



GenAI

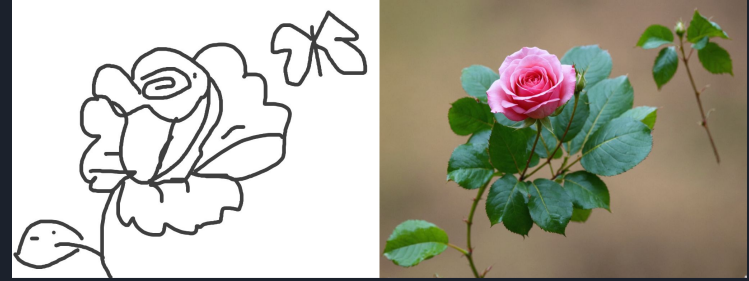
Sketch-to-Image

Yuri V

<https://sketch-image-lpuu.onrender.com/>

Motivation

- **Швидкий перехід від ідеї до результату:** Ескізи часто залишаються на папері через брак часу чи навичок для створення фотореалістичних зображень.
- **Зростаючі можливості генеративного ШІ:** Технології на кшталт WorqHat та OpenAI API дозволяють автоматизувати процес перетворення скетчу на готове зображення.
- **Зручність для дизайнерів і художників:** Спрощення робочого процесу — не потрібно перемикатися між різними програмами й інструментами.
- **Покриття потреб широкої аудиторії:** Як початківці, так і професіонали можуть однаково швидко втілювати свої ідеї.



Introduction

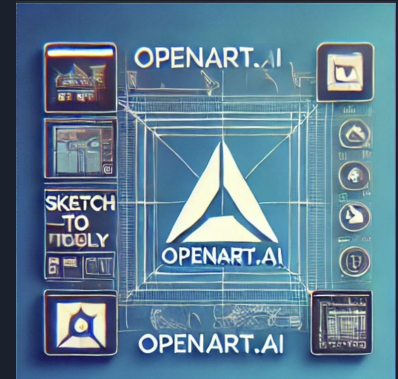
Ріст популярності генеративного ШІ: Поява численних веб-сервісів та дослідницьких робіт, присвячених перетворенню скетчів на фото.



Недоліки готових рішень: Обмежена персоналізація, платна модель або відсутність можливості контролювати метрики.

Існуючі інструменти:

- [OpenArt.ai](https://openart.ai)
- [Canva AI Sketch](https://canva.com/ai-sketch)
- [Adobe Firefly](https://adobe.com/firefly)
- [Freepik AI Sketch-to-Image](https://freepik.com/ai-sketch-to-image)





GenAI Sketch-to-Image: Детальний опис технічного підходу

Бекенд архітектура:

- Flask для серверної частини та API
- PIL для обробки зображень
- AWS S3 для зберігання файлів

AI/ML компоненти:

- WorqHat API: генерація описів та перетворення ескізів
- CLIP для семантичної схожості
- YOLOv8 для детекції об'єктів
- FID і SSIM для оцінки якості згенерованих зображень

Ключові алгоритми та процеси:

- Попередня обробка (нормалізація, масштабування)
- Генерація описів зображень
- Конвертація скетчів у реалістичні зображення
- Обчислення метрик (CLIP, Object Detection, FID, SSIM)

Фронтенд технології:

- HTML/CSS/JavaScript
- Canvas та DrawingTool для інтерактивного малювання

Інтеграції та API:

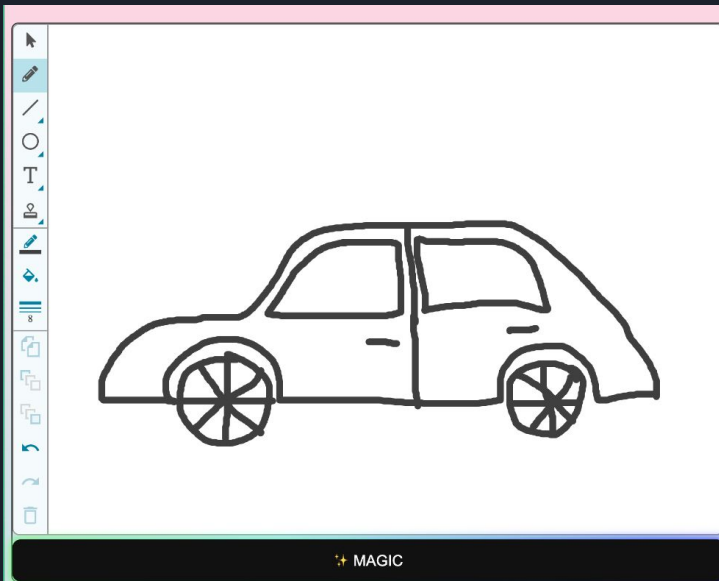
- WorqHat API для AI-функціональності
- AWS S3 API для зберігання і керування файлами

Безпека:

- Змінні середовища для ключів (API Keys)
- Базова валідація даних на бекенді

Демо: Демонстрація роботи системи

<https://sketch-image-lpuu.onrender.com/>



Okay, here is the prompt to generate a realistic image based on your sketch:

"A realistic, side-view image of a silver sedan with a rounded, retro-style design. The car has two doors divided by a visible central pillar, with two large windows on the side. The tires are black with silver rims and show eight spokes each. The car is parked on an asphalt road, casting a soft shadow to the ground with diffuse daylight. The body of the car should look smooth with reflections and highlights emphasizing its metallic sheen, with a glossy finish. The background should be plain and out-of-focus, emphasizing the car. The overall image should have a photo-realistic quality with attention to detail in lighting and reflections."

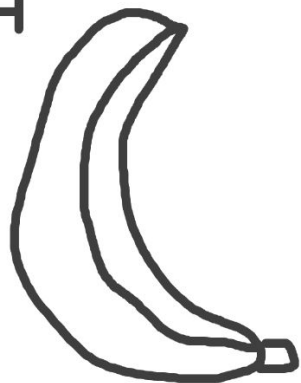
4399b410-a95f-4dae-b9c3-a72e95af2b62

Демо: Демонстрація роботи системи



Демо: Демонстрація роботи системи

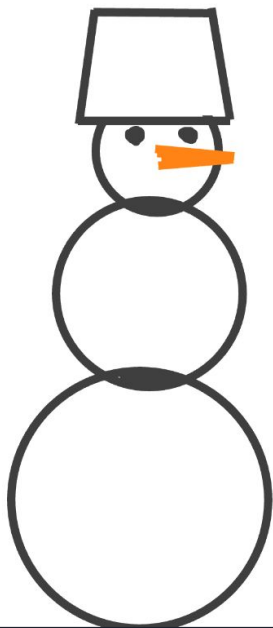
Б*Н*Н



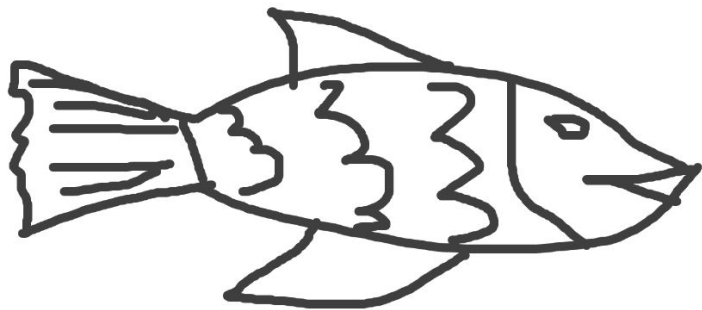
Б*Н*Н



Демо: Демонстрація роботи системи



Демо: Демонстрація роботи системи



Results: Analysis of Metrics and Performance



Метрики:

- CLIP Similarity: 25.28 (Original), 22.82 (Generated)
- Object Detection Match Score: 0.00 (Original), 0.00 (Generated)
- FID Score: 0.21 (Original), 0.30 (Generated)
- SSIM Metric: 0.78 (порівняння оригіналу згенерованого зображення)



Results: Analysis of Metrics and Performance

Ефективність:

- Середня затримка для аналізу ескізу: 2–3 сек
- Генерація реалістичного зображення: 25–30 сек
- Витрати на запити (WorqHat API) є передбачуваними, але можуть зростати з масштабуванням

Сильні сторони:

- Висока точність і якість згенерованих зображень
- Інтеграція кількох метрик (CLIP, FID, SSIM, Object Detection), що дає об'єктивну оцінку результату
- Зручний веб-інтерфейс і чіткий флоу користувача

Слабкі сторони:

- Час генерації може бути надто довгим для нетерплячих користувачів
- Object Detection Match Score показує низький показник (0%), що свідчить про потребу в більш точних методах або кращому передпрацюванні скетча
- Залежність від сторонніх API (WorqHat), що впливає на вартість та можливість кастомізації



Conclusions: Key Insights and Future Improvements

Висновки:

- Досягнуто мети: автоматизація перетворення скетчу на реалістичне зображення з використанням зовнішніх AI-сервісів.
- Реалізовано модуль оцінки якості (CLIP Similarity, Object Detection, FID, SSIM), що надає кількісні показники результатів.
- Користувацький досвід покращується за рахунок веб-інтерфейсу для швидкого завантаження ескізів і отримання зворотного зв'язку.

Майбутні напрямки покращення:

- **Оптимізація часу обробки:** дослідження альтернативних моделей або локального розгортання частини функціоналу для зменшення затримки.
- **Покращення детекції об'єктів:** удосконалення роботи YOLOv8 або інтеграція більш точних методів для підвищення Object Detection Match Score.
- **Більше контексту для описів:** додаткові поля чи підказки для користувача, щоб AI-генерація опису була точнішою та детальнішою.
- **Пакетна обробка:** додати можливість опрацьовувати відразу кілька скетчів для підвищення продуктивності.