

Распознавание текста на основе скелетного представления толстых линий и сверточных сетей

П.Н. Куцевол

kutsevol.pn@phystech.edu

МФТИ

Аннотация Задача распознавания изображений является одной из классических задач машинного обучения, а одна из наиболее распространенных подзадач - задача распознавания текста. В данной статье рассматриваются два подхода в классификации букв латинского алфавита. Один из них - кластеризация растровых изображений с помощью нейронных сетей. Такая задача уже рассмотрена в ряде работ и получены нейронные сети, достаточно точно классифицирующие буквы алфавита. Другой подход заключается в представлении символов в виде графов и обучении сверточной нейронной сети, которая будет оперировать состояниями вершин и ребер графа. Один из вариантов представления изображения буквы графом - скелетное представление толстых линий, описание которого можно также найти в литературе. В рамках данной работы построена нейронная сеть над растровыми изображениями, изучены алгоритмы перехода от растровых изображений к скелетным представлениям и сконструирована сверточная нейронная сеть над скелетными представлениями, распознающая символы. Один из путей дальнейшей работы - комбинация растрового и скелетного представления для увеличения точности классификации.