

ОСНОВНЫЕ ЭТАПЫ СОЗДАНИЯ ДИПФЕЙКОВ С ИСПОЛЬЗОВАНИЕМ FACEFUSION 3

Барковец Р. С.

Научный руководитель – Ковалева И. Л., к.т.н., доцент

Современные технологии искусственного интеллекта открывают новые горизонты в цифровой обработке медиаконтента, и одним из самых популярных и активно развивающихся направлений в этой области является генерация дипфейков – автоматизированных замен лиц в изображениях и видеозаписях. В образовательной сфере такие технологии могут использоваться для создания интерактивных учебных материалов, визуализации исторических событий и даже моделирования различных сценариев в рамках психологических или коммуникативных тренингов.

Для генерации дипфейков можно использовать различные решения. Open-source платформа FaceFusion обладает рядом преимуществ по сравнению с DeepFaceLab, FaceSwap и Reface. По качеству замены лиц FaceFusion уступает лишь DeepFaceLab, который считается «золотым стандартом» в области генерации дипфейков, но при этом он требует значительно больше вычислительных ресурсов и сложен в настройке. FaceFusion, напротив, предлагает хороший баланс между качеством и производительностью, а также отличается гибкостью настроек благодаря открытому исходному коду.

В основе FaceFusion лежат модели, архитектура которых относится к генеративно-состязательным нейросетям (GAN) или свёрточным нейросетям (CNN). Ключевым компонентом системы является InsightFace — ещё один open-source инструмент для распознавания и замены лиц [1,2].

На официальном git-hub репозитории есть ссылка на документацию, где сказано, как происходит установка [1]. Из заметных плюсов можно отметить тот факт, что предусмотрены дополнительные надстройки, позволяющие использовать процессор, cuda ядра на видеокартах NVIDIA и TensorRT ядра, которые можно применять на видеокартах с припиской RTX. После загрузки можно запускать программу с помощью GUI в браузере или CLI при помощи команд.

Если нет необходимости работать в браузере «руками», то это всегда можно сделать из командной строки, обращаясь из неё к python-скрипту, который и запускает программу. Чтобы сделать демонстрационное дипфейк-видео приемлемого качества, можно ввести команду, пример которой приведен на рисунке 1.

```
(facefusion3) d:\ff3\facefusion>python facefusion.py headless-run --source-paths "D:\dataset\people\man1.png" --target-path "D:\dataset\video\5s-going-to-river.mp4" --output-path "D:\dataset\output\output.mp4" --face-detector-size "640x640" --processors face_swapper face_enhancer age_modifier expression_restorer --execution-providers "cuda" --execution-thread-count "32" --execution-queue-count "1"
[FACEFUSION.CORE] Processing step 1 of 1
Downloading: 50%|=====| 164M/325M [00:18<00:41, 4.02MB/s, download_provi
```

Рис.1. Пример команды для обращения к FaceFusion.

С помощью этой команды можно создавать дипфейки с разрешением исходного видео (по умолчанию), используя все доступные cuda ядра видеокарты для максимальной скорости обработки. Также используются основные процессоры, которые достаточно быстро генерируют видео. Превью для итогового результата отображается в правой верхней части экрана. Это позволяет в реальном времени оценить результаты и подогнать настройки, если результат получится неудовлетворительным.

Процесс обработки изображений в FaceFusion начинается с детекции лиц. Для поиска лиц по умолчанию используется модель YOLOv8. С ее помощью выполняется не только поиск лица на изображении, но и выделение ключевых точек ждя точного определения положения и ориентации лица. Затем энкодер преобразует лицо в векторное представление — выполняется эмбединг, который кодирует уникальные черты лица в числовой форме. Следующий этап — замена лица с помощью GAN-модели, например, inswapper_128_fp_16. Этот алгоритм генерирует новое лицо, сохраняя при этом мимику, освещение и другие характеристики исходного изображения. Завершающий этап — блендинг, при котором заменённое лицо естественным образом интегрируется в исходное изображение или видео, сглаживаются границы и корректируются цветовые переходы. Помимо этих трёх этапов в воркфлоу обработки могут включаться модели, которые отвечают за «улучшение» лица, «улучшения» кадра, более точный перенос мимики и возраста. Схема работы представлена на рисунке 2.

Технология дипфейков открывает широкие возможности для студентов и преподавателей. В учебном процессе её можно использовать для создания интерактивных материалов, например, «оживления» исторических личностей в лекциях по истории или литературе. Это позволяет сделать обучение более наглядным и вовлекающим.

Студенты, изучающие компьютерное зрение и машинное обучение, могут применять FaceFusion для генерации тренировочных данных или экспериментов с нейросетевыми моделями. В медиаобразовании технология полезна для освоения навыков видеомонтажа и работы с искусственным интеллектом.

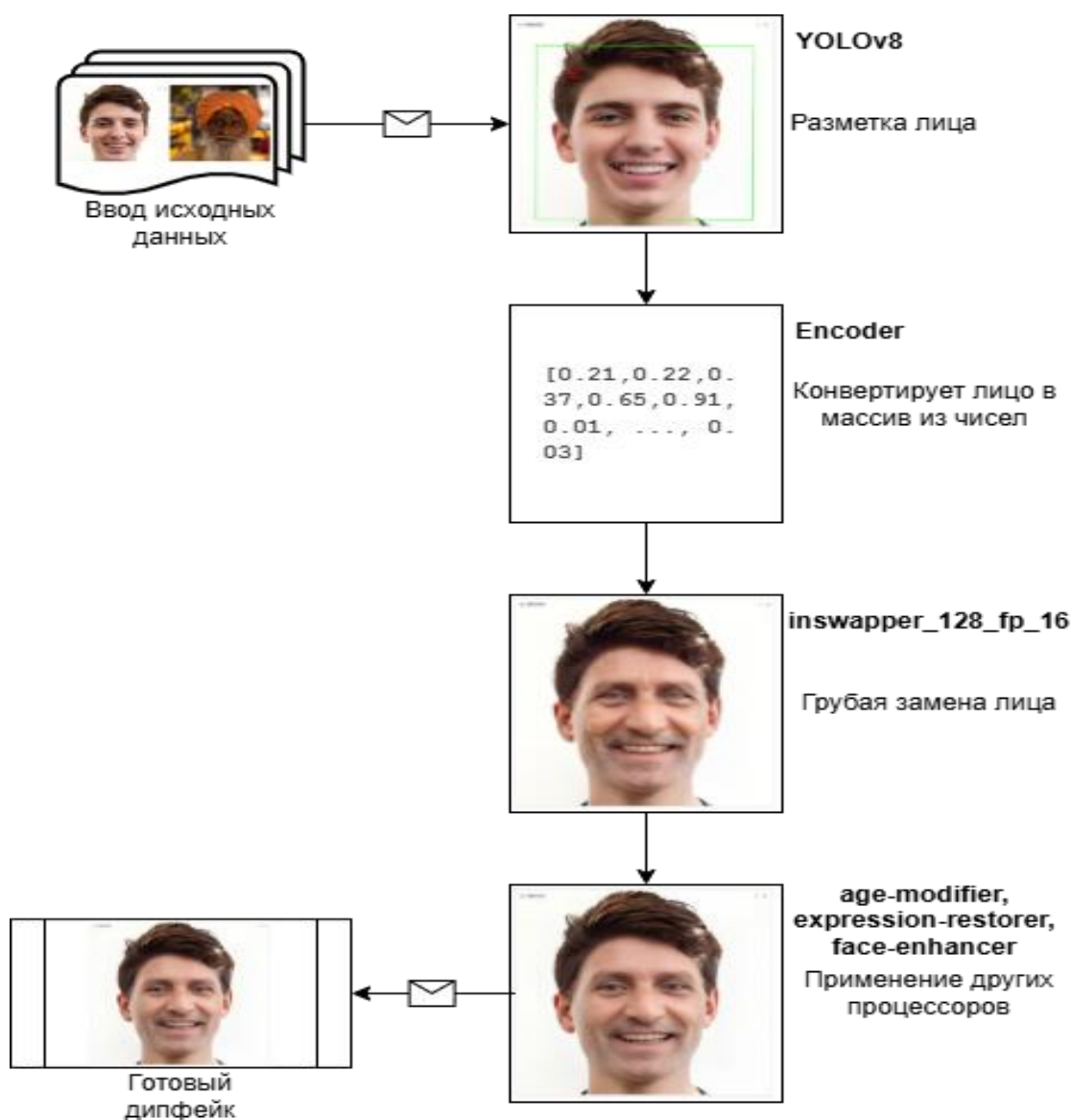


Рис.2. Воркфлоу работы FaceFusion.

Преподаватели могут использовать дипфейки для моделирования различных ситуаций, например, в рамках курсов по психологии, где студенты анализируют эмоции и невербальные сигналы. Также технология может быть полезна при создании учебных видеороликов, где вместо реального преподавателя используется цифровой аватар.

Литература

- 1 Репозиторий Facefusion. [электронный ресурс] URL: <https://github.com/facefusion/facefusion> (Дата обращения: 08.05.2025).
- 2 Документация Facefusion. [электронный ресурс] URL: <https://docs.facefusion.io> (Дата обращения: 08.05.2025).