

Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
Кафедра Медіасистем та технологій
Рівень вищої освіти перший (бакалаврський)
Спеціальність 186 Видавництво та поліграфія
Тип програми Освітньо-професійна
Освітня програма Видавничо-поліграфічна справа
(шифр і назва)

ЗАТВЕРДЖУЮ:
Зав. кафедри МСТ _____
(підпис)
« 20 » травня 2024 р.

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ**

студентові Столярову Іллі Володимировичу
(прізвище, ім'я, по батькові)

1. Тема роботи Використання генеративного сервісу Stable Diffusion
для розробки дизайну друкованої поліграфічної продукції

Затверджена наказом по університету від 20 травня 2024 р. № 459 Ст

2. Термін подання студентом роботи до екзаменаційної комісії 6 червня 2024 р.

3. Вихідні дані до роботи

Генеративна мережа для створення зображень за текстовим описом Stable Diffusion;
науково-технічні публікації; дані Інтернет та відомих проектів; документація
генеративних сервісів; текстові описи ілюстрацій до настільної гри


4. Перелік питань, що потрібно опрацювати в роботі

Вступ; Постановка завдання; Генеративні змагальні нейронні мережі; GAN моделі text-to-
image; Сучасні генеративні нейронні мережі для створення зображень; Процес створення
зображення у Stable Diffusion; Правила створення підказок для управління створенням
зображень; Фільтрація зображень та взаємодія з дизайнером; Критерії для відбору зображень;
Постобробка зображень; Економічна частина; Висновки.

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п. 5 включається до завдання за рішенням випускової кафедри)

Актуальність роботи; Цілі та задачі; Архітектури та порівняльні характеристики
сучасних генеративних змагальних мереж; Робочі налаштування Stable Diffusion; Сценарії
взаємодії дизайнера з генеративним сервісом; Параметри процесу генерації зображень;
Результати проектування ілюстрацій; Економічна частина; Висновки.

6. Консультанти розділів роботи (п. 6 включається до завдання за наявності консультантів згідно з наказом, зазначеним у п. 1)


Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата
Основна частина	проф. Кулішова Н.Є.		03.06.2024
Економічна частина	ас. Помогалова Н.В.		01.06.2024

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Отримання завдання на кваліфікаційну роботу	20.05.2024	Вик.
2	Аналіз завдання на кваліфікаційну роботу	21.05.2024	Вик.
3	Збір теоретичної інформації про генеративні мережі	22.05.2024	Вик.
4	Генерування зображень за допомогою сервісу	23.05.2024	Вик.
5	Редагування зображень за допомогою сервісу	24.05.2024	Вик.
6	Обробка зображень дизайнером	24.05.2024	Вик.
7	Економічна частина	25.05.2024	Вик.
8	Оформлення пояснювальної записки	28.05.2024	Вик.

Дата видачі завдання 20.05.2024 р.

Студент



(підпис)

Столяров І.В.

Керівник роботи



(підпис)

проф. Кулішова Н.Є.
(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 59 с., 4 табл., 18 рис., 3 дод., 17 джерел.

ДИЗАЙН ІЛЮСТРАЦІЙ НАСТІЛЬНОЇ ІГРИ, ГЕНЕРАТИВНИЙ ШТУЧНИЙ ІНТЕЛЕКТ, ПІДКАЗКИ, ГРАФІЧНИЙ ДИЗАЙН, ГЕНЕРАЦІЯ ЗОБРАЖЕНЬ, STABLE DIFFUSION.

Мета кваліфікаційної роботи – дослідити можливість створення і створити зображення для поліграфічної продукції за допомогою генеративної нейронної мережі «Stable Diffusion». У процесі дослідження визначено актуальність обраної теми, вивчені основні теоретичні положення, що пояснюють дію генеративних нейронних мереж; проведено аналіз можливостей сервісу і вимоги до зображень; досліджено аналоги сервісу, переваги, які вигідно відрізняють його від конкурентів, та недоліки, котрі потрібно враховувати при його застосуванні. Під час виконання роботи було детально розглянуто інтерфейс та процес створення підказок сервісу, досліджено процес сумісної праці з дизайнером під час підготовки зображень, згенерованих сервісом. В результаті дослідження описаний процес генерування зображень для друкованої продукції завдяки генеративному сервісу.

Виконано також економічно обґрунтування проекту застосування генеративного інтелекту під час створення дизайну друкованої продукції.

ABSTRACT

Explanatory note of the qualification work: 59 p., 4 tabl., 18 pic., 3 app., 17 sources.

BOARD GAME ILLUSTRATION DESIGN, GENERATIVE ARTIFICIAL INTELLIGENCE, PROMPTS, GRAPHIC DESIGN, IMAGE GENERATION, STABLE DIFFUSION.

The objective of the qualification work is to investigate the possibility of generating and creating images for printed products using a generative neural network «Stable Diffusion». In the process of research, relevance of chosen topic was determined, main theoretical propositions explaining action of generative neural networks were studied; an analysis of service capabilities and requirements for images was carried out; analogues of the service, advantages that favorably distinguish it from competitors, and disadvantages that must be taken into account when using it are studied. During the execution of the work, interface and process of service prompts creating were examined in detail, process of collaborative work with designer during the preparing of images generated by the service was investigated. As a result of research, the process of generating images for printed products using generative service is described.

An economic justification of the project of applying generative intelligence during creation of printed products design was also carried out.

ЗМІСТ

	С.
ВСТУП.....	7
1 ПОСТАНОВКА ПРОБЛЕМИ.....	8
2 ГЕНЕРАТИВНІ НЕЙРОННІ МЕРЕЖІ	10
2.1 Принцип дії генеративних нейронних мереж	10
2.2 Сучасні генеративні нейромережі для створення зображень	13
3 ПРОЦЕС СТВОРЕННЯ ЗООБРАЖЕННЯ У «STABLE DIFFUSION»	26
3.1 Основні параметри генерування зображень у «Stable Diffusion»	26
3.2 Проектні підходи з використанням «GAI»	38
4 ФІЛЬТРАЦІЯ ЗОБРАЖЕНЬ ТА СПІВПРАЦЯ З ДИЗАЙНЕРОМ.....	43
4.1 Критерії вибору зображень	43
4.2 Постобробка.....	44
5 ЕКОНОМІЧНА ЧАСТИНА	50
5.1 Економічні переваги генеративних нейронних мереж	50
5.2 Джерела економії, доходу та фінансування	51
5.3 Порядок генерування зображення	51
5.4 Розрахунок собівартості та ціни генерування зображень	52
ВИСНОВОК	56
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	58
ДОДАТОК А Фінальний вигляд настільної гри.....	60
ДОДАТОК Б «Qr-code» на «Google Disk»	62
ДОДАТОК В Зображення до і після постобробки	63

ВСТУП

Стрімкий розвиток технологій на основі використання штучного інтелекту, зокрема генерації зображень за текстовими запитами, революціонує всі підходи до дизайну. Це впливає на дизайн ігор [1-3], вебдизайн [4], дизайн одягу [5-6], будівель [1], а також мультимедійні та друковані публікації [8]. Поширення подібних досліджень навіть призвело до виникнення окремої галузі – технологій генеративного штучного інтелекту («Generative AI» – «GAI»). Цей підхід безсумнівно приваблює професіоналів своїм практично необмеженим творчим потенціалом, пропонуючи численні можливості для досліджень та скорочуючи час, необхідний для створення художніх рішень.

Однак, незважаючи на ці очевидні переваги, генеративні сервіси мають значні недоліки: запити до мереж формулюються як тексти обмеженої довжини, що змушує графічних дизайнерів працювати в незнайомій сфері створення стиснених наративів. Для генерації зображення дизайнери повинні встановлювати та змінювати значну кількість параметрів, що описують архітектуру генеративної мережі та її процес роботи. Не будучи експертами ні у письмі, ні в штучному інтелекті, професіонали з інших сфер змушені витратити значний час на організацію процесу дизайну за допомогою генеративних сервісів, оцінку численних створених варіантів та пошук найбільш задовільних.

Іноді задовільні зразки потребують подальшого редагування, оскільки вони не повністю відповідають вимогам дизайнера. Щоб ефективно використовувати переваги генеративних мереж та зменшити негативні аспекти їх використання, необхідно розробити стратегію творчого пошуку.

Ця стратегія повинна складатися з чітко визначених кроків налаштування параметрів мережі, однозначних правил формулювання запитів до генеративного сервісу та структурованої послідовності прийняття рішень під час взаємодії з генеративним сервісом.

1 ПОСТАНОВКА ПРОБЛЕМИ

Сучасні генеративні сервіси для редагування або створення зображень, відео та 3D-моделей реалізують підхід генерації багатовимірного результату на основі одновимірної текстової послідовності, поданої до сервісу у формі опису завдання, відомого як «prompt». «prompt» конструюється відповідно до специфічних вказівок і містить кілька фундаментальних елементів, які сприяють створенню найбільш прийнятних результатів та підвищують їхню якість. Завдяки налаштуванню цих елементів, генерація результату може здійснюватися як ітеративна процедура, що дозволяє постійно покращувати та уточнювати кінцевий результат.

Ці генеративні сервіси використовують передові алгоритми та архітектури нейронних мереж для інтерпретації текстового «prompt» і створення високоякісного візуального контенту. Ітеративний характер процесу дозволяє користувачам налаштовувати згенеровані результати, змінюючи «prompt», експериментуючи з різними параметрами та впроваджуючи зворотний зв'язок для досягнення бажаного результату.

Основна мета цієї роботи – розробити процес прийняття рішень для генерації зображень за допомогою генеративного штучного інтелекту для дизайну картонних ігор. Це передбачає створення структурованого підходу до використання інструментів штучного інтелекту для генерації художніх робіт, які відповідають конкретним естетичним та функціональним вимогам настільних ігор.

Для досягнення цієї мети проект буде зосереджений на кількох ключових напрямках. Вивчення принципів функціонування генеративних нейронних мереж, розуміння того, як працюють ці мережі, включаючи основні алгоритми та методи обробки даних, є важливим для ефективного використання інструментів генеративного штучного інтелекту. Вивчення правил побудови текстових «prompt» та розробка ефективних «prompt», які

точно описують бажаний результат, є необхідними для генерації високоякісних зображень. Це включає розуміння синтаксису, семантики та найкращих практик для створення «prompt». Аналіз вимог до дизайну художніх робіт, визначення конкретних потреб і стандартів для художніх робіт настільних ігор, таких як стиль, колірні схеми та тематичні елементи, забезпечує відповідність згенерованих зображень загальному дизайну гри.

Інтегруючи ці напрями дослідження, проект має на меті створити всеосяжну структуру для використання генеративного штучного інтелекту в дизайні художніх робіт для настільних ігор. Ця структура буде керувати процесом прийняття рішень, від створення «prompt» до ітеративного вдосконалення, що в кінцевому результаті призведе до створення візуально привабливих та контекстуально відповідних художніх робіт для ігор.

2 ГЕНЕРАТИВНІ НЕЙРОННІ МЕРЕЖІ

2.1 Принцип дії генеративних нейронних мереж

Технології генерації контенту розробляються для кількох основних типів завдань: генерація природної мови, генерація зображень, мультимодальна генерація (генерація в мультимодальному машинному навчанні). Основна ознака для їх класифікації: унімодальні моделі отримують вхідні інструкції тієї ж модальності, що й очікуваний результат (приклад – генерація тексту за текстовою інструкцією, генерація зображень за малюнком); мультимодальні моделі, навпаки, можуть обробляти вхідні інструкції в різних модальностях і на виході генерувати результат, чия модальність не збігається з вхідною (приклад – генерація програмного коду або тривимірної моделі на основі текстового опису). Схеми перетворення даних у таких системах показані на рис. 2.1 [9].

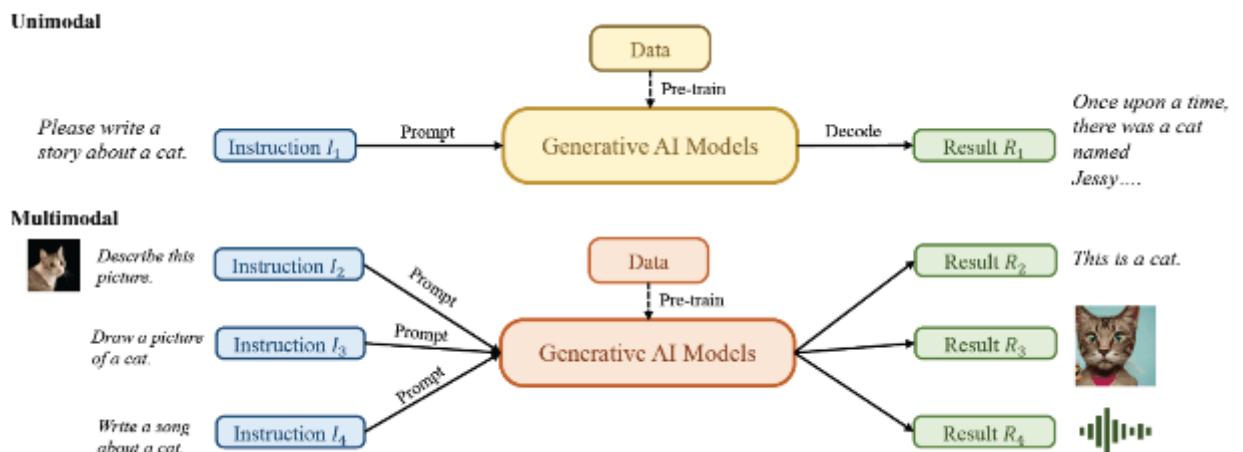


Рисунок 2.1 – Штучні системи інтелекту, які працюють в одній і багатьох модальностях для генерації контенту [9]

Нейронна мережа «Stable Diffusion» базується на мережах генеративної конкуренції («GAN») – методі, опублікованому в 2014 році. Механізм «GAN»

дуже простий: створюються два генератори і дискримінатор. Генератор формує штучні дані. Дискримінатор відрізняє штучні дані, згенеровані генератором, від реальних даних. Дані, згенеровані генератором, намагаються бути якомога ближчими до реальних даних. Це дозволяє створити модель, яка може генерувати дуже точні зображення. Тому цей підхід називають «GAN», оскільки він ставить генератор проти дискримінатора [10].

Далі давайте порівняємо різницю між «GAN» та штучним інтелектом генерації зображень «Stable Diffusion», який набув популярності з 2022 року. Насправді, «Stable Diffusion» базується на моделі, що називається моделлю дифузії, а не «GAN». Іноді це скорочують як «DDPM» («Denoising Diffusion Probabilistic Model»), що буквально перекладається як ймовірнісна модель дифузії з видаленням шуму.

На рисунку 2.2 представлено прогрес у моделях ймовірнісної дифузії. Модель ймовірнісної дифузії (яку ми скоротимо до «моделі дифузії» для кратності) – це параметризований ланцюг Маркова, навчений за допомогою варіаційної інференції для вироблення вибірок, які відповідають даним після закінченого часу. Переходи цього ланцюга вивчаються для зворотного процесу дифузії, який є ланцюгом Маркова, що поступово додає шум до даних у протилежному напрямку від вибірки, доки сигнал не буде знищений. Коли дифузія складається з невеликих порцій гауссівського шуму, достатньо також встановити переходи вибірки ланцюга до умовних гауссівських, що дозволяє особливо просту параметризацію нейронної мережі [10].

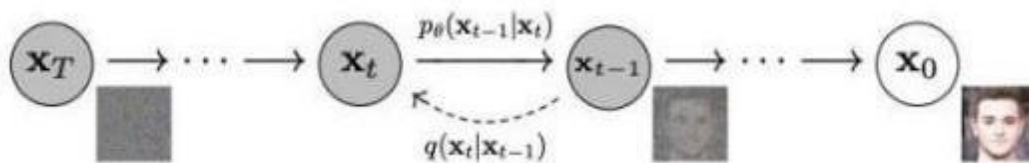


Рисунок 2.2 – «The directed graphical model» розглянуто у цьому розділі [10]

У кінцевому підсумку це техніка навчання для зменшення різниці між зображенням після видалення шуму та оригінальним зображенням. Спочатку

випадковий шум «X0» додається до оригінального зображення, після чого шум видаляється з шумного зображення. Потім зображення після видалення шуму порівнюється з оригінальним зображенням. Наприклад, до зображення людського обличчя додається випадковий шум «X0», і остаточний результат – «XТ», який є повністю шумним.

Ідея моделі дифузії полягає в оптимізації параметрів таким чином, щоб після зворотного процесу додавання цього випадкового шуму шумове зображення було б близьким до оригінального зображення. Вважається, що моделі дифузії перевершують «GAN». Представлені зображення є прикладами «GAN» ліворуч, моделей дифузії посередині та тренувальних даних праворуч (рис. 2.3). Як ми бачимо, моделі дифузії створюють більш різноманітні зображення, ніж «GAN» [11].



Рисунок 2.3 – Порівняння моделей дифузії з генераціями «GAN»

«Stable Diffusion» використовує модель латентної дифузії («LDM») [12], яка поєднує три основні компоненти. (рис. 2.4).

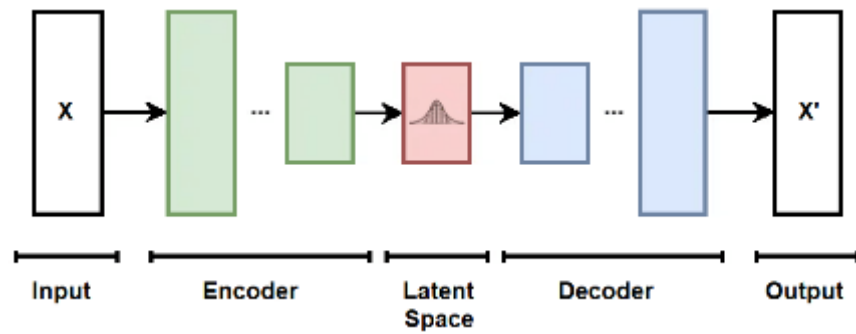


Рисунок 2.4 – Структура моделі «LDM» [12]

Перший компонент, автокодер з варіаційним підходом («VAE»), стискає вхідні зображення до латентного (прихованого) простору, захоплюючи найвиразніші особливості зображень. На цьому етапі також виконується пряма дифузія, де до зображення додається гаусівський шум. Другий компонент, «U-Net», є нейронною мережею, навченою видаляти шум з оригінального сигналу, виконуючи пряму зворотню дифузію на зображенні для виявлення латентного представлення. Третій компонент - декодер «VAE», який генерує кінцеве зображення з пікселів прихованого представлення. Головною перевагою «LDM» є те, що навчання «U-Net» можна налаштовувати відповідно до текстових вхідних даних, асоціюючи накладення шуму з текстовими інструкціями.

Основна відмінність між моделями на основі дифузії та попередніми методами генерації полягає в тому, що моделі на основі дифузії зазвичай навчаються на більших наборах даних зі значно більшою кількістю параметрів, що дозволяє їм вивчати кращі представлення порівняно з іншими методами.

2.2 Сучасні генеративні нейромережі для створення зображень

Порівняння різних генеративних моделей є важливим для просування сфери штучного інтелекту та максимізації потенціалу цих технологій. Кожна архітектура моделі має унікальні переваги й вирішує різні завдання в генерації зображень. Шляхом оцінки їх продуктивності, масштабованості,

обчислювальної ефективності та якості згенерованих зображень. Це порівняння також висвітлює області, де генеративні технології застосовуються найбільш ефективним та впливовим чином.

У цьому розділі буде проведено комплексний аналіз та порівняльний аналіз цих переважаючих генеративних методологій. Ми розглянемо основні принципи, технологічні каркаси, сильні та слабкі сторони різних мереж моделей дифузії. Через це детальне дослідження ми сподіваємося розкрити різноманітність доступних інструментів для генерації зображень та їх наслідки.

«DALL-E»: Розкриття творчості через інтуїтивний дизайн.

Розроблений «OpenAI», «DALL-E» здобув значну увагу завдяки здатності генерувати яскраві, високоякісні зображення за допомогою природних мовних описів. Його основною перевагою є інтуїтивний дизайн, що дозволяє користувачам створювати складні, абстрактні зображення за допомогою простих, схожих на оповідання запитів. «DALL-E» відзначається у тому, що перетворює капризні та уявні ідеї в візуальні реальності, що робить його переважним інструментом для художників та креативних людей, які хочуть надихнутися своїми найфантастичнішими візіями.

Підготовка:

– налаштування середовища: доступ до «DALL-E» зазвичай вимагає використання платформи або «API» «OpenAI», що може включати певні програмні інтеграції або отримання дозволів на доступ;

– розуміння обмежень: усвідомлення сильних та слабких сторін моделі, включаючи її інтерпретацію різних стилів, сюжетів та абстрактних концепцій;

– створення запитів: описи природною мовою: створення запитів за допомогою чіткого, описового мовлення для точного передання бачення дизайнера; художні посилання: включення посилань на мистецькі стилі, історичні періоди або конкретні візуальні елементи для спрямування естетики зображення;

– емоційні та атмосферні деталі: використання мови, що викликає бажаний настрій або атмосферу для покращення візуальної розповіді;

– точне налаштування: ітеративне вдосконалення: цей процес часто включає генерацію зображення, оцінку його відповідності баченню дизайнера та уточнення запиту для наступних ітерацій; регулювання деталей: коригування рівня деталізації або чіткості описів на основі початкових результатів для досягнення більш точної репрезентації;

– генерація та оцінка: генерація зображення: введення запиту в «DALL-E» і ініціація процесу генерації зображення; оцінка: оцінка згенерованого зображення для визначення, чи відповідає воно очікуванням. Якщо необхідно, змінити запит і повторити процес;

– етичні міркування: відповідальне створення: забезпечення того, щоб запити та згенеровані зображення відповідали етичним стандартам і уникали нечутливих або шкідливих зображень; усвідомлення інтелектуальної власності: бути обізнаним щодо питань авторського права, особливо при використанні конкретних творів мистецтва або стилів;

– поширені виклики: балансування креативності з чіткістю: знайдення правильного балансу між уявними описами та чіткими вказівками може бути складним, але є важливим для досягнення ефективних результатів;

– управління очікуваннями: через свою інтерпретаційну природу «DALL-E» може створювати несподівані результати, що вимагає від користувачів адаптувати свої запити або приймати творчі несподіванки;

– основні елементи запиту до «DALL-E»: ефективні запити до «DALL-E» зазвичай включають описову образність: використовуйте яскраву та конкретну мову для опису бажаної сцени або сюжету;

– художні та стилістичні посилання: згадайте конкретні стилі або мистецькі впливи для формування естетики зображення;

– настроїв і емоцій: передайте емоційний тон або атмосферу, яку бажаєте захопити;

– деталі кольору та освітлення: вкажіть кольорові схеми та типи освітлення для покращення настрою та візуального впливу;

– перспектива та кадрування: запропонуйте конкретні кути або елементи композиції для спрямування фокусу моделі;

– відкритість до інтерпретації: залиште простір для творчої інтерпретації «DALL-E» у межах описової рамки Example prompt: «An enchanting forest in autumn, bathed in golden sunlight, with whimsical creatures hidden among the trees, portrayed in a vibrant, impressionistic style» (рис. 2.5).

Цей запит ефективно поєднує описові елементи («enchanting forest», «golden sunlight»), настрій («whimsical»), художній вплив («impressionistic style») та візуальні деталі.



Рисунок 2.5 – Зображення створене «DALL-E». Запит розроблений та складений Робертом Лавінем.

Висновок: «DALL-E» відзначається своєю здатністю розуміти природні мовні запити та створювати зображення, які є уявними і часто інтерпретативними. Важливим аспектом є надання детального і яскравого опису бажаного результату через вербальні інструкції, що дозволяє ШІ заповнити відсутні частини. На відміну від моделей, які вимагають технічного

синтаксису, «DALL-E» розроблений для більш інтуїтивного та зручного користування, адекватно реагуючи на розповідь чи наратив.

«Stable Diffusion»: точність і контроль у генерації зображень. «Stable Diffusion», навпаки, забезпечує високий рівень контролю та специфікації, приваблюючи користувачів, яким потрібна детальна кастомізація зображень. Його розширені синтаксичні опції дозволяють точно маніпулювати елементами зображення, що сприяє більш детальному підходу до створення. Ця модель особливо добре підходить для завдань, що вимагають детальних зображень, технічних ілюстрацій або сценаріїв, де важливе ретельне налаштування елементів.

Підготовка:

- налаштування середовища: для повного використання можливостей «Stable Diffusion» переконайтеся, що налаштоване відповідне середовище, включаючи певне програмне забезпечення, бібліотеки або доступ до хостинг-платформи;

- розуміння обмежень: усвідомлення обмежень моделі, таких як роздільна здатність зображення, складність сюжету та стилістичні обмеження, для ефективного використання;

- створення запитів: позитивні запити: чітко опишіть бажані елементи в зображенні, оскільки специфіка допомагає моделі точно виконувати завдання; негативні запити: вкажіть, що має бути виключено з зображення, щоб уникнути небажаних елементів;

- поєднання елементів: використовуйте природну мову для змішування різних елементів, застосовуючи синтаксис типу «AND» або «BREAK» для структурованих запитів;

- точне налаштування: ітеративний підхід: генерація зображення часто потребує методу проб і помилок. Згенеруйте зображення, оцініть його та вдоскональте запит для кращих результатів;

- деталізація: додайте додаткові деталі або скоригуйте фразування у запиті для покращення результатів;

– генерація та оцінка: генерація зображення: запустіть модель із створеним запитом;

– оцінка: перевірте результат і, за необхідності, змініть запит для кращого відповідності очікуванням дизайнера;

– етичні міркування: обмеження контенту: дотримуйтеся етичних норм і обмежень щодо контенту, уникаючи створення образливих або шкідливих зображень;

– інтелектуальна власність: поважайте права інтелектуальної власності та уникайте створення зображень, які можуть порушувати авторські права;

– поширені виклики: нечіткість у запитах: неоднозначні запити можуть призвести до несподіваних результатів;

– надмірна деталізація: надто детальні запити можуть заплутати модель, що призведе до створення надто складних зображень;

– інтерпретація обмежень моделі: розуміння того, що модель може і не може робити, є важливим для ефективного застосування.

Основні елементи запиту для «Stable Diffusion». Для ефективного використання «Stable Diffusion» врахуйте ці основні елементи при створенні запиту:

– сюжет: визначте основний фокус бажаного зображення, наприклад, пейзаж, людину, об'єкт або абстрактну концепцію;

– описові деталі: використовуйте описові прикметники та прислівники, як-от «fluffy, orange cat with bright green eyes» замість просто «cat»;

– налаштування і контекст: опишіть налаштування або оточення, включаючи внутрішні/зовнішні простори, конкретні місця та відповідні час або погодні умови;

– настрої або атмосфера: виразіть бажаний емоційний тон зображення за допомогою слів, що передають настрої;

– кольорова палітра та освітлення: вкажіть важливі кольори та опишіть тип освітлення для створення атмосфери;

- стиль та художній вплив: згадайте будь-який конкретний художній стиль або вплив у вашому Prompt, враховуючи питання авторських прав;
- перспектива та композиція: зазначте, чи є певна перспектива або композиційний елемент, який хочете підкреслити;
- «Negative Prompt» (опційно): чітко вкажіть елементи, які бажаєте виключити зі свого зображення.

Приклад Запиту («Prompt»):

- «Positive Prompt»: «Serene, sunlit forest with tall pine trees and a gentle stream, covered in early morning mist. The atmosphere is peaceful and untouched, with soft, warm lighting»;
- «Negative Prompt»: «No animals or people in the scene».

Цей «prompt» охоплює сюжет («forest»), описові деталі («tall pine trees, gentle stream»), налаштування («early morning, misty»), настроїв («peaceful, untouched»), освітлення («soft, warm») та негативний елемент («no animals or people») (рис. 2.6).

Ось детальний аналіз деяких ключових варіантів синтаксису, які можуть бути використані з «Stable Diffusion». Позитивні та негативні ключові слова:

- позитивні ключові слова: безпосередньо описують те, що дизайнер хоче включити в зображення;
- негативні ключові слова: вказують, що має бути виключено, часто з префіксом мінус (наприклад, «-no people »).

Логіка «AND / OR»:

- «AND»: використовується для забезпечення того, щоб всі перелічені елементи були включені в зображення (наприклад, «cats AND dogs» гарантуватиме, що на зображенні будуть як коти, так і собаки);
- «OR»: пропонує альтернативи, надаючи моделі вибір («cats OR dogs»);
- ключові фрази з вагою. Синтаксис: («Phrase: Weight») – дозволяє призначати різну вагу різним елементам «prompt», впливаючи на їхню виразність у згенерованому зображенні. Приклад: «cats: 0.7, dogs: 0.3» надасть більше значення котам, ніж собакам у зображенні;



Рисунок 2.6 – Зображення створене «Stable Diffusion».
«Prompt» розроблений та складений Робертом Лавінем.

– послідовні або фазові елементи. Синтаксис: («Phrase: Phrase: Step») вказує на перехід від одного елемента до іншого на певному етапі процесу генерації. Приклад: («cityscape: nature: 0.5») може починатися з міського пейзажу і переходити до природи на півдорозі процесу;

– дужки для акцентування. Використання дужок може підвищити ймовірність включення фраз (наприклад, («sunset, ocean»)) або підкреслити певні елементи чи теми в «prompt»;

– квадратні дужки для зменшення акценту. Квадратні дужки можуть використовуватися для зменшення акценту (наприклад, [crowded, streets]);

– включення елементів, але з меншим значенням;

– стиль і художній вплив. Включення назви стилю мистецтва, епохи або художника для впливу на естетику зображення. Приклад: «Van Gogh style landscape»;

– специфічні модифікатори. Слова, такі як «blurry, detailed, zoomed in, wide angle», можуть спрямовувати фокус, чіткість і перспективу зображення;

– поєднання концепцій. Творче поєднання різних концепцій (наприклад, «steampunk Eiffel Tower»).

Синтаксичні опції, запропоновані «Stable Diffusion», пропонують надійний набір інструментів для вдосконалення «prompts». Ефективність цих «prompts» залежить від досягнення балансу між специфікою, креативністю та розумінням інтерпретаційних можливостей моделі. Продуктивність моделі може змінюватися в залежності від складності запиту та точності формулювання «prompt».

«Midjourney»: художня інтерпретація та злиття стилів. «Midjourney» вирізняється своїми художніми інтерпретаціями та здатністю унікально поєднувати стилі та концепції. Відомий створенням абстрактних або стилістично виразних зображень, ця модель демонструє творчий потенціал ШІ. Вона часто дивує користувачів своїми інтерпретаціями, надаючи результати, які не лише візуально приголомшливі, але й багаті художнім виразом.

Підготовка:

– налаштування середовища: використання «Midjourney» передбачає доступ до його призначеної платформи або інтерфейсу, що потенційно вимагає спеціального програмного забезпечення або членства в певних онлайн-спільнотах;

– розуміння обмежень: розуміння творчих меж та стилістичних нюансів «Midjourney», включаючи його унікальний підхід до художніх стилів та абстрактного рендерингу, є важливим для ефективного використання.

Створення «prompts»:

– описова мова: використовуйте яскраві та уявні описи, щоб направити модель до вашої уявної сцени;

– художні стилі та впливи: інтегруйте посилання на мистецькі стилі, історичні періоди або художні техніки для формування візуального стилю;

– емоційні та атмосферні ключові слова: додавайте слова, що передають настрій або атмосферу, щоб поглибити емоційний вплив зображення.

Точне налаштування:

– експерименти: «Midjourney» часто вимагає експериментів, оскільки модель може інтерпретувати «prompts» у несподівані, але творчо винагороджуючі способи;

– коригування: вдосконалюйте «prompts» на основі початкових результатів, змінюючи або додаючи описові елементи та художні посилання.

Генерація та оцінка:

– генерація зображення: введіть «prompt» і дозвольте «Midjourney» згенерувати зображення;

– оцінка: оцініть результат, щоб визначити його відповідність баченню дизайнера, за необхідності коригуйте «prompt» для наступних ітерацій.

Етичні міркування:

– чутливість контенту: будьте обережні при створенні зображень, що дотримуються поваги та етичних стандартів;

– інтелектуальна власність та оригінальність: переконайтеся, що «prompts» відповідають оригінальним творам та законам про авторські права, навіть коли «Midjourney» генерує унікальні інтерпретації.

Поширені виклики:

– абстрактні інтерпретації: «Midjourney» може створювати зображення більш абстрактно, ніж очікувалося, що вимагає балансу між специфікою та відкритістю до творчих інтерпретацій;

– балансування деталей та креативності: знаходження правильного балансу між детальними описами та наданням художньої свободи може бути складним, але призводить до винагороджуючи результатів.

Основні елементи «Midjourney Prompt»: ефективні «prompts» для «Midjourney» повинні включати:

– яскраві описи: використовуйте багату, виразну мову для яскравого зображення сцени або концепції;

– художній вплив: посилайтеся на конкретні мистецькі стилі або художників для спрямування естетичного напрямку;

– настрої і тон: використовуйте слова, що встановлюють бажаний емоційний або атмосферний тон;

– колір та освітлення: вкажіть конкретні кольори або освітлювальні ефекти для покращення настрою зображення;

– перспектива та композиція: запропонуйте точки зору або композиційні елементи для впливу на інтерпретацію моделі;

– творче поєднання: інтегруйте різні елементи або концепції для натхнення на створення унікальних зображень.

Приклад «prompt»: «A bustling Renaissance marketplace at dawn, filled with colorful stalls, lively merchants, and a backdrop of majestic mountains, painted in the style of Botticelli» (рис. 2.7).



Рисунок 2.7 – Зображення створене «Midjourney». «Prompt» розроблений та складений Робертом Лавінем

Цей «prompt» інтегрує яскраві описи («Renaissance marketplace», «colorful stalls»), художній вплив («Botticelli»), настрої («lively, majestic») та унікальні композиційні елементи.

Ключові підходи та стратегії для створення «prompts» у «Midjourney»:

- специфічні модифікатори для деталей та фокусу: використовуйте терміни на кшталт «zoomed-in», «wide-angle», або «vibrant colors», щоб спрямувати фокус та стиль зображення. Приклад: «Close-up of a raindrop on a leaf, with vivid colors and intricate detail»;

- поєднання елементів: творче: об'єднуйте різні концепції або елементи унікальним чином. Приклад: «A futuristic city with Gothic architecture, under a starlit sky»;

- підказуючі фрази для включення та виключення: використовуйте позитивні формулювання для підкреслення бажаних елементів. Вказівка на те, що потрібно уникати, іноді може бути ефективною, навіть якщо це не так прямо, як з іншими моделями. Приклад: «A bustling medieval marketplace, full of life and color, without any modern elements»;

- колір та освітлення: вкажіть кольори та типи освітлення для створення настрою та візуального впливу. Приклад: «A cityscape at twilight with pastel colors and soft, diffused lighting»;

- перспектива та композиція: опишіть точку зору або композиційні елементи. Приклад: «Bird's-eye view of a winding river through a lush valley»;

У «Midjourney» акцент робиться на описову та уявну мову, а не на технічний синтаксис. «Prompts» повинні яскраво описувати бажаний результат, використовуючи виразну мову, стилістичні посилення та конкретні дескриптори для спрямування ШІ.

Широка інтерпретація моделі означає, що багатство і специфічність мови можуть суттєво вплинути на результат.

Також буде корисно порівняти розглянуті генеративні сервіси за іншими показниками, зокрема, за ціною, основною технологією, складністю використання, можливістю автономного використання, можливостями редагування зображень у рамках сервісу, форматом і роздільною здатністю створених ілюстрацій тощо. Результати порівняння наведено в таблиці 2.1.

Таблиця 2.1 – Порівняльні характеристики генеративних сервісів: «MidJourney», «Stable Diffusion», «DALLE-3»

Характеристика	«MidJourney»	«Stable Diffusion»	«DALLE-3»
Ціна	10-120\$/місяць	Безкоштовно («Open Source»)	20\$/місяць
Навігація	Простий інтерфейс з широкими можливостями	Складний інтерфейс з широкими можливостями та численними додатковими функціями	Простий інтерфейс з обмеженим функціоналом
Можливість використання локально	Ні	Так («Open source»)	Ні
Базується на	«Generative adversarial network»	«Latent diffusion model»	«Generative pre-trained transformer»
Складність використання	Легкий та інтуїтивний, але зі специфічними характеристиками використання	Інтуїтивно зрозумілий, але потребує значного часу на освоєння	Легкий та інтуїтивний у використанні
Стиль «prompt»	«prompts» зав'язанні на ключових словах «prompts»	«prompts» зав'язанні на ключових словах з додатковими опціями уточнення	Вільноформатні «prompts»
Найкраще застосування	Реалістичні зображення	Більш точні результати, але вимагають більше зусиль	Загальні, легкі «prompts»
Редагування зображень	Можливість редагування ділянок згенерованих зображень	Можливість редагування як завантажених, так і згенерованих зображень та їх ділянок	Немає можливості редагування
Формат та роздільна здатність	Обмежені формати та фіксована роздільна здатність	Необмежений вибір форматів та роздільної здатності	Один формат з фіксованою роздільною здатністю

Висновок. Сервіс «Stable Diffusion» був обраний завдяки своїм універсальним функціям та широкій доступності через розповсюдження з відкритим вихідним кодом. Під час генерації зображень можуть виникати проблеми з авторськими правами, але спільнота з відкритим вихідним кодом пропонує різні моделі з різними дозволами на авторські права. Ці фактори допомагають знизити витрати на друк і забезпечують відповідність зображень різним форматам та роздільним здатностям. Використання платформ з відкритим вихідним кодом є вирішальним для вирішення питань авторських прав і забезпечення високої якості результатів.

3 ПРОЦЕС СТВОРЕННЯ ЗООБРАЖЕННЯ У «STABLE DIFFUSION»

3.1 Основні параметри генерування зображень у «Stable Diffusion»

Створення зображень за допомогою нейронних мереж відбувається в різних режимах, таких як створення зображень з текстового опису («txt2img»), трансформація одного зображення в інше («img2img») та відновлення зображень («Inpaint»).

Генерація зображень за допомогою нейронної мережі «Stable Diffusion» має різноманітні налаштування. Деякі з них встановлюються один раз для групи зображень, тоді як інші потребують налаштування для кожного окремого зображення, залежно від очікуваного результату. Ці параметри використовуються для генерації, трансформації та відновлення зображень за допомогою нейронної мережі «Stable Diffusion» [3].

Розглянемо доступні налаштування для створення («txt2img»), трансформації («img2img») та відновлення («Inpaint») зображень. (рис. 3.1-3.5):

– «Stable Diffusion checkpoint» (модель) (1) – з випадального списку обирається модель «Stable Diffusion», яка буде використовуватись для генерації або обробки зображень (рис. 3.1, 3.2, 3.4);

– вкладки режимів та налаштувань (2) – за допомогою вкладок можна перемикаєти режими роботи або налаштування. Вкладки призначені для наступного: «txt2img» – генерація зображень за допомогою нейронної мережі Stable Diffusion; «img2img» – перетворення зображення в зображення, наприклад, перетворення фото на малюнок або картину, обробка фото та інше; «extras» – збільшення зображень за допомогою «upscaler»; «PNG info» – на цій вкладці можна подивитись інформацію про згенероване зображення, яка зберігається у самому зображенні в форматі png (збереження цієї інформації вмикається у налаштуваннях); «checkpoint merger» – об'єднання (злиття) кількох моделей в одну; «train» – тренування моделей (поки не працює у

Windows на відеокартах AMD); «settings» – налаштування програми; «extensions» – керування модулями та розширеннями;

– «Stable Diffusion prompt» (3) – текстовий запит до нейронної мережі по якому буде генеруватись зображення (рис. 3.1, 3.2, 3.4);

– «Stable Diffusion negative prompt» (4) – негативний запит до нейронної мережі, який виключає вказані слова з генерації (рис. 3.1, 3.2, 3.4);

– перетворення зображення в текст («Interrogate») (5) – доступно для «img2img», «Sketch», «Inpaint», «Inpaint sketch», «Inpaint upload». Кнопки «Interrogate CLIP» та «Interrogate DeepBooru» відповідають за генерацію текстового опису зображення. Нейронна мережа «Stable Diffusion» може за допомогою алгоритмів «CLIP» та «DeepBooru» проаналізувати зображення та описати те, що вона бачить на зображенні (рис. 3.2, 3.4);

– «Кнопка Generate» (генерувати) (6) – кнопка «Generate» запускає процес генерації зображення. Під час генерації кнопка «Generate» замінюється двома іншими кнопками: «interrupt» – перервати генерацію; «skip» – пропустити поточне зображення і перейти до наступного (використовується, якщо увімкнена пакетна генерація/обробка «Batch» (27) (рис. 3.1, 3.2, 3.4);

– «Stable Diffusion Styles» (7) – це збережені параметри генерації, в ці параметри входить запит, негативний запит, «seed», тощо. Можна зберегти введені параметри, а потім їх використовувати (рис. 3.1, 3.2, 3.4);

– вкладки керування додатковими модулями: «generation» (8) – вкладка генерації зображення; «textual Inversion» – вкладка додавання встановлених «Embeddings» до запиту; «hypernetworks» – вкладка додавання встановлених «hypernetworks» до запиту; «checkpoints» – вкладка вибору встановлених моделей (те саме, що і «stable diffusion checkpoint» (1)); «lora» – вкладка додавання встановлених моделей lora до запиту (рис. 3.1, 3.4);

– перемикання режимів «img2img» (доступно для «img2img», «Sketch», «Inpaint», «Inpaint sketch», «Inpaint upload») (9) – перемикає різні режими перетворення зображення в зображення: «img2img» – режим перетворення зображення в зображення; «sketch» – режим створення зображення з ескізу;

«inpaint» – режим перемальовування частини зображення за допомогою маски та текстового запиту. Дозволяє вручну замалювати частину зображення (нанести маску), яка буде перемальовуватись нейронною мережею; «inpaint sketch» – режим перемальовування частини зображення шляхом малювання кольорового скетчу в місцях перемальовування. Тобто, потрібно намалювати грубий кольоровий малюнок по фотографії, а нейронна мережа перетворить намальоване та текст запиту на нормальне зображення і гармонійно впише його у фотографію; «inpaint upload» – те саме, що і «Inpaint», але маска не малюється на зображенні, а завантажується у вигляді файлу. Використовується, коли маска для перемальовування була створена в іншій програмі або розширенні; «batch» – пакетне перетворення зображення в зображення (рис. 3.1, 3.4);

– вхідне (еталонне) зображення (доступно для «img2img» , «Sketch», «Inpaint», «Inpaint sketch», «Inpaint upload») (10) – можна завантажити зображення, яке буде аналізуватись нейронною мережею «Stable Diffusion» та використовуватись для генерації або перемальовування (рис. 3.1, 3.4);

– «image mask» (Доступно лише для «Inpaint upload») (11) – можна завантажити зображення-маску, яка буде використовуватись для перемальовування частини зображення (рис. 3.4);

– згенероване нейронною мережею «Stable Diffusion» зображення (12) – відображає останні згенеровані зображення. Зображення можна збільшувати. Також, згенеровані, оброблені або збільшені зображення за замовчуванням зберігаються в папці «outputs», яка знаходиться в папці з програмою «AUTOMATIC1111/stable-diffusion-webui», (шлях збереження можна змінити в налаштуваннях, на вкладці «Settings» (2)) – «Paths for saving»): згенеровані зображення – в папку «txt2img-image», при пакетній генерації («Batch») – сітки зображень зберігаються в «txt2img-grids»; оброблені в модулі «img2img» зображення – в папку «img2img-images», сітки в «img2img-grids»; Збільшені через вкладку «Extras» зображення – в папку «extras-images» (рис. 3.1,3.2, 3.4);

– керування згенерованим зображенням (13) – під зображенням є кнопки, за допомогою яких можна: відкрити папку збереження зображень; зберегти зображення в іншу папку; заархівувати зображення; передати зображення до перетворення зображення в зображення «img2img» (9)); передати зображення в перемальовування зображень (вкладка «Inpaint» (9)); передати зображення у збільшення зображень (вкладка «Extras» (2)) (рис. 3.1,3.2, 3.4);

– «Copy image to» (14) – доступно для «img2img», «Sketch, Inpaint», Inpaint sketch. Кнопки передачі зображення інші вкладки модуля «img2img» (9): «img2img» – до вкладки «img2img»; «sketch» – до вкладки «sketch»; «inpaint» – до вкладки «Inpaint»; «inpaint sketch» – до вкладки «Inpaint sketch» (рис. 3.2, 3.4);

– «Resize mode» (доступно для «img2img», «Sketch», «Inpaint», «Inpaint sketch», «Inpaint upload») (15) – Режим зміни розміру зображення під час генерації: «just resize» – збільшити без збереження пропорцій; «crop and resize» – обрізати згідно пропорцій та збільшити; «resize and fill» – збільшити і доповнити, зберігши пропорції; «just resize» («latent upscale») – збільшити без збереження пропорцій допомогою «latent upscale» (рис. 3.3, 3.5);

– «Mask blur» (доступно для Inpaint, Inpaint sketch, Inpaint upload) (16) – Mask blur визначає скільки пікселів навколо перемальованого зображення буде розмито для поєднання результату перемальовування з зображенням (рис. 3.5);

– «Mask transparency» (доступно лише для «Inpaint sketch») (17) – прозорість намальованої маски у процентах, впливає на те, як буде зміщуватись намальована маска з зображенням під нею під час генерації (рис. 3.5);

– «Mask mode» (доступно для «Inpaint», «Inpaint sketch», «Inpaint upload») (18) – режим перемальовування по масці, можливі значення: «inpaint masked» – перемальовувати те, що під маскою; «inpaint not masked» – перемальовувати те, що навколо маски (рис. 3.5);

– «Masked content» доступно для «Inpaint», «Inpaint sketch», «Inpaint upload» (19) – вміст під маскою, з якого буде починатись перемальовування,

можливі варіанти: «fill» – зображення під маскою буде сильно розмито та використано для генерації (використовується для збереження кольорової палітри); «original» – оригінальне зображення (використовується для незначних змін); «latent noise» – область під маскою буде заповнено випадковим шумом на основі «Stable Diffusion Seed» (30) (використовується для повної зміни області під маскою без прив'язки до композиції і кольорової палітри); Latent nothing – область під маскою буде заповнено одним кольором, який є сумішшю всіх кольорів (використовується для часткової зміни області під маскою без прив'язки до композиції, але зі збереженням кольорової палітри) (рис. 3.5);

– «Inpaint area» (доступно для «Inpaint», «Inpaint sketch», «Inpaint upload») (20) – що нейронна мережа буде використовувати для аналізу під час перемальовування: «whole picture» – все зображення, результат перемальовування буде набагато краще і гармонійніше вписаний в зображення; «only masked» – тільки зображення під маскою, краще підходить для невеликого перемальовування, наприклад, для покращення деталей. (рис. 3.5);

– «Only masked padding, pixels» (доступно для «Inpaint», «Inpaint sketch», «Inpaint upload») (21) – визначає скільки додаткових пікселів буде використовуватись для аналізу зображення, працює тільки для режиму «Inpaint area» (20) – «Only masked» (рис. 3.5);

– «Stable Diffusion Sampler» (22) – «Sampling method» – це метод розбиття на кроки «Sampling steps» (23) генерації зображень. Розбиття на кроки процесу генерації зображень є частиною алгоритму генерації «Stable Diffusion». Він впливає на «випадковість» результату та кількість кроків «Sampling steps» (23), які потрібно зробити для генерації деталізованого зображення. Різні семплери, при тих же параметрах генерації, можуть генерувати зображення які трохи відрізняються одне від одного. Кожен семплер працює по різному, має різну швидкість та їм потрібна різна мінімальна кількість кроків для генерації якісного зображення. Поглянемо ближче на деякі з семплерів: «Stable Diffusion LMS» – працює доволі швидко,

дає хороші результати при 40-50 кроках, але потрібно зробити правильний запит. Для покращення результату можна збільшити кількість кроків до 70-80; «Stable Diffusion DDIM» – «DDIM» по роботі схожий на «LMS», але він робить більш грубі зображення з меншою кількістю деталей. Він може генерувати зображення з доволі малою кількістю кроків 10-20, на ньому можна тестувати запити; «Stable Diffusion Euler» – працює трохи повільніше, гарні зображення виходять на 30-50 кроках; «Stable Diffusion Euler a» – працює доволі швидко, може давати результати при малій кількості кроків (8-15). Показує гарні результати при 20-40 кроках. Ідеально підходить для тестів. Часто використовується для генерації аніме та зображень з мультиків; «Stable Diffusion DPM++ 2M» («Karras») – це сімейство семплерів, вони схожі на «Euler», генерують якісні зображення при великій кількості кроків 40-80 (рис. 3.1, 3.3, 3.5);

– «Sampling steps» – кількість кроків на які семплер «Stable Diffusion Sampler» (22) розбиває процес генерації. Від кількості кроків залежить кількість деталей на зображенні. При занадто малій кількості кроків зображення може бути з шумом або артефактами. А при занадто великій – можуть з'являтися непотрібні деталі. Кількість кроків потрібно обирати виходячи з обраного семплера та потрібної кількості деталей методом спроб.

– «Hires. Fix» (24) – доступно лише для «txt2img». «Hires. Fix» – це функція збільшення зображення під час генерації, її використання, параметри та приклади описані на сторінці «Збільшення зображення під час генерації – «Hires. Fix». (рис. 3.1);

– «Refiner» – дозволяє генерувати зображення двома моделями. У випадяючому списку «Checkpoint» можна обрати другу модель, а в параметрі «Switch at» вказати з якого кроку перемкнутись на другу модель. «Switch at» задається у % від загальної кількості кроків «Sampling steps» (23), 0.35 це після 35% кроків, а 0.7 – після 70% (рис. 3.1, 3.3, 3.5);

– розмір зображення (26) – можна задати розмір генерованого зображення у пікселях. «Width» – довжина, «Height» – висота. Розмір

зображення потрібно обирати виходячи з моделі, яка використовується, наприклад, для «Stable Diffusion 1.5» краще обрати 512×512 пікселів або наблизений до цього розмір. Можна збільшувати зображення по одній із сторін у 1.5-2 рази (наприклад, 768×512 або 512×1024), але не більше ніж у двічі, щоб не було подвоєння зображення. Для генерації зображення більшого розміру використовуйте «Hires. Fix» (24) під час генерації, або «Upscaler» після генерації. Для «img2img», «Sketch», «Inpaint», «Inpaint sketch», «Inpaint upload». На вкладці «Resize by» можна задати коефіцієнт збільшення «Scale», замість розміру зображення (рис. 3.1, 3.3, 3.5);

– «Stable Diffusion Batch» (27) – «AUTOMATIC1111/stable-diffusion-webui» дозволяє робити пакетну («Batch») генерацію/обробку зображень. Є два параметри для керування пакетною генерацією/обробкою, «Batch count» – скільки раз повторити генерацію/обробку з заданими параметрами, та «Batch size» – скільки зображень генерувати/обробляти паралельно (для паралельної генерації/обробки потрібно більше відеопам'яті) (рис. 3.1, 3.3, 3.5);

– «Stable Diffusion CFG Scale» (28) – параметр того, як сильно генероване зображення повинно бути схожим на опис. Чим вище значення – тим більше нейронна мережа «Stable Diffusion» буде намагатись зобразити на зображенні точно те, що вказано у запиті. Нижчі значення будуть давати нейронній мережі вільний простір для творчості. Оптимальні значення 6-10 (рис. 3.1, 3.3, 3.5);

– «Denoising strength» (Доступно для «img2img», «Sketch», «Inpaint», «Inpaint sketch», «Inpaint upload») (29) – «Denoising strength» регулює силу змінення зображення під час перетворення зображення в зображення («img2img») або перемальовування («Inpaint»). Впливає на те, наскільки сильно нейронна мережа «Stable Diffusion» може змінити вхідне зображення під час генерації. Нижчі значення – менші зміни, вищі значення – більші зміни (рис. 3.3, 3.5);

– «Stable Diffusion Seed» (30) – це зерно для початку генерації випадкових чисел. Воно впливає на згенерований ряд випадкових чисел. У

«Stable Diffusion», «Seed» використовується для повторної генерації того самого зображення. Вказавши точний запит, параметри та «Seed» згенерованого зображення, нейронна мережа згенерує таке саме зображення. Значення «-1» – означає випадковий «Seed». Поруч є дві кнопки, кнопка з «кубиком» ставить значення «Seed» в «-1», а кнопка з зеленими стрілочками бере «Seed» з попереднього запиту (рис. 3.1, 3.3, 3.5);

– модулі і розширення (31) – відображає встановлені модулі і розширення, такі як «Adetailer», «ControlNet» та інші. Їх можна вмикати/вимикати та налаштовувати для кожної генерації (рис. 3.1, 3.3, 3.5);

– «Stable Diffusion Scripts» (32) – відображає встановлені скріпти. Обраний скріпт буде виконуватись під час генерації зображення (рис. 3.1, 3.3, 3.5).

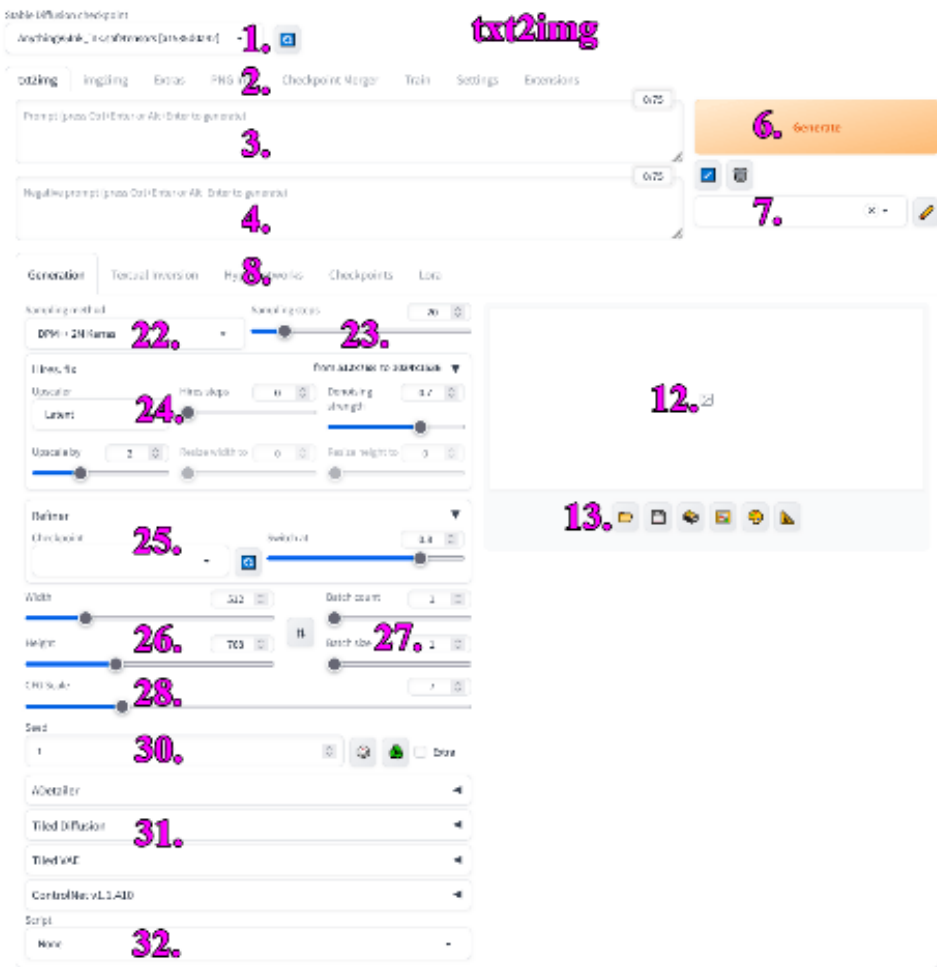


Рисунок 3.1 – Генерація з текстового запиту («txt2image»)

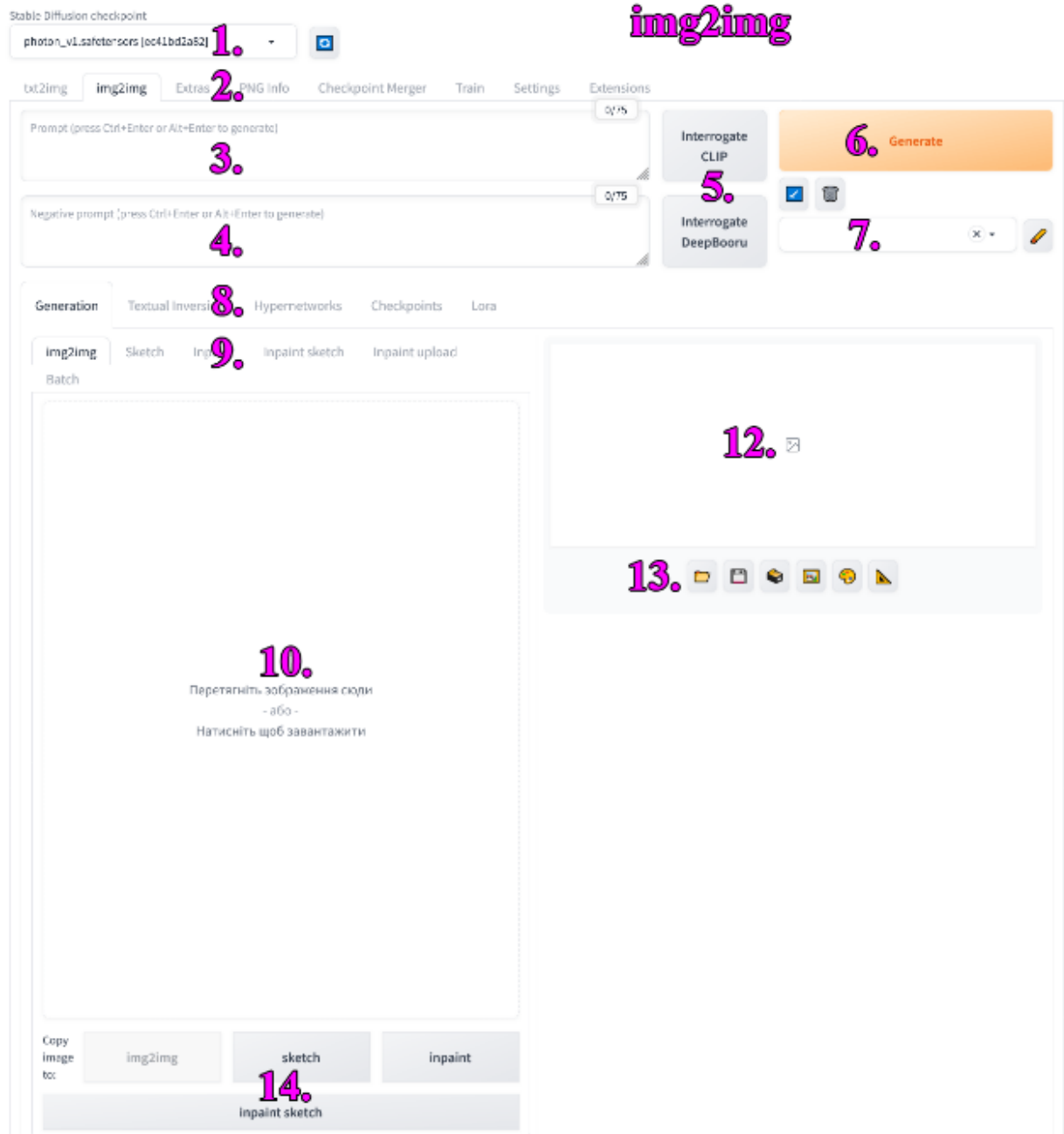


Рисунок 3.2 – Генерація з іншого зображення («img2img»)

img2img

Resize mode
 Just resize Crop and resize Resize and fill
 Just resize (latent upscale) **15.**

Sampling method: Euler a **22.** Sampling steps: 20 **23.**

Refiner: Checkpoint **25.** Switch at: 0.8

Resize to / Resize by
Width: 512
Height: **26.** 512

Batch count: 1
Batch size: **27.**

CFG Scale: **28.** 7

Denoising strength: **29.** 0.75

Seed: -1 **30.** Extra

ADetailer
Tiled Diffusion **31.**
Tiled VAE
ControlNet v1.1.410

Script: None **32.**

Рисунок 3.3 – Вкладка налаштувань («img2img»)

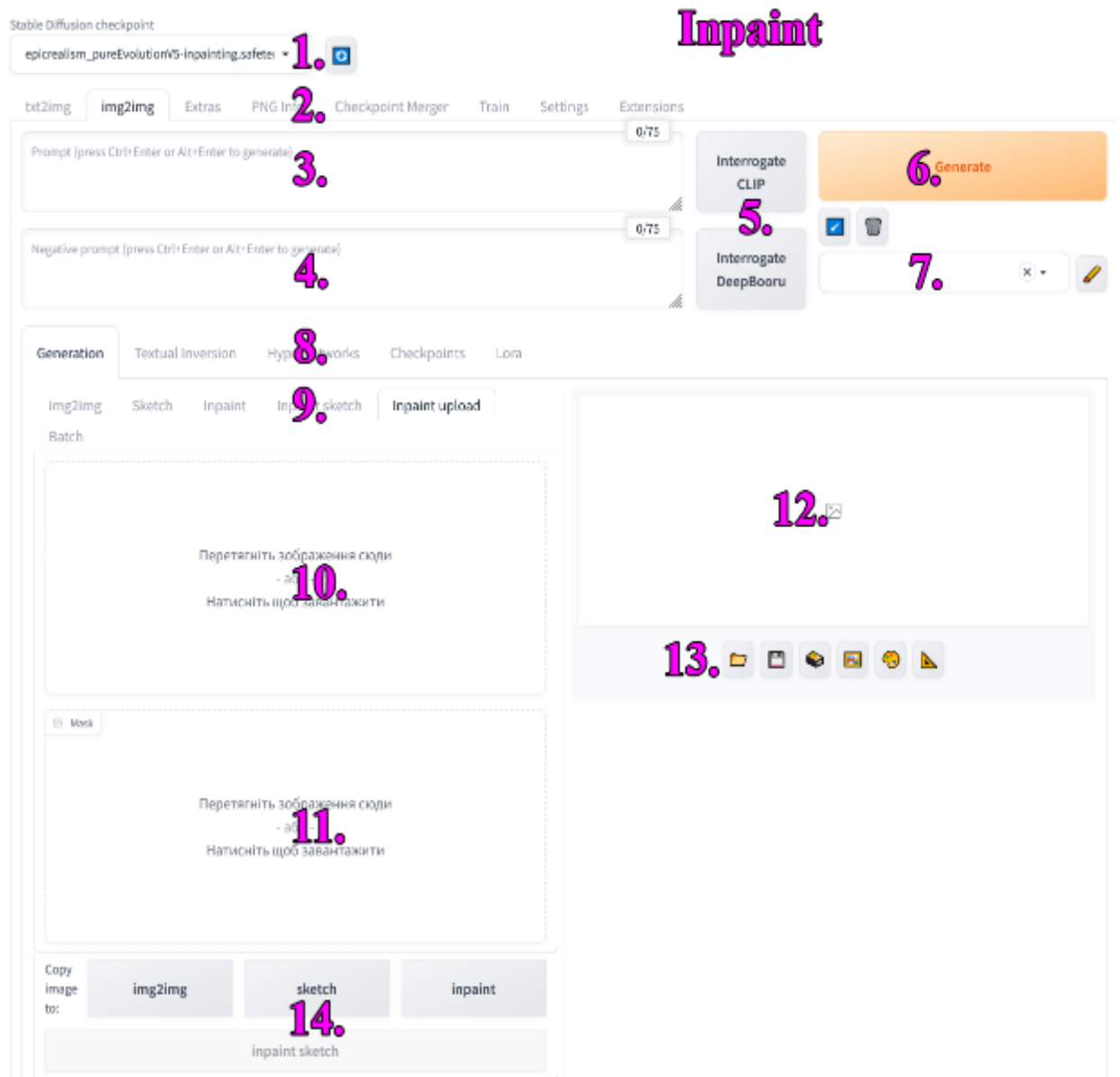


Рисунок 3.4 – Перемальовування наявного зображення («inpaint»)

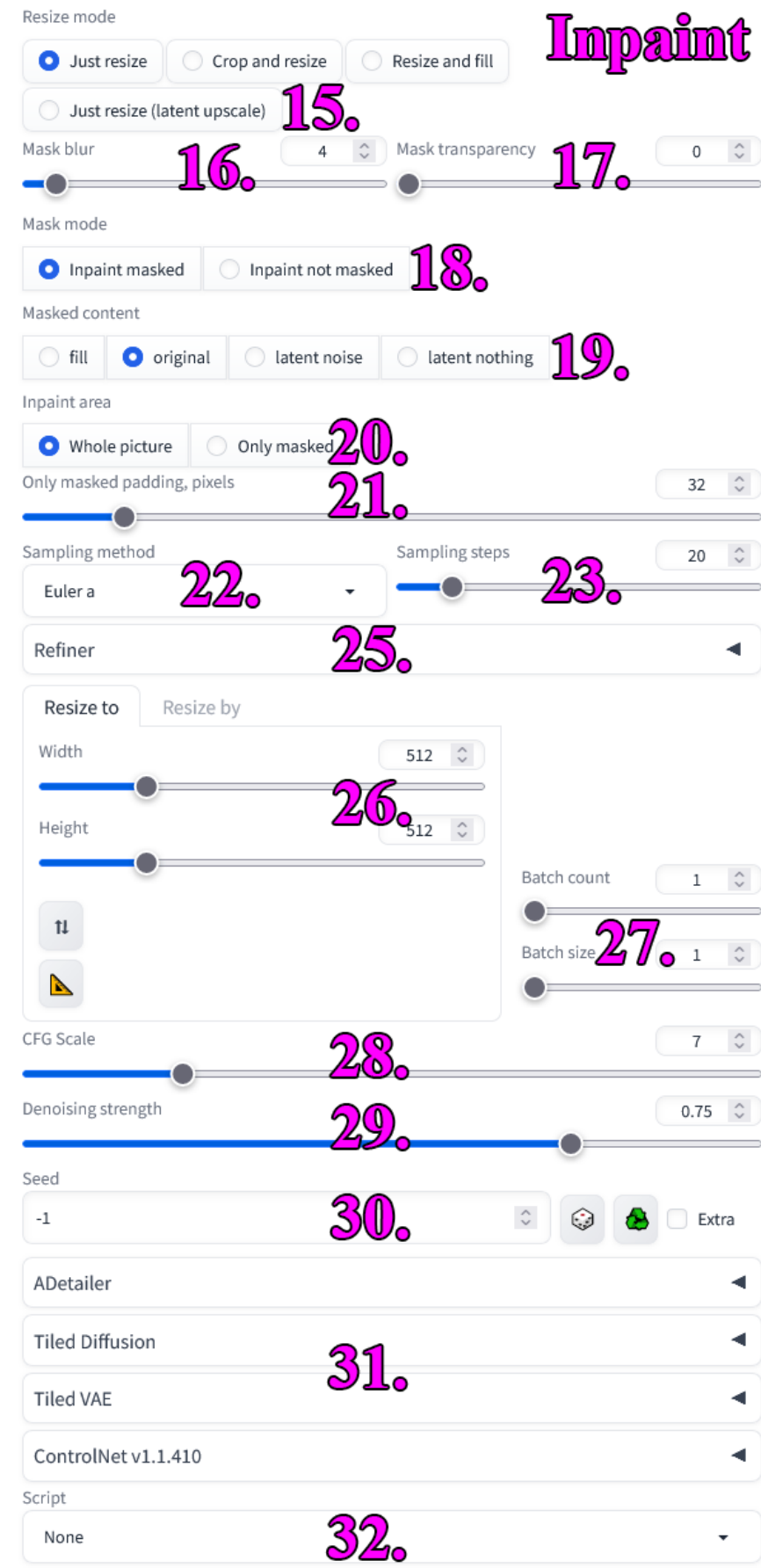


Рисунок 3.5 – Вкладка налаштувань («inpaint»)

3.2 Проектні підходи з використанням «GAI»

Оскільки «Generative AI» («GAI») може створювати унікальні графічні артефакти для різних завдань, важливо визначити його основні властивості як складову технологій розробки друкованого або мультимедійного контенту. Багато дослідників досліджували, як знайти оптимальний дизайн у багатовимірному просторі можливих рішень, які пропонують системи «GAI» [13-15]. Ці архітектури, які лежать в основі генеративних послуг, є ймовірнісними і призначені для генерування широкого діапазону можливих результатів. Ця генеративна варіабельність забезпечує унікальність вихідних даних, навіть коли вхідні дані та налаштування системи ідентичні [15].

Стаття [2] пропонує класифікацію всіх можливих сценаріїв взаємодії між модним дизайнером та генеративною службою. Ця класифікація включає п'ять сценаріїв (в статті вони називаються шаблонами взаємодії): Кураторство, Дослідження, Еволюція, Умовлення та Переписування. Ці сценарії охоплюють сім основних операцій взаємодії між дизайнером та системою «GAI» (рис. 3.6):

- ініціалізація; («initialization»);
- навчання; («training»);
- додавання обмежень («adding restrictions»);
- створення («creation»);
- вибір («selection»);
- адаптація («adaptation»);
- комбінація. («combination »).

При створенні ілюстративного контенту дизайнери часто стикаються з невизначеністю щодо оптимального художнього стилю та зовнішнього вигляду. Тому вони зазвичай починають з проведення декількох пробних генерацій, слідуючи за сценарієм дослідження. Під час цього процесу дизайнери налаштовують налаштування системи, вибираючи типи моделей, методи вибірки для навчання, кількість кроків, масштабування класифікатора та інші параметри.

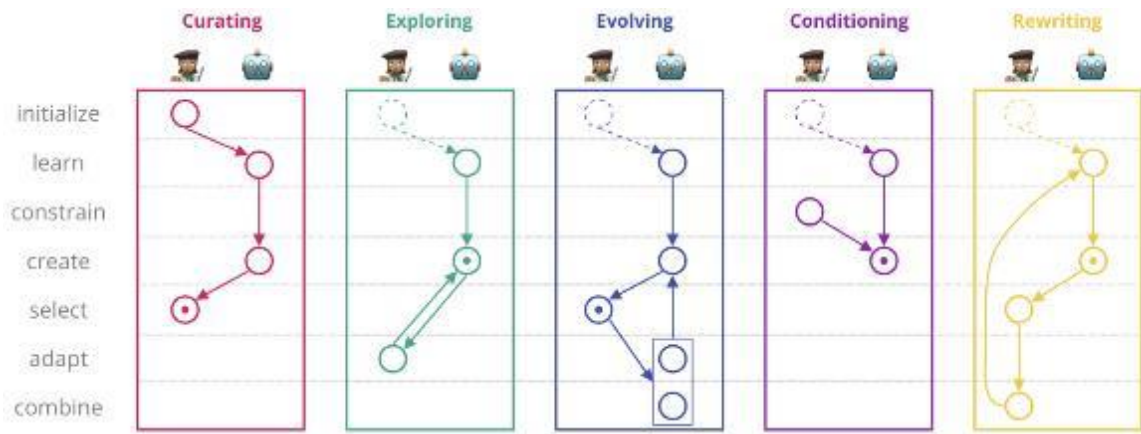


Рисунок 3.6 – Базові сценарії операції дизайнера і спеціаліста з «GAI» [5]

Вони формулюють базовий «prompt», часто без зразків (називається «zero-shot prompt»), і включають «negative prompts» для підвищення реалістичності зображених об'єктів. Після початку процесу генерації дизайнери оцінюють відповідність результатів поставленому завданню та налаштовують параметри мережі на основі як позитивного, так і негативного зворотного зв'язку. Відповідно, початкова фаза дослідження включає кілька ітераційних дій, включаючи (рис. 3.7):

- ініціалізація («initialization»);
- додавання обмежень («adding restrictions»);
- навчання («training»);
- створення («creation»);
- вибір («choice»).

Як відомо, «Stable Diffusion» має функції редагування зображень, що дозволяють користувачам вибірково фарбувати елементи, застосовувати ефекти розмиття без зміни основного вмісту ілюстрації та зберігати оригінальний художній стиль. Використовуючи попередньо визначений набір основних операцій, викладених раніше, можна зробити висновок, що редагування та часткове покращення згенерованих зображень буде відбуватися на етапі оцінки через такі дії, як адаптація та комбінування.

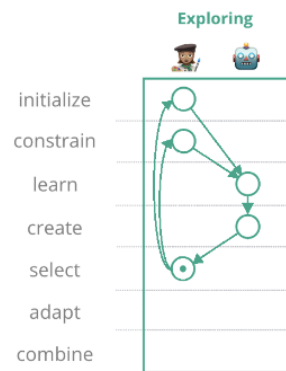


Рисунок 3.7 – Запропонована діаграма сценарію взаємодії дизайнера з Генеративними нейронами системами на етапі

Процес комбінування включає злиття поз об'єктів, згенерованих на етапі додавання умов, з найбільш задовільними результатами етапу дослідження. Після комбінування відбувається адаптація, що включає доопрацювання ілюстрацій за допомогою інструментів редагування на етапі оцінки. Таким чином, загальну схему технології розробки ілюстрацій для картонних ігор можна проілюструвати на рис. 3.8.

Щоб зрозуміти процес створення зображень за допомогою сервісу «Stable Diffusion», візьмемо технічне завдання від дизайнера, який планує пост-обробку цих зображень і включення їх у свій проєкт.

Технічне завдання від дизайнера: «Потрібен портрет літнього чоловіка у стилі кіберпанк до пояса».

Виконання завдання міститиме наступні дії:

- вибір моделі: «Lyriel_v16», оскільки вона створює хороший контраст між темним і світлим у цьому стилі;
- введення наступних «prompts»: «man in neon blue jacket, sci-fi clothes, mature, (futuristic:1.2), cyberpunk»;
- введення наступних «negative prompts»: «dof, grayscale, black and white, bw, 3d, cartoon, anime, sketches, (worst quality:2), (low quality:2), (normal quality:2), lowres, normal quality, ((monochrome)), ((grayscale)), skin spots, acnes, skin blemishes, bad anatomy, girl, loli, young, large breasts, red eyes, muscular, badhandsv5-neg, By bad artist-neg (1), monochrome, nsfw».

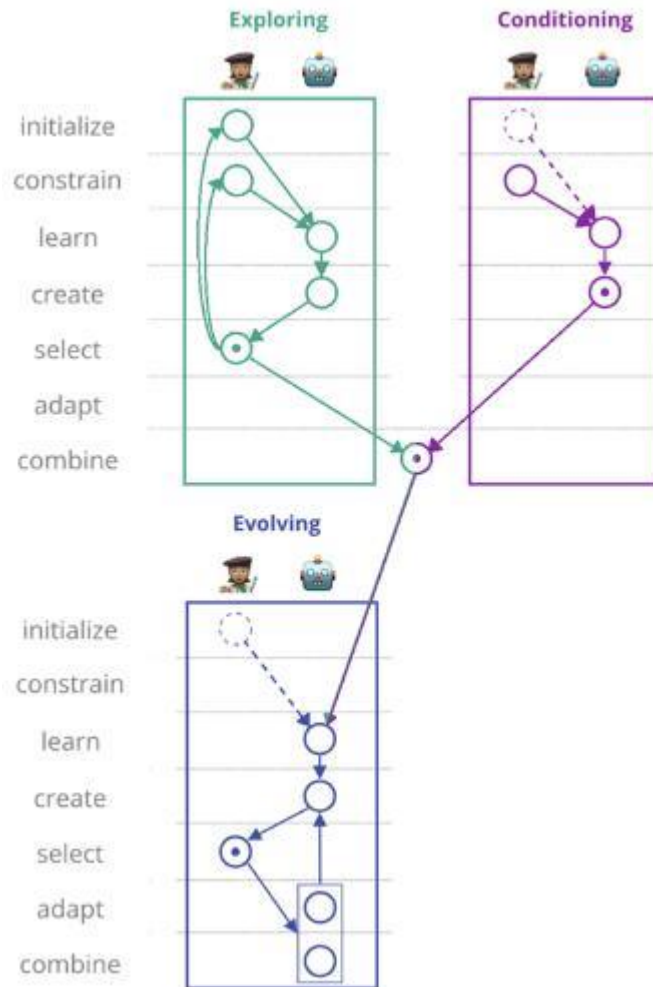


Рисунок 3.8 – Запропонована схема технології розробки дизайну ілюстрацій для настільних ігор

– встановити розмір зображення: 512x712 для створення портрету персонажа. Також ширина та висота повинні бути кратними 8 для коректної обробки нейронною мережею;

– вибрати метод семплування: «DPM++ 2M Karras», оскільки цей семплер рекомендується для моделі «Lyriel_v16» у її документації і є одним з найшвидших семплерів;

– встановити кількість кроків семплування: для фінальної версії зображення обрано 50 кроків, але для прискорення процесу вибору зазвичай встановлюється 20 – 30 кроків;

– «CFG Scale»: залишити значення за замовчуванням 7, оскільки цей параметр зазвичай не потребує налаштування;

– «seed»: після кількох спроб згенерувати бажане зображення було обрано значення 1688159780.

Коментарі під час процесу створення від дизайнера: «Зображення не нагадує кіберпанк стиль. Чоловік виглядає занадто старим. Чоловік не одягнений у потрібний одяг. Фон не відповідає темі кіберпанку» (рис. 3.9).



Рисунок 3.9 – Кінцевий результат генерації

4 ФІЛЬТРАЦІЯ ЗОБРАЖЕНЬ ТА СПІВПРАЦЯ З ДИЗАЙНЕРОМ

4.1 Критерії вибору зображень

При виборі зображень, згенерованих за допомогою сервісу «Stable Diffusion», для використання у друкованому графічному дизайні необхідно враховувати специфічні характеристики цього методу генерації зображень. Ось деякі ключові критерії, які можна використовувати при виборі зображень:

– реалізм і автентичність: зображення повинні мати достатній реалізм, щоб бути придатними для використання у друкованих матеріалах. Усі зображення потрібно перевіряти на наявність ірраціональних і нелогічних аспектів, таких як додаткові відростки або неправильне розташування об'єктів. Вони повинні точно відтворювати деталі та аспекти об'єктів з високою точністю;

– естетичний вигляд і композиція: зображення повинні мати естетично приємний вигляд і хорошу композицію, навіть якщо вони штучно створені. Це допомагає привертати увагу глядачів і підвищувати їхній інтерес;

– відповідність стилю і темі: зображення повинні відповідати стилю і темі друкованого матеріалу, для якого вони призначені. Наприклад, якщо це стиль кіберпанк, зображення повинні містити відповідні елементи та атмосферу;

– якість і роздільна здатність: зображення повинні бути високої якості і мати достатню роздільну здатність для забезпечення якісного друку. Це важливо для того, щоб зображення виглядали чіткими та привабливими у друку;

– унікальність і оригінальність: важливо, щоб зображення були унікальними і відповідали конкретним потребам проєкту. Вони не повинні бути загальними або застарілими.

Звертаючись до цих критеріїв, можна ефективно вибирати високоякісні та відповідні зображення, згенеровані нейронною мережею «Stable Diffusion», для використання у друкованому графічному дизайні.

4.2 Постобробка

Після генерації зображень за допомогою дифузійних моделей, які часто дають випадкові результати, виникає потреба в подальшій пост-обробці для досягнення бажаного результату. Дизайнерам зазвичай доводиться втручатися у зображення, щоб покращити їх естетичний вигляд і забезпечити відповідність заданому контексту. Ось деякі аспекти, які часто потребують уваги під час пост-обробки зображень:

- зміна кольорів для відповідності основним кольорам: на основі визначеного основного кольору або палітри важливо адаптувати кольорову гаму зображення, щоб забезпечити єдність і гармонію з загальним візуальним стилем;

- кадрування зображення: іноді зображення, згенеровані дифузійними моделями, можуть включати непотрібні елементи або мати незручну композицію. Кадрування зображення дозволяє виділити головний об'єкт і посилити його вплив на глядача;

- ретушування артефактів і небажаних об'єктів на зображенні: під час генерації можуть з'являтися артефакти або додаткові об'єкти, які слід виправити або видалити з зображення для покращення його чистоти та якості;

- підготовка колажів зображень для обробки нейронною мережею за допомогою «img2img»: для деяких складних зображень може знадобитися попередня підготовка та обробка кожного елемента окремо перед об'єднанням їх у єдине ціле за допомогою нейронних мереж з «img2img».

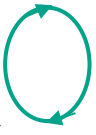
Пост-обробка зображень є важливим етапом у використанні дифузійних моделей для генерації зображень, оскільки вона допомагає покращити якість і виразність кінцевого результату, забезпечуючи відповідність заданим

дизайнерським вимогам і критеріям. На цьому етапі процес дизайну слідує схемі на рисунку 3.2, яка включає фази дослідження, умовності та еволюції. Налаштування конфігурацій генеративної системи описані в таблиці 2.1.

Таким чином, незважаючи на очевидний прогрес у генерації ілюстрацій, дизайнеру довелося вручну доводити зображення до бажаного стану: малювати фон, додавати ігрові позначки на карти та змінювати кольорове оформлення елементів зображення відповідно до ігрових позначок (рис. 4.1).

Згідно з результатами експериментального дослідження, очевидно, що генеративний сервіс ефективно виконує завдання дизайнера щодо створення ілюстрацій для картонної гри (рис. Б.1-Б.4).


Таблиця 4.1 – Детальні параметри генерації зображень із візуальним результатом

Дія	Параметри	Значення параметрів	Результат	Висновок дизайнера
Дослідження («Exploring»);				
Ініціалізація («initialization»)	«Model»	«Reliberate_v10»		
	«Zero-shot prompt»	«king, cyberpunk, holding the sphere, old man, neons, long beard»		
	Розмір зображення	512x712		
	«Sampling Method»	«DPM++ 2M Karras»		
	«Sampling steps»	20		
	«CFG Scale»	За замовченням 7		
	«Seed»  Через декілька ітерацій	1050799033		
Додавання обмежень («Adding restrictions»)	«Negative Prompts»	«female, low quality, worst quality, bad anatomy, bad hands, mutated, fewer digits, extra digits, (worst quality:1.2), (low quality:1.2), (low-res:1.1), (mono-chrome:1.1), (greyscale)»		

Продовження таблиці 4.1

1	2	3	4	5
Зміна умов («Conditioning»);				
Ініціалізація («initialization») «ControlNet» плагіну	«Few-shot prompt»			
Навчання «Learning»	«Open Pose»	«OpenPose_Full»		Поза, вік і верхня сфера персонажа відповідні, але одяг персонажа не одягнений. Замість нижньої сфери повинна бути корона.
Стадія покращення, ітерація 1 («Evolving stage, iteration 1»)				
Ініціалізація («initialization») «Stable Diffusion» у «Inpaint mode»	«Model»	«RevAnimatedInp»		
	«Prompts»	«standing, king, cyberpunk, holding the sphere, old man, neons»		
	«Sampling Method»	«DPM++ 2M Karras»		
	«Denoising Strength»	1.0		
	«Seed»	3263849708		
Додавання обмежень («Adding restrictions»)	«Negative Prompts»	«female, low quality, worst quality, bad anatomy, bad hands, mutated, fewer digits, extra digits, (worst quality:1.2), (low quality:1.2»		Майже все відповідає завданню, замість жовтої сфери внизу має бути корона

Продовження таблиці 4.1

1	2	3	4	5
Стадія покращення, ітерація 2 («Evolving stage, iteration 2»)				
Ініціалізація («initialization») «Stable Diffusion» у «Inpaint mode»	«Model»	«RevAnimatedInp»		
	«Prompts»	«crown»		
	«Sampling Method»	«DPM++ 2M Karras»		
	«Denoising Strength»	0.7		
	«Seed»	1209726908		
Added restrictions	«Negative Prompts»	«female, low quality, worst quality, bad anatomy, bad hands, mutated, fewer digits, extra digits, (worst quality:1.2), (low quality:1.2), (lowres:1.1), (monochrome:1.1)»		
	«PhotoPea App»			
	«Inpaint Area»	«Only Masked»		Фон не відповідає стилю кіберпанк, сфера занадто маленька і не розташована посередині зображення, колір корони не відповідає гамі зображення, корона тьмяна, персонаж не асоціюється з образом короля.

Однак спостерігається, що дизайнер непослідовно формулює «prompts», що призводить до випадків, коли згенеровані дизайни не відповідають заданим специфікаціям дизайнера. Для підвищення ефективності генеративних систем у створенні ілюстративного контенту дизайнери повинні ретельно

документувати «prompts» та їхні оцінки згенерованих результатів на кожному етапі. Це сприяє узгодженню критеріїв оцінки та інструкцій з можливостями системи [16].



Рисунок 4.1 – Кінцевий результат дизайну зображення для картонної гри після ручного доопрацювання дизайнером

Штучно згенеровані зображення можуть слугувати основою для подальшого вдосконалення дизайну. Це передбачає внесення коригувань у процес розробки ілюстрацій, включаючи додавання дій адаптації та інтеграції на етапі еволюції. Ці дії виконуються не тільки автоматизованою системою, але й вручну дизайнером, як показано на рисунок 4.2.

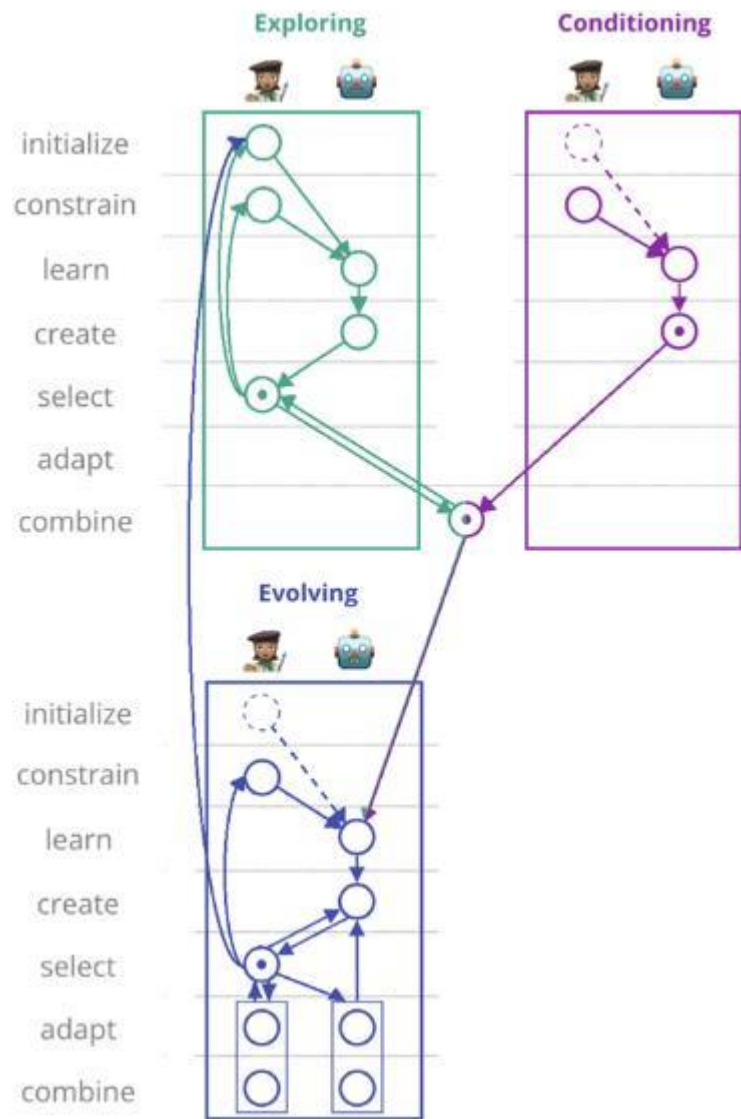


Рисунок 4.2 – Схема технології дизайну ілюстрацій для настільних ігор, включаючи ручне доопрацювання зображень дизайнером

Крім того, важливо враховувати зворотний зв'язок на основі вибору прийнятних варіантів дизайну на етапі оцінки.

Надалі може виникнути потреба або редагувати зображення за допомогою генеративної системи на етапі оцінки, або переглянути «prompts» та згенерувати їх знову в межах цього ж етапу, або ж, можливо, повернутися до початкових етапів процесу дизайну для ініціювання генерації з новими параметрами, налаштуваннями та «prompts».

5 ЕКОНОМІЧНА ЧАСТИНА

Використання генеративних сервісів, таких як штучний інтелект для створення контенту, моделей або рішень, має значну економічну вигоду для підприємств та організацій.

Загалом, використання генеративних сервісів сприяє підвищенню ефективності, зниженню витрат та покращенню якості продуктів і послуг, що в сукупності забезпечує значну економічну вигоду для компаній та організацій.

5.1 Економічні переваги генеративних нейронних мереж

Наведемо економічні переваги.

1. Зниження витрат на трудові ресурси за допомогою автоматизації рутинних завдань, таких як обробка даних, створення звітів, написання текстів і навіть розробка дизайну. Це дозволяє скоротити кількість співробітників, зайнятих цими завданнями, або перенаправити їх на більш важливі проєкти.

2. Зменшення потреби у вузькоспеціалізованих фахівцях. Наприклад, сервіси для автоматичного створення графіки можуть зменшити потребу в художниках. Це особливо вигідно для малих і середніх підприємств, які не завжди можуть дозволити собі наймати дорогих спеціалістів.

3. Підвищення продуктивності завдяки швидкому створенню контенту, чи то текст, чи то зображення або моделі. Це дозволяє співробітникам зосередитися на аналізі, креативному мисленні та інших стратегічних завданнях.

4. Автоматизація процесів, які раніше займали години або дні, можуть бути завершені за лічені хвилини. Це дозволяє швидше впроваджувати нові продукти та послуги.

5. Інновації та конкурентоспроможність – швидке впровадження нових ідей. Генеративні сервіси можуть швидко створювати прототипи нових

продуктів та ідей, що дозволяє компаніям експериментувати та впроваджувати інновації швидше, ніж конкуренти.

6. Конкурентні переваги. Використання передових технологій дозволяє компаніям залишатися на передовій ринку, пропонуючи більш інноваційні та привабливі продукти.

7. Економія часу завдяки автоматизації та скороченню циклів розробки. Генеративні сервіси дозволяють швидше завершувати проєкти, зменшуючи тривалість циклів розробки продуктів та впровадження нових послуг.

8. Персоналізація та масштабованість. Генеративні сервіси дозволяють компаніям легко масштабувати свої послуги відповідно до зростання попиту, не витрачаючи значні ресурси на наймання додаткових працівників.

5.2 Джерела економії, доходу та фінансування

Джерелом економії при використанні сервісів генерування зображень за допомогою нейронних мереж є зменшення часу на створення проєкту та заощадження коштів через відсутність працевлаштування робітників.

Для компанії, що використовує ці сервіси, джерелом доходу є можливість створення якісного візуального контенту, який підвищує привабливість продуктів і послуг, залучає більше клієнтів і збільшує обсяги продажів за низькою ціною завдяки генеративним нейронним мережам.

Витрати компанії включають закупівлю новітньої техніки для генерування зображень, витрати на інтеграцію з існуючими системами, а також витрати на навчання персоналу для ефективного використання сервісу.

5.3 Порядок генерування зображення

Загалом генерування зображень поділяється на такі етапи:

– отримання технічного завдання. На даному етапі відбувається початкова зустріч або консультація з клієнтом для визначення рамок

майбутнього проєкту. Обговорюються основні вимоги та очікування клієнта, включаючи стиль зображень, бюджет проєкту та строки виконання. Визначаються особливості співпраці, такі як частота зв'язку, формат зворотного зв'язку, необхідність проміжних звітів та інші організаційні моменти;

- формулювання теми для генерації зображень на основі отриманої від клієнта інформації, що може включати в себе збір референсів і прикладів, які допоможуть точніше передати бажаний стиль і настрій зображень;

- проводиться ескізування, тобто створення попередніх начерків або схем зображень, які будуть відправлені замовнику для схвалення. Це дозволяє переконатися, що очікування клієнта і художнє бачення команди збігаються;

- процес генерування зображень, в ході якого використовується сервіс генерування зображень на основі нейронних мереж. Важливо підібрати потрібні налаштування, такі як параметри стилю, рівень деталізації, кольорова гама тощо. Після цього генерується велика кількість зразків зображень, з яких проводиться відбір найбільш вдалих варіантів, що відповідають вимогам клієнта;

- постобробка вибраних зображень дизайнером. Відібрані зображення передаються дизайнеру для подальшої обробки. Це може включати корекцію кольорів, додавання деталей, покращення якості зображень та інші маніпуляції для досягнення найкращого результату;

- створення додаткових генерацій у разі необхідності;

- збір зображень у готовий продукт. Це може бути набір ілюстрацій, дизайн рекламних матеріалів, контент для веб-сайтів або будь-який інший продукт, що був замовлений клієнтом;

- презентація готового продукту замовнику. На цьому етапі клієнт може вносити свої корективи або затверджувати кінцевий результат.

5.4 Розрахунок собівартості та ціни генерування зображень

Реалізацією проєкту займається два фахівці: «Prompt Engineer» генерує зображення, а дизайнер їх обробляє. Заробітна плата «Prompt Engineer»

спеціаліста складає 125,00 грн/год, а дизайнера – 100,00 грн/год. Робочий день складає 8 годин мінус одна година на перерву. Отже, 5 робочих днів на тиждень по 7 годин.

Проект є комплексним, включає внесення правок замовником та узгодження остаточних варіантів кожного елемента.

До собівартості розробки проєкту входить основна заробітна плата, єдиний соціальний внесок та інші витрати.

Розрахунок основної заробітної плати наведений у таблиці 5.1.

Таблиця 5.1 – Розрахунок витрат на заробітну плату

Етап	Вид робіт	Годинна ставка, грн/год	Тривалість виконання, години	Заробітна плата, грн
Отримання технічного завдання	Обговорення основних вимог та очікування клієнта	225,00	2	450,00
Визначення рамок майбутнього проєкту	Формулювання теми для генерації	225,00	4	900,00
Генерування зображень	Створення попередніх начерків або схем зображень	125,00	50	6250,00
Відбір зображень	Огляд усіх варіантів зображень та відбір тих, що підходять за параметрами	225,00	2	450,00
Постобробка	Дизайнер прибирає непотрібні елементи робить корекцію кольорів, додає деталі та покращує якість зображень	100,00	30	3000,00
Презентація проєкту замовнику	Показуються усі аспекти і особливості проєкту замовнику	225,00	2	450,00
Усього			90	11500,00

Ставка єдиного соціального внеску становить 22 % від величини основної і додаткової заробітної плати:

$$11500,00 * 0,22 = 2530,00 \text{ грн.}$$

Розрахуємо інші витрати до яких входять:

- обслуговування техніки;
- оплата рахунків за електроенергію.

Витрати на електроенергію розраховуються виходячи зі споживаної потужності використовуваних пристроїв і тарифу на електроенергію. У даному випадку передбачається використання одного комп'ютера потужністю 0,28 кВт/год. Вартість однієї кВт/год електроенергії прийнято у розмірі 2,64 кВт/грн. Отже, плата за електроенергію складе:

$$0,28 * 2,64 * 90 = 66,53 \text{ грн.}$$

Витрати на обслуговування комп'ютера визначаються з його вартості та часу експлуатації, після закінчення якого він підлягає заміні (звичайно цей час не перевищує 3-х років). Комп'ютер коштує 84000,00 грн та протягом року використовується 254 робочих дні. Отже, витрати на його обслуговування складатимуть:

$$(84000,00 / (3 * 8 * 254)) * 90 = 1240,16 \text{ грн.}$$

Складемо отримані дані та отримаємо собівартість проєкту:

$$11500,00 + 2530,00 + 66,53 + 1240,16 = 15336,69 \text{ грн.}$$

Проєкт розроблюється для одного замовника, тому собівартість розробки становитиме 15336,69 грн:

$$15336,69 / 1 = 15336,69.$$

Розрахуємо суму прибутку від реалізації розробки (виходячи з рівня рентабельності 35 %):

$$15336,69 * 0,35 = 5367,84 \text{ грн.}$$

Розрахуємо ціну розробки проєкту без ПДВ:

$$15336,69 + 5367,84 = 20704,53 \text{ грн.}$$

Розрахуємо суму ПДВ, вона дорівнює 20% від ціни:

$$20704,53 * 0,2 = 4140,91 \text{ грн.}$$

Розрахуємо ціну проєкту з урахуванням ПДВ:

$$20704,53 + 4140,91 = 24845,43 \text{ грн.}$$

Результати розрахунків наведено у таблиці 5.2.

Таблиця 5.2 – Розрахунок витрат на розробку та ціни проєкту

Показник	Сума, грн
Основна заробітна плата	11500,00
Єдиний соціальний внесок	2530,00
Витрати на обслуговування ЕОМ	1240,16
Витрати на електроенергію	66,53
Собівартість розробки проєкту	15336,69
Прибуток (рівень рентабельності 35 %)	5367,84
Ціна без ПДВ	20704,53
ПДВ	4140,91
Ціна з урахуванням ПДВ	24845,43

Виходячи з виконаних розрахунків, вартість розробки проєкту з уживанням генеративної нейронної мережі складатиме 24845,43 грн. Розробка здійснюватиметься 90 годин, тобто приблизно 12 днів одним дизайнером та одним «Prompt Engineer». Очікувана сума прибутку складе 5367,84 грн, що підтверджує доцільність реалізації проєкту.

ВИСНОВОК

У цьому проекті розглядається проблема прийняття рішень у процесі проектування художніх творів настільних ігор з використанням генеративних систем штучного інтелекту. Розглянуто сутність та послідовність етапів взаємодії дизайнера та спеціаліста «GAI» під час розробки ілюстративного контенту. На основі встановлених особливостей модифіковано кілька типових сценаріїв взаємодії: сценарій дослідження, сценарій кондиціонування, сценарій розвитку. Для окремих дій, об'єднаних цими сценаріями, було встановлено певний порядок і запропоновано додавання зворотних зв'язків між діями окремих сценаріїв. Введені зворотні зв'язки дозволили створити безперервну трьохетапну схему дизайну художнього оформлення настільних ігор з використанням генеративних систем.

Проведений експеримент підтвердив ефективність розробленої схеми та додав рекомендації щодо більш точного формулювання текстових інструкцій до генеруючої системи. Запропонований підхід дозволяє значно пришвидшити роботу дизайнера при оформленні ілюстративного наповнення видань та підвищити якість унікального креативного наповнення.

За результатами проходження передатестаційної практики були створені ілюстрації до друкованих матеріалів за допомогою інноваційного інструменту – нейронної мережі «Stable Diffusion». Були детально вивчені механізми роботи нейронних мереж, їх переваги та недоліки, а також особливості інтерфейсу та налаштувань сервісу.

У процесі аналізувалися такі аспекти, як нюанси генерації зображення, якість результатів і можливість подальшого редагування. Як ми бачили, постобробка, зокрема виправлення артефактів, зміна кольорів, кадрування та ретушування, є важливою частиною використання зображень, створених службою, у друкованих продуктах. Без втручання дизайнера зображення,

створені штучним інтелектом, часто мають колірний дисонанс, можуть бути нелогічними або побудовані з помилковою композицією.

Особливу увагу було приділено інтеграції зображень у готові макети друкованих матеріалів. Важливо було забезпечити гармонійне поєднання створюваних ілюстрацій з іншими елементами дизайну при збереженні загального стилю та естетики виробу. Це вимагало додаткового налаштування та підгонки зображень під конкретні вимоги проекту.

Підсумовуючи, використання «Stable Diffusion» для створення ілюстрацій дозволило створювати високоякісні зображення, які після додаткової обробки можна успішно використовувати в поліграфічній продукції. Цей досвід продемонстрував значний потенціал нейронних мереж у дизайні та відкрив нові можливості їх застосування у творчих проектах.

В економічній частині роботи було розраховано вартість проекту розробки ілюстрацій до настільної гри з використанням генеративної нейронної мережі.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. ConvAI Convai. URL: <https://www.convai.com/> (дата звернення: 17.05.2024).
2. CSM. CSM AI. URL: <https://csm.ai/>» (дата звернення: 17.05.2024).
3. PixelVibe. URL: <https://www.pixelvibe.com/> (дата звернення: 17.05.2024).
4. Kikuchi K., Inoue N., Otani M., Simo-Serra E., Yamaguchi K. Generative Colorization of Structured Mobile Web Pages // 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). 2023. P. 3639-3648. doi 10.1109/WACV56688.2023.00364.
5. Grabe I., Zhu J. Towards. Co-Creative Generative Adversarial Networks for Fashion Designers // arXiv:2304.09477v1. 2023. URL: <https://doi.org/10.48550/arXiv.2304.09477> (дата звернення: 17.05.2024).
6. Kulkarni C., Druga S., Chang M., Fiannaca A., Cai C., Terry M. A word is worth a thousand pictures: Prompts as ai design material // arXiv:2303.12647v1. 2023. URL: <https://doi.org/10.48550/arXiv.2303.12647> (дата звернення: 17.05.2024).
7. Chen R., Zhao J., Yao X., Jiang S., He Y., Bao B., Luo X., Xu S., Wang C. Generative Design of Outdoor Green Spaces Based on Generative Adversarial Networks // Buildings. 2023. № 13. P. 1083. URL: <https://doi.org/10.3390/buildings13041083> (дата звернення: 17.05.2024).
8. Kong X., Jiang L., Chang H., Zhang H., Hao Y., Gong H., Essa I. Bidirectional Layout Transformer for Controllable Layout Generation // arXiv:2112.05112v2. 2021. URL: <https://doi.org/10.48550/arXiv.2112.05112> (дата звернення: 17.05.2024).
9. Cao Y., Li S., Liu Y., Yan Z., Dai Y., Yu P.S., Sun L. A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT // arXiv:2303.04226v1. 2023. URL: <https://doi.org/10.48550/arXiv.2303.04226> (дата звернення: 17.05.2024).
10. Goodfellow I. J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative adversarial networks //

Communications of the ACM. 2014. Глава 63. С. 139-144. URL: <https://doi.org/10.1145/3422622> (Дата звернення: 17.05.2024).

11. Ho J., Jain A., Abbeel P. Denoising Diffusion Probabilistic Models // arXiv:2006.11239v2. 2020. URL: <https://doi.org/10.48550/arXiv.2006.11239> (дата звернення: 17.05.2024).

12. Rombach R., Blattmann A., Lorenz D., Esser P., Ommer B. High-Resolution Image Synthesis with Latent Diffusion Models // International Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, LA. 2022. P. 10684-10695. arXiv:2112.10752. URL: <https://doi.org/10.48550/arXiv.2112.10752> (Дата звернення: 17.05.2024).

13. Liao Q.V., Gruen D., Miller S. Questioning the ai: informing design practices for explainable ai user experiences // Proceedings of the 2020 CHI – Conference on Human Factors in Computing Systems». 2020. P. 1-15.

14. Morris M.R., Cai C.J., Holbrook J., Kulkarni C., Terry M. The design space of generative models // Proceedings of the NeurIPS 2022. Workshop on Human-Centered AI (NeurIPS2022). 2022.

15. Weisz J.D., Muller M., He J., Houde S. Toward General Design Principles for Generative AI Applications // arXiv:2301.05578. 2023. URL: <https://doi.org/10.48550/arXiv.2301.05578> (Дата звернення: 17.05.2024).

16. Кулішова Н.Є, Столяров І.В, Цикало С.С. Decision making in process of board games artwork design using generative artificial intelligence applications // Технологія і техніка друкарства. 2024. № 1(83). С. 17-25.

17. Полозова Т.В. Методичні вказівки до виконання економічної частини кваліфікаційної роботи для здобувачів першого (бакалаврського) рівня вищої освіти спеціальності 186 Видавництво та поліграфія усіх форм навчання. Харків: ХНУРЕ, 2022. 47 с.