

## Ученые из AIRI улучшили ИИ-стилизацию фотографий и выложили алгоритм в открытый доступ

Современные инструменты стилизации фотографий значительно продвинулись в качестве за последние несколько лет. Во многом прогресс ИИ-редактуры фото обязан развитию генеративно-состязательных архитектур в задаче контролируемого изменения изображений лиц. Пользователи уже могут примерять прически, добавлять к селфи очки, усы и другие элементы. Однако существующие подходы сталкиваются с проблемой: при изменении изображения или исходное лицо редактируется неправильно, например, добавляются лишние детали в тех местах, которые не подвергались редактированию, или теряются детали исходного изображения, такие как макияж, татуировки, аксессуары и т.п.

Команда «Контролируемый генеративный ИИ» лаборатории FusionBrain Института AIRI занимается задачами персонализации при создании мультимедийного контента. В актуальном исследовании, принятом на ведущую конференцию по компьютерному зрению CVPR-2024 (A\*), ученые представили StyleFeatureEditor (SFE) – фреймворк, который позволяет быстро редактировать фотографии, сохраняя качество реконструкции и хорошую редактируемость результата.

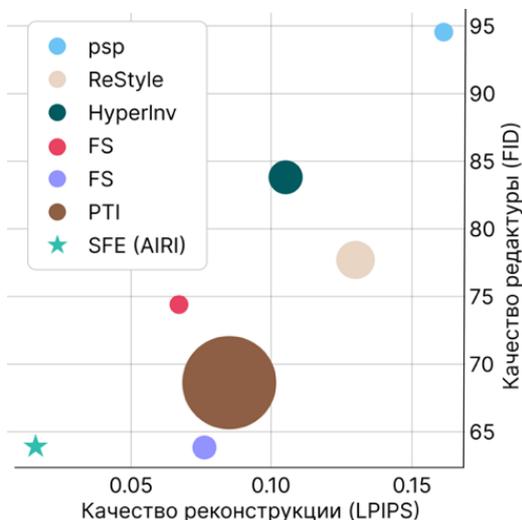


Современные подходы к инверсии изображений – поиску внутренних элементов, которые можно редактировать – делятся на две большие группы: методы, основанные на оптимизации, и методы на базе энкодеров. Оптимизаторы запускают свой процесс для каждой картинке заново, а энкодеры – могут работать с любой картинкой. Проблема энкодеров заключается в том, что сложнее обучить одну нейросеть для всех картинок, чем для каждой картинке запускать процесс оптимизации. Разница между энкодерами и методами оптимизации предоставляет выбор: хотим мы получить быстрый, но не очень точный метод, или медленный, но довольно хороший. И в том, и в другом случае остается риск изменения исходного изображения.

Коллектив ученых из AIRI решил эту проблему благодаря новой процедуре обучения. На первом этапе нейросеть учится реконструировать саму фотографию без какого-либо редактирования – она учится как правильно вынуть из изображения нужные ей признаки. На втором этапе, зафиксировав все знания с первого этапа, нейросеть учится, как правильно редактировать выученные признаки. В этом ей помогают предыдущие методы – с помощью них можно собрать примеры различных редактирований простых изображений, и показать

эти примеры нашей нейросети. Таким образом, мы показываем нашему методу, как реконструировать сложные (реальные) фотографии, как должны редактироваться простые (синтетические), и нейросеть сама начинает понимать, как должны редактироваться сложные фото.

Используя полученные на простых примерах знания, модель обучилась работать со сложными кадрами реальных лиц.



StyleFeatureEditor сравнили с 11 наиболее известными аналогами на примере портретов, включая сложные для анализа снимки с гримом и тату, и автомобилей. Подход не только успешно справляется с добавлением новых стиливых элементов, но и бережно относится к деталям исходных изображений. Качество реконструкции было улучшено в 4 раза по сравнению с SOTA-моделями по стандартной метрике LPIPS

*«Предложенный метод не только позволяет восстановить более мелкие детали изображения, но и обеспечивает их устойчивость при редактировании. Мы также представляем новый пайплайн обучения, специально разработанный для задачи редактирования изображений в латентном пространстве. В сравнении работы с современными подходами мы отметили лучшее качество реконструкции и способность редактировать даже сложные примеры. Важно учитывать, что представленный метод в первую очередь сфокусирован на улучшении метрик и внутренних систем ИИ-генераторов, потому что будет полезен разработчикам и профессиональным дизайнерам, работающим с инструментами на основе генеративного ИИ», — отметил директор лаборатории FusionBrain Института AIRI Андрей Кузнецов*

Найти выложенную в открытый доступ модель и применить ее для своих сервисов можно [репозитории](#) команды на GitHub.

---

**Вопросы:** [pr@airi.net](mailto:pr@airi.net)

*Научно-исследовательский Институт искусственного интеллекта AIRI — автономная некоммерческая организация, занимающаяся фундаментальными и прикладными исследованиями в области искусственного интеллекта. На сегодняшний день более 150 научных сотрудников AIRI задействовано в исследовательских проектах Института для работы совместно с глобальным сообществом разработчиков, академическими и индустриальными партнерами.*