

УДК 004

СВЕРТОЧНЫЕ СЕТИ ДЛЯ СЕМАНТИЧЕСКОЙ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ

Григорьев Д.С.

Научный руководитель: Спицын В.Г., д.т.н., профессор

Национальный Исследовательский Томский политехнический университет,
634050, Россия, г. Томск, пр. Ленина, 30
E-mail: trygx@tpu.ru

The survey describes different techniques for semantic segmentation.

Key words: *Semantic segmentation, neural network, convolution neural network,*

Ключевые слова: *Семантическая сегментация, нейронная сеть, сверточная нейронная сеть*

Введение

На текущий момент в задачах классификации, детектирования, локализации объектов на изображениях или видео сформировался стандарт в основе которого лежит сверточная нейронная сеть. Архитектура искусственных нейронных сетей, была предложена Яном Лекунном для эффективного распознавания изображений. В ее основа имеет биологически правдоподобную структуру, использующую некоторые особенности зрительной коры, в которой были открыты так называемые простые клетки, реагирующие на прямые линии под разными углами, и сложные клетки, реакция которых связана с активацией определенного набора простых клеток. Фактически, идея сверточных сетей заключается в чередовании сверточных слоев и слоев подвыборки.

Название архитектура сети получила из-за наличия операции свертки, суть которой в том, что каждый фрагмент изображения умножается на ядро свертки поэлементно, а результат суммируется и записывается в аналогичную позицию выходного изображения.

На сегодняшний день существует достаточно большое количество алгоритмов, успешно решающих задачи семантической сегментации и использующие сверточную архитектуру [1, 2].

Однако большинство моделей достаточно тяжелы с точки зрения использования их в системах реального времени и для обработки видеопотока.

Исходя из представленного выше, была поставлена задача разработать и реализовать алгоритм, позволяющий распознавать и сегментировать объекты на изображении с приемлемым качеством и эффективностью.

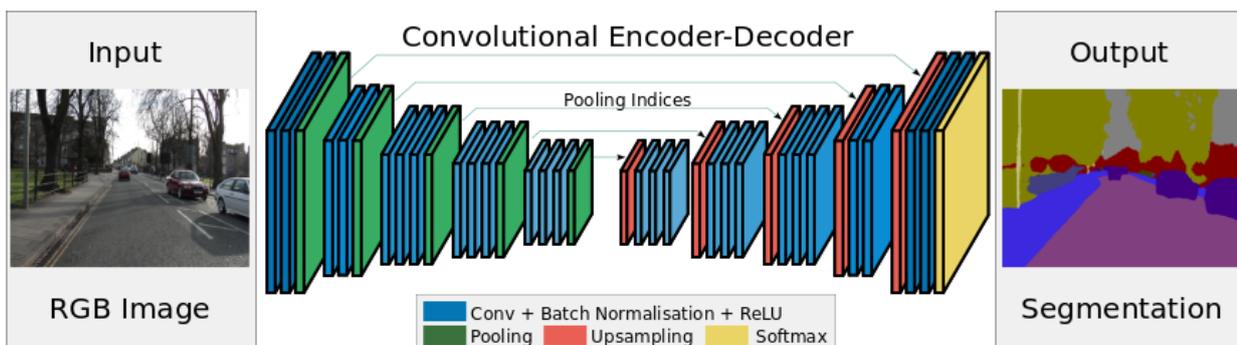


Рис. 1. Пример архитектуры сети (SegNet)[1]

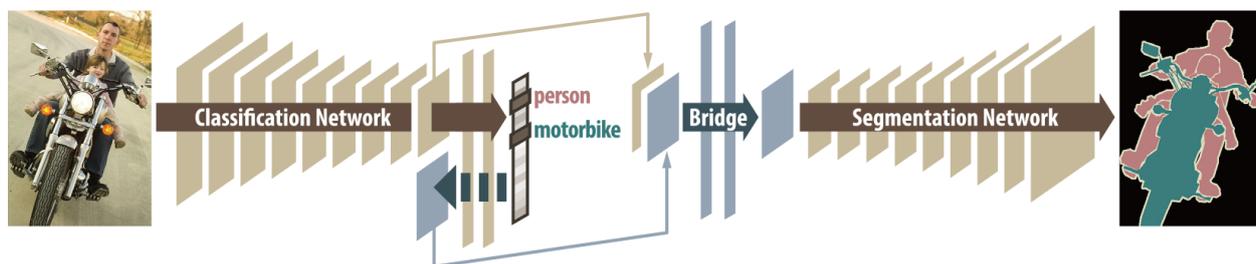


Рис. 2. Пример архитектуры сети (DecoupledNet) [2]

В данной работе представлен алгоритм сегментации использующего в основе компактную сверточную сеть (Fully convolutional network) в сочетании с CRF. С целью повышения качества работы алгоритма решено использовать несколько сверточных сетей объединённых в композицию. В качестве активационной функции была использована параметризованная ReLU. Особенностью реализации алгоритма также является ее исполнение в формате с половинной точностью без существенной потери в качестве по сравнению с оригинальным (с одиночной точностью) исполнением.

СПИСОК ИСТОЧНИКОВ

1. Badrinarayanan V., Kendall A., Cipolla R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. [Электронный ресурс] URL: <http://arxiv.org/abs/1511.00561>
2. Hong S., Noh H., Han B. Decoupled Deep Neural Network for Semi-supervised Semantic Segmentation. [Электронный ресурс] URL: <http://arxiv.org/abs/1506.04924>
3. Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. [Электронный ресурс] URL: <https://arxiv.org/abs/1506.01497.html>
4. Dai J., He K., Sun J. BoxSup: Exploiting Bounding Boxes to Supervise Convolutional Networks for Semantic Segmentation. [Электронный ресурс] URL: <https://arxiv.org/abs/1503.01640.html>
5. David G. Lowe Distinctive Image Features from Scale-Invariant Keypoints. [Электронный ресурс] URL: <http://link.springer.com/article/10.1023/B:VISI.0000029664.99615.94>
6. Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., Rabinovich A. Going deeper with convolutions. In Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2015.
7. Cun Y. Le, Boser B., Denker J.S., Henderson D., Howard R.E., Hubbard W., Jackel L.D. et al. Handwritten digit recognition with a back-propagation network. In Advances in neural information processing systems, 1990.