

Создание персонализированных генераций изображений

Tuesday, 20 May 2025 13:16 (12 minutes)

Генеративные модели большого масштаба продемонстрировали выдающиеся результаты в задаче синтеза изображений по текстовому описанию. Одним из ключевых направлений их применения является генерация персонализированных изображений. Распространённые подходы к персонализации, как правило, предполагают использование дополнительных входных данных, таких как элементы контроля или множество ракурсов объекта. В данной работе мы предлагаем метод, обеспечивающий высококачественную генерацию персонализированных изображений исключительно на основе текстового ввода, без необходимости в дополнительной информации. Основная идея заключается в обучении латентного вектора в рамках вариационного автокодировщика (VAE), который кодирует сведения о телосложении и пропорциях человека. Наша модель, *Body Lightning ID Diffusion* (BoLID), расширяет архитектуру IP-Adapter, сохраняя неизменной основную диффузионную модель. Вместо изменения базового генератора мы обогащаем латентное представление до обучения адаптеров. Последние могут быть обусловлены как выученным латентным вектором, так и входным изображением. Предлагаемый подход демонстрирует конкурентоспособные результаты по стандартным метрикам оценки качества, таким как Fréchet Inception Distance (FID) и Inception Score (IS).

Primary author: KAZACHKOV, Daniil

Co-author: Mr FILATOV, Andrei (Skolkovo Institute of Science and Technology)

Presenter: KAZACHKOV, Daniil

Session Classification: 20-Машинное обучение и нейросети

Track Classification: Машинное обучение и нейросети