

<https://doi.org/10.31891/2219-9365-2025-82-37>

УДК 004.932.2

ЛИПОВИЙ Арсен

Національний університет «Львівська політехніка»

<https://orcid.org/0009-0009-2891-3139>

arsen.y.lipovyi@lpnu.ua

АЛГОРИТМИ ПОКРАЩЕННЯ ЯКОСТІ ЗОБРАЖЕННЯ ЗА ДОПОМОГОЮ ГЛИБИННИХ НЕЙРОННИХ МЕРЕЖ: ПОРІВНЯЛЬНИЙ АНАЛІЗ СУЧАСНИХ МЕТОДІВ

У роботі розглянуто сучасні підходи до покращення якості зображень за допомогою методів глибинного навчання та штучного інтелекту. Зокрема, досліджено алгоритми Super-Resolution Convolutional Neural Network (SRCNN), Generative Adversarial Networks (GAN), Denoising Convolutional Neural Networks (DnCNN), Enhanced Super-Resolution GAN (ESRGAN) та інші. Проведено порівняльний аналіз ефективності згаданих методів при різних умовах та типах спотворень зображень, проаналізовано метрики оцінювання, а також запропоновано підходи до покращення якості зображення, орієнтовані на практичне застосування в системах комп'ютерного зору. Отримані результати демонструють перспективність використання глибинних мереж для підвищення чіткості, контрастності та роздільної здатності зображень, що може бути корисним для задач розпізнавання об'єктів, медичної візуалізації, відеоспостереження та інших суміжних сфер.

Ключові слова: покращення якості зображень, глибинне навчання, SRCNN, GAN, DnCNN, надроздільна здатність, шумозаглушення.

LIPOVYI Arsen

Lviv Polytechnic National University

ALGORITHMS FOR IMPROVING IMAGE QUALITY USING DEEP NEURAL NETWORKS: A COMPARATIVE ANALYSIS OF MODERN METHODS

This paper presents a comprehensive analysis of contemporary deep learning algorithms aimed at enhancing image quality. The study focuses on state-of-the-art methods such as Super-Resolution Convolutional Neural Network (SRCNN), Generative Adversarial Networks (GAN), Denoising Convolutional Neural Networks (DnCNN), and Enhanced Super-Resolution GAN (ESRGAN). These algorithms are evaluated for their effectiveness in improving image clarity, contrast, and resolution under various conditions and types of distortions.

The research delves into the architectural nuances of each algorithm, highlighting their unique approaches to image enhancement. For instance, SRCNN utilizes a straightforward convolutional framework for super-resolution tasks, while GAN-based methods, including ESRGAN, employ adversarial training to generate high-fidelity images with realistic textures. DnCNN focuses on removing noise from images using deep convolutional layers, demonstrating significant improvements in denoising performance.

Evaluation metrics such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) are employed to quantitatively assess the performance of these algorithms. The comparative analysis reveals that while traditional methods like SRCNN offer substantial improvements over baseline techniques, advanced models like ESRGAN achieve superior results in preserving fine details and textures, albeit sometimes at the cost of introducing artifacts.

The study also explores the practical applications of these algorithms in various domains, including medical imaging, surveillance, and autonomous vehicles. In medical imaging, enhanced image quality can lead to more accurate diagnoses. In surveillance, clearer images improve object recognition and tracking. For autonomous vehicles, high-resolution images contribute to better environment perception and decision-making.

Furthermore, the paper discusses the computational complexities associated with each algorithm, considering factors such as processing time and resource requirements. This analysis is crucial for real-world applications where computational efficiency is paramount.

In conclusion, the paper underscores the significant advancements in image quality enhancement achieved through deep learning techniques. While challenges remain, particularly concerning computational demands and potential artifacts, the progress in this field holds promise for numerous practical applications. Future research directions include optimizing these algorithms for real-time processing and further improving their robustness across diverse image types and conditions.

Keywords: image quality enhancement, deep learning, SRCNN, GAN, DnCNN, super-resolution, noise reduction.

Стаття надійшла до редакції / Received 16.04.2025

Прийнята до друку / Accepted 11.05.2025

ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

У сучасних умовах розвитку обробки зображень надзвичайно важливою задачею є підвищення якості візуальних даних. Покращення зображень має широкий спектр застосувань: від медичної візуалізації та дистанційного зондування Землі до систем відеоспостереження та автоматизованих транспортних засобів. Для ефективного вирішення цієї проблеми необхідні алгоритми, здатні відновлювати деталі зображень, підвищувати їх роздільну здатність, покращувати контрастність, зменшувати рівень шуму та артефактів. Традиційні методи покращення якості зображення, такі як інтерполяційні алгоритми або фільтраційні

підходи, мають низку обмежень. З появою глибинного навчання та конволюційних нейронних мереж (CNN) стало можливим створення моделей, які автоматично вилучають ознаки та патерни зображень, оптимізують процес відновлення та покращення їхньої якості. Зокрема, алгоритми, побудовані на базі глибинних мереж, такі як SRCNN для збільшення роздільної здатності, GAN для генерування фотореалістичних високоякісних зображень, DnCNN для шумозаглушення, продемонстрували значне підвищення ефективності у порівнянні з класичними методами[1].

ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Проблема покращення якості зображення полягає у відновленні або поліпшенні візуальних характеристик вхідного зображення. До основних аспектів відносять надроздільну здатність (Super-Resolution, SR), яка являє собою перехід від низькороздільного зображення до його високоякісного аналога. Задача SR актуальна при збереженні дрібних деталей, які було втрачено через низьку роздільну здатність сенсора чи сильну компресію. Шумозаглушення (Denoising), тобто видалення випадкового шуму, який виникає під час зйомки (зокрема, при низькому освітленні або недосконалих сенсорах), без суттєвої втрати важливих деталей. Та відновлення зображень (Restoration), при якому відбувається згладжування, видалення артефактів стиснення та інших спотворень[2]. Традиційні підходи до розв'язання цих задач часто обмежені у своєму функціоналі та продуктивності. Застосування глибинних нейронних мереж дозволило суттєво покращити результати, оскільки такі моделі здатні самостійно навчатися відповідних відновлювальних перетворень за великою вибіркою прикладів.

У контексті сучасних методів покращення якості зображень одним із перших рішень у галузі надроздільної здатності стала модель Super-Resolution Convolutional Neural Network (SRCNN). У ній передбачено використання глибинної згорткової нейронної мережі для відновлення високороздільного зображення з низькороздільного вхідного сигналу. SRCNN навчається безпосередньо від пар «низька/висока роздільна здатність», що дає можливість відтворювати як текстури, так і дрібні деталі зображення.

Поява Generative Adversarial Networks (GAN) відкрила нові перспективи у створенні фотореалістичних зображень. У напрямі підвищення їхньої якості найбільшу увагу привертає метод SRGAN, де завдяки змагальному навчанню між генератором та дискримінатором досягається висока деталізація та природність результатів[3]. Подальші вдосконалення, зокрема ESRGAN, оптимізують як архітектуру мережі, так і функції втрат, що забезпечує ще вищий рівень візуальної якості.

Для автоматичного шумозаглушення існує Denoising Convolutional Neural Networks (DnCNN). Ця модель навчається на значній кількості штучно зашумлених прикладів і відновлює їх до стану «чистого» зображення, використовуючи потенціал глибокої згорткової мережі. Такий підхід дає змогу моделювати залежності між сусідніми пікселями та ефективно усувати шум без надмірного розгладження.

Окрім згаданих методів, існує низка інших архітектур для підвищення якості зображень. Зокрема можна виділити RCAN, який використовує механізми уваги, що допомагає зосередитися на важливих просторових і каналових ознаках [4]. В Таблиці 1 проведений порівняльний аналіз даних моделей.

Аналіз проводився для декількох типових сценаріїв підвищення якості зображень.

Задача підвищення роздільної здатності (Super-Resolution, SR) оцінювалася на стандартних наборах зображень (наприклад, Set5, Set14 тощо) з штучно пониженими шляхом бікубічної інтерполяції зображеннями. Метрики PSNR та SSIM обчислювалися між отриманим високороздільним зображенням та еталонним оригіналом. Для SR-методів, орієнтованих на максимізацію PSNR (SRCNN, RCAN), було зафіксовано високі значення PSNR/SSIM (близько 27–33 дБ PSNR, SSIM ~0.75–0.90 залежно від набору даних), що вказує на їхню здатність точно відтворювати піксельну структуру оригіналу.

Задача придушення шумів (denoising) аналізувалася на зображеннях із доданим штучним гаусовим шумом (наприклад, $\sigma=25$) з відомими еталонами без шуму (BSD68, Set12 тощо). Класичний алгоритм BM3D показав високу якість очищення (PSNR ~28–29 дБ при $\sigma=25$) і слугував базовим орієнтиром. Сучасна CNN-модель DnCNN була натренована на синтетичних прикладах шуму й досягла вищого PSNR (близько 0.5–0.8 дБ виграву над BM3D) та кращого збереження структури зображень. Водночас DnCNN в «сліпому» режимі (blind denoising) здатна адаптуватися до різних рівнів шуму, якщо була відповідно навчена.

Задача усунення артефактів стиснення (JPEG-деблокінг) в аналізі була частково врахована через моделі на кшталт Real-ESRGAN, які в рамках свого навчання імітують компресійні спотворення. Спеціалізована версія DnCNN також може навчатися для видалення JPEG-артефактів – тобто, архітектура DnCNN універсальна для різних типів покращення якості за умови належного добору тренувальних даних.

Вибір метрик. PSNR (Peak Signal-to-Noise Ratio) обрано як базову кількісну метрику точності відновлення – вона чутлива до середньоквадратичної похибки та добре відображає близькість відновленого зображення до еталону на піксельному рівні. SSIM (Structural Similarity Index) доповнює аналіз, оскільки краще корелює зі сприйняттям людиною локальних структур та контрасту. Високі значення PSNR/SSIM у моделей SRCNN/RCAN підтверджують їхню здатність мінімізувати спотворення щодо оригіналу. Натомість для моделей з GAN-навантаженням низький PSNR при високій суб'єктивній якості вказує на те, що ці методи жертвують точністю заради фотореалістичності зображення. Тому в оцінці розглянуто компроміс між

точністю та візуальною якістю: метрики на прикладі LPIPS або вимірювання перцептивної дистанції тут могли б краще відобразити людське сприйняття, але вони виходять за рамки основних показників (PSNR/SSIM), на яких зосереджений аналіз.

Таблиця 1

Порівняльна таблиця глибоких моделей покращення якості зображення

Метод	PSNR (дБ)	SSIM	Віз. якість	Швидкодія (на 1 зобр.)	Парам. (М)	Артефакти	Дані навч.
SRCNN	~27.5 (×4 SR Set14)	~0.75	Краща за бікубічну, розмиті деталі	Висока (реальний час, мала модель)	~0.06	Низькі	~91 зобр. (дуже мало)
RCAN	~28.9 (×4 Set14)	~0.79	Краща деталізація, ніж EDSR	Низька (дуже глибока, сотні шарів)	16	Низькі	~800 зобр. (DIV2K)
ESRGAN	~26.3 (×4 Set14)	~0.78	Дуже чітке, насичене текстурами зображення	Помірна (16М модель, прийнятна на GPU)	16.7	Помірні	~3450 зобр. (DIV2K+Flickr2K)
Real-ESRGAN	Низький (не оптимізов. на PSNR)	Низький	Висока якість на реальних фото: усуває шум, артефакти стиснення, додає деталізацію	Помірна (≈ як ESRGAN з оптимізаціями)	~16	Помірні	~800+ зобр. різн. деград.
DnCNN	~29.4 (σ=25 шум)	~0.88	Ефективно прибирає шум, зберігаючи краї	Висока (реальний час на GPU; ~0.1с CPU)	0.56	Низькі	~400 зобр. (BSD, синт. шум)
BM3D	~28.6 (σ=25 шум)	~0.87 ¹	Знижує рівень шуму, але трохи розгладжує дрібні структури; ледь помітні блокові артефакти	Невисока (≈1 с для 512×512 на CPU; без GPU-прискорення)	0	Низькі	Немає (алгоритм. метод)

Практична ефективність. Отримані результати висвітлюють, що вибір моделі залежить від сценарію використання. Якщо критично важливо зберегти максимальну точність (наприклад, в медичних зображеннях або для подальшої обробки), перевагу має PSNR-орієнтований метод (RCAN), який відтворює дрібні деталі без значних помилок. Його недоліком є дещо згладжений вигляд – текстури можуть виглядати плоскими. Коли ж на перший план виходить перцептуальна якість – наприклад, фотореалістичність для кінцевого глядача – більш доцільно застосувати ESRGAN або його похідні, що генерують деталі (траву, волосся тощо) ціною менших PSNR. В реальних умовах, де характер спотворень невідомий (розмитість фокуса, шум матриці, JPEG-стиснення), найбільш ефективним виявився «blind SR» підхід (Real-ESRGAN): він продемонстрував високу стійкість – здатність покращувати зображення при різних комбінованих деградаціях, на яких традиційні моделі різко втрачають якість. Таким чином, проведений порівняльний аналіз підтвердив, що жоден єдиний показник не характеризує «кращість» моделі однозначно: PSNR/SSIM відображають точність відновлення, тоді як візуальні спостереження та модельовані метрики сприйняття відображають перцептивну якість та реалістичність результату. Поєднання кількісних (PSNR, SSIM) і якісних критеріїв у таблиці дозволяє оцінити практичну придатність моделей: від швидких легковагих рішень для простих завдань до складних GAN-моделей для фотореалістичного покращення зображень у найвимогливіших сценаріях.

ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

У підсумку дослідження було здійснено аналіз провідних сучасних методів покращення якості зображень на основі глибоких нейронних мереж. Порівняння таких архітектур, як SRCNN, ESRGAN, DnCNN та RCAN, підтвердило ефективність глибокого навчання у задачах надроздільної здатності, шумозаглушення й відновлення деталей.

Перспективними векторами подальшого розвитку є оптимізація згорткових структур задля зменшення обчислювальних витрат, а також інтеграція механізмів уваги задля вдосконалення узагальнюючих можливостей. Важливим кроком залишається адаптація моделей до специфічних доменів, зокрема медичних та супутникових зображень, а також впровадження слабонаглядних і безнаглядних методів навчання, що дає

зможу суттєво знизити залежність від великих розмічених вибірок. Загальні результати дослідження свідчать про актуальність і перспективність подальшого розвитку глибоких нейромережових підходів до підвищення якості візуальних даних.

Література

1. Липовий А. Застосування штучного інтелекту для підвищення якості фотофіксації. Матеріали конференції «Комп'ютерні ігри та мультимедіа як інноваційний підхід до комунікації», ОНТУ, 2024, 293-294с.
2. Ledig, C., Theis, L., Huszár, F. et al. (2017). Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4681–4690.
3. Білий М. Д. Дослідження методів обробки зображень. Методи реалістичного заміщення об'єктів. Радіоелектроніка та молодь у XXI столітті : матеріали 28-го Міжнар. молодіж. форуму. – Харків : ХНУРЕ, 2024. – Т. 6 – 472-474с.
4. Kupyn, O., Martyniuk, T., Wu, J., Wang, Z. (2019). DeblurGAN-v2: Deblurring (Orders-of-Magnitude) Faster and Better. In IEEE International Conference on Computer Vision (ICCV), pp. 8878–8887.

References

1. Lypovyi, A. (2024). *Application of artificial intelligence for improving the quality of photo documentation*. In Proceedings of the Conference "Computer Games and Multimedia as an Innovative Approach to Communication", ONTU, pp. 293–294.
2. Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., & Shi, W. (2017). *Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network*. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4681–4690.
3. Bilyi, M. D. (2024). *Research of image processing methods. Methods of realistic object replacement*. In *Radioelectronics and Youth in the 21st Century: Proceedings of the 28th International Youth Forum*, Vol. 6, Kharkiv: KhNURE, pp. 472–474.
4. Kupyn, O., Martyniuk, T., Wu, J., & Wang, Z. (2019). *DeblurGAN-v2: Deblurring (Orders-of-Magnitude) Faster and Better*. In *IEEE International Conference on Computer Vision (ICCV)*, pp. 8878–8887.