

## **ГЕНЕРАЦІЯ ЗОБРАЖЕННЯ ЗА ОПИСОМ ЗА ДОПОМОГОЮ ГЛИБОКОГО НАВЧАННЯ**

*Щербина Є. Д., Шпінарева І. М.*

Одеський національний університет ім. І. І. Мечникова

Люди перетворюють отриману інформацію на символи, такі як літери, слова, що дозволяє зберігати та обробляти її в майбутньому, а також передавати між людьми. Коли людина отримує позначення, вона відтворює інформацію, а не шукає серед збережених даних, як це роблять комп'ютери. Це дозволяє гнучко розуміти контекст середовища, а також використовувати вже знайомі патерни при обставинах, в яких вони не використовувались, що сприяє більш ефективному вивченню навколишнього середовища та спільній роботі над розв'язанням проблем. Контекстів може бути багато. Генерація зображень за описом покриває собою один із таких контекстів – текст в зображення, а також дозволяє більш детально вивчати, як розуміють наш світ штучний інтелект та методи машинного навчання.

Окрім розвитку галузі штучного інтелекту генерація зображень може використовуватися для вирішення прикладних задач.

Мистецтво. Методи для генерації зображень за текстом дозволяють звичайним людям займатися творчістю та виражати себе.

Маркетинг та реклама. Для якісного просування товарів та послуг потрібно багато контенту, який має бути цікавим для клієнтів. Створення зображень та іншої рекламної інформації є затратним з точки зору грошей та часу, тому сучасні маркетологи все активніше використовують методи генерації зображень за описом для створення прототипів або навіть повноцінних рекламних продуктів.

Синтетичні датасети. Зараз однією із найбільших проблем для розробників моделей комп'ютерного зору є недостатня кількість якісних датасетів для повноцінного навчання моделі. Генерація зображень за описом може допомогти створювати датасети відповідно до задач.

Сучасні моделі генерації зображень за описом доволі непогано виконують зазначені вище завдання, але мають свою недоліки:

створені зображення не завжди гарно відображають текстовий запит, а також мають значну кількість артефактів, які іноді вдається виправити за допомогою модифікації запиту, але неконсистентність в якості згенерованих зображень є найбільшою проблемою моделей;

великі затрати ресурсів обчислювальних машин на генерацію зображень та на навчання моделей.

Метою даної роботи є підвищення ефективності генерації зображень за описом шляхом розробки методу глибокого навчання. Під ефективністю розуміється покращення якості і швидкості виконання алгоритмів генерації зображення та обробки природної мови.

За основну модель нейронної мережі для вивченням обрано DALL-E-2 від OpenAI, оскільки вона займає лідируючі позиції на ринку генерації зображень та має відкриту документацію [1]. Для порівняння ефективності також використовується Stable Diffusion, Midjourney та Imagen [2-4].

Для обробки натуральної мови та зіставлення із зображеннями використовуються різні варіації моделі CLIP, що показала одні з найкращих показників розпізнавання об'єктів на зображеннях [5].

Для генерації зображень використовуються такі методи: diffusion models, generative adversarial networks (GANs). Ці методи показують найбільш якісні результати, а їхні комбінації активно досліджуються та інтегруються в DALL-E-2, Stable Diffusion, Midjourney, Imagen та багато інших [6].

Для навчання моделей використовуються датасети, які мають найбільшу кількість різноманітних зображень з описом до них, а саме: MS-COCO, Imagenet, MNIST Database, Google's Open Images та інші [7].

Для валідації отриманих результатів використовується, як експертна оцінка (людина оцінює якість згенерованих зображень), так і методи машинного навчання. Наприклад, маючи модель, яка може розпізнавати зображення та описувати об'єкти на них, можна порівняти точність на реальному датасеті та датасеті, що був створений за допомогою розробленої моделі генерації зображень за описом.

Отже, результатом цієї роботи є модель, що може генерувати якісні зображення за відносно невеликий час та потребує як можливо менше обчислювальних ресурсів.

### **Література**

1. Hierarchical Text-Conditional Image Generation with CLIP Latents [Електронний ресурс] – Режим доступу: <https://cdn.openai.com/papers/dall-e-2.pdf>
2. Stable Diffusion Online [Електронний ресурс] – Режим доступу: <https://stable-diffusionweb.com/>
3. Midjourney [Електронний ресурс] – Режим доступу: <https://www.midjourney.com/home/?callbackUrl=%2Fapp%2F>
4. Imagen. Unprecedented photorealism x deep level of language understanding. [Електронний ресурс] – Режим доступу: <https://imagen.research.google/>

5. Learning Transferable Visual Models From Natural Language Supervision [Електронний ресурс] – Режим доступу: <https://arxiv.org/pdf/2103.00020.pdf>
6. Diffusion Models Vs GANs: Which one to choose for Image Synthesis [Електронний ресурс] – Режим доступу: <https://analyticsindiamag.com/diffusion-models-vs-gans-which-one-to-choose-for-image-synthesis/>
7. Top Used Datasets for Text to Image Synthesis Models [Електронний ресурс] – Режим доступу: <https://analyticsindiamag.com/top-used-datasets-for-text-to-image-synthesis-models/>

## **ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ АДАПТУВАННЯ ЛЕКАЛ ОДЯГУ НА ОСНОВІ МЕТОДІВ ШТУЧНОГО ІНТЕЛЕКТУ**

*Жмакіна А.С., Малахов Є.В.*

Одеський національний університет імені І.І. Мечникова

*Ключові слова:* інформаційна технологія, швейна промисловість, лекала, штучний інтелект.

У ХХІ столітті однією з головних завдань швейної промисловості полягає в тому, щоб покращити точність, ефективність, продуктивність процесу виготовлення деталей одягу, які потім використовуються для виготовлення різноманітних виробів. Інформаційні технології та методи штучного інтелекту внесли свій вклад в переосмислення і розробку нового підходу в розробці одягу та продовжують активно розвиватись у цієї сфері [1]. Згідно з дослідженням McKinsey, використання методів штучного інтелекту в текстильній промисловості може призвести до підвищення ефективності виробництва на 15-20% та показник інвестування модних компанії у технології зросте від 1,6-1,8 до 3,0–3,5 відсотків до 2030 року [2].

Важливим та обов'язковим інструментом розробки сучасного одягу являється лекало, так як дозволяють створювати елементи нарядів в більшій кількості та забезпечуючи високу точність посадки та ідеальне поєднання різних елементів.

Лекало використовується в одному з основних етапів створювання швейного виробу – розкроювання основних деталей [3]. Цей етап включає в себе підготовку матеріалів до розкрою (якісна та кількісна перевірка матеріалу, визначення особливостей малюнка, лицьової та виворітної сторони, напрямки нитки основи, наявності дефектів) та процес розкрою. Допущення помилок на цієї стадії може привести до значних грошових витрат за рахунок марнування матеріалу та незадоволенням клієнта. Деякі причини тому: невчасне виявлення дефектів тканини та неоптимальне викладання лекал на тканині.