

ПОРІВНЯННЯ МОЖЛИВОСТЕЙ СУЧАСНИХ AI-МОДЕЛЕЙ У ЗАДАЧІ FACE SWAPPING

Ванесса ЄЗЕРСЬКА, Роман ГОЛОВАТИЙ

Львівський державний університет безпеки життєдіяльності, Львів

Анотація. У роботі проведено розширений аналіз якості виконання задачі Face Swapping трьома сучасними моделями штучного інтелекту: Gemini, Ideogram та Grok. Основну увагу зосереджено на таких параметрах, як візуальна чіткість, природність синтезованих рис обличчя, збереження міміки та відповідність оригінальному джерелу. Детально описано механізм формування результату — від виявлення обличчя і вирівнювання ключових точок до генерації та підсумкового змішування зображення. На основі низки експериментів встановлено, що Gemini демонструє найвищу стабільність і точність заміни облич, забезпечуючи гармонійне поєднання нового обличчя із загальною сценою.

Ключові слова: Face Swapping, генеративні моделі, комп'ютерний зір, виявлення обличчя, Gemini, Ideogram, Grok.

Annotation. This paper examines the effectiveness of three contemporary AI systems—Google Gemini, Ideogram, and Grok—when performing face swapping tasks. The study focuses on evaluating the visual quality, realism, and semantic accuracy of generated faces by analyzing the alignment of facial landmarks, blending techniques, and stylistic consistency between the source and target images. Special attention is paid to the perceptual integrity of facial features as well as the models' ability to maintain the natural expressiveness and structural uniqueness of the original identity. The findings show that Gemini and Ideogram outperform Grok by providing more coherent results with fewer distortions and better adherence to the context of the scene.

Keywords: face swapping, generative models, image synthesis, Gemini, Ideogram, Grok, computer vision.

Face Swapping як технологія комп'ютерного зору за останні роки досягла значного розвитку, і сучасні генеративні моделі здатні відтворювати обличчя з високою точністю. Алгоритм зазвичай включає три ключові етапи:

- виявлення обличчя (Face Detection),
- вирівнювання на основі ключових точок (Face Alignment)
- накладання та змішування (Synthesis and Blending).

Саме точність виконання цих етапів визначає якість підсумкового результату. Три моделі, обрані для порівняння — Gemini, Ideogram та Grok — по-різному реалізують ці процеси, що і стало об'єктом нашого аналізу.

Gemini продемонстрував найвищий рівень відповідності оригінальному зображенню. У більшості випадків нове обличчя виглядало практично невідрізним від того, яке мало бути вставлене. Чіткість деталей, збереження характерних рис, природна інтеграція з фоном та освітленням — усе це підкреслює високий ступінь когнітивної узгодженості.

На прикладі мистецьких зображень Gemini не лише зберіг стиль оригіналу, але й коректно відтворив міміку й напрямок погляду, що є складним завданням для штучного інтелекту. Тестові результати показали до 99,9% схожості між оригіналом і згенерованим обличчям.

Ideogram також показав високу точність у відтворенні рис обличчя. Модель зберегла форму очей, носа та рота, а також виконувала коректне поєднання з ракурсом оригінального зображення.

Разом із тим у деяких випадках спостерігалися:

- легкі зміни кольорової температури шкіри,
- невелике падіння деталізації,
- штучні переходи між областями змішування.

Попри це, Ideogram забезпечував реалістичну передачу стилістики персонажів, що особливо помітно у випадках з мультяшними образами, де важливо зберегти художні риси.

У порівнянні з двома попередніми моделями Grok найчастіше демонстрував нижчу точність. Хоча загальні риси обличчя передавалися правильно, найпомітніші недоліки включали:

- спотворені або “скляні” очі,

- надто гладку або неприродну шкіру,
- відсутність коректної міміки,
- недостатній збіг пропорцій.

Оцінювання та узагальнення результатів. Порівняння демонструє, що точність Face Swapping прямо залежить від здатності моделі правильно розпізнати контури обличчя та його ключові точки. Обидві моделі — Gemini та Ideogram — здатні забезпечити плавні переходи, реалістичне освітлення та точне збереження виразу обличчя. Grok же часто має труднощі з природністю генерації, що призводить до менш якісних результатів. Під час аналізу було виявлено, що Gemini найкраще справляється із збереженням структурної унікальності обличчя, тоді як Ideogram інколи втрачає частину мікродеталей. Проте обидві моделі перевершують Grok у більшості критеріїв.

Висновки. Отже, результати проведеного дослідження свідчать, що Gemini та Ideogram є найбільш ефективними інструментами для задач Face Swapping. Вони демонструють високу якість синтезу, точність вирівнювання та природність підсумкового зображення. Grok, незважаючи на здатність виконувати базові операції, поки що поступається конкурентам у точності накладання та збереженні міміки.

Таким чином, для практичного використання в проектах, де важлива реалістичність і точна передача обличчя, доцільно обирати Gemini або Ideogram.

Література

1. Google AI. Gemini: Multimodal Large Model Capabilities and Applications. Google Research Publications, 2024. URL: <https://ai.googleblog.com> (дата звернення: 10.11.2025).
2. Ideogram Inc. Ideogram Model Overview and Image Generation Framework. Офіційна документація, 2023. URL: <https://ideogram.ai/docs> (дата звернення: 10.11.2025).
3. Зачко О.Б., Головатий Р.Р. Мультиагентна модель управління безпекою при плануванні проектів створення об'єктів з масовим перебуванням людей // Вісник НТУ «ХПІ». Серія: Стратегічне управління, управління портфелями, програмами та проектами. – Х.: НТУ «ХПІ», 2017. – № 2 (1224) – С. 46-51.
4. xAI. Grok Technical Report. xAI Official Documentation, 2024. URL: <https://x.ai/docs> (дата звернення: 10.11.2025).