

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 ранее было изучено множество нейросетевых (НС) подходов, в которых картинку чаще всего рассматривают, как растровое изображение, что важно с точки зрения прикладной значимости алгоритма. На вход таких НС приходят пиксели изображения для каждого канала. Однако существуют методы получения векторного представления растрового изображения, например, скелетное представление. В таком случае, входом НС будет некое векторное представление графа т.н. "скелет". В данной работе предлагается сравнение качества классификации букв латинского алфавита модели НС, обученной на растровых изображениях, с моделью НС, построенной над скелетным представлением букв. Кроме того, приведено сравнение точности классификации при использовании разных алгоритмов построения скелета и сворачивания графа.

Ключевые слова: сверточные нейронные сети, распознавание текста, скелетное представление, граф, сворачивание графа, растровое изображение