

УДК 004.032.26

О.А. Дмитрієва, д-р техн. наук, проф.,
А.С. Мисько, магістр
Донецький національний технічний університет, м. Покровськ, Україна
olga.dmytriyeva@donntu.edu.ua

Вдосконалення методів масштабування зображень за допомогою нейронних мереж

Роботу присвячено питанням вдосконалення методів масштабування зображень для покращення інтерфейсу застарілих ігрових додатків за допомогою згорткових нейронних мереж. Розроблено модель глибокої рекурсивної мережі, яка може самостійно налаштовуватися та продовжувати навчання з урахуванням специфічних особливостей вихідних зображень комп'ютерних ігрових додатків. Розглянуто можливість поєднання методів масштабування для досягнення кращих практичних результатів без підвищення кількості параметрів системи. Спроековано програмний додаток, що дозволяє використовувати розроблений підхід з метою масштабування зображень частин інтерфейсу ігрових додатків для підвищення їх якості. Виконано попереднє навчання мережі та визначені оптимальні параметри для його виконання. Проведено тестування за різними параметрами навчання.

Ключові слова: масштабування зображень, нейронна мережа, глибоке навчання, матриця особливостей, рекурсивні алгоритми.

DOI: 10.31474/1996-1588-2020-1-30-4-11

Вступ

Класичним завданням комп'ютерного бачення є завдання формування зображення надвисокої розподільної здатності на основі одиночного вхідного зображення низької розподільної здатності SISR (Single-Image Super-Resolution), яке широко використовується у криміналістиці, системах спостереження, медичній візуалізації, де отримання найтонших деталей зображення є дуже важливим, та, останнім часом, у комп'ютерних ігрових додатках. Такі зображення, зазвичай, є текстурами, вони не мають цифрового шуму, який присутній у цифрових фотографіях, і потребують збереження загального вигляду та якості при масштабуванні. Для реалізації в комп'ютерних іграх можуть бути використані методи класичної інтерполяції, перевагою яких слід визначити простоту та можливість роботи у реальному часі [1]. Але головним недоліком цих методів є багата кількість дефектів, які виникають під час роботи, що не є прийнятним для використання у комп'ютерних іграх. На відміну від методів класичної інтерполяції [2-3], методи, що використовують нейронні мережі, дають значно кращі результати та створюють меншу кількість дефектів. Також вони дають можливість проводити навчання мережі на специфічних зображеннях конкретної предметної області. Але недоліками цих методів є їх складна реалізація та довгий час роботи. Навчання мережі може займати до декількох днів. Між тим, на даний час, методи на основі глибокого навчання штучних

згорткових нейронних мереж використовуються для створення моделей відображення частин зображень низької розподільної здатності на зображення високої розподільної здатності, після цього методи сусіднього вкладення інтерполюють підпростір між частинами [4]. Такі методи зазвичай мають стратегії, засновані на зображеннях-прикладях, використовуючи внутрішні подібності одного й того ж зображення, або використовуючи навчання функцій відображення частин на парах зовнішніх зображень низької та високої розподільних здатностей. Такі методи навчання на зовнішніх зображеннях-прикладях можуть бути використані для надвисокої розподільної здатності загального ряду зображень або специфічних завдань, відповідно до наданих навчальних зразків [5].

Мета роботи полягає у розробці моделі глибокої рекурсивної нейронної мережі масштабування, яка може самостійно налаштовуватися з визначенням оптимальних параметрів та продовжувати навчання з урахуванням специфічних особливостей вихідних зображень комп'ютерних ігрових додатків.

Сучасні підходи до масштабування зображень

Методи масштабування зображень можна поділити на такі категорії, як класичні методи інтерполяції, та методи, що використовують штучні нейронні мережі для створення апроксимацій значень пікселів масштабованого зображення. До неадаптивних методів інтерполяції можна віднести такі методи, як

інтерполяція методом найближчого сусіда [2], білінійна або бікубічна інтерполяції [3], або метод інтерполяції за допомогою фільтру Ланцоша [2-3]. Адаптивні методи інтерполяції змінюються в залежності від вихідних даних пікселів зображення, наприклад, наявності різких границь або градієнтів [2]. Оскільки такі методи базуються на апроксимації значень пікселів, у вихідному масштабованому зображенні можуть бути присутні дефекти, такі як розмиття зображення, аліасинг – проблема нечіткості безперервних сигналів при їх дискретизації, ефект «ступінчастості» зображення, який виникає при відображенні кривих ліній та використанні методу найближчого сусіда, дефекти дзвону (феномен Гіббса), які виникають при різкому переході інтенсивності у зображенні та візуально проявляються у вигляді смуг поблизу місць різкого переходу інтенсивності [3].

Урахування перелічених недоліків класичних методів інтерполяції визначило спрямованість масштабування зображень з використанням згорткових нейронних мереж, які за допомогою навчання на зображеннях-прикладках досягають більш точної апроксимації значення пікселів масштабованого зображення. Згорткові штучні нейронні мережі є категорією штучних нейронних мереж та здебільшого використовуються для аналізу зображень та розпізнавання образів [6].

Для об'єктивного визначення якості масштабування цифрових зображень в роботі обрано ряд характеристик, до яких слід віднести середню інтенсивність MI (1), яка визначає якість зображення за його контрастністю. Вважається, що чим більше це значення, тим більш якісне зображення.

$$MI = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |x(i, j)|, \quad (1)$$

де M , N – відповідно ширина та висота зображення у пікселях, $x(i, j)$ – значення інтенсивності у точці зображення.

До наступних обраних характеристик слід віднести середню AD (2) та максимальну MD (3) різниці порівняння значення сигналу двох зображень X і Y , які можуть бути визначені співвідношеннями:

$$AD = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (x(i, j) - y(i, j)), \quad (2)$$

$$MD = \max |x(i, j) - y(i, j)|. \quad (3)$$

Середня абсолютна помилка MAE (4) є середнім значенням абсолютної різниці між сигналами зображень, та визначається як

$$MAE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |x(i, j) - y(i, j)|. \quad (4)$$

Для визначення якості масштабування зображення у роботі також використовується пікове співвідношення між максимальним значенням сигналу та потужністю шуму, який його спотворює [7] $PSNR$ (Peak signal-to-noise ratio). Цей показник було використано в роботі через простоту його обчислення та зручність оптимізації. $PSNR$ визначається через середньоквадратичне відхилення σ згенерованого зображення від оригінального

$$PSNR = 20 * \log_{10}(MAX_I) - 10 * \log_{10}(\sigma), \quad (5)$$

де MAX_I – максимальне значення пікселю оригінального зображення.

Для визначення показників (1)-(5) проводилася низка експериментів з використанням спеціальних наборів вхідних зображень, таких як Set5, Set14, BSD100, що містять 5, 14 та 100 зображень відповідно. Для отримання результатів вхідні зображення зменшувалися у 4 рази, після чого збільшувалися у 4 рази з використанням обраного методу масштабування. Значення $PSNR$, наведені у табл. 1, отримувалися як результат такого масштабування при порівнянні з оригінальним зображенням. Як видно з табл. 1, методи, засновані на штучних нейронних мережах SRCNN (Super resolution convolutional neural network) і DRCN (Deeply-recursive convolutional network), досягають кращих результатів на усіх наборах зображень. Враховуючи це, було прийняте рішення про дослідження саме методів, заснованих на використанні згорткових нейронних мереж.

Таблиця 1 – Порівняння методів масштабування на основі $PSNR$

Вхідні дані	Найближчий сусід	Бікубічна інтерполяція	SRCNN	DRCN
Set5	26.26	28.42	30.49	31.53
Set14	24.64	26.00	27.50	28.02
BSD100	25.02	25.94	26.68	27.23

Головною ідеєю використання нейронних мереж в роботі була спрямованість на отримання зображень надвисокої розподільної здатності з використанням можливості навчання для визначення залежності між зображеннями низької та високої розподільної здатностей, наданими для навчання, для подальшого застосування їх для довільних зображень.

Підвищення якості масштабування зображень за допомогою глибокої згорткової нейронної мережі

Для підвищення якості вихідного зображення було запропоновано використовувати уніфіковану нейронну мережу глибокого

навчання DJSR (Deep joint super resolution) [8], яка використовує частини-приклади, отримані з поширених зовнішніх баз даних зображення разом з частинами-прикладми, отриманими з самого вхідного зображення.

На початку роботи, мережа тренується за допомогою зовнішньої бази даних зображень, після цього вона налаштовується, використовуючи приклади з самого вхідного зображення, кожній такій частині-прикладу під час навчання присвоюється вага довіри. Таким чином це пом'якшує недостатність надійних прикладів. Після цього мережа може бути розширена на декілька спеціалізованих під-мереж, та проведено вибіркове навчання та обробка частин-прикладів. Але при використанні частин-прикладів лише самого вхідного зображення виникають дві проблеми [8]: недостатність інформативних частин-прикладів та обмежена надійність. Для вирішення цих проблем було запропоновано будувати ієрархію частин-прикладів з декількох масштабувань вхідного зображення, кожному з яких надається вага довіри, яка обчислюється похибкою збігу найближчого сусіда. Головна ідея полягає у використанні надлишків частин-прикладів між різними сусідніми масштабами зображення [8]. У порівнянні з звичайним методом, що використовує згорткові нейронні мережі (SRCNN), DJSR надає зображення з більш вираженими границями та текстурями, а також краще стримує дефекти зубчатості та блочності.

Спільною рисою багатьох методів на основі згорткових мереж є невелика кількість шарів згорткування, наприклад SRCNN використовує лише 3 шари. При збільшенні кількості шарів, зазвичай, виникає проблема труднощів навчання глибоких нейронних мереж, що негативно впливає на кінцевий результат, не надаючи значних переваг [9]. Для вирішення цих проблем було запропоновано створення рекурсивної згорткової мережі DRCN. Головна ідея такого підходу полягала в рекурсивному застосуванні одного й того ж шару згортки стільки разів, скільки потрібно. Таким чином, число параметрів мережі не змінювалося при виконанні рекурсії. Для поліпшення навчання мережі було запропоновано декілька модифікацій [10-12].

Метод створення кінцевого зображення є однаковим для усіх рекурсій. Так як кожна рекурсія створює свою версію вихідного зображення, усі проміжні результати поєднуються у один для підвищення точності кінцевого результату. Іншим підходом є запровадження прямого з'єднання пропуску від вхідного шару до шарів створення вихідного зображення. У завданні створення надвисокої розподільної здатності, вхідне та вихідне зображення в значній мірі відображають одну й ту ж інформацію, але

при проходженні багатьох рекурсій точна копія вхідного зображення, скоріш за все, буде розріджена. Тому вхідний шар напряму з'єднується з шарами створення вихідного зображення. Це особливо ефективно коли вхідне та вихідне зображення сильно корелюють.

Також було досліджено, що зі збільшенням кількості рекурсій продуктивність та значення PSNR збільшуються, при цьому кількість параметрів мережі не змінюється, а у порівнянні з іншими методами, DRCN надає більш чіткі границі об'єктів [12]. Для уникнення дефектів шуму та розмитості було запропоновано використання додавання шару навчання деградації без нагляду для створення зображень низької розподільної здатності за допомогою мережі DNSR [11]. Це надало можливість створювати більш реалістичні зображення низької розподільної здатності, ніж звичайна інтерполяція.

Запропонована мережа має 3 частини: модель деградації, деградаційний дискримінатор та модель реконструкції. Модель деградації використовується для отримання реалістичних зображень низької розподільної здатності. Деградаційний дискримінатор забезпечує подібність частин-прикладів отриманого зображення низької розподільної здатності до реальних зображень. Модель реконструкції використовується для безпосередньо створення кінцевих зображень високої розподільної здатності, використовуючи отримані зображення низької розподільної здатності для навчання. Виконання навчання мережі використовувалося на вхідних зображеннях саме предметної області комп'ютерних ігрових додатків, що також дозволило підвищити якість результатів.

Проектування архітектури нейронної мережі масштабування зображень

У рамках даної роботи розглядався процес збільшення цифрових растрових зображень з мінімальною втратою якості. При збільшенні зображення підвищується кількість пікселів, що потребує інтерполяції значень нових пікселів, яке здійснюється за допомогою використання нейронних мереж. Загальна схема масштабування зображень показана на рис. 1. У масштабуванні зображень у більшу сторону, нейронна мережа виконує функції інтерполяції нових пікселів зображення.

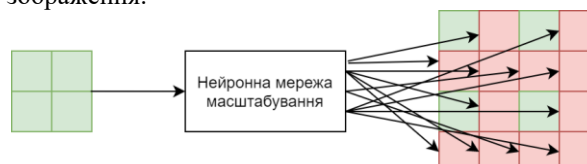


Рисунок 1 – Процес масштабування зображення

Для навчання нейронної мережі масштабування потребується надання вхідних даних. У якості вхідних даних було прийняте рішення використовувати зображення-текстури з ігрових додатків. Це дозволило підвищити якість навчання нейронної мережі для специфічної задачі масштабування зображень з ігрових додатків.

У якості джерела зображень-текстур з ігрових додатків було прийнято рішення використовувати інтернет-ресурс безкоштовних ігрових додатків «itch.io» [13]. Зображення-текстури були отримані з безкоштовних ігор, що відповідали тегам «ретро», «16-біт» та «8-біт». Зображення з ігор з такими тегами були обрані, оскільки таким чином можна отримати нові зображення у високій розподільній здатності, які відповідають стилю старих ігрових додатків.

На вхід мережі подаються вхідні зображення, які було зменшено на деякий коефіцієнт масштабування, а потім збільшено на той же коефіцієнт за допомогою нейронної мережі. На виході згорткової нейронної мережі очікувалося отримання масштабованого зображення, яке у подальшому використовувалося для порівняння з початковим зображенням та для визначення похибки, яка дозволяла корегувати значення ваг та зміщень для фільтрів згорткової мережі [14]. Принцип використання вхідних зображень наведено на рис. 2.

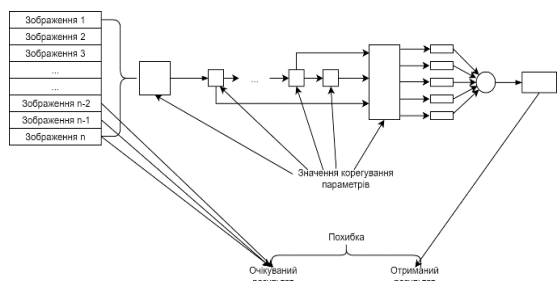


Рисунок 2 – Схема роботи нейронної мережі

Також формувалися окремі набори зображень, що не використовувалися при навчанні. На етапі тестування такі зображення дозволяли впевнитися, що мережа не лише запам'ятала особливості зображень, які були надані для навчання. Загальний алгоритм роботи нейронної мережі наведено на рис. 3. В алгоритмі можна визначити наступні етапи: отримання особливостей зображення, відображення, реконструкція. Етап відображення є рекурсивним [15] та виконується декілька разів. Після визначення структури мережі виконувалася ініціалізація усіх ваг та відхилень. При тренуванні після кожної ітерації ваги та відхилення корегувалися в залежності від отриманих результатів. Кількість ітерацій навчання залежала

від кількості вхідних даних та структури мережі. Тренування припинялося, якщо значення похибки не змінювалося протягом декількох ітерацій.



Рисунок 3 – Загальний алгоритм роботи мережі

Результати тренування, а саме, структура, значення фільтрів згорткових шарів, ваг та відхилень зберігалися у файл. Для використання мережі дані завантажувалися зі збереженого файлу та виконувався процес масштабування, ідентичний процесу навчання але за відсутністю корегування ваг та відхилень.

Навчання глибокої нейронної мережі

Процес навчання мережі полягав в отриманні вхідних зображень високої розподільної здатності, зменшенні їх у кількість разів, на яку виконується масштабування, масштабування отриманого зображення за допомогою нейронної мережі та розрахування похибки між отриманим та вхідним зображенням. Корегування значень фільтрів згорткових шарів мережі виконувалося за допомогою процесу зворотного поширення [15]. Цей процес можна поділити на наступні етапи:

- прохід вперед, визначає прохід вхідного зображення через шари нейронної мережі з поточними значеннями фільтрів, ваг та зміщень та отримання результату;
- функція похибки, визначається величина розходження між отриманим результатом та еталонним зображенням з датасету навчання, зазвичай використовувалася функція середньоквадратичної похибки, головною метою

навчання було зменшення цієї похибки;

- прохід назад, визначає, які саме значення фільтрів та ваг найбільше сприяли похибці, та формує рекомендації щодо мінімізації похибки;
- оновлення значень, значення фільтрів та ваг корегуються з метою зменшення похибки.

Для контролювання швидкості навчання прийнято використовувати коефіцієнт швидкості навчання [16], який визначається розробником. Більший коефіцієнт швидкості означає, що при корегуванні значень фільтрів та ваг будуть використовуватися більші кроки корегування, що підвищує швидкість навчання мережі, але це також може призвести до занадто великих кроків змін значень, які не надаватимуть очікувану точність для досягнення потрібних значень. Занадто мале значення цього коефіцієнту може призвести до зупинення навчання системи через незначні зміни ваг.

Процес зворотного поширення є ітераційним процесом та виконується певну кількість разів, яка визначається розробником для кожного зображення з множини початкових зображень. На початку кожної ітерації поточна похибка також порівнюється з попередньою. Для запобігання неправильного навчання мережі та отримання неоптимального рішення, коефіцієнт швидкості навчання змінюється у процесі навчання мережі.

Для визначення оптимальних параметрів навчання нейронної мережі було створено попередню модель мережі з коефіцієнтом масштабування 4, кількістю ітерацій 30, кількістю зображень для навчання 100. При навчанні мережі розраховувалося значення функції похибки, яка безпосередньо впливає на оновлення ваг шарів мережі [17]. У випадку мережі масштабування зображень в якості функції похибки обиралося середнє значення середньоквадратичної похибки між значеннями пікселів результату масштабування та оригінальним вхідним зображенням, для рекурсивної мережі – середньоквадратична похибка між кожним шаром рекурсії мережі.

Тестування глибокої рекурсивної мережі

В якості джерел зображень-текстур для проведення експериментів були обрані застарілі ігрові додатки року випуску до 2003. У якості змінюваних параметрів мережі виступали початкові значення ваг та зміщень шарів, кількість фільтрів, коефіцієнт швидкості навчання та глибина рекурсії шару визначення частин зображення. При запуску навчання значення похибки та значення PSNR між еталоном та вихідним зображенням на кожній ітерації зберігалися для подальшого аналізу та візуалізації, початкове значення коефіцієнту швидкості навчання 0.1, кількість фільтрів – 48.

На рис. 4-5 наведено вплив глибини рекурсії з показниками 4, 8 та 16 на хід навчання нейронної мережі при незмінних значеннях коефіцієнту швидкості навчання та кількості фільтрів. На рисунках показано графіки залежності похибки та значення характеристики PSNR від ітерації навчання при різних значеннях параметру, що досліджується.

Як можна побачити з результатів, глибина рекурсії незначно зменшує швидкість навчання, але має позитивний вплив на якість зображення результату за характеристикою PSNR, тому було прийняте рішення подальше використання глибини рекурсії згорткового шару мережі 16.

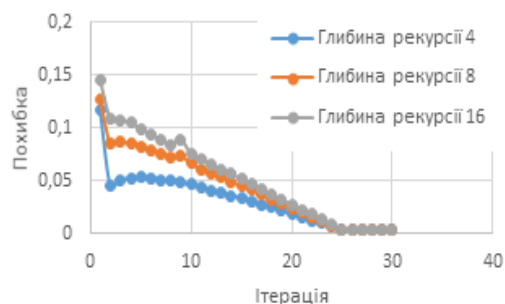


Рисунок 4 – Вплив глибини рекурсії на похибку при навчанні мережі

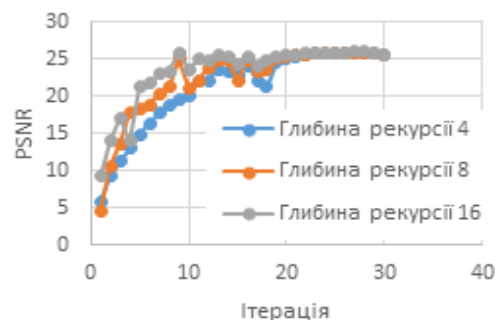


Рисунок 5 – Вплив глибини рекурсії на PSNR при навчанні мережі

Також проводилася група експериментів, які оцінювали вплив кількості фільтрів згорткового шару нейронної мережі на процес навчання та результати. На рис. 6-7 наведено динаміку змін похибки та показника PSNR при попередньому навчанні для відповідних кількостей фільтрів.

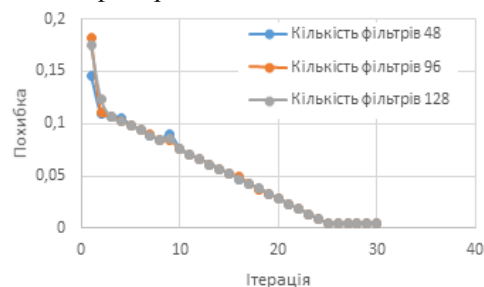


Рисунок 6 – Вплив кількості фільтрів на похибку при навчанні мережі

Як видно з результатів, кількість фільтрів майже не впливає на похибку або якість зображення за характеристикою *PSNR*, але, оскільки кількість фільтрів значно збільшує кількість нейронів, ваг та зміщень, це суттєво збільшує час навчання та роботи нейронної мережі, а також її споживання ресурсів.

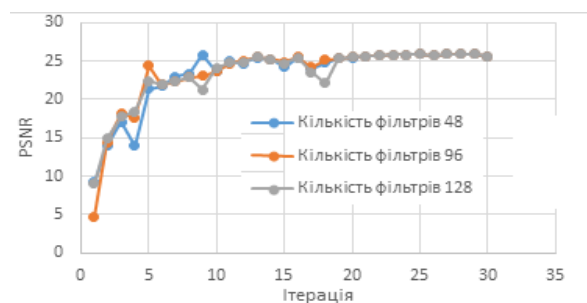


Рисунок 7 – Вплив кількості фільтрів на *PSNR* при навчанні мережі

Оскільки планується робота мережі на звичайному користувальницькому обладнанні, було прийняте рішення використовувати 48 фільтрів у подальшому навчанні мережі. Це надаватиме приблизно однакову якість зображення за проведенням експериментом, але навчання займатиме значно менше часу та ресурсів, що є важливим для роботи на не дуже потужному користувальницькому обладнанні.

Також проводилися експерименти щодо визначення впливу початкового значення коефіцієнту швидкості навчання мережі на її навчання. У якості значень, що перевірялись, було обрано 0.1, 0.01 та 0.001. Доведено, що менший коефіцієнт швидкості навчання дає більш стабільне навчання на меншій кількості вхідних зображень, але його надмірне зменшення негативно впливає на якість зображення результату за характеристикою *PSNR*. Оскільки коефіцієнт швидкості зменшується у процесі навчання, завдання занадто низького коефіцієнту може призвести до перенавчання при великій кількості вхідних зображень.

На основі проведених досліджень для етапу

попереднього навчання можна стверджувати, що збільшення величини глибини рекурсії збільшує час виконання навчання, але в той же час підвищує якість зображення результату за характеристикою *PSNR*. Збільшення кількості фільтрів згорткового шару мережі призводить до значного збільшення складності мережі та використання нею ресурсів та не дає значного підвищення якості зображення результату за характеристикою *PSNR*. При невеликій кількості вхідних зображень слід використовувати менший коефіцієнт швидкості навчання мережі.

Висновки

В ході роботи проведений порівняльний аналіз сучасних підходів до масштабування зображень, обґрунтовано застосування глибокої згорткової нейронної мережі, що самостійно налаштовується. Проведено проектування та моделювання нейронної мережі за допомогою середовищ моделювання. Реалізовано процес навчання нейронної мережі, досліджено вплив параметрів навчання на його хід. Спроєктовано програмний додаток, що дозволяє використовувати розроблений підхід з метою масштабування зображень частин інтерфейсу ігрових додатків для підвищення їх якості. Проведене попереднє навчання мережі та визначені оптимальні параметри для його виконання. У процесі тестування визначено функції ініціалізації та подальшої оптимізації ваг, що були розглянуті, підібрані оптимальні параметри навчання розробленої нейронної мережі масштабування зображень та проведене навчання мережі на навчальному та тестовому наборах даних. Оцінювання отриманих результатів проводилося за характеристикою визначення якості зображення.

Подальша робота за цим напрямком може бути спрямована на визначення оптимальних показників коефіцієнту швидкості навчання мережі, кількості фільтрів, глибини рекурсії та на використання додаткової нейронної мережі попередньої обробки зображень.

Список літератури

1. Shan Q., Li Z., Jia J., Tang C.-K. Fast image/video up sampling. *ACM Transactions on Graphics*. 2008. Vol. 27, № 5. Art. 153. P. 1-8.
2. Sun H., Zhang F., Zheng N. An edge-based adaptive image interpolation and its VLSI architecture. *Asia-Pacific Signal and Information Processing Association Annual Summit*. 2012. Vol. 2. P. 1-6.
3. Yang C., Huang J., Yang M.H. Exploiting self-similarities for single frame super-resolution. *Asian Conference on Computer Vision*. 2012. Vol. 6494. P. 497-510.
4. Kim J., Lee J.K., Lee K. M. Accurate image super-resolution using very deep convolutional networks. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2016. P. 1646-1654.
5. Dong C., Loy C.C., He K. Image Super-Resolution Using Deep Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2016. Vol. 38, № 2. P. 295-307.

6. Zeiler M., Fergus R. Visualizing and understanding convolutional networks. *Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics*. 2014. P. 818-833.
7. Al-najjar Y., Soong, D. Comparison of Image Quality Assessment. *International Journal of Scientific & Engineering Research*. 2012. Vol. 3, №8. P. 1-5.
8. Tai Y., Yang J., Liu X. Image Super-Resolution via Deep Recursive Residual Network. *IEEE Conference on Computer Vision and Pattern Recognition*. 2017. Vol. 17. P. 2790-2798.
9. Wang Z., Yang Y., Wang Z. Self-tuned deep super resolution. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. 2015. Vol. 15. P. 1-8.
10. Dong C., Loy C. C., He K. Learning a deep convolutional network for image super-resolution. *European Conference on Computer Vision*. 2014. Vol. 8692. P. 184-199.
11. Kim J., Lee J.K., Lee K. M. Deeply-recursive convolutional network for image super-resolution. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2016. Vol. 16. P. 1637-1645.
12. Paper D. Gradient Descent: Data Science Fundamentals for Python and MongoDB. A press Media LLC. 2018. 213 p.
13. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition. *3-rd International Conference on Learning Representations*. 2015. Vol. 15. P. 1-14.
14. Nguyen Q., Hein M. The loss surface of deep and wide neural networks. *34th International Conference on Machine Learning*. 2017. Vol. 17. P. 3995-4008.
15. Eigen D., Rolfe J., Fergus R. Understanding deep architectures using a recursive convolutional network. *2nd International Conference on Learning Representations*. 2014. Vol. 14. P. 1-9.
16. Kubat M. *An Introduction to Machine Learning*. Heidelberg: Springer, 2017. 348 p.
17. Haeffele B., Vidal R. Global optimality in neural network training. *Proceedings 30th IEEE Conference on Computer Vision and Pattern Recognition*. 2017. Vol. 17. P. 4390-4398.

References

1. Shan, Q., Li, Z., Jia, J, Tang, C.-K. (2008), "Fast image/video up sampling", *ACM Transactions on Graphics*, pp. 1-8.
2. Sun, H., Zhang, F., Zheng, N. (2012), "An edge-based adaptive image interpolation and its VLSI architecture", *Asia-Pacific Signal and Information Processing Association Annual Summit*, Vol. 2, pp. 1-6.
3. Yang, C., Huang, J., Yang, M.H. (2012), "Exploiting self-similarities for single frame super-resolution", *Asian Conference on Computer Vision*, pp. 497-510.
4. Kim, J., Lee, J.K., Lee, K. (2016), "Accurate image super-resolution using very deep convolutional networks", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1646-1654.
5. Dong, C., Loy, C.C., He, K. (2016), "Image Super-Resolution Using Deep Convolutional Networks", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 38, № 2, pp. 295-307.
6. Zeiler, M., Fergus, R. (2014), "Visualizing and understanding convolutional networks", *Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics*, pp. 818-833.
7. Al-najjar, Y., Soong, D. (2012), "Comparison of Image Quality Assessment", *International Journal of Scientific & Engineering Research*, Vol. 3, №8, pp. 1-5.
8. Tai, Y., Yang, J., Liu, X. (2017), "Image Super-Resolution via Deep Recursive Residual Network", *IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 17, pp. 2790-2798.
9. Wang, Z., Yang, Y., Wang, Z. (2015), "Self-tuned deep super resolution", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, Vol. 15, pp. 1-8.
10. Dong, C., Loy, C.C., He, K. (2016), "Learning a deep convolutional network for image super-resolution", *European Conference on Computer Vision*, Vol. 17, pp. 184-199.
11. Kim, J., Lee, J. K., Lee, K. M. (2016), "Deeply-recursive convolutional network for image super-resolution", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 16, pp. 1637-1645.
12. Paper, D. (2018), *Gradient Descent: Data Science Fundamentals for Python and MongoDB*, A press Media LLC, 213 p.
13. Simonyan, K., Zisserman, A. (2015), "Very deep convolutional networks for large-scale image recognition", *3rd International Conference on Learning Representations*, Vol. 15, pp. 1-14
14. Nguyen, Q., Hein, M. (2017), "The loss surface of deep and wide neural networks", *34th International Conference on Machine Learning*, Vol. 17, pp. 3995-4008.

15. Eigen, D., Rolfe, J., Fergus, R. (2014), "Understanding deep architectures using a recursive convolutional network", *2nd International Conference on Learning Representations*, Vol. 14, pp. 1-9.
16. Kubat, M., (2017), *An Introduction to Machine Learning: An Introduction to Machine Learning*, Springer, Heidelberg, 348 p.
17. Haeffele, B. D., Vidal, R. (2017), "Global optimality in neural network training", *Proceedings 30th IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 17, pp. 4390-4398.

Надійшла до редакції 05.01.2020

О.А. ДМИТРИЕВА, А.С. МИСЬКО

ГБУЗ «Донецкий национальный технический университет», г. Покровск, Украина
olga.dmytriyeva@donntu.edu.ua

СОВЕРШЕНСТВОВАНИЕ МЕТОДОВ МАСШТАБИРОВАНИЯ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ

Работа посвящена вопросам совершенствования методов масштабирования изображений для улучшения интерфейса устаревших игровых приложений с помощью сверточных нейронных сетей. Разработана модель глубокой рекурсивной сети, которая может самостоятельно настраиваться и продолжать обучение с учетом специфических особенностей исходных изображений компьютерных игровых приложений. Рассмотрена возможность сочетания методов масштабирования для достижения лучших практических результатов без повышения количества параметров системы. Спроектировано программное приложение, которое позволяет использовать разработанный подход с целью масштабирования изображений частей интерфейса игровых приложений для повышения их качества. Выполнено предварительное обучение сети и определены оптимальные параметры для его выполнения. Проведено тестирование различных параметров обучения сети, выполнен сравнительный анализ результатов.

Ключевые слова: масштабирование изображений, нейронная сеть, глубокое обучение, матрица особенностей, рекурсивные алгоритмы.

O.A. DMITRIEVA, A.S. MYSKO

Donetsk National Technical University, Pokrovsk, Ukraine
olga.dmytriyeva@donntu.edu.ua

IMPROVEMENT OF IMAGE SCALING METHODS USING NEURAL NETWORKS

The work is devoted to improving image scaling methods to improve the interface of legacy gaming applications using convolutional neural networks. A model of a deep recursive network has been developed, which can independently configure and continue training, taking into account the specific features of the source images of computer gaming applications. The possibility of combining scaling methods to achieve the best practical results without increasing the number of system parameters is considered. A software application has been designed that allows you to use the developed approach to scale images of the interface parts of gaming applications to improve their quality. The network has been pre-trained and the optimal parameters for its implementation have been determined. Various network training parameters were tested, a comparative analysis of the results was performed.

As can be seen from the results, the number of filters does not practically affect the error or image quality according to the PSNR characteristic, but since the number of filters significantly increases the number of neurons, weights and displacements, this significantly increases the time of training and operation of the neural network, as well as its resource consumption. Since it is planned to operate the network on the ordinary user equipment, it was decided to use 48 filters in further network training. This will provide acceptable image quality, but training will take much less time and resources, which is important for working on the not very powerful user equipment. Also, experiments were conducted to determine the initial value effect of the network learning rate coefficient. It is proved that a lower coefficient of learning speed gives a more stable result on a smaller number of the input images. At the same time, setting the learning rate coefficient too low can lead to retraining with a large number of incoming images. With a small number of input images, a lower coefficient of network learning speed should be used. Based on the studies for the preliminary training stage, it can be argued that the recursion depth increasing rises the learning time, but at the same time, improves the image quality of the result according to the PSNR characteristic. An increase in the number of the convolutional layer filters of the network leads to a significant increase in the complexity of the network and the number of used resources, but does not increase the image quality of the result according to the PSNR characteristic.

Key words: image scaling, neural network, deep learning, matrix of features, recursive algorithms.