

УДК (004.932:004.032.26):528.8

д.т.н., професор Терейковський І.А., магістрант Роговий Д.С.

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

ПОРІВНЯЛЬНИЙ АНАЛІЗ МЕТОДІВ ПОКРАЩЕННЯ РОЗДІЛЬНОЇ ЗДАТНОСТІ СУПУТНИКОВИХ ЗНІМКІВ

Abstract

Ihor Tereikovskiy, professor, Dr.Sc; Dmytro Rohovyi, student
Comparative Analysis of Methods for Improving the Resolution of Satellite Images

This paper presents a comparative analysis of various baseline and neural network-based methods for satellite image upscaling, often referred to as Super-Resolution. The study evaluates the performance of several methods and models using the xView satellite imagery dataset as a benchmark. The effectiveness of these methods is compared based on key image quality metrics to identify the most suitable approaches for practical satellite image processing. This research provides valuable insights for tasks in remote sensing that require high-resolution data.

Вступ

Сучасні системи дистанційного зондування Землі генерують величезні обсяги даних, які є критично важливими для вирішення задач в оборонній сфері, агромоніторингу, картографії та моніторингу надзвичайних ситуацій. Ефективність аналізу цих даних безпосередньо залежить від їхньої просторової роздільної здатності. Однак, отримання знімків надвисокої чіткості пов'язане зі значними апаратними та фінансовими обмеженнями. У зв'язку з цим, актуальною науково-практичною задачею є програмне покращення якості (апскейлінг) наявних супутникових зображень за допомогою методів постобробки.

Короткий аналіз останніх досліджень показує, що стандартні алгоритмічні методи інтерполяції, такі як білінійна, бікубічна або інтерполяція Ланцоша, хоч і є обчислювально "легкими", але нездатні відновити складні текстури та неминуче призводять до розмиття зображень. Справжній прорив у цій галузі пов'язаний з використанням методів глибокого навчання, зокрема згорткових нейронних мереж (CNN) та генеративно-змагальних мереж (GAN). Такі архітектури, як SRCNN, VDSR, а особливо SRGAN, продемонстрували значно вищу якість відновлення зображень на загальних датасетах. Ці підходи активно адаптуються і для

специфічних завдань обробки супутникових знімків, де важливе точне відтворення дрібних об'єктів [2].

Разом з тим, залишається невирішеною частина загальної проблеми. Попри велику кількість запропонованих моделей, бракує їх систематизованого порівняльного аналізу саме в контексті обробки супутникових даних. Багато досліджень представляють нову нейромережеву архітектуру, але порівнюють її лише з 1-2 базовими (часто алгоритмічними) методами. Відкритим залишається питання, наскільки сучасні нейромережеві підходи перевершують класичні не лише за візуальною якістю, але й за формальними метриками (PSNR/SSIM), та який з усіх розглянутих методів забезпечує найкращий баланс між якістю відновлення, обчислювальною складністю та схильністю до генерації артефактів при роботі з такими складними наборами даних, як xView.

Постановка задачі

Метою даної роботи є проведення порівняльного аналізу (бенчмаркінгу) репрезентативних алгоритмічних та нейромережевих методів покращення роздільної здатності (Super-Resolution) для задачі апскейлінгу супутникових знімків. Для досягнення поставленої мети необхідно вирішити декілька ключових задач. Насамперед слід провести огляд та вибрати для дослідження як репрезентативні класичні алгоритмічні методи (напр., бікубічна інтерполяція, Ланцош), так і ключові нейромережеві архітектури (напр., SRCNN, SRGAN), що застосовуються для задачі Super-Resolution. Наступним кроком є підготовка експериментального середовища та уніфікація процесу тестування обраних моделей на репрезентативному супутниковому датасеті xView [1]. Після цього необхідно провести експериментальне дослідження, зібравши кількісні показники якості зображення (наприклад, PSNR та SSIM) для кожної моделі. На завершення, на основі ґрунтовного аналізу отриманих результатів, потрібно надати обґрунтовані рекомендації щодо вибору оптимального методу для практичного застосування в задачах нейромережевого аналізу супутникових знімків.

Методологія дослідження та експериментальні дані

Для забезпечення об'єктивного та відтворюваного порівняння методів було розроблено єдине експериментальне середовище. В якості основного набору даних було обрано датасет xView[1], що містить широкий спектр супутникових знімків високої роздільної здатності (HR). Для експерименту HR-зображення були зменшені з коефіцієнтом 4x для отримання вхідних зображень низької роздільної здатності (LR).

Для кількісної оцінки якості відновлених зображень (SR) було обрано три ключові метрики. PSNR (Peak Signal-to-Noise Ratio) вимірює рівень спотворень між відновленим та оригінальним зображенням (вимірюється в дБ, вище – краще). SSIM (Structural Similarity Index) оцінює структурну подібність зображень (в діапазоні від 0 до 1, вище – краще). LPIPS (Learned Perceptual Image Patch Similarity): перцептивна метрика на основі глибокої нейронної мережі, що оцінює візуальну схожість так, як її бачить людина (в діапазоні від 0 до ~1, нижче – краще) [3]. Окрім якості, вимірювався середній час обробки одного зображення (в секундах, нижче – краще) для оцінки обчислювальної ефективності.

Огляд досліджуваних методів

даному дослідженні було проведено порівняння двох категорій методів покращення роздільної здатності (для коефіцієнту 4x).

Класичні алгоритмічні методи. Ця група представлена стандартними методами інтерполяції, що слугують базовим рівнем (baseline): Nearest-neighbor, Bilinear, Bicubic та Lanczos.

Нейромережеві методи (Super-Resolution). Ця група включає широкий спектр архітектур глибокого навчання, від легких та швидких (FSRCNN, ESPCN, CARN) до більш складних та обчислювально "важких" моделей, що демонструють найвищу якість (LapSRN, EDSR, RCAN, RDN). Також до дослідження включена полегшена версія FSRCNN-small.

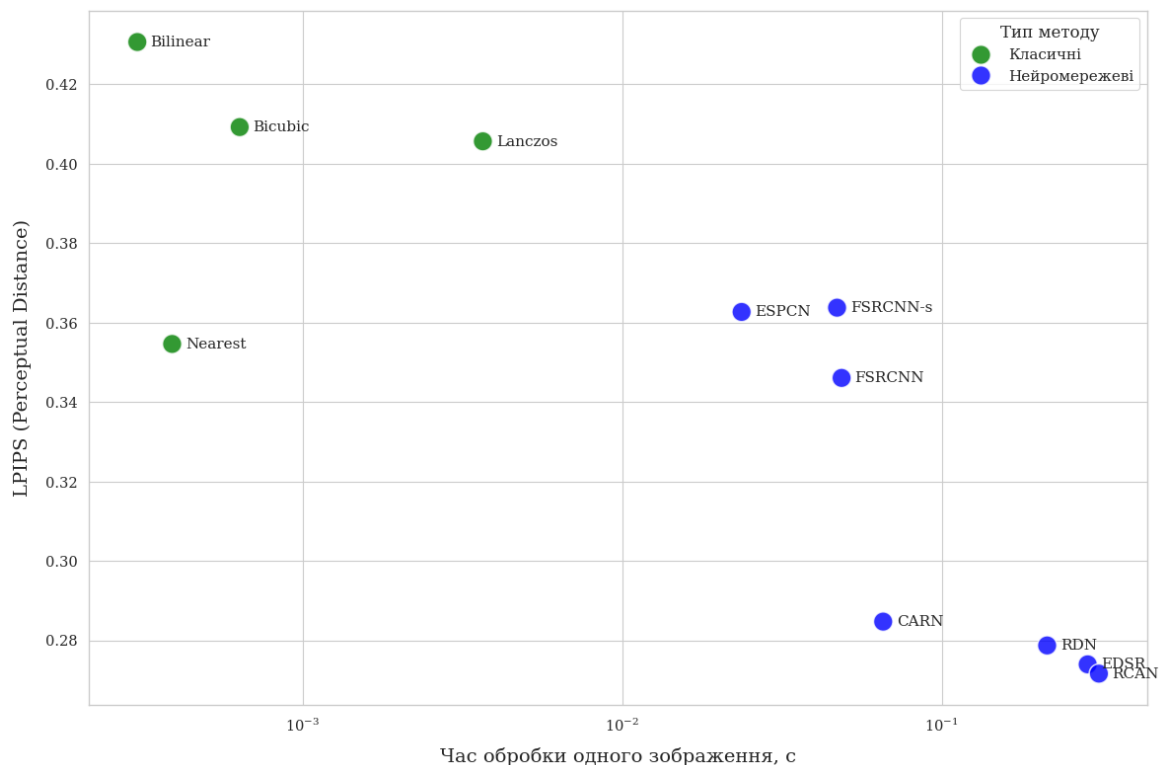


Рис. 1. Співвідношення перцептивної якості (LPIPS) та часу обробки зображення

Результати експериментів та їх аналіз

Найбільш наочно результати бенчмаркінгу демонструє точкова діаграма (Рис. 1), що відображає компроміс (trade-off) між якістю та швидкістю. На ній візуалізовано всі протестовані для масштабування 4x методи: по осі Y відкладено перцептивну якість (метрика LPIPS, де нижче значення є кращим), а по осі X – середній час обробки в логарифмічній шкалі. Така візуалізація дозволяє миттєво класифікувати моделі. Наприклад, найбажаніший квадрант "швидкі та якісні" знаходиться у лівому нижньому куті діаграми.

Для детального числового аналізу, ключові репрезентативні результати згруповані в Таблиці 1.

Таблиця 1.

Порівняльна характеристика методів апскейлінгу

Метод	PSNR (дБ) ↑	SSIM ↑	LPIPS ↓	Час (с/зобр.) ↓
Bicubic (Baseline)	32.64	0.836	0.409	0.0008
ESPCN	32.45	0.827	0.363	0.0237
CARN	33.10	0.859	0.285	0.0656
RCAN	32.86	0.858	0.272	0.3104
EDSR	32.99	0.859	0.274	0.2863

Аналіз Таблиці 1 показує, що сучасні нейромережеві архітектури (CARN, RCAN, EDSR) демонструють значну перевагу над класичним методом Bicubic за всіма трьома метриками якості. Класичні методи, хоч і найшвидші, дають "розмиті" результати, що підтверджує найгірше значення LPIPS (0.409).

Водночас, не всі нейромережі показали перевагу: моделі ESPCN та LapSRN, хоч і покращили перцептивну якість (LPIPS), фактично поступилися базовій інтерполяції за класичними метриками PSNR та SSIM.

Серед моделей спостерігається чіткий компроміс між швидкістю та якістю. "Важкі" моделі (EDSR, RCAN) демонструють найкращу перцептивну якість, причому RCAN досягла абсолютного мінімуму за LPIPS (0.272). Модель CARN представляє видатний баланс: вона здобула найвищий PSNR (33.10 дБ) та високу перцептивну якість (LPIPS 0.285) при в 4-5 разів меншому часі обробки, ніж RCAN. Це робить її оптимальним кандидатом для практичних завдань, де потрібен компроміс між якістю та швидкістю.

Висновки

У ході дослідження було проведено бенчмаркінг алгоритмічних та нейромережевих методів апскейлінгу супутникових знімків на датасеті xView, який довів беззаперечну перевагу моделей глибокого навчання над

класичними підходами за всіма ключовими метриками (PSNR, SSIM та LPIPS). Результати демонструють чіткий компроміс між швидкістю та якістю: "важкі" архітектури, як-от RCAN, досягають найкращої перцептивної якості (найнижчий LPIPS), тоді як "легша" модель CARN пропонує оптимальний баланс, показавши найвищий PSNR при відносно високій швидкодії, що робить її практичним вибором. Подальші розвідки можуть включати оцінку впливу цих методів апскейлінгу на ефективність вирішення прикладних задач, зокрема детектування об'єктів на покращених знімках.

Література

1. Lam D. xView: Objects in context in overhead imagery [Text] / D. Lam, R. Kuzma, K. McGee et al. // arXiv preprint arXiv:1802.07856. — 2018.
2. Chen H. Real-world Image Super-resolution with Deep Learning: A Survey [Text] / H. Chen, X. He, L. Qing, Y. Wu et al. // arXiv preprint arXiv:2103.02368. — 2021.
3. Zhang R. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric [Text] / R. Zhang, P. Isola, A. A. Efros, E. Shechtman et al. // arXiv preprint arXiv:1801.03924. — 2018.