

Ігор СКЛЯРОВ

кандидат технічних наук, доцент,
доцент кафедри металевих і дерев'яних конструкцій
Київського національного університету будівництва
і архітектури (Київ)
skliarov.io@knuba.edu.ua

Тетяна СКЛЯРОВА

асистент кафедри металевих і дерев'яних конструкцій
Київського національного університету будівництва
і архітектури (Київ)
skliarova.ts@knuba.edu.ua

ВИКОРИСТАННЯ ІНСТРУМЕНТІВ ГЕНЕРАТИВНОГО ШТУЧНОГО ІНТЕЛЕКТУ В АРХІТЕКТУРНІЙ ОСВІТІ

Постановка проблеми. Моделі генеративних нейромереж на основі прихованої дифузії (Latent Diffusion Models, або LDM) стали важливим досягненням останніх років у створенні графічного контенту та мають значний потенціал для використання в архітектурній освіті. Ця технологія може стати корисною для швидкого прототипування, надаючи архітекторам можливість швидко генерувати численні варіанти дизайну, експериментуючи зі стилями, формами та матеріалами, і пропонуючи інноваційні рішення, які складно уявити за допомогою традиційних підходів. LDM сприяють розробці реалістичних архітектурних візуалізацій, створенню фасадів, інтер'єрів та інших елементів будівель, а також дослідженню різноманітних дизайнерських концепцій [1]. Важливими напрямками є оптимізація об'ємно-планувальних рішень, аналіз освітлення та акустичних характеристик приміщень. Крім того, ці моделі можна адаптувати для роботи з конкретними архітектурними стилями, створюючи дизайн, що відповідає визначеним вимогам. Такий підхід має значення для розширення можливостей архітекторів та фахівців у будівельній галузі.

Аналіз досліджень та результати. Генеративні нейромережі, які є потужним інструментом штучного інтелекту, активно розвиваються та привертають значну увагу. Серед найбільш популярних моделей можна виокремити Stable Diffusion, Midjourney, DALL·E та Flux, кожна з яких має унікальні особливості. Наприклад, Midjourney відома своїм акцентом на художньому стилі та високою деталізацією створюваних зображень. DALL·E, розроблена OpenAI, вражає своєю здатністю генерувати реалістичні зображення, що

відповідають текстовим описам. Flux, хоч і є відносно новою моделлю, вже довела свою ефективність завдяки інтуїтивному управлінню процесом генерації та вирішенню типових проблем, таких як коректна кількість пальців або точна генерація тексту на зображеннях. Однак у контексті архітектури та будівництва найбільш корисною є Stable Diffusion, оскільки вона забезпечує високий рівень контролю над генерацією зображень і дозволяє адаптуватися до специфічних задач завдяки таким інструментам, як ControlNets та IPAdapters.

Основою технології прихованої дифузії [3, 4] є моделі глибокого навчання, спрямовані на створення реалістичних зображень. Цей метод базується на використанні ланцюгів Маркова [2], які навчаються за допомогою варіаційного висновку. Процес дифузії можна уявити як поступове додавання шуму до початкового зображення, аж поки воно не перетвориться на випадковий шум. Нейромережа навчається виконувати зворотну операцію – видаляти шум, поступово відновлюючи вигляд зображення. Вхідними даними для моделі є зашумлене зображення та рівень шуму, на основі яких мережа прогнозує вигляд попереднього етапу із меншою кількістю шуму.

Під час генерації зображення навчена модель починає з чистого шуму, поступово застосовуючи процес денойзингу. Цей підхід дозволяє крок за кроком відновлювати деталі зображення, рухаючись у прихованому просторі можливих варіантів до локального мінімуму функції, що оцінює відповідність бажаним властивостям. Різні початкові точки (рівні шуму) дають змогу отримати різноманітні результати, які відповідають заданим характеристикам.

У прихованому просторі пошук напрямку для кожного наступного «стрибка» контролюється завдяки застосуванню шарів перехресної уваги, інтегрованих у денойзингові шари UNet. Ці шари забезпечують мультимодальне управління процесом, дозволяючи поєднувати різні принципи опису зображень, як-от текстовий опис чи відповідність композиції.

Ефективним інструментом для управління стилем згенерованих зображень є так звані зображення-текстові адаптери. Це спеціалізовані нейромережеві моделі, які виділяють із зображення характеристики, що можуть бути описані текстом, і інтегрують ці властивості під час генерації разом з основним текстовим запитом. Такі характеристики можуть включати стиль, композицію або кольорову гаму, розпізнані моделлю. Зображення-текстові адаптери дозволяють використовувати кілька різних властивостей одночасно, що робить їх мультимодальними. На рис. 1 наведені приклади, де для

генерації застосовувалися текстовий опис, ескіз та стилеві характеристики, взяті з двох зразків.

Сучасні моделі нейромереж тренуються на величезних датасетах, що містять мільярди зображень. Однак навіть такий масштаб даних не гарантує наявності достатньої кількості прикладів певного типу зображень або їх якісної анотації. Як наслідок, моделі чудово справляються із генерацією популярних об'єктів, таких як Ейфелева вежа, вікторіанські будинки чи архітектура мінімалізму, але можуть зазнавати труднощів із відтворенням менш поширених споруд чи стилів.

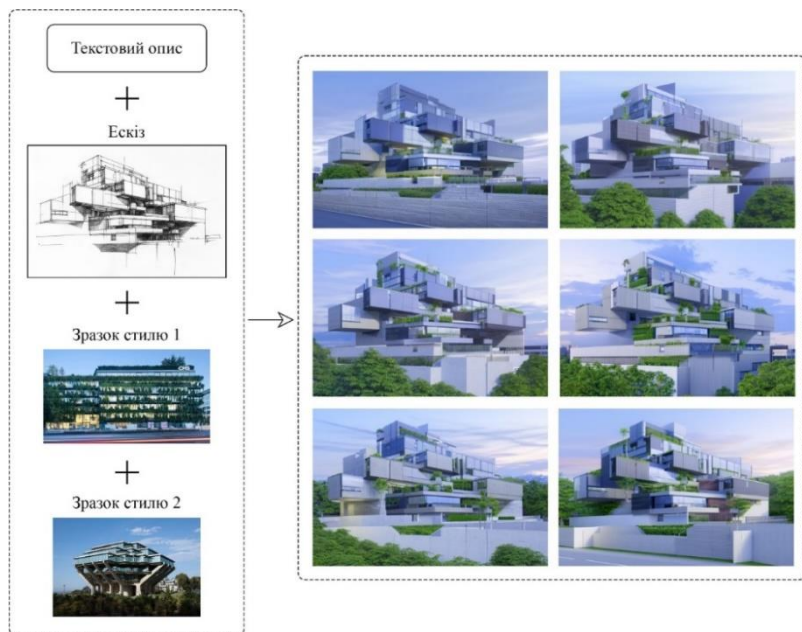


Рисунок 1 – Приклад генерації за текстовим описом, ескізом, а також двома референсними зображеннями для використання елементів стилю з них за допомогою *IP-adapt*

Для адаптації моделей до специфічних завдань широко застосовується стратегія донавчання. На відміну від початкового навчання, яке вимагає створення моделі з нуля, донавчання дозволяє скористатися вже наявними знаннями моделі, скоригувавши її для роботи з новими даними. Це значно знижує потребу в часі та обчислювальних ресурсах. Процес донавчання можна порівняти зі спеціалізацією: якщо базове навчання формує загальні знання, то донавчання спрямоване на їх поглиблення у вузькій сфері. Саме

розвиток у цьому напрямку вважається перспективним для подальшого вдосконалення нейромереж.

Висновки. Використання генеративних інструментів штучного інтелекту в архітектурному проектуванні відкриває нові горизонти для творчості та інновацій. Завдяки цим технологіям архітектори можуть швидше та ефективніше створювати оригінальні й адаптовані до сучасних вимог проекти. Проте їх успішність значною мірою залежить від якості та різноманітності навчальних даних. Подальші дослідження мають зосередитися на вдосконаленні моделей, здатних створювати більш реалістичні та деталізовані проекти, а також на їх інтеграції з іншими цифровими інструментами проектування.

Попри певні обмеження, дифузійні моделі вже сьогодні надають архітекторам потужний засіб для пошуку нових форм і концепцій, сприяючи появі інноваційних рішень у галузі архітектури. Водночас генеративні нейромережі не замінюють архітектора, а допомагають удосконалити, урізноманітнити й збагатити його роботу. Важливо наголосити, що остаточний вибір і відповідальність за проект завжди залишаються за архітектором.

Література:

1. Гетун Г. В., Куліков П. М., Плоский В. О., Чернишев Д. О. Конструкції будівель і споруд. Книга 2. Нежитлові будівлі. *Підручник*. К. : Рута, 2023.
2. Гетун Г. В., Буценко Ю. П., Лабжинський В. А., Баліна О. І., Безклубенко І. С., Соломін А. В. Прогнозування ситуацій та оптимізація прийняття рішень на основі скінченних ланцюгів Маркова в районах з промисловим забрудненням. *Стаття в н. т. збірнику «Опір матеріалів і теорія споруд»*, № 104. К. : 2020. С. 164–174. ISSN 2410-2547. DOI: 10.32347/2410-2547.2020.104.164-174
3. Сонг Й., Соул-Дікстейн Дж. Н., Кінгма Д. П., Кумар А., Ермон С. та Пул (2020). Генеративна модель на основі оцінки за допомогою стохастичних диференціальних рівнянь. arXiv, abs/2011.13456.
4. Ромбах Р., Блаттманн А., Лоренц Д., Ессер П. та Оммер Б. (2022). Синтез зображень високої роздільної здатності за допомогою латентних дифузійних моделей. Конференція IEEE/CVF з комп'ютерного зору та розпізнавання образів (CVPR), Новий Орлеан, Луїзіана, США, 2022, с. 10674–10685, DOI: 10.1109/CVPR52688.2022.01042