоновано двоетапний алгоритм, а саме просторову декомпозицію зображення та перевірку гіпотези щодо наявності об'єкта на рівні окремого фрагмента. Це дало змогу трансформувати ресурсоемну задачу семантичної сегментації у задачу локальної класифікації з нижчими обчислювальними витратами.
2. Запропоновано чотири промпт-стратегії виявлення заданого об'єкта на супутникових та аерофотознімках, що дає змогу експериментально визначити оптимальну за точністю стратегію для прикладної задачі залежно від просторової роздільної здатності.
3. Розроблено алгоритм моделювання обмеженої просторової роздільної здатності знімків, що базується на контрольованому масштабуванні та інтерполяції локальних фрагментів зображення. Встановлено закономірності впливу деградації просторової інформації в діапазоні 1–2 м/піксель на ефективність автоматизованої ідентифікації. Показано, що зі зменшенням деталізації зображення знижується інформативність локальних текстурних ознак та змінюється режим ефективності стратегій класифікації.
4. Проведено обчислювальні експерименти з порівняння стратегій формування критеріїв ідентифікації об'єктів сонячної генерації за різної просторової роздільної здатності. Встановлено, що за високої деталізації, 0.1–0. м/піксель, найвищу точність забезпечує стратегія «бінарна класифікація на прикладах зображень» (F1-score = 0.9523), тоді як при збільшенні масштабу пікселя, 1–2 м/піксель, стійкішими виявляються «класифікація на основі покрокового аналізу ознак» (F1-score = 0.6801) та «класифікація на основі гіпотез» (F1-score = 0.6502). Підтверджено необхідність адаптації логіки ідентифікації до умов просторової деградації зображення.

Отримані результати доводять можливість застосування мультимодальних мовних моделей для масштабованої автоматизованої інвентаризації об'єктів розподіленої сонячної генерації без спеціалізованого навчання моделей, що знижує обчислювальні витрати за рахунок відсутності етапу спеціалізованого навчання та підвищує практичну придатність методу для обробки великих територій.

ПОСИЛАННЯ

1. Center for European Policy Analysis. A rebirth in flame: Ukraine's beleaguered energy system. 2024. URL: <https://cepa.org/article/a-rebirth-in-flame-ukraines-beleaguered-energy-system>.
2. Bondarenko D., Matiakh S., Surzhyk T., Sheiko I. Aspects of the further development of photoenergy according to the materials of the scientific and practical conference «Renewable energy and energy efficiency in the 21st century» 2023. Vidnovluvana Energetika. 2023. No. 4(75). Pp. 39–44. [https://doi.org/10.36296/1819-8058.2023.4\(75\).39-44](https://doi.org/10.36296/1819-8058.2023.4(75).39-44)
3. Bondarenko D., Matiakh S., Surzhyk T., Sheiko I., Kravchenko M. DEVELOPMENT TRENDS OF SOLAR POWER ENGINEERING BASED ON THE MATERIALS OF THE SCIENTIFIC AND PRACTICAL CONFERENCE «RENEWABLE ENERGY AND ENERGY EFFICIENCY IN THE 21. Vidnovluvana Energetika. 2024. № 3(78). Pp. 76–83. [https://doi.org/10.36296/1819-8058.2024.3\(78\).76-83](https://doi.org/10.36296/1819-8058.2024.3(78).76-83)
4. Bondarenko D., Matiakh S., Surzhyk T., Sheiko I. ACTUAL PATHWAYS FOR SOLAR ENERGY DEVELOPMENT BASED ON THE MATERIALS OF THE SCIENTIFIC AND PRACTICAL CONFERENCE «RENEWABLE ENERGY AND ENERGY EFFICIENCY IN THE 21ST CENTURY» 2025. Vidnovluvana Energetika. 2025. № 3(82). Pp. 82–88. [https://doi.org/10.36296/1819-8058.2025.3\(82\).82-88](https://doi.org/10.36296/1819-8058.2025.3(82).82-88)
5. Matiakh S., Ausheva N., Kardashov O., Bondarenko D. . DETERMINATION OF THE PHOTOVOLTAIC POTENTIAL OF TERRITORIAL COMMUNITIES OF UKRAINE USING GEOINFORMATION SYSTEMS. Vidnovluvana Energetika. 2025. № 4(83). Pp. 165–188. [https://doi.org/10.36296/1819-8058.2025.4\(83\).165-188](https://doi.org/10.36296/1819-8058.2025.4(83).165-188)
6. Уряд спростив процедури встановлення сонячних електростанцій на будівлях [Електронний ресурс]. Режим доступу: <https://www.kmu.gov.ua/news/uriad-sprostyv-protsedury-vstanovlennia-soniachnykh-elektrostantsii-na-budivliakh>
7. Ma L., Liu Y., Zhang X. Deep learning in remote sensing applications: A meta-analysis and review. ISPRS Journal of Photogrammetry and Remote Sensing. 2019. Vol. 152. Pp. 166–177. DOI: <https://doi.org/10.1016/j.isprsjprs.2019.04.015>.

-
8. Richards J. Remote Sensing Digital Image Analysis. Berlin: Springer, 2013. 494 p. DOI: <https://doi.org/10.1007/978-3-642-30062-2>.
 9. Schowengerdt R. Remote Sensing: Models and Methods for Image Processing. Burlington: Academic Press, 2007. 560 p.
 10. Long J., Shelhamer E., Darrell T. Fully convolutional networks for semantic segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2015. Vol. 37. No. 11. DOI: <https://doi.org/10.48550/arXiv.1411.4038>.
 11. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition. Proceedings of CVPR. 2016. DOI: <https://doi.org/10.1109/CVPR.2016.90>.
 12. Alshammari R., Khan S., Alqahtani A. Automated Rooftop Solar Panel Detection Through Convolutional Neural Networks. International Journal of Remote Sensing. 2024. DOI: <https://doi.org/10.1080/07038992.2024.2363236>.
 13. Li X., Zhang Y., Wang H. Enhancing Rooftop Photovoltaic Segmentation Using Spatial Feature Reconstruction and Multi-scale Feature Aggregation. Energies. 2025. DOI: <https://doi.org/10.3390/en18010119>.
 14. Radford A., Kim J. W., Hallacy C. Learning Transferable Visual Models From Natural Language Supervision. Proceedings of ICML. 2021. DOI: <https://doi.org/10.48550/arXiv.2103.00020>.
 15. Vaswani A., Shazeer N., Parmar N. Attention Is All You Need. Advances in Neural Information Processing Systems. 2017. DOI: <https://doi.org/10.48550/arXiv.1706.03762>.
 16. Dosovitskiy A. Beyer L., Weissenborn D. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. ICLR. 2021. DOI: <https://doi.org/10.48550/arXiv.2010.11929>.
 17. Ouyang L. et al. Training language models to follow instructions with human feedback. NeurIPS. 2022. DOI: <https://doi.org/10.48550/arXiv.2203.02155>.
 18. Shapovalova S., Holovakin M. Prompt Strategies for Multimodal LLM-Based Rooftop Solar Panel Detection, Zenodo. 2026 DOI: <https://doi.org/10.5281/zenodo.18670195>.
 19. Land Nordrhein-Westfalen, Geobasis NRW. Digital Orthophotos via WMS Service (WMS NW DOP). Düsseldorf, 2024. URL: https://www.wms.nrw.de/geobasis/wms_nw_dop.
 20. EPSG Geodetic Parameter Dataset. EPSG:25832 – ETRS89 / UTM zone 32N. URL: <https://epsg.org/>.
 21. OpenAI. GPT-4o System Card. 2024. DOI: <https://doi.org/10.48550/arXiv.2410.21276>.
-

AUTOMATED IDENTIFICATION OF ROOFTOP SOLAR POWER PLANTS USING LARGE MULTIMODAL MODELS UNDER LIMITED VISUAL INFORMATION

Received Feb. 22, 2026; accepted Mar. 23, 2026
Available online Mar. 31, 2026

Shapovalova S.¹, Matiakh S.², Holovakin M.³

Author for correspondence: Tkach Dmytro,
e-mail: DmitriyT@krok.edu.ua

¹ Cand. of Sciences (Tech.), Associate Professor
<https://orcid.org/0000-0002-3431-5639>

² Cand. of Sciences (Tech.)
<https://orcid.org/0000-0002-1707-3519>

³ Postgraduate Student
<https://orcid.org/0009-0006-1456-8926>

^{1,3} National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

² Institute of Renewable Energy of the National Academy of Sciences of Ukraine, Kyiv, Ukraine,

Abstract. A method for detecting rooftop photovoltaic systems in satellite and aerial imagery under degraded spatial resolution is proposed. Object detection is based on solving an image classification task using a large language model to reduce computational load and accelerate the identification process. The proposed method is intended for rapid inventory of solar generation facilities to enhance the resilience of Ukraine's energy infrastructure. Unlike semantic segmentation approaches, which require specialized model training and significant computational resources, the proposed method reformulates the task as binary classification of local image fragments using a large language model. This approach eliminates the need for additional training and reduces overall computational costs by replacing the resource-intensive semantic segmentation task with a binary classification problem applied to local image patches. Four prompt strategies for target object detection are developed and evaluated. An algorithm for simulating limited spatial resolution through controlled scaling and interpolation is introduced. Computational experiments were performed using the GPT-4o large language model to assess alternative strategies for formulating identification criteria for solar generation assets at different levels of spatial resolution. Degradation in the range of 1–2 m/pixel significantly affects detection accuracy. At high detail levels (0.1 – 0.3 m/pixel), the highest performance is achieved by the "binary classification based on image examples" strategy (F1-score = 0.9523). At lower resolutions (1–2 m/pixel), the more robust approaches are "classification based on step-by-step feature analysis" (F1-score = 0.6801) and "classification based on hypotheses" (F1-score = 0.6502). The results demonstrate that multimodal language models can support scalable automated inventory of distributed solar installations over large territories without task-specific training.

Keywords: aerial photograph, satellite image, solar panels, object detection in images, LLM, GPT, spatial resolution of aerial photographs, energy infrastructure.

List of Symbols and Abbreviations

RS – Remote – Sensing
SPP – Solar Power Plant
CEPA – Center for European Policy Analysis
WMS – Web Map Service
OGC – Open Geospatial Consortium
CRS – Coordinate Reference System

EPSG – Coordinate reference system code (EPSG:25832 used in this study)
GeoTIFF – Georeferenced raster image format
GetMap – WMS request type for retrieving a raster map fragment
LLM – Large Language Model
API – Application Programming Interface

Introduction

Automated identification of rooftop solar installations has become increasingly important in the context of large-scale destruction of Ukraine's energy infrastructure resulting from military actions. Under such conditions, rapid assessment of distributed generation assets over extensive territories is essential for evaluating renewable energy potential and planning emergency measures aimed at strengthening energy system resilience. According to CEPA [1], missile and drone attacks between 2022 and 2024 led to substantial losses in electricity generation capacity. In 2024, approximately 9 GW of capacity was lost,

corresponding to nearly one-third of the pre-war electricity consumption level, while the share of thermal generation declined from 23.5% to about 5%. In this context, decentralized renewable installations – particularly rooftop photovoltaic systems – play an increasingly significant role, as they enhance local energy resilience and reduce pressure on damaged infrastructure.

The implementation of rooftop solar power plants in Ukraine is rapidly developing as a means of energy independence. Their use is economically beneficial in the long term, allowing communities to reduce electricity costs and increase independence from the central power grid.

Scientific research on the implementation of rooftop solar power plants in Ukraine focuses on assessing the technical potential, optimizing structures, reducing energy losses, determining the efficiency of placing solar power plants on the roofs of residential and industrial facilities, assessing their impact on the stability of distribution networks and the possibility of reducing the load on the networks [2-5].

Since the implementation of rooftop solar power plants in Ukrainian communities is a priority for increasing energy resilience, especially during blackouts, the government simplified the procedure for their installation on the roofs and facades of buildings, eliminating the need for many permits, which requires an operational inventory of solar generation facilities in order to increase the resilience of Ukraine's energy infrastructure [6]

However, the absence of up-to-date inventories and operational monitoring tools complicates their integration into energy planning processes. Traditional approaches to aerial image analysis [7-9], including semantic segmentation methods [10], are computationally intensive and rely on specialized annotated datasets, limiting their applicability under variable imaging quality and reduced spatial resolution. Consequently, the development of scalable methods for rapid identification of rooftop photovoltaic systems that can operate under degraded visual conditions without prior model training remains an important research challenge.

Most contemporary studies perform rooftop solar panel detection in aerial imagery using convolutional neural networks (CNNs) [11]. These approaches primarily aim to improve detection accuracy by refining network architectures and incorporating multi-scale visual features.

In [12], the authors propose a semantic segmentation framework based on a deep convolutional neural network trained on a specialized annotated dataset. The study reports high accuracy in detecting solar panels in high-resolution imagery. While the results demonstrate the effectiveness of CNN-based approaches for this task, the method requires prior model training, substantial computational resources, and high-quality input data.

In [13], the authors propose an enhanced convolutional neural network architecture incorporating multi-scale feature analysis and spatial reconstruction modules to improve solar panel segmentation accuracy. While this approach achieves superior performance compared to baseline models, the improvement comes at the cost of increased architectural complexity and higher computational demand.

More broadly, contemporary research on automated solar panel detection focuses primarily on maximizing accuracy through increasingly sophisticated CNN architectures and additional feature processing mechanisms, often assuming high-quality imagery. Although such methods report strong performance, they remain computationally intensive and require specialized training datasets, limiting their suitability for rapid large-scale analysis or for operation under reduced spatial resolution conditions.

Semantic segmentation methods based on deep learning require substantial computational resources and processing time for accurate object detection across large image datasets. This creates the need for a computationally efficient preliminary screening approach capable of identifying image subsets that are highly likely to contain the target object.

This study proposes an efficient alternative that eliminates task-specific model training and shifts the emphasis from increasing architectural complexity to formalizing decision criteria for visual interpretation. The method leverages LLMs [14-17] for automated object classification in aerial imagery.

Unlike conventional computer vision approaches, where classification logic is embedded in trained model parameters, LLM-based analysis relies on explicit prompt-driven specification of visual criteria. This enables rapid preliminary screening of large aerial image datasets without additional training, facilitating efficient selection of candidate image subsets with a high likelihood of containing solar panels. As a result, the overall computational burden of subsequent semantic segmentation is substantially reduced.

The aim of this study is to develop a method for detecting solar generation facilities in satellite and aerial imagery under conditions of degraded spatial resolution.

To achieve the stated aim, the study must address the following tasks:

1. Formalize the task of identifying solar generation facilities in aerial imagery for the use LLM.
2. Develop prompt strategies for the LLM.
3. Prepare a set of images for computational experiments.
4. Develop an algorithm for simulating conditions of limited spatial resolution to study the impact of visual information degradation on identification results.
5. Conduct a comparative analysis of the results of computational experiments on solar panel detection under different prompt strategies and resolution conditions.

The object of the study is methods and tools for solving the problem of binary classification of objects in aerial imagery.

The subject of the study is the application of multimodal language models for object classification in local image fragments at different spatial resolutions.

Research methods: trained large language models, GPT-4o, prompt-oriented image identification strategies, modeling of different levels of aerial image detail.

1. The task of detecting the presence of a solar panel in an aerial image based on an LLM

The task of detecting solar arrays in an aerial image is formulated as a binary image classification problem.

For proper image analysis, aerial imagery must be pre-processed. Therefore, solar panel detection is carried out in two stages:

1. Spatial decomposition.

The region of interest in the image is divided into a set of local fragments of fixed size, ensuring localization of the analysis and unification of input data for further processing.

2. Classification of the presence or absence of a solar panel in local aerial image fragments.

For each image fragment, the presence of solar generation facilities is evaluated using an LLM.

For classification, each fragment is provided to the LLM together with a prompt describing the visual characteristics of solar panels and the detection task. The model outputs a confidence score in the range [0,1]. A fragment is classified as containing a solar panel if the score exceeds 0.6.

Object detection in aerial imagery is complicated by the variability of input data, which may differ significantly in:

1. Spatial resolution – the size of the ground area represented by a single pixel.
2. Image quality – the presence of missing fragments or blurred regions.
3. Visual similarity to other objects, as characteristic features of solar panels may be partially or completely lost, or may resemble other rooftop structures.

Under such conditions, direct image classification without explicit feature specification may lead to unstable predictions and increased misclassification rates.

When defining the visual characteristics of solar panels, it is important to consider that rooftop PV systems may share similarities with agricultural greenhouses with translucent coverings, metal or reflective roofs, skylights, rooftop mechanical structures, and structural shadows.

The use of an LLM enables selective analysis of local image fragments rather than full-scene processing. Computation is restricted to fragments for which the binary classifier indicates a sufficiently high likelihood of containing solar panels. This approach transforms the problem from resource-intensive semantic segmentation requiring task-specific training into localized binary classification, thereby reducing overall computational cost and improving scalability for large-area satellite imagery processing.

2. Prompt Strategies for the LLM

The effectiveness of automated solar panel detection depends on how the criteria for visual interpretation are specified within the prompt. Four identification strategies were evaluated, differing in the decision logic used for object recognition:

1. Classification based on basic feature interpretation.
2. Binary classification based on image examples.
3. Classification based on step-by-step feature analysis.
4. Classification based on hypotheses.

Classification based on basic feature interpretation. In this strategy, LLM was instructed to act as an expert. A textual description of typical solar panel characteristics was provided, including a dark rectangular structure, regular geometry, and contrast with the surrounding background.

This approach approximates intuitive visual interpretation without detailed verification of individual features. At high spatial resolution, it provides acceptable accuracy. However, as spatial detail decreases, the false-positive rate increases, particularly due to similarity with dark roofs or shadows.

Binary classification based on image examples. In this strategy, the LLM was prompted with the role: *“You are a specialized AI for detecting solar arrays in satellite imagery.”* Two example images were provided, representing typical cases of the presence and absence of solar panels. The prompt included explanations of the classification decision for each example.

Here, the identification criteria are implicitly defined through representative target-class examples and borderline cases. The classification decision is based on similarity between the analyzed fragment and the reference samples.

For high-quality imagery, this method significantly improves accuracy because structural panel elements remain distinguishable. However, performance degrades as spatial resolution decreases, since fine textural details essential for similarity comparison become partially or completely lost.

Classification based on step-by-step feature analysis. In this strategy, the LLM was instructed to act as an analytical expert. Object identification is performed in three stages:

1. Identification of key visual features: luminance, uniformity, and edge structure.
2. Deductive inference of class membership based on feature combinations.
3. Conversion of the final decision into a confidence score in the range [0,1].

At the first stage, relative brightness and geometric structure are evaluated. At the second stage, spatial grouping of elements is analyzed.

This approach reduces reliance on fine textural details and shifts evaluation toward more general features that remain stable under moderate spatial resolution degradation. Experimental results showed improved stability for medium-quality imagery.

Classification based on hypotheses. In this strategy, the LLM was prompted to act as an expert researcher (*“You are a Lead Scientist in Remote Sensing.”*). The classification process is organized as parallel evaluation of alternative hypotheses regarding the rooftop object:

1. Formulation and evaluation of Hypothesis A (bright reflective object / greenhouse). High luminance, glare, and internal structure are analyzed.

2. Formulation and evaluation of Hypothesis B (isolated small object / skylight). Object isolation and relative size compared to the roof are evaluated.
3. Formulation and evaluation of Hypothesis C (solar panels). Lower relative luminance and spatial clustering of modules or regular structure are analyzed.
4. Synthesis of results. The hypotheses are compared, and the class “solar panels” is assigned only if Hypothesis C is more strongly supported than A and B.

This strategy demonstrated the highest robustness to decreasing spatial resolution, as it relies on comparison of alternative explanations rather than a single feature. However, in cases of dark uniform roofs, false positives may increase due to overestimation of Hypothesis C.

The full prompt formulations for each strategy are provided in [18].

3. Dataset Preparation

The image dataset was constructed using high spatial resolution digital aerial imagery obtained from a national open Web Map Service (WMS) [19] via API access. Images were retrieved using the *GetMap* request in the EPSG:25832 projection [20], ensuring correct metric referencing and consistent image area dimensions.

For each selected territorial zone, a bounding rectangle of fixed size (in meters) was defined. The corresponding raster image was downloaded with specified width and height parameters to achieve a target spatial resolution of 0.1 m/pixel. The images were subsequently converted to GeoTIFF format with embedded spatial referencing (CRS and geotransformation), ensuring reproducibility of the experiments and enabling further spatial processing.

The experimental dataset consists of 1000 local image fragments, including 390 fragments containing solar panels and 610 without solar panels. The class distribution is imbalanced and was considered during the interpretation of performance metrics.

In real-world remote sensing applications, image characteristics depend on the imaging platform, sensor parameters, flight altitude, atmospheric conditions, and post-processing procedures. Different satellite and aerial systems produce imagery with varying spatial resolution, resulting in differences in scene detail, object boundary clarity, and spatial smoothing.

4. Modeling Conditions of Limited Spatial Image Resolution

In real-world conditions, input imagery does not always contain clearly defined object boundaries, and fine structural elements may be partially or completely lost. Therefore, image analysis algorithms must be robust to variations in spatial resolution and capable of operating reliably with data of varying quality.

To address this issue, a representative dataset was constructed to reproduce typical limitations of real observation systems. Modeling limited spatial resolution involves

artificially degrading high-detail aerial imagery to simulate data obtained from different imaging platforms. This approach enables analysis of the impact of spatial resolution reduction on object classification performance and supports validation of the proposed method under conditions close to real-world scenarios.

Modeling limited spatial resolution is implemented as a two-stage algorithm:

1. Spatial decomposition of the aerial image into local fragments.
2. Controlled reduction of the spatial resolution of each fragment.

The original aerial image is treated as a multichannel raster image. To enable localized analysis and standardized input formatting, the image is divided into equal-sized fragments arranged in a grid. The partitioning ensures complete spatial coverage of the region of interest, including boundary areas. Each fragment has a size of 512×512 pixels. The baseline spatial resolution of the original imagery is $r_0 = 0.10$ m/pixel.

To simulate imaging conditions, a finite set of spatial resolution values

$$\mathcal{R} = \{0.10, 0.20, 0.30, 0.40, 0.50, 0.70, 1.00, 1.30, 1.50, 2.00\} \text{ m/pixel is used.}$$

For each value $r \in \mathcal{R}$ a pixel correspondence coefficient is determined:

$$s_r = \frac{r}{r_0} \quad (1)$$

where r – the reference spatial resolution of the image fragment (m/pixel),

r_0 – the value of the target spatial resolution (m/pixel).

The discrete multichannel image has the form

$$I: \{0, \dots, H-1\} \times \{0, \dots, W-1\} \times \{1, \dots, C\} \rightarrow \mathbb{R} \quad (2)$$

де H – image height (number of rows),

W – image width (number of columns),

C – number of spectral channels,

$x \in \{0, \dots, W-1\}$ – horizontal pixel coordinate,

$y \in \{0, \dots, H-1\}$ – vertical pixel coordinate,

$c \in \{1, \dots, C\}$ – spectral channel index.

The downscaled image with scale r is defined by the operator

$$I^{(r)} = D_s(I) \quad (3)$$

where D_s – is an averaging operator with scale factor s , which transforms an image of size $H \times W \times C$ into an image of size $H' \times W' \times C$.

The dimensions of the downscaled image are determined as

$$H' = \left\lfloor \frac{H}{s} \right\rfloor, W' = \left\lfloor \frac{W}{s} \right\rfloor, \quad (4, 5)$$

where H – image height in pixels,

W – image width in pixels,

s – scale factor.

To ensure a unified input data format, inverse scaling is performed:

$$\tilde{I}^{(r)} = u_s(I^{(r)}), \quad (6)$$

where u_s – is a bilinear interpolation operator that rescales an image of size $H' \times W' \times C$ to the fixed size $H \times W \times C$.

During the analysis of urban aerial imagery, particular attention was given to edge cases, i.e., objects whose

morphological characteristics partially resemble those of solar panels.

To identify visually similar objects, a comparative analysis of geometric and textural features in urban images was performed. Structural regularity, spatial proportions, intensity uniformity within the object, and contrast with the surrounding background were evaluated. In addition, the behavior of such objects under controlled spatial resolution degradation was examined to assess how loss of image detail affects preservation of their characteristic structure.

5. Computational Experiments

Computational experiments were conducted using the GPT-4o model [21] on the full dataset of 1000 aerial image fragments. For each image, the model generated predictions according to the four strategies described in Section 1. Performance was evaluated using precision, recall, and F1-score metrics. The results are presented in Table.

Table. Results of Solar Panel Classification in Images

Spatial resolution, m/pixel	Identification strategy	Precision	Recall	F1-score
0.1	Classification based on step-by-step feature analysis	0.8365	0.959	0.8936
0.1	Binary classification based on image examples	0.957	0.9477	0.9523
0.1	Classification based on hypotheses	0.6225	0.9431	0.75
0.1	Classification based on basic feature interpretation	0.7906	0.9573	0.866
0.5	Classification based on step-by-step feature analysis	0.69	0.828	0.7527
0.5	Binary classification based on image examples	0.9404	0.7029	0.8045
0.5	Classification based on hypotheses	0.5603	0.9566	0.7067
0.5	Classification based on basic feature interpretation	0.6494	0.8511	0.7367
1	Classification based on step-by-step feature analysis	0.5529	0.6575	0.6007
1	Binary classification based on image examples	0.5853	0.4029	0.4773
1	Classification based on hypotheses	0.5106	0.9153	0.6555
1	Classification based on basic feature interpretation	0.5205	0.8654	0.65
2	Classification based on step-by-step feature analysis	0.552	0.8857	0.6801
2	Binary classification based on image examples	0.6653	0.2972	0.4109
2	Classification based on hypotheses	0.6991	0.6077	0.6502
2	Classification based on basic feature interpretation	0.4573	0.5769	0.5102

Comparative analysis of different prompt strategies demonstrates a strong dependence of detection accuracy on spatial resolution.

Figure illustrates the variation of F1-score with spatial resolution for all evaluated strategies. At high levels of detail, the highest F1-score is obtained with the “binary classification based on image examples” strategy. However, at 1 – 2 m/pixel, a sharp performance decline is observed.

In contrast, the “classification based on step-by-step feature analysis” and “classification based on hypotheses”

strategies show a more gradual performance decrease, indicating greater robustness to spatial information loss.

The results indicate that the effectiveness of the evaluated strategies depends on pixel scale. At high spatial resolution, approaches based on similarity to reference examples achieve the best performance. However, as image detail decreases, strategies relying on structured and hierarchical feature analysis demonstrate greater robustness.

Methods focused on local textural features are optimal for high-resolution imagery, whereas with increasing pixel size,

approaches based on broader geometric and photometric characteristics become more reliable. These findings

highlight the importance of adapting the identification strategy to spatial resolution degradation.

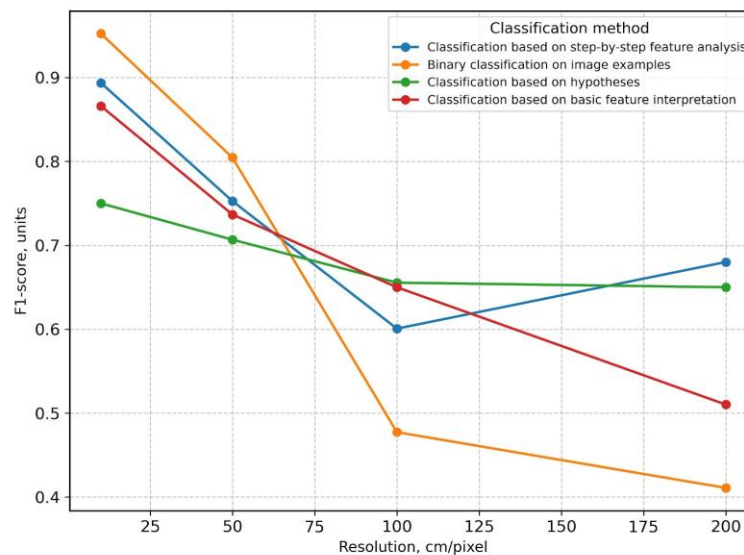


Figure. Dependence of F1-score on Spatial Resolution

Conclusion

1. The automated identification of solar panels in aerial imagery was formulated as a binary classification problem applied to local image fragments using a large multimodal model without parameter modification. A two-stage algorithm was developed, consisting of spatial decomposition and fragment-level hypothesis testing. This approach replaces resource-intensive semantic segmentation with localized classification, resulting in reduced computational cost.
2. Four prompt strategies for target object detection in satellite and aerial imagery were developed and experimentally evaluated to determine the most accurate approach under different spatial resolution conditions.
3. An algorithm for modeling limited spatial resolution based on controlled scaling and interpolation of local image fragments was developed. The impact of spatial information degradation in the range of 1 – 2 m/pixel on detection performance was analyzed. The results show that as image detail decreases, the discriminative power of local textural features declines, and the relative effectiveness of classification strategies changes.
4. Computational experiments compared the proposed identification strategies under different spatial resolution conditions. At high detail levels (0.1 – 0.3 m/pixel), the “binary classification based on image examples” strategy achieved the highest accuracy (F1-score = 0.9523). However, at lower resolutions (1 – 2 m/pixel), the more robust approaches were “classification based on step-by-step feature analysis” (F1-score = 0.6801) and “classification based on hypotheses” (F1-score = 0.6502). These results highlight the importance of adapting the identification strategy to spatial resolution degradation.

The results demonstrate that multimodal language models can be effectively applied for a scalable automated inventory of distributed solar installations without task-specific model training. By eliminating the training stage, the proposed approach reduces overall computational cost and improves practical applicability for large-area image processing.

REFERENCES

1. Center for European Policy Analysis. A rebirth in flame: Ukraine’s beleaguered energy system. 2024. URL: <https://cepa.org/article/a-rebirth-in-flame-ukraines-beleaguered-energy-system>
2. Bondarenko D., Matiakh S., Surzhyk T., Sheiko I. Aspects of the further development of photoenergy according to the materials of the scientific and practical conference “Renewable energy and energy efficiency in the 21st century” 2023. Vidnovluvana Energetika, 2023. No. 4(75), pp. 39–44. [https://doi.org/10.36296/1819-8058.2023.4\(75\).39-44](https://doi.org/10.36296/1819-8058.2023.4(75).39-44)
3. Bondarenko D., Matiakh S., Surzhyk T., Sheiko I., Kravchenko M. Development Trends of Solar Power Engineering Based on the Materials of the Scientific and Practical Conference “Renewable Energy and Energy Efficiency in the 21st Century”. Vidnovluvana Energetika, 2024, No. 3(78), pp. 76–83. [https://doi.org/10.36296/1819-8058.2024.3\(78\).76-83](https://doi.org/10.36296/1819-8058.2024.3(78).76-83)
4. Bondarenko D., Matiakh S., Surzhyk T., Sheiko I. Actual Pathways for Solar Energy Development Based on the Materials of the Scientific and Practical Conference “Renewable Energy and Energy Efficiency in the 21st Century” 2025. Vidnovluvana Energetika, 2025, No.3 (82), pp.82–88. [https://doi.org/10.36296/1819-8058.2025.3\(82\).82-88](https://doi.org/10.36296/1819-8058.2025.3(82).82-88)

5. Matiakh S., Ausheva N., Kardashov O., Bondarenko D. (2025). Determination of the Photovoltaic Potential of Territorial Communities of Ukraine Using Geoinformation Systems. *Vidnovlyuvana Energetika*, № 4(83), 165-188. [https://doi.org/10.36296/1819-8058.2025.4\(83\).165-188](https://doi.org/10.36296/1819-8058.2025.4(83).165-188)
6. Government simplifies procedures for installing solar power plants on buildings [Electronic resource]. – Available at: <https://www.kmu.gov.ua/news/uriad-sprostyv-protsedury-vstanovlennia-soniachnykh-elektrostantsii-na-budivliakh>
7. Ma L., Liu Y., Zhang X. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2019. Vol. 152. Pp. 166–177. DOI: <https://doi.org/10.1016/j.isprsjprs.2019.04.015>
8. Richards J. *Remote Sensing Digital Image Analysis*. Berlin: Springer, 2013. 494 p. DOI: <https://doi.org/10.1007/978-3-642-30062-2>
9. Schowengerdt R. *Remote Sensing: Models and Methods for Image Processing*. Burlington: Academic Press, 2007. 560 p
10. Long J., Shelhamer E., Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2015. Vol. 37. No. 11. DOI: <https://doi.org/10.48550/arXiv.1411.4038>
11. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition. *Proceedings of CVPR*. 2016. DOI: <https://doi.org/10.1109/CVPR.2016.90>
12. Alshammari R., Khan S., Alqahtani A. Automated Rooftop Solar Panel Detection through Convolutional Neural Networks. *International Journal of Remote Sensing*. 2024. DOI: <https://doi.org/10.1080/07038992.2024.2363236>
13. Li X., Zhang Y., Wang H. Enhancing Rooftop Photovoltaic Segmentation Using Spatial Feature Reconstruction and Multi-scale Feature Aggregation. *Energies*. 2025. DOI: <https://doi.org/10.3390/en18010119>
14. Radford A., Kim J. W., Hallacy C. Learning Transferable Visual Models from Natural Language Supervision. *Proceedings of ICML*. 2021. DOI: <https://doi.org/10.48550/arXiv.2103.00020>
15. Vaswani A., Shazeer N., Parmar N. Attention Is All You Need. *Advances in Neural Information Processing Systems*. 2017. DOI: <https://doi.org/10.48550/arXiv.1706.03762>
16. Dosovitskiy A., Beyer L., Weissenborn D. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *ICLR*. 2021. DOI: <https://doi.org/10.48550/arXiv.2010.11929>
17. Ouyang L. et al. Training language models to follow instructions with human feedback. *NeurIPS*. 2022. DOI: <https://doi.org/10.48550/arXiv.2203.02155>
18. Shapovalova S., Holovakin M. Prompt Strategies for Multimodal LLM-Based Rooftop Solar Panel Detection, Zenodo. 2026 DOI: <https://doi.org/10.5281/zenodo.18670195>
19. Land Nordrhein-Westfalen, Geobasis NRW. Digital Orthophotos via WMS Service (WMS NW DOP). Düsseldorf, 2024. URL: https://www.wms.nrw.de/geobasis/wms_nw_dop
20. EPSG Geodetic Parameter Dataset. EPSG:25832 – ETRS89 / UTM zone 32N. URL: <https://epsg.org/>
21. OpenAI. GPT-4o System Card. 2024. DOI: <https://doi.org/10.48550/arXiv.2410.21276>