

УДК 004.912

РАЗРАБОТКА ГЕНЕРАТИВНОЙ IMAGE-TO-TEXT МОДЕЛИ ДЛЯ ЗАДАЧ МАШИННОГО ПЕРЕВОДА

Синенкова А.А. (ИТМО)

Научный руководитель – Михайленко А.Ю. (ИТМО)

Введение. Машинный перевод - это процесс перевода текстов с одного естественного языка на другой при помощи компьютерных программ. Задача поиска эффективных программ для перевода текста, представленного в цифровом формате, решается с появлением первых ЭВМ в 1940-х годах [1], и в настоящее время существует множество решений, демонстрирующих достаточно точные результаты (Google Translate, Яндекс Переводчик). Однако в реальности часто необходим перевод текста, цифровой аналог которого труднодоступен, поэтому существует потребность в программе, способной качественно переводить текст с изображений. С развитием области глубокого обучения для задач машинного перевода ищут решения, основанные на нейронных сетях. Этот подход оказался крайне эффективным, и за относительно небольшой период времени такие системы, как DeepL [2], превзошли показатели передовых переводчиков на некоторых парах языков. Однако при переводе с изображений даже нейросетевые программы далеки от совершенства, так, например, они плохо распознают текст на некоторых языках, а также проявляют высокую чувствительность к неидеальному качеству изображений, в то время как в реальности нередко отсутствует возможность получить более качественную картинку.

Основная часть. С момента своего появления в 2013 и 2015 годах соответственно вариационный автоэнкодер и диффузионная модель демонстрируют крайне успешные результаты при решении многих прикладных задач, таких как изменение изображений, генерация различных типов данных и т.д. Целью данной работы является исследование эффективности этих нейросетей в области машинного перевода на изображениях с текстом. Задача состоит в том, чтобы реализовать выбранные модели, сравнить их качество, затем выбрать наиболее удачную и попытаться улучшить её качество.

Выводы. В рамках первого этапа работы было проведено исследование существующих решений для задач машинного перевода.

Список использованных источников:

1. История машинного перевода. [Электронный ресурс]. URL: <https://www.langust.ru/etc/history.shtml?ysclid=lsbxoobgkb373000157>
2. Почему именно DeepL? [Электронный ресурс]. URL: <https://www.deepl.com/ru/blog/how-deepl-work>