

Отчет по результатам А/В тестирования моделей генерации стикеров

Введение

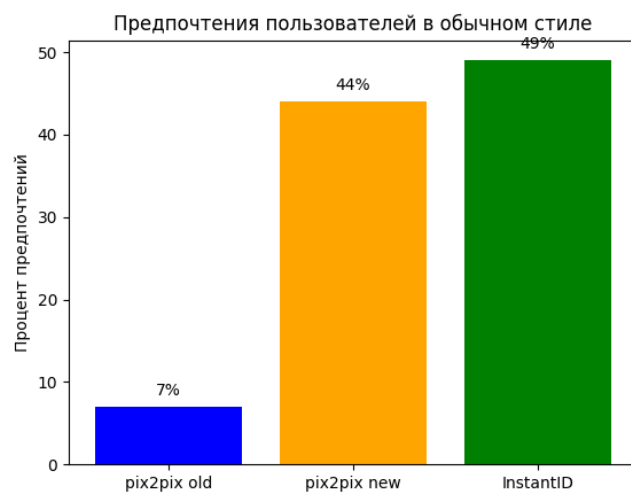
Для исследования предпочтений пользователей в отношении различных моделей генерации стикеров мы провели опрос среди 50 человек. Было протестировано три модели: pix2pix old, pix2pix new и InstantID. Пользователи оценивали модели в двух стилях: обычный стиль и стиль кошек. Исходя из этих оценок, мы рассчитали процентное соотношение предпочтений между моделями.

Методология

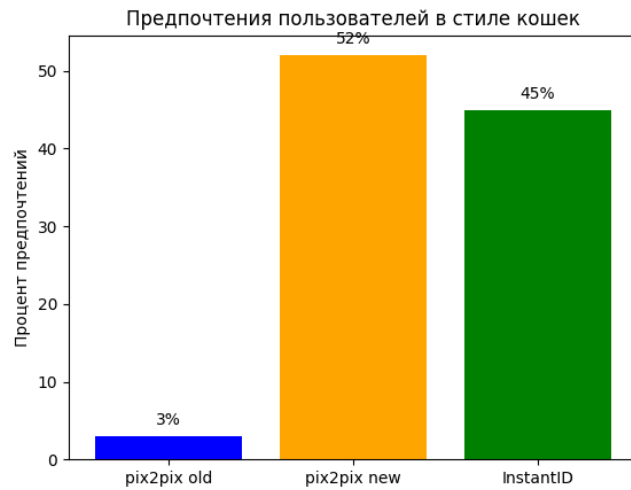
Каждому участнику было показано 20 одинаковых фотографий, сгенерированных каждой из моделей. Участники голосовали за ту модель, которая, по их мнению, лучше справилась с задачей. Голос участника был засчитан в пользу той модели, которая набрала больше голосов среди показанных изображений.

Примечание: Поскольку опрос всех участников по всем стилям был бы трудоемким, мы выбрали только один из доступных стилей – стиль кошек.

Результаты



(a) Предпочтения пользователей в обычном стиле



(b) Предпочтения пользователей в стиле кошек

Рис. 1: Распределение предпочтений пользователей

Заключение

Результаты показали, что модель InstantID получила наибольшее количество голосов в обычном стиле, но в стиле кошек немного уступила нашей модели pix2pix new. Модель pix2pix old набрала наименьшее количество голосов в обоих стилях.

Модель InstantID продемонстрировала лучшую генерацию в обычном стиле, поскольку в её датасете использовались данные, схожие со стандартными стикерами. Модель pix2pix new, благодаря улучшениям, практически не уступает InstantID в обычном стиле и даже немного превосходит её в пользовательском. Кошки – довольно простой стиль. Если рассматривать более сложные стили, с которым модель InstantID сталкивалась реже, наша модель будет показывать еще лучше результаты. Следует отметить, что старая модель pix2pix old иногда генерирует неплохие обычные стикеры, но при работе с пользовательскими стилями её результаты значительно хуже.

Приложение

Для сравнения использовался следующий датасет:
https://huggingface.co/datasets/Alexator26/AB_dataset