

## ИСПОЛЬЗОВАНИЕ МЕТРИКИ РАСПОЗНАВАНИЯ ЛИЦ ДЛЯ ОЦЕНКИ ВИЗУАЛЬНОЙ ИДЕНТИЧНОСТИ ПЕРСОНАЖЕЙ В ГЕНЕРАТИВНОЙ ГРАФИКЕ

Дашинамжилов Б. Б.<sup>1</sup>, Меженин А. В.<sup>1</sup>

Научный руководитель – канд. техн. наук, доцент Меженин А. В.<sup>1</sup>

<sup>1</sup>Университет ИТМО

312259@edu.itmo.ru

### Введение

В задачах генеративной графики всё чаще возникает необходимость сохранения идентичности персонажа — при изменении ракурса, освещения, стиля или текстового описания. Существуют различные zero-shot методы [1], позволяющие интегрировать лицо человека в генерацию используя лишь одну фотографию, без дополнительного обучения. Однако наличие методов сохранения идентичности само по себе не решает ключевую методологическую проблему — оценку качества их работы.

Субъективная оценка требует привлечения большого числа участников, значительных затрат по времени, а также может зависеть от индивидуальных различий каждого человека. Автоматические метрики (например, CLIP или DINO) не ориентированы на индивидуальные лицевые признаки [2]. ArcFace — модель распознавания лиц [3], разработанная для задач верификации и идентификации личности, и изначально не предназначенная для оценки сгенерированных изображений, что делает актуальным вопрос о том, насколько его показатели действительно отражают сохранение визуальной идентичности в задачах генеративной графики.

### Основная часть

Был проведен эксперимент, в котором сравнивались методы сохранения визуальной идентичности. Для 15 различных персонажей, представляющих публично известных людей, с различающимися возрастом, полом и визуальными характеристиками, были сгенерированы изображения, с использованием четырёх методов (Vanilla, FaceID, InstantID, PullID). Для каждого сгенерированного изображения фиксировалось значение ArcFace. Далее был проведен пользовательский опрос с участием 32 респондентов. Участникам показывались пары изображений (реальное фото и результат генерации), представленные в случайном порядке, без указания используемого метода. Респонденты субъективно оценивали визуальную схожесть и эстетическое качество генерации по шкале Лайкерта.

### Выводы

Была выявлена статистически значимая положительная корреляция между метрикой ArcFace и пользовательской оценкой визуальной схожести, что подтверждает соответствие автоматической метрики с субъективной оценкой.

### Литература

1. InstantID: Zero-shot Identity-Preserving Generation in Seconds / Q. Wang, X. Bai, H. Wang, Z. Qin, A. Chen // arXiv:2401.07519. – 2024
2. A Tale of Two Features: Stable Diffusion Complements DINO for Zero-Shot Semantic Correspondence / J. Zhang, C. Herrmann, J. Hur, L. P. Cabrera, V. Jampani, D. Sun, M.-H. Yang // arXiv preprint arXiv:2305.15347. – 2023. – <https://arxiv.org/pdf/2305.15347>
3. ArcFace: Additive Angular Margin Loss for Deep Face Recognition / J. Deng, J. Guo, N. Xue, S. Zafeiriou // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR '19), June 16–20, 2019, Long Beach, CA, USA. – 2019.