

ГЕНЕРАЦИЯ КАРАНДАШНОГО ЭСКИЗА ПО ФОТОГРАФИИ С ПОМОЩЬЮ НЕЙРОСЕТИ

Карякин М.М.

Южный Федеральный Университет, Россия, 344090, г. Ростов-на-Дону, ул. Мильчакова 8а, +79508652364, mkariakin@sfnedu.ru

Задачи стилистического преобразования изображений (Image-to-Image translation) представляют значительный интерес как с теоретической, так и с прикладной точек зрения. Среди существующих подходов архитектура CycleGAN [1] демонстрирует ключевое преимущество, позволяя обучаться на непарных наборах данных, что критически важно задач, где создание множества строго соответствующих пар «исходное изображение – целевое изображение» затруднительно.

Реализована и оптимизирована модель на основе CycleGAN [1] для решения задачи генерации карандашных эскизов по фотографиям портретов. В качестве исходных данных использованы датасеты CUNHK Face Sketch Database (эскизы) и полутонная версия CelebAMask-HQ (фотографии). Для достижения устойчивого обучения и стилизации текстуры в модели применены ключевые механизмы CycleGAN: состязательные потери, потери циклической и идентификационной консистентности, а также патч-дискриминатор (PatchGAN [2]).

В процессе обучения исходная конфигурация модели столкнулась с генерацией структурных артефактов. Для их устранения была проведена оптимизация архитектуры, направленная на снижение избыточной сложности сети и повышение стабильности обучения. Внесённые изменения позволили устранить критические дефекты выходных изображений, сохранив при этом высокое качество стилизации и циклической согласованности. Эффективность оптимизации подтверждается снижением метрики FID с 46.1715 до 42.8885 (Fréchet inception distance [3]), что указывает на близость сгенерированного распределения вектора признаков к целевому.

Литература.

1. Zhu, J., Park, T., Isola, P., Efros, A. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks // 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy, 2017. Pp. 2242-2251.
2. Isola, P., Zhu, J., Zhou, T., Efros, A. Image-to-Image Translation with Conditional Adversarial Networks // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA, 2017. Pp. 5967-5976.
3. Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., GANs Trained by a Two Time Scale Update Rule Converge to a Local Nash Equilibrium // Neural Information Processing Systems (NIPS), Long Beach, California, 2017. Vol. 30. Pp. 6626-6637