

УДК 621.372.061

Рестаурація зображень методом super-resolution з використанням згорткових нейронних мереж

Недзельський О. Ю., Лащевська Н. О.

Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", м. Київ, Україна

E-mail: sashanedzel@gmail.com

Головна мета методу супер роздільної здатності (super-resolution) полягає у створенні зображення більш високої роздільної здатності з зображень нижчої роздільної здатності. Зображення високої роздільної здатності забезпечують високу щільність пікселів, отже, більше деталей на вихідному зображенні. Необхідність високої роздільної здатності широко поширена у методах комп'ютерного зору, в програмах для розпізнавання образів або звичайного аналізу зображень. Проте зображення високої роздільної здатності не завжди є доступними. Це пов'язано з тим, що процеси перетворення та методи для обробки вимагають надпотужних процесів, тому і обладнання для отримання зображень високої роздільної здатності виявляється дорогим. Ці проблеми можуть бути подолані за допомогою алгоритмів обробки зображень, які є відносно недорогими, що призвело до появи концепції надрозв'язання. Це дає перевагу, тому що може коштувати дешевше, а існуючі системи візуалізації з низькою роздільною здатністю є достатньо доступними. Висока роздільна здатність має велике значення у медичній візуалізації для діагностики. Багато програм вимагають масштабування конкретної області зображення, при цьому висока роздільна здатність стає необхідною, наприклад, для спостереження, криміналістики та супутникової візуалізації. Наведений в роботі метод з використанням згорткової нейронної мережі для відтворення зображень супер роздільної здатності напряму виконує перетворення з низького зображення на зображення подібне до оригіналу. Щоб прискорити час виходу, запропонований метод виконує більшість обчислювальних операцій у просторі з низьким дозволом та при цьому зменшення дискретизації не призводить до втрати інформації. Головна задача роботи нейронної мережі полягає в реконструкції спотвореного зображення та пошуку ідеальної функції відтворення, по якій, власне, нейронна мережа простої структури створює якісні зображення з кращими показниками, таким як роздільна здатність, співвідношення сигнал/шум, менші часові витрати на відновлення зображення. Під час експерименту було визначено алгоритм, по якому запропонована нейронна мережа може реконструювати будь-яке зображення, з різними видами спотворень. Метод super-resolution був реалізований з використанням мови програмування python 3.6 та програмних модулів для згорткових нейронних мереж tensorflow та tensorlayer. Графічні дані співвідношення сигнал/шум, структурної подібності та графіки втрат були отримані за допомогою модулю tensorboardX.

Ключові слова: супер роздільна здатність; згорткові нейронні мережі; співвідношення сигналу до шуму; втрата MSE; втрата VGG; коефіцієнт зменшення вибірки; кодер; декодер; десубпіксель

DOI: [10.20535/RADAP.2023.91.79-86](https://doi.org/10.20535/RADAP.2023.91.79-86)

Вступ

У даній статті розглядається згорткова нейронна мережа для поліпшення якості зображення, яка може бути навчена як для збільшення роздільної здатності зображення, так і для поліпшення інформативності зображень, отримання точних і в той же час візуально приємних зображень.

Роздільна здатність зображення описує деталі зображення та визначає кількість пікселів на одиницю площі. Чим вища роздільна здатність, тим більше деталей зображення. Супер роздільна здатність (Super-Resolution, SR) – це процес реконструкції зображення з високою роздільною здатністю (high-resolution, HR) із заданого (вхідного) зображення з

низькою роздільною здатністю (low-resolution, LR). Super-Resolution є складною задачею, і вирішення цієї проблеми є активною областю дослідження в галузі комп'ютерного зору та глибинного навчання.

Процес відновлення високої роздільної здатності зображень з низькорозмірних зображень – це завдання відновлення втраченої інформації. Очевидно, що при застосуванні функції деградації ми отримуємо зображення LR із зображення HR. Одним із простих методів отримання даних LR є погіршення даних HR різними втратами та шумами. Маючи інформацію про точну функцію деградації, застосовуючи її зворотно до зображення LR, можна створити функцію відтворення та відновити зображення HR [1]. Зазвичай при вирішенні зворотної

задачі функція деградації наперед невідома. Проблематика Super-Resolution полягає в тому, що при зменшенні розміру зображення втрачається багато деталей та інформації. Якщо ми намагаємося відновити високоякісне зображення з низької роздільної здатності, то виникає проблема відновлення цих втрачених деталей.

Сьогодні для покращення якісних характеристик багатовимірних сигналів (зображень) широко використовують згорткові нейронні мережі [2–4], щоб відновлювати високоякісні зображення з низької роздільної здатності. Структура мережі – одно-

спрямована (без зворотних зв'язків), багатопшарова зображена на Рис. 1. Однак, процес відновлювання високоякісних зображень за допомогою згорткової нейронної мережі може бути пов'язаний з багатьма проблемами, такими як складність тренування мережі, погана якість відновлення та високі вимоги до обчислювальної потужності. Також можуть виникати проблеми, пов'язані з відновленням деталей, які не були присутні в низькороздільних зображеннях, і зі створенням вірогідних деталей, які відсутні на вихідному зображенні, але відповідають реальним деталям.

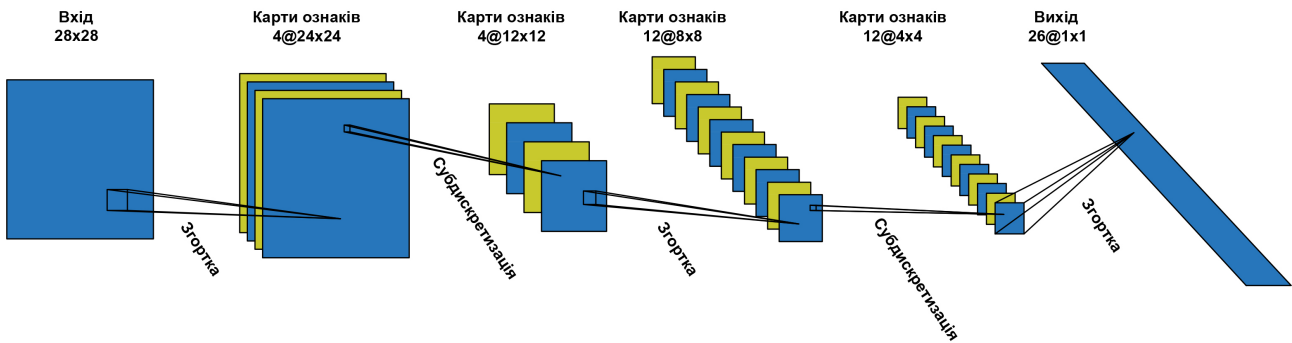


Рис. 1. Архітектура згорткової нейронної мережі

Для вирішення цих проблем було застосовано підхід глибокого навчання (deep learning) згорткової нейронної мережі.

1 Навчання згорткової нейронної мережі

Навчання функції наскрізного відображення F вимагає оцінки параметрів мережі

$$\theta = \{W_1, W_2, W_3, B_1, B_2, B_3\}. \quad (1)$$

Це досягається шляхом мінімізації втрат між відновленими зображеннями $F(Y; \theta)$ і відповідними зображеннями високої роздільної здатності X . Для набору зображень високої роздільної здатності $\{X_i\}$ та відповідних зображень із низькою роздільною здатністю $\{Y_i\}$ використовується середньоквадратична помилка (MSE) як функція втрат:

$$L(\theta) = \frac{1}{n} \sum_{i=1}^n F(Y_i; \theta) - X_i^2, \quad (2)$$

де n – кількість навчальних вибірок. Використання MSE як функції втрат сприяє високому PSNR (пікове співвідношення сигналу до шуму). PSNR є широко використовуваною метрикою для кількісної оцінки якості відновлення зображення. Незважаючи на те, що запропонована модель тренується на користь високого PSNR, під час навчання все

ще спостерігається задовільна продуктивність, коли модель оцінюється за допомогою альтернативних показників оцінки, наприклад, SSIM, MSSIM [5]. Втрати мінімізуються за допомогою стохастичного градієнтного спуску зі стандартним зворотним поширенням. Зокрема, вагові матриці оновлюються як:

$$\Delta_{i+1} = 0.9 \cdot \Delta_i - \eta \cdot \frac{\partial L}{\partial W_i^l}, \quad (3)$$

$$W_{i+1}^l = W_i^l + \Delta_{i+1}, \quad (4)$$

де $l \in \{1, 2, 3\}$ та i – індекси шарів та ітерацій, η – швидкість навчання, $\frac{\partial L}{\partial W_i^l}$ – похідна. Ваги фільтрів кожного шару ініціалізуються шляхом випадкового вибору гаусівського розподілу з нульовим середнім і стандартним відхиленням 0,001. Швидкість навчання становить 10^{-4} для перших двох шарів і 10^{-5} для останнього. Емпіричним шляхом встановлено, що менша швидкість навчання на останньому шарі важлива для сходження мережі (подібно до випадку зменшення шуму). На етапі навчання основні зображення істини X_i готуються як підзображення $f_{sub} \times f_{sub} \times c$ - піксель, випадково обрізані з навчальних зображень. Під «підзображеннями» мається на увазі, що ці зразки розглядаються як невеликі «зображення», а не «фрагменти», у тому сенсі, що «фрагменти» перекриваються і вимагають певного усереднення як постобробку, а «підзображення» не потребують. Щоб синтезувати вибірки з низькою роздільною здатністю Y_i , підзображення

розвиваються за допомогою ядра Гауса, підбірка за допомогою коефіцієнта збільшення, і масштабується воно на той самий фактор за допомогою бікубічної інтерполяції. Щоб уникнути граничних ефектів під час навчання, усі згорткові шари не мають заповнення, і мережа видає менший вихід $((f_{sub} - f_1 - f_2 - f_3 + 3)^2 \times c)$. Функція втрат MSE оцінюється тільки за різницею між центральними пікселями X_i та вихідним сигналом мережі. Навіть при використанні фіксованого розміру зображення під час навчання, згорткова нейронна мережа може бути застосована до зображень

довільного розміру під час тестування. Більшість методів створення, навчання та роботи згорткових нейронних мереж спрямовані на поліпшення лише таких якісних показників оцінки зображення як пікове співвідношення сигнал/шум і середнє значення ваг зв'язків без будь-якого врахування часу оброблення. В наслідок цього навіть до комп'ютерів високого класу потужності пред'являються високі обчислювальні вимоги, не кажучи вже про мобільні пристрої [6]. На Рис. 2 представлена архітектура запропонованого методу.

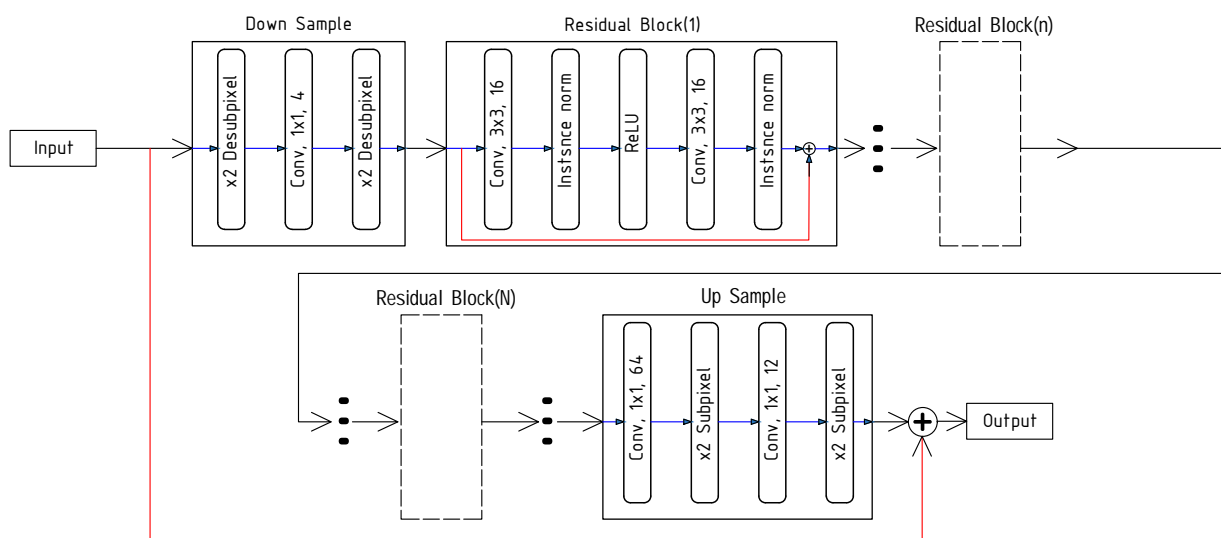


Рис. 2. Архітектура методу для реставрації зображень з використанням згорткових нейронних мереж

Upsampling і downsampling є операціями зміни роздільної здатності (розміру) зображення. Downsampling (зменшення роздільної здатності) – це процес зменшення кількості пікселів у зображенні. Це зазвичай досягається шляхом згортки (згущення) зображення з використанням фільтра, що зменшує розмір зображення та виключає частину високочастотних складових. Наприклад, якщо ми зменшуємо розмір зображення вдвічі, ми можемо використовувати фільтр, який видаляє кожен другий рядок і стовпець, зменшуючи розмір зображення у два рази. Upsampling (збільшення роздільної здатності) – це процес збільшення кількості пікселів у зображенні, який може бути досягнутий шляхом відновлення пропущених пікселів з використанням інтерполяції. Наприклад, ми можемо збільшити розмір зображення у два рази, використовуючи методи інтерполяції, такі як білінійна інтерполяція, яка використовує набір сусідніх пікселів, щоб відновити нові пікселі між ними. Для відновлення високоякісного зображення з низької роздільної здатності, спочатку зображення зменшують до певного розміру, а потім застосовують методи upsampling для відновлення деталей та підвищення роздільної здатності

[7,8]. Конволюційний шар (Convolutional layer) – це основний будівельний блок в згорткових нейронних мережах, який відповідає за обробку вхідного зображення. Convolutional layer зазвичай містить набір фільтрів (які також називають ядром або матрицею зважених параметрів), які «скользять» по вхідному зображенні, виконуючи операцію згортки. Кожен фільтр має вагові коефіцієнти, які використовуються для змішування пікселів вхідного зображення, щоб отримати нове значення пікселя у вихідному зображенні. Використовуючи набір фільтрів, Convolutional layer видає декілька карт ознак, що представляють вихідну інформацію. Після проходження вхідного зображення через Convolutional layer, вихідні карт ознак передаються в наступний Convolutional layer або інший тип шару, що дозволяє відновлювати більш складні функції зображення. У CNN (convolutional neural network) зазвичай використовують декілька Convolutional layers з різними наборами фільтрів, що дозволяє отримати все більш і більш складні ознаки зображення. Застосування Convolutional layers дозволяє виявляти просторові залежності в зображеннях [9], що є корисним в багатьох завданнях обробки зображень, таких як

класифікація, виявлення об'єктів, відновлення зображень високої роздільності тощо. ReLU (Rectified Linear Unit) є однією з найбільш поширених функцій активації в глибокому навчанні. Ця функція активації має вигляд: $f(x) = \max(0, x)$, де x – вхідний сигнал, а $f(x)$ – вихідний сигнал після застосування функції активації. Результатом функції ReLU є 0 для всіх вхідних значень, які менші за 0, а для всіх значень, які більші або рівні 0, результатом є саме це значення. Це означає, що функція ReLU "активується" тільки тоді, коли вхідний сигнал більший за 0, інакше вона не дає ніякої відповіді. Використання функції ReLU має декілька переваг у порівнянні з іншими функціями активації, зокрема вона швидше обчислюється і має менше проблем з градієнтами в процесі навчання нейронної мережі [10]. Subpixel є технікою збільшення роздільної здатності зображення шляхом використання інтерполяції кольору між пікселями зображення. Замість того, щоб просто розтягувати пікселі, що може призвести до використання більшої кількості пікселів, subpixel розділяє кожен піксель на більшу кількість частин, кожна з яких містить часткову інформацію про колір. Це дозволяє досягти більшої роздільної здатності без збільшення кількості пікселів. Desubpixel, з іншого боку, є процесом зменшення роздільної здатності зображення. Це може бути корисно при зменшенні розміру зображення або при зменшенні розміру деяких його елементів, наприклад, при зменшенні розміру шрифту на зображенні. Загалом, subpixel та desubpixel – це техніки, які використовуються для оптимізації розміру зображень з метою покращення якості візуального відображення та зменшення розміру файлу зображення. Залишковий блок (Residual block) є типом блоку, який дозволяє пропускати вхідні дані через нейронну мережу, додаючи їх до вихідного сигналу блоку, що називається "skip connection" або "shortcut connection". Це дозволяє нейронній мережі "залишатися на шляху" до вхідних даних, що може допомогти покращити якість навчання та зменшити ймовірність зникаючих градієнтів. Residual block може бути визначено як: $y = F(x) + x$, де x – вхідний сигнал, y – вихідний сигнал, а F – функція, яка навчається блоком. Якщо вхідний та вихідний сигнали мають різну кількість функціональних просторів, то вхідний сигнал x може проходити через додатковий шар, щоб привести його до відповідного розміру перед додаванням до вихідного сигналу $F(x)$. Residual block дозволяє нейронній мережі "вчитися" на різниці між вхідними та вихідними даними, замість того, щоб "вчитися заново" на тих самих даних, що може призвести до перенавчання. Шари нормалізації (Instance Normalization) – це метод нормалізації (рівняння, що перетворює вхідні дані в стандартизовану форму) для глибоких нейронних мереж, який застосовується до кожного елемента вхідного шару окремо (на відміну від Batch Normalization, де

нормалізація виконується для всіх зразків в пакеті). У Instance Normalization для кожного зразка вхідних даних відбувається нормалізація по каналах (глибина), шляхом віднімання середнього значення і ділення на стандартне відхилення. Це допомагає уникнути зміщення вихідних значень відносно точки розподілу вхідних даних, що може призвести до занадто великих або занадто малих значень ваг. Це може зменшити проблеми зі збіжністю, забезпечити більш стійку роботу мережі та поліпшити якість роботи на нових даних. За допомогою Downsampling зображення спочатку знижується з коефіцієнтом 4 за допомогою двох запропонованих $\times 2$ модулів desubpixel. Конволюційний шар 1×1 використовується для налаштування кількості каналів в потрібне значення. Після зниження дискретизації в простір з низькою роздільною здатністю, ознаки надходять в модуль з N залишкових блоків, кожен з яких складається з двох конволюційних шарів 3×3 з наступною нормалізацією і ReLU активації. Зазначимо, що шари нормалізації екземплярів використовуються для покращення зображення лише для нормалізації варіації контрасту між зразками. Блок вихід N -го залишкового блоку переходить в блок Up Sample, де потрапляє в модуль $\times 2$ Subpixel, перед підсумовуванням з низькоякісним вхідним зображенням для отримання прогнозованого високоякісного вихідного зображення [11]. Зменшення кроку дискретизації зазвичай призводить до втрати інформації. Модуль реверсивного зниження дискретизації, так званий десубпіксель, виконує зниження дискретизації таким чином, що його вхідний сигнал може бути відновлений, як показано на Рис. 3. Запропонований модуль десубпікселя систематично перебудовує просторові характеристики в канали для зменшення просторових розмірів без втрати інформації. Нехай U і D відповідно позначають функцію збільшення субпіксельної дискретизації та зменшення дискретизації десубпікселя [12]. Операції по зменшенню та збільшенню кроку дискретизації призводять до перетворення ідентичності таким чином, що:

$$U(D(X)) = X. \quad (5)$$

Запропонований метод є швидким та ефективним методом з наступних трьох причин. По-перше, розглянутий метод виконує більшість своїх обчислень у просторі з низькою роздільною здатністю. Обчислювальна складність набагато нижча, ніж у мереж кодер-декодер із незмінною роздільною здатністю. По-друге, кількість каналів, що використовуються в залишкових блоках, невелика, що дозволяє нейронній мережі бути глибокою. По-третє, даний метод виконує зниження дискретизації, яке називається десубпіксель, що не призводить до втрати інформації [13].

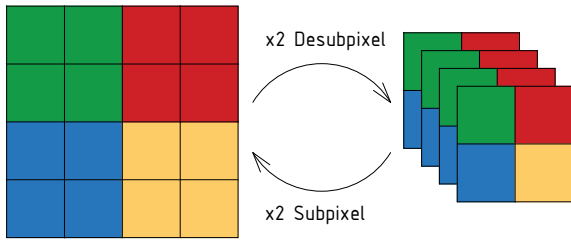


Рис. 3. Субпіксель та запропонований десубпіксель

2 Функції втрат

Mean Squared Error (MSE) – це метрика, яка використовується для оцінки якості прогнозів моделі. Вона широко використовується в статистичному аналізі, машинному навчанні та інших галузях. MSE розраховується як середнє значення квадратів відхилень прогнозованих значень від фактичних значень. Чим менше значення MSE, тим краще модель прогнозує дані. Математично, формула MSE виглядає наступним чином:

$$\text{MSE} = \left(\frac{1}{n}\right) * \sum_{(i=1, n)} (y_i - \hat{y}_i)^2, \quad (6)$$

де n – кількість прикладів у вибірці, y_i – фактичне значення для i -го прикладу, \hat{y}_i – прогнозоване значення для i -го прикладу. MSE може бути корисною метрикою для порівняння різних моделей та для визначення того, яку модель слід використовувати для подальшого прогнозування.

VGG Loss – це метрика, яка використовується для оцінки якості, зокрема для задач генерації зображень, використовуючи глибокі нейронні мережі. Ця метрика отримала свою назву на честь нейронної мережі VGG, яка використовується для виконання завдань комп'ютерного зору. VGG Loss вимірює різницю між відтвореним зображенням та оригінальним зображенням. Для цього використовуються відображення об'єктів на різних рівнях абстракції зображення, які знаходяться в нейронній мережі VGG. Формула VGG Loss може виглядати наступним чином:

$$\text{VGG}_{\text{Loss}} = \lambda_1 \cdot L_1 + \lambda_2 \cdot L_2 + \lambda_3 \cdot L_3 + \lambda_4 \cdot L_4, \quad (7)$$

де $L_1 - L_4$ – це середньоквадратичні помилки між відображеннями об'єктів на різних рівнях абстракції зображення відтвореного зображення та оригінального зображення. Коефіцієнти $\lambda_1 - \lambda_4$ використовуються для збалансування ваг кожної складової метрики. У загальному, VGG Loss допомагає забезпечити, що відтворені зображення мають подібні

характеристики до оригінальних зображень, такі як контрастність, насиченість кольорів та форми об'єктів на зображенні. Це допомагає забезпечити більш якісну генерацію зображень та покращити рівень задоволення користувачів.

Параметри компромісу використовуються для оптимізації алгоритмів та моделей машинного навчання, коли необхідно збалансувати різні параметри для досягнення найкращої загальної продуктивності. Параметри компромісу включають в себе збалансування різних аспектів моделі, наприклад точності та швидкості обчислень. Остаточна функція втрат складається з втрат середньоквадратичної помилки (MSE) \mathcal{L}_M і втрати VGG \mathcal{L}_V з параметрами компромісу α_M і α_V відповідно:

$$\mathcal{L} = \alpha_M \mathcal{L}_M + \alpha_V \mathcal{L}_V. \quad (8)$$

3 Експеримент

Для тестування були використані еталонні набори даних, ми використовуємо набір із 100 зображень з обличчям людей. Всі зображення були завантажені з електронного ресурсу, де користувачі діляться своїми зображеннями безкоштовно. Запропонована в роботі нейронна мережа навчається на еталонних зображеннях в вузьконаправленому профілі, для даного експерименту було обрано обличчя людей в різних ракурсах. Наступним етапом були обрані функції спотворення для подальшого реконструювання зображення. Для початку закладалися вхідні дані, на яких і працює нейронна мережа. Під час проведення дослідження потрібно було прорахувати кількість циклів навчання, згорткових шарів, залишкових блоків та інші характеристики описані вище. Результатом тестування було проведення порівняння співвідношення пік-сигнал-шум (PSNR) та структурної подібності (SSIM) запропонованих методів з різними варіантами налаштування втрат та кількості циклів навчання. Даний метод був реалізований з використанням мови програмування python 3.6 та модулів: tensorflow 1.10, tensorlayer 1.9, tensorboardX 1.4. Експерименти проводяться на одному графічному процесорі NVidia GeForce GTX 1650 і на це йде близько 5 годин навчання.

Для дослідження було виконано три методи навчання: MSE_{s_4} , MSE_{s_2} та $Full_{s_4}$. Метод MSE_{s_4} має коефіцієнт зменшення вибірки 4 та параметри компромісу $\alpha_M = 1$ та $\alpha_V = 0$. Метод MSE_{s_2} має коефіцієнт зменшення вибірки 2 та параметри компромісу $\alpha_M = 1$ та $\alpha_V = 0$. Метод $Full_{s_4}$ має коефіцієнт зменшення вибірки 4 та параметри компромісу $\alpha_M = 1$ та $\alpha_V = 10^{-4}$. Кількість циклів навчання 20, 50 та 100. На Рис. 4 показані графіки порівняння результатів досліджень, в ході експерименту ми побачили що, метод MSE_{s_2} краще справляється з реконструкцією зображень при впливі завад але якщо дивитися з точки зору ви-

хідних зображень, то метод $Full_{s4}$ краще працює з темними зображеннями.

Тому далі було розроблено об'єднаний метод MSE_{s2Full} з коефіцієнтом зменшення вибірки 2 та

параметрами компромісу $\alpha_M = 1$ та $\alpha_V = 10^{-4}$. Даний метод краще працює з різними зображеннями і відтворює спотворене зображення близьке до оригіналу Рис. 4.

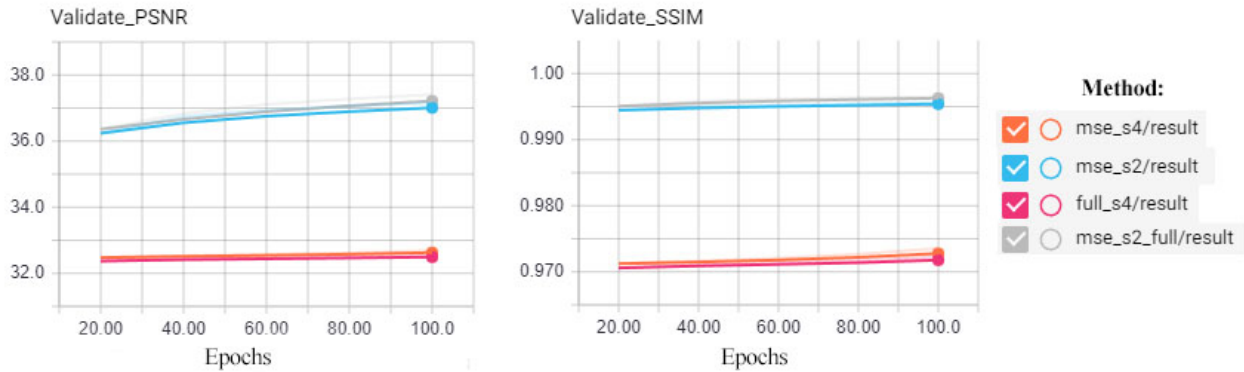


Рис. 4. Порівняння PSNR та SSIM запропонованих методів з різними варіантами налаштування

Табл. 1 Порівняння PSNR і SSIM запропонованого методу з різними варіантами налаштування.

Метод	PSNR	SSIM
MSE_{s4}	32.69	0.9736
MSE_{s2}	37.16	0.9956
$Full_{s4}$	32.53	0.9723
MSE_{s2Full}	37.41	0.9965

Тестування мережі показало що, запропонований метод MSE_{s2Full} досягає кращих показників в порівнянні з методами MSE_{s4} , MSE_{s2} та $Full_{s4}$ як за PSNR, так і за SSIM.

4 Результати

В результаті експерименту визначено алгоритм, по якому запропонована нейронна мережа може по-

кращувати будь-яке зображення, з різними видами спотворень. На Рис. 5 показані покращені зображення за допомогою даної нейронної мережі. Зображення низької та високої якості зроблені за допомогою камери xiaomi mi 10t pro та дзеркальної фотокамери Canon 250D відповідно. При покращенні зображень контрасти зазвичай можуть різнитися між низькоякісними та високоякісними зображеннями, що має бути усунене за допомогою методу нормалізації шарів екземплярів [14]. Обмеження методу полягає в недосконалісті ресурсів мобільних пристроїв, це обмеження на складних зразках неминуче (Рис. 6). Метод вводить артефакти, пов'язані з труднощами розрізнення хрестоподібних або вертикальних ліній. При покращенні зображення, оскільки вхідні дані надходять в умовах поганого освітлення, запропонований алгоритм нездатний посилити яскраві кольори.

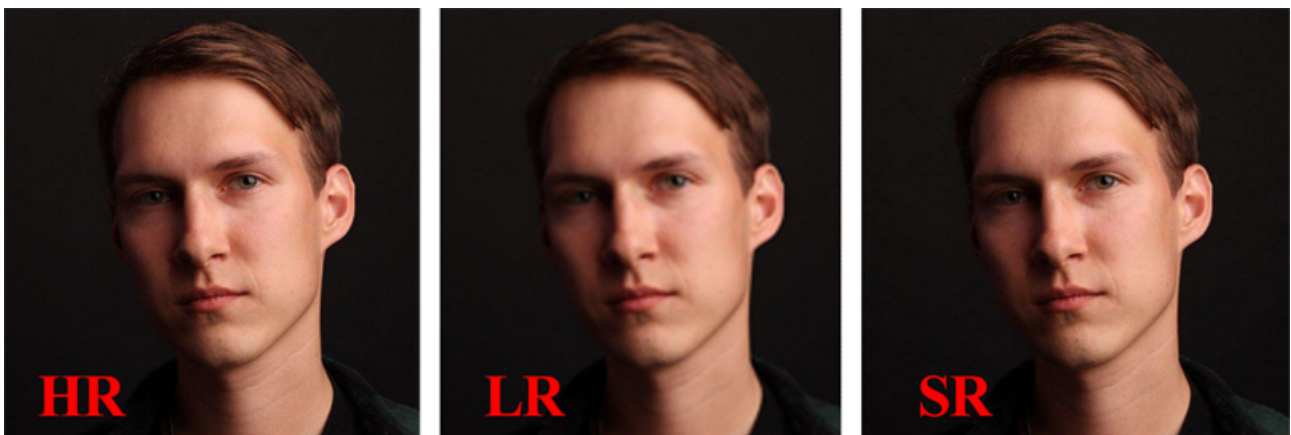


Рис. 5. Зображення з високою роздільною здатністю (HR), зображення з низькою роздільною здатністю (LR) та відновлене зображення з високою роздільною здатністю (SR) методу MSE_{s2Full}



Рис. 6. Обмеження методу, до та після обробки нейронної мережі

Висновки

Супер роздільна здатність є досить новим науковим напрямом в обробці зображень, що постійно розвивається та покращується. Хоча деякі методи SR були розроблені вже давно, їхній розвиток та застосування стали можливими тільки з використанням сучасних алгоритмів та технологій, таких як глибинні нейронні мережі.

Метод для відтворення та поліпшення якості зображень є швидким та ефективним методом. Проведення експерименту з використанням даного методу показало його наступні переваги:

- більшість обчислень проводиться у просторі з низькою роздільною здатністю, що значно спрощує обчислювальні операції;

- кількість каналів, що використовуються в залишкових блоках, невелика, що дозволяє мережі бути глибшою в стільки разів в скільки разів зменшується кількість каналів;

- метод виконує зменшення дискретизації, що не призводить до втрати інформації.

Оскільки мова йде про роботу згорткової нейронної мережі в мобільних пристроях, які мають обмежений обчислювальний ресурс, то слід також враховувати і обмеження, які накладаються на використання даного методу. Так, наприклад, під час покращення зображення, коли вхідне зображення виконано з поганим освітленням, метод не може покращити яскраві кольори.

Щоб прискорити час виходу, запропонований метод виконує більшість обчислювальних операцій у просторі з низьким дозволом. Особливості низької роздільної здатності, отримані за допомогою запропонованого десубпикселя, який забезпечує ефективний спосіб зменшення відбору зображень з високою роздільною здатністю. Просторові характеристики систематично переставляються в канали, зберігаючи значення незмінними, за рахунок чого забезпечують достатню інформацію для наступних шарів згортки. Для підвищення точності реконструкції конволюційна архітектура мережі є більш глибокою, що робить її придатною для глибинного навчання, з невеликою кількістю каналів. Експериментальні результати за стандартними наборами даних

демонструють прогресуючі досягнення з точки зору якості зображення та часу роботи запропонованого методу порівняно з подібними вже відомими методами реставрації зображень. Отже, можна сказати, що супер роздільна здатність продовжує бути активною областю досліджень, яка постійно розвивається та вдосконалюється.

References

- [1] Yang J. and Huang T. (2017). Image Super-Resolution: Historical Overview and Future Challenges. *in Super-Resolution Imaging*, CRC Press, pp. 1–34. doi: 10.1201/9781439819319-1.
- [2] Kim P. (2017). Convolutional Neural Network. *in MATLAB Deep Learning*, Berkeley, CA: Apress, pp. 121–147. doi: 10.1007/978-1-4842-2845-6.
- [3] Li Z., Liu F., Yang W., Peng S., and Zhou J. (2022). A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. *IEEE Trans Neural Netw Learn Syst*, Vol. 33, No. 12, pp. 6999–7019. doi: 10.1109/TNNLS.2021.3084827.
- [4] Albawi S., Mohammed T. A., and Al-Zawi S. (2017). Understanding of a convolutional neural network. *2017 International Conference on Engineering and Technology (ICET)*, pp. 1–6. doi: 10.1109/ICEngTechnol.2017.8308186.
- [5] Zhao N., Wei Q., Basarab A., Dobigeon N., Kouame D., and Tourneret J.-Y. (2016). Fast Single Image Super-Resolution Using a New Analytical Solution $l_2 - l_2$ Problems. *IEEE Transactions on Image Processing*, Vol. 25, No. 8, pp. 3683–3697. doi: 10.1109/TIP.2016.2567075.
- [6] Ignatov A. et al. (2018). PIRM Challenge on Perceptual Image Enhancement on Smartphones: Report. *Cornell University*. doi: 10.48550/arXiv.1810.01641.
- [7] Fattal R. (2007). Image upsampling via imposed edge statistics. *ACM SIGGRAPH 2007 papers*, p. 95. doi: 10.1145/1275808.1276496.
- [8] Zhang Y., Zhao D., Zhang J., Xiong R. and Gao W. (2011). Interpolation-Dependent Image Downsampling. *IEEE Transactions on Image Processing*, Vol. 20, No. 11, pp. 3291–3296. doi: 10.1109/TIP.2011.2158226.
- [9] Bayar B. and Stamm M. C. (2016). A Deep Learning Approach to Universal Image Manipulation Detection Using a New Convolutional Layer. *Proceedings of the 4th ACM Workshop on Information Hiding and Multimedia Security*, pp. 5–10. doi: 10.1145/2909827.2930786.
- [10] Ide H. and Kurita T. (2017). Improvement of learning for CNN with ReLU activation by sparse regularization. *2017 International Joint Conference on Neural Networks (IJCNN)*, pp. 2684–2691. doi: 10.1109/IJCNN.2017.7966185.
- [11] Ngerplubpla J. and Chitsobhuk O. (2019). Neuro-fuzzy profile clustering in image enhancement. *2019 7th International Electrical Engineering Congress (iEECON)*, pp. 1–4. doi: 10.1109/iEECON45304.2019.8938965.
- [12] Govil, R. (2000). Neural Networks in Signal Processing. In: Ruan, D. (eds) Fuzzy Systems and Soft Computing in Nuclear Engineering. *Studies in Fuzziness and Soft Computing*, Vol 38. doi: 10.1007/978-3-7908-1866-6_11.

- [13] Shi W. et al. (2016). Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1874–1883. doi: 10.1109/CVPR.2016.207.
- [14] Ignatov A., Kobyshev N., Timofte R., and Vanhoey K. (2017). DSLR-Quality Photos on Mobile Devices with Deep Convolutional Networks. *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 3297–3305. doi: 10.1109/ICCV.2017.355.

Super-Resolution Image Restoration Using Convolutional Neural Network

Nedzelskyi O. Yu., Lashchevska N. O.

The main goal of the super resolution method is to create a higher resolution image from a lower resolution image. High-resolution images provide a high pixel density, hence more detail in the original image. The need for high resolution is widespread in computer vision techniques, pattern recognition applications, or general image analysis. However, high-resolution images are not always available. This is due to the fact that the conversion processes and processing methods require ultra-powerful processes, and the equipment for obtaining high-resolution images is expensive. These problems can be overcome by using image processing algorithms that are relatively inexpensive, which has led to the concept of super-resolution. This has the advantage that it can cost less

and existing low-resolution imaging systems are readily available. High resolution is essential in medical imaging for diagnosis. Many applications require zooming into a specific image area, where high resolution becomes essential, such as surveillance, forensics, and satellite imaging. The method is presented in this paper, using a convolutional neural network to reproduce super-resolution images, directly performs the conversion from a low-resolution image to an image similar to the original. To speed up the output time, the proposed method performs most computational operations in low-resolution space, while reducing the sampling does not lead to information loss. The main task of the neural network is to reconstruct the distorted image and find the ideal reconstruction function, according to which, in fact, a neural network of a simple structure creates high-quality images with better performance, such as resolution, signal-to-noise ratio, with less time spent on image restoration. During the experiment, we determined an algorithm by which the proposed neural network can reconstruct any image with different types of distortion. The super-resolution method is implemented using the python 3.6 programming language and the tensorflow and tensorlayer software modules for convolutional neural networks. Graphical data of signal-to-noise ratio, structural similarity, and loss plots are obtained using the tensorboardX module.

Keywords: super resolution; convolutional neural networks; signal-to-noise ratio; MSE loss; VGG loss; sampling reduction factor; encoder; decoder; desubpixel