

ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ
ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
КАФЕДРА ІНЖЕНЕРІЇ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

КВАЛІФІКАЦІЙНА РОБОТА

на тему: «Оптимізація процесу розробки мобільної гри на
основі зображень, створених за допомогою штучного
інтелекту»

на здобуття освітнього ступеня магістра
зі спеціальності 121 Інженерія програмного забезпечення
(код, найменування спеціальності)
освітньо-професійної програми «Інженерія програмного забезпечення»
(назва)

*Кваліфікаційна робота містить результати власних досліджень. Використання
ідей, результатів і текстів інших авторів мають посилання
на відповідне джерело*

(підпис) Ярослав ЄВТАЄВ

Виконав: здобувач вищої освіти групи ПДМ-64
Ярослав ЄВТАЄВ

Керівник: Олесь ДІБРІВНИЙ
доктор філософії (PhD)

Рецензент: _____
*науковий ступінь,
вчене звання* Ім'я, ПРИЗВИЩЕ

Київ 2024

**ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ**
Навчально-науковий інститут інформаційних технологій

Кафедра Інженерії програмного забезпечення

Ступінь вищої освіти Магістр

Спеціальність 121 Інженерія програмного забезпечення

Освітньо-професійна програма «Інженерія програмного забезпечення»

ЗАТВЕРДЖУЮ

Завідувач кафедри

Інженерії програмного забезпечення

_____ Ірина ЗАМРІЙ

« _____ » _____ 2023 р.

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ**

_____ Євтаєву Ярославу Олександровичу _____

1. Тема кваліфікаційної роботи: «Оптимізація процесу розробки мобільної гри на основі зображень, створених за допомогою штучного інтелекту»

керівник кваліфікаційної роботи Олесь ДІБРІВНИЙ доктор філософії (PhD),

затверджені наказом Державного університету інформаційно-комунікаційних технологій від «19» жовтня 2023 р. №145.

2. Строк подання кваліфікаційної роботи «29» грудня 2023 р.

3. Вихідні дані до кваліфікаційної роботи: науково-технічна література, методи генерації зображень, сервіси генерації зображень.

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити)

1. Дослідження розробки мобільних ігор.

2. Аналіз технологій генерації зображень.

3. Розробка методу оптимізації.

4. Розробка програми оптимізації.

5. Перелік графічного матеріалу: *презентація*

1. Аналіз систем для генерації зображень.
2. Математичка модель.
3. Діаграма послідовності.
4. Лінійні діаграми порівняння методів.
5. Діаграма класів.
6. Інтерфейс програми.
7. Експертні оцінки.
8. Числові показники ефективності.

6. Дата видачі завдання «19» жовтня 2023 р.

КАЛЕНДАРНИЙ ПЛАН

| № з/п | Назва етапів кваліфікаційної роботи | Строк виконання етапів роботи | Примітка |
|-------|--------------------------------------------------------------------------------------------------------------------------|-------------------------------|----------|
| 1 | Аналіз наявної науково-технічної літератури | 19.10-05.11.23 | |
| 2 | Дослідження розробки мобільних ігор | 06.11-12.11.23 | |
| 3 | Аналіз існуючих методів та систем генерації зображень | 13.11-19.11.23 | |
| 4 | Розробка методу для зміни існуючих зображень за певним стилем та розміром. | 20.11-26.11.23 | |
| 5 | Розробка програми яка використовує методи генерації зображень для зміни існуючих зображень за певним стилем та розміром. | 27.11-03.12.23 | |
| 6 | Аналіз та тестування результатів виконаної роботи. | 04.12-10.12.23 | |
| 7 | Оформлення роботи: вступ, висновки, реферат | 11.12-20.12.23 | |
| 8 | Розробка демонстраційних матеріалів | 21.12-29.12.23 | |

Здобувач вищої освіти

_____ (підпис)

Ярослав ЄВТАЄВ

Керівник
кваліфікаційної роботи

_____ (підпис)

Олесь ДІБРІВНИЙ

РЕФЕРАТ

Текстова частина кваліфікаційної роботи на здобуття освітнього ступеня магістра: 84 стор., 2 табл., 30 рис., 30 джерел.

Мета роботи – зменшення витрат на розробку мобільної гри, за рахунок автоматизованої генерації візуального контенту.

Об'єкт дослідження – процес розробки мобільної гри, зосереджений на використанні штучного інтелекту для генерації зображень.

Предмет дослідження – оптимізація процесу створення графічного контенту мобільної гри.

Короткий зміст роботи: У роботі розглядається використання штучного інтелекту в процесі розробки мобільних ігор. Здійснено аналіз існуючих методів генерації зображень та їхню ефективність при створенні графічних елементів в іграх. Розроблено метод та програму для зменшення витрат на розробку, за рахунок автоматизованої генерації візуального контенту. Розроблено мобільну гру з використанням зображень, створених за допомогою штучного інтелекту. Проведено порівняльний аналіз швидкості та якості розробки гри з використанням штучного інтелекту порівняно з традиційними методами.

КЛЮЧОВІ СЛОВА: ШТУЧНИЙ ІНТЕЛЕКТ, МОБІЛЬНІ ІГРИ, ГРАФІЧНІ ЕЛЕМЕНТИ, СЕРВІСИ, DALL-E 3, STABLE DIFFUSION

ABSTRACT

Text part of the master's qualification work: 84 pages, 30 pictures, 2 table, 30 sources.

The purpose of the work is to reduce the costs of developing a mobile game by automating the generation of visual content.

Object of research – the process of developing a mobile game, focused on the use of artificial intelligence for image generation

Subject of research – optimization of the process of creating graphic content for a mobile game.

Summary of the work: The work examines the use of artificial intelligence in the process of developing mobile games. An analysis of existing methods of image generation and their effectiveness in creating graphic elements in games is conducted. A method and a program have been developed to reduce development costs through automated generation of visual content. A mobile game was developed using images created with the help of artificial intelligence. Central to the study is the exploration and utilization of the Dalle 3 service for AI-driven image creation. The research presents a novel program designed to transform and adapt multiple images into user-defined styles, aligning them with the specific requirements of various mobile gaming platforms. A significant application of this research is demonstrated in the successful development of a mobile game within the Trading Card Game (TCG) genre. This practical implementation not only validates the effectiveness of the proposed AI integration strategy but also highlights the potential of artificial intelligence to revolutionize the mobile gaming industry. The study underscores the transformative capabilities of AI in streamlining game development processes, enhancing creative output, and offering scalable solutions that can adapt to diverse gaming platforms. The findings of this research provide valuable insights into the future of mobile game development, emphasizing the role of AI in driving innovation, efficiency, and diversity in game design.

KEYWORDS: ARTIFICIAL INTELLIGENCE, MOBILE GAMES, GRAPHIC ELEMENTS, SERVICES, DALL-E 3, STABLE DIFFUSION

ЗМІСТ

| | |
|-------------------------------------------------------------------------------|----|
| ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ..... | 9 |
| ВСТУП..... | 10 |
| РОЗДІЛ 1 ОСНОВИ ОПТИМІЗАЦІЇ ПРОЦЕСУ РОЗРОБКИ МОБІЛЬНОЇ ГРИ..... | 14 |
| 1.1. Огляд сучасного стану розробки мобільних ігор..... | 14 |
| 1.2. Критичний огляд підходів до оптимізації розробки | 17 |
| 1.3. Штучний інтелект у розробці ігор | 19 |
| 1.4. Етапи процесу розробки мобільної гри..... | 21 |
| РОЗДІЛ 2 АНАЛІЗ ТЕХНОЛОГІЙ ТА СИСТЕМ ДЛЯ ГЕНЕРАЦІЇ ВІЗУАЛЬНОГО КОНТЕНТУ | 25 |
| 2.1. Аналіз існуючих технологій генерації візуального контенту..... | 25 |
| 2.2. Аналіз систем для генерації зображень..... | 40 |
| 2.3. Аналіз та розробка шкали оцінювання зображень..... | 47 |
| 2.4. Експертна оцінка критеріїв якості зображень | 49 |
| 2.5. Математична модель оптимізаційного рішення..... | 53 |
| РОЗДІЛ 3 РОЗРОБКА ТА ВПРОВАДЖЕННЯ ОПТИМІЗАЦІЙНОГО РІШЕННЯ..... | 57 |
| 3.1. Схема оптимізації розробки мобільної гри..... | 57 |
| 3.2. Опис використаних програмних засобів | 59 |
| 3.3. Опис структури проекту | 63 |
| 3.4. Опис розроблених класів | 67 |
| 3.5. Економічний аналіз зменшення витрат на розробку | 71 |
| 3.6. Практичне застосування розробленого методу..... | 78 |
| 3.7. Пропозиції щодо подальшого вдосконалення та застосування методу.. | 82 |
| ВИСНОВКИ..... | 84 |
| ПЕРЕЛІК ПОСИЛАНЬ | 85 |
| ДЕМОНСТРАЦІЙНІ МАТЕРІАЛИ (Презентація)..... | 89 |

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

| | | |
|------|---|--------------------------------------------------------------------------------|
| ШІ | - | Штучний Інтелект (Artificial Intelligence) |
| API | - | Графічний Інтерфейс (Graphical User Interface) |
| GUI | - | Graphical User Interface |
| SD | - | Стабільна Дифузія (Stable Diffusion) |
| UX | - | Досвід Користувача (User Experience) |
| UI | - | Інтерфейс Користувача (User Interface) |
| TCG | - | Колекційна Карткова Гра (Trading Card Game) |
| ML | - | Машинне Навчання (Machine Learning) |
| GA | - | Генетичний Алгоритм (Genetic Algorithm) |
| VR | - | Віртуальна Реальність (Virtual Reality) |
| AR | - | Доповнена Реальність (Augmented Reality) |
| GAN | - | Генеративно-змагальні Мережі (Generative Adversarial Networks) |
| VAE | - | Варіаційні Автокодувальники (Variational Autoencoders) |
| MIT | - | Массачусетський Технологічний Інститут (Massachusetts Institute of Technology) |
| GPT | - | Генеративний Передбачувальний Трансформер (Generative Pretrained Transformer) |
| iGPT | - | Генеративний Трансформер Зображень (Image Generative Pretrained Transformer) |

ВСТУП

У сучасному світі динамічно розвиваються технології штучного інтелекту, вони відіграють ключову роль у різноманітних сферах життя суспільства. Однією з таких сфер, де штучний інтелект демонструє свій величезний потенціал, є індустрія мобільних ігор. Кожного року ринок мобільних ігор розширяється, пропонуючи все більше якісного контенту. Така тенденція ставить перед розробниками нові виклики, зокрема, пов'язані з необхідністю зменшення витрат на розробку і одночасного підвищення якості графічного контенту. В цьому контексті оптимізація процесу розробки із застосуванням штучного інтелекту для генерації зображень набуває особливої актуальності.

Актуальність теми полягає у необхідності пошуку нових підходів до зниження витрат на розробку, при одночасному збереженні або навіть підвищенні якості графічних елементів в іграх. Це особливо важливо для України, де індустрія мобільних ігор знаходиться на стадії активного зростання і потребує інноваційних підходів розробки та покращення якості контенту для конкуренції на міжнародному рівні. У цьому контексті, використання штучного інтелекту для створення зображень в іграх відкриває нові можливості і дозволяє сильно прискорити розробку проектів, значно зменшити їх кінцеву вартість і навіть іноді покращити якість візуального контенту.

У роботі проводиться критичний аналіз існуючих методів розробки мобільних ігор, з особливим акцентом на роль штучного інтелекту в цьому процесі. Проведено порівняльний аналіз традиційних підходів до створення візуального контенту та інноваційних методів, що базуються на використанні штучного інтелекту. Це дозволить обґрунтувати доцільність та ефективність впровадження новітніх технологій у процес розробки мобільних ігор.

Про аналізувавши роботу [1] було визначено декілька ключових проблем:

1. Складність розробки ігор: багато дисциплінарний характер розробки ігор, який включає розробку дизайну гри.

2. Баланс творчості та тестування: зменшення часу на тестування проекту через нехватку часових ресурсів.
3. Висока вартість проекту
4. Високі вимоги до ресурсів

Ці проблеми можна вирішити прискоривши розробку ігор за допомогою використання ШІ для створення зображень. Інструменти штучного інтелекту можуть швидко створювати прототипи різних візуальних стилів, що полегшує тестування та вдосконалення. Таким чином, штучний інтелект у створенні зображень може оптимізувати процес розробки, особливо на етапі тестування, зменшивши складність, вартість і прискорить розробку проекту.

Мета дослідження: зменшення витрат на розробку мобільної гри, за рахунок методу автоматизованої генерації візуального контенту.

В процесі дослідження вирішувалися наступні завдання:

1. Аналіз сучасного стану ринку мобільних ігор: Вивчення останніх тенденцій, проблем та можливостей у галузі розробки мобільних ігор.
2. Дослідження штучного інтелекту в графічному дизайні: Огляд існуючих технологій ШІ для створення візуального контенту, їхні переваги та обмеження.
3. Розробка методу створення графічних зображень використовуючи ШІ: Адаптація алгоритмів генеративного штучного інтелекту для ефективного використання у розробці мобільних ігор.
4. Розробка мобільної гри для тестування методу: Реалізація розробленого методу в реальних умовах розробки ігор.
5. Оцінка розробленого методу та порівняння його з аналогами: Аналіз ефективності методу оптимізації в контексті зниження витрат та часу на розробку. Порівняння методу з звичайними методами розробки додатків.
6. Формулювання рекомендацій для вдосконалення процесу розробки мобільних ігор: Розробка практичних рекомендацій для подальшого

7. використання штучного інтелекту у розробці мобільних ігор.

Об'єктом дослідження є процес розробки мобільних ігор з використанням штучного інтелекту для генерації зображень. Цей процес включає планування, дизайн, програмування, тестування та випуск мобільної гри, з особливим акцентом на інтеграцію алгоритмів ШІ для автоматизації та оптимізації створення візуального контенту. Основні характеристики об'єкта охоплюють взаємодію між різними аспектами розробки, включаючи графічний дизайн, програмування, управління проектами та користувацький досвід.

Предметом дослідження у цій кваліфікаційній роботі є оптимізація процесу створення графічного контенту мобільних ігор. Цей предмет зосереджується на конкретних проблемах та викликах, які виникають при інтеграції ШІ-технологій у створення візуальних елементів ігор. Основна увага приділяється вивченню методів, які можуть підвищити ефективність та знизити витрати в процесі розробки, одночасно зберігаючи високу якість кінцевого продукту.

Методи дослідження було проведено детальний аналіз літератури для вивчення сучасних тенденцій у розробці мобільних ігор та впровадженні штучного інтелекту. Емпіричні методи були застосовані для аналізу реальних процесів розробки та впровадження ШІ-технологій. Комп'ютерне моделювання допомогло оцінити ефективність впроваджених рішень. Кількісні та якісні методи аналізу були використані для оцінки змін у витратах та продуктивності. Також були проведені експертні оцінки для додаткового збору даних та верифікації результатів.

Наукова новизна полягає в дослідженні невивчених раніше аспектів інтеграції ШІ у розробку мобільних ігор, особливо в контексті оптимізації графічного дизайну. Застосовані інноваційні методи та алгоритми дозволили розробити новітній підхід до створення візуального контенту. Практична значущість роботи полягає у можливості використання цих результатів у реальних проектах розробки мобільних ігор, що сприятиме підвищенню ефективності та зниженню витрат всіх ресурсів проекту.

Апробація результатів пройшла на наукових конференціях. На конференції "Проблеми комп'ютерної інженерії" у Києві, організованій ДУІКТ у 2023 році, була представлена доповідь "Штучний інтелект в мобільних іграх: порівняльний аналіз інструментів для створення зображень", яка пізніше була подана до публікації. Також, результати дослідження були виставлені на Всеукраїнській науково-практичній конференції "Telecommunication: Problems and Innovation", що відбулася в Києві в тому ж році, з доповіддю на тему "Оптимізація процесу розробки мобільної гри на основі зображень, створених за допомогою штучного інтелекту", яка також була подана до публікації.

Теоретична, методична та практична значущість отриманих результатів. Теоретично, робота розширює розуміння та використання штучного інтелекту на практичному прикладі, а саме на генерації графічних елементів для ігор. Методична значущість полягає в тому що, були розроблені нові підходи та алгоритми для інтеграції ШІ у графічний дизайн, а саме програма для автоматичної генерації зображень на основі інших зображень з використанням певного стилю. Практично, результати мають важливе значення для галузі мобільних ігор, оскільки пропонують ефективні рішення для зниження витрат та підвищення якості розробки.

1 ОСНОВИ ОПТИМІЗАЦІЇ ПРОЦЕСУ РОЗРОБКИ МОБІЛЬНОЇ ГРИ

1.1. Огляд сучасного стану розробки мобільних ігор

З огляду на динамічний світ технологій, сфера розробки мобільних ігор постійно еволюціонує, пропонуючи нові горизонти для креативності та інновацій. У цьому контексті, надзвичайно важливим є аналіз сучасних тенденцій та викликів, що стоять перед розробниками. Зокрема, ринок мобільних ігор останніми роками демонструє значний інтерес до використання штучного інтелекту, що відкриває нові можливості для автоматизації та підвищення якості графічного контенту.

Проаналізувавши опитування розробників 2023 року [2] рис. 1.1, можна зробити наступні висновки:

- Труднощі з Пошуком Талантів: Найбільшою проблемою, вказаною 32% респондентів, є залучення та утримання талантів, що відображає глобальну нестачу робочої сили.
- Зменшення Інтересу до AR/VR: Зниження розробки ігор для платформ AR та VR, незважаючи на попередні прогнози.
- Переваги Платформ: ПК залишається найпопулярнішою платформою для розробки ігор, залучаючи 90% респондентів. Відзначено зниження інтересу до платформ потокового відео.
- Само видавництво: Значна частина студій вибирають самостійно публікувати свої ігри.
- Інструменти та ПЗ для Розробки: Unreal Engine - лідер серед ігрових движків на другому місці Unity[5]. Важливими інструментами у розробці ігор є Photoshop, Maya та Blender.
- Інновації при розробці, використанні ШІ: Звіт висвітлює виклики, такі як фінансування, час на впровадження інновацій, а також відображає зацікавленість індустрії, щодо технологічних досягнень, таких як ШІ.

- Перспективи Індустрії: Розробники виражають бажання мати різноманітну індустрію, підкреслюючи важливість балансу між прибутковістю та оригінальністю та переходу від прибуткових моделей до якісного користувацького досвіду.

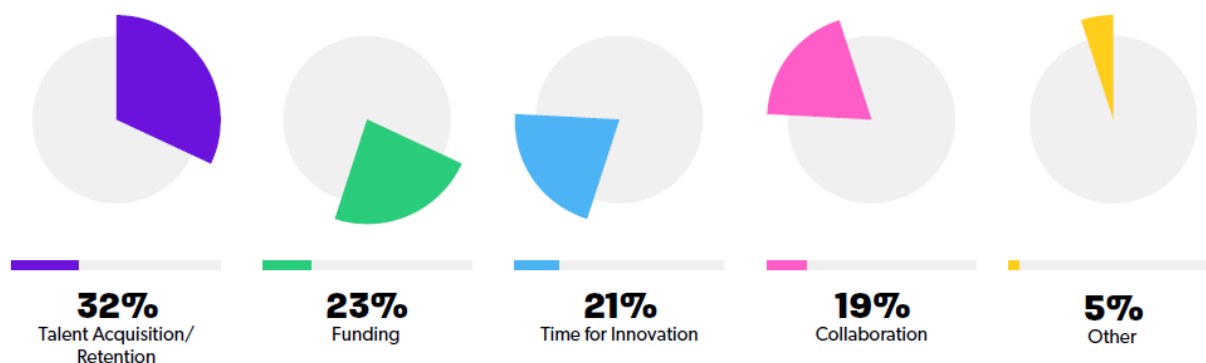


Рис. 1.1. Що є найбільшим викликом для гри під час процесу розробки

Ці висновки надають цінну інформацію для стратегічного планування та розуміння еволюції сфери розробки та дизайну ігор. В основі прогресу розробки мобільних ігор лежить гостра увага до зменшення часу та грошових інвестицій, необхідних для створення захоплюючих мобільних ігор. Розробники використовують передові технології та інноваційні методології, щоб спростити процес створення. Ця зміна є не просто перевагою, а необхідністю, викликаною постійно зростаючою конкуренцією та зростаючими очікуваннями різноманітної та вимогливої аудиторії.

Проаналізувавши ще одну доповідь від компанії Unity [3], можна зробити такі висновки: розробники ігор зосередились на підвищенні ефективності та оптимізації, а також на підтримці ігор після їх випуску. Розробники прагнуть до стійкості та адаптивності у відповідь на економічні зміни та виклики. Основні елементи доповіді включають:

- Ефективність та Оптимізація: Більшість малих студій змогли випустити ігри менш ніж за рік, працюючи менше годин. Це показує зростання ефективності у сфері мобільних ігор.

- Стратегічний Підхід до Вибору Платформ: Великі студії збільшили кількість багато платформних та крос платформних розробок, тоді як більшість малих студій (90%) вирішили розробляти ігри лише для однієї платформи.
- Підтримка Ігор після Випуску: Розробники активно працюють над тим, щоб продовжити життєвий цикл своїх ігор на 33% довше, щоб утримати існуючу базу гравців.
- Використання Генеративного Штучного Інтелекту: Очікується, що генеративний штучний інтелект буде більш активно включатися у процес розробки, щоб збільшити продуктивність творців і скоротити час та ресурси, необхідні для створення ігор.

Ці висновки від компанії Unity [3] відображають, як індустрія мобільних ігор адаптується до змінюваних умов та використовує інноваційні підходи для підтримки стабільного розвитку.

Найважливіше те, що всі розробники хочуть зменшити витрати ресурсів на свої проекти, а саме часових та фінансових. Однією з основних стратегій скорочення термінів розробки є - правильне розподілення фінансових ресурсів яке відіграє вирішальну роль у поточному стані розробки мобільних ігор. Розробники все частіше використовують економічно ефективні інструменти та платформи, які пропонують надійні можливості без високої ціни [4]. Хмарні рішення, програмне забезпечення з відкритим вихідним кодом і модульні ігрові механізми [5] – це деякі ресурси, які дають змогу розробникам створювати високоякісні ігри, не виснажуючи свої фінансові резерви. Попри жагу до скороченням термінів і бюджетів, підтримання якості ігрового досвіду залишається на першому місці. Розробники знаходять винахідливі способи збалансувати швидку розробку з високоякісними результатами.

Стає важливим адаптувати ігри до широкого спектру пристроїв, враховуючи їхні технічні характеристики. Для адаптація ігор до різноманітних платформ треба використовувати сервіси для генерації зображень, це дозволяє швидко створювати візуальний контент, який оптимізований під різні роздільні здатності та характеристики екранів. Автоматизований підхід не лише знижує час та витрати на

розробку, але й забезпечує високу якість зображень на всіх пристроях, від смартфонів до планшетів. Така універсальність є ключовою для досягнення широкої аудиторії у сучасному різноманітному цифровому середовищі.

Дивлячись у майбутнє, сектор розробки мобільних ігор готовий продовжити свій шлях до більш ефективних та економічно життєздатних методів. Ця еволюція корисна не лише для розробників, але й для ігрової спільноти, оскільки вона обіцяє майбутнє, де цікаві та інноваційні ігри доступніші, ніж будь-коли раніше. Цей шлях складний і вимогливий, але завдяки креативності, здатності до адаптації та стратегічному мисленню галузь добре споряджена, щоб пройти цю захоплюючу еру цифрових розваг.

1.2. Критичний огляд підходів до оптимізації розробки

Для того, щоб провести огляд підходів до оптимізації розробки, необхідно спочатку зрозуміти ключові аспекти та цілі, які переслідуються при такій оптимізації. Основними цілями є підвищення продуктивності, ефективності та якості розробленого продукту, а також скорочення часу і витрат на його створення [28].

Агільні методології, такі як Scrum [10] та Kanban, були розроблені з метою оптимізації процесів розробки. Вони забезпечують гнучкість, швидкість та адаптивність до змін у проекті. Ключові елементи цих методологій включають короткі цикли розробки (ітерації), постійний зворотний зв'язок від користувачів та гнучке планування.

Lean розробка фокусується на мінімізації витрат та видаленні всього зайвого з процесу розробки. Основна ідея полягає в тому, щоб створювати більше цінності для клієнта з меншими витратами. Ключовими принципами є ідентифікація та усунення втрат, оптимізація потоку роботи та неперервне вдосконалення.

Автоматизація тестування є критично важливою для скорочення часу та зусиль, необхідних для перевірки якості продукту. Це включає в себе використання різноманітних інструментів та фреймворків для автоматизації тестових сценаріїв,

що дозволяє швидко виявляти помилки та недоліки.

Неперервна інтеграція (CI) та неперервне розгортання (CD) забезпечують постійне об'єднання та тестування коду, що сприяє швидшому виявленню та виправленню помилок. Ці практики також сприяють кращій колаборації в команді та зменшують ризики, пов'язані з розгортанням.

Неперервна інтеграція (CI) та неперервне розгортання (CD) забезпечують постійне об'єднання та тестування коду, що сприяє швидшому виявленню та виправленню помилок. Ці практики також сприяють кращій колаборації в команді та зменшують ризики, пов'язані з розгортанням.

Модульне програмування допомагає оптимізувати процес розробки шляхом розбиття складного проекту на менші, незалежні модулі. Це спрощує управління кодом, підвищує його пере використання та полегшує тестування та обслуговування.

Шаблони проектування є стандартними рішеннями для загальних проблем у проектуванні програмного забезпечення. Вони допомагають створювати масштабовані, легко обслуговувані та ефективні системи.

Оптимізація процесів розробки за допомогою штучного інтелекту (ШІ) відкриває нові можливості для підвищення ефективності, якості та швидкості виробництва програмного забезпечення. ШІ може вплинути на різні аспекти розробки, від автоматизації рутинних завдань до покращення прийняття рішень. Ось декілька ключових аспектів цього підходу: ШІ може використовуватися для автоматичного генерування коду або рефакторингу існуючого коду. Це включає в себе ідентифікацію та виправлення помилок, оптимізацію коду для підвищення його ефективності та читабельності, та автоматизацію рутинних аспектів кодування. Застосування ШІ в тестуванні може значно підвищити його ефективність, зокрема через автоматизацію створення тестових сценаріїв і виявлення складних помилок. ШІ також може аналізувати результати тестів, виявляючи закономірності та проблемні аспекти, які можуть залишитися непоміченими людиною.

Прогнозування та управління ризиками

ШІ може використовуватися для аналізу даних проекту та прогнозування потенційних ризиків, допомагаючи командам своєчасно реагувати на виклики та уникати затримок у розробці. Алгоритми ШІ можуть допомогти в ідентифікації оптимального розподілу ресурсів, таких як час команди та обчислювальні потужності, для максимізації продуктивності. ШІ може аналізувати великі обсяги даних, що дозволяє розробникам краще розуміти поведінку користувачів та ринкові тенденції, а також приймати більш обґрунтовані рішення щодо розробки продукту.

Інтеграція ШІ у процес розробки вимагає чіткого розуміння його потенціалу та обмежень. Важливо враховувати етичні аспекти, а також забезпечувати надійність та безпеку використання алгоритмів ШІ. Зрештою, успіх оптимізації розробки за допомогою ШІ залежить від гармонійного поєднання людських та технічних ресурсів.

1.3. Штучний інтелект у розробці ігор

В наш час ШІ стає все більш важливою складовою у розробці ігор, оскільки він відкриває нові можливості для інновацій та покращення геймплея. Штучний інтелект[27] може використовуватися в різних аспектах розробки ігор, від створення більш переконливих неперсональних персонажів (NPC) до генерації контенту та оптимізації ігрового процесу.

- Процедурна генерація контенту

ШІ може бути застосований для процедурної генерації контенту, що дозволяє створювати великі та різноманітні ігрові світи без великої витрати ручної праці. Це включає генерацію ландшафтів, місій, історій, персонажів та інших елементів гри. Наприклад, в грі "No Man's Sky" використовується процедурна генерація для створення майже безкінечної кількості унікальних планет та екосистем.

- Покращення поведінки NPC

ШІ може надати NPC більш реалістичні та переконливі моделі поведінки. За

допомогою машинного навчання NPC можуть адаптуватися до дій гравця, вчитися з власного досвіду та навіть демонструвати емоційні реакції. Це створює більш занурюючий ігровий досвід, де дії гравця мають значний вплив на ігрове середовище.

- Адаптивний геймплей

ШІ може адаптувати ігровий процес під конкретного гравця, змінюючи рівень складності, сюжетні повороти або навіть стиль гри в залежності від дій та переваг гравця. Така адаптація робить ігровий процес більш особистим та залучаючим.

- Оптимізація ігрового балансу

ШІ може використовуватися для аналізу ігрових даних та оптимізації ігрового балансу. Алгоритми можуть виявляти невірноважені елементи гри та запропонувати зміни, що поліпшують ігровий досвід. Це особливо важливо в мультиплеєрних іграх, де підтримання справедливого балансу між гравцями є ключовим.

Практичні приклади застосування генеративних ШІ в іграх:

AI Dungeon by Latitude [21]: Використовуючи GPT-3 OpenAI, AI Dungeon — це текстова пригодницька гра, яка може генерувати необмежені наративи на основі введення користувача, пропонуючи динамічну реакцію та персоналізовані пригоди. Цей підхід забезпечує інтерактивний ігровий процес із високим ступенем відтворення.

Pingle Studios [22] дослідила потенціал штучного інтелекту в розробці ігор, зокрема у процедурному створенні контенту. Вони використовували алгоритми ШІ для створення ігрових ресурсів, таких як спрайти, 3D-моделі, текстури та звукові ефекти. Крім того, вони використали штучний інтелект для створення динамічного світу та рівнів, створення квестів і місій, а також створення оповідань. Це дозволило отримати більш персоналізований ігровий досвід, наприклад налаштувати рівні складності та адаптувати ігрову механіку на основі стилю гри гравця.

Штучний інтелект відіграє ключову роль у сучасній розробці ігор [26], пропонуючи нові можливості для творчості та інновацій. Використання ШІ дозволяє створювати більш динамічні, реалістичні та персоналізовані ігрові досвіди. Однак, важливо забезпечити правильне балансування між технічними можливостями ШІ та збереженням творчого контролю з боку розробників.

1.4. Етапи процесу розробки мобільної гри

Розробка мобільної гри є комплексним процесом, який включає в себе кілька ключових етапів [29]. Кожен з них відіграє важливу роль у створенні кінцевого продукту.

1. Концептуалізація

На цьому етапі розробляється основна ідея гри, визначаються її цілі, цільова аудиторія та геймплейні особливості. Важливо чітко сформулювати концепт, щоб у подальшому розробка була зосереджена на досягненні поставлених цілей.

2. Планування

На цьому етапі розробляється детальний план проекту, включаючи розподіл ресурсів, графік робіт та визначення ключових віх. Також відбувається вибір інструментів та технологій для розробки.

3. Дизайн

Етап дизайну включає розробку геймплея, історії, персонажів, рівнів та інтерфейсу. Тут також відбувається створення візуального стилю гри, включаючи художнє оформлення, анімацію та графічні елементи.

4. Розробка

На цьому етапі програмісти створюють код гри, реалізуючи всі заплановані функції та механіки. Відбувається інтеграція візуальних та аудіо елементів, створених художниками та дизайнерами.

5. Тестування

Тестування є ключовим етапом, на якому виявляються та виправляються помилки. Це включає в себе як тестування функціональності, так і юзабіліті-

тестування.

6. Випуск та підтримка

Після завершення розробки та тестування гра випускається на ринок. Подальша підтримка включає випуск оновлень, виправлення помилок та розширення контенту.

Фокус на створенні візуального контенту

Оскільки у роботі буде оптимізуватись процес створення візуальних елементів в мобільних іграх, розберемо цей етап детальніше. Створення візуального контенту охоплює кілька ключових аспектів:

1. Концепт-арт: На цьому етапі художники створюють первісні ескізи та концепції, які визначають візуальний стиль гри, включаючи персонажів, фони та інші графічні елементи.
2. Спрайти та текстури: Після затвердження концепт-арту, наступним кроком є створення детальних 2D спрайтів та текстур. Цей процес включає детальне прорисовування та оцифрування зображень, що будуть використовуватися в грі.
3. Анімація: Важливим аспектом є також створення анімації для 2D спрайтів. Це може включати традиційну кадр-за-кадром анімацію або використання спеціалізованих програм для анімування 2D об'єктів.
4. Інтерфейс користувача: Оформлення елементів інтерфейсу гри також здійснюється в 2D форматі. Це включає дизайн меню, іконок, кнопок, індикаторів та інших елементів управління та відображення інформації.

Кожен з цих етапів створення дизайну гри ділиться ще на декілька пунктів:

1. Концепт-арт

- Задум та Ідея: На цьому етапі художники та дизайнери розробляють первісні ідеї для візуального стилю гри. Це включає персонажів, середовище, предмети, фони та атмосферу.
- Скетчі та Наброски: Розробка базується на створенні первинних набросків, які відображають основні ідеї та концепції. Ці наброски є невід'ємною частиною процесу творчості, дозволяючи швидко

візуалізувати та обговорювати ідеї.

- Фінальний Концепт: Після обговорення та редагування, створюються фінальні версії концепт-арту. Вони демонструють деталізований вигляд елементів гри, з чітко визначеними кольорами, текстурами та стилем.

2. Спрайти та текстури

- Створення Спрайтів: Спрайти - це 2D графічні об'єкти, які використовуються для створення персонажів, предметів, та елементів середовища в грі. Кожен спрайт малюється окремо, з увагою до деталей та стилю.
- Текстуризація: Для надання спрайтам більшої реалістичності та виразності, їм додаються текстури. Текстуризація включає в себе створення та налаштування поверхонь, що імітують матеріали та освітлення.

3. Анімація

- Кадр-за-Кадром: Традиційна анімація виконується шляхом створення послідовності окремих кадрів. Для кожного руху персонажа або об'єкта малюється окремий кадр.
- Твінінг та Інші Техніки: Сучасні програми дозволяють використовувати твінінг - процес, при якому програма автоматично генерує проміжні кадри між двома позами, що спрощує процес анімації.

4. Інтерфейс користувача

- Дизайн Елементів: Елементи інтерфейсу, такі як меню, кнопки, іконки та індикатори, мають бути чіткими, зрозумілими та зручними для користувача. Дизайнери створюють ці елементи, забезпечуючи їхню консистенцію з загальним стилем гри.
- Інтерактивність: Важливо, щоб елементи інтерфейсу були не тільки візуально привабливими, але й інтерактивними, реагуючи на дії

користувача.

- Адаптація під різні пристрої: Оскільки мобільні пристрої мають різні розміри та роздільні здатності екранів, інтерфейс має бути оптимізованим для різноманітних умов використання.

В середньому кожен етап займає багато часу, це відображено на рис.

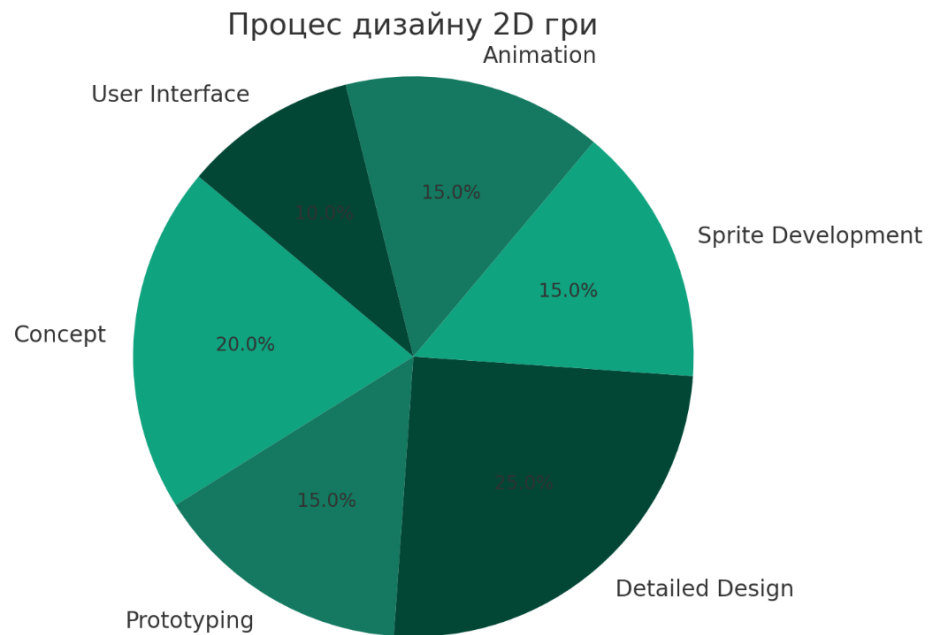


Рис. 1.2. Середній час для створення дизайну гри

Висновки та перспективи оптимізації:

Створення візуального контенту в мобільних іграх є складним та багатогранним процесом, що включає в себе ряд різноманітних завдань та вимог до креативності та технічних навичок. Використовуючи новітні технології генерації зображень можна зменшити витрати в декілька разів, тому в роботі буде розглядатись саме це питання. Застосування інструментів на базі штучного інтелекту для автоматизації створення візуального контенту дозволяє не тільки скоротити час на його розробку, але й значно підвищити якість та унікальність кінцевого продукту. Використання ШІ в процесі розробки мобільних ігор відкриває нові можливості для інновацій та креативності.

2 АНАЛІЗ ТЕХНОЛОГІЙ ТА СИСТЕМ ДЛЯ ГЕНЕРАЦІЇ ВІЗУАЛЬНОГО КОНТЕНТУ

2.1. Аналіз існуючих технологій генерації візуального контенту

Протягом останніх двох років генеративні моделі зображень на основі штучного інтелекту стрімко набули популярності, ставши значно вдосконаленими порівняно з попередніми роками [6], коли вони були менш ефективними. Спочатку, хоча технологія, що стоїть за ними, вважалася революційною та викликала захоплення серед наукової спільноти, якість генерованих зображень часто залишала бажати кращого. Раніше результати були далекими від ідеалу, що обмежувало їхнє практичне застосування. Однак, завдяки стрімкому прогресу в області обчислювальних технологій та алгоритмів глибокого навчання, ці системи тепер здатні генерувати зображення, які вражають своєю деталізацією та реалістичністю. Цей стрибок у якості та можливостях відкрив нові перспективи для використання AI у візуальному мистецтві, дизайні та інших галузях, де візуальна креативність є ключовою.

Щоб проаналізувати існуючі методи генерації візуального контенту та визначити найкращі, потрібно розглянути різні аспекти, такі як якість зображення, універсальність, доступність для користувачів, а також діапазон стилів і концепцій, які моделі можуть інтерпретувати та візуалізувати.

1. **Generative Adversarial Networks (GAN)** (Генеративні змагальні мережі): відомі тим, що створюють високоякісні реалістичні зображення, широко використовуються для завдань, де візуальна точність має вирішальне значення. Однак вони можуть потребувати інтенсивних обчислень і значного тонкого налаштування для конкретних програм.
2. **Variational Autoencoders (VAE)** (Варіаційні автокодері): Ефективні у

створенні нових зображень, які є варіантами вхідних даних. Вони часто використовуються для завдань, які вимагають певного рівня креативності, наприклад створення мистецтва. Однак якість зображення не завжди може відповідати рівню GAN.

3. **Transformers** (Трансформери) : архітектура Transformer[14] відома своєю здатністю генерувати різноманітні та образні зображення з тексту описи. Він поєднує розуміння природної мови з генерацією зображень, що робить його надзвичайно універсальним. Сила полягає в його здатності інтерпретувати та візуалізувати широкий спектр текстових підказок, включаючи абстрактні та складні ідеї.
4. **Diffusion Models**[15] (Моделі дифузії): ці моделі, як і ті, що використовуються, відомі тим, що створюють високоякісні та різноманітні зображення. Вони працюють шляхом поступового перетворення шуму в детальні зображення та набувають популярності завдяки балансу між якістю зображення та свободою творчості.

1. **Generative Adversarial Networks (GAN) Генеративні змагальні мережі.**

GAN це клас систем машинного навчання, розроблений для створення нових екземплярів даних, схожих на навчальні дані. Мережа стала важливим знаряддям у світі штучного інтелекту та машинного навчання, особливо в області генерації зображень. Ця технологія базується на ідеї змагання між двома нейронними мережами: генератором та дискримінатором. Генератор має за завдання створювати зображення, які виглядають достатньо реалістично, щоб обдурити дискримінатор, тоді як дискримінатор намагається відрізнити справжні зображення від синтетичних, створених генератором.

Основний принцип GAN є: складається з двох основних компонентів: генератора (G) та дискримінатора (D). Генератор намагається створити дані, що не можна відрізнити від справжніх, тоді як дискримінатор намагається відрізнити справжні дані від синтетичних. Принцип роботи GAN методу зображений на рис. 2.1. В статті Generative Adversarial Networks [23] можна знайти більш детальний

аналіз GAN методу стаття описує його як алгоритми штучного інтелекту, призначені для розв'язання задач генеративного моделювання, зокрема для створення реалістичних зображень високої роздільної здатності.

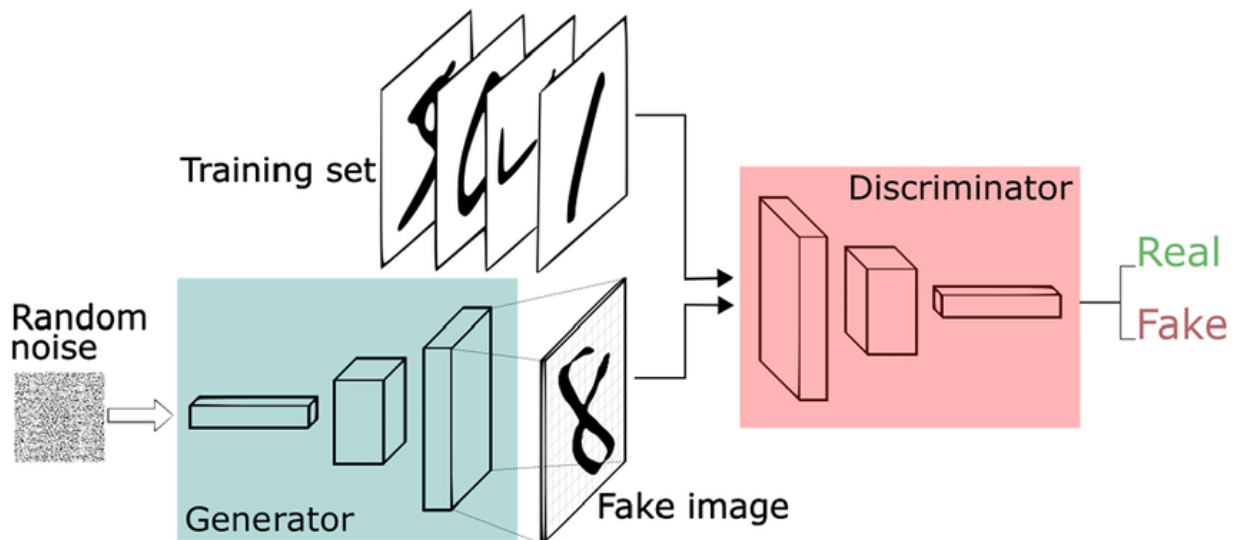


Рис. 2.1. Метод генерації зображень GAN

Математичні формули GAN

Фундаментальною частиною GAN є їх функція втрати, яка представляє собою мінімаксну гру. В математичному виразі ця функція виглядає так:

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log (1 - D(G(z)))], \quad (2.1)$$

E - означає очікуване значення,

X - представляє справжні дані,

Z - введеним шумом для генератора,

p_{data} - є розподілом справжніх даних, а

p_z - озподілом введеного шуму.

Основна ідея полягає в тому, що дискримінатор намагається максимізувати ймовірність вірного вгадування, тоді як генератор прагне створювати дані, які будуть виглядати достатньо переконливо, щоб дискримінатор їх прийняв за реальні.

Архітектура GAN включає дві основні частини. Генератор, який зазвичай є глибокою нейронною мережею, відповідає за трансформацію введеного шуму в дані, схожі на справжні. Дискримінатор, також глибока нейронна мережа, але націлена на класифікацію, визначає, чи є вхідні дані справжніми або синтетичними.

Процес тренування GAN включає чергування між оновленням генератора та дискримінатора. Однак, тренування GAN може бути складним через динаміку мінімакських ігор, і часто виникають проблеми зі збіжністю. З цієї причини багато дослідників працюють над розробкою методів та стратегій для стабілізації процесу тренування і покращення якості генерації зображень.

GAN знайшли застосування в різних областях, включаючи генерації зображень, перенесення стилів та інші. Розвиток спеціалізованих варіантів GAN, таких як Conditional GAN, CycleGAN, StyleGAN, дозволяє ще більше розширити можливості цієї технології. Незважаючи на існуючі виклики, GAN продовжують відігравати ключову роль у сучасному штучному інтелекті, відкриваючи нові горизонти в області генерації даних.

Приклади сервісів та застосунків, які використовують GAN:

- DeepArt: Сервіс, який використовує GAN для перетворення звичайних фотографій у твори мистецтва, імітуючи стилі відомих художників. Це один із прикладів використання GAN у перенесенні стилю зображень.
- ThisPersonDoesNotExist: Вебсайт, який генерує реалістичні зображення облич людей, які насправді не існують. Використовуючи GAN, сайт створює фотографії, що важко відрізнити від справжніх, показуючи потужність GAN у генерації облич.
- GANPaint Studio: Розроблений MIT, цей інструмент дозволяє користувачам редагувати зображення за допомогою глибоких мереж. Він демонструє, як GAN можуть використовуватися для розуміння та модифікації вмісту зображень.
- StyleGAN: Розроблений дослідниками з NVIDIA, StyleGAN здатний створювати дуже реалістичні обличчя та інші зображення. Цей алгоритм став популярним у дослідницькому співтоваристві та використовується

для створення різних типів зображень.

- **MakeApp:** Мобільний додаток, який використовує технологію GAN для нанесення або видалення макіяжу з обличчя на фотографіях.
- **Deepfakes:** Хоча deepfake технології часто обговорюються у контексті їх потенційного зловживання, вони також використовують GAN для створення реалістичних відеозаписів, де обличчя однієї людини замінюється на обличчя іншої.
- **RunwayML:** Платформа, яка надає різноманітні інструменти машинного навчання, включаючи GAN, для творців, дизайнерів та художників.

Ці приклади показують різноманітність способів, якими GAN можуть бути використані для створення, модифікації та покращення зображень та відео.

2. Variational Autoencoders(VAE) Варіаційні автокодері.

Модель є одним з ключових аспектів сучасного машинного навчання, особливо в областях, де важливе генерування або моделювання даних. VAE поєднують ідеї глибокого навчання та баєсовської статистики, щоб створити потужні моделі генерації.

VAE відрізняються від традиційних автокодерів своєю здатністю кодувати вхідні зображення не у фіксований код в латентному просторі, а у параметри статистичного розподілу: середнє значення та дисперсію. Це означає, що припускається, що вхідне зображення було згенероване статистичним процесом, і що випадковість цього процесу повинна враховуватися під час кодування та декодування які зображені на рис. 2.2 . VAE використовує ці параметри для випадкового вибору одного елемента розподілу, а потім декодує цей елемент назад у вхідне зображення.

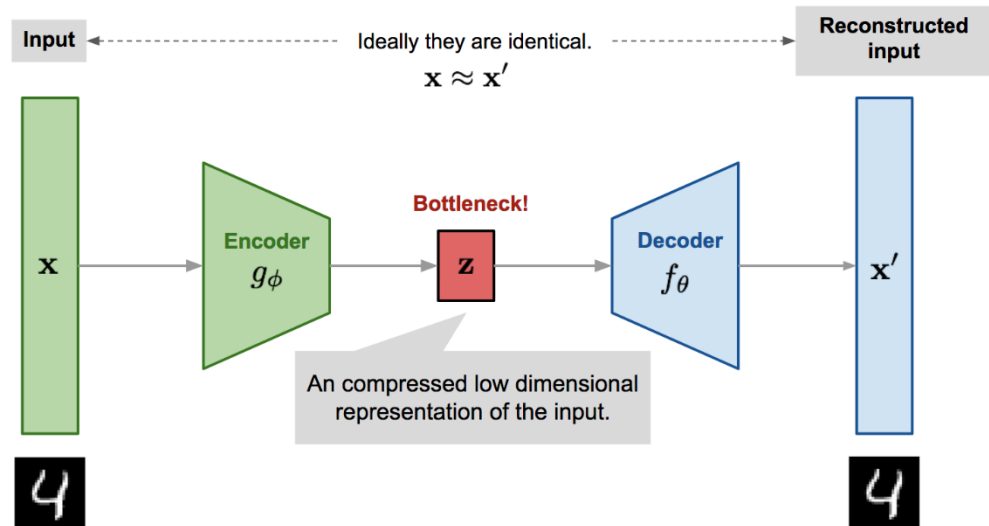


Рис. 2.2. Метод генерації VAE

Для тренування VAE використовується дві функції втрати: "втрата реконструкції" на рис. 2.2, яка змушує декодовані зразки відповідати початковим вхідним даним, та "регуляризаційна втрата", яка допомагає вивчати добре сформовані латентні простори та зменшувати перенавчання на навчальних даних. Параметри VAE тренуються шляхом використання цих двох функцій втрат [2].

Основною відмінністю VAE від традиційних автокодерів є те, що вони використовують стохастичний процес для кодування вхідних даних у прихованому просторі. Це дозволяє генерувати нові дані, виходячи з розподілу в прихованому просторі.

Основна математична формула, яка лежить в основі VAE, пов'язана з максимізацією доказовості нижньої межі (ELBO - Evidence Lower Bound). Ця формула виглядає так:

$$\mathcal{L}(\theta, \phi; x) = E_{q_{\phi}(z|x)}[\log p_{\theta}(x|z)] - D_{KL}(q_{\phi}(z|x)|p(z)), \quad (2.2)$$

\mathcal{L} - є функцією втрати для VAE.

θ та ϕ - параметрами декодера та енкодера відповідно.

x - вхідними даними.

z -прихованою змінною.

$q_{\phi}(z|x)$ – апроксимацією апостеріорного розподілу енкодера.

$p_{\theta}(x|z)$ – умовною ймовірністю декодера.

D_{KL} – розходженням Кульбака-Лейблера, яке міряє відмінність між двома розподілами.

Ця формула виражає баланс між точністю відтворення даних (перший термін) та регуляризацією розподілу прихованих змінних (другий термін). Це дозволяє VAE ефективно генерувати нові дані, зберігаючи при цьому реалістичність та різноманітність.

Варіаційні автокодери (VAE) використовуються в різноманітних додатках і сервісах, що включають генерацію та модифікацію зображень, рекомендаційні системи, та навіть у наукових дослідженнях. Ось деякі з найпопулярніших сервісів та областей застосування, де використовуються VAE:

- Академічні проекти та дослідження: В університетах та науково-дослідних інститутах VAE часто використовуються в рамках дослідницьких проектів, спрямованих на вивчення генерації та відтворення зображень.
- Adobe Firefly [20]:, особливо в їх функціях штучного інтелекту, таких як Content-Aware Fill, використовуються технології VAE, для аналізу зображень та заповнення порожніх областей у відповідності до контексту.
- Artbreeder: Платформа, що дозволяє користувачам створювати унікальні художні зображення шляхом комбінування характеристик різних зображень, використовуючи принципи, VAE.
- FaceApp: Цей додаток використовує технології глибокого навчання VAE, для зміни обличчя на фотографіях, наприклад, зміни віку, статі, зачісок тощо.

Ці сервіси і додатки демонструють широкий спектр застосувань VAE у сфері генерації зображень, від художньої обробки до автоматичного заповнення зображень та створення реалістичних обличь. Однак, важливо відмітити, що багато

сучасних систем генерації зображень використовують комбінації різних технік глибокого навчання, і не завжди можна точно визначити, чи використовується саме VAE.

3. Transformers

Трансформери, які спочатку були розроблені для завдань обробки природної мови, знайшли широке застосування у генерації зображень. Ця адаптація трансформерів для візуальних завдань є одним з найважливіших напрямків у галузі штучного інтелекту та машинного навчання. "Вступ до трансформаторів" Річард Е. Тернер [12].

Трансформери складаються з *Encoder* та *Decoder* вони є основними компонентами трансформатора. Кожен з них складається зі стопки однакових шарів. Encoder обробляє та вивчає представлення вхідних даних, а декодер використовує ці уявлення для створення вихідних даних.

Encoder: його роль полягає в обробці вхідних даних (наприклад, речення мовою оригіналу в завданні перекладу) і кодуванні їх у ряд векторів, які містять вивчене представлення вхідних даних. Енкодер не розуміє контекст або значення даних, натомість він вивчає закономірності та зв'язки всередині даних. На кожному рівні енкодера вхідні дані проходять через два підрівні: механізм самоконтролю та нейронну мережу прямого зв'язку. Механізм самоконтролю дозволяє моделі враховувати інші слова у вхідних даних під час кодування кожного слова, тоді як мережа прямого зв'язку перетворює дані далі. Вихідні дані цих рівнів потім пропускаються через залишкові з'єднання і нормалізацію рівня для покращення процесу навчання.

Decoder: його роль починається після того, як енкодер обробить весь вхідний сигнал. Декодер генерує вихідні дані (наприклад, речення цільовою мовою в завданні перекладу) одну частину за раз. Він бере вихід енкодера (закодовані вхідні дані) і використовує його для генерації виходу, звертаючи на нього вибіркочу увагу, використовуючи інший набір вивчених ваг.

Кожен рівень декодера також складається з двох підрівнів, як і енкодер, але з додатковим третім підрівнем. Це механізм самоуважності, механізм уваги

енкодера-декодера та нейронна мережа прямого зв'язку. Увага енкодера-декодера дозволяє декодеру зосередитися на різних частинах вхідної послідовності на основі поточного стану виведення.

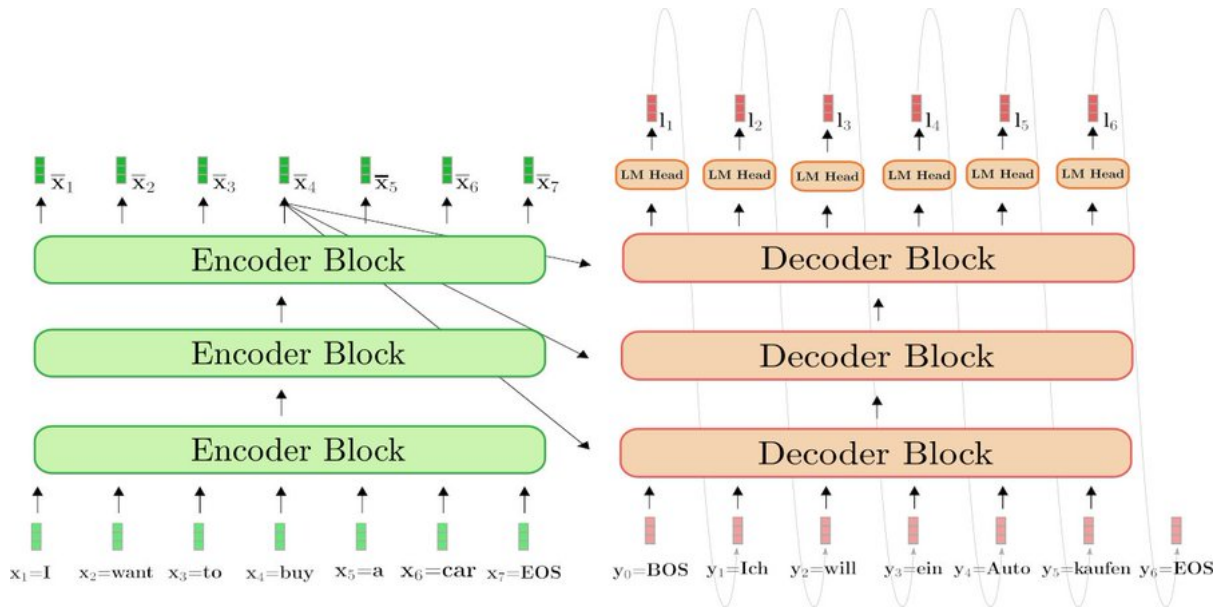


Рис. 2.3. Процес авторегресивної генерації трансформерних моделей

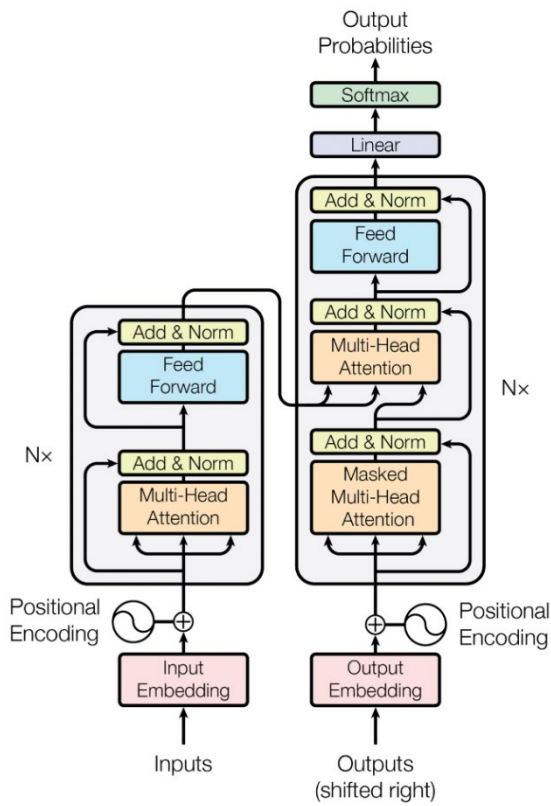


Рис. 2.4. Модель оригінального трансформера

Однією з ключових особливостей трансформерів є їхня здатність до паралельної обробки даних, на відміну від послідовної обробки у традиційних рекурентних нейронних мережах. Це дозволяє ефективніше обробляти великі зображення та створювати складні візуальні взаємозв'язки.

Основна математична формула, яка лежить в основі трансформерів, пов'язана з механізмом уваги.

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \quad (2.3)$$

Q, K, V представляють запити (queries), ключі (keys) та значення (values) відповідно, які є виводами з попередніх шарів нейронної мережі.

Функція softmax - застосовується для нормалізації ваг уваги.

d_k - розмірністю ключів, яка використовується для масштабування добутку, що забезпечує більш стабільне навчання.

Ця формула уваги дозволяє моделі фокусуватися на важливих частинах вхідних даних (наприклад, на певних ділянках зображення), що значно покращує якість генерації зображень. Трансформери, завдяки своїй здатності моделювати довгострокові залежності та робити це паралельно для всіх частин зображення, відкривають нові горизонти у генерації високоякісних та реалістичних зображень.

Технологія Transformers знайшла застосування у ряді сервісів для генерації зображень. Ці сервіси використовують складні механізми самоуваги та паралельної обробки даних, що дозволяє їм створювати високоякісні та інноваційні візуальні зображення. Ось деякі з найвідоміших прикладів:

- Google DeepMind's VQ-VAE-2: Хоча це не чистий трансформер, VQ-VAE-2 від Google DeepMind поєднує ідеї варіаційних автокодерів (VAE) з трансформерами для створення високоякісних зображень. Він використовує трансформери для моделювання залежностей у великих масштабах та генерує деталізовані зображення.
- OpenAI's DALL-E[17] від OpenAI є одним з найвідоміших прикладів

застосування трансформерів у генерації зображень. Цей сервіс може створювати унікальні зображення з текстових описів, демонструючи вражаючу здатність до креативності та розуміння складних запитів.

- **OpenAI's GPT-3 with Image Generation Capabilities:** Хоча GPT-3 був спочатку розроблений для обробки тексту, його архітектура трансформера була адаптована для генерації зображень. Це дозволяє GPT-3 створювати візуальний контент на основі текстових описів.
- **Image GPT (iGPT):** Цей проект від OpenAI демонструє, як трансформери можуть бути використані для генерації зображень, використовуючи аналогічну архітектуру до тієї, що використовується в GPT для тексту.

Ці сервіси відкривають нові горизонти в можливостях генерації зображень, показуючи, як інноваційні підходи до машинного навчання можуть змінити спосіб, яким користувач створює та взаємодіє з візуальним контентом. Використання трансформерів у цих контекстах демонструє їхню високу гнучкість та потенціал для творчих та комплексних завдань.

4. Diffusion Models (Моделі дифузії)

Моделі дифузії для генерації зображень представляють собою один з найновіших і найбільш захоплюючих напрямків у сфері генеративних моделей машинного навчання. Ці моделі використовують процес, який поступово перетворює структуроване зображення на випадковий шум, а потім навчаються відновлювати оригінальне зображення з цього шуму. Цей підхід дозволяє генерувати високоякісні, деталізовані зображення.

Diffusion Models це тип імовірнісної генеративної моделі, яка функціонує шляхом поступового руйнування даних шляхом ін'єкції шуму, а потім навчання змінювати цей процес у зворотному напрямку для створення нових зразків. Цей підхід включає два ключові процеси: прямий процес, коли дані поступово збурюються випадковим шумом, і зворотний процес, коли цей шум послідовно видаляється для створення нових зразків даних. Попередній процес зазвичай заздалегідь визначений, перетворюючи будь-який розподіл даних у більш простий попередній розподіл. З іншого боку, зворотний процес вивчається та зазвичай

представлений серією ланцюгів Маркова, які перетворюють шум назад у дані [5].

Принцип Роботи моделі дифузії складаються з двох основних процесів: процесу дифузії та процесу зворотної дифузії. Процес дифузії поступово додає шум до зображення, перетворюючи його на випадковий шумовий сигнал. Зворотний процес намагається відновити оригінальне зображення з цього шуму, використовуючи навчену модель.

Математична формула, яка описує процес дифузії, зазвичай виражається через ряд кроків, де кожен крок додає трохи шуму до зображення.

Процес дифузії починається зі зображення x_t до якого додається шум, і потім поступово відновлюється оригінальне зображення x_{t-1} шляхом видалення шуму. Нейронна мережа ϵ_θ використовується для оцінки та корекції шуму на кожному кроці, що дозволяє моделі генерувати деталізовані та реалістичні зображення. Це можна виразити наступним чином:

Формула дифузії зображення:

$$x_t = \sqrt{\alpha_t}x_{t-1} + \sqrt{1 - \alpha_t}\epsilon, \quad (2.4)$$

Формула зворотної дифузії:

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{1 - \alpha_t}{\sqrt{1 - \alpha_t}} \epsilon_\theta(x_t, t) \right), \quad (2.5)$$

x_{t-1} та x_t - це зображення на попередньому та поточному кроках дифузії відповідно.

α_t - це коефіцієнт, який контролює рівень "шуму" або змін, які вносяться на кожному кроці дифузії.

$(\alpha_t)^{-}$ - це накопичений коефіцієнт дифузії до поточного часу t

ϵ - це випадковий шум, який додається до зображення.

$\epsilon_\theta(x_t, t)$ - це нейронна мережа, яка намагається передбачити шум, доданий

на кроці t_* .

Параметри навчаються під час процесу θ тренування моделі.

Зворотний процес дифузії є ключовим для генерації зображень. Він використовує нейронну мережу для вгадування оригінального зображення з шумового сигналу. Цей процес поступово видаляє шум, відновлюючи зображення до його первісного стану.

Моделі дифузії виявилися особливо ефективними в генерації реалістичних зображень, текстур та інших складних візуальних патернів. Вони здатні створювати деталізовані та реалістичні зображення, перевершуючи багато інших підходів у якості генерованих зображень.

Однак незважаючи на свій успіх при генерації зображень, дифузійні моделі часто передбачають дорогі процедури вибірки та можуть зіткнутися з труднощами в оптимальній оцінці ймовірності. Дослідницьке співтовариство активно працює над підвищенням ефективності цих моделей у різних аспектах. У майбутньому існує потенціал для подальшого розвитку та інновацій у цій галузі, зокрема щодо підвищення ефективності цих моделей і розширення їх застосування для більш широкого кола завдань [5].

Моделі дифузії стали значним проривом у генерації зображень, використовуючи принципи статистичної фізики для створення деталізованих та реалістичних візуальних зображень. Ось декілька прикладів сервісів, які використовують методи дифузійних моделей:

- OpenAI's DALL-E 2: Це оновлена версія оригінального DALL-E, яка використовує дифузійні моделі для генерації високоякісних зображень з текстових описів. DALL-E 2 відомий своєю здатністю створювати дуже деталізовані та різноманітні зображення на основі складних запитів [6].
- Google Research's Imagen: Цей сервіс від Google також використовує дифузійні моделі для генерації фотореалістичних зображень. Imagen демонструє здатність створювати високоякісний візуальний контент, що точно відповідає заданим текстовим описам.

- Stable Diffusion[19]: Це ще один приклад застосування дифузійних моделей для генерації зображень. Stable Diffusion здатен генерувати деталізовані зображення, використовуючи відносно невелику кількість обчислювальних ресурсів, роблячи технологію більш доступною.
- Midjourney: Цей сервіс також використовує дифузійні моделі для створення зображень. Midjourney[18] фокусується на генерації візуального контенту, який має художній стиль та високий рівень креативності.

Ці сервіси показують потенціал дифузійних моделей у створенні високоякісних зображень. Вони відіграють ключову роль у сучасних інноваціях у генерації зображень, пропонуючи нові можливості для творчості та візуалізації.

Застосування моделей штучного інтелекту у генерації зображень та перетворенні тексту в візуальний контент є вражаючим поєднанням технологій штучного інтелекту, комп'ютерного бачення та обробки природної мови. Цей синтез технологій відкриває двері до нових можливостей у численних сферах, від художнього творчості до автоматизації процесів створення контенту. З розвитком досліджень у цій області ми стоїмо на порозі ери, де ще більш чудові та неймовірні інновації стануть реальністю. Ці технології вже змінюють пейзаж творчості, дозволяючи створювати точні візуальні відтворення з текстових описів та розробляти унікальні художні зображення. Вплив цих розробок у сфері штучного інтелекту невпинно формує нові горизонти у сфері творчості та інноваційних технологій.

Порівняння та аналіз методів генерації зображень, які включають Генеративно-змагальні мережі (GAN), Варіаційні автокодери (VAE), Дифузійні моделі та Трансформери, виявляє унікальні характеристики, переваги та недоліки кожного підходу. Ось короткий аналіз цих методів:

1. Генеративно-змагальні мережі (GAN)

Переваги:

- Здатність створювати високоякісні, реалістичні зображення.
- Ефективність у застосуваннях, де важлива деталізація та

різноманітність візуального контенту.

Недоліки:

- Висока складність та вимогливість до обчислювальних ресурсів.
- Тренування може бути нестабільним (наприклад, проблема mode collapse).

Застосування:

- Генерація мистецьких та фото-реалістичних зображень, моделювання стилів.

2. Варіаційні автокодери (VAE)

Переваги:

- Стабільне та ефективне тренування.
- Здатність моделювати розподіл вхідних даних.

Недоліки:

- Згенеровані зображення можуть бути менш чіткими порівняно з GAN.
- Іноді борються з втратою важливих деталей у зображенні.

Застосування:

- Генерація варіативних зображень, реконструкція зображень, робота зі зменшеним розміром даних.

3. Дифузійні моделі

Переваги:

- Висока якість генерованих зображень з додаванням деталей в кінці генерації.
- Висока здатність створювати різноманітні зображення на основі заданих умов.

Недоліки:

- Відносно нова технологія з обмеженою кількістю досліджень.
- Може вимагати великих обчислювальних ресурсів для тренування.

Застосування:

- Генерація реалістичних портретів, візуалізація даних, творче мистецтво.

4. Трансформери

Переваги:

- Велика гнучкість при написанні запитів та можливість обробки великих наборів даних.
- Ефективне моделювання складних візуальних взаємозв'язків.

Недоліки:

- Високі вимоги до обчислювальних ресурсів.
- Складність та витрати часу на тренування моделі.

Застосування:

- Генерація деталізованих зображень, інтерпретація візуальних контекстів, робота зі складними даними.

Кожен з цих методів має свої особливості, що робить їх придатними для різних задач у генерації зображень. Вибір методу залежить від специфіки завдання, якості необхідних зображень та доступних ресурсів.

У дослідженні було зосереджено увагу на пошуку методів та систем, які дозволяють генерувати зображення за допомогою штучного інтелекту на основі текстових запитів. Існують цікаві інструменти, які пропонують завантажити серію власних фотографій, а потім створюють портрети за допомогою AI, зазвичай з використанням алгоритму Stable Diffusion, однак такі інструменти не відповідали критеріям даного аналізу через їхню специфічність. Аналіз зосереджувався на самостійних генераторах зображень на основі AI, а не на інструментах, створених на їх базі.

2.2. Аналіз систем для генерації зображень

Для використання передових технологій генерації, такі як дифузійні моделі та трансформери потрібна система або платформа в якій це можна зробити. Тому доцільно виконати порівняльний аналіз цих систем генерації зображень. Особлива увага приділяється здатності цих систем до точного відтворення заданих текстових описів, їхньої ефективності, ціни, якості генерованих зображень та широті

застосування у різних сферах.

Серед існуючих систем, три найвідоміші платформи — DALL·E 3, Midjourney, та Stable Diffusion — виділяються з різних причин[25]. Ці три платформи вважаються лідерами у сфері генерації зображень за допомогою штучного інтелекту, кожна з яких пропонує унікальний підхід та функціонал.

Паралельно, відзначаються дві інші програми, які були розроблені великими корпораціями, такими як Adobe. Цікаво, що ці компанії спочатку вважали, що генеративні інструменти стануть перешкодою, але в результаті інтегрували їх у свої продукти, розширюючи можливості власних платформ.

Ці чотири системи стали єдиними, які відповідали критеріям для виконання поставленої мети роботи, і це з двох основних причин. По-перше, вони демонструють вражаючий рівень технологічного прогресу у сфері AI-генерованих зображень. По-друге, кожна з цих платформ має свої унікальні характеристики та специфікації, які відповідають різним потребам користувачів, від професіоналів до аматорів [24].

Таблиця 2.1.

Коротке порівняння систем для генерації зображень

| Система | Переваги | Ціна | Компанія |
|--------------------------------|--------------------------------------------------------------------|---------------------------------------------------------------------------------------------|-----------------------------------------|
| DALL·E 3 | Простота використання та якість | Включено з ChatGPT Plus за 20 доларів США на місяць | OpenAI |
| Midjourney | Якісні результати | Від \$10/місяць за ~200 зображень/місяць | Midjourney(доступ тільки через Discord) |
| DreamStudio (Stable Diffusion) | Налаштування та контроль генерації | Безкоштовно за 25 кредитів; від \$10 за 1000 кредитів, або безкоштовно на власному пристрої | Stability AI |
| Adobe Photoshop (Firefly) | Інтеграція створених III зображень у сервіси Adobe та в фотографії | від \$19,99/місяць у рамках плану Creative Cloud Photography | Adobe |

Тестове зображення в порівнянні цих систем можна побачити на рис. 2.5, проаналізувавши ці зображення можна зробити висновок що найгірше з задачею впоралась Adobe Firefly [20], а найкраще відображення потрібного стилю у DALL-E 3. У Midjourney найреалістичніше але не співпадає з потрібним стилем.



Рис. 2.5. Запит – Дівчинка з котиком в стилі мультфільм

DALL-E 3

Спочатку визначимось з перевагами та недоліками сервісу генерації зображень DALL·E 3 [7], він використовує трансформерні моделі генерації зображень. Однією з ключових переваг DALL·E 3 є його надзвичайна простота у використанні, а також інтеграція з ChatGPT Plus, що забезпечує значний обсяг функціоналу AI за вкладені кошти. Проте, серед недоліків варто зазначити, що керування через ChatGPT може бути нестабільним, відсутність безкоштовної пробної версії, а також відносно висока вартість у \$20 на місяць, якщо не потрібен GPT. DALL·E 3, як попередник DALL·E 2, заслужив репутацію провідного генератора зображень AI завдяки своїм значним поліпшенням. Зокрема, DALL·E 3 виробляє більш цікаві, реалістичні та послідовні результати для будь-якого текстового запиту. Завдяки цьому, OpenAI змогла знову зайняти лідируючі позиції серед конкурентів у сфері генераторів зображень AI.

DALL·E 3 відрізняється своєю простотою у використанні. Потрібно лише надати ChatGPT текстовий запит, і через декілька моментів користувач отримає

чотири варіанти зображень, створених за допомогою AI. Це досягається завдяки розумінню мови GPT, яке розширює запити, роблячи кожен результат унікальним. Водночас, DALL·E 3 не пропонує безкоштовного пробного періоду і доступний лише для підписників ChatGPT Plus. Щодо DALL·E 2, хоча його можливості простіші, він пропонує потужні функції, такі як редактор зображень для додавання додаткових кадрів або видалення елементів із зображення.

У підсумку, DALL·E 3, є одним з кращих засобів для генерації зображень, особливо завдяки його інтуїтивно зрозумілому інтерфейсу та здатності виробляти більш досконалі візуальні результати. Особливо цінною DALL·E 3 може бути для генерації зображень у галузі ігор, де якість та оригінальність візуального контенту мають вирішальне значення.

Midjourney

У дослідженні було проаналізовані переваги та недоліки сервісу генерації зображень Midjourney[18] цей сервіс базується на дифузійних моделях. Серед позитивних аспектів варто виділити здатність Midjourney стабільно створювати високоякісні зображення, що виглядають найкраще серед AI-генераторів. Крім того, спільнота користувачів сервісу є чудовим джерелом натхнення. Однак, серед негативних сторін виявлено, що Midjourney доступний лише через Discord, що може здатися незвично; зображення, створені користувачами, за замовчуванням є публічними; наразі призупинено безкоштовні пробні періоди.

Зображення, створені за допомогою Midjourney, відрізняються високою якістю та привабливістю, особливо при зображенні людей та об'єктів реального світу. Це підтверджується тим, що Midjourney став першим AI-генератором, який переміг у мистецькому конкурсі. Однак користування сервісом через Discord та публічний доступ до згенерованих зображень може стати перешкодою для комерційного використання.

На рис.2.6 Можна побачити порівняння генерації 3 сервісів, повний запит: симетричне, дуже близьке лице жінки половина лиця якої органічна і воно плачуть. Можна побачити що Midjourney зробив якісне зображення порівнянно з Firefly, однак тільки зображення від Dalle повністю відповідає заданому запиту.



Рис. 2.6. Запит – лице жінки, яка наполовину органічна та плаче

Загалом, Midjourney пропонує досить доступні тарифні плани, починаючи від \$10 на місяць, що дозволяє генерувати приблизно 200 зображень на місяць з правами на комерційне використання. Це робить Midjourney цінним інструментом для творчих особистостей і професіоналів, які шукають високоякісні візуальні рішення. Однак, у контексті генерації зображень для ігор, DALL·E, завдяки своїй більшій універсальності та можливостям, може вважатися більш переважним варіантом.

DreamStudio (Stable Diffusion)

У роботі було розглянуто DreamStudio, який використовує методи дифузії, зокрема, відкритий код Stable Diffusion[19]. Цей підхід дозволяє технічно підготовленим користувачам завантажувати та запускати модель локально на власному комп'ютері, а також навчати та точно налаштовувати модель для конкретних завдань [30]. Стабільність і гнучкість Stable Diffusion роблять його відмінним вибором для розробників, які хочуть створювати щось особливе за допомогою AI.

Серед переваг DreamStudio варто відзначити, що це єдиний великий AI-генератор зображень, який все ще пропонує безкоштовні кредити, володіє високою доступністю та можливістю налаштування, а також демонструє загалом високу якість результатів. Проте серед недоліків - складніше навчання порівняно з іншими генераторами мистецтва AI та не надто інтуїтивно зрозумілі інструменти

редагування.

DreamStudio пропонує значний контроль над різними аспектами генерації зображення з AI. Користувачі мають можливість налаштувати розмір кінцевого зображення, точність відповідності запиту, кількість кроків дифузійної моделі та кількість генерованих зображень. DreamStudio також пропонує інструменти in-painting та out-painting, доступні через Chrome, та обіцяють додати більше редакційних можливостей. Сервіс працює на системі кредитів. При реєстрації користувачі отримують 25 безкоштовних кредитів, що вистачає приблизно на 30 запитів або 120 зображень за замовчуваннями. За використання потужніших моделей, генерацію більших або більшої кількості зображень, або їхнє додаткове ітерування витрачаються кредити швидше.

Загалом, DreamStudio та Stable Diffusion надають користувачам найбільший простір для кастомізації та контролю над процесом генерації зображень AI. Вони дають можливість заглиблюватися в AI настільки глибоко, наскільки це потрібно, та навіть створювати власні AI-сервіси. Однак у підсумку, DALL·E виявляється більш підходящим для розробки графічних елементів у галузі розробки ігор, завдяки своїй високій якості зображень, можливості створювати різні стилі не змінюючи модель генерації та інтуїтивності у використанні, адаптивності до різноманітних запитів та здатності створювати унікальний візуальний контент.

Adobe Photoshop (Firefly)

У дослідженні було розглянуто Adobe Firefly [20], модель штучного інтелекту від Adobe, яка використовує методи трансформери для генерації зображень з текстових описів. Firefly, запущений у бета-версії цього року, доступний через веб-додаток Adobe Express та останню бета-версію Photoshop, що вимагає підписки на Creative Cloud для доступу. Firefly має кілька унікальних можливостей: він може генерувати нові зображення з детальних текстових описів, створювати текстові ефекти з написаних підказок, переколірувати векторне мистецтво або додавати елементи, створені AI, до зображень. Особливо вражає інтеграція Firefly з Photoshop, стандартним інструментом для редагування зображень у галузі.

Особливість Firefly полягає в функції Generative Fill, яка дозволяє використовувати звичайні інструменти Photoshop для вибору області зображення та замінювати її іншим контентом, використовуючи текстовий запит. Generative Fill розуміє контекст зображення, забезпечуючи гармонійне вбудовування нових елементів. Що стосується якості генерації зображень, результати Firefly можуть бути різними. У деяких випадках він може конкурувати з DALL·E або Stable Diffusion, але в інших - результати можуть викликати питання.

Завершуючи аналіз сервісу Firefly можна сказати: зображення створенні за допомогою сервісу не рекомендуються для комерційного використання. З огляду на можливості та інтеграцію з професійними додатками Adobe, Firefly може стати важливим інструментом для мільйонів професіоналів. Однак, якщо розглядати специфіку генерації зображень для ігор, DALL·E, з його більш універсальними можливостями, може бути кращим вибором.

Якщо порівнювати ці сервіси можна дізнатись деякі деталі про них, наприклад, серед всіх генеративних мереж тільки у DALL-E 3 не має проблем з малюванням рук це можна побачити на рис. 2.7.



Рис.2.7. Запит - Фотографія мирної сімейної обстановки вдома

Про інші сервіси для генерації зображень:

У дослідженні було вивчено різноманітність генераторів зображень, створених за допомогою штучного інтелекту. Зосереджено увагу лише на чотирьох

AI-моделях генерації зображень, оскільки інтерес полягав у самій моделі AI, а не в додатках, створених на їх основі.

Існує багато інших AI-генераторів зображень, які включають різні моделі, інтегровані у інші інструменти, такі як додатки для написання текстів, редагування фотографій або сайти зі стоковими зображеннями. Деякі з них дозволяють вибирати з кількох моделей, і кожен має свій підхід до генерації зображень AI.

Висновок полягає в тому, що серед розглянутих опцій DALL·E виділяється як найкращий варіант для генерації зображень, зокрема у контексті розробки ігор. Це обумовлено його високою якістю зображень, здатністю відтворювати складні запити та широкими можливостями кастомізації.

2.3. Аналіз та розробка шкали оцінювання зображень

Для того, щоб визначити якість зображень, згенерованих штучним інтелектом, необхідно розглянути декілька ключових параметрів, які визначають основні аспекти якості цих зображень. Розглянемо головні параметри якості, такі як Реалістичність, Деталізація, Оригінальність, Відповідність Стилю, Технічна Якість та Розпізнаваність Об'єктів, оскільки саме вони найбільше впливають на сприйняття та ефективність генеративних зображень. Інші параметри, не включені до основних, оскільки вони мають менш важливий вплив на загальне сприйняття та згенерованих зображень.

1. Реалістичність (Realism):

Реалістичність визначає ступінь точності, з якою зображення відтворює заданий об'єкт чи сцену. Цей параметр оцінює, наскільки відтворені деталі, кольори та освітлення відповідають реальному об'єкту чи сцені. Оцінка цього показника надає можливість визначити, наскільки добре зроблено відтворення об'єкту за допомогою штучного інтелекту. Оцінка в 5 балів визначає таку високу реалістичність, що важко відрізнити створене зображення від того, що створено людиною.

2. Деталізація (Detail):

Деталізація визначає рівень точності та якості відтворення дрібних елементів на зображенні. Цей параметр оцінює, наскільки відображені деталі на зображенні відповідають реальним об'єктам чи сцені. Висока деталізація показує, що навіть найдрібніші деталі відображені чітко і без розмитості, не втрачаючи при цьому якості. Оцінка в 5 балів свідчить про високий рівень деталізації, де дрібні елементи відображені максимально точно та зроблені з великою якістю.

3. Оригінальність (Originality):

Оригінальність визначає ступінь унікальності та креативності генерованого зображення. Цей параметр оцінює, наскільки створене зображення відрізняється від інших інтерпретацій об'єкта чи сцени. Вища оцінка в 5 балів свідчить про те, що зображення є абсолютно оригінальним і не схоже на жодне інше, воно виражає унікальний погляд або створене з великою дозою креативності.

4. Відповідність Стилю (Style Match):

Відповідність стилю визначає, наскільки зображення відповідає заданому стилю або тематиці. Цей параметр оцінює, наскільки вдало вдалося втілити певну естетику або атмосферу у створеному зображенні. Оцінка в 5 балів, свідчить про те, що зображення відтворює обраний стиль або тематику з великою точністю, відображаючи всі основні риси і характеристики цього стилю.

5. Технічна Якість (Technical Quality):

Технічна якість визначає відсутність технічних дефектів на зображенні, таких як артефакти, шум або розмитість. Цей параметр оцінює, наскільки добре вдалося уникнути помилок та забезпечити чистоту та чіткість зображення. Оцінка в 5 балів свідчить про те, що зображення має високу технічну якість, без будь-яких помітних дефектів, і відображається чисто та ідеально з технічної точки зору.

6. Розпізнаваність Об'єктів (Object Recognition):

Розпізнаваність об'єктів визначає ступінь легкості, з якою людське око може ідентифікувати та розуміти об'єкти на зображенні. Цей параметр оцінює, наскільки чітко та зрозуміло відображені об'єкти на зображенні, як легко можна впізнати їхні форми та розміри. Оцінка в 5 балів свідчить про те, що об'єкти на зображенні відображені точно і ясно.

Кожен показник може мати власну вагу в залежності від специфіки застосування зображення. Наприклад, для ігрової індустрії можуть бути більш важливими ре

2.4. Експертна оцінка критеріїв якості зображень

Було розроблено модель для автоматизованої генерації зображень, що збільшує швидкість створення графічного контенту. Було створено інструмент, який дозволяє розробникам швидко перетворювати первинні зображення в різноманітні художні стилі, сприяючи більш ефективному візуальному розробленню ігор. В рамках проекту була проведена експертна оцінка якості зображень, які були створені за допомогою цієї моделі.

Експертна оцінка критеріїв якості зображень, створених штучним інтелектом створена *Google form* вигляд форми зображений на рис. 2.8, де учасники оцінювали зображення трьох об'єктів: сови, восьминога і дракона, представлених у семи різних стилях: *LowPoly* приклад на рис. 2.12, *Vector*, *Neon PixelArt*, *Ultrarealistic* приклад восьминога в цьому стилі на рис. 2.9, *Fantasy-realistic* приклад на рис. 2.13, *Hand-drawn pencil*. Учасники оцінювали зображення за такими параметрами, як відповідність стилю, деталізація, оригінальність, відсутність дефектів та розпізнаваність об'єктів. Дані були зібрані з студентів вищих навчальних закладів, та з дизайнерів.

Стиль: Червоний неон, у вогні



| | 1 | 2 | 3 | 4 | 5 |
|--------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Відповідність Стилю | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Деталізація | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Оригінальність | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Відсутність дефектів | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Розпізнаваність Об'єктів | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Рис. 2.8. Гугл форма якості неонового зображення по шкалі



Рис. 2.9. Восьминіг в Ультра реалістичному стилі

Результати оцінки показали, приклад результатів можна побачити на рис. 2.10, що модель відмінно справляється з деталізацією та розпізнаваністю об'єктів, як це видно з високих середніх оцінок у цих категоріях на рис. 2.11. Однак було також виявлено, що деякі стилі, як-от hand-drawn pencil, отримали нижчі оцінки за відповідність стилю та відсутність дефектів. Це вказує на потребу подальшого удосконалення моделі для певних художніх стилів. З іншого боку, стилі, такі як

Ultrarealistic та Neon, були високо оцінені за якість, демонструючи потенціал моделі для створення реалістичних і привабливих візуалізацій.

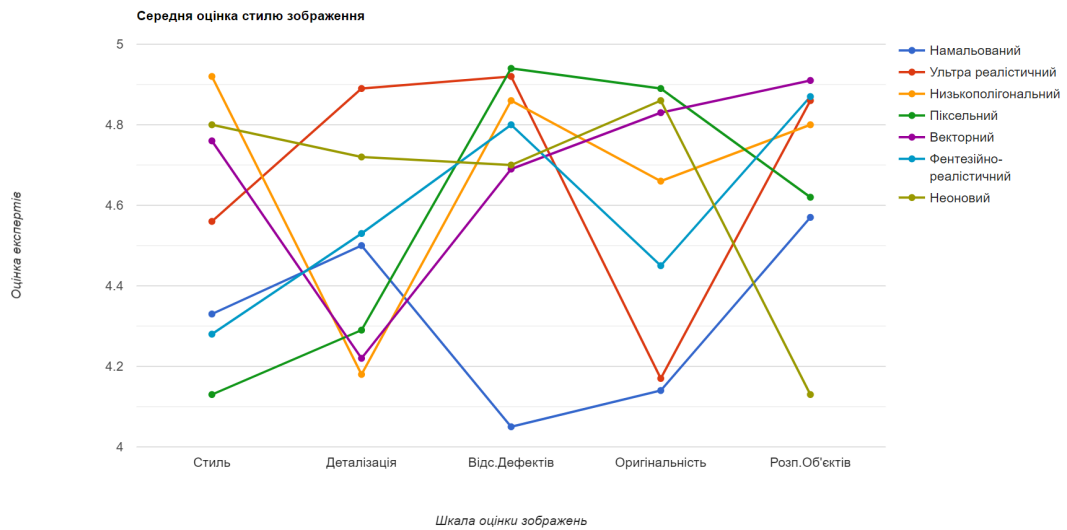


Рис. 2.10. Середня оцінка стилів експертами

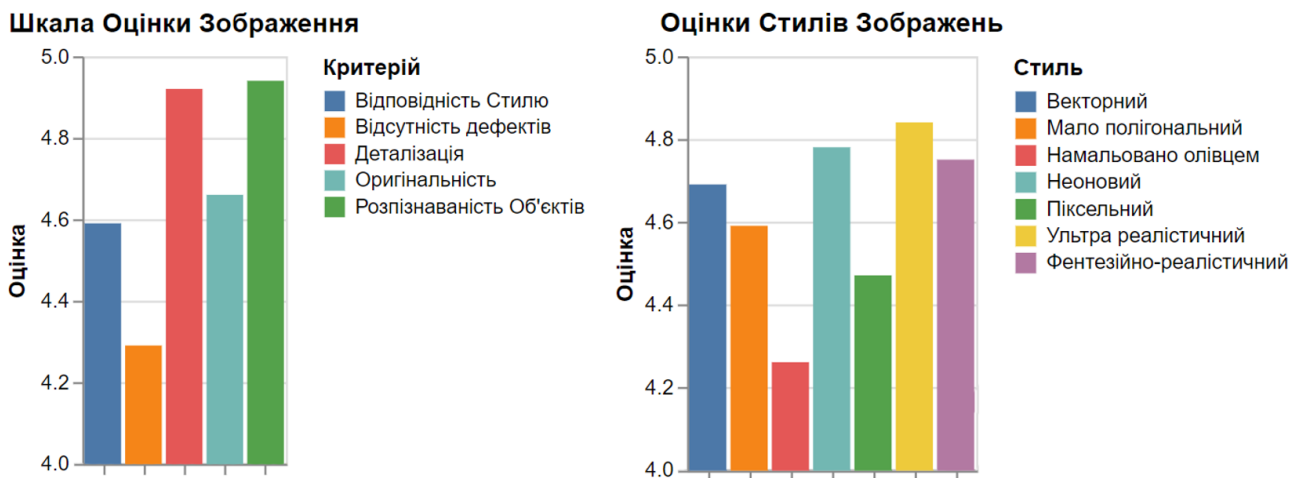


Рис. 2.11. Середня якість стилів, та якість параметрів

Оригінальність зображень була оцінена як досить висока. Відповідність стилю була оцінена трохи нижче, що вказує на те, що іноді стиль зображень не повністю відповідає очікуванням користувачів.



Рис. 2.12. Дракон в *Low-Poly* стилі



Рис.2.13. Сова в фентезійному стилі

Додатковий аналіз показав, що якість різних стилів є досить високою, особливо ультра реалістичного стилю, та неоновому стилю, як це видно на рис. 2.11. Стиль *hand-drawn pencil* отримав найнижчу оцінку, це пов'язано з тим, що модель не завжди могла ізолювати зображення від зайвих деталей. *PixelArt* теж був оцінений не дуже високо, оскільки часто в зображеннях з'являлися небажані рамки.

Загалом, ці дані допомагають зрозуміти, як різні стилі впливають на загальну якість зображення та які аспекти потребують додаткової уваги при генерації зображень за допомогою штучного інтелекту для використання у мобільних іграх. Таким чином, розроблена програма та проведені дослідження дозволяють зробити висновок про ефективність застосування штучного інтелекту в процесі розробки графіки для мобільних ігор. Автоматизація генерації зображень зменшує час, необхідний для дизайну, і водночас підвищує креативні можливості розробників, що є ключовим аспектом у конкурентному світі мобільних ігрових додатків.

2.5. Математична модель оптимізаційного рішення

У процесі розробки математичної моделі, яка застосовує методи генерації зображень для модифікації існуючих зображень згідно з заданим стилем та розміром, було вибрано поєднання AI-сервісів, а саме GPT Vision та DALL·E 3. Ця модель передбачає кілька ключових кроків: починаючи з введення зображень користувачем, їхньої обробки з використанням GPT Vision для створення текстових описів, після чого ці описи використовуються в DALL·E 3 для генерації нових зображень. На завершальному етапі, користувач отримує змінні зображення, які можна додатково обробити.

GPT Vision є частиною цієї моделі, яка відіграє ключову роль у аналізі та описі вихідних зображень. Цей сервіс базується на принципах штучного інтелекту, використовуючи розширені алгоритми для розпізнавання та інтерпретації зображень. GPT Vision здатен не лише розпізнавати різноманітні об'єкти на зображенні, але й аналізувати їх контекст, стиль та інші характеристики, що важливо для точного текстового опису. Після отримання опису, він використовується як основа для генерації нових зображень в DALL·E 3, що дозволяє створити модифіковані версії оригінальних зображень згідно з заданими параметрами стилю та розміру. Таким чином, GPT Vision відіграє важливу роль у цій моделі, надаючи можливість глибоко аналізувати та інтерпретувати візуальний контент, що є необхідним для точного та ефективного процесу генерації зображень.

Базова структура математичної моделі:

1. Введення даних

Нехай $X = x_1, x_2, \dots, x_n$, буде набором зображень, які користувач завантажує в програму. Користувач також визначає стиль

S – текстовий опис стилю зображення

R – розмір зображень

T – кількість слів для опису зображень (токени)

2. Генерація текстового опису

Кожне зображення x_i відправляється до GPT Vision, який генерує опис

зображення d_i з урахуванням вказаного стилю S , та кількості слів для опису T . Опис етапу може бути представлений як функція від зображення та стилю:

$$d_i = GPTVision(x_i, S, T), \quad (2.6)$$

3. Генерація нового зображення

Отриманий опис d_i та вказані параметри розміру R використовуються для генерації нового зображення y_i за допомогою DALL·E 3:

$$y_i = DALLE3(d_i, R), \quad (2.7)$$

4. Вивід зображень

Після генерації, користувач отримує набір зображень

$Y = y_1, y_2, \dots, y_n$, кожне з яких відповідає описам та стилю, заданому користувачем.

5. Обробка зображень

На останньому етапі користувач може вибрати опції для обрізання, зменшення або збільшення роздільної здатності зображень. Це можна виразити як функцію обробки P :

$$y'_i = P(y_i), \quad (2.8)$$

y'_i - це кінцеве оброблене зображення.

6. Повна формула математичної моделі:

Ця формула включає всі етапи процесу - від завантаження вхідних зображень та їх опису за допомогою GPT Vision, до генерації нових зображень за допомогою DALL·E 3 та їх подальшої обробки. Таким чином, вона надає інтегроване математичне представлення всього процесу генерації та обробки зображень.

$$Y' = P(\{DALLE3(GPTVision(x_i, S), R) \mid x_i \in X\}) \quad (2.9)$$

Висновок для математичної моделі:

Використання комбінації GPT Vision для аналізу та опису зображень та DALL·E 3 для їх перетворення дозволяє ефективно створювати нові зображення у заданому стилі та розмірі. Така модель є особливо корисною для генерації зображень для ігор, де потрібна висока гнучкість у створенні візуального контенту. Модель надає можливість автоматизувати процес створення візуального контенту, значно підвищуючи ефективність та творчі можливості у різних сферах.

Можливі Сфери Застосування:

- **Графічний Дизайн та Реклама:** Модель може бути використана для швидкого створення візуальних концептів та рекламних матеріалів, забезпечуючи унікальність та креативність візуального представлення.
- **Розробка Ігор та Віртуальних Світів:** Використання моделі для генерації зображень забезпечує швидке створення ігрових асетів, персонажів та елементів оточення, що важливо для розробників ігор.
- **Архітектура та Інтер'єрний Дизайн:** Модель може застосовуватися для візуалізації архітектурних проектів та інтер'єрних рішень, дозволяючи клієнтам бачити кінцевий результат до його реалізації.
- **Мистецтво та Розваги:** В області мистецтва модель може слугувати інструментом для створення нових художніх творів, а в кіно - для генерації візуальних ефектів та концепт-арту.

Плюси Моделі:

- **Гнучкість та Адаптивність:** Модель забезпечує високий рівень гнучкості, дозволяючи користувачам генерувати зображення з різноманітних текстових описів та у різних стилях.
- **Економія Часу та Ресурсів:** Автоматизація процесу створення зображень може значно знизити час та витрати, особливо у професійних сферах, де потрібне швидке прототипування.
- **Підвищення Творчих Можливостей:** Модель надає користувачам інструменти для експериментування з візуальними ідеями, які можуть

бути складно або неможливо реалізувати традиційними методами.

- Інтеграція з Існуючими Інструментами: Включення моделі у програми, такі як Photoshop, розширює можливості цих інструментів, додаючи AI-засновану генерацію зображень до їх функціоналу.
- Персоналізація та Унікальність: Модель забезпечує високий рівень персоналізації вихідних зображень, дозволяючи створювати унікальний візуальний контент.

Завдяки своїй гнучкості, здатності до швидкого прототипування та високому рівню персоналізації, розглянута математична модель є важливим інструментом у сферах, де потрібно швидко та ефективно генерувати візуальний контент. Вона особливо корисна для розробки ігор, де потрібна висока якість зображень та креативність у візуальному дизайні.

3 РОЗРОБКА ТА ВПРОВАДЖЕННЯ ОПТИМІЗАЦІЙНОГО РІШЕННЯ

3.1. Схема оптимізації розробки мобільної гри

В роботі було розроблено схему оптимізації розробки мобільної гри, що включає автоматизацію створення та модифікації графічного контенту. Центральним елементом схеми є використання програми з штучним інтелектом для генерації зображень на основі текстових описів, які вводяться користувачем.

Схема оптимізації розробки мобільної гри:

1. Завантаження Зображення:

- Розпочинається з завантаження оригінальних зображень до програми через інтерфейс користувача.
- Виконується вибір потрібних зображень із файлової системи.

2. Опис Стилю, розміру, кількості слів опису:

- Опис стилю вибирається з готових або вводиться користувачем, що включає інформацію про бажані кольори, форми та інші естетичні властивості.
- Вибір розміру зображення
- Вибір Кількості слів для опису

3. Аналіз Зображення:

- Після завантаження зображення аналізуються за допомогою GPT Vision, щоб поєднати об'єкт на зображенні та стиль користувача, після цього отримати детальний текстовий опис.

4. Генерація Зображень:

- На основі отриманого текстового опису програма звертається до моделі DALL·E 3, яка генерує нові зображення з врахуванням заданих параметрів.

5. Перегляд і Відбір Зображень:

- Генеровані зображення відображаються в інтерфейсі програми, де користувач може переглянути та відібрати найбільш підходящі.
6. Обробка зображення:
- Можливість змінити роздільну зображення кожного зображення одночасно, обрізати всі зображення за певною висотою та шириною, та деякі додаткові функції обробки, як повороти зображень.
7. Додаткова обробка в сторонніх програмах:
- Зображення імпортуються до графічного редактора, такого як Adobe Photoshop або GIMP, де вони можуть бути додатково відредаговані для використання в мобільній грі.

Весь процес відбувається в контрольованому середовищі з мінімальним втручанням користувача, що забезпечує швидкість та ефективність розробки. Оптимізація полягає в тому, що програма максимально автоматизує повторювані та трудомісткі завдання, звільняючи час та ресурси розробників для креативних аспектів розробки гри. Схема покращує ітераційний процес дизайну, дозволяючи швидше досягати бажаного візуального ефекту та вносити зміни без значних затрат часу. Діаграма Послідовності для цієї схеми оптимізації відображення на рис. 3.1

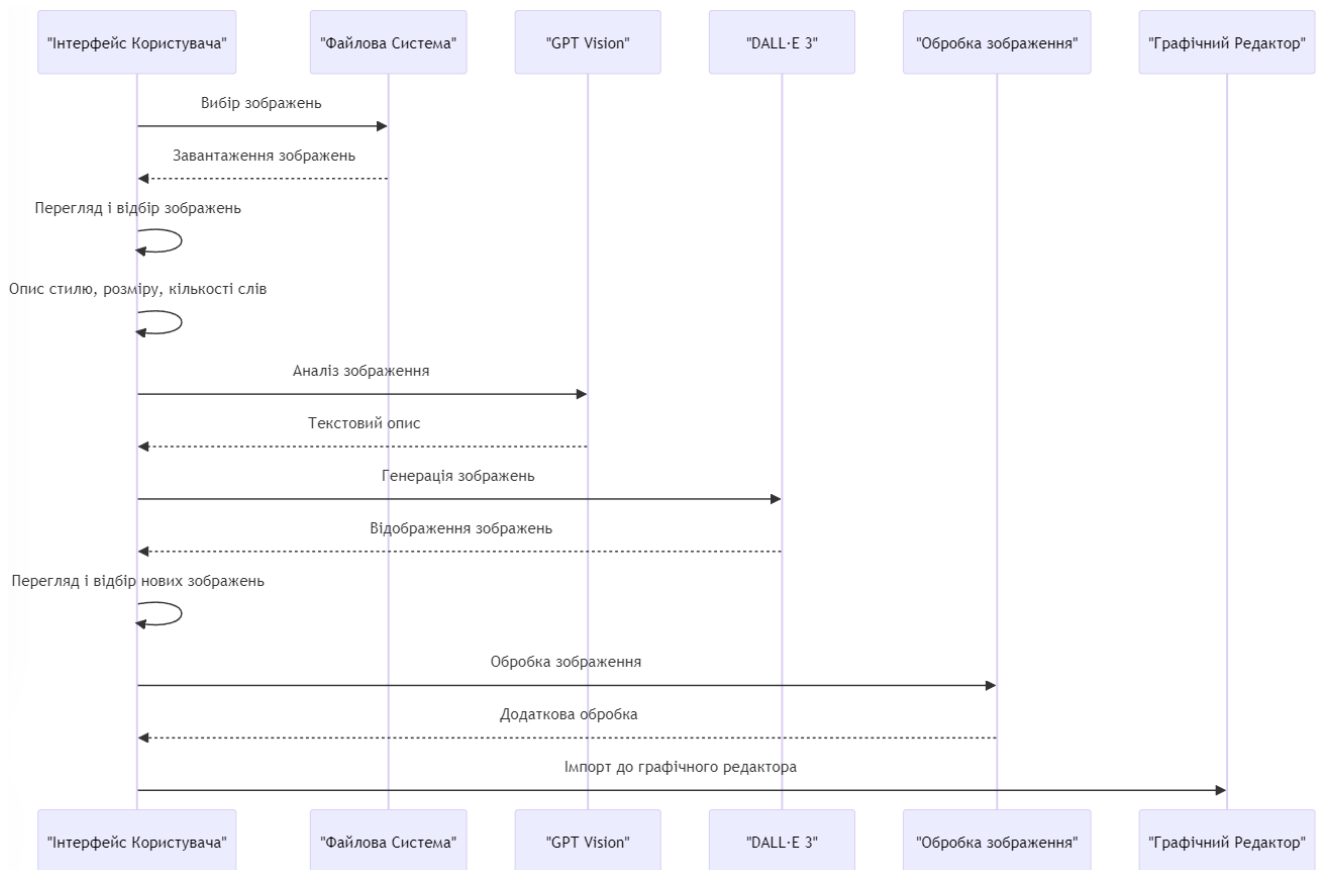


Рис.3.1. Діаграма послідовності

3.2. Опис використаних програмних засобів

Для створення та виконання математичної моделі програми яка використовує методи генерації зображень для зміни існуючих зображень за певним стилем та розміром, було використано ряд програмних засобів, кожен з яких відіграє ключову роль у процесі.

Огляд основних інструментів:

- Мова програмування Python:

Для розробки моделі було використано мову програмування Python, яка за останні роки здобула велику популярність та визнання серед розробників у всьому світі. Python відрізняється своєю універсальністю та високим рівнем читабельності коду, що робить його ідеальним вибором для широкого спектру програмних проектів - від веб-розробки до наукових досліджень. Однією з ключових особливостей Python є його простота у використанні. Синтаксис мови є чітким та

інтуїтивно зрозумілим, що дозволяє новачкам швидко навчитися основам програмування. Ця характеристика також сприяє зниженню витрат на розробку та підтримку програм, оскільки код легше читати та розуміти.

Python підтримує різні парадигми програмування, включаючи об'єктно-орієнтоване, процедурне та функціональне програмування, що робить його вкрай гнучким. Така багатофункціональність дозволяє розробникам використовувати Python для різноманітних проектів, адаптуючи стиль кодування до конкретних потреб проекту. Python також славиться своєю масштабованістю та продуктивністю. Він здатен ефективно обробляти великі обсяги даних і виконувати складні обчислення, що робить його популярним вибором у сферах наукових досліджень, машинного навчання, штучного інтелекту, аналізу даних та великих даних. Ще однією перевагою Python є величезна спільнота розробників та багата екосистема. Велика кількість бібліотек та фреймворків, доступних для Python, робить мову ще більш потужною. Бібліотеки, такі як NumPy для наукових обчислень, Pandas для аналізу даних, TensorFlow та PyTorch для машинного навчання, значно спрощують розробку складних програм.

Завдяки своїм перевагам, Python широко застосовується у різноманітних галузях, включаючи розробку веб-додатків, систем автоматизації, наукових досліджень, геймдев, інтернету речей та багато інших. Його універсальність, легкість у вивченні, підтримка різних парадигм програмування, масштабованість та активна спільнота роблять Python однією з найбільш затребуваних та впливових мов програмування сучасності.

- **Tkinter Python library**

Tkinter, бібліотека Python, потрібна для створення графічних інтерфейсів користувача (GUI). Це стандартний інструмент для розробки GUI в Python, який використовується як професійними розробниками для створення інтерактивних та візуально привабливих додатків. Завдяки своїй простоті та гнучкості, Tkinter став популярним вибором для багатьох проектів, що потребують графічного інтерфейсу, включаючи невеликі особисті проекти, освітні додатки та прототипи програмного забезпечення.

Tkinter - це бібліотека Python для створення графічних інтерфейсів користувача. Її назва походить від Tk GUI Toolkit, з яким вона інтегрована, та Python. Tkinter дозволяє розробникам швидко та ефективно створювати інтерактивні GUI, використовуючи стандартні віджети, такі як кнопки, текстові поля, мітки та інші.

- **Pillow Python library**

Бібліотека Pillow Python була застосована у розробці моделі програми, використовувалася для обробки зображень. Зокрема, вона дозволяла змінювати розміри зображень, виконувати основне редагування, таке як обрізання або ротація, та підготовляти зображення до подальшої обробки за допомогою AI-моделей.

Pillow - це вдосконалена версія бібліотеки PIL (Python Imaging Library), яка є відкритою та широко використовується для роботи з зображеннями у Python. Бібліотека Pillow надає широкий спектр можливостей для обробки зображень, включаючи читання, запис, обробку та відображення зображень у різних форматах.

- **Asyncio Python library**

Бібліотекою, яка надає можливість написання конкурентного коду з використанням синтаксису `async/await`. Вона є частиною стандартної бібліотеки Python і призначена для побудови асинхронних мережевих програм. Asyncio дозволяє ефективно виконувати багато операцій вводу/виводу та інші операції, які можуть блокувати потік виконання, без необхідності використання багато поточності. Основні можливості asyncio включають підтримку асинхронних викликів функцій, встановлення та управління циклами подій, створення задач та обробка ф'ючерсів. Вона реалізує патерн програмування "event loop", де події (наприклад, вхідні дані, закінчення операцій) керують потоком програми.

Для розробки методу asyncio була використана для керування асинхронними запитами до AI-сервісів, таких як GPT Vision та DALL·E. Використання asyncio дозволило оптимізувати обробку множини запитів до цих сервісів, особливо важливо у випадку обробки великої кількості зображень або при високій навантаженості на систему. Asyncio допомогла ефективному управлінні асинхронними мережевими запитами та відповідями, не блокуючи основний потік

програми під час очікування завершення цих операцій. Це поліпшує загальну продуктивність системи, зменшуючи час відгуку та підвищуючи пропускну здатність.

- OpenAI Python library

Бібліотека є набором інструментів, розроблених компанією OpenAI, які дозволяють розробникам інтегрувати можливості штучного інтелекту, зокрема машинного навчання та глибокого навчання, у свої програми. Ця бібліотека включає доступ до різних AI-моделей та API, зокрема до таких популярних інструментів, як GPT (Generative Pretrained Transformer) та DALL·E.

OpenAI Python library відіграла важливу роль для створення методу. В програмі, спочатку використовувалася бібліотека для аналізу вхідних зображень і створення текстових описів їх стилів за допомогою GPT Vision, який є частиною OpenAI. Після цього, ці описи відправлялися до моделі DALL·E, також доступної через OpenAI Python library, для генерації нових зображень, що відповідали заданим стилям та розмірам.

Використання цієї бібліотеки дозволило програмі ефективно комбінувати аналіз зображень та генерацію зображень, забезпечуючи таким чином створення змінених зображень за певними параметрами. Це підкреслює важливість інтеграції сучасних AI-технологій у програмні рішення, надаючи розробникам могутні інструменти для створення інноваційних додатків.

- GPT Vision (від OpenAI)

GPT Vision використовувався для генерації детальних описів вхідних зображень. Це дозволило створити точний текстовий контекст, який відповідав стилю та характеристикам кожного зображення. Такі описи стали ключовими для подальшої обробки зображень за допомогою моделі DALL·E від OpenAI, яка спеціалізується на генерації зображень із текстових описів. Інтеграція GPT Vision у процес розробки дозволила створити ефективну та інноваційну систему, яка забезпечує гнучке та точне перетворення візуальних даних.

- DALL·E 3 (від OpenAI)

Штучний інтелект, який здатний генерувати високоякісні зображення з

текстових описів, у моделі він використовується для створення нових зображень на основі текстових описів, отриманих від GPT Vision, з урахуванням заданих параметрів розміру, використовуючи передові методи трансформерів для інтерпретації та візуалізації текстових запитів.

3.3. Опис структури проекту

Структура була організована для ефективної роботи з генерацією зображень. Проект включав кілька ключових компонентів, кожен з яких виконував певну роль у загальному процесі рис.3.2.



Рис.3.2. Діаграма: Mind Map

- **Модуль Завантаження Зображень:**

Цей модуль було використано для імпортування існуючих зображень, які користувач бажав змінити. Він дозволяв користувачам завантажувати зображення у різних форматах і готував їх до подальшої обробки.

- **Інтеграція з GPT Vision:**

Після завантаження зображень, програма використовувала GPT Vision для аналізу зображень та генерації текстових описів їх стилів. Цей процес забезпечував створення точних описів для кожного зображення, що слугували вхідними даними

для генерації нових зображень.

- Використання DALL·E 3 для Генерації Зображень:

З використанням текстових описів, отриманих від GPT Vision, DALL·E 3 генерував нові зображення, що відповідали запитам користувачів. Цей модуль забезпечував перетворення текстових описів на візуальні зображення.

- Модуль Обробки Зображень:

Отримані від DALL·E 3 зображення подавалися у модуль обробки, де вони могли бути додатково модифіковані, включаючи зміну розмірів, обрізання чи інші необхідні корективи.

- Інтерфейс Користувача:

Програма має інтуїтивно зрозумілий інтерфейс користувача, який дозволяє легко між різними модулями, завантажувати зображення, вводити параметри запиту та переглядати генеровані зображення.

- Модуль Зберігання та Експорту:

Нарешті, було включено функціонал для зберігання та експорту готових зображень, дозволяючи користувачам зберігати або використовувати згенеровані зображення за межами програми, побачити як працює метод покроково можна на діаграмі.

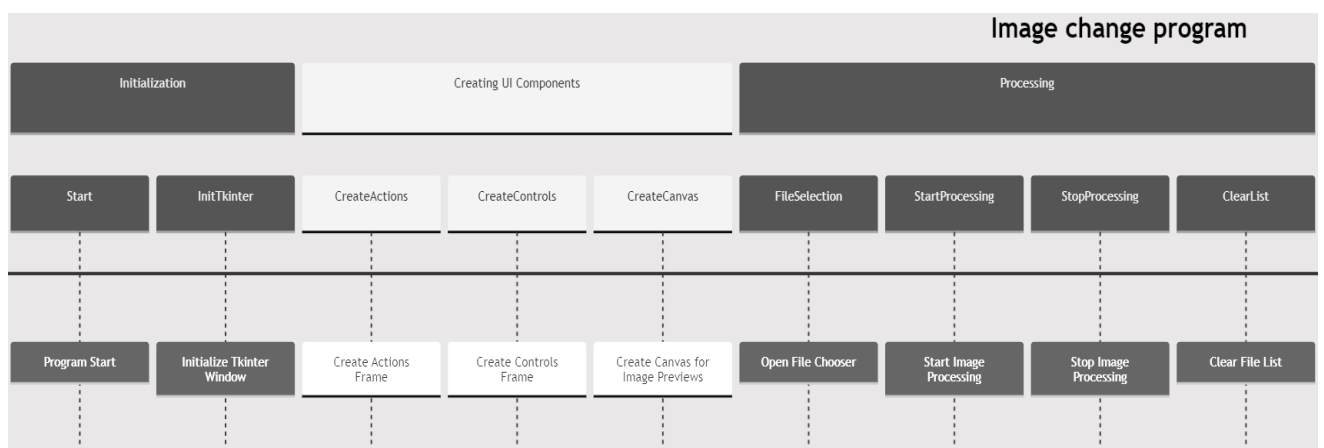


Рис. 3.3. Діаграма: Roadmap частина 1

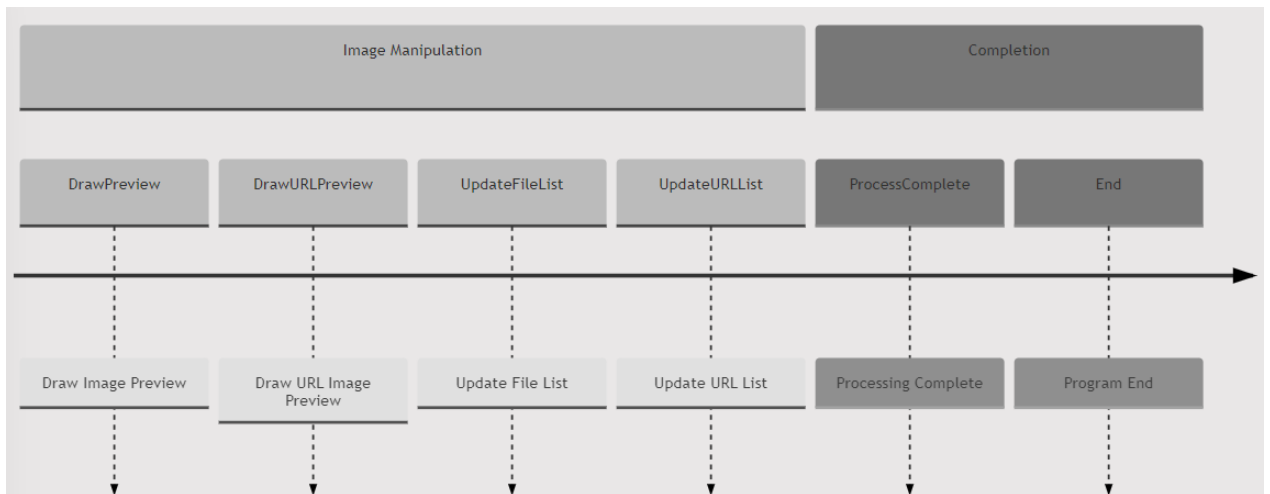


Рис. 3.4. Діаграма: Roadmap частина 2

Така структура проекту забезпечувала гладку взаємодію між різними компонентами, від завантаження зображень до їх остаточного експорту, створюючи ефективний та зручний процес для генерації зображень за заданими користувачем стилем і розміром.

Опис інтерфейсу

Інтерфейс користувача програми можна побачити на рис. 3.5. Він був створений за допомогою набору графічних бібліотек та фреймворків, які сприяють розробці інтуїтивно-зрозумілих і зручних для користувача графічних інтерфейсів. Зокрема, для створення візуальних елементів інтерфейсу могли бути використані такі інструменти, як Tkinter, який є стандартною бібліотекою для створення GUI в Python.

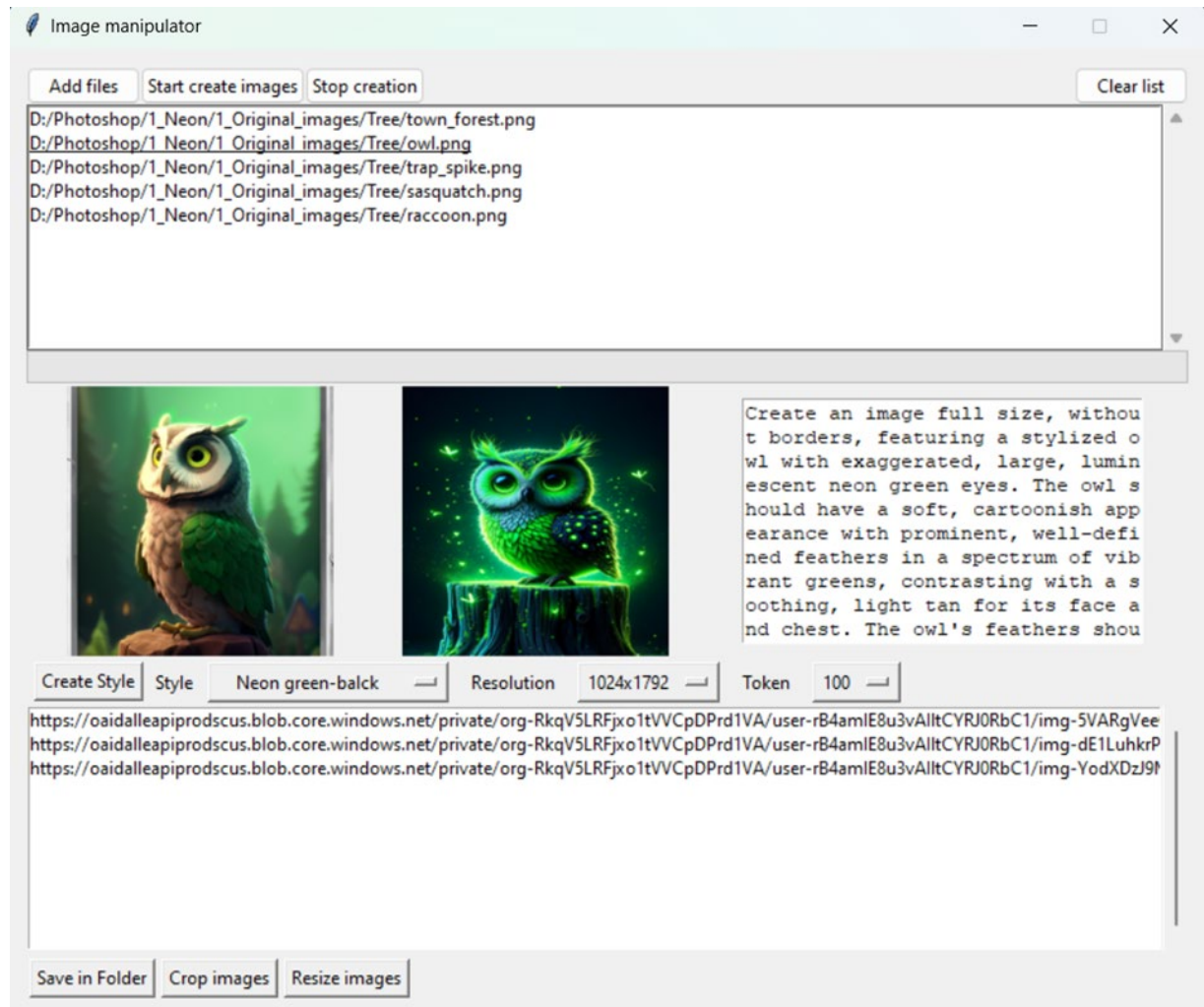


Рис. 3.5. Інтерфейс програми для генерації зображень

- Список Файлів:

У верхній лівій частині інтерфейсу розташований список файлів, які користувач додав до програми для обробки. Є можливість "Add files" для додавання нових зображень, "Start create images" для початку процесу створення або модифікації зображень, "Stop creation" для зупинки процесу та "Clear list" для очищення списку завантажених файлів.

- Перегляд Зображень:

Центральна область інтерфейсу показує вже оброблені зображення. Тут користувач може візуально оцінити результати генерації чи модифікації зображень.

- Панель Керування Стилем:

У нижній частині інтерфейсу розміщені елементи керування для визначення

стилю зображень. Тут можна вибрати "Create Style", що дозволяє створити новий стиль, або вибрати вже створений "Style" з наявних опцій (наприклад, "Neon green-black"), визначити роздільну здатність зображення ("Resolution") та кількість токенів, які будуть використані для генерації ("Token").

- **Текстове Поле для Опису:**

У правій частині інтерфейсу є текстове поле, де користувач може ввести або редагувати текстовий опис для генерації зображення. Цей опис є інструкцією для AI-моделі про те, як повинне виглядати кінцеве зображення.

- **Кнопки Дій для Зображень:**

Внизу є кнопки "Save in Folder", що дозволяють зберегти готові зображення, "Crop images" для їх обрізки та "Resize images" для зміни розміру.

- **URL-адреси для Запитів:**

У нижній частині, під полем для текстового опису, розташовані URL-адреси, зображень які згенерував DALL-E 3, зображення по цьому електронному ресурсу можна завантажити собі на пристрій.

3.4. Опис розроблених класів

У розробленому проекті були створені класи, що відповідають за різні аспекти роботи програми для обробки та генерації зображень діаграму класів можна побачити на рис. 3.6 – рис. 3.7.

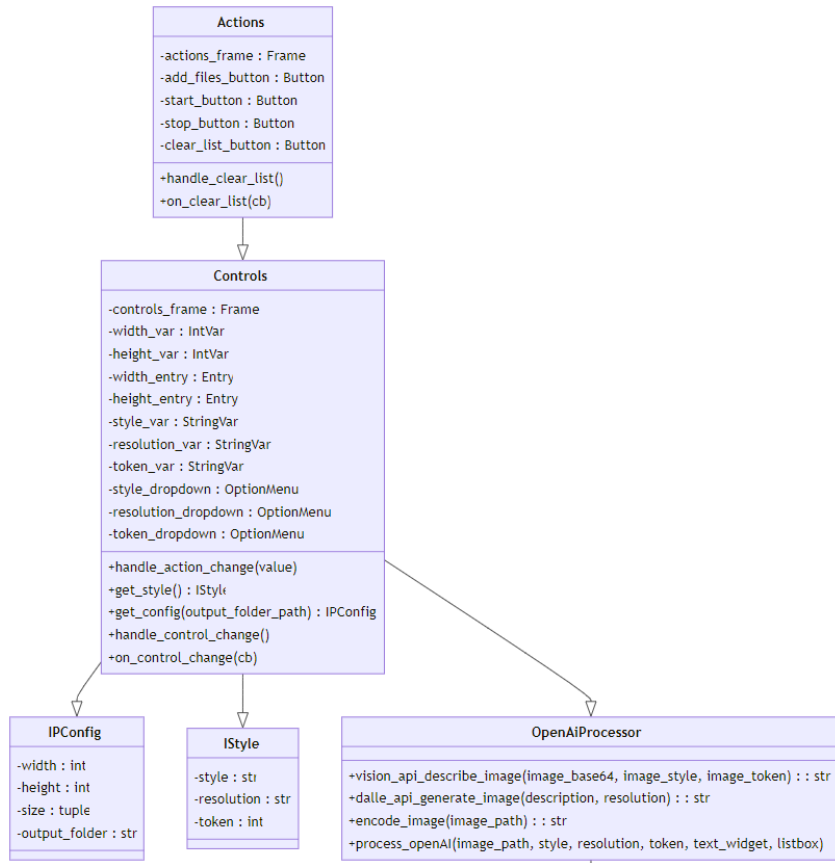


Рис. 3.6. Діаграма класів 1 частина

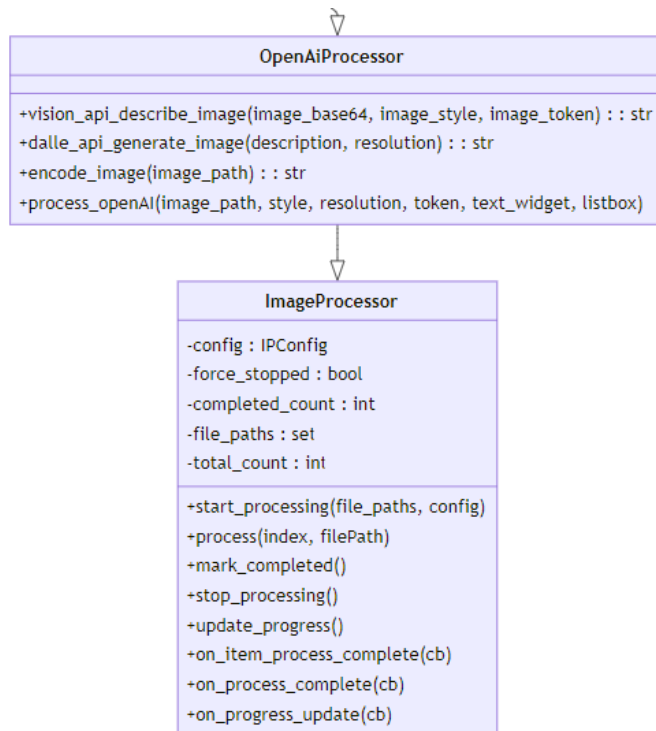


Рис. 3.7. Діаграма класів 2 частина

Клас Actions займається керуванням дій користувача, таких як додавання файлів, запуск та зупинка процесу генерації зображень, а також очищення списку завантажених зображень. Він включає у себе метод `handle_clear_list`, який очищує список завантажених файлів, та `on_clear_list`, який дозволяє зв'язувати зовнішню функцію з кнопкою очищення.

Клас Controls відповідає за управління параметрами генерації зображень, такими як стиль, роздільна здатність та кількість токенів. Він дозволяє користувачам вибрати або створити новий стиль за допомогою діалогового вікна та управляє вхідними даними для генерації зображень.

IPConfig та **IStyle** є класами конфігурації, які зберігають параметри для обробки зображень та стилі, обрані користувачем. Вони містять інформацію про розміри, стиль та токени, які будуть використовуватися під час генерації.

Клас OpenAiProcessor виконує інтеграцію з API OpenAI для отримання описів зображень та генерації нових зображень за цими описами. Він включає методи для кодування зображень у формат, прийнятний для OpenAI, відправлення описів зображень та отримання нових зображень з API, для прикладу коду ось метод для доступу до сервісу опису зображень GPTVision.

```
class OpenAiProcessor:
    openai.api_key = ""

    def vision_api_describe_image(image_base64, image_style, image_token):
        response = openai.chat.completions.create(
            model = "gpt-4-vision-preview",
            messages=[
                {
                    "role": "user",
                    "content": [
                        {
                            "type": "image_url",
                            "image_url": {
                                "url": image_base64
```

```

    }
  },
  {
    "type": "text",
    "text": f"new style: '{image_style}',
    Describe in detail the image for dalle3,subject in the image,
    camera location,
    so as to full combine the style and elements of that image and the
    new style.
    Don't create an image, start straight away with the text for dally3
    start with the word
    'Create an image full size, without borders ...'"
  }
]
},
],
max_tokens=image_token
)
description_text=response.choices[0].message.content
return description_text

```

ImageProcessor є класом, що відповідає за безпосередню обробку зображень, таку як зміна розміру або обрізання. Він управляє процесом обробки та зберігання зображень, надаючи методи для запуску та зупинки обробки, оновлення прогресу та зворотного виклику по завершенню процесу.

Всі ці класи взаємодіють між собою в рамках головного інтерфейсу користувача, створеного за допомогою tkinter, що включає віджети для вибору файлів, відображення прогресу та управління параметрами генерації. Взаємодія з API OpenAI реалізована через асинхронні запити, що дозволяє забезпечити ефективну обробку без блокування графічного інтерфейсу.

3.5. Економічний аналіз зменшення витрат на розробку

Для аналізу економічної ефективності зменшення витрат на розробку було проведено порівняння вартості створення зображень різними фахівцями з різними рівнями кваліфікації та з використанням штучного інтелекту. У ході аналізу були розглянуті витрати на годину праці різних категорій спеціалістів але однієї професії *2D Artist* - від *Junior* до *Senior*. Для кожної категорії визначено середній час, необхідний для створення одного зображення, та відповідну вартість цієї роботи.

Паралельно було враховано вартість використання технологій штучного інтелекту для генерації зображень, що дозволило оцінити потенційну економію від їх застосування. Розрахунок економії базувався на порівнянні часу та вартості створення зображень людьми і з використанням ШІ, з урахуванням додаткових витрат на використання цих технологій.

Данні для визначення середньої зарплати взяті з [9],[8]. Визначення середнього часу розробки одного зображення були взяті з [7].

Середня зарплата художника за місяць та за годину:

Junior - отримує 7 537 грн на місяць або 47,11 грн в годину.

Middle - отримує 44 328 грн місяць, 277 грн в годину

Senior - отримує 84 334 грн місяць, 535 грн в годину.

Середній час для створення середнього розміру ілюстрації:

Junior - від 2 до 5 год.

Middle – від 2 до 3 год.

Senior – від 1 до 2 год.

Ціна за генерацію одного зображення 4.4 грн, генерується в середньому за 1 хв, але після цього іноді треб додаткова обробка людиною від 30хв-2 год.

Математичні формули розрахунків:

Середня ставка спеціаліста (C_s):

$$C_s = \frac{\sum_{i=1}^n C_{s_i}}{n}, \quad (3.1)$$

Середній час на створення оригінального зображення (T_s):

$$T_s = \frac{\sum_{i=1}^n T_{s_i}}{n}, \quad (3.2)$$

Середня ціна для створення оригінального зображення (C_o):

$$C_o = \frac{\sum_{i=1}^n C_{o_i}}{n}, \quad (3.3)$$

Час на створення готового зображення з використанням генерації (T_{sai}):

$$T_{sai} = (T_{ai} \times N) + \left(\frac{T_s}{S_k}\right), \quad (3.4)$$

Ціна за створення готового зображення з використанням генерації (C_{sai}):

$$C_{sai} = (T_{sai} \times C_s) + (C_{ai} \times N), \quad (3.5)$$

Якщо:

N – Кількість згенерованих зображень яка потрібна.

S_k – Технічна якість згенерованого зображення, 5 бальна система.

C_{ai} – Вартість генерації зображення = 4.4 грн.

T_{ai} - Час генерації зображення = від 40сек, до 2 хв.

Зарплати розраховані в порядку зростання від найдешевшого Junior до найдорожчого Senior.

Таблиця 3.1.

Порівняння витрат для створення зображення

| Ім'я | Ціна за годину (грн/год) | Час (год) | Ціна за зображення (грн) | Час з методом (год) | Ціна з методом (грн) |
|-------------|---------------------------------|------------------|---------------------------------|----------------------------|-----------------------------|
| Elena | 72.06 | 3.33 | 240.26 | 2.97 | 261.99 |
| Maria | 83.68 | 2.57 | 214.96 | 1.14 | 107.48 |
| Dmitry | 117.00 | 2.05 | 239.82 | 3.01 | 393.17 |
| Lena | 149.62 | 3.48 | 520.82 | 1.48 | 235.66 |
| Max | 170.06 | 2.31 | 392.76 | 0.37 | 67.63 |
| Alex | 189.97 | 2.53 | 479.86 | 0.76 | 157.52 |
| John | 217.00 | 2.00 | 434.20 | 1.66 | 375.51 |
| Nikolay | 222.45 | 2.08 | 462.37 | 2.73 | 651.99 |
| Olga | 253.29 | 2.16 | 547.15 | 3.02 | 783.68 |
| Ivan | 262.56 | 3.46 | 908.20 | 0.65 | 173.85 |
| Pavel | 274.18 | 2.76 | 757.72 | 2.51 | 718.36 |
| Sasha | 355.12 | 2.39 | 850.06 | 0.79 | 288.30 |
| Anna | 355.29 | 2.35 | 833.86 | 0.51 | 183.22 |
| Irina | 371.74 | 2.83 | 1051.86 | 1.80 | 687.05 |
| Yuri | 373.74 | 3.15 | 1177.91 | 2.51 | 950.03 |
| Viktor | 414.62 | 3.22 | 1334.59 | 1.51 | 634.21 |
| Katya | 472.82 | 2.48 | 1174.00 | 0.64 | 312.19 |
| Olena | 477.64 | 3.29 | 1570.86 | 0.20 | 96.81 |
| Sophia | 497.65 | 2.92 | 1451.75 | 2.21 | 1121.36 |
| Sergey | 533.97 | 3.03 | 1619.34 | 0.66 | 357.40 |

Аналіз цієї таблиці, а саме порівняння ціни створення зображення на рис. 3.8, порівняння часу який потрібен для створення зображення на рис. 3.9 Проаналізувавши ці данні можемо вже побачити, що використання розробленого методу для створення зображень є дуже ефективним.

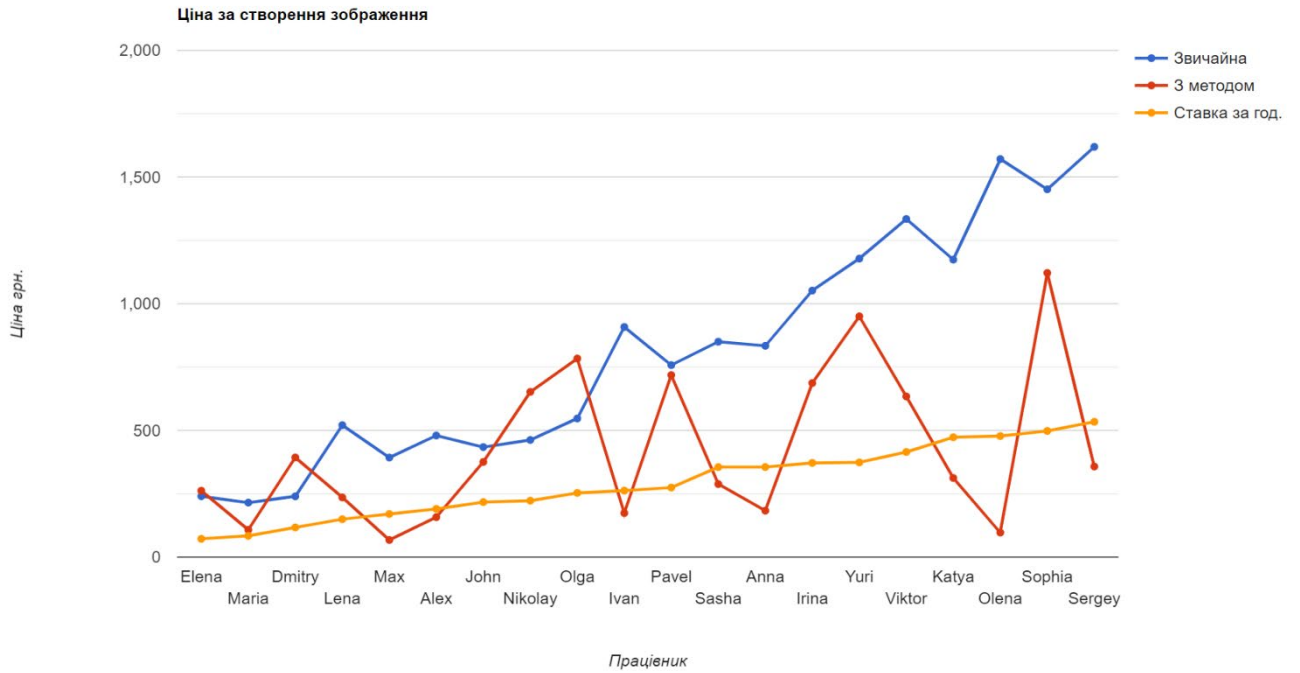


Рис. 3.8. Лінійна діаграма порівняння ціни для створення зображення

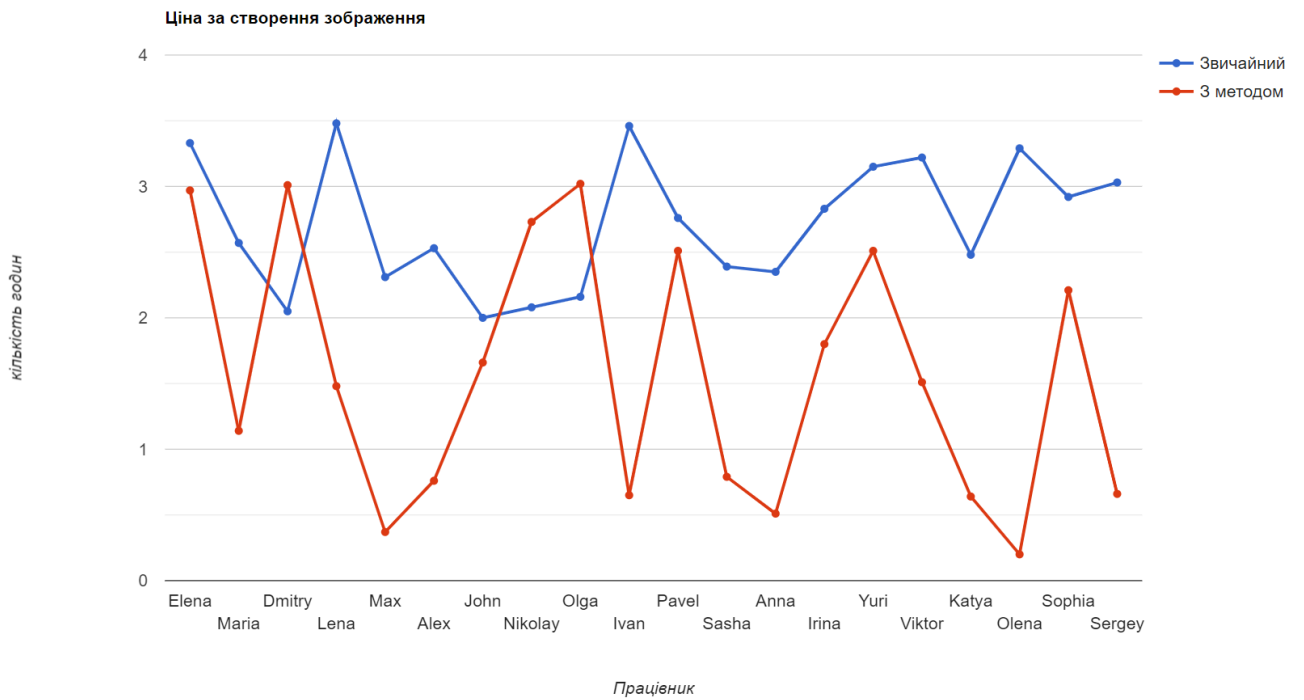


Рис. 3.9. Лінійна діаграма порівняння часу для створення зображення

Підставляємо значення в формулу, виходить:

N – велика кількість згенерованих зображень = 4

S_k - Дуже багато дефектів які треба виправляти, для виявлення мінімального приросту оптимізації = 1.5

C_{ai} –Вартість генерації зображення = 4.4 грн.

T_{ai} - Час генерації зображення в середньому = 0.025 години

Середня ставка спеціаліста:

$$C_s = \frac{\sum_{i=1}^n C_{s_i}}{20} = 292.66 \text{грн/год}, \quad (3.6)$$

Середній час на створення оригінального зображення:

$$T_s = \frac{\sum_{i=1}^n T_{s_i}}{20} = 2.77 \text{год}, \quad (3.7)$$

Середня ціна для створення оригінального зображення:

$$C_o = \frac{\sum_{i=1}^n C_{o_i}}{20} = 819.83 \text{грн}, \quad (3.8)$$

При мінімальному значенні оптимізації, коли треба декілька згенерованих зображень і вони мають багато дефектів, значення оптимізації буде таким:

Час на створення готового зображення з використанням генерації зображень:

$$T_{sai} = (0.025 \times 4) + \left(\frac{2.77}{1.5}\right) = 1.95 \text{години}, \quad (3.9)$$

Ціна за створення готового зображення з використанням генерації:

$$C_{sai} = (1.95 \times 292.66) + (4.44 \times 4) = 588.29 \text{грн}, \quad (3.10)$$

Порівняння ціни та часу для самостійного створення зображення, та з використанням розробленого методу генерації після якого йде обробка зображення людиною, мінімальний показник оптимізації:

Ціна створення зображення зменшується на $= \frac{819.83 - 588.29}{819.83} \times 10 = 28.24\%$

В гривнях ціна зменшилась на $= 819.83 - 588.29 = 231.54$

Час створення зображення зменшується на $= \frac{2.77 - 1.95}{2.77} \times 100 = 29.60\%$

В годинах час зменшується на $= 2.77 - 1.95 = 0.82$

Отже, використання розробленого методу для створення зображень мінімально зменшує вартість створення на **28.24%** та час створення на **29.60%** в середньому.

При максимальному значенні оптимізації, коли з першої спроби зображення вийшло ідеальним і немає дефектів числові показники будуть такими:
Час на створення готового зображення з використанням генерації зображень:

$$T_{sai} = (0.025 \times 1) + \left(\frac{2.77}{5}\right) = 0.579 \text{ години}, \quad (3.11)$$

Ціна за створення готового зображення з використанням генерації:

$$C_{sai} = (0.579 \times 292.66) + (4.44 \times 1) = 173.90 \text{ грн}, \quad (3.12)$$

Порівняння ціни та часу для самостійного створення зображення, та з використанням розробленого методу генерації, максимальний показник оптимізації:

Ціна створення зображення зменшується на $= \frac{819.83 - 173.90}{819.83} \times 10 = 78.80\%$

В гривнях ціна зменшилась на $= 819.83 - 173.90 = 645.93$

Час створення зображення зменшується на $= \frac{2.77 - 0.579}{2.77} \times 100 = 79.09\%$

В годинах час зменшується на $= 2.77 - 0.579 = 2.191$ години

Отже, використання розробленого методу для створення зображень максимально зменшує вартість створення на **78.80%** та час створення на **79.09%** в середньому.

Гістограми які відображають мінімальну та максимальну оптимізації процесу створення готового зображення на рис. 3.10, можемо побачити що середня кількість хв. для створення оригінального зображення це 166хв, при використанні оптимізаційного методу, при максимальній його ефективності потрібно 35хв, при мінімальній ефективності 117хв. Середня кількість грн. для створення зображення 819.83грн, якщо метод максимально ефективний ціна становить 173.90грн, якщо мінімально ефективний то 588.29грн.

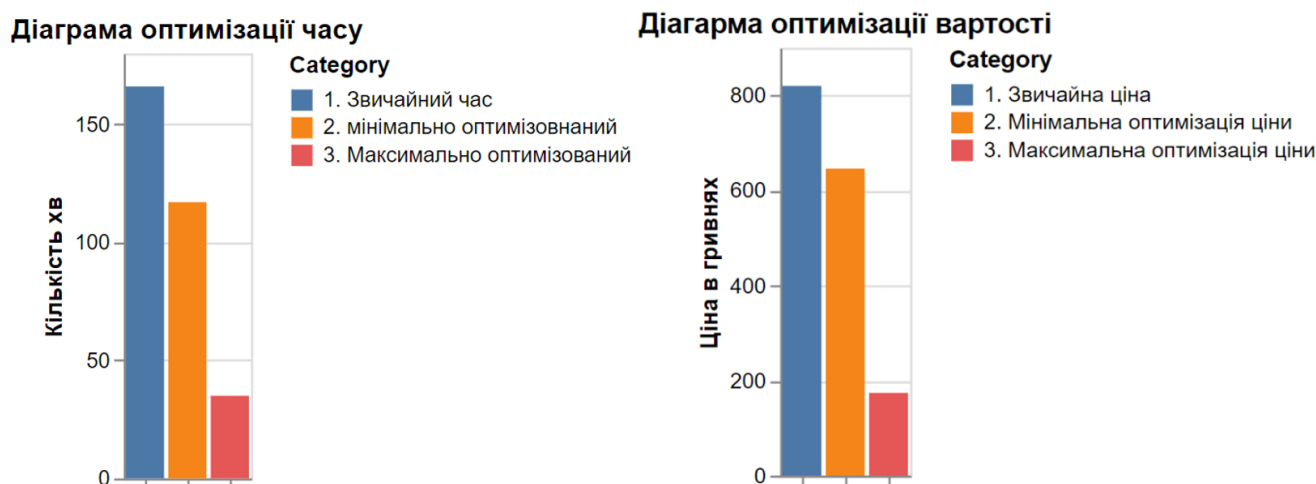


Рис. 3.10. Гістограми оптимізації

Проаналізувавши ці данні можемо побачити, що час для створення зображення використовуючи розроблений метод в середньому, зменшився на 49.2%, а ціна використовуючи розроблений метод теж зменшилась на 50.85%. Таким чином, економічний аналіз зменшення витрат на розробку, враховуючи розроблений метод для генерації зображень, демонструє значний потенціал для підвищення ефективності та зниження витрат в різних областях дизайну та розробки, особливо при створенні мобільних ігор.

3.6. Практичне застосування розробленого методу

Метод був застосований при розробці мобільної гри жанру TCG, або гра з колекційними картами (Trading Card Game). Назва гри “*Neon Card Rush*” логотип розробленої гри можна побачити на рис.3.11. Всі візуальні елементи такі як: картки персонажів, UI елементи, ігрове поле і таке інше, були розроблені використовуючи цей метод.



Рис. 3.11. Логотип гри жанру TCG, з назвою “Neon Card Rush”

Гра жанру TCG - це тип ігор, в яких гравці збирають, обмінюються та грають картами, кожна з яких має свої унікальні властивості та характеристики. Гравці створюють колекції карт і використовують їх для створення стратегій та боротьби проти інших гравців. Цей жанр став дуже популярним в ігровій індустрії завдяки своїй глибокій геймплейній механіці та можливості створювати власні ігрові стратегії.

Схему на якій відображенні головні меню гри можна побачити на рис. 3.12. Гра починається з головного меню, з якого гравці можуть перейти до одного з 3 під меню. Вибір режиму гри – гравці вибирають між двома режимами, гра з своїм відображенням, в якого також колода карт, або гра з босами в кожного з яких є свої колоди і патерни поведінки. Гравці можуть змінити свою колоду, додати нові картки або видалити не потрібні. Також є меню відкриття наборів карток в яких гравці можуть вигравати нові картки та додавати їх до колекції.

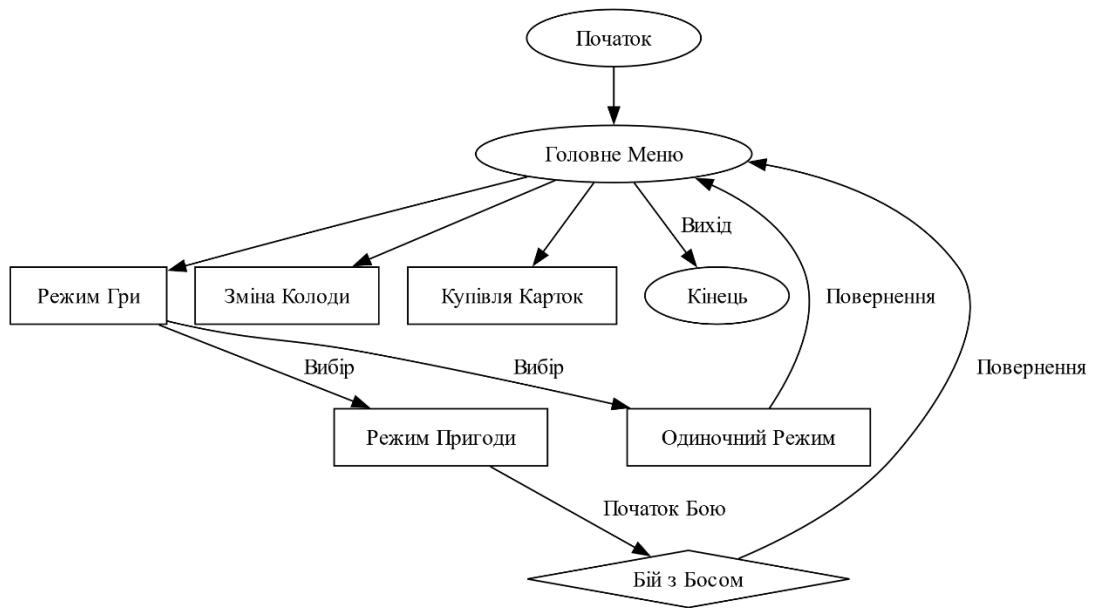


Рис. 3.12. Головні меню гри “Neon Card Rush”

Застосування розробленого методу в грі жанру TCG відкриває безліч можливостей. Головними компонентами таких ігор є ілюстрації карт, приклад таких зображень можна побачити на рис. 3.13 , інтерфейс гри на рис. 3.15, а також інші візуальні елементи, які визначають стиль та атмосферу гри. Завдяки методу, який був розроблений, вдалося значно спростити процес створення цих візуальних елементів. Зображення карт персонажів, фони ігрового поля рис. 3.14., анімації та інші компоненти гри були розроблені з використанням цього методу, що дозволило прискорити процес розробки та забезпечити високу якість графіки, на рис. 3.15 відображено геймплей мобільної гри.



Рис. 3.13. Приклад карток персонажів для “Neon Card Rush”

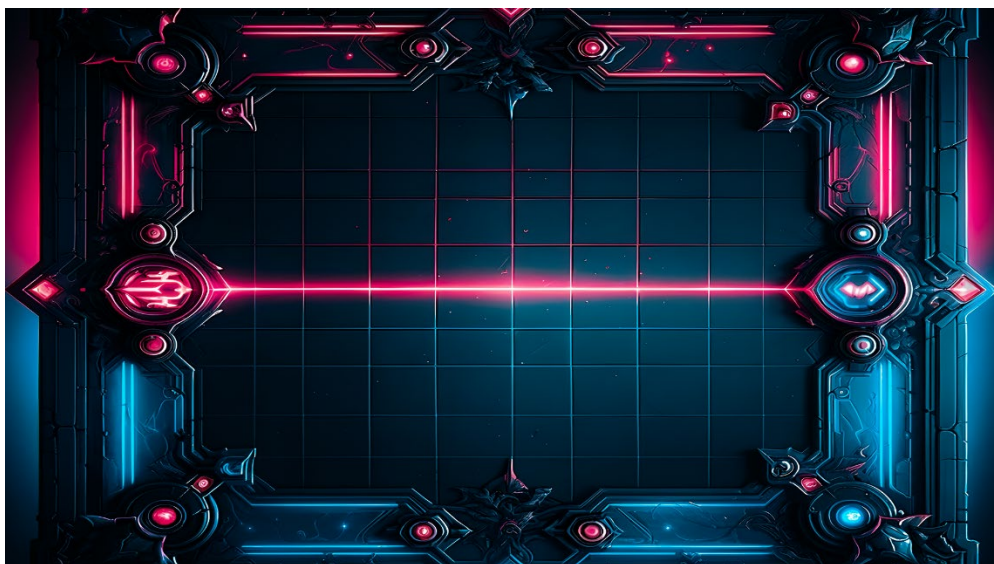


Рис. 3.14. Ігрове поле для гри “Neon Card Rush”



Рис. 3.15. Геймплей гри “Neon Card Rush”

Спеціалісти професії *2 Artist*, створюють оригінальні ілюстрації для дизайну або для мобільних ігор. Для порівняння методів обрано найдешевшого спеціаліста рівня *Junior* з середньою заробітною платою в **47.11 грн** в годину. Такий спеціаліст в середньому створює ілюстрацію **від 2 до 5 год.**

Використовуючи формули виявленні раніше в роботі можна дізнатися :

N – кількість згенерованих зображень.

Sk – кількіість дефектів які треба виправляти.

Cai –вартість генерації зображення = **4.4 грн.**

Tai - час генерації зображення в середньому = **0.025 години**

Cs – 47.11 грн.

Ts - від 2 до 4 год.

При Максимальних витратах на генерацію: $N = 5$, $Sk=1.5$

При Мінімальних витратах на генерацію: $N = 3$, $Sk=3$

Формули для створення зображення по розробленому методу:

$$\text{Час} - Tsai = (Tai \times N) + \left(\frac{Ts}{Sk}\right)$$

$$\text{Ціна} - C sai = (Tsai \times Cs) + (Cai \times N)$$

Формула звичайного методу:

$$\text{Ціна} - Co = Cs * Ts$$

Для створення мобільної гри знадобилось **109 зображень** на створення яких потрібно:

Junior спеціаліст, створення зображення звичайним методом:

Максимальні витрати: 20 540 грн. 436 год = 36 днів.

Мінімальні витрати: 10 270 грн. 218 год = 18 днів

Junior спеціаліст, створення зображення оптимізованим методом:

Максимальні витрати: 12 101.6 грн. 204.92 год = 17 днів.

Мінімальні витрати: 6 959 грн. 117 год = 10 днів

Самостійне створення зображення оптимізованим методом, якщо є базові навички обробки зображень:

Максимальні витрати: 2 848 грн. 449.6 год = 38 днів.

Мінімальні витрати: 1 665 грн. 226.1 год = 19 днів.

Таким чином при використанні розробленого методу спеціаліст рівня Junior в середньому виконав завдання на **38.14% Дешевше**, та на **50% Швидше**.

Самостійне використання методу буде на **76.32%**. Дешевше ніж використання спеціаліста, але на **111.11% Довше**.

3.7. Пропозиції щодо подальшого вдосконалення та застосування методу

У результаті впровадження розробленого методу, ігри жанру TCG стали ще більш привабливими для гравців завдяки вражаючій візуальній стороні та докладній проробці деталей. Використання розробленого методу в ігровій індустрії є лише одним з прикладів його успішного застосування, і цей метод має потенціал для використання в інших галузях, де важливе створення вражаючої візуальної естетики.

Для подальшого вдосконалення методу можна додати нові функції в програму, такі як генерування стиля по другому зображенню, таким чином користувач зможе завантажити одне зображення на якому є потрібний йому об'єкт а на другому потрібний йому стиль. Крім того, важливим аспектом є розвиток

можливостей кастомізації зображень, що надасть дизайнерам більше контролю над кінцевим продуктом та дозволить створювати ще більш унікальні та персоналізовані візуальні елементи.

Що стосується розширення сфер застосування методу, його можна використовувати у таких галузях, як реклама, кіноіндустрія, архітектура, модний дизайн та інші, де візуальна привабливість відіграє ключову роль. Використання цього методу в рекламній галузі може допомогти в створенні унікального та залучаючого контенту, тоді як у кіноіндустрії він може бути використаний для створення реалістичних спецефектів та фонів.

Особливу увагу варто приділити дослідженню взаємодії між людиною та штучним інтелектом у процесі дизайну, щоб максимально використати креативний потенціал людини, поєднуючи його з технічними можливостями ШІ. Такий підхід дозволить створювати продукти, які будуть відображати як технічне досконалість, так і глибокий творчий вираз.

ВИСНОВКИ

1. Для створення оптимізаційного рішення в роботі було здійснено глибокий аналіз існуючих методів та сервісів для генерації зображень. Вибраний сервіс DALL-E 3, який використовує метод Трансформерів, забезпечуючи високу якість та гнучкість для створення візуального контенту.

2. Розроблено метод для автоматизації процесу створення та модифікації графічних елементів. Перетворюючи вхідні зображення в стилізовані версії з урахуванням параметрів, заданих користувачем.

3. Розроблена програма по цій моделі яка значно зменшує час, значно скорочує час, потрібний для розробки візуального контенту, автоматизуючи деякі процеси, які традиційно вимагають зусиль від художників та дизайнерів. Це відкриває можливості для швидкого тестування та реалізації різноманітних візуальних стилів.

4. Розроблено графічний інтерфейс, який спрощує керування процесом вибору, завантаження та обробки зображень.

5. Оптимізаційне рішення було успішно впроваджено в процес розробки мобільної гри жанру TCG, демонструючи його ефективність та практичність на конкретному прикладі. Всі зображення в грі розроблені за допомогою цього методу.

6. Проведений детальний економічний аналіз виявлено значне зменшення витрат та часу на розробку, підкріплюючи математичними даними не тільки переваги використання штучного інтелекту в процесі створення візуального контенту, а й використання саме розробленого методу для цього.

ПЕРЕЛІК ПОСИЛАНЬ

1. Testing Methods for Mobile Game Development : автореф. дис. на здобуття наук. ступеня докт. "2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH)" / . – Vienna, Austria, 2018. – 8 с.
2. State of Game Development & Design [Електронний ресурс] //Perforce. – 2023. – Режим доступу до ресурсу: <https://www.perforce.com/resources/vcs/game-development-report>.
3. Unity Gaming Report [Електронний ресурс] // Unity. – 2023. – Режим доступу до ресурсу: <https://investors.unity.com/news/news-details/2023/Unity-Gaming-Report-2023-Highlights-the-Resilience-of-the-Gaming-Industry/default.aspx>.
4. Best Practices, Pitfalls and Platforms [Електронний ресурс] // MobileApps.com. – 2020. – Режим доступу до ресурсу: <https://www.mobileapps.com/blog/mobile-game-development>.
5. Unity Plug-ins [Електронний ресурс] // Unity Documentation. – 2022. – Режим доступу до ресурсу: <https://docs.unity3d.com/Manual/Plugins.html>.
6. Mbalaka B. Epistemically violent biases in artificial intelligence design: the case of DALLE-2 and Starry AI [Електронний ресурс] / Blessing Mbalaka // Digital Transformation and Society. – 2023. – Режим доступу до ресурсу: <https://www.emerald.com/insight/content/doi/10.1108/DTS-01-2023-0003/full/html>.
7. Popovic D. What Makes Fast Graphic Design & How Long Does It Take?. ManyPixels â Unlimited Graphic Design | Try Now. URL: <https://www.manypixels.co/blog/graphic-design/process> (date of access: 20.11.2023).
8. Games Salary in Ukraine | PayScale. Payscale - Salary Comparison, Salary Survey, Search Wages. URL: <https://www.payscale.com/research/UA/Skill=Games/Salary> (date of access: 20.11.2023)

9. Зарплати українських ІТ-спеціалістів [Електронний ресурс]. – 2023. – Режим доступу до ресурсу: <https://dou.ua/lenta/articles/salary-report-tech-nontech-winter-2023/>.

10. Essential elements of Agile Scrum [Електронний ресурс] - 2023. – Режим доступу до ресурсу: <https://medium.com/agilemania/essential-elements-of-agile-scrum-d5d5d8cafe8d>.

11. Thakur A. Generative Adversarial Networks. International Journal for Research in Applied Science and Engineering Technology. 2021. Vol. 9, no. 8. P. 2307–2325. URL: <https://doi.org/10.22214/ijraset.2021.37723> (date of access: 16.11.2023).

12. Chollet F. Deep Learning with Python, Second Edition. Manning Publications Co. LLC, 2021. 400 p.

13. Matthews J., Fastnedge D., Nairn A. The future of advertising campaigns: The role of AI-generated images in advertising creative. Journal of pervasive media. 2023. Vol. 8, no. 1. P. 29–49. URL: https://doi.org/10.1386/jpm_00003_1 (date of access: 15.11.2023).

14. An Introduction to Transformers. ar5iv. URL: <https://ar5iv.labs.arxiv.org/html/2304.10557> (date of access: 16.11.2023)

15. Diffusion Models: A Comprehensive Survey of Methods and Applications / L. Yang et al. ACM Computing Surveys. 2023. URL: <https://doi.org/10.1145/3626235> (date of access: 16.11.2023).

16. Rayner A. Can Google’s Deep Dream become an art machine?. the Guardian. URL: <https://www.theguardian.com/artanddesign/2016/mar/28/google-deep-dream-art> (date of access: 16.11.2023).

17. DALL·E 3. OpenAI. URL: <https://openai.com/dall-e-3> (date of access: 17.11.2023).

18. Midjourney Quick Start Guide. Midjourney Documentation and User Guide. URL: <https://docs.midjourney.com/docs/quick-start> (date of access: 17.11.2023)

19. GitHub - Stability-AI/stablediffusion: High-Resolution Image Synthesis with Latent Diffusion Models. GitHub. URL: <https://github.com/Stability-AI/stablediffusion> (date of access: 17.11.2023).

20. Огляд Adobe Firefly. Adobe Help Center. URL: <https://helpx.adobe.com/ua/firefly/using/firefly-overview.html> (дата звернення: 17.11.2023).

21. AI Dungeon by Latitude [Електронний ресурс] – Режим доступу до ресурсу: <https://play.aidungeon.com/>.

22. Pinglestudio [Електронний ресурс] – Режим доступу до ресурсу: <https://pinglestudio.com/>.

23. Foster D. Generative Deep Learning: Teaching Machines to Paint, Write, Compose, and Play. O'Reilly Media, Incorporated, 2023.

24. Guinness H. Stable Diffusion vs. DALL·E 3: Which is better? [2024] | Zapier. Zapier | Automation that moves you forward. URL: <https://zapier.com/blog/stable-diffusion-vs-dalle/> (date of access: 16.11.2023).

25. Guinness H. The best AI image generators in 2024 | Zapier. Zapier | Automation that moves you forward. URL: <https://zapier.com/blog/best-ai-image-generator/> (date of access: 16.11.2023).

26. Limano F. Implementation of artificial intelligence based image creation technology for conceptual ideas in 3D visual modeling. 2023 international conference on information management and technology (icimtech), Malang, Indonesia, 24–25 August 2023. 2023. URL: <https://doi.org/10.1109/icimtech59029.2023.10278051> (date of access: 15.11.2023).

27. Ramdurai B., Adhithya P. The impact, advancements and applications of generative AI. International journal of computer science and engineering. 2023. Vol. 10, no. 6. P. 1–8. URL: <https://doi.org/10.14445/23488387/ijcse-v10i6p101> (date of access: 15.12.2023).

28. Testing methods for mobile game development a case study on user feedback in different development phases / E. Annanpera et al. 2018 IEEE 6th international

conference on serious games and applications for health (segah), Vienna, 16–18 May 2018. 2018. URL: <https://doi.org/10.1109/segah.2018.8401372> (date of access: 15.12.2023).

29. Improving the process for mobile games development / H. Inam et al. 2017 International Conference on Communication Technologies (ComTech), Rawalpindi, Pakistan, 19–21 April 2017. 2017. URL: <https://doi.org/10.1109/comtech.2017.8065748>

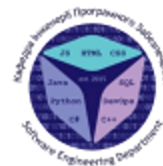
30. Leonardo.Ai [Электронный ресурс] – Режим доступа до ресурсу: <https://app.leonardo.ai/ai-generation>

ДЕМОНСТРАЦІЙНІ МАТЕРІАЛИ

(Презентація)



ДЕРЖАВНИЙ УНІВЕРСИТЕТ ТЕЛЕКОМУНІКАЦІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ
ТЕХНОЛОГІЙ



Кафедра інженерії програмного забезпечення

МАГІСТЕРСЬКА РОБОТА

«ОПТИМІЗАЦІЯ ПРОЦЕСУ РОЗРОБКИ МОБІЛЬНОЇ ГРИ НА ОСНОВІ ЗОБРАЖЕНЬ, СТВОРЕНИХ ЗА ДОПОМОГОЮ ШТУЧНОГО ІНТЕЛЕКТУ»

Виконав: студент групи ПДМ-64, Євтаєв Ярослав Олександрович

Керівник: к.т.н., доц., доцент кафедри ІПЗ Дібрівний Олесь Андрійович

Київ - 2023

МЕТА, ОБ'ЄКТ, ПРЕДМЕТ ДОСЛІДЖЕННЯ

Мета роботи зменшення витрат на розробку мобільної гри, за рахунок автоматизованої генерації візуального контенту.

Об'єкт дослідження: процес розробки мобільної гри.

Предмет дослідження: оптимізація процесу створення графічного контенту мобільної гри.

АНАЛІЗ ІСНУЮЧИХ СИСТЕМ ДЛЯ ГЕНЕРАЦІЇ ЗОБРАЖЕНЬ

| Система | Переваги | Недоліки |
|---------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Adobe Photoshop (Firefly) | Інтеграція створених ШІ зображень у сервіси Adobe та в фотографії задовільна якість зображень , зокрема у відтворенні текстур та кольорів. | Висока вартість підписки (\$20) на місяць, має обмежені можливості персоналізації зображень. |
| Midjourney | Створює унікальні, творчі зображення та має активну спільноту користувачів, яка сприяє обміну знаннями та досвідом. | Вартість підписки(\$10), вимагає часу та експериментування для досягнення бажаного результату, обмеження у комерційному використанні |
| DreamStudio (Stable Diffusion) | Налаштування та контроль генерації, має відкритий доступ, швидка генерація зображень в залежності від потужності системи, не має обмежень політик конфіденційності. Може генерувати зображення з інших зображень. | Може мати проблеми з точністю деталей у складних зображеннях, складний спосіб встановлення та інтерфейс, деякі зображення можуть мати проблеми з авторськими правами. |
| DALL -E 3 | Простота використання, висока здатність до деталізації та різноманіття, здатність аналізувати складні текстові запити перетворюючих на конкретні зображення. | Вартість підписки (\$20), обмеження на кількість генерації зображень в день, не має функції змінювати вже створених зображень. |

3

МАТЕМАТИЧНА МОДЕЛЬ

Ця формула включає всі етапи процесу - від завантаження вхідних зображень та їх опису за допомогою GPT Vision, до генерації нових зображень за допомогою DALL·E 3 та їх подальшої обробки. Таким чином, вона надає інтегроване математичне представлення всього процесу генерації та обробки зображень.

$$Y' = P(\{DALLE3(GPTVision(x_i, S), R) \mid x_i \in X\})$$

Y' - це кінцеве оброблене зображення.

x_i - кожне зображення які змінюються

S – текстовий опис стилю зображення

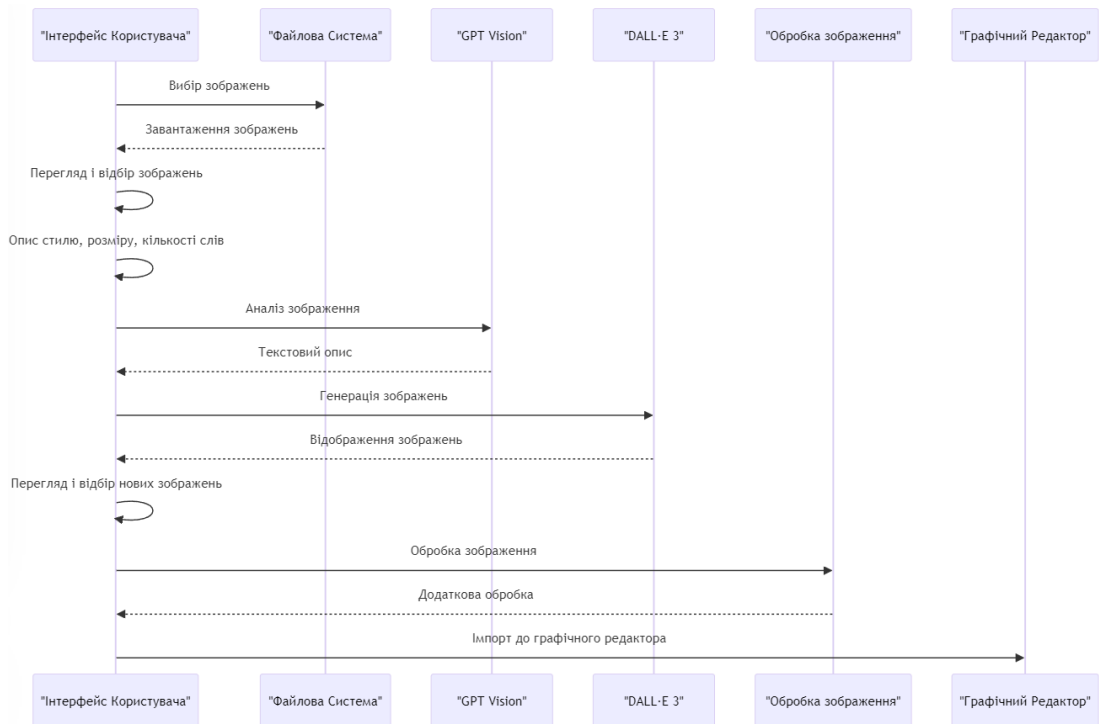
R – розмір зображень

T – кількість слів для опису зображень (токени)

P – метод зміни зображень після генерації, обрізання, зменшення або збільшення роздільної здатності зображення

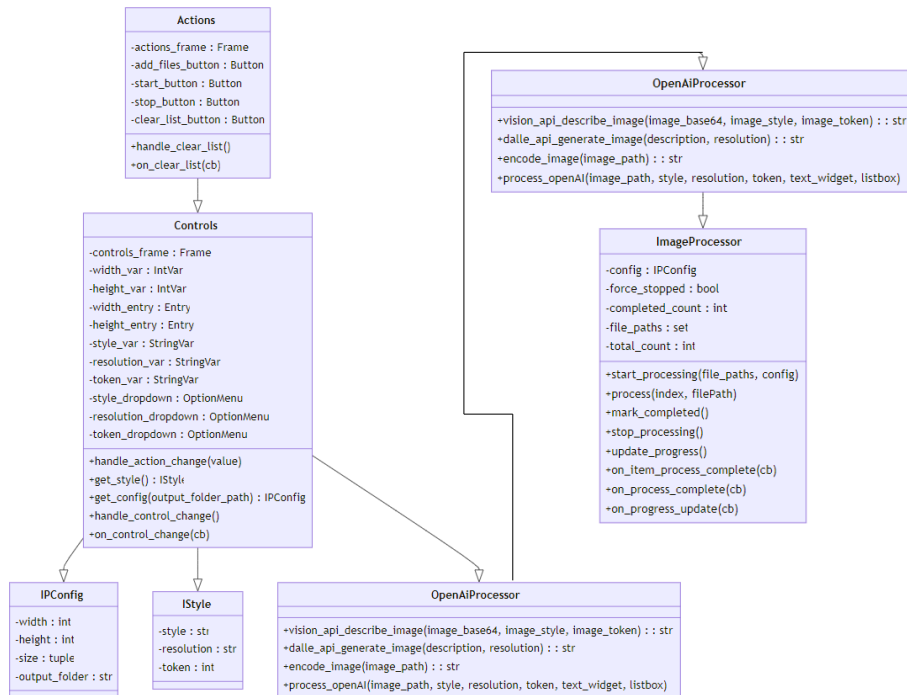
4

ДІАГРАМА ПОСЛІДОВНОСТІ



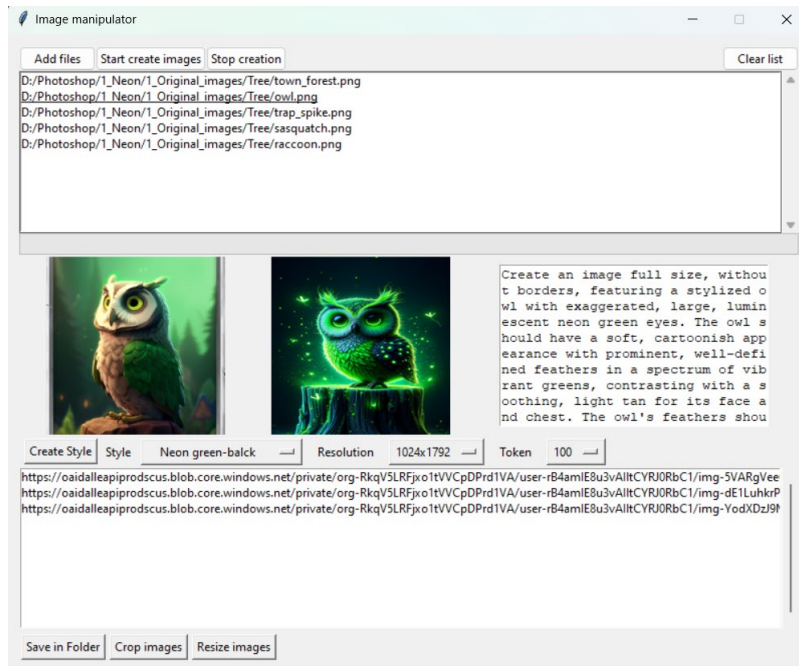
5

ДІАГРАМА КЛАСІВ



6

ІНТЕРФЕЙС ПРОГРАМИ



7

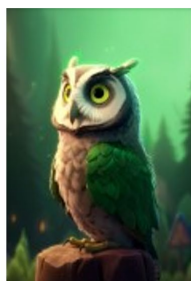
ПРИКЛАДИ ЗГЕНЕРОВАНИХ ЗОБРАЖЕНЬ



Оригінальне



В мало полігональному стилі



Оригінальне

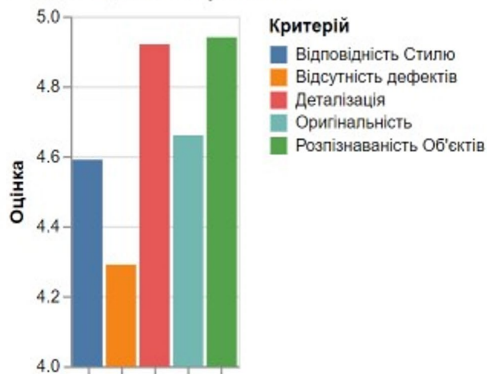


В векторному стилі

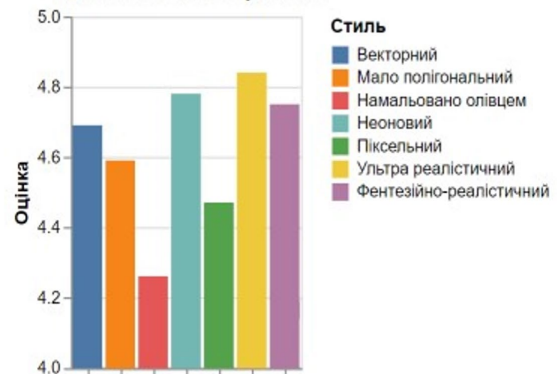
8

ЕКСПЕРТНІ ОЦІНКИ

Шкала Оцінки Зображення



Оцінки Стилів Зображень



9

МАТЕМАТИЧНІ ФОРМУЛИ ВАРТОСТІ ЗОБРАЖЕНЬ

Середня ставка спеціаліста (C_s):

$$C_s = \frac{\sum_{i=1}^n C_{s_i}}{n}$$

Середній час на створення оригінального зображення (T_s):

$$T_s = \frac{\sum_{i=1}^n T_{s_i}}{n}$$

Середня ціна для створення оригінального зображення (C_o):

$$C_o = \frac{\sum_{i=1}^n C_{o_i}}{n}$$

Час на створення готового зображення з використанням генерації (T_{sai}):

$$T_{sai} = (T_{ai} \times N) + \left(\frac{T_s}{S_k} \right)$$

Ціна за створення готового зображення з використанням генерації (C_{sai}):

$$C_{sai} = (T_{sai} \times C_s) + (C_{ai} \times N)$$

N – Кількість згенерованих зображень яка потрібна.

S_k –Технічна якість згенерованого зображення, 5 бальна система.

C_{ai} –Вартість генерації зображення = 4.4 грн.

T_{ai} - Час генерації зображення = від 40 сек., до 2 хв.

10

ЧИСЛОВІ ПОКАЗНИКИ ЕФЕКТИВНОСТІ ВИКОРИСТАННЯ ГЕНЕРАЦІЇ ЗОБРАЖЕНЬ

Junior-ілюстратор заробляє **47.11 грн/год.**

Створення ілюстрації займає **від 2 до 5 годин**

Для мобільної гри потрібно створити **109** зображень, для цього потрібно :

Junior ілюстратор, створення зображення звичайним методом:

Макс. витрати : 20 5 40 грн. 436 год. (36 днів)

Мін. витрати : 10 270 грн. 218 год. (18 днів)

Junior ілюстратор, створення зображення оптимізованим методом

Макс. витрати : 12 10 2 грн. 205 год. (17 днів)

Мін. витрати : 6 95 9 грн. 117 год. (10 днів)

Використання оптимізованого методу, зменшує час та вартість створення зображення мінімально на **28%**, максимально на **79%**.

11

ВИСНОВКИ

1. У роботі здійснено аналіз методів генерації зображень, використано сервіс DALL-E 3 на основі трансформерів, для створення візуального контенту завдяки його високій якості та гнучкості.
2. Розроблено метод автоматизації процесу створення та модифікації графічних елементів, перетворюючи зображення в стилізовані версії з урахуванням користувацьких параметрів.
3. Створена програма яка значно скорочує час, необхідний для розробки візуального контенту, автоматизуючи процеси та відкриваючи можливості для швидкого тестування різних візуальних стилів.
4. Розроблено графічний інтерфейс, що полегшує керування процесом вибору, завантаження та обробки зображень.
5. Оптимізаційне рішення впроваджено в процес розробки мобільної гри, демонструючи ефективність і практичність. Всі зображення в грі створені за допомогою цього методу.
6. Економічний аналіз виявив значне зменшення витрат на розробку, підтверджуючи переваги використання розробленого методу з використанням штучного інтелекту для створення візуального контенту.

12

АПРОБАЦІЯ РЕЗУЛЬТАТІВ ДОСЛІДЖЕННЯ

Стаття: Дібрівний О.А., Євтаєв Я.О. Аналіз сервісів для генерації зображень та їх використання в геймдеві// Зв'язок. Подано до друку.

Тези доповідей:

1. Дібрівний О.А., Євтаєв Я.О. Штучний інтелект в мобільних іграх: порівняльний аналіз інструментів для створення зображень. // Науково-практична конференція «проблеми комп'ютерної інженерії» – Київ : ДУІКТ, 2023. – С. 20.

2. Дібрівний О.А., Євтаєв Я.О. Оптимізація процесу розробки мобільної гри на основі зображень, створених за допомогою штучного інтелекту // Всеукраїнська науково-практична конференція «telecommunication : problems and innovation» - Київ: ДУІКТ, 2023. – С. 22.

ДЯКУЮ ЗА УВАГУ!