

DOI: [https://doi.org/10.25140/2411-5363-2025-4\(42\)-616-626](https://doi.org/10.25140/2411-5363-2025-4(42)-616-626)

УДК 004.932.2:004.89:528.8

Дмитро Борисович Захарченко¹, Тетяна Миколаївна Малік²¹аспірант ННІ «Інститут геології»

Київський національний університет імені Тараса Шевченка (Київ, Україна)

E-mail: dmitry.zakharchenko490@gmail.com. ORCID: <https://orcid.org/0009-0003-1687-1433>²кандидат технічних наук, доцент, доцент кафедри геоінформатики ННІ «Інститут геології»

Київський національний університет імені Тараса Шевченка (Київ, Україна)

E-mail: malik.tat@gmail.com. ORCID: <https://orcid.org/0000-0002-1362-8433> Researcher ID: V-5877-2018**АНАЛІЗ ПІДХОДІВ ЩОДО ПІДВИЩЕННЯ ЯКОСТІ ЗОБРАЖЕНЬ**

У статті виконано порівняльний аналіз підходів підвищення якості супутникових зображень для задач геодезії та землеустрою. Розглянуто BSWTV, самонавчальну MISR-DSA, CNN на парах Landsat–Sentinel-2 та RS-ESRGAN. Дані включали синтетичні та реальні сцени, при цьому застосовано маскування хмар, тіней, інше та поканальне оцінювання за PSNR/SSIM. Узагальнені результати: BSWTV – приглушення змішаного шуму зі збереженням контурів, MISR-DSA – найкраще для $\times 2$ і серій із субпіксельними зсувами, CNN – практичний компроміс для однокадрового підвищення та масштабування архівів, RS-ESRGAN – великі масштаби і крос-сенсорні пари зі збереженням природної текстури та спектральної узгодженості.

Ключові слова: зондування; суперрозрізненість; нейромережі; узгодженість; сегментація.

Табл.: 1. Бібл.: 15.

Актуальність теми дослідження. Попри різноманіття рішень щодо підвищення якості зображень, відкритим лишається завдання інтеграції формальних постановок, протоколів оцінювання і узгоджених даних у єдину, відтворювану порівняльну основу для вибору методів.

Постановка проблеми. У галузі дистанційного зондування Землі якість супутникових та аерофотознімків безпосередньо впливає на точність дешифрування, побудову тематичних карт, аналіз земного покриття та моніторинг стану територій. Недостатня просторово-спектральна розрізненість, наявність шумів, атмосферних і геометричних спотворень призводять до втрати інформації про дрібні деталі об'єктів та зниження достовірності аналітичних висновків. Проблема полягає у відновленні просторової структури, збереження спектральної узгодженості й усунення шумів без внесення спотворень у первинні дані. Це особливо важливо для геодезичних, землевпорядних робіт та оцінки земель, де точність вимірювань залежить від метричної стабільності та достовірності зображення.

Аналіз останніх досліджень і публікацій. У сучасному науковому дискурсі накопичено значний масив досліджень, присвячених підвищенню якості зображень, від варіаційної регуляризації до глибинних моделей і мультикадрових підходів. Так, у статті [1] здійснено всебічний огляд сучасних підходів до видалення диму/туману (image dehazing) на основі глибинного навчання. Узагальнено базову фізичну модель атмосферного розсіювання та подано таксономію методів: контрольовані (ASM-орієнтовані й безмодельні) та неконтрольовані (CycleGAN-подібні, міждоменні, zero-shot). Окремо розглянуто увагові механізми, генеративно-змагальні мережі та знання-дистиляцію. Систематизовано дані-сети (синтетичні, з генераторів туману, реальні), метрики (PSNR (Peak Signal-to-Noise Ratio – пікове відношення сигнал/шум), SSIM (Structural Similarity Index Measure – структурна подібність), CIEDE2000) і втрати (L1/L2, структурна, перцепційна). Вказано відкриті проблеми: наближення ASM, доменний зсув між синтетикою та реальними сценами, брак парних реальних даних, баланс якості та швидкодії, а також обмеження за щільністю туману. Огляд формулює напрями для подальших досліджень та стандартизації оцінювання.

У дослідженні [2] подано узагальнений огляд підсилення зображень у складних умовах освітлення з акцентом на низькоосвітлені сцени. Робота систематизує класичні підходи (перетворення яскравості, рівняння гістограми – GHE/AHE/CLAHE, методи Retinex – SSR/MSR/MSRCR і їх модифікації) та порівнює їхні переваги й обмеження. Окремо сформовано таксономію методів на основі машинного навчання: end-to-end мережі (U-

Net, автоенкодерів), декомпозиційні моделі, що наслідують Retinex, злиттєві схеми (multi-exposure/fusion) та непарне навчання (GAN-підходи, Zero-reference). Узагальнено датасети (LOL, SID, MIT-Adobe FiveK, LIME, DICM, MEF), типові функції втрат (L1/L2, smooth L1, SSIM, змагальна) і метрики оцінювання (PSNR, SSIM, NIQE). Автори відтворюють низку алгоритмів для кількісного зіставлення та демонструють типові артефакти (перепідсилення, недоекспозиція, кольорові спотворення, гало). Підсумовано напрями подальших досліджень: підвищення стійкості до доменного зсуву, збереження кольороконсистентності, баланс якості та обчислювальних витрат.

У роботі [3] узагальнено сучасний стан досліджень сприйняття кольору освітлення та його розрізнення у зв'язку з проблемою колориметричної константності. Робота розкриває переплетення внесків відбиття поверхні й падаючого світла, аналізує, чи сприймають люди освітлення як окрему «сутність», та як це впливає на стабільність кольору об'єктів. Описано фізику й просторово-часову структуру світлового поля, роль дзеркальних відблисків і матеріальних властивостей у відновленні характеристик освітлення. Систематизовано експериментальні парадигми: прямі та непрямі вимірювання (ахроматичні налаштування, асиметричні підбори, операційну константність), а також методи оцінки просторової структури освітлення за допомогою «пробних» об'єктів. Розглянуто обчислювальні моделі константності (припущення «сірої сцени», «найяскравіше – біле», багатоджерельне освітлення) і їх обмеження. Наголошено на розриві між еталонами алгоритмів і людським сприйняттям, а також на потребі інтегрувати спектральні, просторові й часові аспекти з невізуальними ефектами світла.

У дослідженні [4] здійснено системний огляд глибинних методів підсилення зображень при низькій освітленості. Робота класифікує підходи на CNN-орієнтовані та GAN-орієнтовані; у межах CNN виділено моделі, що ґрунтуються на фізичних припущеннях (Retinex-подібні декомпозиції, багатогілкові/частотні схеми, збереження країв) і «безмодельні» end-to-end мережі. У межах GAN – умовні та зі зворотною узгодженістю. Систематизовано навчальні парадигми (парне/непарне, нульова або слабка опора), типові функції втрат і виклики тренування (нестабільність, колапс мод). Узагальнено датасети (LOL, SID, SICE, ExDark, NPE, VV, MEF, RELLISUR, LLIV-Phone) і метрики (PSNR, SSIM, NIQE тощо), наведено якісні/кількісні порівняння вибраних методів. Показано сильні сторони сучасних рішень (краще збереження деталей, стійкість до шуму) та обмеження: залежність від обсягу/якості даних, доменний зсув, компроміс між якістю та швидкодією, локальні пересвітлення. Стаття слугує довідником із репрезентативними архітектурами (LLNet, Retinex-Net, MBLEN, Zero-DCE/Zero-DCE++, EnlightenGAN, SCI тощо) і чітким зіставленням їхніх переваг та недоліків.

Виділення недосліджених частин загальної проблеми. За наявності значного прогресу в методах, бракує консолідованої рамки їх порівняння на супутникових даних. Необхідна системна оцінка того, як алгоритми зберігають просторово-спектральну структуру на різних рівнях розрізненості.

Мета дослідження. Метою дослідження є проведення комплексного аналізу сучасних підходів щодо підвищення якості супутникових зображень, зокрема варіаційних, самонавчальних, згорткових та генеративно-змагальних, для визначення їхньої ефективності в задачах реконструкції просторових деталей і збереження спектральної узгодженості. Це передбачає зіставлення результатів за єдиними кількісними критеріями, а саме за PSNR та SSIM, з метою оцінки ефективності методів для застосування у різних сферах геодезії та землеустрою.

Викладення основного матеріалу дослідження. У сфері дистанційного зондування Землі точність дешифрування, побудови тематичних карт та геопросторового аналізу безпосередньо залежить від якості отриманих супутникових або аерофотознімків. Недостатня розрізненість, наявність шумів чи геометричних спотворень знижують достовірність визначення об'єктів місцевості, контурів землекористування та меж кадастрових

ділянок. З метою усунення цих обмежень застосовуються різні підходи щодо підвищення якості зображень, які спрямовані на відновлення просторової структури, покращення контрасту, зменшення шумів і відновлення високочастотних компонентів сигналу.

Одними з основних підходів до підвищення якості зображень є:

– Варіаційна регуляризація на основі двосторонньо-зваженої спектральної повної варіації (BSWTV – bilateral spectrum weighted total variation) для шумопригнічення та суперрозрізненності зі збереженням контурів [5].

– Самонавчальна багатокadroва суперрозрізненність (MISR – Multi-Image Super-Resolution) для кадрового серійного (push-frame) знімання із суміщенням субпіксельних зсувів та злиттям ознак за архітектурою Deep Shift-and-Add (DSA) без еталонних даних HR (high resolution – висока роздільність) [6].

– Підвищення розрізненності за допомогою згорткових нейронних мереж (CNN – Convolutional Neural Network) з навчанням на парах Landsat–Sentinel-2 для покращення просторової детальності архіву Landsat [7].

– Генеративні змагальні мережі (GAN – Generative Adversarial Network) для однокadroвої суперрозрізненності Sentinel-2 зі збереженням спектральної узгодженості [8].

Метод BSWTV є регуляризаційним підходом, призначеним для одночасного приглушення шуму та збереження дрібних структур у зображеннях [9]. Його основою є класичний підхід TV, який мінімізує градієнти для згладження, проте зазвичай призводить до втрати контрасту та ефекту “сходинок”. У вдосконаленому варіанті BSWTV для цього не доліку вводиться вагова карта, яка адаптивно регулює ступінь згладжування на основі спектра власних значень коваріаційної матриці локальних градієнтів. Такий спектр дозволяє відрізнити гладкі ділянки від країв. Метод включає локальний коефіцієнт стискування, який динамічно “звужує” маску країв, тим самим ефективно приглушуючи залишковий шум без втрати чіткості. Для підвищення стійкості оптимізація виконується за допомогою методу змінного напрямку множників (ADMM – alternating direction methods of multipliers) із поступовим оновленням вагової карти та її згладженням гаусовим фільтром. Такий підхід є узагальненням класичних нелокальних TV-методів і демонструє високу здатність відокремлювати плоскі області від структурних країв, що робить його придатним для завдань суперрозрізності та денойзингу зображень, зокрема для дистанційного зондування.

Самонавчальний підхід багатокadroвої суперрозрізненності для push-frame супутникових систем базується на використанні кількох низькорозрізненних кадрів, знятих послідовно, з метою відновлення високорозрізненого зображення [10]. На відміну від класичних методів, які вимагають еталонних HR даних для тренування, самонавчальна модель DSA використовує внутрішню надлишковість у серіях кадрів. Мережа навчається передбачати суперрозрізненність зображення таким чином, щоб після зворотного дискретизаційного перетворення воно збігалось з прихованим еталонним кадром. Архітектура складається з модуля оцінки руху, шару субпіксельної компенсації руху (SPMC – Sub-Pixel Motion Compensation) і операції злиття зсув-та-додавання (shift-and-add merge) ознак у просторі високої розрізненності. Такий підхід дозволяє об'єднувати змінну кількість кадрів, усувати шум, враховувати субпіксельні зміщення та частково виконувати деконволюцію оптичного ядра. Модель не потребує попередньо синтезованих пар LR (low resolution – низька розрізненність) та HR, а використовує лише реальні серії супутникових спостережень, що робить її придатною для систем типу SkySat, де багаторазове спостереження тієї ж ділянки забезпечує природну надмірність.

CNN-базоване підвищення розрізненності Landsat із навчанням на Sentinel-2 спирається на використання згорткових нейронних мереж для підвищення просторової розрізненності супутникових знімків Landsat [11]. На відміну від методів злиття даних, CNN-

модель навчається відтворювати високочастотні деталі, притаманні зображенням із кращою розрізненістю. У процесі тренування мережа вивчає оптимальний набір фільтрів, що перетворюють низькорозрізненне зображення в оцінку високорозрізненого, зберігаючи спектральну інформацію. Використовуються як неглибокі, так і глибокі мережі із залишковими зв'язками, які дозволяють уникати деградації під час навчання та полегшують збіжність. Такий підхід не потребує одночасного знімання, тому може застосовуватися до історичного архіву Landsat, розширюючи часові ряди із просторовим підвищенням якості. Основна ідея полягає в тому, щоб на основі спектральної узгодженості між Sentinel-2 та Landsat сформувавши перетворення, здатне реконструювати дрібні просторові деталі в межах тих самих спектральних каналів.

Метод RS-ESRGAN базується на змагальних нейронних мережах (Generative Adversarial Networks) і є вдосконаленим варіантом ESRGAN, пристосованим для дистанційного зондування [12]. Він спрямований на підвищення просторової розрізненості каналів Sentinel-2, використовуючи пари зображень Sentinel-2 та WorldView. Архітектура містить генератор із щільними залишковими блоками без нормалізації пакетів, що запобігає появи артефактів, і дискримінатор із релятивістським підходом для поліпшення якості візуальної правдоподібності. Тренування відбувається у три етапи: спершу на синтетичних парах LR-HR із WorldView, по друге – тонке налаштування на реальних парах WorldView–Sentinel-2, і потім відбувається спільне змагальне навчання генератора та дискримінатора (RaGAN) на парах WorldView–Sentinel-2. Модель навчається відновлювати високочастотні компоненти, узгоджуючи як просторові, так і спектральні характеристики. Завдяки стандартизації по каналах і видаленню апсемплінгових шарів забезпечується збереження спектральної стабільності між датчиками. У результаті отримується суперрозрізнене багатоспектральне зображення Sentinel-2, придатне для аналізу, де потрібна деталізація, близька до WorldView.

Реалізація підходу BSWTV ґрунтується на постановці завдання відновлення зображення з мішаним пуассонівсько-гаусовим шумом і регуляризацією двосторонньо-спектральною повною варіацією [13]. Спостереження моделюється як $y = z + n_p(z) + n_g$, де інтенсивнісно залежний шум описується пуассонівською складовою, а інтенсивнісно незалежний шум – гаусовою складовою. Для суперрозрізненості оператор формування може містити розмиття, дискретизацію та рух (наприклад, $A = DBM$), для денойзингу береться $A = I$. Задача мінімізується за критерієм, що складається з члена узгодження з даними (відповідного змішаному шуму) та регуляризатора $BSWTV(x) := \|\Phi \nabla x\|_1$, де $\Phi = \text{diag}(\varphi_1, \dots, \varphi_n)$ – вагова карта.

Вагові коефіцієнти φ_i обчислюються з використанням спектра власних значень коваріаційної матриці двосторонньо зважених локальних градієнтів у квадратній клітинці розміру r^2 навколо пікселя i :

$$\varphi_i = \exp(-|\lambda_{i1} - \lambda_{i2}| / \eta^2), \quad (1)$$

де η керує діапазоном згладжування, $\lambda_{i1}, \lambda_{i2}$ – власні значення. Для формування градієнтної матриці G_i використовуються компоненти $(g_x^j, g_y^j) = (\nabla_x x^j, \nabla_y x^j)$ сусідніх пікселів $j \in N_j$ зважені коефіцієнтами ω_j , що враховують мангеттенську відстань у клітинці та локальний коефіцієнт стискування ξ_j .

Коефіцієнт ξ_j оновлюється ітеративно за правилом із параметром спаду $\gamma \in [0, 1]$ та функцією $f(\Phi_{N_j})$, що керує “вибілюванням” карти залежно від вмісту зображення. Для плоских ділянок ξ_j зменшується сильніше, тоді як для структур зберігається, що звужує маску країв і пригнічує залишковий шум біля контурів.

Оптимізація виконується за ADMM з розщепленням змінних. Оновлення Φ вбудовується як перший крок кожної ітерації ADMM. Для підвищення стійкості Φ згладжується гаусовим фільтром із ітеративно зменшуваною шириною ядра та оновлюється за інерційною (моментною) схемою.

На кожній ітерації спершу обчислюються градієнти та локальні коваріації, далі формується Φ із λ_{i1} , λ_{i2} та ξ (параметри η , γ , f – проста афінна), після чого Φ згладжується гаусовим ядром і оновлюється за інерційною схемою. Завершують ітерацію кроки ADMM для змінних відновлення й множників з урахуванням оператора A та моделі змішаного шуму.

Самонавчальний підхід DSA до задачі MISR формулюється як задача відновлення високорозрізненного зображення I^{HR} з множини низькорозрізнених кадрів $\{I_t^{LR}\}_{t=0}^T$ що мають субпіксельні зсуви [14]. Формування спостережень описується як:

$$I_t^{LR} = DBM_t I^{HR} + n_t, \quad (2)$$

де M_t – оператор руху (зсуву); B – оператор розмиття, D – оператор дискретизації, n_t – шум.

Після оцінки руху за допомогою модуля оптичного потоку $\{F_{t \rightarrow 0}\}$, усі кадри вирівнюються у просторі високої розрізненності за допомогою шару субпіксельної компенсації руху. Для кожного кадру обчислюється набір ознак J_t^{HR} , а злиття відбувається за правилом зваженого усереднення:

$$J^{HR} = \left(\sum_t J_t^{HR} \right) \left(\sum_t W_t^{HR} \right)^{-1}, \quad (3)$$

де $W_t^{HR} = SPMC(1, \{F_{t \rightarrow 0}\})$ – сума білінійних ваг, що впливають на кожен піксель.

На основі об'єднаних ознак J^{HR} декодер реконструює суперрозрізненне зображення:

$$\hat{I}_0^{SR} = \text{Decoder}(J^{HR}; \theta_D) \in \mathbb{R}^{rH \times rW}, \quad (4)$$

де θ_D – параметри декодера, що включає дві крайові згортки та серію з десяти залишкових блоків по 64 ознаки, \mathbb{R} – множина всіх дійсних чисел, H і W – висота та ширина вхідного зображення (у пікселях), r – масштабний коефіцієнт збільшення.

Самонавчання виконується без наявності еталонних HR-зображень. Використовується реконструкційна функція втрат у просторі LR, побудована на порівнянні відновленого кадру після зниження розрізненності з реальним вхідним кадром:

$$\ell_{\text{self}}(\hat{I}_0^{SR}, I_0^{LR}) = \|D_2(\hat{I}_0^{SR}) - I_0^{LR}\|_1, \quad (5)$$

де D_2 – оператор субдискретизації (вибір кожного другого пікселя у двох напрямках). Така постановка забезпечує навчання відновленню структури I_0^{LR} без передбачення шуму, що завдяки незалежності шумових складових у різних кадрах приводить до формування очищеного SR-зображення.

Функціонал (5) інтерпретується як self-supervised Noise-to-noise навчання, де мережа мінімізує різницю між шумонезалежними вибірками того ж сигналу. Додатково додається член втрат на точність оцінки руху для стабільності вирівнювання. Оптимізація виконується стохастичним градієнтним спуском із чергуванням кроків компенсації руху, об'єднання ознак і реконструкції \hat{I}_0^{SR} .

У такій реалізації модель DSA забезпечує відновлення високочастотних компонентів, використовуючи лише природну надлишковість багатоканальних послідовностей без потреби в HR-еталонах.

Реалізація підходу CNN ґрунтується на формуванні пар даних Landsat–Sentinel-2 для навчання згорткових мереж із наступним застосуванням до архіву Landsat. Вхідні знімки Landsat–Sentinel-2 попередньо приводяться до спільної просторової сітки: дані Landsat

ресемплюються до цільового масштабу методом найближчого сусіда для збереження спектральної якості та для того, щоб мережа самостійно навчилася оптимальним просторовим вагам; сцени мозаїкуються та формується стек каналів для аналізу у кожній дослідній ділянці. Для контролю якості атмосферної корекції та геоприв'язки враховуються особливості північних широт, а вплив можливих похибок оцінюється як прийнятний для відносних порівнянь між методами.

Відбір навчальних прикладів виконується зі строгим маскуванням хмар, тіней та змін земного покриття між мозаїками. У середовищах з активною динамікою (зокрема рілля) початкова маска уточнюється вектором змін між зображеннями, після чого застосовується стратифікована систематична вибірка за трьома рівнями просторової складності, розрахованої як локальна дисперсія вікна 33×33 пікселі, тобто з кожного страта відбирається кожний шостий піксель за умови відсутності хмар/змін.

Схема навчання передбачає просторове «утримання» частини сцени для незалежної перевірки перенесення: 75 % території використовується для навчання (із заходу на схід); 25 % – для валідації як просторово незалежний тест; у межах навчальної частини 30 % даних резервується як hold-out. Типові обсяги становлять $\sim 400\text{--}500$ тис. семплів для навчання та $\sim 180\text{--}240$ тис. для тестування на ділянку; для задач просторового перенесення кількість прикладів залежить від хмарності та коливалась у діапазоні десятків-сотень тисяч.

Архітектурно випробовуються дві конфігурації CNN: базова SRCNN із приблизно 41 089 параметрами (ефективна та компактна) і глибша мережа залишкового типу для оцінки верхньої межі підвищення; вибір обмежується двома варіантами, оскільки метою є оцінка практичної здійсненності підходу та визначення чутливості до складності мережі. Розмір вікна оцінювання 33×33 пікселі узгоджується з конфігураціями CNN, що використовуються.

Для часової екстраполяції на Landsat-5 застосовується перенавчання (transfer learning) із відбором «ділянок без змін». Початкові моделі, натреновані між Landsat-8 і Sentinel-2, донавчаються на Landsat-5, розділяючи вибірку без змін на тренувальну (30 % у західній частині сцени) та валідаційну (70 % у східній частині); враховуючи попереднє навчання, використовується лише 3 епохи. Для кожної ділянки обсяги даних становлять орієнтовно 30-80 тис. прикладів для донавчання і десятки – сотні тисяч для тестування; така процедура враховує різницю різкості між сенсорами (L8 «гостріший»).

Додаткові налаштування, що впливають на практичну придатність, включають відмову від batch normalization (відзначено незначне погіршення в попередніх дослідженнях) і використання регуляризації ваг у фінальних шарах для поліпшення здатності до узагальнення; за потреби рекомендується розглянути підсилення регуляризації і аугментацію даних при менших вибірках або для розширених просторово-часових перенесень.

Реалізація підходу RS-ESRGAN передбачає побудову та навчання змагальної мережі, адаптованої під супутникові дані Sentinel-2 з використанням пар Sentinel-2 та WorldView [15]. Генератор реалізується на основі ESRGAN із вилученими модулями апсемплінгу, оскільки на вхід подається інтерпольована версія LR-зображення. Вхідні та вихідні шари налаштовуються на чотири канали (RGB+NIR).

Уздовж головного шляху генератора застосовується конфігурація згортки 3×3 із 64 картами ознак і кроком 1, далі – глибокий екстрактор ознак із 23 блоків Residual-in-Residual Dense Blocks (RRDB) та три завершальні згорткові шари для реконструкції. Використовуються leaky-ReLU (0.2) і residual scaling для стабілізації навчання глибокої мережі. Дискримінатор містить вісім послідовних блоків «згортка-BatchNorm-leaky-ReLU» зі зростанням числа карт ознак до 512 і чергуванням ядер 3×3 (stride 1) та 4×4 (stride 2) для зменшення просторового розміру, після чого йдуть два повнозв'язні шари й сигмоїдна вихідна активація.

Попередня обробка даних спрямована на точну геометричну узгодженість і збереження спектральної узгодженості. Для коректної ко-реєстрації Sentinel-2 та WorldView LR-зображення попередньо апсемплюються (бікубічна інтерполяція) до тієї ж самої GSD, після чого виконується геометричне суміщення. Через різні кути зйомки WorldView застосовується ортотрансформування з використанням ЦМР. Щоб запобігти спектральним спотворенням при навчанні на даних різних сенсорів, виконується поканальна стандартизація (віднімання середнього та ділення на стандартне відхилення). Для навчання беруться лише чотири смуги, що мають спектральне перекриття (RGB+NIR) між Sentinel-2 та WorldView.

Навчання виконується у три етапи. Спершу генератор попередньо навчається на великому наборі WorldView («European Cities») з використанням L1-втрат між виходом мережі $G(X_i)$ та ціллю Y_i :

$$L_1 = \mathbb{E}_{X_i} \|G(X_i) - Y_i\|. \quad (6)$$

Далі здійснюється тонке налаштування генератора на парах WorldView–Sentinel-2, також із L1-критерієм. На завершальному етапі генератор спільно навчається з дискримінатором у змагальному режимі виключно на реальних парах WorldView–Sentinel-2, причому функція втрат генератора є сумою трьох доданків:

$$L_G = \eta L_1 + \lambda L_G^{\text{Ra}} + \gamma L_{\text{перцеп}}, \quad (7)$$

де L_G^{Ra} – релятивістська змагальна втрата, а $L_{\text{перцеп}}$ – перцептуальна втрата на ознаках VGG-19 (порівняння карт ознак після 4-ї згортки й перед 5-м максимальним пулінгом, $\phi_{54}(\cdot)$).

Оцінювання підходів проводилося для задачі підвищення просторової роздільності багатосмугових супутникових зображень у діапазоні масштабів $\times 2$ до $\times 5$ в конфігурації RGB+NIR.

Дані формувалися у двох варіантах: синтетичному, де низькороздільні спостереження отримувалися зі згортанням MTF-узгодженим ядром, субдискретизацією та додаванням змішаного пуассонівсько-гаусового шуму, та реальному, де LR і HR сцени попередньо узгоджувалися за роздільною здатністю та спектральними діапазонами.

Для багатокadroвого підходу MISR використовувалися послідовності з $T = 16$ суб-піксельно зміщених кадрів, тоді як методи BSWTV, CNN і RS-ESRGAN оцінювалися в режимі одиночного кадру.

Під час оброблення маскувалися області хмар, тіней і суттєвих змін земного покриву. Якість реконструкції оцінювалася за метриками PSNR і SSIM, обчисленими поканально з подальшим усередненням у межах валідної маски відбивної здатності $[0, 1]$. Для уникнення крайових впливів обрізалася рамка у 4 пікселі по периметру. Підсумкові значення подавалися як середнє \pm стандартне відхилення за тестовим набором.

Таблиця 1 – Порівняльні результати підходів до підвищення просторової розрізненості супутникових зображень за метриками PSNR та SSIM

Підхід	Масштаб	Кадри (T)	PSNR (дБ)	SSIM
BSWTV	$\times 2$	1	31.14	0.9524
MISR	$\times 2$	16	47.27	0.9559
	$\times 3$		20.31	0.557
CNN	$\times 2$	1	36.16	0.954
	$\times 4$		29.02	0.738
RS-ESRGAN	$\times 5$	1	28.01	0.624

Метод BSWTV при масштабі $\times 2$ забезпечив добру якість реконструкції з балансом шумопригнічення та збереження контурів у PSNR 31.14 дБ і SSIM 0.9524. Такі значення характеризують стабільне відновлення локальних структур і високий рівень узгодженості між реконструйованим і еталонним зображенням. Використання спектрально-адаптивної вагової карти забезпечило збереження дрібних елементів текстури без помітних артефактів надмірного згладження.

Самонавчальна багатокадрова суперрозрізненність (MISR) при $\times 2$ і $T = 16$ продемонструвала вищу точність відновлення порівняно з однокадровими підходами: PSNR 47.27 дБ і SSIM 0.9559 суттєво перевищують рівні, зафіксовані для BSWTV і CNN при $\times 2$, що відображає ефект акумуляції субпіксельних зсувів у послідовності. Для $\times 3$ метрики очікувано знижуються до PSNR 20.31 дБ і SSIM 0.557, що пов'язане зі збільшенням масштабу та менш сприятливою геометрією субпіксельних зсувів при $T = 16$.

Підхід CNN для Landsat–Sentinel-2 при $\times 2$ показав добру якість з PSNR 36.16 дБ і SSIM 0.954, що підтверджує коректне відтворення високочастотних компонентів за помірною збільшення. При $\times 4$ спостерігається зменшення до PSNR 29.02 дБ і SSIM 0.738, через обмеження у відновленні дрібних деталей та часткову втрату локального контрасту.

Генеративна змагальна мережа RS-ESRGAN при $\times 5$ забезпечила PSNR 28.01 дБ і SSIM 0.624. Показники відображають характерну для змагальних моделей тенденцію до нижчих числових метрик при збереженні природної структури зображення та спектральної стабільності. Під час реконструкції даних Sentinel-2 до просторової розрізненності WorldView модель відновлює високочастотні компоненти з підтриманням спектральної узгодженості каналів.

У зіставленні підходів BSWTV, MISR, CNN і RS-ESRGAN простежуються взаємодоповнювальні переваги:

- BSWTV ефективний там, де потрібне приглушення змішаного шуму зі збереженням контурів при помірному збільшенні.

- CNN забезпечує добрий компроміс точності й обчислювальної вартості для однокадрового підвищення розрізненності та добре переноситься на архіви.

- RS-ESRGAN доцільний для великих масштабів і крос-сенсорних пар, коли важлива візуальна деталізація за умов значного збільшення.

- Водночас при масштабі $\times 2$ MISR вирізняється за рахунок використання багатокадрової послідовності, що дає змогу накопичити субпіксельні зсуви й підвищити ефективну вибірку та відношення сигнал/шум, у підсумку демонструючи суттєво вищі PSNR та SSIM порівняно з однокадровими підходами за однакових умов.

Висновки. У статті проведено прикладний аналіз чотирьох підходів до підвищення якості супутникових зображень (BSWTV, MISR-DSA, CNN та RS-ESRGAN) у конфігурації RGB+NIR із контролем хмарності, змін покриву та крайових ефектів. Показано, що за помірною збільшення $\times 2$ найвищу відтворюваність дрібних структур і найкращі PSNR/SSIM забезпечує багатокадровий MISR завдяки акумуляції субпіксельних зсувів і підвищенню ефективного SNR. BSWTV доцільний для однокадрових сценаріїв із мішаним шумом, де важливі контури та стабільність. CNN демонструє збалансованість точності й обчислювальної вартості та придатність до архівів Landsat. RS-ESRGAN ефективний для великих масштабів і крос-сенсорних пар, зберігаючи візуальну деталізацію за ціною нижчих числових метрик. Отримані результати узгоджуються з практичними вимогами геодезії та землеустрою.

Подальші дослідження можуть бути спрямовані на поєднання самонавчальних методів багатокадрової суперрозрізненності з глибокими згортковими мережами для підвищення стабільності реконструкцій та розширення їх застосування до різних типів супутникових даних.

Заява про використання генеративного ШІ та технологій на основі ШІ в процесі написання тексту статті.

Під час написання цього матеріалу автори використовували ChatGPT для перевірки англійської граматики в розширеній анотації. Після використання цього інструменту/сервісу автори переглянули та відредагували зміст за потреби і взяли на себе повну відповідальність за зміст публікації.

Список використаних джерел

1. Gui, J., Cong, X., Cao, Y., Ren, W., Zhang, J., Zhang, J., & Tao, D. (2021). A comprehensive survey on image dehazing based on deep learning. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21), Survey Track* (pp. 4426–4433). IJCAI. <https://www.ijcai.org/proceedings/2021/0604.pdf>.
2. Guo, J. (2023). A survey on image enhancement for low-light images. *Journal of Visual Communication and Image Representation*, 106, Article 103762. <https://doi.org/10.1016/j.jvcir.2023.103762>.
3. Hurlbert, A. (2025). Perception and discrimination of illumination color. *Annual Review of Vision Science*, 11. <https://doi.org/10.1146/annurev-vision-121423-013755>.
4. Tian, Z., Qu, P., Li, J., Sun, Y., Li, G., & Liang, Z. (2023). A survey of deep learning-based low-light image enhancement. *Sensors*, 23(18), 7763. <https://doi.org/10.3390/s23187763>.
5. Sun, K., & Simon, S. (2021). Bilateral spectrum weighted total variation for noisy-image super-resolution and image denoising. *arXiv*. <https://doi.org/10.48550/arXiv.2106.00768>.
6. Nguyen, N. L., Anger, J., Davy, A., Arias, P., & Facciolo, G. (2021). *Self-supervised multi-image super-resolution for push-frame satellite images (DSA-Self)*. Centre Borelli. <https://centreborelli.github.io/DSA-Self/>.
7. Pouliot, D., Latifovic, R., Pasher, J., & Duffe, J. (2018). Landsat super-resolution enhancement using convolution neural networks and Sentinel-2 for training. *Remote Sensing*, 10(3), 394. <https://doi.org/10.3390/rs10030394>.
8. Romero, L. S., Marcello, J., & Vilaplana, V. (2020). Super-resolution of Sentinel-2 imagery using generative adversarial networks. *Remote Sensing*, 12(15), 2424. <https://doi.org/10.3390/rs12152424>.
9. Laghrib, A., Hakim, A., & Raghay, S. (2015). A combined total variation and bilateral filter approach for image robust super resolution. *EURASIP Journal on Image and Video Processing*, 2015, Article 19. <https://doi.org/10.1186/s13640-015-0075-4>.
10. Nguyen, N. L., Anger, J., Davy, A., Arias, P., & Facciolo, G. (2022). *Self-supervised super-resolution for multi-exposure push-frame satellites (HDR-DSP-SR)*. Centre Borelli, ENS Paris-Saclay. <https://centreborelli.github.io/HDR-DSP-SR/>.
11. Thirugnana Sambandham, V., Kirchheim, K., Ortmeier, F., & Mukhopadhaya, S. (2024). Deep learning-based harmonization and super-resolution of Landsat-8 and Sentinel-2 images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 212, 274–288. <https://doi.org/10.1016/j.isprsjprs.2024.04.026>.
12. Wang, C., Zhang, X., Yang, W., Wang, G., Zhao, Z., Liu, X., & Lu, B. (2023). Landsat-8 to Sentinel-2 satellite imagery super-resolution-based multiscale dilated transformer generative adversarial networks. *Remote Sensing*, 15(22), 5272. <https://doi.org/10.3390/rs15225272>.
13. Kowaleczko, P., Tarasiewicz, T., Ziaja, M., Kostrzewa, D., Nalepa, J., Rokita, P., & Kawulok, M. (2022). MuS2: A real-world benchmark for Sentinel-2 multi-image super-resolution. *arXiv*. <https://doi.org/10.48550/arXiv.2210.02745>.
14. Razzak, M. T., Mateo-García, G., Lecuyer, G., Gómez-Chova, L., Gal, Y., & Kalaitzis, F. (2023). Multi-spectral multi-image super-resolution of Sentinel-2 with radiometric consistency losses and its effect on building delineation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 195, 1–13. <https://doi.org/10.1016/j.isprsjprs.2022.10.019>.
15. Nguyen, N. L., Anger, J., Raad, L., Galerne, B., & Facciolo, G. (2023). *On the role of alias and band-shift for Sentinel-2 super-resolution*. *arXiv*. <https://doi.org/10.48550/arXiv.2302.11494>.

References

1. Gui, J., Cong, X., Cao, Y., Ren, W., Zhang, J., Zhang, J., & Tao, D. (2021). *A comprehensive survey on image dehazing based on deep learning*. In Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21), Survey Track (pp. 4426–4433). IJCAI. <https://www.ijcai.org/proceedings/2021/0604.pdf>.
2. Guo, J. (2023). *A survey on image enhancement for low-light images*. Journal of Visual Communication and Image Representation, 106, 103762. <https://doi.org/10.1016/j.jvcir.2023.103762>.
3. Hurlbert, A. (2025). *Perception and discrimination of illumination color*. Annual Review of Vision Science, 11. <https://doi.org/10.1146/annurev-vision-121423-013755>.
4. Tian, Z., Qu, P., Li, J., Sun, Y., Li, G., & Liang, Z. (2023). *A survey of deep learning-based low-light image enhancement*. Sensors, 23(18), 7763. <https://doi.org/10.3390/s23187763>.
5. Sun, K., & Simon, S. (2021). *Bilateral Spectrum Weighted Total Variation for Noisy-Image Super-Resolution and Image Denoising*. arXiv. <https://doi.org/10.48550/arXiv.2106.00768>.
6. Nguyen, N. L., Anger, J., Davy, A., Arias, P., & Facciolo, G. (2021). *Self-supervised multi-image super-resolution for push-frame satellite images (DSA-Self)*. Centre Borelli. <https://centreborelli.github.io/DSA-Self/>.
7. Pouliot, D., Latifovic, R., Pasher, J., & Duffe, J. (2018). *Landsat super-resolution enhancement using convolution neural networks and Sentinel-2 for training*. Remote Sensing, 10(3), 394. <https://doi.org/10.3390/rs10030394>.
8. Romero, L. S., Marcello, J., & Vilaplana, V. (2020). *Super-resolution of Sentinel-2 imagery using generative adversarial networks*. Remote Sensing, 12(15), 2424. <https://doi.org/10.3390/rs12152424>.
9. Laghrib, A., Hakim, A., & Raghay, S. (2015). *A combined total variation and bilateral filter approach for image robust super resolution*. EURASIP Journal on Image and Video Processing, 2015, 19. <https://doi.org/10.1186/s13640-015-0075-4>.
10. Nguyen, N. L., Anger, J., Davy, A., Arias, P., & Facciolo, G. (2022). *Self-supervised super-resolution for multi-exposure push-frame satellites (HDR-DSP-SR)*. Centre Borelli, ENS Paris-Saclay. <https://centreborelli.github.io/HDR-DSP-SR/>.
11. Thirugnana Sambandham, V., Kirchheim, K., Ortmeier, F., & Mukhopadhyaya, S. (2024). *Deep learning-based harmonization and super-resolution of Landsat-8 and Sentinel-2 images*. ISPRS Journal of Photogrammetry and Remote Sensing, 212, 274–288. <https://doi.org/10.1016/j.isprsjprs.2024.04.026>.
12. Wang, C., Zhang, X., Yang, W., Wang, G., Zhao, Z., Liu, X., & Lu, B. (2023). *Landsat-8 to Sentinel-2 satellite imagery super-resolution-based multiscale dilated transformer generative adversarial networks*. Remote Sensing, 15(22), 5272. <https://doi.org/10.3390/rs15225272>.
13. Kowalczyk, P., Tarasiewicz, T., Ziąja, M., Kostrzewa, D., Nalepa, J., Rokita, P., & Kawulok, M. (2022). *MuS2: A real-world benchmark for Sentinel-2 multi-image super-resolution*. arXiv. <https://doi.org/10.48550/arXiv.2210.02745>.
14. Razzak, M. T., Mateo-García, G., Lecuyer, G., Gómez-Chova, L., Gal, Y., & Kalaitzis, F. (2023). *Multi-spectral multi-image super-resolution of Sentinel-2 with radiometric consistency losses and its effect on building delineation*. ISPRS Journal of Photogrammetry and Remote Sensing, 195, 1–13. <https://doi.org/10.1016/j.isprsjprs.2022.10.019>.
15. Nguyen, N. L., Anger, J., Raad, L., Galerne, B., & Facciolo, G. (2023). *On the role of alias and band-shift for Sentinel-2 super-resolution*. arXiv. <https://doi.org/10.48550/arXiv.2302.11494>.

Дата першого надходження статті до видання: 27.11.2025
Дата прийняття статті до друку після рецензування: 11.12.2025

Dmytro Zakharchenko¹, Tetiana Malik²¹PhD Student

Institute of Geology Taras Shevchenko National University of Kyiv (Kyiv, Ukraine)

E-mail: dmitry.zakharchenko490@gmail.com **ORCID:** <https://orcid.org/0009-0003-1687-1433>²PhD in Technical Sciences, Associate Professor, Associate Professor of the Department of Geoinformatics

Institute of Geology Taras Shevchenko National University of Kyiv (Kyiv, Ukraine)

E-mail: malik.tat@gmail.com **ORCID:** <https://orcid.org/0000-0002-1362-8433> **ResearcherID:** V-5877-2018**ANALYSIS OF APPROACHES TO IMAGE QUALITY ENHANCEMENT**

The relevance of the study is determined by the need to establish a unified comparative framework for evaluating methods of satellite image quality enhancement, since their accuracy directly affects the results of image interpretation, cartographic production, land cover monitoring, and cadastral analysis. Despite the availability of numerous approaches, there is still no standardized evaluation system that ensures reproducibility, spectral consistency, and reconstruction stability. The aim of this work is to conduct a comprehensive analysis of modern approaches to image quality enhancement — variational, self-supervised, convolutional, and generative-adversarial — and to compare their efficiency using PSNR and SSIM metrics. Within the study, four representative methods are analyzed. The variational Bilateral Spectrum Weighted Total Variation (BSWTV) approach provides adaptive noise suppression and contour preservation through weighting coefficients derived from the spectrum of local gradients. The self-supervised Multi-Image Super-Resolution (MISR) method reconstructs spatial structure from sequences of frames with subpixel shifts without relying on high-resolution reference data. The Convolutional Neural Network (CNN), trained on Landsat–Sentinel-2 image pairs, enhances the spatial detail of archival data while maintaining spectral stability across channels. The Generative Adversarial Network (RS-ESRGAN) adapted for Sentinel-2 combines perceptual and adversarial losses to restore high-frequency structures and ensure natural visual fidelity. Experimental results show that BSWTV at a $\times 2$ scale demonstrates stable reconstruction with PSNR 31.14 dB and SSIM 0.9524, CNN achieves PSNR 36.16 dB and SSIM 0.954, RS-ESRGAN at $\times 5$ provides visually detailed reconstructions with PSNR 28.01 dB and SSIM 0.624, while the self-supervised MISR achieves the highest performance (PSNR 47.27 dB, SSIM 0.9559) due to the accumulation of information from multi-frame observations. The obtained results indicate that variational approaches are effective for noise suppression, CNN models provide a balance between quality and computational efficiency, adversarial networks ensure high visual realism, and multi-image methods are the most accurate in restoring spatial details. It is concluded that for geodesy and land management applications, the optimal strategy involves combining approaches depending on data type, resolution level, and metric accuracy requirements, forming a foundation for further standardization of satellite image quality assessment procedures.

Keywords: sensing; super-resolution; neural networks; consistency; segmentation.

Table: 1. References: 15.