

Распознавание текста на основе скелетного представления толстых линий и сверточных сетей

Коробов Н.С.

Московский физико-технический институт (Государственный университет)
korobov.ns@phystech.edu

Аннотация Задача распознавания текста на изображении - одна из традиционных задач машинного зрения. В более узкой постановке такая задача сводится к классификации изображений букв. Для решения этой задачи существует множество нейросетевых (НС) подходов, в которых картинку обычно рассматривают, как растровое изображение, что важно с точки зрения прикладной значимости алгоритма. Однако существуют методы получения векторного представления растрового изображения, например, скелетное представление. В таком случае, входом НС будет некое векторное представление графа - "скелет". В данной работе предлагается модель классификатора изображений с символами, построенная над их скелетным представлением. А также сравнение качества классификации модели НС, обученной на растровых изображениях, с предложенной моделью на датасете MNIST. Кроме того, приведено сравнение точности классификации при использовании разных алгоритмов построения скелета и graph-embedding на том же датасете.

Ключевые слова: сверточные нейронные сети, распознавание текста, скелетное представление, graph-embedding

1 Введение

2 Обзор литературы