



ПОКРАЩЕННЯ ЯКОСТІ ЗОБРАЖЕНЬ ЗАСОБАМИ НЕЙРОННИХ МЕРЕЖ ПРИ РІЗНИХ ТИПАХ СПОТВОРЕНЬ

Ю. Романишин¹ [ORCID: 0000-0003-2794-432X], С. Єлманов² [ORCID: 0000-0002-6097-9527], М. Дуркот [ORCID: 0000-0003-4147-3077],
О. Теглівець [ORCID: 0000-0003-3880-2159], В. Мельник [ORCID: 0009-0004-0739-6842], А. Юреня¹ [ORCID: 0009-0009-2488-1213]

¹ Національний університет “Львівська політехніка”, вул. С. Бандери, 12, Львів, 79013, Україна
² Спеціальне конструкторське бюро телевізійних систем, вул. В. Великого, 4, Львів, 79026, Україна

Відповідальний за рукопис: Юрій Романишин (e-mail: yurii.m.romanyshyn@lpnu.ua).

(Подано 10 травня 2025)

В статті розглянуто завдання покращення якості зображень засобами нейронних мереж за різних типів спотворень кількох рівнів – погіршення контрасту, додавання різних шумів, компресія зображень тощо. Як базу зображень для навчання нейронних мереж використано поширену базу TID2013, що містить як оригінальні зображення, так і змінені за допомогою різних типів спотворень (25 базових зображень, 24 типи спотворень та 5 їх рівнів). Цю базу розділено на навчальні (480 зображень), перевірні (валідаційні, 360) та тестові (120) зображення. Для покращення вибрано зображення з найбільшим рівнем спотворень 5. Для дослідження завдання покращення якості зображень використано модифікацію нейронної мережі MIRNet та нейронну мережу на основі лише згорткових шарів. Наведено особливості побудови та параметри нейронних мереж, структуру згорткової нейронної мережі, графіки втрат та точності під час навчання, часові параметри навчання. Модифікація мережі MIRNet полягала у налаштуванні процедури формування навчальних та валідаційних зображень на основі бази TID2013, зміні деяких параметрів нейронної мережі. Кількість шарів модифікованої нейронної мережі дорівнює 1339 (ураховуючи вхідний шар, шари Concatenate, Pooling2D, Reshape тощо). Основними шарами є згорткові шари Conv2D. Як міру втрат під час навчання нейронної мережі використано середньоквадратичну похибку, а як міру точності пікове співвідношення сигнал / шум, а також середню абсолютну похибку. В нейронній мережі на основі згорткових шарів використано сім згорткових шарів з кількістю каналів 128 та 256 (крім останнього згорткового шару із трьома каналами для отримання результуючого покращеного зображення). Середня абсолютна похибка для навчальної вибірки становила 0,0491, а для валідаційної 0,0527. Використано згорткові шари з однаковою кількістю каналів у кожному. Не спостерігався ефект перенавчання, у разі використання нерегулярної структури згорткових шарів (наприклад, з 32 та 128 каналами) відзначалося явище перенавчання. Для комп'ютерних експериментів із навчання нейронних мереж використано графічний відеоадаптер NVIDIA GeForce RTX 4080.

Ключові слова: покращення якості зображень, нейронна мережа, CNN, MIRNet, TID2013, Keras.

УДК: 004.93+004.8

Вступ

Якісне розв'язання різних задач опрацювання зображень, таких як розпізнавання зображень та розпізнавання елементів на цих зображеннях, класифікація, кластеризація тощо, може потребувати попереднього покращення якості цих зображень. Задачі попереднього поліпшення якості

зображень виникають під час дистанційного зондування Землі, опрацювання супутникових знімків, у медицині тощо. Складнощі застосування відомих формальних методів та алгоритмів покращення зображень зумовлюють значний інтерес до застосування нейронних мереж для розв'язання такого класу задач. Широкі можливості нейронних мереж з різними структурами шарів, різною кількістю нейронів, різними функціями активації нейронів, різними методами та алгоритмами навчання нейронних мереж забезпечують потенційні можливості ефективного розв'язання задач поліпшення якості зображень. Водночас доцільно враховувати причини погіршення якості зображень (наприклад, низький рівень освітлення, наявність шумів, різних типів спотворень) та рівень спотворення. У зв'язку з цим для побудови та тестування нейронних мереж покращення якості зображень необхідні бази із великою кількістю зображень та різних типів спотворень. Одна з таких баз зображень – поширена база TID2013 з різними типами спотворень та різними їх рівнями. Незважаючи на значну кількість публікацій щодо покращення якості зображень, наявні методи та програмні засоби, переважно, орієнтовані на певний тип спотворень або обмежений клас типів спотворень.

2. Аналіз публікацій за темою покращення якості зображень засобами нейронних мереж

Із тематики нейромережевого поліпшення якості зображень, зокрема покращення їх контрасту, опубліковано чимало робіт. Ці публікації умовно можна розділити на кілька груп.

Оглядові статті, в яких аналізують та порівнюють різні підходи до застосування нейронних мереж (першочергово глибоких нейронних мереж) для покращення якості зображень

У [1] наведено короткий огляд методів покращення якості зображень, зокрема засобами глибоких нейронних мереж. Розглянуто основні причини погіршення якості зображень (низький рівень освітлення під час отримання зображень, шуми різних типів тощо), деякі класичні методи покращення якості (контрасту) зображень (вирівнювання гістограми та його модифікацій, використання різних типів фільтрів, лінійні та нелінійні розтяги тощо), порівняння кількох моделей глибоких нейронних мереж та результатів їх використання для оцінювання якості зображень за різних метрик. Зазначено, що одними із останніх сфер застосування нейронних мереж в опрацюванні зображень були саме задачі нейромережевого поліпшення їх якості. Однак в роботі немає детальнішої інформації щодо структур та кількісних параметрів аналізованих глибоких нейронних мереж.

Статті з використанням зміни розрізняювальної здатності зображень

В [2] розглянуто згорткову нейронну мережу для поліпшення якості зображення. На початковому етапі здійснюється зменшення розрізняювальної здатності зображень, які надходять на вхід згорткових шарів нейронної мережі, а на кінцевому етапі розрізняювальна здатність вихідних зображень повертається до початкових значень. Наведені експериментальні результати підтверджують істотне покращення якості зображень порівняно із кількома іншими методами на низці стандартних наборів зображень. Крім того, використана структура нейронної мережі з невеликою кількістю каналів зумовлює порівняно невеликі обчислювальні затрати під час навчання нейронної мережі, що дає змогу використати таку нейронну мережу в мобільних пристроях. Однак, як правило, суттєве покращення якості зображень потребує значних обчислювальних затрат.

Публікації зі стандартним використанням згорткових нейронних мереж

Основними результатами в [3] є деякі висновки щодо експериментів, виконаних із використанням згорткових нейронних мереж (ЗНН, CNN – Convolutional Neural Network) та оператора Лапласа для поліпшення півтонових зображень, що дало змогу підвищити швидкість навчання. Крім того, під час тонкого налаштування попередньо навченої моделі CNN використано традиційні методи обмеженого адаптивного вирівнювання гістограми (CLAHE), вейвлет-перетворення.

Використання комбінованих архітектур нейронних мереж

В [4] розглянуто архітектуру згорткової нейронної мережі з урахуванням як локальних, так і глобальних особливостей для покращення зображення. Для оброблення як локальних, так і глобальних функцій використано архітектуру, яка складається з трьох мереж: локального кодера, глобального кодера і декодера. Отримані експериментальні результати порівняно з використовуваними методами покращення якості зображень.

Приклади використання нейронних мереж у пакеті Keras

У пакеті Keras [5] є п'ять прикладів [6–10] програмної реалізації покращення якості кількох класів зображень. У [6] розглянуто приклад застосування глибокого згорткового автокодера (декодера) для усунення шумів на зображенні (Convolutional autoencoder for image denoising). Використано відому базу зображень рукописних десяткових цифр MNIST, в які попередньо внесено адитивні нормально розподілені шумові завади. Нейронна мережа містить три згорткові шари (Conv2D), два згорткові транспонувальні шари (conv2d_transpose, Conv2DTr) (які дають змогу реалізувати функції, обернені до функцій згорткових шарів), два об'єднувальні шари max_pooling2d. Вхідні дані – зображення малого розміру (півтонові зображення 28 на 28 пікселів). Тому загальна кількість параметрів для навчання невелика і становить 28353. Використано функції активації RELU, у разі застосування згорткових шарів – опцію **same** для отримання зображення того самого розміру. Недоліком цієї нейронної мережі є спеціалізація лише на одному типі спотворення – зашумлені зображення, малі розміри зображень, прості півтонові зображення (зображення цифр). Завдяки транспонувальним згортковим шарам розміри вихідних зображень однакові із розмірами вхідних.

В [7, 11] описано нейронну мережу ESPCN (Efficient Sub-Pixel CNN), яка реконструює версію зображення з високою розрізнявальною здатністю на основі версії з низькою розрізнявальною здатністю. У мережі використано ефективні шари “субпіксельної згортки”, які навчаються як набори фільтрів масштабування зображення.

В [8, 12] побудована нейронна мережа MIRNet, в якій вхідне зображення передається також на вихід нейронної мережі, де додається до даних, отриманих перетвореннями нейронної мережі, як це звичайно використовують у “залишковій” (Residual) нейронній мережі. Нейронна мережа містить три послідовні “залишкові” блоки RRG (Recursive Residual Group) RRG1, RRG2, RRG3, в кожному з яких використовується група з двох мультимасштабних “залишкових” блоків MRB (Multiscale Residual Block) MRB1, MRB2. Мультимасштабність полягає в паралельному використанні оригінального масштабу та масштабування в два та чотири рази зі зменшенням розрізнявальної здатності. В блоках MRB реалізовано блоки просторової “уваги” (Spatial Attention) та “уваги” на рівні каналів згорткових шарів (Channel Attention).

В [9, 13] реалізовано глибоку “залишкову” нейронну мережу для покращення якості зображень із використанням підвищення розрізнявальної здатності зображення. Архітектура мережі ґрунтується на архітектурі SRResNet (Super Resolution Residual Network) і складається з кількох “залишкових” блоків. Незвичним є використання замість функції втрати L2 (середня квадратична похибка) функції втрати L1 (середня абсолютна похибка). Нейронна мережа містить 16 “залишкових” блоків з 64 каналами.

В [10, 14] завдання покращення зображення в умовах слабкого освітлення сформульовано як завдання оцінювання специфічних для зображення “тональних кривих”, які використовують для коригування динамічного діапазону цього зображення. На вході нейронної мережі задають зображення, отримані за слабкого освітлення, а на виході отримують коригувальні “тональні криві”, які надалі використовують для попиксельного налаштування динамічного діапазону вхідного сигналу та одержання зображення покращеної якості. Цей метод близький до коригування кривих, використовуваного в програмному забезпеченні для редагування фотографій, наприклад, Adobe Photoshop.

У наведених варіантах нейронних мереж для покращення якості зображень ефективність їх роботи була проілюстрована лише на певних типах спотворення зображень – отриманих за низьких рівнів освітлення, із додаванням до зображень певних типів шумів тощо. Отже, можна зробити висновок про актуальність побудови та дослідження нейронних мереж для покращення якості зображень за наявності різних типів спотворень.

Мета статті – представлення деяких варіантів нейронних мереж та дослідження результатів їх навчання та функціонування при використанні бази зображень з різними типами спотворень та різними їх рівнями.

3. База зображень TID2013 для дослідження нейронних мереж покращення їх якості

Як і для задач оцінки якості зображень, для дослідження нейронних мереж покращення якості зображень була використана база TID2013 [15, 16], особливостями якої є її цільове призначення для оцінки та порівняння якості 25 еталонних зображень та зображень, отриманих за допомогою 24 типів спотворень з 5 рівнями цих спотворень. Крім того, важливим є наявність числових експертних оцінок якості зображень. Дані для проведення числових експериментів та дослідження особливостей нейронних мереж покращення якості зображень структуровані таким чином. Навчальна та валідаційна (перевірні) вибірки містять 480 зображень, з яких 360 зображень є навчальними, а 120 зображень валідаційними. Використовуються не всі зображення бази TID2013 для всіх типів спотворень, а лише зображення з найбільшими рівнями спотворення та відповідні їм еталонні зображення. Навчальна вибірка містила вхідні зображення для їх покращення i01_01_5.bmp, ..., i01_24_5.bmp, i02_01_5.bmp, ..., i02_24_5.bmp, ..., i15_01_5.bmp, ..., i15_24_5.bmp (15 базових зображень з 24 типами спотворень з найбільшим рівнем спотворень 5). Відповідні їм еталонні зображення i01.bmp, i02.bmp, ..., i15.bmp утворюють вихідний масив (з 24 копіями відповідно до кількості типів спотворень). Валідаційна група зображень містила 120 вхідних зображень для візуальної та числової оцінки покращення їх якості i21_01_5.bmp, ..., i21_24_5.bmp, i22_01_5.bmp, ..., i22_24_5.bmp, ..., i25_01_5.bmp, ..., i25_24_5.bmp (5 базових зображень з 24 типами спотворень з найбільшим рівнем спотворень 5) та відповідні їм еталонні зображення i21.bmp, i22.bmp, ..., i25.bmp.

4. Використані апаратні та програмні засоби для побудови, навчання та дослідження функціонування нейронних мереж покращення якості зображень

Практично всі задачі нейромережевого опрацювання зображень, в тому числі і ті, які використовуються для оцінки та покращення їх якості, потребують значних обчислювальних ресурсів, в першу чергу тих, які надають змогу паралельно виконувати частину операцій, зокрема, тензорних операцій. Для побудови, навчання та дослідження особливостей функціонування нейронних мереж в цій статті використовувався центральний процесор 12th Gen Intel(R) Core(TM) i7-12700KF при тактовій частоті 3.60 ГГц, оперативна пам'ять 64,0 ГБ, відеоадаптер NVIDIA GeForce RTX 4080. Наявність графічного процесора дає можливість суттєво підвищити ефективність процедур навчання глибоких нейронних мереж.

Для побудови нейронної мережі та тестування задач покращення якості зображень використовувалися програмні бібліотеки Keras та TensorFlow [5, 17].

5. Побудова та використання нейронних мереж для дослідження покращення якості зображень бази TID2013

Використання модифікованої нейронної мережі MIRNet

Одним з варіантів нейронної мережі для покращення якості зображень була модифікація мережі MIRNet [8, 12]. Модифікація мережі полягала у налаштуванні процедури формування навчальних та валідаційних зображень на основі бази TID2013, зміні деяких параметрів нейронної мережі.

Кількість шарів нейронної мережі дорівнює 1339 (включаючи вхідний шар, шари Concatenate, Pooling2D, Reshape тощо). Основними шарами є згорткові шари Conv2D. Формат вхідних зображень має вигляд (4, 128, 128, 3), розмір блоків (4) кольорових зображень (остання вісь тензора 3) при навчанні з випадковим вибором блоків розміром 128 на 128 пікселів. Вхідними зображеннями є сформовані з TID2013 360 навчальних зображень високої та низької якості, 120 валідаційних зображень високої та низької якості. З врахуванням блокування по 4 зображення це 90 навчальних наборів зображень і 30 валідаційних.

Кількість каналів в базових згорткових шарах дорівнювала 64, кількість налаштовуваних параметрів 44 157 007.

Як міра втрат (loss) при навчанні нейронної мережі використовувалася середньоквадратична похибка mse (Mean Square Error):

$$mse = \frac{1}{N} \sum_{i=1}^N (y_i - p_i)^2 \quad (1)$$

де y_i – значення елементів, отримувані на виході шарів нейронної мережі; p_i – значення елементів, задані як еталонні; N – кількість елементів.

Як міра точності в параметрах нейронної мережі було вибране пікове співвідношення сигнал/шум у децибелах між двома зображеннями I_1 та I_2 PSNR (Peak Signal-to-Noise Ratio), як використано в [8]:

$$PSNR = 10 \log_{10} \frac{\max(R)^2}{MSE} \quad (2)$$

де (в термінах зображень):

$$MSE = \frac{\sum_{m,n} [I_1(m,n) - I_2(m,n)]^2}{M \times N} \quad (3)$$

M та N – розміри зображень; $R = 255$ – максимальне значення величини пікселів (при 8-бітовому кодуванні зображень).

В якості функції оптимізації була вибрана функція Adam.

На рис. 1, а зображено графіки залежності втрат від кількості епох в процесі навчання для навчальної та перевіркої (валідаційної) вибірки, а на рис. 1, б – аналогічна залежність PSNR.

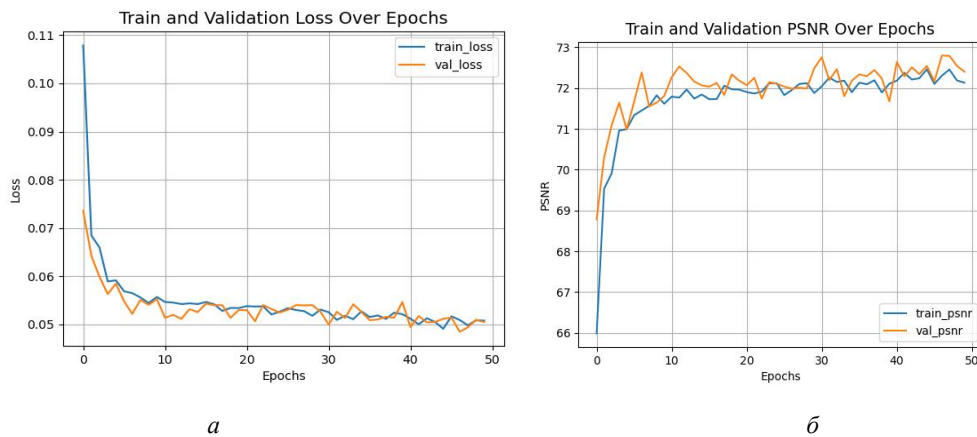


Рис. 1. Параметри втрат та точності для навчальних та перевірних зображень в процесі навчання

Як видно з наведених графіків навчання, при кількості ітерацій до 10 втрати швидко зменшуються, а PSNR швидко зростає, після чого ці процеси дуже сповільнюються. Час виконання однієї ітерації становив 32 сек.

Аналогічний характер залежностей був отриманий при базовій кількості каналів 32. Оскільки ці залежності практично не відрізнялися від залежностей на рис. 1, кількість каналів можна зменшити в порівнянні з використаними в [8] значеннями 64. При цьому суттєво зменшується кількість параметрів для навчання: з 44 157 007 до 11 059 295. Значно зменшився також і час виконання однієї ітерації, який дорівнював 22 сек.

Інша широко використовувана метрика точності (асигасу) в параметрах навчання нейронних мереж (звичайно використовується в задачах регресії) *mae* (Mean Absolute Error):

$$mae = \frac{1}{N} \sum_{i=1}^N |y_i - p_i| \quad (4)$$

з параметрами, аналогічними співвідношенню (1).

При цьому величини точності PSNR та втрат *mse* можуть бути перераховані одна в одну.

На рис. 2, *а* та 2, *б* представлені залежності *mse* та *mae* для навчальної та валідаційної множини зображень в процесі навчання від кількості ітерацій для зображень бази TID2013.

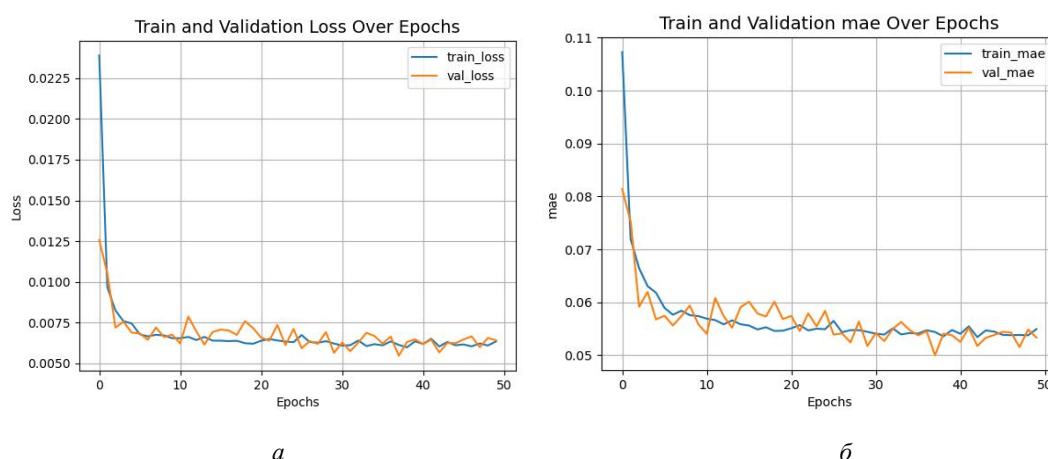


Рис. 2. Параметри *mse* та *mae* для навчальних та перевірних зображень в процесі навчання

Побудова та використання нейронної мережі лише на основі згорткових шарів

В табл. 1 наведені структура та параметри нейронної мережі, побудованої лише на основі згорткових шарів з кількістю каналів 128.

Таблиця 1

**Структура та параметри нейронної мережі
на основі згорткових шарів з кількістю каналів 128**

Параметри нейронної мережі			
Шари нейронів	Функції активації	Формат вихідних даних	К-сть парам.
Conv2D (128, (3, 3)) input_shape = (384, 512, 3)	relu	(None, 384, 512, 128)	3584
Conv2D (128, (3, 3))	relu	(None, 384, 512, 128)	147584
Conv2D (128, (3, 3))	relu	(None, 384, 512, 128)	147584
Conv2D (128, (3, 3))	relu	(None, 384, 512, 128)	147584
Conv2D (128, (3, 3))	relu	(None, 384, 512, 128)	147584
Conv2D (128, (3, 3))	relu	(None, 384, 512, 128)	147584
Conv2D (3, (3, 3))	relu	(None, 384, 512, 3)	3459
			Σ: 744963

Структура вхідних даних: (8, 384, 512, 3). З врахуванням розміру блоків по 8 розмір навчальних блоків становить 45, розмір валідаційних блоків дорівнює 15.

На рис. 3, а, б представлено графіки залежності втрат для навчальних (mse) та валідаційних (val mse) даних та точності для навчальних (mae) та валідаційних (val mae) даних від номера ітерацій в процесі навчання нейронної мережі.

Час виконання однієї ітерації становив 13–14 с, результати на останній ітерації: loss = 0,0052; mae = 0,0491; val_loss = 0,0056; val_mae = 0,0527.

Як видно з цих графіків, аналогічно модифікованій програмі MIRNet, після 10 кроків ітерацій втрати та точність змінюються дуже повільно, параметри точності mae та val_mae близькі до аналогічних параметрів програми MIRNet. Зміна валідаційних параметрів втрат та точності з номером ітерацій є більш плавною, ніж для MIRNet.

При використанні нерегулярної структури згорткових шарів (наприклад з 32, 128 каналами) спостерігалось явище перенавчання, при якому точність для валідаційних зображень була значно нижчою, ніж точність для навчальних зображень.

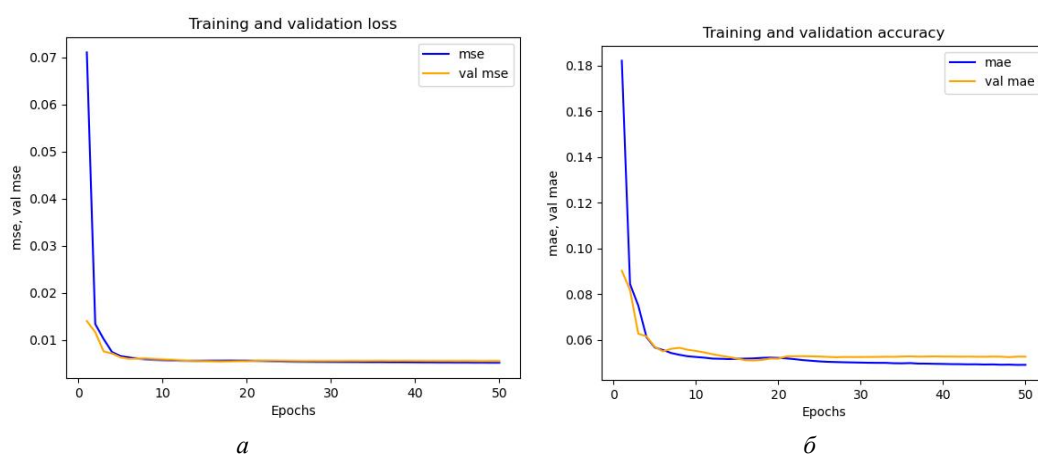


Рис. 3. Параметри втрат та точності для навчальних та перевірних зображень в процесі навчання

В табл. 2 наведені структура та параметри нейронної мережі, побудованої лише на основі згорткових шарів з кількістю каналів 256.

Таблиця 2

**Структура та параметри нейронної мережі
на основі згорткових шарів з кількістю каналів 256**

Параметри нейронної мережі			
Шари нейронів	Функції активації	Формат вихідних даних	К-сть парам.
Conv2D (256, (3, 3)) input_shape=(384, 512, 3)	relu	(None, 384, 512, 256)	7168
Conv2D (256, (3, 3))	relu	(None, 384, 512, 256)	590080
Conv2D (256, (3, 3))	relu	(None, 384, 512, 256)	590080
Conv2D (256, (3, 3))	relu	(None, 384, 512, 256)	590080
Conv2D (256, (3, 3))	relu	(None, 384, 512, 256)	590080
Conv2D (256, (3, 3))	relu	(None, 384, 512, 256)	590080
Conv2D (3, (3, 3))	relu	(None, 384, 512, 3)	6915
			Σ: 2964483

На рис. 4 представлено залежності втрат та точності для навчальної та валідаційної множини зображень від номера ітерації.

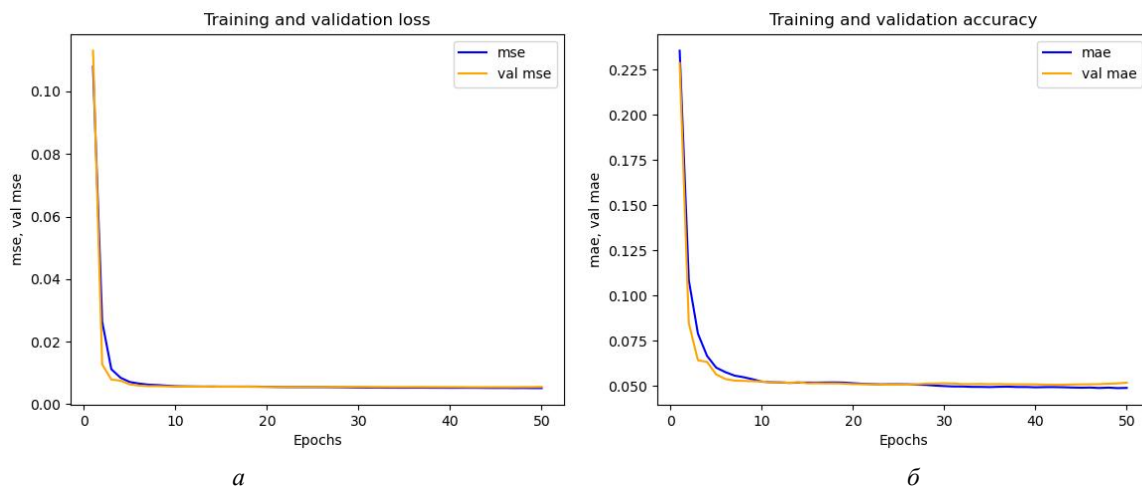


Рис. 4. Параметри mse та mae для навчальних та перевірних зображень під час навчання з кількістю каналів 256

Час виконання однієї ітерації становив близько 54 с, результати на останній ітерації: $loss = 0,0052$, $mae = 0,0490$, $val_loss = 0,0055$, $val_mae = 0,0519$.

Висновки

Нейронні мережі є важливим універсальним апаратним та програмним інструментом для опрацювання зображень, зокрема, для покращення їх якості перед використанням цих зображень для завдань розпізнавання, класифікації, кластеризації зображень тощо. Процедури навчання нейронних мереж передбачають необхідність використання відповідних баз зображень з різними типами спотворень та різними їх рівнями. Серед наявних баз зображень вибрано базу TID2013 з 25 базовими зображеннями, 24 типами спотворень та п'ятьма їх рівнями. Відомі реалізації нейронних мереж звичайно орієнтовані на покращення конкретних типів спотворень, причому, наприклад, MIRNet, відзначаються великою складністю, у них використано комбіновані структури (згорткові шари, “залишкові” (residual) структури, блоки “з увагою” (attention), різні масштаби з покращенням (або погіршенням) розрізняювальної здатності тощо). На підставі обчислювальних експериментів показано, що порівняно істотно простіші регулярні нейронні структури на основі лише згорткових шарів можуть забезпечити отримання результатів, близьких до одержаних за допомогою значно складніших структур нейронних мереж. Подальші можливості розвитку нейронних мереж для покращення якості зображень пов'язані із оптимізацією структур мереж (кількість шарів нейронів, кількість нейронів у шарах, параметри згорткових шарів, функції активації, функції оптимізації тощо) із використанням генетичних та еволюційних алгоритмів.

Список використаних літературних джерел

- [1] Kaur, A., and Gagandeep (2018), “A Review on Image Enhancement with Deep Learning Approach”, *ACCENTS Transactions on Image Processing and Computer Vision*, Vol. 4 (11), pp. 16–20. DOI: 10.19101/TIPCV.2018.411002
- [2] Vu, T., Nguyen, C. V., Pham, T. X., Luu, T. M., and Yoo, C. D. (2019), “Fast and Efficient Image Quality Enhancement via Desubpixel Convolutional Neural Networks”, *LNCS, Springer, Cham*, Vol. 11133, pp. 243–259. DOI: 10.1007/978-3-030-11021-5_16
- [3] Chen, X. (2019), “Image Enhancement Effect on the Performance of Convolutional Neural Networks”, *Thesis for the degree of Master of Science in Computer Science, Blekinge Institute of Technology, Sweden*, 40 p. Available: <https://www.diva-portal.org/smash/get/diva2:1341096/FULLTEXT02.pdf> (accessed 08.05.2025)

- [4] Kinoshita, Y., and Kiya, H. (2019), "Convolutional Neural Networks Considering Local and Global Features for Image Enhancement". Available: <https://arxiv.org/pdf/1905.02899>, 5 p. (accessed 08.05.2025). DOI: 10.48550/arXiv.1905.02899
- [5] <https://keras.io/> (accessed 08.05.2025)
- [6] Convolutional Autoencoder for Image Denoising. Available: <https://keras.io/examples/vision/autoencoder/> (accessed 08.05.2025)
- [7] Image Super-Resolution Using an Efficient Sub-Pixel CNN. Available: https://keras.io/examples/vision/super_resolution_sub_pixel/ (accessed 08.05.2025)
- [8] Low-Light Image Enhancement Using MIRNet. Available: <https://keras.io/examples/vision/mirnet/> (accessed 08.05.2025)
- [9] Enhanced Deep Residual Networks for Single-image Super-Resolution. Available: <https://keras.io/examples/vision/edsr/> (accessed 08.05.2025)
- [10] Zero-DCE for Low-Light Image Enhancement. Available: https://keras.io/examples/vision/zero_dce/ (accessed 08.05.2025)
- [11] Shi, W., Caballero, J., Huszar, F., Totz, J., Aitken, A. P., Bishop, R., Rueckert, D., and Wang, Z. (2016), "Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network". Available: <https://arxiv.org/pdf/1609.05158>, 10 pp. (accessed 08.05.2025). DOI: 10.48550/arXiv.1609.05158
- [12] Zamir, S. W., Arora, A., Khan, S., Hayat, M., Khan, F. S., Yang, M.-H., and Shao, L. (2020), "Learning Enriched Features for Real Image Restoration and Enhancement". Available: <https://arxiv.org/pdf/2003.06792>, 20 p. (accessed 08.05.2025). DOI: 10.48550/arXiv.2003.06792
- [13] Lim, B., Son, S., Kim, H., Nah, S., and Lee, K. M. (2017), "Enhanced Deep Residual Networks for Single Image Super-Resolution", Available: <https://arxiv.org/pdf/1707.02921>, 9 p. (accessed 08.05.2025). DOI: 10.48550/arXiv.1707.02921
- [14] Guo, C., Li, C., Guo, J., Loy, C. C., Hou, J., Kwong, S., and Cong, R. (2020), "Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement". Available: <https://arxiv.org/pdf/2001.06826>, 10 p. (accessed 08.05.2025). DOI: 10.48550/arXiv.2001.06826
- [15] Ponomarenko, N., Jin, L., Ieremeiev, O., Lukin, V., Egiazarian, K., Astola, J., Vozel, B., Chehdi, K., Carli, M., Battisti, F., and Kuo, C.-C. J. (2015), "Image database TID2013: Peculiarities, results and perspectives", *Signal Processing: Image Communication*, Vol. 30, January, pp. 57–77. DOI: 10.1016/j.image.2014.10.009
- [16] Ponomarenko, N., Ieremeiev, O., Lukin, V., Jin, L., Egiazarian, K., Astola, J., Vozel, B., Chehdi, K., Carli, M., Battisti, F., and Kuo, C.-C. J. (2013), "A new color image database TID2013: Innovations and results", *Proceedings of ACIVS, Poznan, Poland, Oct.*, pp. 402–413. DOI: 10.1007/978-3-319-02895-8_36
- [17] Chollet, F. *Deep Learning with Python*, 2nd ed., Manning Publications Co., 2021, 478 p.

IMAGE QUALITY ENHANCEMENT USING NEURAL NETWORKS FOR VARIOUS TYPES OF DISTORTION

**Yuriy Romanyshyn¹, Sergei Yelmanov², Markian Durkot, Oles Teglivets,
Volodymyr Melnyk, Andrii Yurenia¹**

¹ Lviv Polytechnic National University, 12, S. Bandery str., Lviv, 79013, Ukraine

² Special Design Office of Television Systems, 4, V. Velykoho st., Lviv, 79026, Ukraine

The paper considers the problem of enhancing image quality using neural networks with different types of multi-level distortions – contrast degradation, adding noise of various natures, image compression, etc. The widespread TID2013 database, which contains both original images and images modified using various types of distortions (25 basic images, 24 types of distortions, and 5 of their levels), was used as the image database for training neural networks. This database was divided into training (480 images), validation (360), and test (120) images. The images with the

highest level of distortion, 5, were selected for enhancement. To study the problem of image quality enhancement, a modification of the MIRNet neural network and a neural network based only on convolutional layers were used. The features of the construction and parameters of neural networks, the structure of the convolutional neural network, loss and accuracy graphs in the learning process, and time parameters of learning are presented. The modification of the MIRNet network consisted of adjusting the procedure for forming training and validation images based on the TID2013 database, and changing some parameters of the neural network. The number of layers of the modified neural network was 1339 (including the input layer, Concatenate, Pooling2D, Reshape layers, etc.). The main layers were Conv2D convolutional layers. The mean square error was used as a measure of loss when training the neural network, and the peak signal-to-noise ratio and the mean absolute error were used as measures of accuracy. In the neural network based on convolutional layers, 7 convolutional layers with the number of channels of 128 and 256 were used (except for the last convolutional layer with 3 channels to obtain the resulting enhanced image). The mean absolute error for the training images was 0.0491, and for the validation, 0.0527. Convolutional layers with the same number of channels in each layer were used. In this case, no overfitting effect was observed. When using an irregular structure of convolutional layers (for example, with 32 and 128 channels), the phenomenon of overfitting was observed. When conducting computer experiments on training of neural networks, an NVIDIA GeForce RTX 4080 graphics adapter was used.

Keywords: *image quality enhancement, neural network, CNN, MIRNet, TID2013, Keras.*