

УДК 004.932.4:77.05]:004.89

ДОСЛІДЖЕННЯ СУЧАСНИХ МЕТОДІВ ПОЛІПШЕННЯ ЯКОСТІ ФОТОЗОБРАЖЕНЬ, ЗРОБЛЕНИХ У ПОГАНИХ УМОВАХ, З ВИКОРИСТАННЯМ ШТУЧНОГО ІНТЕЛЕКТУ

Верєпа Д.С.

e-mail: denys.verepa@nure.ua

Науковий керівник – канд. техн. наук, ст. викл. Кобилін І.О.
Харківський національний університет радіоелектроніки, каф. ІНФ
м. Харків, Україна

This work examines AI methods for enhancing images captured in poor conditions – low light, noise, blur. It reviews major deep-learning approaches, including CNNs, GANs, and Retinex-inspired models. A hybrid attention-based UNet is highlighted, reaching about 28.4 dB PSNR and 0.88 SSIM on benchmark datasets. The work compares supervised, zero-shot, and unsupervised strategies, noting their key trade-offs. It also discusses PSNR/SSIM and the gap between metric scores and real downstream-task performance. The main goal is to identify the most effective real-time enhancement methods for edge devices.

Формування якісних зображень є фундаментальною проблемою для багатьох застосунків комп'ютерного зору. В умовах низької освітленості, сильного шуму, розмиття та некоректної експозиції цифрові камери формуються зображення, що суттєво відрізняються від реальної сцени. Такий спад якості перешкоджає як суб'єктивному сприйняттю, так і коректній роботі алгоритмів розпізнавання, навігації або діагностики. Традиційні методи попередньої обробки (наприклад, вирівнювання гістограми, корекція гами чи класичні варіації Retinex) можуть частково покращити контраст або яскравість, проте не враховують складні просторові залежності та часто залишають шум і артефакти. Тому останнім часом зростає інтерес до методів поліпшення якості зображень на основі штучних нейронних мереж.

Одним із найпоширеніших підходів є пряме навчання нейронної мережі від низькоякісного до високоякісного зображення. Глибокі згорткові мережі (CNN) та різноманітні архітектури типу «кодер–декодер» використовуються для відновлення деталей, видалення шуму й підвищення контрасту. Наприклад, дослідження [1] на основі модифікованої UNet, доповненої механізмами уваги, демонструє можливість досягнути пікового відношення сигнал/шум (PSNR) близько 28,4 дБ та індексу структурної подібності (SSIM) $\approx 0,88$ на наборах даних LOL і SID при обсязі моделі $\sim 1,3$ млн параметрів. Такі моделі навчаються за допомогою комбінації перцептивних та структурних функцій втрат, а тренувальні пари формуються як зі справжніх низькоякісних фотографій, так і зі штучно створених зображень шляхом моделювання недостатньої експозиції та шуму.

Кількісна оцінка результатів покращення здійснюється за допомогою метрик якості. Найпоширенішою є PSNR (Peak Signal-to-Noise Ratio), що виражає співвідношення між максимальною потужністю сигналу та потужністю шуму. Вона обчислюється через середньоквадратичну похибку (MSE) між оригінальним зображенням f та відновленим g :

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (f_{ij} - g_{ij})^2 \quad (1)$$

де f_{ij} – значення інтенсивності (яскравості) пікселя в оригінальному зображенні; g_{ij} – значення інтенсивності пікселя у відновленому зображенні; m і n – висота та ширина зображення:

$$PSNR = 20 \cdot \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right) \quad (2)$$

де MAX_I – максимальне можливе значення інтенсивності пікселя (наприклад, 255 для 8-bit).

Велике значення PSNR означає меншу похибку та кращу реконструкцію. Для обліку особливостей сприйняття людини використовують також індекс структурної подібності SSIM, що оцінює схожість двох зображень за компонентами яскравості, контрасту та структури. SSIM обчислюється локально і усереднюється по всьому зображенню, що дозволяє враховувати як глобальні, так і локальні характеристики.

У роботі [2] наведено огляд сучасних методів підвищення якості зображень у темних сценах. Автори підкреслюють, що хоча керовані (supervised) моделі забезпечують високу зорову якість, вони дають лише помірне підвищення точності в наступних задачах комп'ютерного зору; навпаки, методи zero-shot та безучкового навчання, попри нижчі значення PSNR/SSIM, часто покращують роботу систем сегментації, детекції чи класифікації. Крім того, в огляді відзначено перспективність доменної адаптації, яка дає значні покращення на сегментаційних датасетах при відсутності розмічених даних, що свідчить про розрив між існуючими метриками якості і реальними потребами.

Іншу групу методів становлять моделі, що використовують фізично обґрунтовану теорію Retinex, за якою зображення розкладається на компоненти освітленості та відбиття. Нещодавнє рішення RetinexMamba [3] комбінує традиційний підхід з архітектурою на основі Mamba та замінює механізм багато-головної уваги на сплавну увагу. Така модель зберігає інтерпретованість, підвищує швидкість обробки та перевершує попередні Retinex-орієнтовані мережі на датасеті LOL як за кількісними, так і за якісними показниками. Висновки цього дослідження підтверджують, що

інтеграція фізичних припущень і сучасних глибоких мереж є перспективною для збалансування якості та швидкодії.

Окремий напрямок спрямований на підвищення просторової роздільної здатності. Суть супер-резолюції полягає у відтворенні зображення високої чіткості з його низькороздільного варіанту. Застосування цієї технології охоплює відеоспостереження, медичну діагностику, відновлення старих фотографій та мікроскопію. Сучасні методи SISR реалізуються на основі згорткових мереж, трансформерів, генеративно-змагальних мереж або дифузійних моделей і навчаються відтворювати деталі, що були згладжені або втрачено. Відповідно до оглядового дослідження [4], різні підкласи відрізняються складністю, вимогами до даних та можливістю узагальнення, а дослідники звертають увагу на необхідність балансування між точністю та швидкістю.

Загалом, методи покращення якості зображень на основі штучних нейронних мереж демонструють значний прогрес порівняно з класичними підходами. Вони здатні видаляти шум, відновлювати дрібні деталі та підвищувати контраст у складних умовах зйомки. Водночас кожна група методів має свої обмеження: керовані моделі потребують великих датасетів, генеративні мережі можуть породжувати артефакти, а нові архітектури вимагають оптимізації для роботи на ресурсно обмежених пристроях. Перспективними напрямками є легковагові мережі з урахуванням фізичних моделей, підходи zero-shot і доменна адаптація, а також впровадження дифузійних та трансформерних моделей. Подальший розвиток цього напрямку сприятиме появі більш універсальних та ефективних систем підвищення якості зображень для реальних застосувань.

Список використаних джерел:

1. Manuel J. C. S. Reis. Low-Light Image Enhancement Using Deep Learning: A Lightweight Network with Synthetic and Benchmark Dataset Evaluation. Applied Sciences. 2025. URL: <https://www.mdpi.com/2076-3417/15/11/6330>
2. Liu F., Fan L. A review of advancements in low-light image enhancement using deep learning. Neurocomputing. 2025. URL: <https://arxiv.org/abs/2505.05759>
3. Bai J., Yin Y., He Q., Li Y., Zhang X. RetinexMamba: Retinex-based Mamba for Low-Light Image Enhancement. 2024. URL: <https://arxiv.org/abs/2405.03349>
4. Zhang L., Li A., Hou Q., Zhu C., Eldar Y. Deep Learning Empowered Super-Resolution: A Comprehensive Survey and Future Prospects. 2025. URL: <https://arxiv.org/html/2509.22692v1>