

УДК 004.8:004.93

DOI <https://doi.org/10.32782/2663-5941/2025.5.2/31>

Романишин Ю.М.

Національний університет «Львівська політехніка»

Єлманов С.О.

Спеціальне конструкторське бюро телевізійних систем

Дуркот М.Ю.

Національний університет «Львівська політехніка»

Теглівець О.В.

Національний університет «Львівська політехніка»

Мельник В.Р.

Національний університет «Львівська політехніка»

Юреня А.В.

Національний університет «Львівська політехніка»

ПОКРАЩЕННЯ ЯКОСТІ ЗОБРАЖЕНЬ МОДИФІКАЦІЯМИ НЕЙРОННОЇ МЕРЕЖІ MIRNET

У статті розглянуто задачу покращення якості зображень на основі використання нейронної мережі MIRNet та її модифікацій. Наведені структури та особливості функціонування окремих блоків мережі MIRNet з форматами вхідних та вихідних даних (тензорів) складових блоків: блоку зменшення та збільшення роздільної здатності проміжних активацій в 2 рази; блоку «канальної уваги» з встановленням вектора з максимальним впливом каналів (фільтрів); дуальної «канальної та просторової уваги». Модифікації нейронної мережі MIRNet полягали у налаштуванні операторів вводу даних використаної бази зображень TID2013 з іншими форматами зображень, використанні різної кількості багатомасштабних «залишкових» блоків (Multi-Scale Residual Blocks) та повторюваних «залишкових» груп (Recursive Residual Groups), використанні в якості метрики в процесі навчання, крім PSNR, середньої абсолютної похибки та деяких інших налаштувань загального характеру. В якості бази тестових зображень для дослідження особливостей навчання та функціонування нейронної мережі була вибрана база TID2013 з 25 еталонними зображеннями високої якості, отриманими на їх основі погіршеними зображеннями з 24 типами спотворень та 5 рівнями цих спотворень. Розглянуто 2 варіанти покращення зображень бази TID2013: 1) при навчанні мережі на зображеннях бази TID2013 (еталонних та з найвищим рівнем 5 спотворень) та оцінці якості зображень з цієї бази та 2) при навчанні нейронної мережі на базі зображень LoL dataset та оцінці якості зображень бази TID2013 за допомогою навченої таким чином нейронної мережі. При навчанні нейронної мережі були використані дві метрики якості зображень: PSNR (як в класичному варіанті використання мережі MIRNet) та MAE (середня абсолютна похибка). Наведені графіки похибок та метрик в залежності від номера ітерації в процесі навчання нейронної мережі для навчальних та перевірних (валідаційних) вибірок зображень. Для побудови, навчання та дослідження особливостей функціонування нейронних мереж 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.

Ключові слова: покращення якості зображень, нейронні мережі, глибоке навчання, згорткові нейронні шари, нейронні шари «з увагою», MIRNet, база зображень TID2013.

Постановка проблеми. Задача покращення якості зображень є однією з важливих складових загальної проблеми опрацювання зображень. Низька якість зображень може бути обумовлена низьким рівнем освітлення при їх отриманні, низьким рівнем контрастності, втратами якості при передачі чи компресії зображень, розмитістю внаслідок руху об'єкта зображення, наявністю шумів тощо. При цьому значно обмежуються можливості якісного розв'язання інших задач опрацювання зображень – розпізнавання та класифікація зображень, виявлення, розпізнавання та локалізація елементів на зображеннях, сегментація, колірна корекція тощо. Такі ситуації можливі при опрацюванні медичних зображень, аерофотозніманні, супутникових зображень та інших. У зв'язку з цим для таких зображень важливою є їх попередня обробка з метою покращення їх якості. Більшість відомих методів покращення зображень, зокрема, покращення їх контрастності, орієнтовані на числові параметри якості зображень і конкретні причини та особливості погіршення якості. В результаті такі методи ефективні лише для певних типів спотворень.

У зв'язку з цим широкого використання набули методи на основі нейронних мереж, в першу чергу глибоких нейронних мереж, та методів і процедур глибокого навчання. Ці технології покращення якості зображень звичайно не орієнтовані на використання формалізованих процедур перетворення зображень (фільтрації, методів покращення локального чи глобального контрасту), а реалізують процедури навчання при наявності еталонних зображень та їх екземплярів погіршеної якості. Серед відомих структур нейронних мереж для покращення якості зображень слід виділити нейронну мережу MIRNet (Multi-scale Image Restoration Network), яка є одним з прикладів, наведених на платформі Keras [1]. Однак ця нейронна мережа досліджувалася, в основному, на прикладах покращення зображень з низьким рівнем освітлення. У зв'язку з цим виникає актуальна задача дослідження використання нейронної мережі MIRNet та її модифікацій для покращення якості зображень з широкою різноманітністю типів спотворень, що розглядається в даній статті.

Аналіз останніх досліджень і публікацій. Нижче проведений аналіз особливостей нейронної мережі MIRNet, а також результатів її використання для покращення зображень певних типів та кількох баз зображень.

В [2] представлена нова архітектура нейронної мережі для відновлення та покращення якості

зображень. Основою нейронної мережі є багатомасштабний «залишковий» блок, що містить кілька ключових елементів: а) паралельні згорткові потоки з багатомасштабною роздільною здатністю для виділення багатомасштабних ознак; б) обмін інформацією між потоками з різною роздільною здатністю; в) просторові та каналні механізми уваги для виділення контекстної інформації та г) агрегацію ознак на основі уваги. Нейронна мережа, названа MIRNet, була протестована на п'яти базах тестових зображень: DND (Darmstadt Noise Dataset), SIDD (Smartphone Image Denoising Dataset), RealSR (Real-World Single Image Super-Resolution), LoL (LoW-Light dataset), MIT-Adobe FiveK.

В [3] розглядається архітектура нейронної мережі MPRNet (Multi-Stage Progressive Restoration Network), яка в значній мірі близька до архітектури MIRNet. Нейронна мережа містить такі ключові компоненти: 1) на попередніх етапах використовується кодер-декодер для навчання багатомасштабної контекстної інформації, тоді як останній етап працює з оригінальною роздільною здатністю зображення для збереження дрібних просторових деталей; 2) модуль керування увагою підключається між етапами, щоб забезпечити прогресивне навчання; 3) додано механізм міжетапного об'єднання ознак, який допомагає поширювати багатомасштабні контекстні ознаки від попередніх до пізніших етапів.

В [4] представлено архітектуру MIRNet-v2 з метою підтримки просторово точних представлень високої роздільної здатності отриманням додаткової контекстної інформації з представлень низької роздільної здатності. Експерименти, проведені на шести реальних наборах зображень, продемонстрували високі результати для різних завдань обробки зображень: усунення розмиття зображень (для бази DPDD – Dual-Pixel Defocus Deblurring), зниження рівня шумів (для баз DND та SIDD), покращення якості зображень з високою роздільною здатністю (для бази RealSR), покращення якості зображень при низькому рівні освітлення (для баз LoL та MIT-Adobe FiveK).

В [5] представлено метод, заснований на згортковій нейронній мережі, блоці подвійної уваги та синтезі ознак з різними ядрами згортки (SKFS – Selective Kernel Feature Synthesis), аналогічно мережі MIRNet, який поєднується з моделлю на основі теорії Retinex для покращення затемнених зображень, отриманих в умовах слабкого освітлення. Модель являє собою багатомасштабний залишковий блок, що склада-

ється з кількох важливих компонентів, еквівалентних згортковій нейронній мережі з архітектурою VGG16 та різними ядрами згортки. Обчислювальні експерименти проводилися на графічному процесорі Tesla T4 в середовищі Google Colab на наборах даних CID (College Image Dataset), ELD (Extreme Low-light Denoising Dataset) та ExDark. Розроблена нейронна мережа названа MSR-MIRNet (Multi-Scale Residual – Multi-scale Image Restoration Network) як модифікація нейронної мережі MIRNet. Нейронна мережа MSR-MIRNet показала кращі результати, ніж окремо нейронні мережі MSR та MIRNet.

В [6] використано кілька етапів покращення якості зображень, отриманих під водою. Спочатку оригінальне зображення попередньо обробляється алгоритмом балансу білого для корекції кольору, потім контрастність зображення покращується за допомогою методу покращення контрастності. Попередньо опрацьоване таким чином зображення передається нейронній мережі MIRNet для покращення. Якість отриманого зображення вимірюється за допомогою параметрів пікового співвідношення сигнал/шум (PSNR), середньоквадратичної похибки (RMSE) та індексу структурної подібності (SSIM).

В [7] пропонується метод покращення підводних зображень під назвою Underwater Image Colorfulness Enhancement MIRNet (UICE-MIRNet) для покращення видимості малих, численних, щільних об'єктів з подальшим виявленням підводних об'єктів за допомогою YOLOv4. UICE-MIRNet – це спеціалізована версія класичного MIRNet. UICE-MIRNet включає блок покращення колірності підводних зображень (UI-CEB – Underwater Image Colorfulness Enhancement Block). Цей блок дає змогу виділяти ділянки з низькою кольоровістю з підводних зображень та виконує корекцію кольору без впливу на контекстну інформацію. Числова ефективність методу оцінювалася на основі стандартних метрик UIQM (Underwater Image Quality Measure), UCIQE (Underwater Color Image Quality Evaluation Metric), ентропії та PSNR. Числові експерименти проводилися на основі двох баз зображень Brackish (Brackish Underwater Dataset) та Trash-ICRA19 і продемонструвати ефективність запропонованого методу порівняно з існуючими методами.

В [8] розглядається задача розпізнавання виразів обличчя (FER – Face Expression Recognition) при використанні чотирьох моделей попереднього покращення зображень – MIRNet, RetinexNet, Retinex та AGC (Adaptive Gamma Correction).

Найкращий результат за точністю розпізнавання (67,6%) в умовах слабкого освітлення при використанні попереднього покращення зображень був отриманий для комбінації MIRNet та ResNet50.

В [9] представлено поєднання нейронної мережі MIRNet-v2, вдосконаленої методики попередньої обробки зображень, з автокодерами для зменшення шуму в зображеннях. Як параметри точності (якості вихідного зображення) використовувалися: PSNR, SSIM та MSE. Для навчання та тестування нейронної мережі використовувалися бази зображень LoL, SIDD, DND.

В [10] розглядаються приклади застосування архітектури глибокої нейронної мережі MIRNet, призначеної для покращення якості зображень. Наведені результати 19 експериментів з використанням наборів даних LoL та SICE для перевірки та оптимізації продуктивності MIRNet. Показана ефективність MIRNet та деяких її архітектурних модифікацій.

З аналізу цих та інших публікацій щодо нейронної мережі MIRNet можна зробити висновки, що дослідження ефективності використання нейронної мережі MIRNet та її модифікацій для покращення якості зображень базувалося, в основному, на тестових базах зображень з обмеженою кількістю типів спотворень, що обумовлює доцільність дослідження нейронної мережі при розширенні типів спотворень зображень.

Метою статті є дослідження особливостей використання нейронної мережі MIRNet та деяких її модифікацій для покращення якості зображень бази TID2013, яка містить базові зображення високої якості та їх спотворені варіанти 24-х типів.

Виклад основного матеріалу. Нейронна мережа MIRNet виділяється серед інших нейронних мереж покращення якості зображень тим, що вона містить значну кількість різноманітних структур, що в сукупності утворюють ефективний інструмент для отримання якісних зображень з їх низькоякісних варіантів.

Як і для різних інших задач опрацювання зображень, широко використовуються згорткові структури, які забезпечують зниження обчислювальних затрат при навчанні нейронної мережі. Потік даних вхідних зображень розпаралелюється на потоки з основною та зменшеною в 2 та 4 рази роздільною здатністю. Структура блоку зменшення роздільної здатності (**down sampling module**) в 2 рази зображена на рис. 1.

Блок містить два шляхи – основний і «залишковий» (residual), результати яких додаються на виході блоку. Вхідний тензор має формат

$[k \times m \times n \times c]$, де k – розмір блоку, m та n – ширина та висота каналу (зображення), c – кількість каналів (в прикладі на рис. 1 $[4 \times 128 \times 128 \times 64]$). Основний шлях блоку містить 3 згорткові шари **Conv2D** з розмірами вікон (1, 1), (3, 3) та (1, 1), функціями активації **relu** зі збільшенням кількості каналів в 2 рази в останньому шарі та шар **MaxPooling2D** зі зменшенням розмірів каналів в 2 рази. Залишковий блок містить шар **MaxPooling2D** зі зменшенням розмірів каналів в 2 рази та згортковий шар **Conv2D** з розміром вікна (1, 1) зі збільшенням кількості каналів в 2 рази. Результати двох шляхів додаються та отримується вихідний тензор формату $[k \times m/2 \times n/2 \times 2c]$.

Структура блоку збільшення роздільної здатності (**up sampling module**) в 2 рази з відповідним зменшенням кількості каналів зображена на рис. 2.

На вхід блоку подається тензор з форматом $[k \times m \times n \times c]$ (в прикладі на рис. 2 $[4 \times 128 \times 128 \times 64]$). Вхідний потік розпаралелюється на 2 шляхи: основний і залишковий. Основний шлях містить 3 згорткові шари **Conv2D** з розмірами вікон (1, 1), (3, 3) та (1, 1), функціями активації **relu** та шаром **UpSampling2D** збільшення роздільної здатності в 2 рази, в останньому згортковому шарі кількість каналів зменшується в 2 рази. Залишковий шлях містить шар **UpSampling2D** збільшення розділь-

ної здатності в 2 рази та згортковий шар **Conv2D** з розміром вікна (1, 1) та зменшенням кількості каналів в 2 рази. Результати двох шляхів додаються з утворенням тензора формату $[k \times 2m \times 2n \times c/2]$ (в прикладі на рис. 2 $[4 \times 256 \times 256 \times 32]$). На виході нейронної мережі результати паралельних шляхів з різною роздільною здатністю приводяться до початкової роздільної здатності.

З метою підвищення ефективності навчання та функціонування нейронної мережі **MIRNet** використовуються блоки подвійної уваги (**Dual Attention Unit – DAU**) для виділення ознак у згорткових потоках. Ці блоки виділяють більш інформативні ознаки. Використовуються два варіанти блоків уваги – з просторовою та каналною увагою. Вхідні дані, які поступають на вхід блоку **DAU**, мають вигляд тензора $[k \times m \times n \times c]$ (наприклад $[4 \times 128 \times 128 \times 3]$ для вхідного зображення при початковому масштабі, при подвійному масштабі $[4 \times 256 \times 256 \times 3]$ та $[4 \times 64 \times 64 \times 3]$ при зменшеному масштабі).

Структура блоку каналної уваги (**channel attention block**) зображена на рис. 3.

Операції блоку зводяться до наступного. Блок каналної уваги містить шар **GlobalAveragePooling2D**, який усереднює всі значення карти вхідних ознак $[k \times m \times n \times c]$ (в даному прикладі $[4 \times 128 \times 128 \times 64]$) по кожному

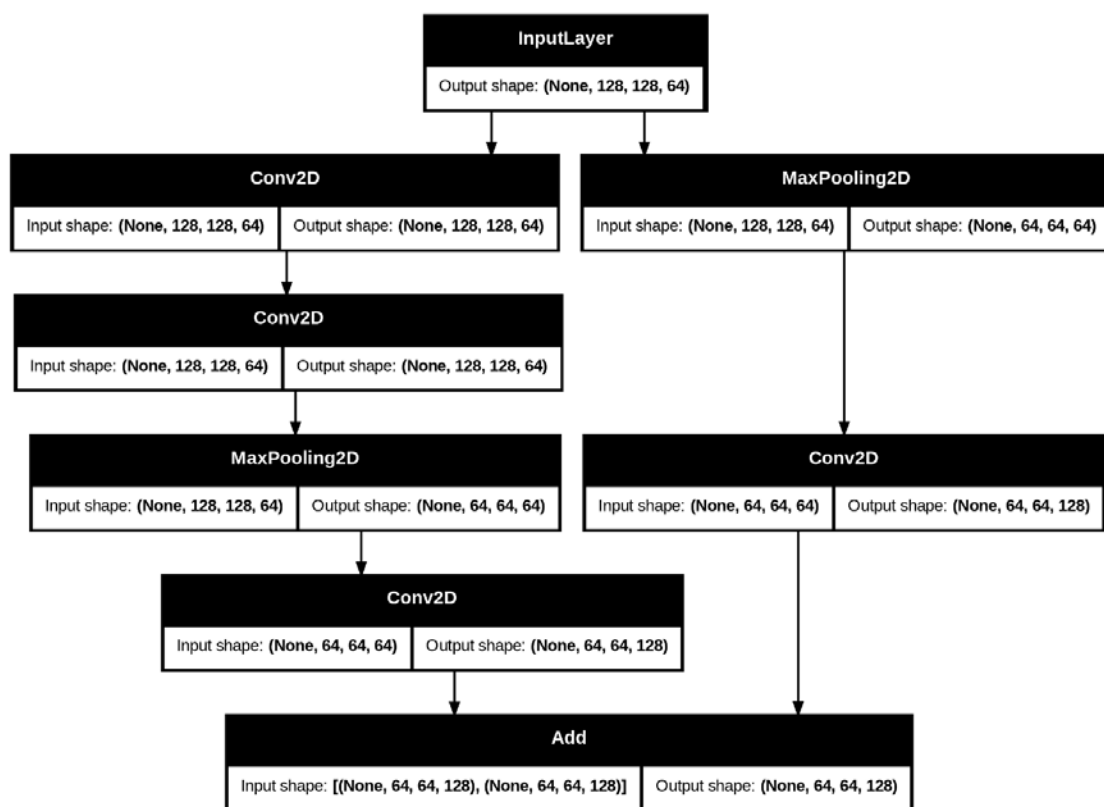


Рис. 1. Структура блоку зниження роздільної здатності в 2 рази

каналу. Це приводить до формату $[k \times c]$. Оператор **Reshape (1, 1, channels)** приводить дані до формату $[k \times 1 \times 1 \times c]$, необхідного для подальшого узгодження форматів. Наступні 2 згорткові шари з розмірами вікон (1, 1), функціями активації **relu** та **sigmoid**, зменшенням кількості каналів у 8 разів та відновленням кількості каналів до кількості каналів вхідного тензора формують множники в межах від 0 до 1 відповідно до середніх значень карт ознак кожного каналу. Ці множники виділяють операцією множення найбільш важливі значення карти ознак вхідного тензора за важливістю усередненого значення каналу.

Блок просторової уваги (**spatial attention block**) працює аналогічним чином, однак замість множників у вигляді вектора з розміром кількості каналів для блоку каналної уваги, в блоці просторової уваги множники мають матричну структуру $[m \times n]$ відповідно до важливості кожного елемента вхідної карти ознак по всіх каналах.

Блок дуальної уваги об'єднує виходи блоків каналної та просторової уваги за допомогою шару **Concatenate**. Структура цього блоку зображена на рис. 4. Об'єднання виходів обох блоків уваги приводить до збільшення кількості каналів в 2 рази, наступний згортковий шар **Conv2D** з розміром вікна (1, 1) зменшує кількість каналів

вдвічі, що дає змогу шаром **Add** додати результати блоку дуальної уваги до вхідного тензора.

Модифікації нейронної мережі MIRNet полягали у налаштуванні операторів вводу даних використаної бази зображень TID2013 з іншими форматами зображень, використанні різної кількості багатомасштабних залишкових блоків (Multi-Scale Residual Block) та повторюваних залишкових груп (Recursive Residual Group), використанні в якості метрики в процесі навчання, крім PSNR, середньої арифметичної похибки (MAE) та деяких інших налаштувань загального характеру.

Використання бази зображень TID2013 для тестування нейронної мережі MIRNet та її модифікацій щодо покращення якості зображень. Для дослідження особливостей функціонування нейронної мережі MIRNet та її модифікацій щодо покращення якості зображень зі спотвореннями різних типів була використана база зображень TID2013 [11]. Ця база містить 25 еталонних зображень високої якості та отриманих на їх основі погіршених зображень за допомогою 24 різних процедур спотворень (всі зображення мають розмір 512×384 пікселів в форматі **.bmp**). Для кожного типу спотворень зображення є 5 рівнів. Таким чином, база зображень TID2013 містить 25 еталонних та 3000 зображень

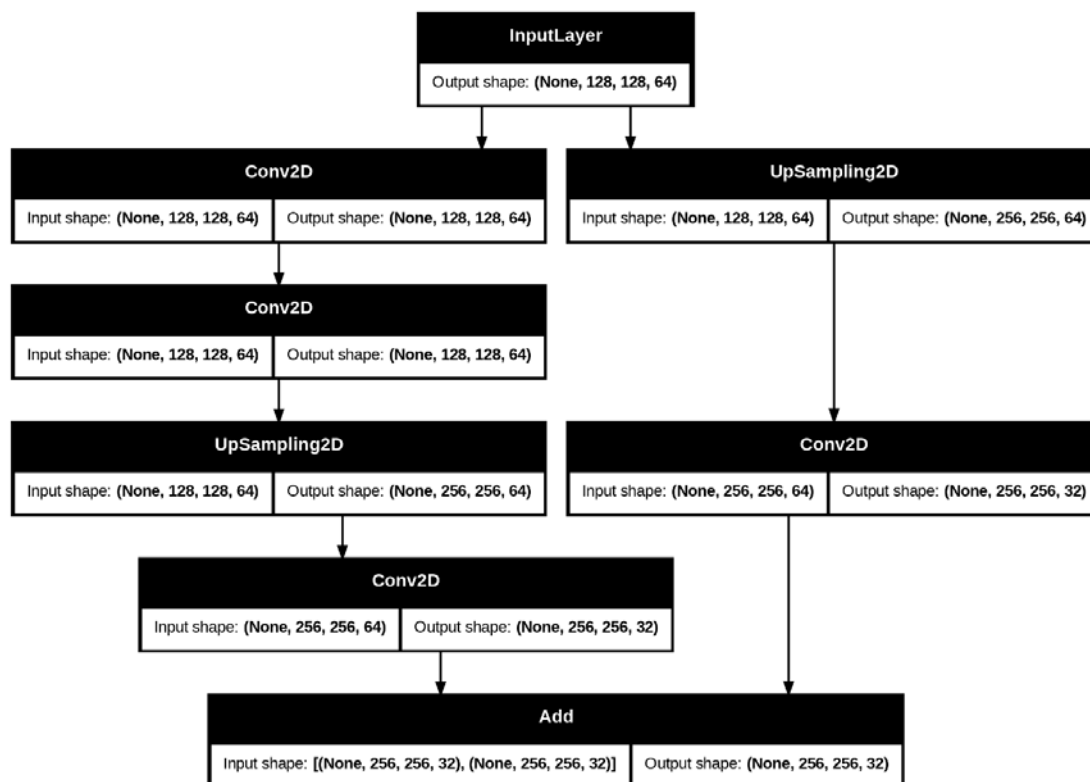


Рис. 2. Структура блоку збільшення роздільної здатності в 2 рази

погіршеної якості. Для всіх зображень погіршеної якості наведені експертні оцінки їх якості.

Для проведення числових експериментів використано 480 зображень з найгіршою якістю (рівнем спотворень 5), з яких 360 зображень утворюють навчальну вибірку, а 120 зображень є перевірними (валідаційними). Крім того, в процесі навчання використано відповідні їм еталонні зображення. До навчальної вибірки ввійшли вхідні зображення (в позначеннях бази TID2013) з найбільшим рівнем спотворень i01_01_5.bmp, ..., i01_24_5.bmp, ..., i15_01_5.bmp, ..., i15_24_5.bmp, а також відповідні їм вихідні еталонні зображення i01.bmp, ..., i15.bmp.

До валідаційної вибірки ввійшли вхідні зображення i21_01_5.bmp, ..., i21_24_5.bmp, ..., i25_01_5.bmp, ..., i25_24_5.bmp, а також відповідні їм вихідні еталонні зображення i21.bmp, ..., i25.bmp.

Було реалізовано 2 варіанти покращення якості зображень. В першому варіанті навчання та тестування нейронної мережі проводилося на навчальних вибірках з відображенням результатів для валідаційних вибірок бази зображень TID2013.

Були використані також 2 варіанти міри втрат при навчанні нейронної мережі та метрики якості зображень. В першій парі як міра втрат *loss* при навчанні нейронної мережі використовувалася величина, пов'язана з середньоквадратичною похибкою *mse* (Mean Square Error):

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

де y_i та x_i – отримане та еталонне значення на виході шарів нейронної мережі; N – кількість елементів.

Як метрика використовувалося значення *PSNR* (Peak Signal-to-Noise Ratio), як в класичному варіанті нейронної мережі MIRNet:

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

де $R = 255$ – максимальне значення пікселів зображення.

В другій парі мірою втрат була величина *mse*, а метрикою була *mae* (Mean Absolute Error) – середня абсолютна похибка:

$$mae = \frac{1}{N} \sum_{i=1}^N |y_i - x_i|. \quad (3)$$

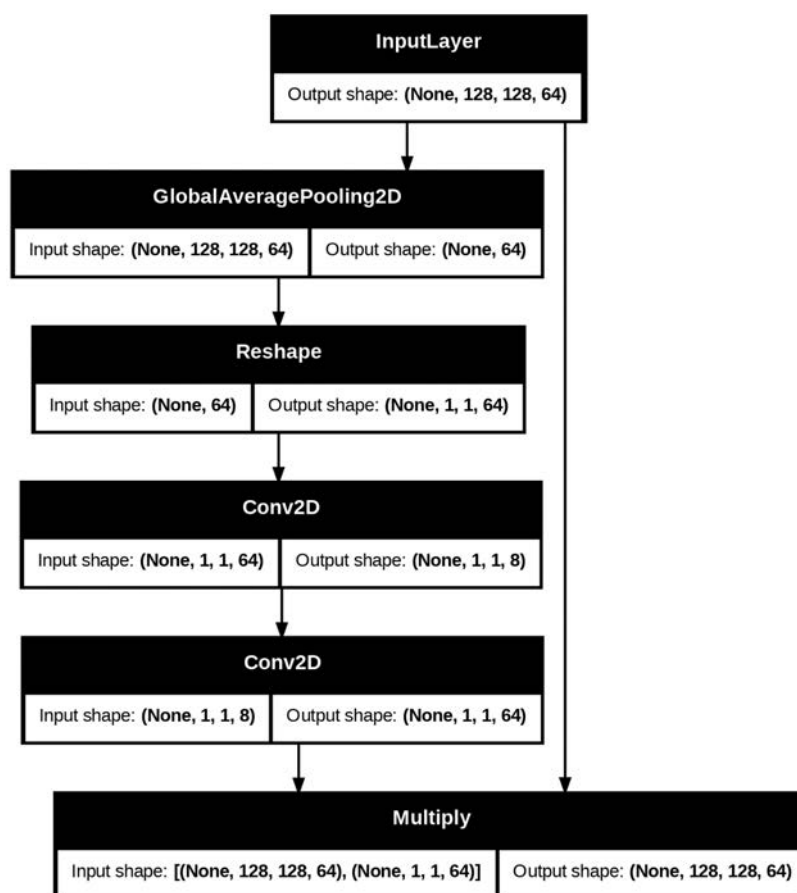


Рис. 3. Структура блоку каналної уваги (channel attention block)

При проведенні обчислювальних експериментів навчання нейронних мереж використовувалися відеоадаптер NVIDIA GeForce RTX 4080, центральний процесор 12th Gen Intel(R) Core(TM) i7-12700KF при тактовій частоті 3.60 ГГц, оперативна пам'ять 64,0 ГБ. Крім того, для побудови нейронних мереж покращення якості зображень використовувалися програмні бібліотеки Keras та TensorFlow.

На рис. 5 представлено залежності втрат та метрики PSNR від номера ітерації в процесі навчання нейронної мережі на базі зображень TID2013, а на рис. 6 подано аналогічні залежності для втрат *mse* та метрики *mae*.

Як видно з цих графіків, похибка навчання достатньо швидко зменшується протягом перших 10 ітерацій, значення PSNR швидко зростає, після чого коливається в межах 72-73 як для навчальних вибірок, так і для перевірних. Це свідчить про відсутність ефекту перенавчання та можливості покращення результатів. Час виконання однієї ітерації становив 32-33 сек. Кількість шарів нейронної мережі дорівнює 1339 (включаючи вхідний шар, шари **Concatenate**, **Pooling2D**, **Reshape** тощо), кількість налаштовуваних параметрів 44 157 007.

Наступні експерименти полягали у навчанні нейронної мережі на зображеннях бази LoL (переважно отриманих при низькому рівні освітлення) та оцінці якості перевірних зображень бази TID2013 за допомогою навченої таким чином мережі. При тій же кількості ітерацій 50 в процесі навчання нейронної мережі на зображеннях бази LoL отримане середнє значення PSNR для перевірних зображень становило 57. Таким чином, навчання на навчальних зображеннях та перевірка на валідаційних зображеннях однієї й тієї ж бази TID2013 дає вищу точність, ніж навчання на зображеннях однієї бази LoL та тестування на зображеннях іншої бази TID2013.

Висновки. На підставі отриманих результатів можна зробити такі висновки. Нейронна мережа MIRNet, завдяки комплексному використанню різних блоків опрацювання вхідних зображень



Рис. 4. Структура блоку дуальної уваги

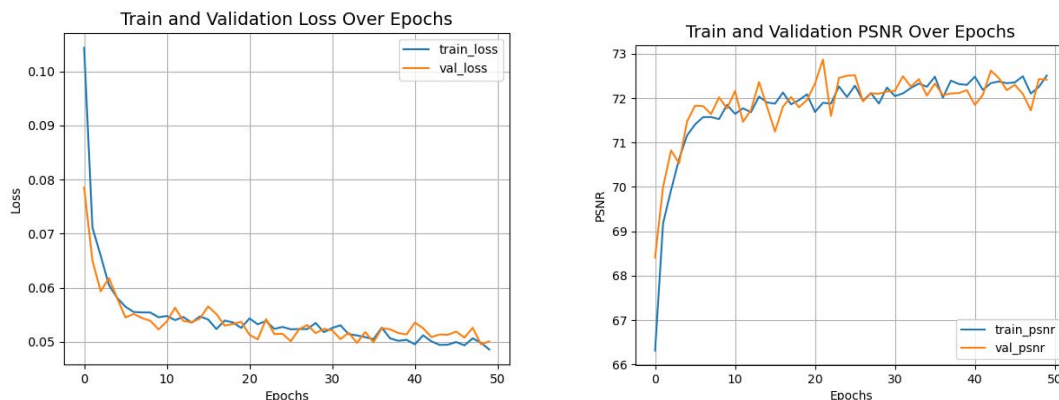


Рис. 5. Залежність втрат (loss) та метрики PSNR від номера ітерацій в процесі навчання зображень бази TID2013

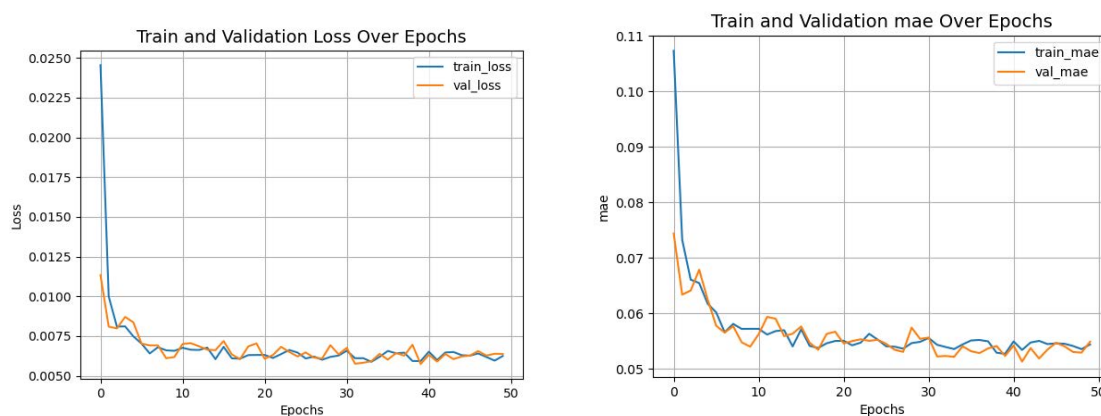


Рис.6. Залежність втрат *mse* та метрики *mae* від номера ітерацій в процесі навчання зображень бази TID2013

(використання блоків різномасштабних потоків обробки даних, використання залишкових механізмів в паралельних потоках обробки, застосування методу дуальної уваги, що поєднує просторову та каналну увагу з встановленням та використанням множників, які підсилюють відповідні фрагменти тензорів, тощо), може достатньо ефективно покращувати якість зображень. Однак відомі результати тестування цієї нейронної мережі стосувалися переважно баз зображень з невеликою кількістю типів спотворення зобра-

жень. Використання в якості бази зображень TID2013 показало можливості нейронної мережі MIRNet та її модифікацій ефективно покращувати якість зображень зі значною кількістю типів спотворень (24). При використанні для навчання зображень бази TID2013 отримані кращі результати (вища точність) в порівнянні з навчанням на класичних зображеннях бази LoL, отриманих з низьким рівнем освітлення, та застосуванні таким чином навченої мережі для перевірних зображень бази TID2013.

Список літератури:

1. Low-Light Image Enhancement Using MIRNet. URL: <https://keras.io/examples/vision/mirnet/> (дата звернення: 20.08.2025).
2. Zamir S.W., Arora A., Khan S., Hayat M., Khan F.S., Yang M.-H., Shao L. Learning Enriched Features for Real Image Restoration and Enhancement. URL: <https://arxiv.org/pdf/2003.06792> (дата звернення: 20.08.2025).
3. Zamir S.W., Arora A., Khan S., Hayat M., Khan F.S., Yang M.-H., Shao L. Multi-Stage Progressive Image Restoration. URL: <https://arxiv.org/pdf/2102.02808> (дата звернення: 20.08.2025).
4. Zamir S.W., Arora A., Khan S., Hayat M., Khan F.S., Yang M.-H., Shao L. Learning Enriched Features for Fast Image Restoration and Enhancement. URL: <https://arxiv.org/pdf/2205.01649> (дата звернення: 20.08.2025).

5. Das Mou T., Alam S.B., Rahman H., Srivastava G., Hasan M., Uddin M.F. Multi-Range Sequential Learning Based Dark Image Enhancement with Color Upgradation. *Applied Sciences*. 2023. 13(2). 1034. P. 2–14. DOI: 10.3390/app13021034.
6. Thamarai M., Aruna S.P., Sonti K., Chakravarthi P.S. Underwater Image Enhancement Using MIRNet. *Journal of Electronic & Information Systems*. 2023. Vol. 5. Issue 1. P. 36–44. DOI: 10.30564/jeis.v5i1.5600.
7. Sarkar P., De S., Gurung S., Dey P. UICE-MIRNet guided image enhancement for underwater object detection. *Scientific Reports*. 2024. 14 (1). DOI: 10.1038/s41598-024-73243-9.
8. Theofilus R.B., Kusuma G.P. Improving Low-Light Face Expression Recognition Using Combination of Mirnet and Resnet50 Models. *Journal of Theoretical and Applied Information Technology*. 2023. Vol. 101. No 5. P. 1791–1805.
9. Rath S., Priyadarshini S.B.B., Sahu P.K., Panda M., Patel D.K., Jagadev N., Sahoo S., Patra N. Enhancing Visual Content Quality in Education and Religion: a Fusion Approach of Mirnet-v2 and Autoencoders for Image Preprocessing and Noise Reduction. *Journal of Theoretical and Applied Information Technology*. 2024. Vol. 102. No. 15. P. 5783–5793.
10. Chen E., Yasrab R., Saha P. Using MIRNet for Low Light Image Enhancement. *Proceedings of the 18th International Joint Conference on Biomedical Engineering Systems and Technologies* (Porto, Portugal, 20-22 February 2025). Vol. 1. P. 280–287. DOI: 10.5220/0013099600003911.
11. Ponomarenko N., Jin L., Ieremeiev O., Lukin V., Egiazarian K., Astola J., Vozel B., Chehdi K., Carli M., Battisti F., Kuo C.-C.J. Image database TID2013: Peculiarities, results and perspectives. *Signal Processing: Image Communication*. 2015. Vol. 30. P. 57–77. DOI: 10.1016/j.image.2014.10.009.

Romanyshyn Yu.M., Yelmanov S.O., Durkot M.Yu., Teglivets O.V., Melnyk V.R., Yurenia A.V. IMAGE QUALITY ENHANCEMENT USING NEURAL NETWORK MIRNET MODIFICATIONS

The paper discusses the image quality enhancement task based on the MIRNet neural network and its modifications. The structures and functional features of individual MIRNet blocks are presented, along with the input and output data formats (tensors) of the block components: a block for reducing and increasing the resolution of intermediate activations by a factor of 2; a “channel attention” block with the establishment of a vector with the maximum influence of channels (filters); dual “channel and spatial attention”. Modifications to the MIRNet neural network consisted of configuring the data input operators of the TID2013 image database used with other image formats, using different numbers of multi-scale “residual” blocks and recursive residual groups, using as metrics in the training process, in addition to PSNR, mean absolute error, and some other general settings. The TID2013 database with 25 high-quality reference images was selected as the test image database for studying neural network training and functioning characteristics. Based on these images, degraded images with 24 types of distortions and 5 levels of these distortions were obtained. Two options for enhancing the images in the TID2013 database were considered: 1) training the network on images from the TID2013 database (reference images and images with the highest level 5 of distortion) and evaluating the quality of image enhancements from this database, and 2) training the neural network on images from the LoL dataset and evaluating the quality of images from the TID2013 database using the neural network trained in this way. Two image quality metrics were used when training the neural network: PSNR (as in the classic version of the MIRNet network) and MAE (mean absolute error). The graphs of the errors and metrics depending on the iteration number in the neural network training process for training and validation image samples are presented. To build, train, and study the features of neural networks, a 12th Gen Intel(R) Core(TM) i7-12700KF central processor with a clock speed of 3.60 GHz, 64.0 GB of RAM, and an NVIDIA GeForce RTX 4080 video adapter were used.

Key words: image quality enhancement, neural networks, deep learning, convolutional neural layers, attention-based neural layers, MIRNet, TID2013 image database.

Дата надходження статті: 27.08.2025

Дата прийняття статті: 11.09.2025

Опубліковано: 16.12.2025