

Сравнение нейросетевых и непрерывно-морфологических методов в задаче детекции текста (Text Detection)

Гайдученко Н.Е., Труш Н.А., Торлак А.В., **Миронова Л.Р.**, **Акимов К.М.**, Гончар Д.А.

October 29, 2018

Обзор проблемы

В связи с распространенностью мобильных устройств, роботизированных систем, поиска картинок в Интернете, задача распознавание текста имеет широкий спектр применения, однако, фундаментальным исследованием все же остается вопрос определения наличия самого текста на изображении и, если таковой имеется, определить его и локализовать. В литературе различные этапы этих фундаментальных задач имеют разные названия, включая локализацию текста [1] , целью которой является определение области изображения, потенциально содержащей текст; детекция текста, определяющая является ли данная область текстом или нет, использует процедуры локализации и тестирования. Решение этой задачи остается по-прежнему проблематичным из-за таких факторов как ориентация текста, шрифт, цвет и т.д.

В нашей работе рассматривается детекция текста на документах и фотографиях, сделанных в реальных условиях и обычно на встроенные в телефоны камеры. Вторая задача усложняется из-за наличия множества факторов. Шум датчика от камеры на телефоне обычно выше, чем у традиционного сканера, а отсутствие стабилизации приводит к размытию фотографий. Неправильное освещение также сильно затрудняет обработку текста: изображение может быть слишком темным или засвеченным ярким солнцем или включенной вспышкой. Вдобавок естественный текст сцены вполне может иметь углы обзора, которые не параллельны тексту, а на поверхности неплоских объектов текст деформируется.

Обзор методов

В процессе исследования мы тестируем нейросетевые алгоритмы: SSD, EAST, CTPN, SynthText и морфологические методы на pdf-документах и фотографиях, сделанных нами. Нейронным сетям требуется предоставлять большой объем данных для обучения, морфологические методы же, напротив, сразу применяют определенные действия для анализа данного изображения.

Использованная литература

[1] { R. Lienhart, A. Wernicke, "Localizing and segmenting text in images and videos", IEEE Trans. Circuits Syst. Video Technol., vol. 12, no. 4, pp. 256-268, Apr. 2002. }