

## 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.*

Детектирование текста на фотографиях является первоклассной, многогранной задачей, которая предоставляет неограниченные возможности для прикладных задач. В отличие от распознавания текста на отсканированных документах, мы сталкиваемся с такими проблемами как: различные текстуры фона, сложное освещение, загроможденность, вариативность шрифтов. В вышеупомянутых случаях детектирование сводится к нетривиальным методам ( conditional random fields [1], pictorial structures [2] ), которые разделяют процесс детектирования на несколько этапов, необходимых для получения полной системы.

Однако, в данной работе мы будем рассматривать такие методы как сверточные нейронные сети [3][4], сводящие процесс к одному этапу - детектированию из необработанных цветных пикселей. Использование данного метода позволяет построить end-to-end систему с примитивными техниками пост-обработки, такими как Non-Maximum Suppression [5] и Beam Search [6].

Опираясь на недавние разработки в сфере детектирования объектов [5][6], рассмотрим метод, основывающийся на детектировании текста с помощью ограничивающих четырехугольников [7], размер которых прогнозируется через единую обучаемую нейронную сеть. Далее, используя CRNN [8] - open-source модуль для распознавания текста, обнаружение текста осуществляется с осведомленностью его смыслового уровня, тем самым значительно повышая точность определения слов.

### References:

- [1] A. Mishra, K. Alahari, and C. V. Jawahar. *Top-down and bottom-up cues for scene text recognition. In CVPR, 2012.*
- [2] K. Wang, B. Babenko, and S. Belongie. *End-to-end scene text*
- [3] Tao Wang, David J., Adam Coates, Andrew Y. *End-to-End Text Recognition with Convolutional Neural Networks*

- [4] Manolis Delakis, Christophe Garcia *TEXT DETECTION WITH CONVOLUTIONAL NEURALNETWORKS*
- [5] A. Neubeck and L. Gool. *Efficient non-maximum suppression. In ICPR, 2006.*
- [6] S. J. Russell, P. Norvig, J. F. Candy, J. M. Malik, and D. D. Edwards. *Artificial intelligence: a modern approach. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1996.*
- [7] Minghui Liao, Baoguang Shi, Xiang Bai *A Single-Shot Oriented Scene Text Detector*
- [8] B. Shi, X. Bai, and C. Yao, *An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition IEEE TPAMI, vol. 39, no. 11, pp. 2298–2304, 2017.*