

Распознавание текста на основе скелетного представления толстых линий и сверточных сетей*

А. С. Лукоянов¹

lukoyanov.as@phystech.edu

¹МФТИ

Задача оптического распознавания символов уже стала классической среди задач компьютерного зрения. Несмотря на то, что качество существующих моделей довольно высоко, каждый год выходит множество научных работ, посвященных именно классификации символов. Традиционно входом таких алгоритмов является растровое изображение, что, безусловно, имеет большую прикладную значимость чем классификация символов в векторном представлении. Тем не менее, существует более комплексный подход, при котором растровое изображение сначала переводится в векторное представление путем построения скелета символа, то есть графовой структуры, а потом подается на вход обучаемой модели.

Такой подход имеет несколько недостатков, один из них заключается в неприменимости традиционных сверточных нейронных сетей на графовых структурах. В данной работе мы предлагаем способ свертывания графовых структур, позволяющий породить информативное описание скелета толстой линии. Так же в работе приводится сравнительный анализ архитектур, работающих непосредственно на растровом представлении символов и архитектур, использующих графовое представление символов. Нам удалось добиться значимого повышения качества распознавания толстых линий за счет нового способа порождения их описаний.

Ключевые слова: классификация символов; распознавание текста; графовые структуры; скелетное представление; толстые линии; свертки.

DOI: 10.21469/22233792

*Работа выполнена при финансовой поддержке РФФИ, проекты № 00-00-00000 и 00-00-00001.