

1. Вступ

Ціль проєкту: Створення автоенкодерної моделі для відновлення зображень облич із шумом, розмиттям, JPEG-артефактами та частковою втратою частин зображення. Модель навчається розрізняти чисте та пошкоджене зображення й реконструювати оригінал.

Мотивація: Задача відновлення якості облич важлива в медичному аналізі, безпеці, реставрації фотографій і відео. Автоенкодери забезпечують стабільність, швидкість і гнучкість при розгортанні в реальних системах.

2. Процес роботи

Збір і підготовка даних:

- Використано набір чистих зображень облич розміром 1024x1024, кількість - 10 000 зображень.
- Автоматично додано декілька типів шуму: Gaussian noise, розмиття (blur), JPEG-артефакти, часткове обрізання (masking). Також можуть одночасно додаватися два на одне зображення
- Завантаження даних організовано через PyTorch DataLoader, з урахуванням вкладених директорій.
- Dataset source - <https://github.com/NVlabs/ffhq-dataset>

Навчання моделі:

- Реалізовано симетричну архітектуру автоенкодера з Conv2D блоками, batch normalization та ReLU активацією.
- Використано функцію втрат MSE (mean squared error).
- Додано підтримку збереження чекпоінтів після кожної епохи.
- Для кожної епохи логувалась тривалість тренування у форматі hh:mm:ss.

Інтеграція:

- Створено модуль для генерації денойзингових результатів.
 - Прототип
-

3. Виклики та їх вирішення

Технічні проблеми:

- Великі зображення (1024x1024) призводили до браку GPU-пам'яті.
- Структура директорій із вкладеними папками потребувала обробки при завантаженні датасету.

Вирішення:

- Оптимізовано batch_size.
 - Реалізовано функцію обходу вкладених директорій.
 - Можливість запуску на CPU у fallback-режимі.
 - Використання ітератора для зменшення використання ОЗУ при великого розміру датасета
-

4. Результати

Метрики:

- PSNR (Peak Signal-to-Noise Ratio): ~35 дБ
- SSIM (Structural Similarity Index): ~0.90

Порівняння:

- Autoencoder показав вищу якість порівняно з базовими методами фільтрації або Gaussian denoising.
 - Обробка із partial masking краще підтримується автоенкодером ніж простими CNN.
-

5. Висновки

Оцінка:

- Реалізований автоенкодер якісно прибирає шум із зображень 1024x1024 та здатен реконструювати навіть частково втрачені ділянки.
- Простота моделі робить її придатною для мобільного або embedded-застосування.
- Оцінка по конкретним шумам
 - Gaussian noise - 8/10
 - розмиття (blur) - 8/10
 - JPEG-артефакти - 8/10
 - Часткове обрізання (masking) - 2/10 (модель інколи сильно помиляється, приклад https://github.com/pylypen/robotdreams_gen_ai/blob/main/homework_17/data_sample/sample_epoch_498.png)
- Модель не дуже гарно справляється з шумом по типу Masking, можна помітити що модель старається зробити все можливо але все одно частина картини має один колір, іноді модель плутає чорні зони з шумом Masking й таким чином тільки погіршує якість.
- Після першої епохи картинка була розмита але краще ніж при першій епохи GAN моделі
- Починаючи з 100 епохи з деякими шумами модель дуже гарно себе показала

- Взагалом зображення трохи розмите але все одно показує доволі гарні результати
- Приклад згенерованих картинок знаходиться тут - https://github.com/pylypen/robotdreams_gen_ai/tree/main/homework_17/data_sample

Майбутні покращення:

- Додавання attention-механізмів для покращення локального відновлення.
- Збільшення кількості епох для навчання
- Збільшення датасету в декілька разів
- Перехід до багаторівневих автоенкодерів (VAE, U-Net) які можуть значно покращити якість, але деякі моделі будуть навчатися більше часу.