CVPR 2017 “Large Kernel Matters—Improve Semantic Segmentation by Global Convolutional Network”
from Tsinghua University and Megvii Inc. (Face++)

2.1 介绍

这篇文章总结了语义分割中的卷积核尺寸小所带来的感受野小的问题，对其加以改进，并提出了轮廓改善网络以改善分割后物体的边缘。

目前语义分割主要是 FCN 结构的，即将最后的全连接层删除，使用反卷积的方式将高层的语义信息上采样到原图大小，在上采样过程中结合低层特征以尽可能恢复物体轮廓。全连接层会将所有特征连在一起，输出一个分类结果。而语义分割需要对每个像素都给出分类结果，所以必须删除全连接层。

同时还需要删除 Global Pooling 层，以保留特征的位置信息。

2.2 问题与改进

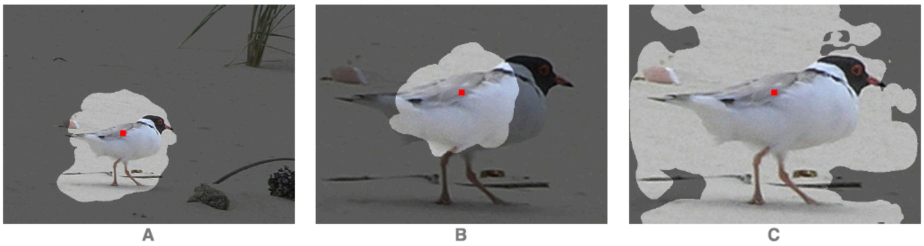

Figure 3. Visualization of *valid receptive field* (VRF) introduced by [38]. Regions on images show the VRF for the score map located at the center of the bird. For traditional segmentation model, even though the receptive field is as large as the input image, however, the VRF just covers the bird (A) and fails to hold the entire object if the input resized to a larger scale (B). As a comparison, our Global Convolution Network significantly enlarges the VRF (C).

上图是对感受野的形象化阐释。图 A 和 B 是小尺寸卷积核的感受野，在网络深度不变的情况下，将目标物体放大，则感受野会变得无法覆盖住整个目标物体。因此增大卷积核尺寸以增大感受野很重要。

但是直接增加卷积核的尺寸会带来计算量的急剧增加，因此作者在这里采用了一种“带状”（k x 1 和 1 x k）的卷积核，通过两次卷积以使得每个像素获得和卷积核为 k x k 时相同的作用范围。但是计算量会降低很多。

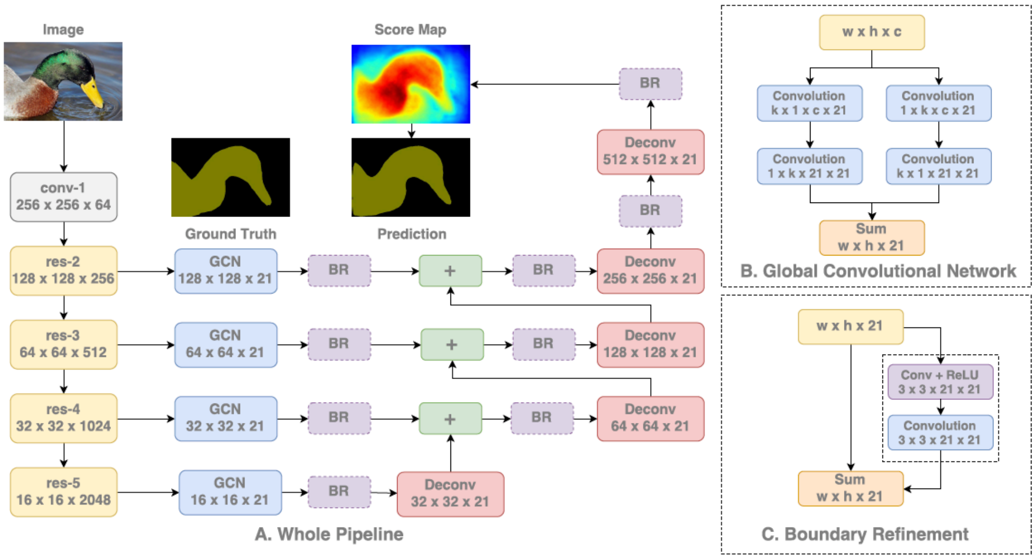

Figure 2. An overview of the whole pipeline in (A). The details of Global Convolutional Network (GCN) and Boundary Refinement (BR) block are illustrated in (B) and (C), respectively.

上图即为网络的整体结构，是一个很经典的 U-Net 结构。但是在卷积部分作者使用 GCN 来代替原本的小卷积核，并加入了 Boundary Refinement 模块，使用残差方式，以增强对物体轮廓的恢复。

2.3 实验结果

（1）大卷积核是有用的

k	base	3	5	7	9	11	13	15
Score	69.0	70.1	71.1	72.8	73.4	73.7	74.0	74.5

Table 1. Experimental results on different k settings of Global Convolutional Network. The score is evaluated by standard mean IoU(%) on PASCAL VOC 2012 validation set.

（2）和“方块状”传统卷积核的性能和参数对比，都是带状卷积核胜出

k	3	5	7	9
Score (GCN)	70.1	71.1	72.8	73.4
Score (Conv)	69.8	70.4	69.6	68.8
# of Params (GCN)	260K	434K	608K	782K
# of Params (Conv)	387K	1075K	2107K	3484K

Table 2. Comparison experiments between Global Convolutional Network and the trivial implementation. The score is measured under standard mean IoU(%), and the 3rd and 4th rows show number of parameters of GCN and trivial Convolution after res-5.

（3）边缘改善模块所带来的增益，对 IoU 不是有特别大帮助，仅 0.1%

Model	Boundary (acc.)	Internal (acc.)	Overall (IoU)
Baseline	71.3	93.9	69.0
GCN	71.5	95.0	74.5
GCN + BR	73.4	95.1	74.7

Table 5. Experimental results on *Residual Boundary Alignment*. The Boundary and Internal columns are measured by the per-pixel accuracy while the 3rd column is measured by standard mean IoU.

（4） 在当时取得了 STOA 的性能

Method	mean-IoU(%)
FCN 8s [29]	65.3
DPN [24]	59.1
CRFasRNN [37]	62.5
Scale invariant CNN + CRF [19]	66.3
Dilation10 [36]	67.1
DeepLabv2-CRF [7]	70.4
Adelaide_context [21]	71.6
LRR-4x [12]	71.8
Our approach	**76.9**

Table 10. Experimental results on Cityscapes test set.