Vol. 7, No. 1, 2025

Олексій Веретюк[1], Богдан Огерук[2], Назарій Андрущак[3]

[1] Кафедра систем автоматизованого проектування, Національний університет "Львівська політехніка", вул. С. Бандери, Львів, Україна, E-mail: oleksii.v.veretiuk@lpnu.ua, ORCID 0009-0009-2340-4458
[2] Кафедра систем автоматизованого проектування, Національний університет "Львівська політехніка", вул. С. Бандери, Львів, Україна, E-mail: bohdan.r.oheruk@lpnu.ua, ORCID 0009-0000-6519-6225
[3] Кафедра систем автоматизованого проектування, Національний університет "Львівська політехніка", вул. С. Бандери, Львів, Україна, E-mail: nazariy.a.andrushchak@lpnu.ua, ORCID 0000-0002-8248-404X

ВИКОРИСТАННЯ ШУМІВ ДЛЯ СТАБІЛЬНОГО ГЕНЕРУВАННЯ ЗОБРАЖЕНЬ У ДИФУЗІЙНИХ МОДЕЛЯХ

Отримано: Лютий 18, 2025 / Переглянуто: Лютий 26, 2025 / Прийнято: Березень 14, 2025

https://doi.org/

Анотація. В основі цієї роботи лежить дослідження процесу генерування зображень за допомогою дифузійних моделей. Продемонстровано можливості покращення стабільності генерування необхідного зображення за допомогою підходу, що передбачає використання постійного шуму як регуляризатора та накладання на нього маски випадкового шуму, яка створюється навколо кожного пікселя вхідного зображення у випадковому радіусі. Реалізація моделей проведена за допомогою мови Python та бібліотек tensorflow, numpy та keras. Показано кінцеві результати, в яких вийшло досягти стабільності в генеруванні при збереженні варіативності.

Ключові слова: дифузійні моделі, генерування зображень, стабільність генерування.

Вступ

Поява генеративного штучного інтелекту відкриває нові можливості в оптимізації та автоматизації процесів, де використання алгоритмічного підходу є неможливим або має обмеження. Вже сьогодні генеративний штучний інтелект може створювати текст, зображення, 3D моделі та анімацію й музику.

Генеративні моделі, такі як GPT [1], DALL·E [2] та Stable Diffusion [3], демонструють високий рівень творчих можливостей, створюючи контент, що раніше потребував значних людських зусиль. Наприклад, художники та дизайнери можуть використовувати штучний інтелект для швидкого створення ескізів, концептів та стилістичних експериментів. У галузі медіа та маркетингу штучний інтелект допомагає генерувати рекламні матеріали, пости для соціальних мереж та відеоконтент.

Проте, незважаючи на значні досягнення в цьому напрямку з технічної сторони, застосування цих технологій у багатьох практичних задачах стикається з проблемами та обмеженнями пов'язаними зі специфікою роботи штучного інтелекту.

Огляд сучасних джерел інформації за тематикою публікації

Одним з найбільш перспективних напрямків штучного інтелекту є технології генерування зображень та іншого графічного контенту [4]. Такі технології здатні докорінно змінити індустрію контенту, автоматизуючи та спрощуючи процес малювання. Впровадження генеративного штучного інтелекту дозволить підвищити ефективність створення контенту й дасть можливість митцям приділяти більше часу сюжету свого творіння, а не роботі над відтворенням матеріалу [5]. Більше того, це дозволить людям, які не мають навиків малювання спробувати себе в цій індустрії.

На сьогодні більшість проєктів та досліджень зосереджені на генеруванні високо деталізованих зображень, а також генеруванні картинок з текстового запиту. Потенціал генеративного штучного інтелекту має великий вплив на контент індустрію особливо ту, яка стосується процесу малювання. Проте поточний підхід використання та роботи генеративного штучного інтелекту не дуже підходить для автоматизації. Сьогодні генеративний штучний інтелект зосереджений на генеруванні кінцевого результату за допомогою текстового запиту або поєднання вхідної картинки та текстового запиту. Тому користувач має дуже обмежений вплив на результат. В той час як процес малювання йде від створення окремих деталей й подальшого об'єднання цих деталей в кінцевий результат. Такий процес вимагає більшого контролю над процесом генерації. Отже для того, щоб вдало застосовувати генеративний штучний інтелект для автоматизації процесу малювання необхідно забезпечити стабільність генерування відповідно до вхідного ескізу разом зі збереженням певного рівня варіативності у вихідному результаті [6].

Виклад основного матеріалу

Генеративний ШІ (він же Generative Artificial Intelligence, GenAI) галузь штучного інтелекту, що швидко розвивається, основна ідея – це створити те, чого раніше не існувало. Насамперед це нові форми контенту: текстовий, аудіо та візуальний. Генеративні моделі використовують як основу для навчання набори даних, проте не просто комбінують їх відповідно до запиту, а створюють фактично з нуля [7]. В цьому полягає головна відмінність від дискримінаційного ШІ, який аналізує різницю між різними типами даних.

Наведемо простий приклад. Якщо поставити дискримінаційному ШІ питання на кшталт «Людина йде дорогою чи їде автомобілем?», він з високою ймовірністю дасть точну та однозначну відповідь. Якщо ж генеративному ШІ запропонувати намалювати, як однією дорогою йде людина, а її обганяє автомобіль – результатом буде саме малюнок, а не текст.

Таким чином, генеративний ШІ створює новий контент на основі того, що він дізнався з раніше створеного кимось контенту, причому відбувається це в процесі безперервного навчання ШІ.

Виділяють кілька моделей роботи GenAI, в основу яких покладені такі перетворення:

- з тексту до тексту;
- з тексту у 2D зображення;
- з тексту у 3D зображення/відео;
- з тексту на дію (відповідь питання, пошук інформації, аналіз даних).

Генеративний штучний інтелект має такі можливості: вести розмови так, ніби по той бік екрана знаходиться така сама людина, як ви, писати програмний код, створювати з нуля зображення та відео за описами. ChatGPT – це також приклад генеративного ШІ, причому в базовому варіанті доступного широкому колу користувачів.

Принцип роботи генеративного ШІ добре видно на рис. 1.

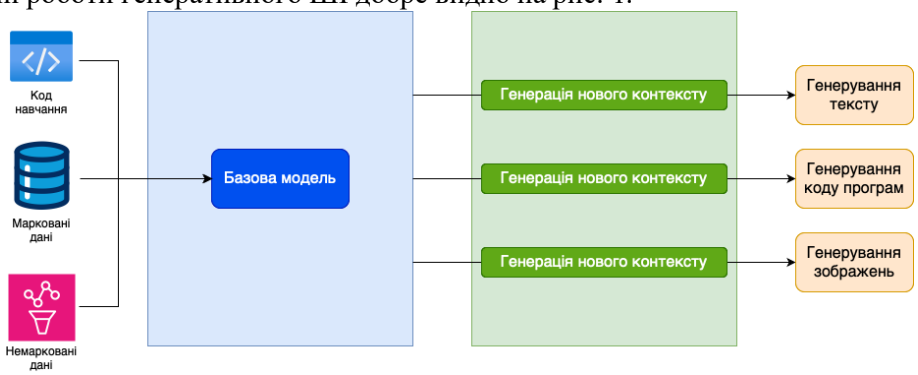


Рис. 1. Схематична діаграма принципу роботи генеративного ШІ.

Про генеративне ШІ активно говорять останні кілька років, хоча цю технологію не можна назвати новою – під її опис підходить тест Алана Тюрінга, запропонований ним ще в 1950 році. Тоді він стверджував, що машину можна назвати розумною, якщо вона почне генерувати відповіді на питання, які нічим не відрізняються від людських. Генеративні моделі почали розробляти ще у 1960-х та 1970-х, проте складніші моделі, зокрема на основі глибокого навчання, почали активно розвиватися лише у 1990-х. Саме вони змогли генерувати досить реалістичний, схожий на людський текст і навіть відтворювати мову. Черговий спалах популярності генеративного ШІ припав на появу GPT-3, створеної компанією OpenAI (ChatGPT – це її високотехнологічний продукт, який використовує саме цю мовну модель).

Якщо розглядати галузі, де вже застосовуються технології ШІ, то можна назвати величезну кількість областей та найнеймовірніших проєктів – від навчання та медицини до обробки великих даних та аналітики. Якщо звузити сегмент до генеративного ШІ, то створення унікального творчого контенту теж буде лише одним із ймовірних сценаріїв його застосування.

В яких завданнях потрібний і вже використовується GenAI:

- покращення якості цифрових зображень та відео;
- створення персоналізованого контенту;
- прототипування у виробничих цілях;
- генерація програмного коду;
- створення чат-ботів, віртуальних помічників;
- виконання візуальних перевірок та контролю якості.

У перспективі генеративний ШІ знайде собі застосування практично в кожній галузі, але вже сьогодні можна виділити декілька пріоритетних напрямків, в яких його використання дає максимальний ефект і користь, а саме генеративно-змагальні мережі та дифузійні моделі.

Генеративно-змагальні мережі (GAN) – це клас алгоритмів машинного навчання, де дві нейронні мережі (генератор та дискримінатор) змагаються одна проти одної. Генератор намагається створити дані, а дискримінатор намагається визначити, чи є ці дані справжніми або синтезованими [8]. Основний принцип роботи GAN полягає в грі між двома мережами як можна бачити на рис. 2. генератор – створює зображення (або інший тип даних) з випадкового шуму, а дискримінатор – отримує зображення як з реального набору даних, так і з генератора, та намагається визначити, яке з них справжнє. Після цього вираховуються втрати для оптимізації мережі під час навчання. Відповідно, метою є максимізація генерування реалістичних даних генератором і правильне розрізнення справжніх і згенерованих даних дискримінатором.

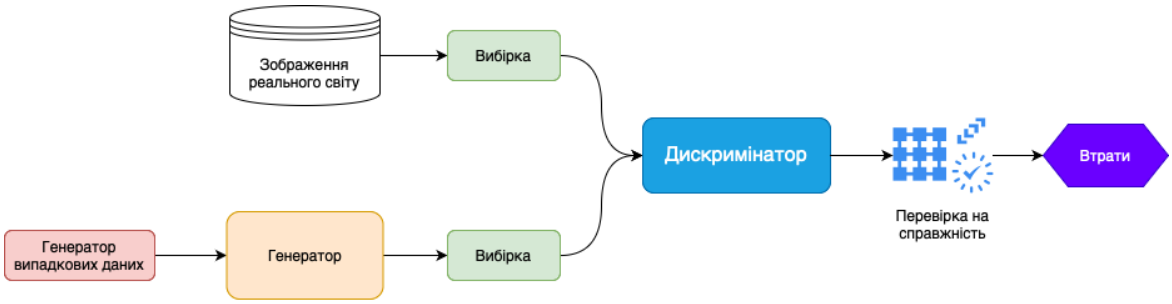


Рис. 2. Схематична діаграма принципу роботи генеративно-змагальної мережі.

Процес навчання GAN полягає в постійному покращенні обох мереж: кожен раз, коли дискримінатор визначає фальшиве зображення, генератор вчиться з помилки, намагаючись створити більш переконливе зображення наступного разу.

Дифузійні моделі (Diffusion Models) – це тип генеративних моделей, що навчаються створювати нові дані через процес, зворотний дифузії (поширення шуму). Вони були розроблені як альтернатива генеративно-змагальним мережам (GAN), і їх основна ідея полягає у використанні поступової "псевдодифузії" [9], яка перетворює реальні дані в шум, а потім через зворотний процес

відновлює їх. На рис. 3 ілюструється принцип роботи стохастичних диференціальних рівнянь (SDE) у процесах дифузії даних, які застосовуються у генеративних моделях (наприклад, дифузійних моделях для генерації зображень).

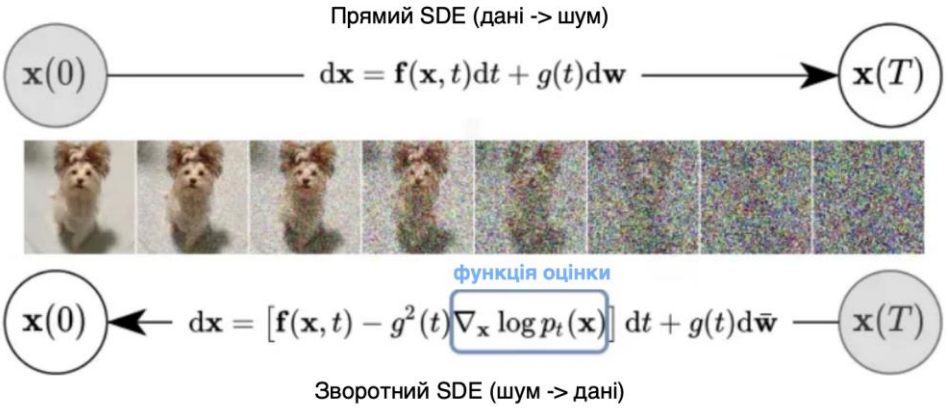


Рис. 3. Принцип роботи стохастичних диференціальних рівнянь (SDE) у процесах дифузії даних.

Верхня частина: Прямий SDE (дані → шум) на ній показано процес поступового додавання шуму до зображення собаки. Від початкового чистого зображення (x(0)) до повністю зашумленого (x(T)).

$$dx = f(x,t)dt + g(t)dw, \tag{1}$$

де *x* – вектор стану, що еволюціонує у часі й описує зображення, *t* – часова змінна, *f(x,t)* – дрейфова функція, *g(t)* – коефіцієнт дифузії, *dw* – приріст білого шуму (Вінерівський процес)

Також у нижній частині рис. 3. показаний Зворотний SDE (шум → дані). Це процес відновлення зображення із шуму назад до вихідних даних. Використовується функція оцінки градієнта лог-правдоподібності ($\nabla_x \log p_t(x)$), яка допомагає моделі "розуміти", як прибрати шум.

$$dx = [f(x,t) - g^2(t)\nabla_x \log p_t(x)]dt + g(t)d\bar{w}\,, \tag{2}$$

де *x* – вектор стану, що еволюціонує у часі й описує зображення, *t* – часова змінна, *f(x,t)* – дрейфова функція, *g(t)* – коефіцієнт дифузії, $\nabla_x \log p_t(x)$ – оцінка градієнта лог-правдоподібності, $d\bar{w}$ – приріст шуму у зворотньому напрямку

Ключова ідея знаходиться у прямому процесі, в результаті якого додається шум до зображення та зворотному процесі, який використовує навчену модель оцінки градієнта, прибирає шум, відтворюючи оригінальні дані або навіть створює нові синтетичні зображення. Такий принцип лежить в основі сучасних генеративних моделей як Stable diffusion або DALL·E 2.

Дифузійні моделі стали потужною альтернативою у сфері генеративного ШІ, продемонструвавши неабиякий успіх у різних галузях, таких як комп'ютерний зір, аудіосинтез та навчання з підкріпленням. На відміну від традиційних методів, таких як генеративні змагальні мережі (GAN) та варіаційні автокодери (VAE), яким притаманні обмеження щодо точності та ефективності обробки даних високої розмірності, дифузійні моделі пропонують надійні та адаптивні рішення для генерації нових зразків, пристосованих до конкретних завдань.

Хоча дифузійні моделі є перспективними, їхня теоретична база є відносно обмеженою, що створює труднощі для подальшого методологічного розвитку. Однак останні дослідження спрямовані на подолання цих обмежень і підвищення продуктивності дифузійних моделей, зокрема, шляхом інтеграції умовних параметрів, які уможливлюють точну генерацію зразків.

Значний внесок у розвиток можливостей дифузійних моделей зробило дослідження, проведене вченими з Принстонського університету та Каліфорнійського університету в Берклі [10]. Методологія дослідження полягає у застосуванні складних моделей умовної дифузії, які використовують керуючі сигнали, щоб спрямувати генерацію зразків даних до бажаних

властивостей. Такий підхід не тільки підвищує точність, але й забезпечує ефективність обробки даних високої розмірності.

Порівняння генеративно-змагальних мереж та дифузійних моделей.

Дифузійні моделі та генеративно-змагальні мережі (GAN) є популярними підходами для генерації зображень, але вони працюють за різними принципами. GAN використовують змагальний процес між двома мережами: генератором і дискримінатором. Генератор створює зображення, а дискримінатор намагається відрізнити реальні зображення від згенерованих. Завдяки цьому генератор навчається створювати все реалістичніші зображення, але GAN часто страждають від проблем з нестабільністю навчання, таких як колапс моделі, коли генератор починає створювати дуже схожі зображення які не різняться між собою в залежності від вхідних даних.

Дифузійні моделі, на відміну від GAN, використовують концепцію поступового додавання шуму до зображень на етапі навчання. Після навчання модель тренують відновлювати ці зображення шляхом зворотного процесу зменшення шуму. Такий підхід дозволяє дифузійним моделям уникнути проблеми колапсу моделі, оскільки процес зворотного дифузії дає більше контролю над генерацією, дозволяючи створювати більш різноманітні й деталізовані зображення. Хоча дифузійні моделі зазвичай потребують більше часу для генерації зображень, їх здатність зберігати високу якість на всіх етапах генерації [11] є значною перевагою. У порівнянні з GAN, дифузійні моделі демонструють більшу стабільність у навчанні, що робить їх перспективними для більш складних задач, таких як генерація фотореалістичних зображень. Однак, генерувати зображення за допомогою дифузійних моделей зазвичай займає більше часу і вимагає значних обчислювальних ресурсів [12].

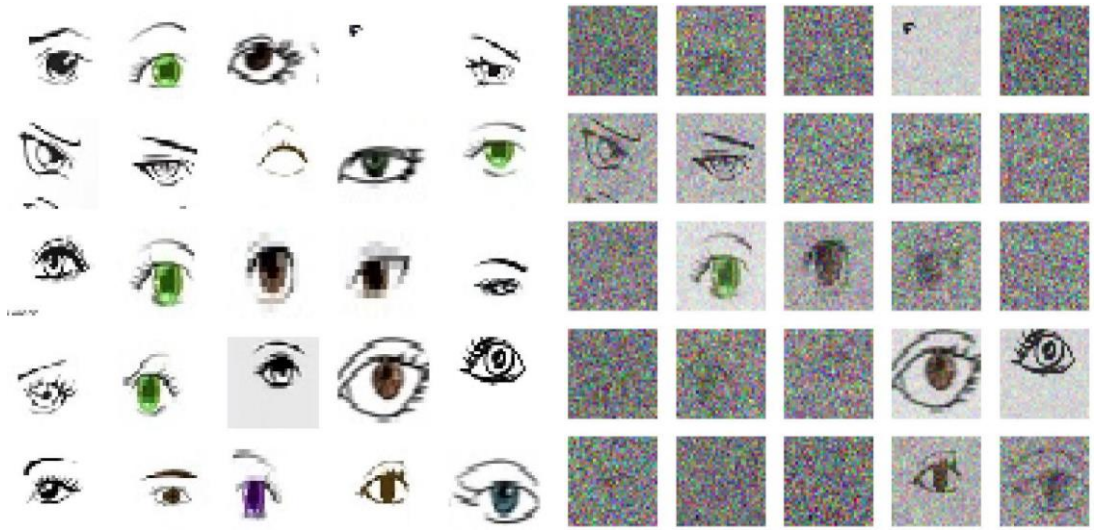
Технології: У контексті моделі дифузії, Keras і TensorFlow [13] є основними бібліотеками, що використовуються для розробки та реалізації моделей глибокого навчання. TensorFlow надає потужну основу для побудови нейронних мереж, включаючи числові обчислення та операції на тензорах, які є необхідними для виконання складних алгоритмів, таких як дифузійні моделі. Зокрема, TensorFlow забезпечує високу гнучкість та масштабованість, що дозволяє ефективно навчати моделі на великих наборах даних.

Keras, як високорівнева бібліотека, є частиною TensorFlow і надає зручний інтерфейс для побудови та тренування моделей. Вона значно спрощує розробку нейронних мереж завдяки абстракціям, які дозволяють зосередитись на архітектурі моделі, а не на складних технічних деталях. У випадку з дифузійними моделями, Keras використовується для побудови та навчання глибоких нейронних мереж, які реалізують зворотний процес дифузії — поступове відновлення зображень із шуму.

Обидві бібліотеки забезпечують інтеграцію з різноманітними алгоритмами оптимізації, такими як Adam чи SGD, для мінімізації функції втрат та вдосконалення параметрів моделі. Це дозволяє моделі дифузії ефективно навчатися на великому обсязі даних, покращуючи точність відновлення зображень із шуму. Крім того, TensorFlow і Keras підтримують використання графічних процесорів (GPU) для пришвидшення обчислень, що особливо важливо при роботі з великими моделями та великими наборами даних.

Завдяки простоті використання Keras і потужності TensorFlow, ці технології стають потужними інструментами для розробки та реалізації складних моделей, таких як дифузійні, де точність і ефективність грають важливу роль у досягненні високоякісних результатів.

Визначення експерименту. У межах дослідження було створено кілька варіантів дифузійних моделей для генерації зображень. Ці моделі були натреновані на невеликому наборі даних, що містить 581 зображення намальованих очей у стилі японських коміксів. Тренувальний набір включає малюнки з різною формою, розміром та рівнем деталізації, що дозволяло моделі навчитися обробляти варіативність вхідних даних. На рис. 4б показано приклад тренувальних даних.



а) б)

Рис. 4. Приклад 25 зображень очей, які були використані для тренування моделі:
а) без накладеного шуму, б) з накладеним шумом

Спочатку було побудовано модель дифузії з використанням випадкового шуму, який накладається на вхідне зображення з випадковим рівнем інтенсивності. Рівні інтенсивності розділені від 0 до 15, де 0 означає повне заміщення картинки шумом, а 15 найнижчу інтенсивність, крізь яку видно початкове зображення. На рис. 4б наведений приклад накладання шумів різної інтенсивності на зображення з тренувального набору даних.

В результаті тренування моделі, було отримано можливість генерування зображень з випадкових шумів. Для перевірки результатів генерування було створено вхідний набір даних, який складається з 25 зображень, які представляють собою випадковий набір шумів показаний на рис. 5. Після обробки вхідних даних було отримано набір згенерованих зображень (рис. 6). Можна побачити, що результат є випадковим й немає явного взаємозв'язку з вхідними даними. Перевагою цього підходу є висока варіативність вихідних зображень. Однак результат генерування є нестабільним.

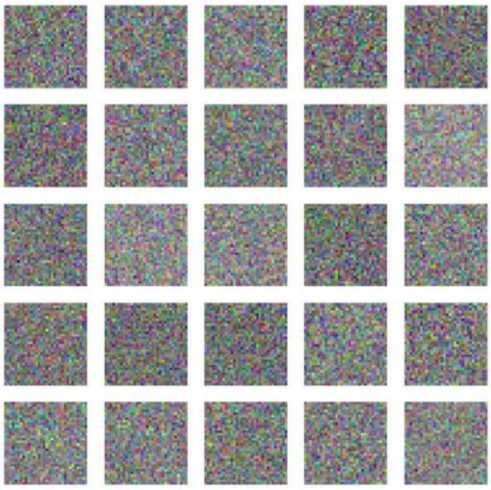


Рис. 5. Приклад випадкових шумів, які подаються на вхід моделі.

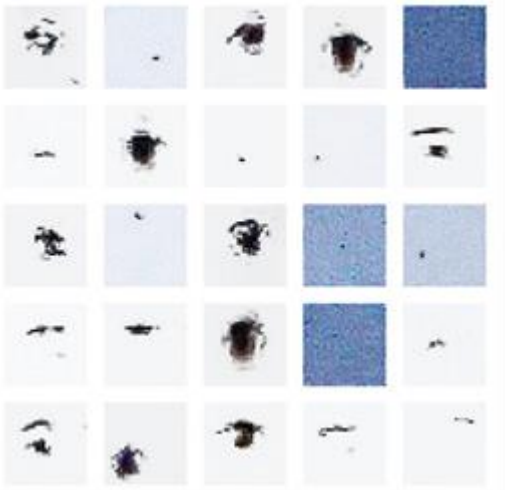


Рис. 6. Результат генерування з випадкових шумів.

Наступним етапом в експерименті була верифікація роботи моделі з вхідними даними, які не є повністю випадковими. Для цього було використано окремо створене зображення (рис. 7), яке є відсутнім в тренувальному наборі даних, в якості вхідних даних. Для цього на створене зображення

були накладені випадкові шуми з максимальною інтенсивністю. Було отримано набір вхідних даних схожий на рис. 5. Результат генерації зображень з новоствореного вхідного набору даних (рис. 8) є схожим на результати генерації показані на рис. 6. Можна стверджувати, що наявність зображення за шумами не впливає на кінцевий результат генерування.

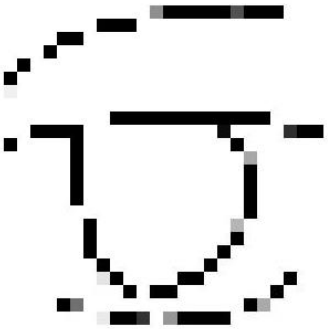


Рис. 7. Вхідне зображення.

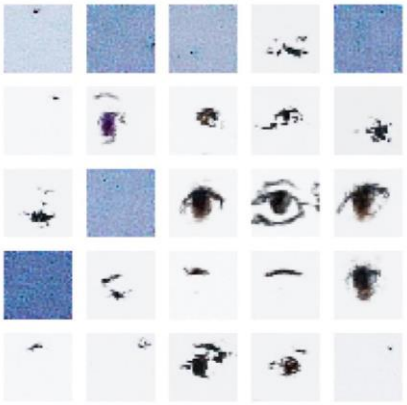


Рис. 8. Результат генерування з випадкових шумів, які закривають вхідне зображення.

Наступний експеримент передбачає перевірку роботи моделі за умов накладання шумів на зображення зі спадною інтенсивністю. Для цього було створено вхідний набір із 25 зображень, у якому кожне зображення містило накладені шуми з поступовим зменшенням їхньої інтенсивності (рис. 9). У результаті отримано набір зображень, які демонструють, що зі зменшенням інтенсивності шумів модель починає генерувати зображення, більш схожі на вхідне (рис. 10). Це забезпечує певну стабільність результату, проте водночас зникає можливість генерування додаткових деталей, що доповнюють зображення

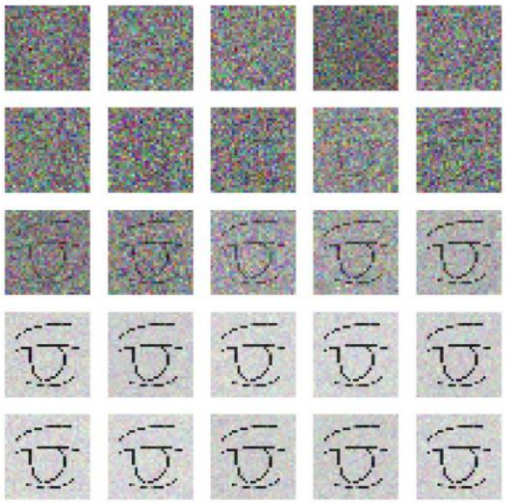


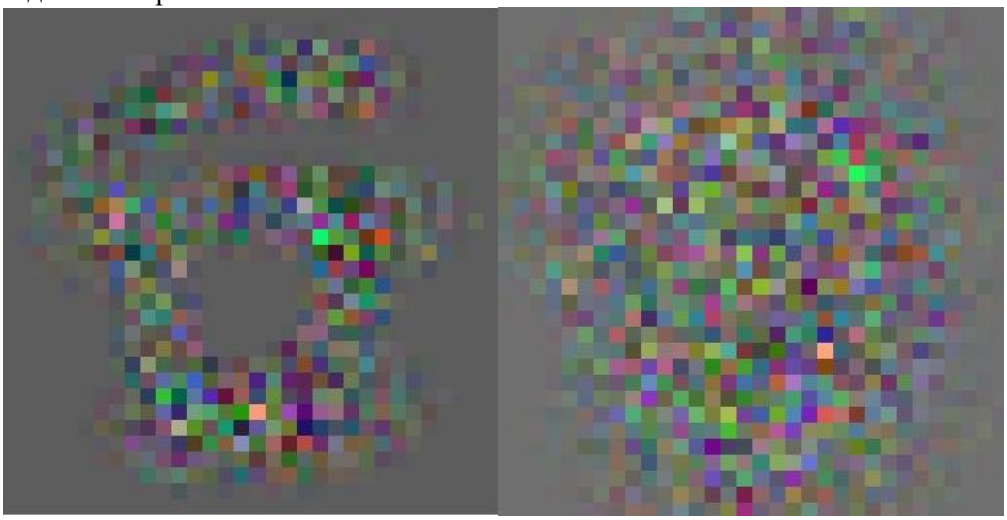
Рис. 9. Вхідне зображення з накладеними шумами з інтенсивністю, яка зменшується.

Рис. 10. Результат генерування з вхідного зображення з накладеними шумами з інтенсивністю, яка зменшується.

Для того, щоб досягти стабільного генерування зображення, нами було вперше запропоновано підхід, який використовує постійний шум як регуляризатор та маску випадкового шуму, яка створюється навколо пікселів вхідного зображення у відповідному радіусі. Під час тренування моделі на відміну від попереднього підходу, замість звичайного накладання випадкових

шумів різної інтенсивності застосовувалися маски шумів із випадково обраним радіусом та варіативною інтенсивністю, тоді як інші етапи тренування залишалися незмінними.

На рис. 11 наведено приклади масок випадкового шуму з різним радіусом, створених навколо зображення, поданого на рис. 8.



а) б)

Рис. 11. Маски випадкового шуму згенеровані навколо зображення:
а) з радіусом 4, б) з радіусом 7.

На рис. 12 показано, як виглядають вхідні дані після накладання маски шумів на постійний шум, що використовується як регуляризатор. Спочатку формується фон із рівномірно розподіленим постійним шумом, після чого на нього накладається маска випадкового шуму з обраним радіусом, сформована на основі вхідного зображення. Варто зазначити, що в цьому прикладі інтенсивність постійного шуму була зменшена для кращої візуалізації меж маски.

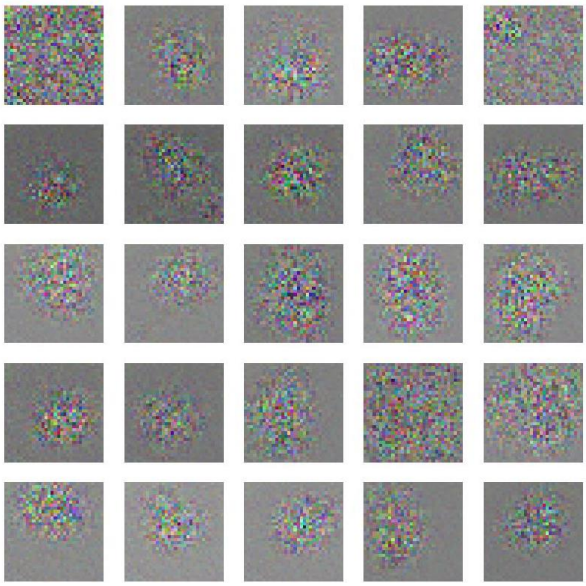


Рис. 12. Маска випадкового шуму накладена на постійний шум.

Для аналізу впливу зміни радіусу створеної маски зображення на фінальний результат було сформовано вхідний набір із 10 зображень, де у першому зображенні радіус маски становить 2, а в останньому – 11. Це дозволяє відстежити зміни в якості генерованих зображень після обробки цих даних моделлю.

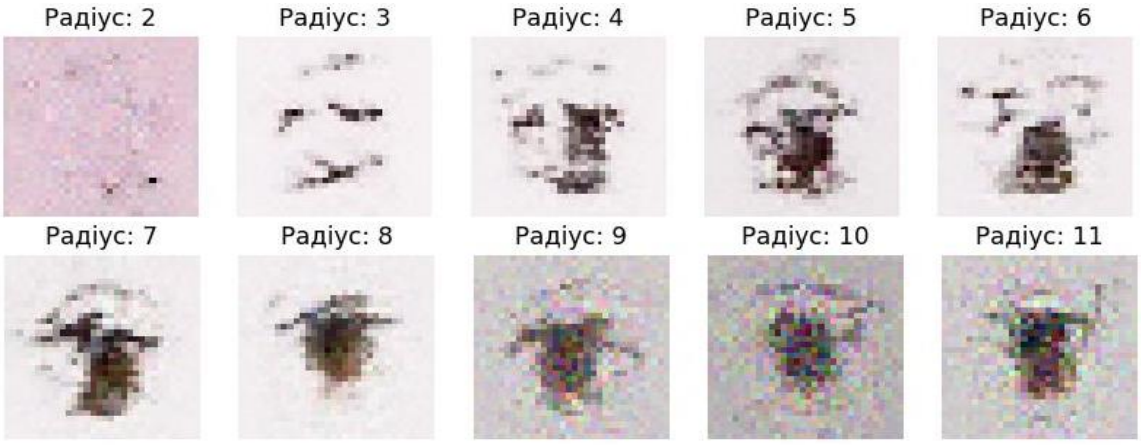


Рис. 13. Генерації зображення з різними радіусом.

Аналізуючи рис. 13, можна побачити, що зображення загалом зберігають загальні розміри та стиль, проте довільність радіусу має значний вплив на кінцевий результат. Якщо радіус маски малий, то результат буде менш деталізованим. В якості експерименту було вирішено визначити постійний радіус 7 і запустити процес генерування. Результат можна побачити на рис. 14. Збільшений та фіксований радіус дозволяє забезпечити більш стабільне генерування кінцевого результату.

Для верифікації запропонованого методу було проведено тестування на іншому зображенні, поданому на рис. 15. Це дозволяє оцінити стабільність та узагальнюючу здатність моделі під час генерування, а також визначити, чи зберігається якість відтворення основних деталей при зміні вхідних даних.

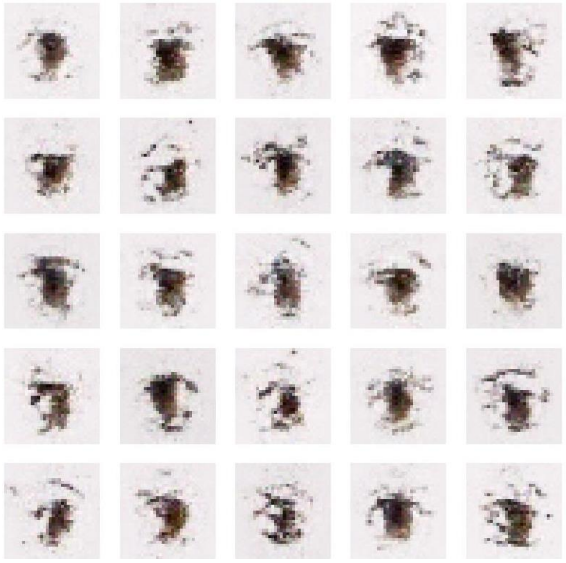


Рис. 14. Генерація зображень з маски випадкового шуму з постійним радіусом 7.

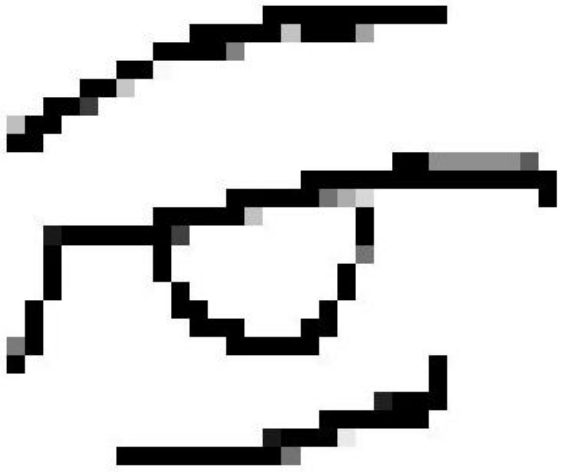


Рис. 15. Зображення для тестування

Аналіз результатів генерації (рис. 16) показує, що всі згенеровані зображення містять спільні елементи та зберігають схожість між собою. Крім того, вони виявляють візуальну відповідність тестовому зображенню, що підтверджує стабільність моделі та її здатність відтворювати ключові риси вхідних даних.

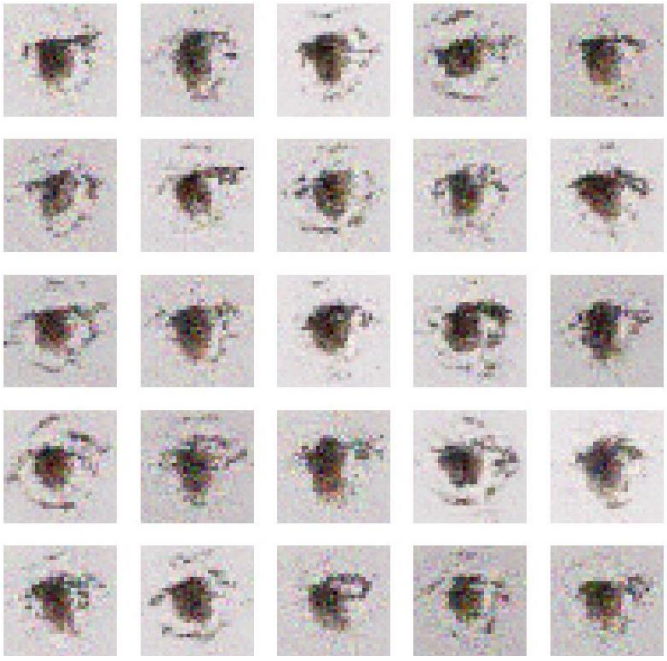


Рис. 16. Результат генерування

Висновки

У ході проведеного дослідження нами вперше було запропоновано підхід для стабільного генерування зображень у дифузійних моделях, що базується на використанні постійного шуму як регуляризатора та маски випадкового шуму, сформованої навколо кожного пікселя вхідного зображення у випадковому радіусі. Реалізація та експериментальні дослідження були проведені з використанням мов програмування Python та бібліотек TensorFlow, NumPy та Keras. У процесі експериментів проводились дослідження впливу накладання шумів на вхідні зображення та їхній вплив на кінцевий результат.

Аналіз результатів показав, що радіус маски шуму має суттєвий вплив на якість генерованих зображень. Зокрема, при використанні масок з малим радіусом спостерігалося зниження деталізації, тоді як застосування масок з постійним радіусом (наприклад, 7 пікселів) дозволяє отримати стабільні та деталізовані результати. Запропонований метод забезпечує баланс між стабільністю та варіативністю: модель зберігала ключові риси та стиль вхідного зображення, водночас генеруючи різноманітні варіації.

Ефективність підходу підтверджена експериментами на прикладі генерації зображень очей. Запропонована нами модель демонструвала стабільність у відтворенні форми та стилю, зберігаючи можливість варіювати деталі та отримувати різноманітні зразки. Отримані результати мають практичну цінність, оскільки можуть бути застосовані для автоматизації процесу створення ескізів та концептуальних ілюстрацій, особливо у сферах мистецтва, дизайну та анімації.

Таким чином, запропонований метод стабілізації генерації у дифузійних моделях демонструє перспективність для подальшого розвитку технологій генеративного штучного інтелекту та має потенціал для широкого застосування у різних галузях цифрової творчості.

Перелік використаних джерел

[1]. Zhang, M., & Li, J. (2021). A commentary of GPT-3 in MIT Technology Review 2021. Fundamental Research, 1(6), 831-833. https://doi.org/10.1016/j.fmre.2021.11.011

[2]. Puteikis, K., & Mameniškienė, R. (2024). Artificial intelligence: Can it help us better grasp the idea of epilepsy? An exploratory dialogue with ChatGPT and DALL· E 2. Epilepsy & Behavior, 156, 109822. https://doi.org/10.1016/j.yebeh.2024.109822

[3]. Borji, A. (2022). Generated faces in the wild: Quantitative comparison of stable diffusion, midjourney and dall-e 2. arXiv preprint arXiv:2210.00586.

[4]. Sun, Z., Fang, H., Cao, J., Zhao, X., & Wang, D. (2024, October). Rethinking Image Editing Detection in the Era of Generative AI Revolution. In Proceedings of the 32nd ACM International Conference on Multimedia (pp. 3538-3547). https://doi.org/10.1145/3664647.3681445

[5]. Archana Balkrishna, Y. (2024). An analysis on the use of image design with generative AI technologies. International Journal of Trend in Scientific Research and Development, 8(1), 596-599.

[6]. Van Daele, D., Decleyre, N., Dubois, H., & Meert, W. (2021, May). An automated engineering assistant: Learning parsers for technical drawings. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 35, No. 17, pp. 15195-15203). https://doi.org/10.1609/aaai.v35i17.17783

[7]. Banh, L., & Strobel, G. (2023). Generative artificial intelligence. Electronic Markets, 33(1), 63. https://doi.org/10.1007/s12525-023-00680-1

[8]. Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative adversarial networks: An overview. IEEE signal processing magazine, 35(1), 53-65. https://doi.org/10.1109/MSP.2017.2765202

[9]. Yang, L., Zhang, Z., Song, Y., Hong, S., Xu, R., Zhao, Y., ... & Yang, M. H. (2023). Diffusion models: A comprehensive survey of methods and applications. ACM Computing Surveys, 56(4), 1-39.

[10]. Batzolis, G., Stanczuk, J., Schönlieb, C. B., & Etmann, C. (2021). Conditional image generation with score-based diffusion models. arXiv preprint arXiv:2111.13606.

[11]. Wang, Z., Zheng, H., He, P., Chen, W., & Zhou, M. (2022). Diffusion-gan: Training gans with diffusion. arXiv preprint arXiv:2206.02262.

[12]. Dhariwal, P., & Nichol, A. (2021). Diffusion models beat gans on image synthesis. Advances in neural information processing systems, 34, 8780-8794.

[13]. Géron, A. (2022). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow. " O'Reilly Media, Inc.".

Oleksii Veretiuk[1], Bohdan Oheruk[2], Nazariy Andrushchak[3]

[1] Department of Computer Design Systems, Lviv Polytechnic National University, S. Bandery street 12, Lviv, Ukraine, E-mail: oleksii.v.veretiuk@lpnu.ua, ORCID 0009-0009-2340-4458

[2] Department of Computer Design Systems, Lviv Polytechnic National University, S. Bandery street 12, Lviv, Ukraine, E-mail: bohdan.r.oheruk@lpnu.ua, ORCID 0009-0000-6519-6225

[3] Department of Computer Design Systems, Lviv Polytechnic National University, S. Bandery street 12, Lviv, Ukraine, E-mail: nazariy.a.andrushchak@lpnu.ua, ORCID 0000-0002-8248-404X

USING NOISES FOR STABLE IMAGE GENERATION IN DIFFUSION MODELS

Received: February 18, 2025 / Revised: February 26, 2025 / Accepted: March 14 2025

Abstract. The foundation of this work lies in the study of the image generation process using diffusion models. The potential for increasing generation stability is demonstrated through an approach that involves using constant noise as a regularizer, with the addition of a mask of random noise constructed around each pixel of the input image within a random radius. The models were implemented using Python and TensorFlow, NumPy, and Keras libraries. The final results are presented, showing that stability was achieved while maintaining a certain level of variability in the generated output.

Keywords: diffusion models, image generation, generation stability.