

DOI: <https://doi.org/10.33216/1998-7927-2025-294-8-5-10>

УДК 004.932:004.942.5

ВДОСКОНАЛЕННЯ МЕТОДІВ ІМПЕЙНТИНГУ НА ОСНОВІ ГЕНЕРАТИВНИХ МОДЕЛЕЙ

Семаньків М.В., Ціхун О.В.

IMPROVEMENT OF IMAGE INPAINTING METHODS BASED ON GENERATIVE MODELS

Semankiv M.V., Tsikhub O.V.

Обсяг візуальної інформації, що нас оточує, постійно зростає, і це зумовлює потребу в інтуїтивно зрозумілих інструментах для її обробки. Дослідження, які спрямовані на удосконалення алгоритмів для додавання або видалення об'єктів із графічного контенту, є актуальними та мають високий практичний потенціал. Системи, які використовують такі алгоритми, можуть знайти застосування в різних галузях – від кінематографа та фотографії до реставрації мистецьких творів і навіть медицини. Технологія Image inpainting дозволяє автоматично відновлювати відсутні частини зображення, економлячи час та зусилля користувача. Попри наявність великої кількості рішень для задач реконструкції все ще бракує систем, які б комплексно використовували генеративні моделі для досягнення високої якості відновлення контенту. Це підкреслює значущість досліджень у даній сфері, спрямованих на вдосконалення існуючих алгоритмів. Оцінка сучасного стану об'єкта дослідження характеризується значним прогресом у розвитку генеративних моделей. Водночас існує потреба в подальшому вдосконаленні для підвищення універсальності, оптимізації використання ресурсів та зменшення обчислювальних витрат. В роботі досліджено методи генеративних конкурентних мереж (GAN) для задач зафарбовування зображень, оцінено ефективність наявних підходів. Удосконалено метод генерації контенту із застосуванням унікальної функції втрат, яка за показниками точності та часу виконання перевищує розглянуті конкурентні реалізації. Розроблено програмне забезпечення для вирішення практичних задач. Проведено роботу над системою відновлення графічного контенту на основі GAN, здійснено її вдосконалення за рахунок математичних інструментів. Результати, що отримані в рамках

дослідження, демонструють покращення якості відновлення зображень за допомогою розробленої системи на основі GAN, хорошу якість видалення структур із зображень з мінімальними артефактами навіть при низькій роздільній здатності за короткий час. Однак, слід зауважити, точність відновлення може знижуватися при роботі із складними чи деталізованими зображеннями. Практична значущість дослідження поширюється на всі галузі, де обробка візуальних даних має ключове значення. Вона полягає у вирішенні проблеми неякісного видалення об'єктів із зображень, розвитку алгоритмів заміни та видалення конкретних ділянок на зображеннях за допомогою генеративних конкурентних мереж.

Ключові слова: inpainting, GAN-мережі.

Вступ. Імпейнтинг (image inpainting) є перспективним напрямом комп'ютерної графіки, що забезпечує відновлення, видалення або генерацію відсутніх фрагментів зображення з високим рівнем реалістичності. Розвиток цієї технології пов'язаний із застосуванням глибоких нейронних мереж, зокрема дифузійних моделей, які здатні враховувати як локальні текстурні особливості, так і глобальний семантичний контекст. Завдяки цьому імпейнтинг має широкі перспективи у сфері цифрової реставрації, медичної візуалізації та автоматизованого редагування мультимедійних даних, формуючи основу для інтелектуальних систем нового покоління.

Одним із актуальних напрямків у дослідженнях імпейнтингу є видалення обраних об'єктів із зображень та заповнення пустих

ділянок реалістичним контентом. Це має важливе практичне значення, оскільки дозволяє автоматизувати редагування фотографій та відео для медіаіндустрії, реклами, кінематографу та систем відеоспостереження, а також сприяє захисту конфіденційності та персональних даних. Водночас цей напрямок є перспективним завдяки можливостям сучасних методів глибокого навчання, зокрема генеративно-змагальних мереж і дифузійних моделей, які дозволяють створювати високоякісний контент, адаптований до конкретного контексту сцени, та стимулюють розвиток суміжних галузей, таких як генеративний дизайн, доповнена реальність і реставрація історичних зображень. Проте він залишається надзвичайно складним через необхідність врахування складних взаємозв'язків між текстурами, освітленням, перспективою та структурою сцени, особливо при видаленні великих або критично важливих об'єктів, де важливо логічно інтегрувати новий контент у композицію зображення, зменшуючи ризик появи артефактів або генерації неправдоподібних деталей. Все це робить об'єктно-контекстний імпейнтинг науково та інженерно вимогливим завданням.

Дослідження напрямку. Розвиток даного напрямку включає розробку нових архітектур нейронних мереж, оптимізацію функцій втрат, використання нормалізаційних потоків і методів трансформерів для підвищення точності, швидкості та ефективності алгоритмів. Результати досліджень демонструють прогрес у розвитку генеративних моделей та підтверджують перспективність подальшого вдосконалення методів імпейнтингу для досягнення високої якості відновлення контенту. В статті [1] запропоновано метод імпейнтингу, який дозволяє видаляти об'єкти із зображень на основі текстових інструкцій. В [2] показано каскадну модифікацію GAN для імпейнтингу, що враховує структуру об'єктів на зображенні. В [3] автор описує метод, що дозволяє безшовно видаляти об'єкти разом із їхніми візуальними ефектами. Ці статті охоплюють різні аспекти задачі видалення об'єктів із зображень, від використання текстових інструкцій до складних нейронних мереж з урахуванням об'єктів. Але описані методи мають ряд недоліків, а саме: у дуже складних сценах або при великій кількості одночасно видалених об'єктів модель може генерувати артефакти або спотворені деталі; використання великих моделей і наборів даних робить навчання та інференс ресурсомісткими [4-5].

Постановка задачі. Сучасні методи імпейнтингу зазвичай базуються на глибоких нейронних мережах, зокрема:

- GAN (Generative Adversarial Networks) – створюють реалістичні деталі для заповнення порожніх ділянок.

- Transformers і Diffusion-моделі – забезпечують більш узгоджене і контекстуально точне заповнення.

- Маскування з контекстним увагою (contextual attention) – враховує найближчі пікселі та структуру зображення для кращої інтеграції заповненого фрагмента.

Для реалізації системи “розумного” видалення об'єктів із зображень, яка базується на використанні генеративних конкурентних мереж із унікальною функцією втрат, необхідно виконати кілька кроків: визначити джерела даних, проаналізувати їх придатність, оцінити характеристики обраного датасету та обґрунтувати вибір методу машинного навчання [5].

GAN мережа побудована на двох основних компонентах: генераторі та дискримінаторі. Генератор складається з двох підмереж:

- SINet (Semantic Inferring Network): Екстрагує семантичні ознаки з вхідного зображення, враховуючи замасковані області. Для цього використовуються згорткові шари зі послідовним збільшенням кількості каналів.

- GPNet (Global Perceiving Network): Захоплює глобальний контекст зображення, тому обробляє як приховані, так і відкриті ділянки. Розширені згортки зберігають просторову інформацію. Схожа будова на SINet, тільки без деконволюційних шарів.

Після обробки вхідних даних у згаданих мережах, вихідні значення комбінуються: оригінальне зображення разом з маскою об'єднується з відновленим SINet результатом, формуючи повне представлення зображення. Далі воно передається в GPNet для подальшого вдосконалення. Така методика забезпечує якість штучних даних, роблячи дискримінатору складнощі у визначенні походження зображення. Тому він має бути достатньо ефективним, щоб обробляти складні патерни.

Дискримінатор функціонує як бінарний класифікатор, архітектурно може бути представлений як традиційна згорткова нейронна мережа (CNNs). Правильно підібраний дизайн нейромережі, є істотним, оскільки його здатність до розмежування ознак прямо впливає на коректну роботу системи (запобігає затухання градієнта) [6].

Дискримінатор мінімізує від'ємну логістичну ймовірність правильної класифікації як реальних, так і згенерованих зразків, що можна виразити через формулу:

$$\nabla \theta_d \frac{1}{m} \sum_{i=1}^m [\log D(x^{(i)}) + \log (1 - D(G(z^{(i)})))] ,$$

де $D(x^{(i)})$ – ймовірності того, що дискримінатор

D класифікує оригінальне зображення $x^{(i)}$ як реальне;

$\log (1 - D(G(z^{(i)})))$ – логарифм ймовірності того, що дискримінатор D класифікує згенероване зображення $G(z^{(i)})$ як не реальне.

На основі результату дискримінатора, функція втрат генератора оцінює свою роботу для нейронної мережі.

Використання генеративних конкурентних мереж (GAN) для задач видалення об'єктів із зображень супроводжується низкою проблем, що обмежують якість відновлення контенту. До основних труднощів належать поява артефактів та розмиття у згенерованих областях, непослідовність кольорових та текстурних характеристик порівняно з навколишнім середовищем, а також семантична невідповідність відновлених фрагментів контексту сцени. Значним викликом є залежність від якості та різноманітності навчальних даних, що обумовлює низьку здатність моделі до узагальнення у випадку нестандартних вхідних даних. Додатково проблемними залишаються висока обчислювальна складність, потреба у великих обсягах даних та нестабільність процесу навчання, що може призводити до колапсу моделі. У сукупності ці фактори вказують на необхідність подальшого вдосконалення архітектур і методів оптимізації GAN для підвищення якості та надійності реконструкції зображень [7-8].

Метою роботи є вдосконалення методу на основі GAN для генерації зображень, використовуючи функції втрат з оптимальними вагами. На основі проведених досліджень вдосконалено метод на основі GAN, який вирішуватиме вище обговорені проблеми шляхом використання нової функції втрат, що комбінує адверсарні L1 та L2 втрати з оптимальними вагами. Такий підхід підвищить якість відновлення частин зображень, до того ж забезпечить стабільність навчання моделі.

Для GAN моделі розроблено наступні формули, які контролюють навчання генератора та дискримінатора:

$$L_{gen} = -E_x[D(G(x))] + \lambda_1 L_2 + \lambda_2 L_1 ,$$

$$L_{gen} = E_x[L_{real}(D(G(x)))] + E_z[L_{fake}(D(G(x)))] + \lambda_1 L_2 + \lambda_2 L_1 ,$$

де L_{gen} – функція помилки генератора;

L_{real} – функція втрат для реальних зображень;

E_x – це очікування, що зображення є реальним;

$D(G(x))$ – вихід дискримінатора D , який приймає згенероване зображення $G(x)$ як вхід;

E_z – очікування за латентними змінними z ;

$\lambda_1 L_2$, $\lambda_2 L_1$ – регуляризаційні члени із коефіцієнтами.

Завдяки системі інтелектуального видалення, розробленої у межах даного дослідження, користувачі мають змогу модифікувати різні зображення із мінімальними артефактами, адаптовуючи їх до власних потреб.

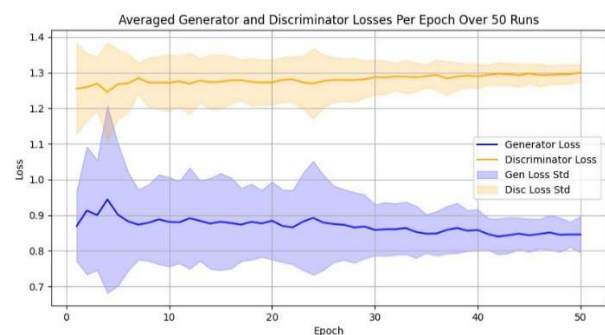


Рис. 1. Графік базової функції втрат під час навчання

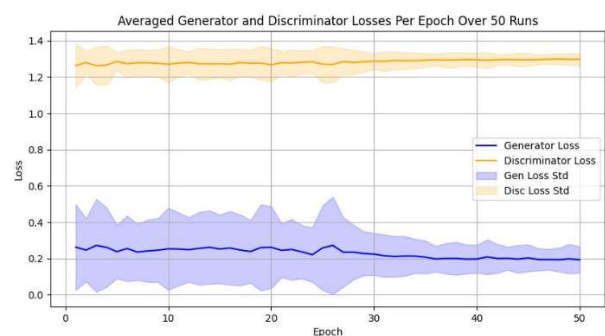


Рис. 2. Графік кастомної функції втрат, розробленої в цьому дослідженні

Проводилося дослідження на платформі Google Colab, вона надає можливість використовувати різні графічні процесори, що підходять для тренування базової генеративної моделі на наборі даних MNIST (складається із зображень рукописних цифр від 0 до 9. Незважаючи на свою простоту, дозволяє швидко перевірити експериментальні концепції).

При тому, що архітектура моделі залишається незмінною протягом навчання, функція втрат визначає, як оптимізатор коригує ваги та параметри алгоритму. Вона вчиться на основі результатів, які отримує з функцією втрат (чим більше значення втрат, тим сильніші будуть зміни в параметрах) чим точніша модель – тим менші втрати. Оновлюються параметри, зменшуючи втрати в наступних ітераціях і покращуючи алгоритм в цілому.

Розробка додає елементи кастомізації, поєднує різні типи втрат з можливістю налаштування ваг для кожної з них. Така комбінація впроваджує новизну для реалізації, хоча основні компоненти широко використовуються.

Створена унікальна функція втрат, може бути використана для задачі image inpainting, бо адверсальна втрати допомагає генератору створювати більш реалістичні зображення. А L1 і L2 втрати можуть допомогти досягати високої якості відновлення, зменшуючи різницю між оригінальними та відновленими ділянками.

Результати, отримані в рамках дослідження, демонструють значне покращення якості відновлення зображень за допомогою розробленої системи на основі GAN. У порівнянні з базовими критеріями втрат для задач відновлення. Зокрема, система показала покращення в середньому на 65% у значеннях функції втрат для генератора, порівнюючи із базовою моделлю (з 0.9 до 0.3). Порівнюючи з методом циклічної втрати, покращення зросли в межах 2-3%. При цьому вона зберегла високу ефективність за часом обчислень та стабільність роботи дискримінатора. Загальний час виконання однієї ітерації в середньому зріс лише на 0.94% (з 635 до 641 секунди) для базової реалізації та на 2% для методу циклічної втрати.

Приклад, наведений на рис.1 демонструє результат видалення об'єкта із зображення. Варто зазначити, що колір собаки схожий із заднім фоном, що частково посприяло якіснішому заповненню. Водночас незважаючи на досить велику маску, система впоралася із завданням на належному рівні.



а



б



в

Рис. 3. Результат усунення великої частини зображення:

а – початкове зображення;
б – зображення з маскою; в – результат видалення

Висновки. Проаналізовано методи, які є основою для розробки системи зафарбовування зображень. Досліджено актуальні підходи для видалення та створення текстури з метою заповнення потрібних частин. Удосконалено метод генерації контенту із застосуванням унікальної функції втрат, яка за показниками точності та часу виконання перевищує розглянуті конкурентні реалізації. Розроблено

систему видалення об'єктів із цифрових зображень, що використовує алгоритм GAN на базі Python та його бібліотек. Система орієнтована на широкий спектр користувачів: від бізнес-компаній до кінцевих споживачів, які прагнуть покращити обробку візуальних даних за допомогою інструментів редагування.

Результати демонструють задовільну якість видалення структур із зображень з мінімальними артефактами, навіть при низькій роздільній здатності за короткий час. Однак, точність відновлення може знижуватися при роботі із складними чи деталізованими зображеннями. У дослідженні була розроблена методика видалення об'єктів за допомогою інтеграції сучасних типів архітектур нейронних мереж з унікальною функцією втрат.

Розроблене рішення дозволяє автоматизувати процеси редагування зображень, що зменшує витрати часу та ресурсів, необхідних для виконання таких завдань вручну. Система забезпечує доступ до інноваційних технологій для широкого кола користувачів та знижує витрати на спеціалізоване програмне забезпечення. Завдяки цьому розробка має не лише наукове, а й соціально-економічне значення, сприяючи доступності обробки візуальної інформації. Здійснено дослідницьку роботу для оптимізації методу GAN та побудовано програмну реалізацію системи, яка дозволяє у короткий термін усувати небажані компоненти із зображень. Результати дослідження підтверджують ефективність розробленого рішення. Таким способом користувачі можуть ефективно редагувати графічний контент відповідно до персональних вподобань.

References

1. Inst-Inpaint: Instructing to Remove Objects via Text Prompts. Cornell University. URL: <https://arxiv.org/abs/2304.03246> (data accessed: 05/09/2025)
2. Image Inpainting with Cascaded Modulation GAN and Object-Aware Training. Cornell University. URL: <https://arxiv.org/abs/2203.11947> (data accessed: 04/09/2025)
3. Remove Objects and Their Effects in Images with Paired Video-Frame Data. Cornell University. URL: <https://arxiv.org/abs/2501.07397> (дата звернення 05.09.2025)
4. Barglazan A.-A., Brad R., Constantinescu C. Image Inpainting Forgery Detection: A Review. J. Imaging. 2024. Vol. 10, No. 2. P. 42. URL: <https://doi.org/10.3390/jimaging10020042>
5. What is a GAN?. Amazon. URL <https://aws.amazon.com/what-is/gan/> (data accessed: 02/09/2025)
6. GPNet: Simplifying Graph Neural Networks via Multi-channel Geometric Polynomials. Cornell University. URL: <https://arxiv.org/abs/2209.15454> (data accessed: 05/09/2025)
7. Goodfellow I., Bengio Y., Courville A. Deep Learning. Cambridge, MA: MIT Press, 2016. P.775.
8. Ma Y., Liu X., Bai S., Wang L., He D., Liu A. Coarse-to-Fine Image Inpainting via Region-wise Convolutions and Non-Local Correlation. The 28th International Joint Conference on Artificial Intelligence (IJCAI-19). 2019. P. 3123–3129. URL: <https://doi.org/10.24963/ijcai.2019/433>

Semankiv M.V., Tsikhub O.V. Improvement of image inpainting methods based on generative models

The volume of visual information surrounding us is constantly growing, which creates the need for intuitive tools for its processing. Research aimed at improving algorithms for adding or removing objects from graphical content is highly relevant and has significant practical potential. Systems employing such algorithms can be applied across various domains—from cinematography and photography to the restoration of artworks and even medicine. Image inpainting technology enables the automatic reconstruction of missing parts of an image, saving users time and effort. Despite the abundance of existing solutions for reconstruction tasks, there is still a lack of systems that comprehensively leverage generative models to achieve high-quality content restoration. This highlights the importance of research in this field, focusing on improving current algorithms. The assessment of the current state of the field shows considerable progress in the development of generative models. At the same time, further refinement is needed to enhance universality, optimize resource usage, and reduce computational costs. This work investigates methods of Generative Adversarial Networks (GANs) for image inpainting tasks and evaluates the effectiveness of existing approaches. A content generation method has been improved by introducing a unique loss function, which surpasses the considered competitive implementations in terms of both accuracy and execution time. Software has been developed to address practical tasks. Work has been carried out on a GAN-based system for graphical content restoration, with improvements achieved through mathematical tools. The results obtained in the course of the research demonstrate enhanced image restoration quality with the proposed GAN-based system, including high-quality object removal with minimal artifacts, even at low resolution and within a short processing time. However, it should be noted that restoration accuracy may decrease when dealing with complex or highly detailed images. The practical significance of this research extends to all fields where visual data

processing is crucial. It lies in addressing the problem of low-quality object removal from images and advancing algorithms for replacing and removing specific regions in images using Generative Adversarial Networks.

Keywords: *inpainting, GAN-network.*

Ціхун Олександр Васильович – магістр, Карпатський національний університет імені Василя Стефаника (м. Івано-Франківськ)
oleksandr.tsikhun@gmail.com

Семаньків Марія Василівна – к.т.н., доц. кафедри комп'ютерних наук та інформаційних систем, Карпатський національний університет імені Василя Стефаника (м. Івано-Франківськ)
mariia.semankiv@pnu.edu.ua

Стаття подана 19.09.2025.