

ИСПОЛЬЗОВАНИЕ КАРТ СЕГМЕНТАЦИИ В СОВРЕМЕННЫХ МОДЕЛЯХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ УЛУЧШЕНИЯ ИЗОБРАЖЕНИЙ

Титаренко М. А. (Университет ИТМО, Санкт-Петербург, Институт физиологии им. И. П. Павлова) Научный руководитель – к.т.н. Малашин Р. О. (Институт физиологии им. И. П. Павлова, Санкт-Петербург)

Наша работа направлена на адаптацию глубоких нейронных сетей для улучшения изображений в условиях шума с применением дополнительной семантической информации. Мы исследуем модель VQGAN, объединяющую подходы свёрточных слоев, трансформеров и генеративных состязательных сетей. В нашем подходе помимо зашумленных изображений в сеть подается дополнительная информация в виде масок сегментации объектов на изображении.

Введение. Все более популярны становятся подходы применения дополнительной семантической информации для решения задачи улучшения изображений искусственными нейронными сетями. В предыдущих работах мы уже исследовали такие методы:

- использование дополнительной сети для извлечения признаков для применения в функции ошибки [1];
- использование дополнительной информации при подаче на вход сети [2].

Мы показали, что дополнительная семантическая информация может быть полезна для улучшения изображений. В новом исследовании мы переходим от простых сверточных моделей нейронных сетей (Unet) к более сложным (VQGAN), чтобы сеть могла выучивать более сложные зависимости.

Целью данной работы является проведение серии экспериментов по обучению улучшению изображений модели сети VQGAN, получающей на вход изображение и карту сегментации объектов на этом изображении.

Основная часть. Мы работаем с моделью VQGAN. Это нейронная сеть включает в себя свёрточные слои и слои внимания, также называемые трансформерами. Свёрточные слои в нейронных сетях обладают особенностью извлекать и обрабатывать признаки локально, а трансформеры же наоборот способны извлекать признаки глобально. При объединении таких слоев удастся получить более универсальный и мощный инструмент для обработки изображений. Авторы [3] создали модель VQGAN для выполнения задач генерации изображений на основе различных семантических представлений таких как семантические маски, карты глубин, карты границ и некоторых других. Мы предлагаем применить эту нейронную сеть для улучшения изображений с дополнительной входной информацией в виде масок сегментации объектов на изображении.

Выводы. В работе исследуется возможность применения нейронной сети VQGAN для улучшения изображений с использованием дополнительной семантической информации в виде масок сегментации. Это исследование может помочь в задачах улучшения и сегментации, например, медицинских изображений таких как снимки образцов крови, МРТ или УЗИ.

1. Титаренко М.А., Малашин Р.О. Метод улучшения изображений с помощью глубоких нейронных сетей при использовании высокоуровневой информации // ОПТИЧЕСКИЙ ЖУРНАЛ - 2020. - Т. 87. - № 10. - С. 59-68
2. Титаренко М.А., Малашин Р.О. Исследование способностей нейронных сетей к извлечению и использованию семантической информации при обучении восстановлению зашумлённых изображений // Оптический журнал - 2022. - Т. 89. - № 2. - С. 25-35
3. Esser P., Rombach R., Bjorn O. Taming Transformers for High-Resolution Image Synthesis. // URL: <https://arxiv.org/pdf/2012.09841.pdf> (accessed: 22.02.2022).

Титаренко М. А. (автор)

Подпись

Малашин Р. О. (научный руководитель)

Подпись