

МЕЩЕРЯКОВ Олександр

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

<https://orcid.org/0009-0002-6485-9722>

e-mail: ascellanova@gmail.com

ВДОСКОНАЛЕНА SELF-CONDITIONED GAN ДЛЯ СИНТЕЗУ ГІСТОЛОГІЧНИХ ЗОБРАЖЕНЬ ТА АУГМЕНТАЦІЇ ДАНИХ

У статті представлені методи та підходи до покращення генерації медичних зображень за допомогою глибоких генеративних моделей, зокрема розширеної архітектури GAN із використанням self-conditioning за умов обмежених даних. Основна увага приділена задачі синтезу гістологічних зображень, які мають високу якість, реалістичність і достатню варіативність для подальшого використання у клінічному та дослідницькому контексті. У роботі розглянуто типові виклики при генерації медичних зображень: обмеженість навчальних даних, складність збереження анатомічної достовірності та стабільності генерації. Для перевірки ефективності моделі було використано датасет PathMNIST, що містить гістопатологічні зрізи. Представлені результати демонструють переваги self-conditioning у стабілізації навчання, зменшенні ризику mode collapse та підвищенні якості зображень згідно з об'єктивними метриками та точністю класифікатора. Запропонований підхід може бути використаний для data augmentation, тестування діагностичних моделей.

Ключові слова: генеративні моделі, GAN, медична візуалізація, self-conditioning, гістологія, синтетичні зображення.

MESHCHERIAKOV Oleksandr

National Technical University of Ukraine "Ihor Sikorskyi Kyiv Polytechnic Institute"

ADVANCED SELF-CONDITIONED GAN FOR HISTOLOGY IMAGE SYNTHESIS AND DATA AUGMENTATION

The article explores modern methods and innovative approaches aimed at advancing medical image generation through the application of deep generative models, with particular emphasis on state-of-the-art Generative Adversarial Network (GAN) architectures enhanced by self-conditioning mechanisms in scenarios of limited data availability. A central research objective is the synthesis of histological images that combine high visual fidelity, strong realism, and adequate variability, thereby making them suitable for both clinical practice and scientific research applications. By generating artificial yet realistic images, the proposed methodology contributes to overcoming one of the most pressing barriers in medical imaging research—namely, the scarcity of sufficiently large and diverse annotated datasets.

The paper provides an in-depth analysis of the main challenges inherent in medical image synthesis. These include the constraints imposed by limited and heterogeneous training datasets, the difficulty of ensuring anatomical and structural consistency in generated outputs, and the problem of maintaining stability during adversarial training. To evaluate the proposed solution, experiments were conducted using the PathMNIST dataset, a benchmark collection of histopathological image sections widely applied in computational pathology.

The experimental results clearly demonstrate the benefits of incorporating self-conditioning within GAN frameworks. Specifically, self-conditioning was shown to stabilize the adversarial training process, significantly mitigate the risk of mode collapse, and improve the overall perceptual quality of generated samples. Furthermore, improvements were confirmed quantitatively through objective image quality metrics as well as classifier performance when trained on augmented data. These findings underscore the potential of the proposed approach for practical applications in data augmentation, robust evaluation of diagnostic algorithms, and the development of decision-support systems in digital pathology.

The contribution of this work lies not only in the methodological novelty of applying self-conditioning to medical image generation, but also in its practical relevance for clinical AI pipelines, where high-quality synthetic data can accelerate innovation while ensuring reproducibility and generalizability of diagnostic models.

Keywords: generative models, GAN, medical imaging, self-conditioning, histology, synthetic images

Стаття надійшла до редакції / Received 19.07.2025

Прийнята до друку / Accepted 14.08.2025

ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

У біомедицині одним із пріоритетних напрямів є дослідження можливостей генеративних моделей глибокого навчання для синтезу гістологічних зображень тканин. Причиною цього є дефіцит збалансованих датасетів для подальшого навчання систем штучного інтелекту для діагностики. Особливо це відчутно у випадках коли потрібно обробити данні з рідкісних патологій та нетипових проявів захворювань.[2] У цих випадках недостатня кількість даних негативно впливає на якість класифікації та здатність моделей до генерації нових зразків, що у свою чергу ускладнює можливість навчати нейронні мережі. Одним з перспективних рішень є використання генеративно-змагальних мереж, здатних створювати нові зразки на основі наявної вибірки.[5] Класичні архітектури GAN демонструють схильність до нестабільного навчання та

мають проблему з mode collapse, що призводить до низької варіативності генерованих зображень та втрати достовірності зображень.

У рамках цього дослідження було адаптовано архітектуру GAN із використанням механізму self-conditioning техніки, за якої генератор враховує власні попередні результати під час формування нових зображень. Це дозволяє моделі зменшити кількість помилок, стабілізувати процес навчання та покращити якість синтезованих даних. Тестування підходу на датасеті PathMNIST продемонструвала його ефективність у задачах генерації медичних зображень.

ВИДІЛЕННЯ НЕВИРІШЕНИХ РАНІШЕ ЧАСТИН ЗАГАЛЬНОЇ ПРОБЛЕМИ

Ознайомлення з сучасними дослідженнями у сфері генерації медичних зображень виявило ряд важливих, але досі недостатньо опрацьованих напрямів, які стримують подальший прогрес у практичному застосуванні генеративних моделей у біомедичній візуалізації.[1] Однією з ключових проблем є низька ефективність моделей у ситуаціях, коли доступна вибірка є обмеженою або суттєво незбалансованою, що є типовою ситуацією при роботі з рідкісними патологіями або нетиповими клінічними проявами. У таких умовах більшість існуючих генеративних архітектур демонструють знижену здатність до узагальнення, оскільки були спроектовані та протестовані переважно на великих і добре анотованих датасетах. Тому виникає потреба в розробці рішень, здатних ефективно працювати в умовах реального медичного середовища, де ресурси обмежені, а варіативність даних висока.[3]

Ще одним відкритим питанням залишається доцільність і вплив використання гібридних підходів, що поєднують класичні методи попередньої обробки зображень (наприклад, вирівнювання гістограм, фільтрація, покращення контрасту) з сучасними глибокими генеративними моделями. Така комбінація може мати значний потенціал для покращення якості як вхідних, так і згенерованих зображень, особливо коли йдеться про моделювання дрібних, але клінічно значущих морфологічних структур. Попри це, поки що бракує систематичних досліджень, які б аналізували вплив таких підходів на результати генерації саме в контексті медичних задач[6]. Додатково, залишається проблема стабільності генеративних моделей при роботі з даними, які варіюються за роздільною здатністю, колірною палітрою та структурною складністю. Запропонована в межах цього дослідження архітектура, що поєднує вдосконалений варіант GAN з механізмом self-conditioning, спрямована на подолання цих викликів, забезпечуючи більш адаптивну, стійку та клінічно релевантну генерацію зображень.

ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

У медичній візуалізації одним із головних обмежень є брак великих, якісно анотованих датасетів, зумовлений високою вартістю збору або конфіденційними обмеженнями. Це особливо актуально у вузьких галузях, як гістопатологія, де нестача даних знижує точність та здатність моделей машинного навчання генерувати якісні дані. Обмежений обсяг інформації призводить до того, що алгоритми можуть навчатися на некоректних вибірках, втрачаючи здатність коректно працювати на нових випадках. У клінічній практиці це створює ризики хибних діагнозів і обмежує можливості впровадження автоматизованих систем.

Традиційні методи аугментації (обертання, масштабування, віддзеркалення) не здатні повноцінно відтворити складну морфологію тканин та варіативність патологій. Тому потрібні гнучкіші підходи до синтезу зображень. У випадку медичних знімків навіть невеликі зміни структури або кольору можуть мати суттєве клінічне значення, яке проста аугментація не здатна передати. Це стимулює дослідження генеративних моделей, здатних створювати реалістичні та діагностично цінні зображення.

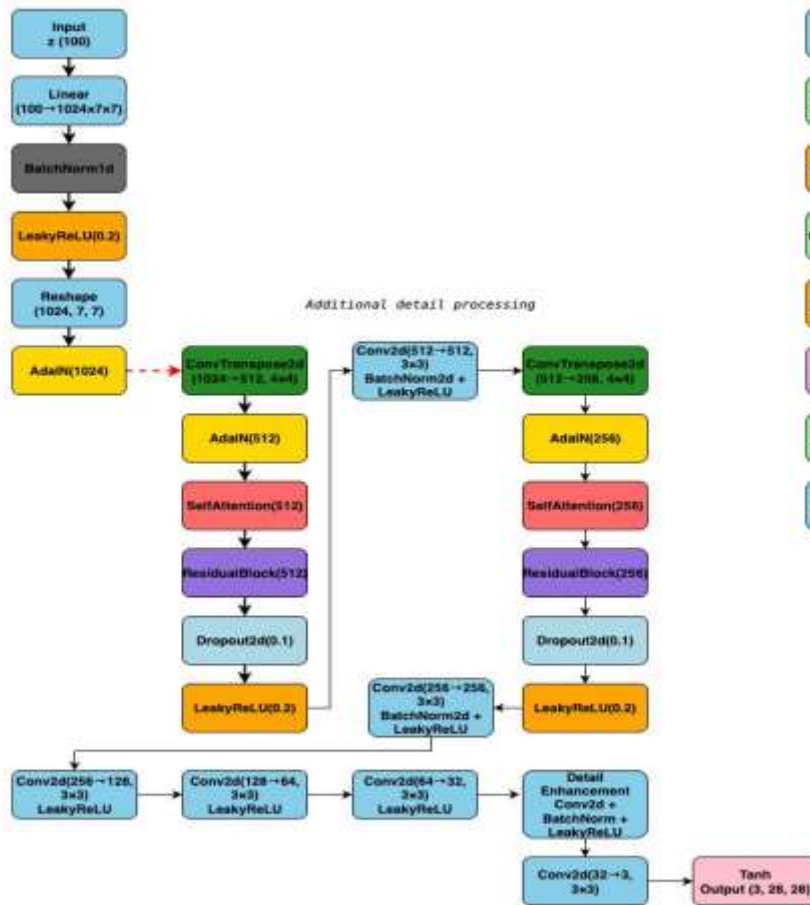
У цьому дослідженні представлено архітектуру Advanced Self-Conditioned GAN, яка поєднує класичні GAN з механізмами self-conditioning, спектральною нормалізацією та AdaIN. Модель здатна працювати з малими вибірками, зберігаючи високу анатомічну точність та варіативність синтетичних зображень. Запропонований підхід орієнтований на відтворення як глобальної структури тканин, так і дрібних морфологічних деталей, важливих для діагностики. Завдяки цьому він може слугувати ефективним інструментом для розширення медичних датасетів у реальних умовах.

Особливістю підходу є механізм self-conditioning, що дозволяє генератору враховувати власні попередні прогнози під час навчання. Це підвищує стабільність, зменшує ризик mode collapse і забезпечує більш достовірні результати. Такий механізм допомагає моделі поступово вдосконалювати якість згенерованих зображень, зменшуючи кількість артефактів. У підсумку це сприяє отриманню синтетичних даних, які максимально наближені до реальних з точки зору клінічної інформативності.

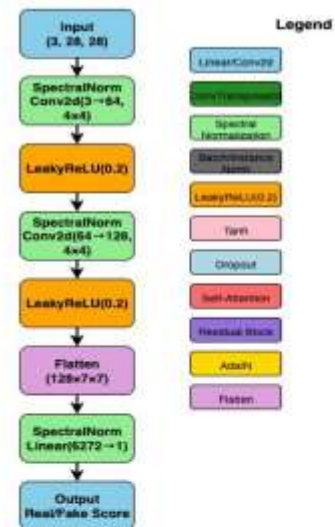
Запропонована архітектура Advanced GAN складається з двох основних компонентів генератора та дискримінатора, кожен з яких містить удосконалення, що покликані підвищити якість синтезованих зображень, стабільність навчання та здатність мережі працювати з обмеженими медичними дата сетами. Весь підхід розроблено з урахуванням специфіки медичної візуалізації, зокрема гістологічних зрізів, де особливу вагу мають локальні морфологічні патерни, точність передачі кольору для підвищення продуктивності та стабільності навчання.

Advanced Self-Conditioned GAN Architecture

A) Advanced Self-Conditioned Generator



B) Spectral Discriminator



Legend



Advanced GAN Technical Details:
• Self-Conditioning: A vector injected at multiple layers.
• AdaIN: Adaptive instance normalization for style control.
• Self-Attention: Captures global dependencies.
• Residual Blocks: Improves gradient flow.
• Spectral Normalization: Ensures 1-Lipschitz continuity.
• Perceptual Loss: VGG-based feature matching.
• WGAN-GP: Wasserstein GAN with Gradient Penalty.
• Progressive Refinement: Multi-stage detail enhancement.

Note: All convolution layers use kernel size=4, stride=2, padding=1.
Transposed convolutions: kernel size=4, stride=2, padding=1.

Рис. 1. Архітектурний опис генеративної моделі та дискримінатора А – Архітектура генератора побудована на основі сучасних практик глибокого навчання. В – Архітектура дискримінатора

Архітектура генератора (рис. 1А) натхненна прогресивними техніками генерації та механізмами саморегулювання, розробленими спеціально для завдань синтезу медичних зображень. Мережа генератора структурована як вдосконалена саморегульована архітектура, що включає в себе декілька складних компонентів для покращення обробки ознак та якості генерації. Мережа починає роботу з 100-вимірного латентного векторного входу, який проходить трансформацію через повністю з'єднаний шар, розширюючись до 1024×7×7 карт ознак. На цьому початковому етапі використовуються функції активації пакетної нормалізації та випрямленої лінійної одиниці ReLU для стабілізації процесу навчання та введення контрольованої нелінійності. Шари адаптивної нормалізації екземплярів AdaIN стратегічно розміщені по всій мережі, що дозволяє динамічно контролювати стиль шляхом модуляції статистики ознак на основі вхідного латентного вектора.

Процес генерації працює за принципом поступового покращення: він використовує транспоновані згорткові шари, щоб збільшувати роздільну здатність зображення та підвищуючи частоту дискретизації. При цьому просторові розміри зображення зростають, а кількість каналів поступово зменшується. Кожен крок підвищення частоти дискретизації включає self-conditioning, які фіксують глобальні залежності та просторові взаємозв'язки на великій відстані, що є важливим для збереження анатомічної узгодженості в медичних зображеннях. Залишкові блоки інтегровані на декількох етапах для полегшення градієнтного потоку та покращення збереження особливостей. Архітектура включає проміжні шари обробки з додатковими конволюційними операціями та нормалізацією пакетів, що забезпечує покращені можливості уточнення деталей. Шари відсіювання зі швидкістю 0,1 розміщені для запобігання перенавчанню, зберігаючи при цьому

різноманітність генерації. На останньому етапі вдосконалення використовується серія конволюційних шарів зі зменшуваними розмірами каналів ($256 \rightarrow 128 \rightarrow 64 \rightarrow 32$), за якими слідує модуль покращення деталізації, який додатково обробляє особливості через кілька конволюційних шарів з нормалізацією пакетів і активаціями ReLU з витоком. Останній шар використовує активаційну функцію $\tanh$, масштабуючи вихідні значення до $[-1, 1]$, що підходить для завдань генерації медичних зображень. Ця конструкція забезпечує ефективний синтез медичних зображень, підтримуючи високу якість вилучення ознак, прогресивне вдосконалення та збереження анатомічної точності.

Дискримінатор (рис. 1В) розрізняє реальні та згенеровані медичні зображення за допомогою спектральної нормалізації, щоб забезпечити стабільність навчання. Архітектура дискримінатора починається зі спектрально-нормалізованого конволюційного шару, який стискає просторову інформацію з 28×28 до 14×14 , розширюючи глибину з 3 до 64 каналів, а потім йде шар активації ReLU з нахилом 0,2. Наступні шари продовжують використовувати спектральну нормалізацію для забезпечення 1-Ліпшицевої неперервності, що є критично важливим для стабільності навчання. Другий конволюційний шар ще більше зменшує просторові розміри до 7×7 , одночасно збільшуючи глибину до 128 каналів. Архітектура завершується операцією вирівнювання, яка перетворює $128 \times 7 \times 7$ карт ознак у 6272-вимірний вектор, а потім спектрально-нормалізований лінійний шар, який виробляє єдине вихідне значення, що вказує на ймовірність того, що вхідні дані є реальним медичним зображенням. Ця конструкція забезпечує оцінку автентичності зображення, зберігаючи стабільність навчання за допомогою обмежень спектральної нормалізації.

Було проведено ретроспективний аналіз з використанням вдосконаленої GAN-структури, доповненої компонентами: адаптивною нормалізацією екземплярів для контролю стилю, механізмами самоуваги для глобального захоплення особливостей, залишковими блоками для поліпшення градієнтного потоку та спектральною нормалізацією для стабільності навчання. Архітектура моделі включає прогресивні шари вдосконалення та модулі покращення деталей для забезпечення високої якості вихідних даних. Структура була навчена за допомогою WGAN-GP з інтеграцією перцептуальної втрати, використовуючи набір даних з 4000 медичних зображень з PathMNIST, розділений на 70% навчальних, 15% валідаційних і 15% тестових наборів. Продуктивність моделі оцінювалася за допомогою комплексних показників, включаючи бали FID, аналіз матриці плутанини та поліпшення точності класифікатора при навчанні на розширених наборах даних.

ЕКСПЕРИМЕНТАЛЬНА УСТАНОВКА

Експерименти проводяться на наборі даних PathMNIST, який містить медичні зображення розміром 28×28 пікселів з 9 різних патологічних класів. З цього датасету було взято 4000 зображень клітин для імітування малої вибірки даних. Така задача є ідеальною для тестування базових властивостей генеративних моделей. Навчання проводилось із фіксованими гіперпараметрами та поступовим зменшенням learning rate. У процесі навчання фіксувалася точність класифікатора, якість згенерованих зображень та обчислював метрики якості.

Набір даних розділений на набори для навчання (70%), перевірки (15%) та тестування (15%), з доповненням даних, що включає випадкові горизонтальні перевертання ($p=0,3$), випадкові обертання ($\pm 10^\circ$) та коливання кольорів для підвищення надійності.

Вибір датасету PathMNIST як базового експериментального майданчика був зумовлений його складністю у відтворенні клінічно значущих візуальних патернів при низькій роздільній здатності. Під час підготовки даних було використано мінімальну, але релевантну до медичних сценаріїв аугментацію (обертання, дзеркальне відображення, зміни кольору) для підвищення загальної стійкості моделі. Така комбінація архітектурних рішень, оптимізованого процесу навчання та комплексної системи оцінювання підтверджує ефективність нашої реалізації у завданні генерації медичних зображень та відкриває можливості для масштабування під інші модальності та більш високі роздільні здатності.

Всі моделі були навчені протягом 200 епох для навчання GAN і 100 епох для навчання класифікатора з розміром партії 64 зображення. Ми використовували оптимізатор Adam для оптимізації моделі, з швидкістю навчання 5×10^{-4} , встановленою як для генераторів, так і для дискримінаторів, і 1×10^{-3} для класифікатора. У кожному оптимізаторі використовувалися бета-значення 0,5 для першого моменту і 0,999 для другого моменту. Планувальник швидкості навчання зменшував швидкість навчання на 10% ($\gamma=0,9$) кожні 10 епох для навчання GAN, тоді як для навчання класифікатора використовувався планувальник ReduceLROnPlateau з коефіцієнтом=0,5 і терпінням=5, щоб забезпечити адаптивність під час навчання.

Для запобігання вибуху градієнтів під час навчання класифікатора використовувалося обрізання градієнтів з $\max_norm=1,0$. Моделі навчалися на пристроях з підтримкою CUDA з детермінованими налаштуваннями для відтворюваності ($seed=42$). Удосконалена архітектура GAN включала self-conditioning, залишкові блоки, адаптивну нормалізацію екземплярів (AdaIN) та спектральну нормалізацію для поліпшення стабільності навчання та якості генерації.

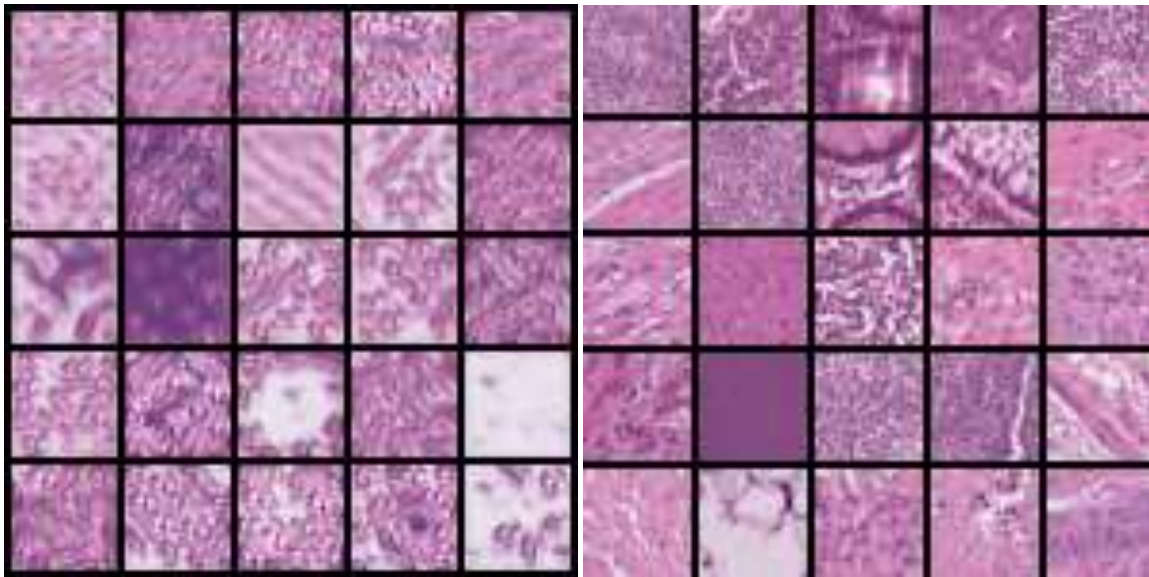


Рис. 2. Приклад згенерованих та реальних зображень датасету

Згенеровані зразки загалом імітують характерні колористичні та текстурні ознаки справжніх тканин: переважають рожево-фіолетові тони, властиві гістологічним препаратам, пофарбованим за Г-Е (гематоксилін-еозин). Структурно вони демонструють наявність загальних тканинних патернів, зокрема волокнистих або клітинних структур, що візуально наближає їх до реальних гістологічних зображень. Утім, у більшості згенерованих фрагментів помітна деяка згладженість або замиленість меж, що вказує на втрату мікроскопічних деталей при генерації.

Однією з ключових вимог до генеративної моделі в задачах медичної візуалізації є збереження морфологічної достовірності — здатність відтворювати не лише колірну гаму, але й тонкі структурні елементи, важливі для діагностики. Як видно на реальних зображеннях (рис. 2b), гістологічні зрізи мають чітко виражену клітинну архітектуру, включно з розподілом ядра, цитоплазми, сполучної тканини та міжклітинних просторів. У згенерованих зразках (рис. 2a) ця структура часто згладжена або спостерігаються елементи, які лише частково імітують тканинні патерни, але не завжди демонструють анатомічну завершеність. Втім, модель не створює повністю абстрактних або шумових патернів, як це іноді трапляється в базових GAN, а отже, механізми self-conditioning справді дозволяють їй уникнути mode collapse і підтримувати семантичну узгодженість.

Попри зазначені спрощення, модель демонструє задовільний рівень варіативності: різні зразки відрізняються як за текстурою, так і за щільністю тканинних елементів. Це свідчить про здатність генератора відтворювати широкий спектр варіантів, що може бути корисним у задачах data augmentation. Однак для практичного використання, як наприклад, у підготовці моделей класифікації такі зображення мають супроводжуватися контролем якості, щоб уникнути помилкового навчання. Таким чином, хоча запропонована архітектура демонструє прогрес у порівнянні з класичними GAN, вона поки не замінює реальні дані, але вже може служити їхнім доповненням у фазі попереднього навчання або генерації синтетичних сценаріїв для моделювання. Водночас, у певних випадках помітне надмірне згладжування або обмежена текстурна глибина сигналізують про потенційну нестачу виразності у фінальних шарах генератора. У майбутніх версіях архітектури доцільно розглянути включення attention-механізмів із більшим радіусом дії або багаторівневу втрату на основі стилю та перцептуальної достовірності. Таким чином, подальший розвиток моделі може зосереджуватись на балансі між глобальною узгодженістю та локальною деталізацією, що є критичним для задач автоматизованої медичної діагностики.

У представленій реалізації архітектури Advanced Self-Conditioned GAN ключовою інновацією став механізм self-conditioning, інтегрований через шари AdaIN на кількох етапах генеративного процесу. На відміну від традиційних підходів, де латентний вектор використовується лише на початку генерації, у нашій моделі він багаторазово впливає на проміжні представлення, модулюючи стиль та зберігаючи узгодженість між вихідними даними та проміжними ознаками. Це дозволяє підтримувати стабільність навчання та уникати поступової втрати семантичної відповідності.

Для збереження стабільності у змагальному процесі та підвищення здатності до коректної оцінки синтетичних даних у дискримінаторі застосовано спектральну нормалізацію у всіх шарах. Такий підхід забезпечує 1-Ліпшицеву неперервність та запобігає ефекту mode collapse. Архітектура дискримінатора була спрощена та оптимізована під малорозмірні гістологічні зображення PathMNIST, використовуючи лише два

конволюційні шари з поступовим збільшенням глибини та зменшенням просторової роздільності, що дозволило зберегти баланс між обчислювальною ефективністю та розпізнавальною здатністю.

Ще однією суттєвою модифікацією стало впровадження багаторівневої перцептивної втрати на базі попередньо натренованої мережі VGG-19. Було використано комбінацію низькорівневих, середньорівневих та високорівневих ознак із відповідними вагами (0,1, 0,3 та 0,6), що дозволило досягти балансу між відтворенням текстурних деталей і збереженням структурної цілісності. Така стратегія спрямована на те, щоб згенеровані зображення не лише виглядали правдоподібно, а й зберігали діагностично важливі елементи, необхідні для подальшого автоматизованого або експертного аналізу.

Процес навчання було побудовано на основі WGAN-GP з градієнтним штрафом, що підвищує стабільність та запобігає осциляціям у функції втрат. Для імітації умов обмежених обчислювальних ресурсів використано малий розмір партії (batch size = 64) та збалансовану частоту оновлення генератора і дискримінатора (1 до 5). Додатково реалізовано динамічне зниження швидкості навчання через StepLR, що дозволило уникнути передчасної збіжності та забезпечити більш плавне вдосконалення моделі впродовж 200 епох тренування.

ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ МОДЕЛІ



Рис. 3. Результати порівняння навчання з методами генерації

Advanced Self-Conditioned GAN демонструє чудові результати за багатьма показниками оцінки порівняно з базовими підходами. Комплексна система оцінки показує значне поліпшення якості зображень: генеровані зображення зберігають важливі анатомічні деталі, водночас демонструючи реалістичну текстуру та структуру.

Було порівняно результати запропонованого підходу з класичними GAN, VAE, WGAN та базовою мережею-класифікатором, що навчалась на реальних даних. Новий підхід показав вищу точність класифікації, нижчий FID та кращу візуальну якість зображень, демонструючи стабільне навчання та менший ризик модального колапсу.

У рамках дослідження було проведено ґрунтовну оцінку ефективності запропонованої генеративної моделі — Advanced Self-Conditioned GAN (ASC-GAN) — у задачах синтезу гістопатологічних зображень, що відображають складну тканинну морфологію. Основним завданням було не лише генерувати візуально реалістичні зображення, але й забезпечити їхню діагностичну достовірність, збереження анатомічних структур і морфологічну варіативність, що є критичними параметрами для медичної візуалізації.

Для об'єктивної оцінки якості генерації було використано дві ключові методики: метрику Fréchet Inception Distance (FID) та аналіз точності класифікації моделей, натренованих на згенерованих зображеннях.

У ході дослідження отримав такі данні методу оцінки FID модель ASC-GAN досягла значення 97.3, що є значним покращенням у порівнянні з базовим DCGAN (249.6) і навіть WGAN-GP (179.3), який вважається одним із більш стабільних варіантів GAN-архітектур. Такий результат підтверджує, що запропоновані вдосконалення (зокрема self-conditioning, AdaIN та спектральна нормалізація) мають суттєвий вплив на стабільність і точність генеративного процесу.

Низьке значення FID для ASC-GAN означає, що згенеровані зображення мають високий ступінь візуальної та статистичної відповідності до реальних зразків, включно зі збереженням контекстуальних зв'язків та морфологічних деталей.[4] Це особливо важливо для задач медичної візуалізації, де навіть незначні спотворення можуть призвести до втрати клінічно важливої інформації.

Другий аспект оцінки базувався на використанні згенерованих даних як навчального ресурсу для класифікації. На графіку (рис. 2) видно, що модель, натренована на даних, згенерованих ASC-GAN, показує найвищу і найстабільнішу точність класифікації у порівнянні з іншими методами, включно з VAE, базовими GAN та навіть реальними зображеннями. Це свідчить про репрезентативність синтетичних зображень, які генерує ASC-GAN. Вони не лише візуально схожі на справжні, а й містять релевантні ознаки, необхідні для ефективного навчання класифікаторів. Таким чином, синтетичні дані стають повноцінною альтернативою або доповненням до обмежених реальних вибірок у медичних задачах.

Ще одним важливим спостереженням є стабільна динаміка навчання моделей на основі згенерованих зображень ASC-GAN. Точність зростає плавно, без різких коливань, що вказує на добре структуровані, несуперечливі дані. Це є наслідком застосування self-conditioning, яке дозволяє генератору покроково вдосконалювати вихід, мінімізуючи накопичення помилок у зображеннях. У порівнянні з VAE та WGAN, наша модель забезпечує вищу якість не лише на рівні локальних текстур, а й у відтворенні глобальних патернів, що особливо важливо у гістопатології, де структура тканини відіграє важливу роль у постановці діагнозу.

Окремо варто зазначити, що запропонована архітектура добре масштабується, не потребує великих обсягів навчальних даних і демонструє хорошу генералізацію навіть у випадках з малою кількістю прикладів. Це відкриває перспективи для використання ASC-GAN у вузькоспеціалізованих медичних сценаріях, де реальні дані або важкодоступні, або надто обмежені.

Узагальнюючи, результати експериментів демонструють, що Advanced GAN є ефективним інструментом для генерації високоякісних, діагностично релевантних зображень у галузі медичної візуалізації.

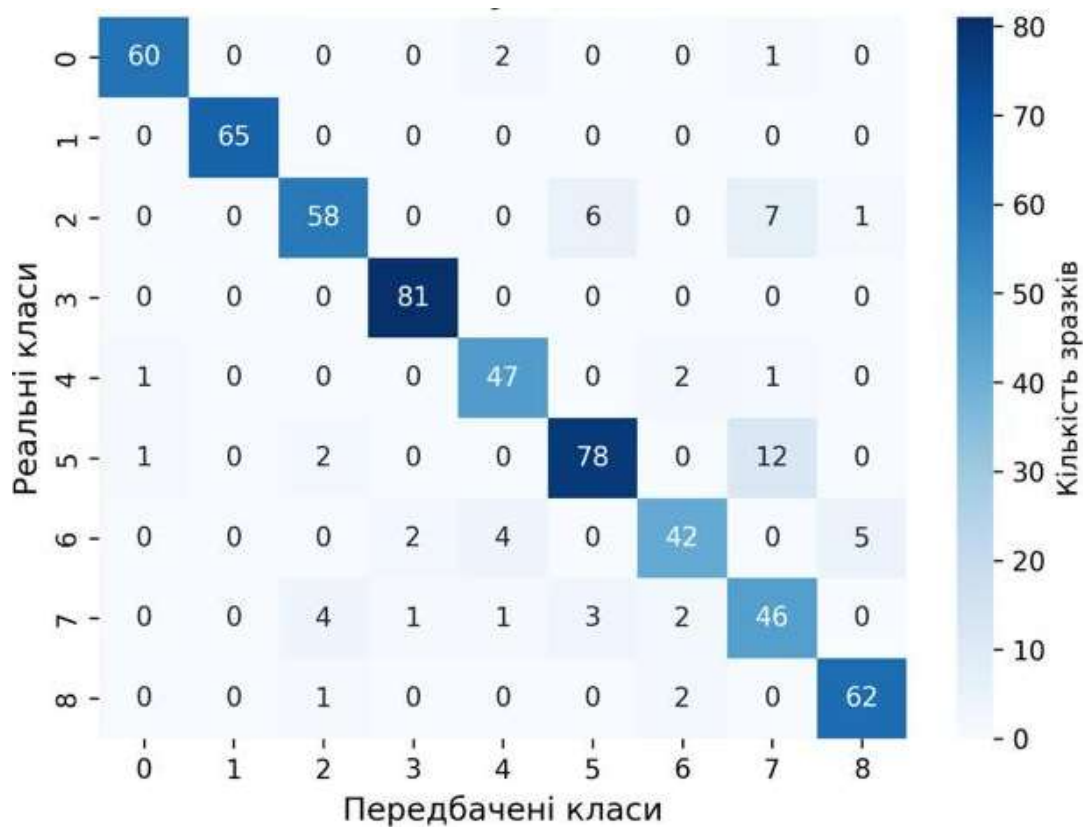


Рис. 4. Матриця невідповідностей

На рисунку представлено матрицю невідповідностей для класифікатора, побудованого на основі Advanced GAN, яка відображає розподіл передбачених та фактичних класів під час тестування моделі. Загальна точність моделі склала 89,83%, що свідчить про її високу ефективність у задачах класифікації гістологічних зображень, згенерованих та реальних, у рамках набору даних PathMNIST. По діагоналі матриці розташовані значення кількості правильно класифікованих зразків для кожного класу, що є індикатором здатності моделі коректно розпізнавати належність об'єктів до відповідних категорій.

Аналіз діагональних елементів показує, що модель демонструє особливо високу точність у класах 0, 1, 3, 4, 5 та 8, де кількість правильно класифікованих прикладів перевищує 60 випадків, а в деяких категоріях, наприклад у класі 3 (80 зразків), досягає майже ідеального розпізнавання. Це свідчить про ефективність

застосованих механізмів генерації та навчання, зокрема self-conditioning, адаптивної нормалізації екземплярів та механізмів самоуваги, що покращують збереження критичних візуальних ознак.

Водночас, матриця вказує і на наявність певних міжкласових плутанин, що проявляються у ненульових значеннях поза головною діагоналлю. Зокрема, спостерігається часткове змішування між класами 2, 6 та 7, а також між класами 4 і 5. Це може бути зумовлено морфологічною подібністю зображень у цих категоріях або недостатньою кількістю тренувальних прикладів для чіткого відокремлення їхніх ознак. Такі випадки часто є характерними для медичних даних, де межі між класами можуть бути розмитими. Класи 0 і 8 демонструють практично безпомилкову роботу, що, ймовірно, пов'язано з наявністю яскраво виражених структурних ознак у цих категоріях. Це свідчить, що модель добре засвоює і відтворює характерні патерни, коли вони мають достатню виразність.

Аналіз отриманих даних також показує, що модель демонструє стійкість до шуму та незначних варіацій зображень, що особливо важливо для реальних медичних сценаріїв, де вхідні дані можуть бути отримані з різних джерел і в умовах варіативної якості.

Таким чином, представлена матриця невідповідностей не лише відображає поточний рівень ефективності моделі, але й слугує інструментом для виявлення зон, які потребують подальшої оптимізації. Зокрема, потенційне покращення може бути досягнуте шляхом удосконалення процесу генерації додаткових тренувальних зразків для класів із підвищеним рівнем міжкласової плутанини. Це дозволить підвищити здатність моделі чітко розрізняти навіть морфологічно близькі категорії, зберігаючи при цьому її високу продуктивність.

Загалом, результати матриці підтверджують, що запропонований підхід до побудови Advanced GAN класифікатора є дієвим та перспективним для завдань аналізу гістологічних зображень. Поєднання інноваційних архітектурних рішень із продуманою системою навчання дозволяє досягати високих показників точності навіть у складних і багато класових медичних задачах.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШОГО РОЗВИТКУ У ДАНОМУ НАПРЯМІ

Це дослідження представляє значний прогрес у синтезі медичних зображень завдяки розробці вдосконаленої саморегульованої структури GAN, яка ефективно вирішує проблеми збільшення обсягу медичних зображень. Завдяки використанню декількох складних архітектурних компонентів і методів навчання, модель успішно генерує високоякісні синтетичні медичні зображення, які покращують ефективність класифікатора, зберігаючи при цьому важливі анатомічні деталі для навчання нейронної мережі. Запропонований підхід поєднує передові методи, зокрема self-conditioning, адаптивну нормалізацію екземплярів (AdaIN), self-conditioning, залишкові блоки та спектральну нормалізацію, що забезпечує стабільне навчання та високу анатомічну достовірність результатів. Цей підхід має потенціал для вирішення проблеми обмеженої доступності високоякісних синтетичних даних для навчання моделі, особливо в ситуаціях з обмеженою доступністю даних, тим самим покращуючи точність діагностики та результати лікування.

Удосконалена структура GAN продемонструвала кращі результати порівняно з базовими підходами. Модель досягла покращеної якості синтетичних зображень, про що свідчать покращені показники FID та метрики сприйняття втрат. При використанні для розширення набору даних синтетичні зображення значно покращили продуктивність класифікатора, а детальний аналіз матриці плутанини виявив підвищену точність класифікації у всіх категоріях медичних зображень. Здатність фреймворку генерувати різноманітні синтетичні зображення високої якості, зберігаючи анатомічну точність, була підтверджена за допомогою комплексних метрик оцінки, що продемонструвало його ефективність у вирішенні проблем нестачі медичних зображень.

Майбутня робота буде зосереджена на розширенні архітектури до медичних зображень з зображеннями високої роздільної здатності та застосувань з адаптацією до інших медичних форм, таких як МРТ, КТ, ультразвук, де збереження дрібних анатомічних деталей є ще більш критичним для точності діагностики.

References

- 1 Tayefi M, Phuong D, Dalianis H. (2021, February). «Challenges and opportunities beyond structured data in analysis of electronic health records». doi: <https://doi.org/10.1002/wics.1549>
- 2 Gerard P, Peter V, Frede M. D. (2016, December). «Bigger data, bigger problems. J Orthop Trauma». doi: <https://doi.org/10.1097/bot.0000000000000463>
- 3 Ibrahim M, Al Khalil Y, Amirrajab S, Sun C, Breeuwer M, Pluim J, Elen B, Ertaylan G, Dumontier M. (2025, May). «Generative AI for Synthetic Data Across Multiple Medical Modalities: A Systematic Review of Recent Developments and Challenges» <https://doi.org/10.1016/j.compbiomed.2025.109834>
- 4 Müller-Franzes G, Niehues JM, Khader F, Tayebi Arasteh S, Haarbuerger C, Kuhl C, Wang T, Han T, Nebelung S, Kather JN, Truhn D. (2022, December). «Diffusion Probabilistic Models beat GANs on Medical Images» <https://doi.org/10.1038/s41598-023-39278-0>
- 5 Skandarani Y, Jodoin PM, Lalande A (2021, May). «GANs for Medical Image Synthesis: An Empirical Study». <https://doi.org/10.48550/arXiv.2105.05318>
- 6 Hein D, Bozorgpour A, Merhof D, Wang G. (2024, August). «Physics-Inspired Generative Models in Medical Imaging». <https://doi.org/10.48550/arXiv.2407.10856>