

Новая модель генерации изображений не справляется с изображением человека.

В среду компания Stability AI — это компания, специализирующаяся на разработке и исследовании в области искусственного интеллекта. Она известна созданием одного из самых популярных инструментов генеративного ИИ — Stable Diffusion, который позволяет пользователям создавать высококачественные изображения на основе текстовых описаний. Stability AI сосредотачивается на разработке открытых и доступных технологий ИИ, способствуя инновациям в различных сферах, включая искусство, дизайн и обработку данных. Компания стремится демократизировать доступ к передовым технологиям ИИ, предоставляя инструменты для широкого круга пользователей, от художников и дизайнеров до исследователей и разработчиков. Stability AI представила новую версию модели синтеза изображений — Stable Diffusion 3 Medium. Эта модель преобразует текстовые запросы в изображения, однако её появление вызвало волну критики из-за плохого качества изображений людей, особенно по сравнению с Midjourney и DALL-E 3. Как оказалось, SD3 Medium часто создает анатомически некорректные изображения, что стало причиной насмешек.

На Reddit появилась ветка под названием «Этот выпуск — шутка? [SD3-2B]», где пользователи делятся примерами неудач SD3 Medium при создании изображений людей, особенно рук и ног.

Другая ветка обсуждает проблемы с генерацией изображений девушек, лежащих на траве.

Создание рук традиционно было сложной задачей для генераторов изображений из-за недостатка примеров в обучающих наборах данных. Современные модели уже научились справляться с этой задачей, но Stable Diffusion 3 Medium сделала шаг назад. Пользователи считают, что причина провалов SD3 Medium — фильтрация контента для взрослых из обучающих данных. «Сильная цензура модели также удаляет анатомические данные, вот что произошло», — отметил один из пользователей.

Такие же проблемы возникли и при выпуске Stable Diffusion 2.0 в 2022 году. Тогда фильтрация контента для взрослых ухудшила способность модели генерировать точную анатомию человека. Stability AI исправила это в версиях SD 2.1 и SD XL, частично

восстановив утраченные способности.

Некоторые пользователи считают, что NSFW-фильтр, использованный для очистки обучающих данных, был слишком строгим. В результате из набора данных могли быть удалены изображения, не содержащие сексуального контента, но важные для обучения модели правильному изображению людей в различных ситуациях. Один из пользователей Reddit написал: «[SD3] работает нормально, пока в кадре нет людей. Кажется, их улучшенный фильтр NSFW решил, что все человекоподобное — это NSFW».