

УСЛОВНАЯ ГЕНЕРАЦИЯ ВЕКТОРНЫХ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ ДИФFUЗИОННОЙ МОДЕЛИ

Касай С.А. (ИТМО)

Научный руководитель – кандидат технических наук, Ефимова В.А. (ИТМО)

Введение. Небольшие компании заинтересованы в выпуске своей продукции, но не имеют свободных средств на разработку ее дизайна. Тематическая продукция может включать футболки, сумки, кружки и прочее. В то же время для печати на материалах требуются сверхвысокоразмерные изображения, то есть векторные, которые задаются кривыми, а не множеством точек, и поэтому легко масштабируются. По описанию компании, ссылке на сайт компании и текстовому описанию желаемого изображения требуется сгенерировать векторное изображение. Для осуществления данной идеи требуется решение специальных задач - обучение диффузионной модели генерировать растровые контекстуальные изображения их их векторизация

Основная часть. С помощью диффузионных моделей решается задача контекстуальной генерации высокоразмерных изображений [1], а специальные методы обучения позволяют экономить вычислительные ресурсы при использовании своих наборов данных [2].

Благодаря использованию дополнительных функций потерь и особых архитектур конечных генеративных моделей решаются следующие задачи:

- 1) Задачи контекстуальной генерации **векторного** изображения путем добавления функции потерь [3] в диффузионную модель. Без внедрения функций потерь становится крайне трудно контролировать генерацию векторных изображений. Текстовое описание предоставляет довольно гибкий инструмент для задания цвета, количество объектов и других параметров изображений. А функция потерь и дифференцируемый рендер [4] позволяет оптимизировать векторное изображение, соединяя тем самым векторное и пиксельное пространства.
- 2) Задачи сохранения параметров предобученной модели при использовании своего датасета для обучения генеративной диффузионной модели [5]

Выводы. Проведен анализ генеративных моделей и методов контроля генерации, разработана методика для контекстуальной генерации сверхвысокоразмерных векторных изображений с помощью диффузионной

Список использованных источников:

1. Podell, Dustin, Zion English, Kyle Lacey, Andreas Blattmann, Tim Dockhorn, Jonas Müller, Joe Penna, and Robin Rombach. Sdxl: Improving latent diffusion models for high-resolution image synthesis. arXiv preprint arXiv:2307.01952 (2023). - С. 3-7.
2. Hu, Edward J., Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. arXiv preprint arXiv:2106.09685 (2021). - С. 2-8.
3. Jain, Ajay, Amber Xie, and Pieter Abbeel. Vectorfusion: Text-to-svg by abstracting pixel-based diffusion models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1911-1920. 2023. - С. 1-5.

4. Li, Tzu-Mao, Michal Lukáč, Michaël Gharbi, and Jonathan Ragan-Kelley. Differentiable vector graphics rasterization for editing and learning. ACM Transactions on Graphics (TOG) 39, no. 6 (2020): 1-15.

5. Zhang, Lvmin, Anyi Rao, and Maneesh Agrawala. "Adding conditional control to text-to-image diffusion models." In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 3836-3847. 2023.

Научный руководитель работы

к. т. н



Ефимова В.А.

«7»_февраля 2024 г

Студент гр. М4250.



Касай С.А.

«7»_февраля 2024г