

4. European Commission. Ethics Guidelines for Trustworthy AI. URL: <https://digital-strategy.ec.europa.eu> (дата звернення: 18.03.2026).
5. IBM. AI Ethics. URL: <https://www.ibm.com/artificial-intelligence/ethics> (дата звернення: 18.03.2026).

УДК 004.8:004.932

РОЗРОБКА ФОТОРЕАЛІСТИЧНОГО ЦИФРОВОГО АВАТАРА З ВИКОРИСТАННЯМ ГЕНЕРАТИВНИХ МОДЕЛЕЙ ШТУЧНОГО ІНТЕЛЕКТУ

*Компанієць Ю. М.
yuliannakompaniets@gmail.com
Черкаський державний фаховий бізнес-коледж
Ночевнов Д. П.
м. Черкаси, Україна*

Цифрові аватари сьогодні активно використовуються у відеоконференціях, онлайн-освіті, іграх, медичних тренажерах і середовищах розширеної реальності, а потреба у якісних і реалістичних аватарах постійно зростає. Якщо раніше їх створення ґрунтувалося переважно на трудомісткому ручному 3D-моделюванні, то поява генеративних моделей штучного інтелекту дала змогу автоматично синтезувати фотореалістичні обличчя та змінювати їх параметри на основі лише кількох вхідних фотографій.

GAN стали першою широко застосованою архітектурою для генерації реалістичних зображень, а StyleGAN – одним із найвідоміших рішень для синтезу облич із високою якістю та керованістю стилем [1]. Variational Autoencoder (VAE) реалізують інший підхід через латентний простір, поступаючись GAN у деталізації, але забезпечуючи стабільніше навчання та зручну інтерполяцію. Отже, GAN і VAE можна розглядати як недифузійні генеративні підходи, що становлять теоретичну основу для порівняння із сучасними дифузійними моделями. Дифузійні моделі виконують генерацію шляхом поетапного видалення шуму, а Latent Diffusion Model, на якій побудований Stable Diffusion, переносить цей процес у стиснутий латентний простір, що пришвидшує обчислення [2]. Додатково ControlNet надає таким

моделям можливість точно керувати просторовою структурою зображення, зокрема позою, контурами обличчя та напрямком погляду, що робить їх ефективним інструментом для синтезу фотореалістичних аватарів [3].

Методи повноцінної 3D-реконструкції, зокрема підходи на основі параметричних 3D-моделей обличчя, таких як FLAME, і сучасні методи рендерингу, зокрема 3D Gaussian Splatting, потребують значних обчислювальних ресурсів і великого обсягу вхідних даних. Внаслідок цього їх використання в практичних сценаріях є ускладненим. З огляду на це в роботі обрано підхід на основі 2D-генеративних моделей, який забезпечує отримання фотореалістичних результатів навіть за обмеженої кількості вхідних зображень.

У практичній частині роботи обробку вхідних фотографій реалізовано із використанням програми в середовищі Google Colaboratory, тоді як генерацію фотореалістичного аватара та отримання фінального результату виконано за допомогою сервісу Ip-Adapter-FaceID, доступного як Hugging Face Space [6]. Процес обробки можна умовно поділити на чотири основні етапи.

На першому етапі виконується детекція обличчя та його вирівнювання. Бібліотека OpenCV у поєднанні з детектором на основі MTCNN або RetinaFace виявляє обличчя на вхідному зображенні, визначає ключові точки (очі, ніс, кути рота) та виконує афінне перетворення, приводячи обличчя до стандартного положення, як на рисунку 1.

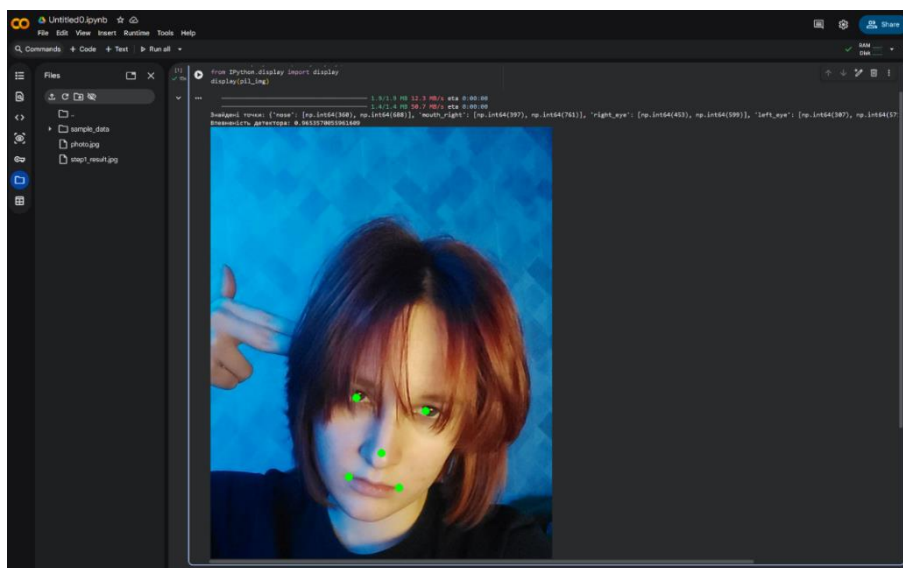


Рисунок 1 – Детекція обличчя

На другому етапі з підготовленого зображення витягується вектор ідентичності за допомогою попередньо навченої моделі розпізнавання облич – зокрема, ArcFace. Цей вектор містить компактне числове представлення унікальних рис обличчя, яке надалі використовується як умова при генерації, щоб забезпечити схожість аватара з реальною особою (див. рисунок 2).

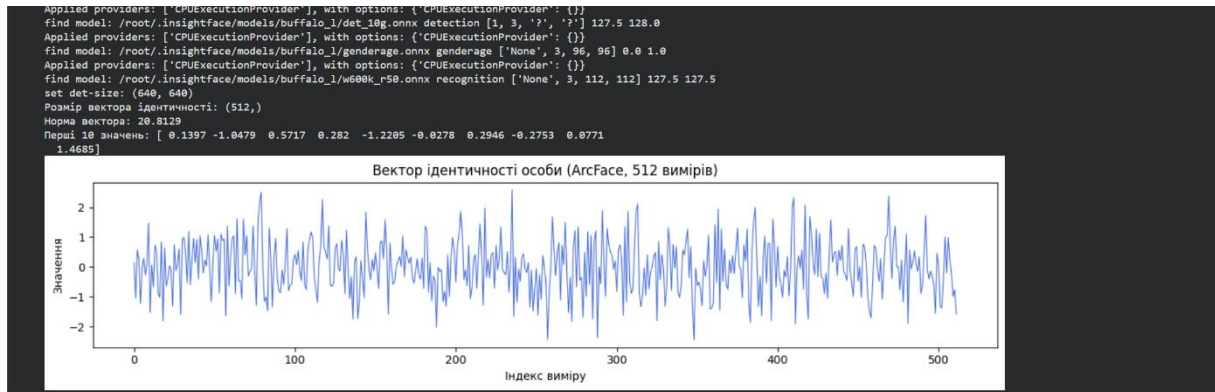


Рисунок 2 – Вектор ідентичності

На третьому етапі відбувається безпосередня генерація аватара за допомогою Stable Diffusion із підключеним ControlNet, який отримує умовне зображення у вигляді карти глибини або скелету обличчя, тоді як текстовий промпт задає бажаний стиль, освітлення та ракурс (рисунки 3, 4). Вектор ідентичності інтегрується через IP-Adapter, запропонований в [5], що дає змогу передавати риси обличчя без донавчання моделі для кожної нової особи.

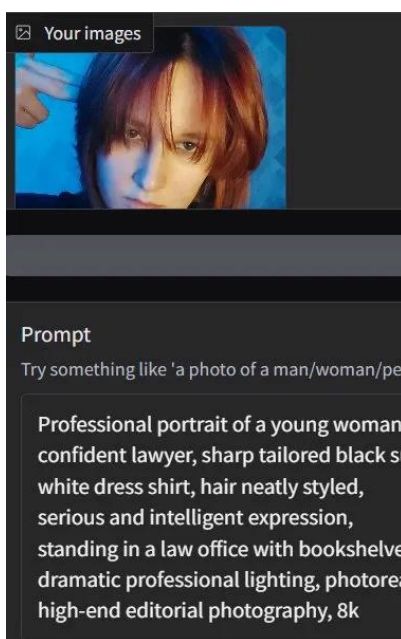


Рисунок 3 – Вхідні фото та промт



Рисунок 4 – Згенерований результат

На четвертому етапі виконується постобробка: підвищення роздільності за допомогою ESRGAN, корекція кольорів, а також опціональне зменшення артефактів за допомогою GFPGAN.

Використаний підхід демонструє практично прийнятну якість за метриками FID, LPIPS та ID Similarity, який отримав позитивну суб'єктивну оцінку від учасників тестування. Технологічний стек реалізації включає Python 3.11, PyTorch 2.x, Hugging Face Diffusers, OpenCV і CUDA, а всі компоненти упаковано у Docker-образ для спрощення розгортання.

Серед ключових напрямків подальшого розвитку – адаптація системи до генерації у реальному часі, розширення функціональності для підтримки анімації аватара на основі мімічного відеосигналу, а також дослідження мультимодальних підходів, що поєднують зображення, звук і текст для ще більш природної взаємодії.

Список використаних джерел

1. Karras T., Laine S., Aila T. A Style-Based Generator Architecture for Generative Adversarial Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2021. Vol. 43, No. 12. P. 4217–4228.

2. Rombach R., Blattmann A., Lorenz D., Esser P., Ommer B. High-Resolution Image Synthesis with Latent Diffusion Models. Proceedings of the IEEE/CVF CVPR. 2022. P. 10684–10695.
3. Zhang L., Rao A., Agrawala M. Adding Conditional Control to Text-to-Image Diffusion Models. Proceedings of the IEEE/CVF ICCV. 2023. P. 3836–3847.
4. Kerbl B., Kopanas G., Leimkühler T., Drettakis G. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. ACM Transactions on Graphics. 2023. Vol. 42, No. 4. Article 139.
5. Ye H., Zhang J., Liu S., Han X., Yang W. IP-Adapter: Text Compatible Image Prompt Adapter for Text-to-Image Diffusion Models. arXiv preprint. 2023. arXiv:2308.06721. URL: <https://doi.org/10.48550/arXiv.2308.06721>, (дата звернення: 20.03.2026).
6. IP-Adapter-FaceID Demo. URL: <https://multimodalart-ip-adapter-faceid.hf.space> (дата звернення: 20.03.2026).

УДК 004.891

ШТУЧНИЙ ІНТЕЛЕКТ ЯК ІНСТРУМЕНТ ОПТИМІЗАЦІЇ АЛГОРИТМІВ ОБРОБКИ ДАНИХ

Кононенко С.А.

vinsofia43@gmail.com

Черкаський державний фаховий бізнес-коледж

Немченко В.Ю.

м. Черкаси, Україна

Сучасний етап розвитку глобальної інформаційної інфраструктури характеризується стрімким переходом від кількісного накопичення даних до необхідності їх якісної та надшвидкої обробки. Традиційні детерміновані алгоритми, що десятиліттями складали фундамент комп'ютерних наук, у сучасних умовах Big Data починають демонструвати обмеженість, зумовлену статичністю їхньої внутрішньої логіки. Проблема полягає в тому, що класичний алгоритм не враховує специфіку та внутрішню структуру конкретного набору даних, орієнтуючись лише на найгірший або середній теоретичний випадок