

*Зіглін А. О., бакалавр, Чембержі Д. А., кандидат мистецтвознавства, доцент  
Київський національний університет технологій та дизайну*

### **ОСОБЛИВОСТІ ІНТЕГРАЦІЇ ШТУЧНОГО ІНТЕЛЕКТУ У РОБОТУ ЦИФРОВОГО ХУДОЖНИКА**

*Анотація.* Розглянуто новітні технології у сфері генеративного штучного інтелекту в сфері зображень, можливості інтеграції нейромереж і алгоритми створення цифрового живопису сучасних цифрових художників та пошук найоптимальнішого процесу використання із застосуванням додаткових плагінів для контролю роботи штучного інтелекту з передбачуваним та контрольованим результатом.

*Ключові слова:* штучний інтелект, цифровий художник, Stable Diffusion, ControlNet, цифрове мистецтво, ілюстрація.

*Ziglin A. O., Chemberzhi D. A.*

*Kyiv National University of Technologies and Design*

### **FEATURES OF THE INTEGRATION OF ARTIFICIAL INTELLIGENCE INTO THE WORK OF A DIGITAL ARTIST**

*Abstract.* The latest technologies in the field of generative artificial intelligence in the field of image processing, the possibility of integrating neural networks into the digital painting algorithms of modern digital artists and the search for the most optimal process of use with the use of additional plugins to control the work of artificial intelligence with a predictable result are considered.

*Key words:* artificial intelligence, digital artist, Stable Diffusion, ControlNet, digital art, illustration.

**Вступ.** У час, коли прогрес не стоїть на місці і технології розвиваються шаленим темпом, неможливо заперечувати технічні досягнення навіть у сфері мистецтва. З розвитком методів машинного навчання, які можуть генерувати зображення, для митців відкрились нові горизонти використання штучного інтелекту у своїх ілюстраціях, перекладаючи на нього частину рутини, і залишаючи собі роботу над концепціями та сенсами. Розвиток ШІ відкриває перед цифровими художниками нові можливості, проте створює і ряд проблем.

Однією з таких і є інтеграція ШІ у творчий процес. Деякі митці вважають, що творчість є виключно людською сферою, і що нейромережі ніколи не зможуть пізнати людську креативність, а отже – і бути сильним конкурентом для художників «старої школи». Інший табір митців вважає, що використання ШІ є обманом та підробкою, і спеціалісти, які інтегрують в свій пайплайн (англ. Pipeline – конвеєр. Це певний порядок дій, згідно з яким виконується робота) нейромережі – не є справжніми художниками.

Останні роки було проведено багато досліджень на тему генеративних нейромереж, хоча більшість з них все ще знаходиться на ранніх етапах розвитку. Одним з найперспективніших інструментом для цифрових художників є Stable Diffusion. На відміну від своїх прямих конкурентів DALL-E та Midjourney – SD має велику перевагу у кастомізації процесу роботи, створення власних моделей та використання плагіну image2image та ControlNet. Проте, в існуючих дослідженнях на цю тему, досить мало уваги приділено саме цим унікальним особливостям SD, які є найважливішими інструментами в руках цифрового художника і значно обходять звичний спосіб text2image.

Зацікавлення новітніми технологіями не проходить і повз великих гравців на ринку комп'ютерних ігор, так Blizzard Entertainment навчає генератор зображень на ресурсах World of Warcraft, Diablo та Overwatch. Оскільки генеративний штучний

інтелект створює мистецтво швидше, ніж будь-яка людина, такі студії, як Blizzard, підрозділ Activision Blizzard, сподіваються, що ця технологія зможе позбавити від роботи дизайн і розробку та зробити створення відеоігор веселішим [1].

Проте індустрія цифрового мистецтва диктує свої правила гри. Швидкість та якість роботи виноситься на перше місце, і спеціалісти вимушені шукати шляхи інтеграції інновацій у сфері ШІ у свої процеси роботи, аби залишатись конкурентними на ринку праці. Але така технічна сфера як машинне навчання – не є простим напрямом для людей мистецтва, а отже залишається проблема використання ШІ на базовому та споживацькому рівні, аби для будь якого фріланс-художника було зрозуміло, як врешті решт інтегрувати нейромережі у свій пайплайн роботи та використовувати інноваційні технології у повному обсязі.

**Постановка завдання.** Основною метою дослідження є пошук сучасних методів якісної інтеграції можливостей нейромережі SD у пайплайн цифрового художника, який зможе пришвидшити роботу та винести її на новий рівень якості. Також супутнім завданням дослідження є вивчення додаткових, більш тонких налаштувань Stable Diffusion, таких як image2image та ControlNet, а також можливостей його препроцесорів, для контролю процесу генерації.

**Результати досліджень.** Незважаючи на великий обсяг можливостей використання Stable Diffusion – більшість користується нею в класичному розумінні, як text2image [2]. Проте, SD скриває за собою дещо більше, ніж text2image, наприклад image2image та ControlNet, про які далі і буде йти мова. Image2image є досить поверхневим використанням першочергового зображення, як основи для майбутньої генерації, у той час як ControlNet дає широкий спектр можливостей. Проте використовуючи ці плагіни одночасно – можна досягти неймовірних результатів. Така зв'язка дає одночасно контроль і над елементами та контурами зображення, так і над кольоровою гаммою, що у підсумку дає найбільш прогнозований результат.

ControlNet – це архітектура нейронної мережі, призначена для вдосконалення попередньо вивчених моделей дифузії зображення на основі конкретних завдань. Він працює шляхом маніпулювання вхідними умовами блоків нейронної мережі, які є наборами нейронних рівнів, які використовуються для побудови нейронних мереж, навчених за допомогою векторів зовнішніх умов. Це дозволяє зберегти оригінальні ваги, уникаючи будь-яких змін через невеликі набори даних, і дозволяє швидше налаштувати, ніж навчати з нуля [3].

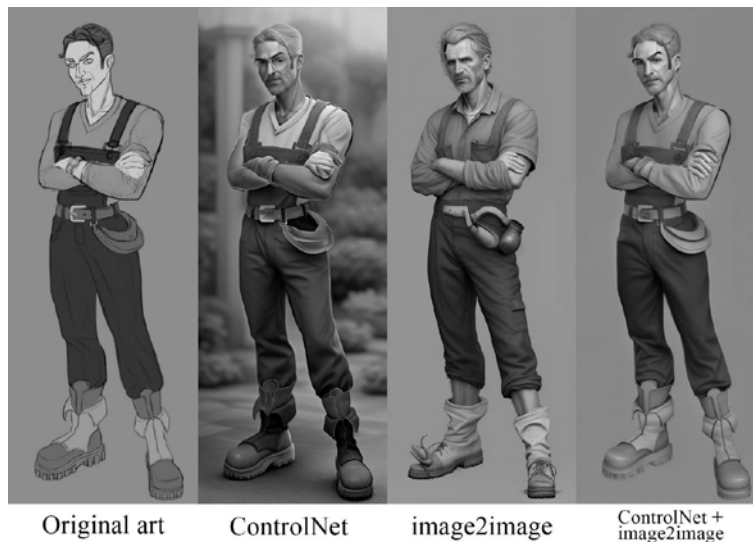
Використання image2image дає змогу генерувати нові зображення з вхідної картини та відповідної текстової підказки. Вихідне зображення зберігає першоджерельний колір і композицію зображення. Важливо зазначити, що вхідне зображення не повинно бути складним або візуально привабливим. Основну увагу слід приділити кольорам і композиції, оскільки ці елементи будуть перенесені на кінцевий результат [4].

Основним інструментом всередині ControlNet є препроцесори. У них є одна мета – отримати із зображення, яке було завантажено як приклад, позу, лінії, композицію або глибину зображення, на основі чого вже за допомогою текстового промта (англ. Prompt – підказка) нейромережа зможе видати фінальне зображення. Їх існує досить велика кількість, адже через відкритий код Stable Diffusion – кожен розробник має можливість створювати та кастомізувати web-інтерфейс на власний розсуд.

Серед основних препроцесорів існують такі, як:

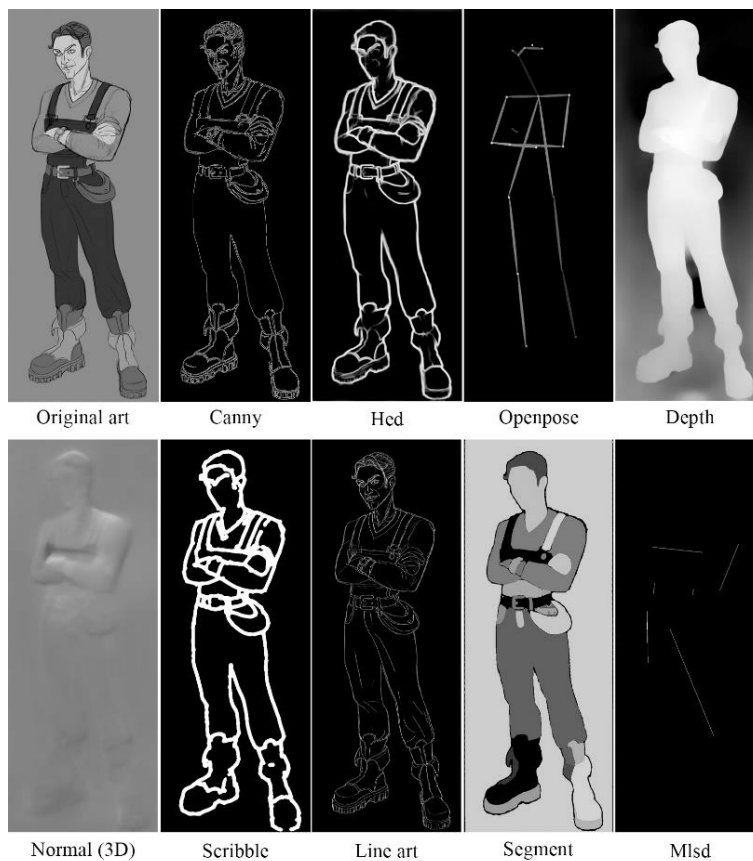
- Canny – витягує із зображення чіткі контури, створюючи умовний лайн-арт;
- Hed – витягує із зображення скетч, з м'якими краями;
- Openpose – бере із зображення лише позу;
- Depth – маска глибини, для витягування композиції;

- Normal – каарта об'єму (більше для 3D художників);
- Scribble – робить дитячі каракулі і створює із них щось нове;
- Mlsd – працює лише з рівними лініями, схожими та інтер'єрні/архітектурні креслення;
- Segment – ділить зображення на сегменти і генерує в них щось нове.



Джерело: на основі власних досліджень автора.

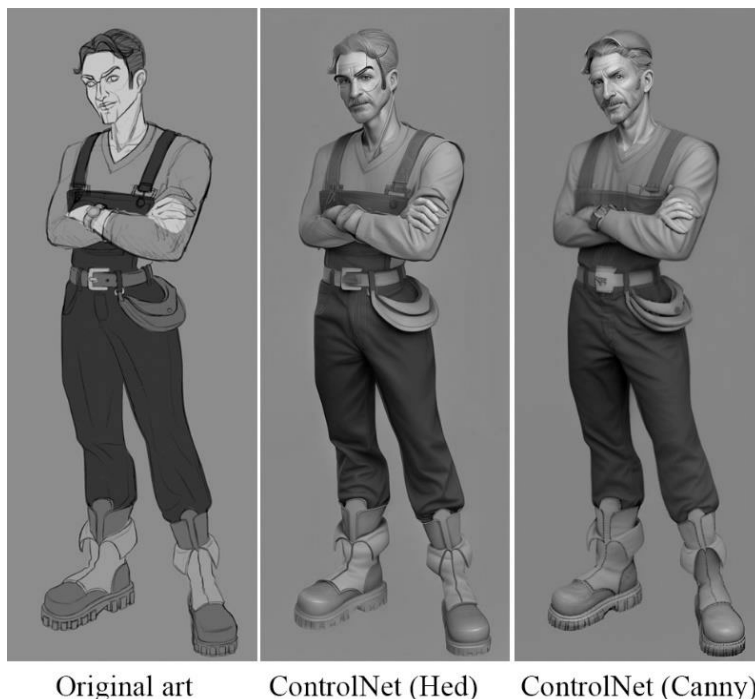
Рис. 1. Порівняння генерацій на різних функціях



Джерело: на основі власних досліджень автора.

Рис. 2. Приклади роботи основних препроцесорів

Серед усіх препроцесорів, для цифрових художників найкориснішими є Canny та Hed, проте навіть між ними існує різниця, яка вже буде залежити від поставлених цілей. Ця різниця вже може бути на розсуд самого спеціаліста, проте за досвідом постійної роботи з різними препроцесорами, найоптимальнішим є Hed для використання в роботі з власними намальовками та Openpose для використання пози із будь якого зображення.



Джерело: на основі власних досліджень автора.

Рис. 3. Приклади готових генерацій препроцесорів Hed та Canny

Окрім використання image2image та ControlNet, одним з найважливіших елементів генерації все ще залишається сама модель для генерації та текстовий промт. Для пошуку моделей головним ресурсом є Civitai. Проєкт Civit AI – це сайт, за допомогою якого можна знайти або опублікувати навчені моделі для нейромережі, а також багато іншого (Lora, Openpose, і тд). Це величезна бібліотека моделей з усього світу, яка дає можливість безкоштовно підшукати для себе потрібну основу для генерації на будь який смак: від гіперреалізму та кіберпанку до аніме та стилістики Pixar. [5] Найоптимальнішими варіантами, які покривають увесь спектр запитів середнього користувача є: SD 1.5, SDXL. (від розробників Stable Diffusion) та користувацька модель Deliberate. Вони дають можливість отримувати найбільш реалістичні результати, та мають певну бібліотеку стилізації, яка контролюється за допомогою промтів.

Стилізація за допомогою промта проходить наступним чином: після текстового опису того, що має бути зображено на картинці, треба вказати «тригерні» слова для певної моделі, на яких вона була навчена. На прикладі моделі Deliberate, набір тригерних слів для роботи в прикладах дослідження виглядатиме так:

- Mj – для використання характерної стилістики Midjourney;
- Cinematic – кінематографічний ефект, за допомогою якого можна отримати якісний кадр;
- Digital painting/illustration – стилістичні особливості уфрового мистецтва;
- Hyperdetailed – підвищення деталізації зображення;

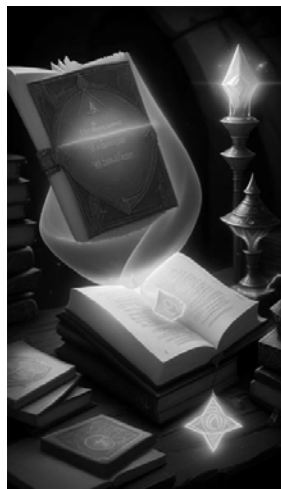
- Trending on artstation – використання сумарного аналізу трендових ілюстрацій на платформі Artstation, де публікуються портфоліо художників з усього світу.

До негативного промта прийнято базово використовувати наступний набір токенів, для запобігання деформації, анатомічних помилок та розмиття: deformed, distorted, disfigured, poorly draw, bad anatomy, wrong anatomy, extra limb, missing limb, floating limbs, mutated hands and fingers, mutation, ugly, blurry, amputation, etc.

Останнім кроком роботи всередині нейромережі є збільшення роздільної здатності зображення. Більшість моделей навчаються на зображеннях 512 на 512 – це найоптимальніший розмір, аби не перевантажувати GPU комп'ютера та зберегти розмір файлу моделі у розумних рамках. Проте використання зображення в такому розмірі може принести багато недоліків та артефактів, наприклад для друку або для відображення зображення на великих екранах. Для цих запитів кращим рішенням буде використання Upscale (висококласний), який дає можливість збільшити зображення у 4 рази, не проходячи знову процес генерації, а лише постобробку вже існуючої генерації [6].

То ж, рекомендований в результаті досліджень пайплайн роботи над артом з використанням нейромереж буде виглядати так:

Пошук референсів – формування текстового опису – перший прогін через нейромережі (text2image) – створення драфтів (англ. Drafts – чернетки) – створення фінального лайн- арту (для ControlNet) – заливка основних елементів арту – нанесення схематично основних світлотіньових моментів та атмосфери (для image2image) – доопрацювання у графічних редакторах.



first generation



Sketch



Second generation



Final art

Джерело: на основі власних досліджень автора.

**Рис. 4. Скорочений приклад інтеграції нейромереж у робочий процес**

**Висновки.** Штучний інтелект все далі стає одним з найперспективніших напрямків сьогодення, зокрема і у цифровому мистецтві. Попри неоднозначне ставлення до ШІ з боку законодавства різних держав – відмовитись від новітніх технологій буде не так просто. Дослідження "Generative AI and Jobs: глобальний аналіз потенційного впливу на кількість і якість робочих місць" показало, що ШІ буде мати переважно доповнювальний вплив на робочі місця. ШІ буде використовуватися для автоматизації деяких завдань, але не замінить людей повністю. Це означає, що більшість робочих місць і галузей будуть змінені, але не знищені [7, 8].

Незважаючи на всі недоліки фінальних результатів генерації Stable Diffusion та зокрема ControlNet – це незамінний інструмент у руках досвідченого цифрового

художника, який розуміє, як правильно інтегрувати у свій пайплайн роботи цей інструмент. За допомогою новітніх технології у сфері машинного навчання, процес та якість створення зображення кратно підвищується, що є переконливим аргументом для використання від передових ігрових студій до дрібного бізнесу.

### Список використаної літератури

1. Shannon, L. (2023). A.I. May Help Design Your Favorite Video Game Character. URL: <https://www.nytimes.com/2023/05/22/arts/blizzard-diffusion-ai-video-games.html?searchResultPosition=5>.
2. Dehouche, N., Dehouche, K. (2023). What's in a text-to-image prompt? The potential of stable diffusion in visual arts education. *Heliyon*.
3. Zhang, L., Rao, A., Agrawala, M. (2023). Adding conditional control to text-to-image diffusion models. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 3836–3847).
4. How to Use img2img in Stable Diffusion (Step-by-Step). *GreatAiPrompts*. URL: <https://www.greataiprompts.com/guide/how-to-use-img2img-in-stable-diffusion/>
5. Civitai: The Home of Open-Source Generative AI. URL: <https://civitai.com>.
6. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10684–10695).
7. Кацімон О. Штучний інтелект навряд чи знищить більшість робочих місць. Звіт ООН. URL: <https://suspilne.media/555807-stucnij-intelekt-navrad-ci-znisit-bilsist-robocih-misc-zvit-oon/>
8. Gmyrek, P., Berg, Ja., Bescond, D. Generative AI and jobs: A global analysis of potential effects on job quantity and quality. URL: [https://www.ilo.org/wcmsp5/groups/public/---dgreports/---inst/documents/publication/wcms\\_890761.pdf](https://www.ilo.org/wcmsp5/groups/public/---dgreports/---inst/documents/publication/wcms_890761.pdf).