

Распознавание текста на основе скелетного представления толстых линий и сверточных сетей

Александр Бойко

boyko.am@phystech.edu

В статье рассматривается задача распознавания текста на изображении при помощи скелетного представления толстых линий и комбинации этого алгоритма с кластеризацией растровых изображений с помощью нейронных сетей (что является классическим решением).

Алгоритм является непрерывным, т.е. оперирующим с формами и фигурами, а второй - дискретным, т.е. оперирующим с пикселями. Мотивацией для использования непрерывных методов является их родство человеческому восприятию. Кроме того, непрерывные методы помогают бороться с искажениями текста на обрабатываемых изображениях, что является серьезной проблемой в области распознавания текста (пример: <http://www.machinelearning.ru/wiki/images/e/ec/Gr2006Masalovitch.pdf>).

В качестве непрерывного представления текста используется скелет, состоящий из набора кривых и радиальной функции, сопоставляющей каждую точку кривой с максимальным радиусом окружности, лежащей внутри фигуры с центром в этой точке. Такое представление называется медиальным.

Планируется реализация классификатора, состоящего из композиции трех алгоритмов: алгоритма медиального представления текста, алгоритма генерации признаков по скелету и классификатора на основе сверточной нейронной сети.