

М. А. Кадан

(УО «ГрГУ им. Я. Купалы», Гродно)

ВОССТАНОВЛЕНИЕ ТЕКСТОВОГО КОНТЕНТА С ПОМОЩЬЮ ГЕНЕРАТИВНО-СОСТЯЗАТЕЛЬНЫХ СЕТЕЙ

Генерирующие состязательные сети (GAN, Generative Adversarial Networks) явились революционным прорывом в создании контента с помощью методов искусственного интеллекта. Особенность GAN в том, что они обучаются создавать синтетические данные, подобные эталонным данным. Примером использования GAN является построение сети, которая анализируя изображения рукописных цифр, учится генерировать новые изображения с нуля – по сути, в этом случае мы обучаем сеть «писать».

В работе ставилась задача обучения GAN способности «читать». Предполагается, что мы располагаем набором цифровых изображений плохого качества (подготовленных при отсутствии оптической стабилизации, плохой резкости, в условиях плохой освещенности, низкой разрешающей способности и т.д.), содержащих текстовую информацию. Необходимо добиться, путем обучения GAN, улучшения качества таких зашумленных изображений текста, которое позволит восстановить этот текст, возможно используя словарь. В качестве эталонного множества используем множество печатных латинских букв, представляющих различные компьютерные шрифты.

Была сформирована обучающая выборка, включающая примеры написания отдельных латинских букв. Обучающая выборка представлена набором графических изображений в градациях серого. Элементы выборки нормированы по размеру и отцентрированы внутри изображения: буквы размером 20x20 пикселей вписаны в квадрат 28x28 и отцентрированы с помощью поиска центра масс. Обучающая выборка представлена в формате IDX – формате хранения векторов и многомерных матриц числовых типов.

Проведенный эксперимент позволил сгенерировать на основе зашумленных текстов более 25000 букв, которые GAN сочла неотличимыми от букв, представленных в обучающей выборке. Несмотря на использование производительной вычислительной системы, эксперимент потребовал значительных временных затрат.

Дальнейшее развитие эксперимента видится в направлении обеспечения возможности выделения из искаженного контента осмысленных фраз, соответствующих фразам эталонного словаря.