Instance-aware Semantic Segmentation Via Multi-task Network Cascades

Qilei Zhang

Jun 28 2018

Abstract

Semantic segmentation research has recently witnessed rapid progress, but many leading methods are unable to identify object instances.

1. Introduction

Since the development of fully convolutional networks, the accuracy of semantic segmentation has been improved rapidly thanks to deeply learned features, large-scale annotations , and advanced reasoning over graphical models. Nevertheless, FCNs and improvements are designed to predict a category label for each pixel, but are unaware of individual object instances [1].
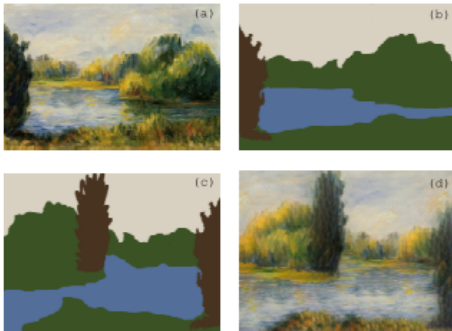

Figure 1. Illustrations of common multi-task learning (left) and our multi-task cascade (right).

2. Related Work

Object detection methods involve predicting object bounding boxes and categories [2].In SPPnet and Fast RCNN, the convolutional layers of CNNs are shared on the entire image for fast computation. Faster R-CNN exploits the shared convolutional features to extract region proposals used by the detector. Sharing convolutional features leads to substantially faster speed for object detection systems [3].

Figure 2. Multi-task Network Cascades for instance-aware semantic segmentation. At the top right corner is a simplified illustration.

3. Multi-task Network Cascades

In MNC model, the network takes an image of arbitrary size as the input, and outputs instance-aware semantic segmentation results.The cascade has three stages: proposing box-level instances, regressing mask-level instances, and categorizing each instance.These three stages are designed to share convolutional features. Each stage involves a loss term, but a later stages loss relies on the output of an earlier stage, so the three loss terms are not independent [4].

References

[1] F. Briggs, X. Z. Fern, R. Raich, and Q. Lou. Instance annotation for multi-instance multi-label learning. *Acm Transactions on Knowledge Discovery from Data*, 7(3):1–30, 2013. 1

[2] M. Jankovic, J. S. L. Cardinal, and J. C. Bocquet. Context management in collaborative decision making in complex design projects. *International Journal of Product Development*, 20(4):286, 2015. 1

[3] Miledi, Woodward, and R. M. How spoken languages work in the absence of an inventory of discrete units. *Language Sciences*, 53(1):623–643, 2016. 1

[4] B. Zhou, E. P?ppel, and Y. Bao. In the jungle of time: the concept of identity as a way out. *Frontiers in Psychology*, 5(4):844, 2014. 1