Instance-sensitive Fully Convolutional Networks

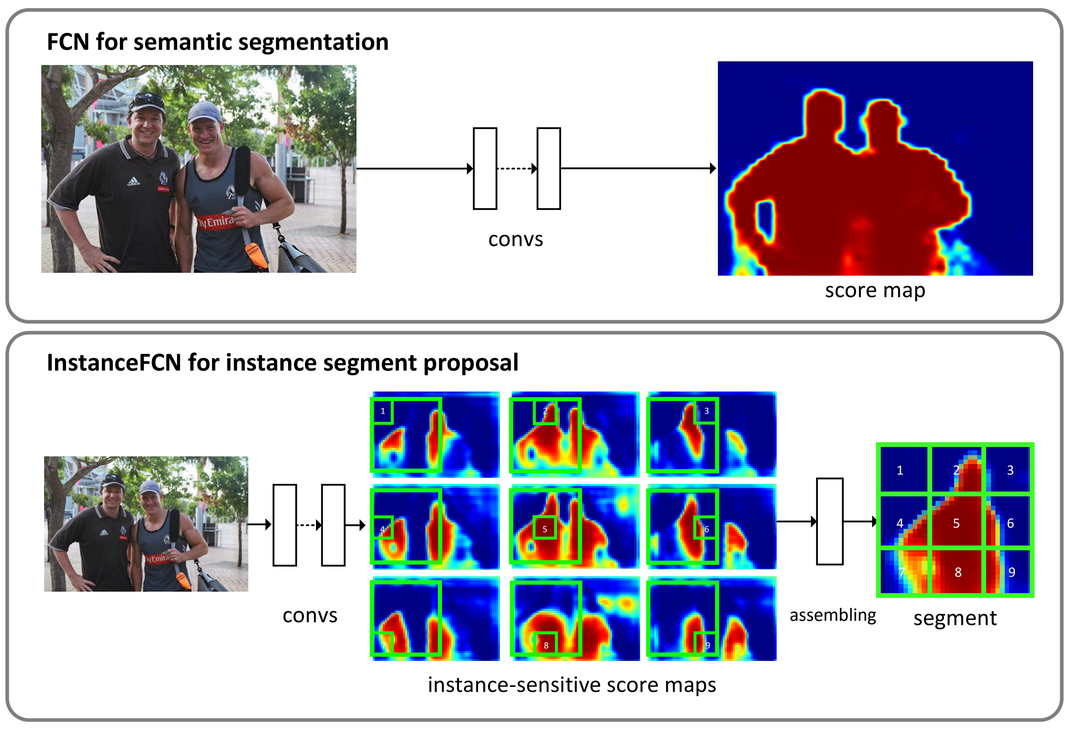
Jifeng Dai, Kaiming He, Yi Li, Shaoqing Ren, Jian Sun
Microsoft Research

Highlights

- FCNs for instance-level segment proposal
- No high-dim layer related to mask resolution
- Competitive results on PASCAL VOC and COCO

From FCN to InstanceFCN

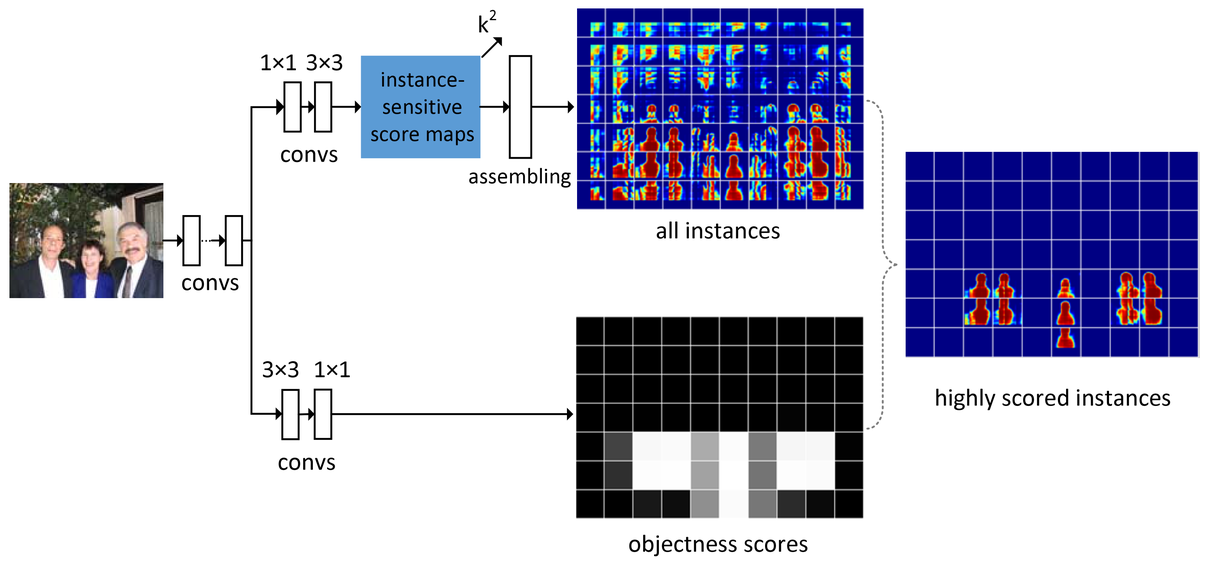
- Conventional FCN for semantic segmentation
  - One score map, unaware of object instances
- InstanceFCN for instance segment proposal
  - Instance-sensitive score maps, each of which represents a pixelwise classifier of a relative position to instances
  - A simple assembling module outputs instance candidate



InstanceFCN v.s. DeepMask

- DeepMask
  - Mapping an image sliding window to $m \times m$-dim mask ($m^2 = 3,136$)
- Our InstanceFCN
  - Each output pixel is a low-dim classifier ($k^2 = 9$)
  - Exploiting local coherence of natural images

InstanceFCN Architecture



Experiments

Results on VOC 2012 validation

method	AR@10 (%)	AR@100 (%)	AR@1000 (%)
SS	7.0	23.5	43.3
MCG	18.9	36.8	49.5
∼DeepMask	31.2	42.9	47.0
MNC	33.4	48.5	**53.8**
ours	**38.9**	**49.7**	52.6

Results on the first 5k images from COCO validation

segment proposals	AR@10 (%)	AR@100 (%)	AR@1000 (%)
GOP	2.3	12.3	25.3
Rigor	-	9.4	25.3
SS	2.5	9.5	23.0
MCG	7.7	18.6	29.9
DeepMask	12.6	24.5	33.1
DeepMaskZoom	12.7	26.1	36.6
ours	**16.6**	**31.7**	**39.2**

(The missed ground-truth objects are marked by red outlines)