

Weekly Report

Adviser: Prof. Yang Wen

Student: Cheng Wensheng

Period: 2018.4.27-5.4

Abstract—This week I mainly put my effort on reading the classic CNN-based semantic segmentation framework SegNet.

I. PAPER OUTLINE

AFTER the pioneering work of Fully Convolutional Networks (FCN), SegNet has been proposed as the best segmentation framework at that time. Its main contributions are as follows.

- Proposed a new decoder network to map the low resolution encoder feature maps to full input resolution feature maps for pixel-wise classification. The novelty of SegNet lies in the manner in which the decoder upsamples its lower resolution input feature map(s).
- Specifically, the decoder uses pooling indices computed in the max-pooling step of the corresponding encoder to perform non-linear upsampling. This eliminates the need for learning to upsample.
- The upsampled maps are sparse and are then convolved with trainable filters to produce dense feature maps.

II. CONCLUSION

The author did corresponding experiments to compare SegNet with other state-of-the-art frameworks, analyzed results and got the following conclusion.

Those architectures which store the encoder network feature maps in full perform best but consume more memory during inference time. SegNet on the other hand is more efficient since it only stores the max-pooling indices of the feature maps and uses them in its decoder network to achieve good performance.

Fig. 1 is the overall framework. Fig. 2 is the output of semantic segmentation.

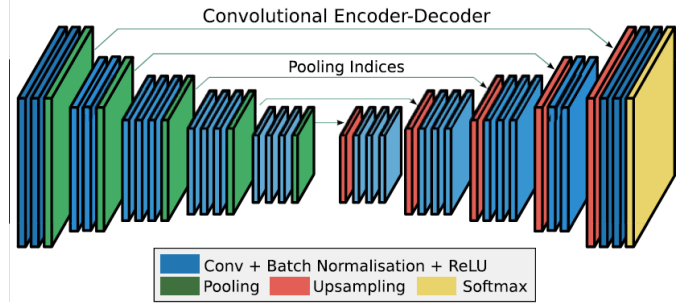


Fig. 1: Framework



Fig. 2: Semantic segmentation results