Instance Segmentation

Shu Kong

CS, ICS, UCI

UCIrvine
University of California, Irvine

Outline

1. problem definition
2. object proposal based instance refinement
3. FCN architecture with smarter label
4. others
5. Conclusion

Outline

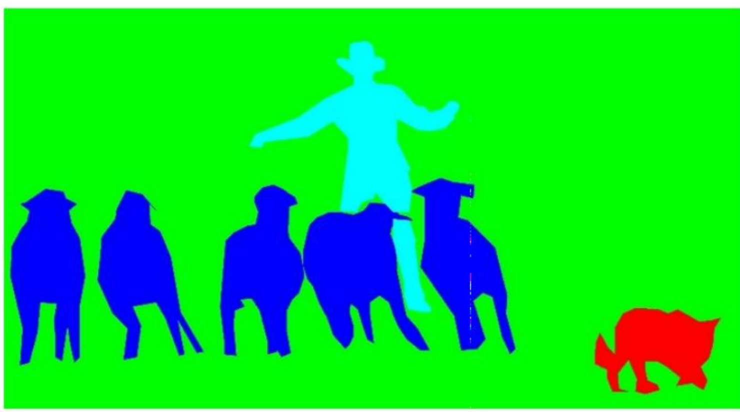
1. problem definition
2. object proposal based instance refinement
3. FCN architecture with smarter label
4. others
5. Conclusion

UCIrvine
UNIVERSITY OF CALIFORNIA, IRVINE

Semantic to Instance Segmentation

semantic segmentation -- find regions belonging to category-level labels by grouping pixels

instance segmentation -- find out all the instances by grouping pixels



Semantic segmentation

Instance segmentation

We had that?!

or similar things for instance segmentation?

We had that?!

or similar things for instance segmentation?

Yes

We had that?!

or similar things for instance segmentation?

Yes

Where?

We had that?!

or similar things for instance segmentation?

Yes

Where?

Here!

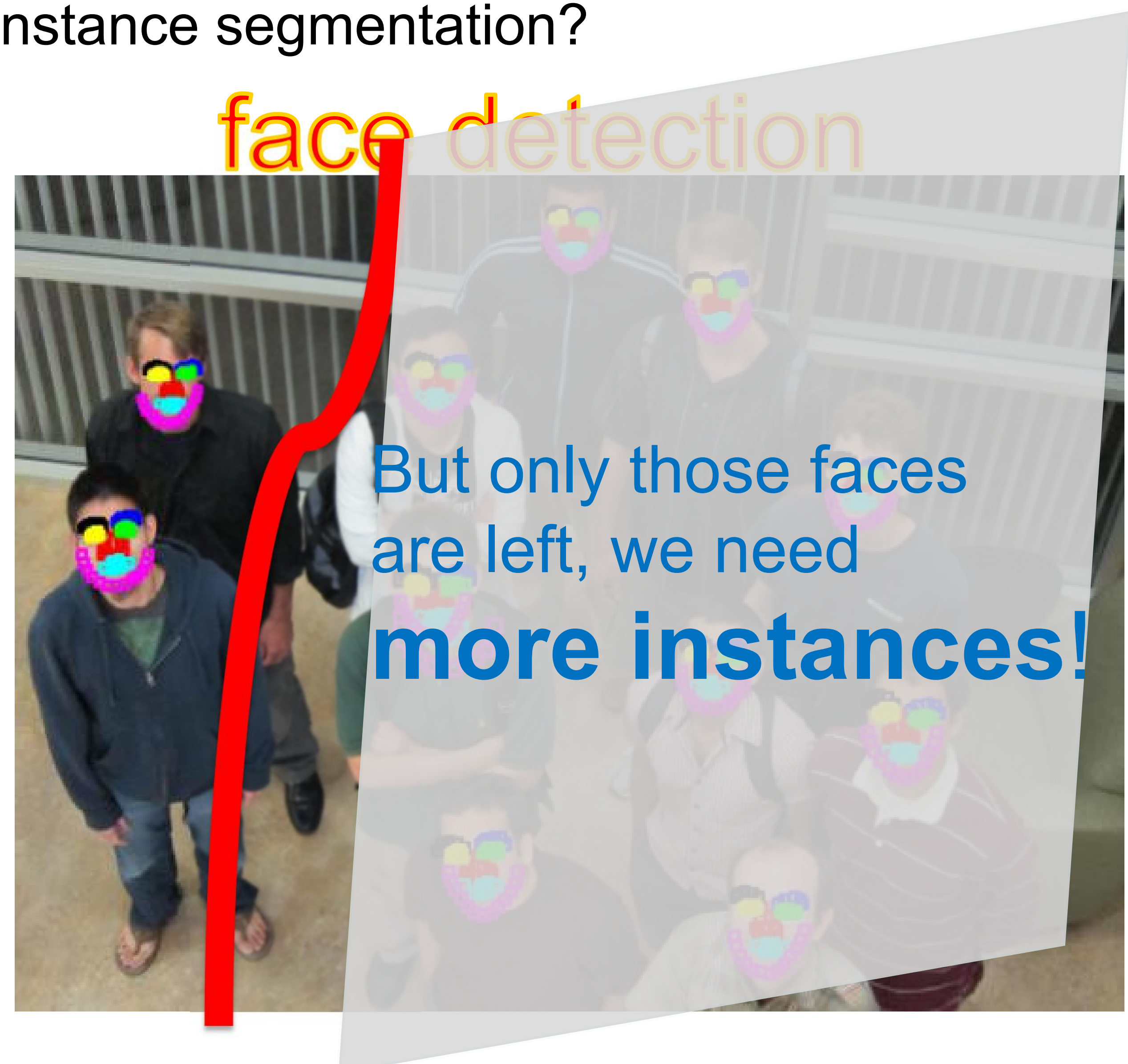
We had that?!

or similar things for instance segmentation?

Yes

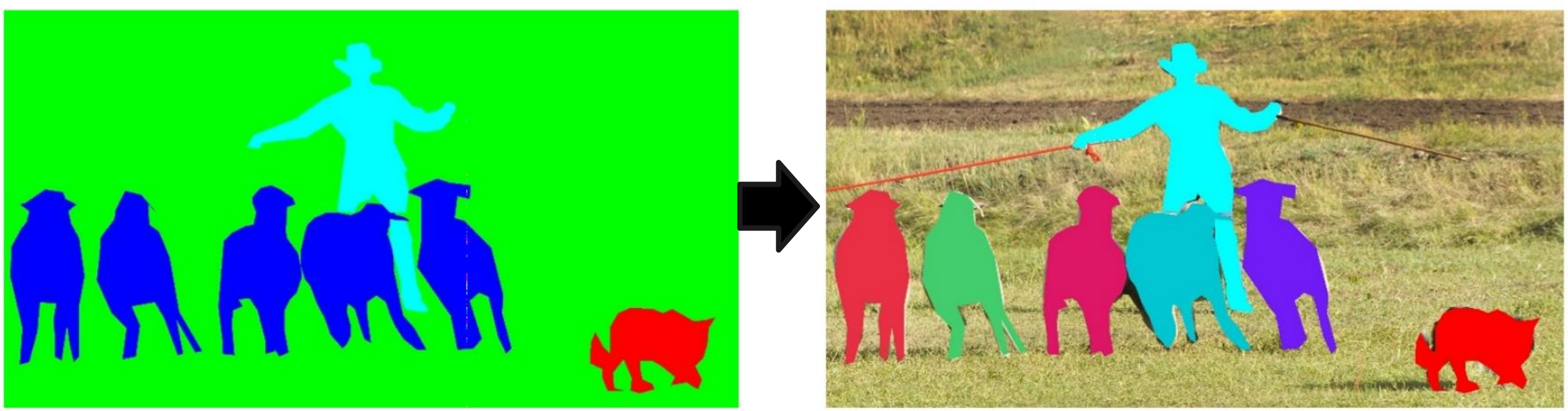
Where?

Here!



Instance Segmentation

for istance segmentation, here is a starter---

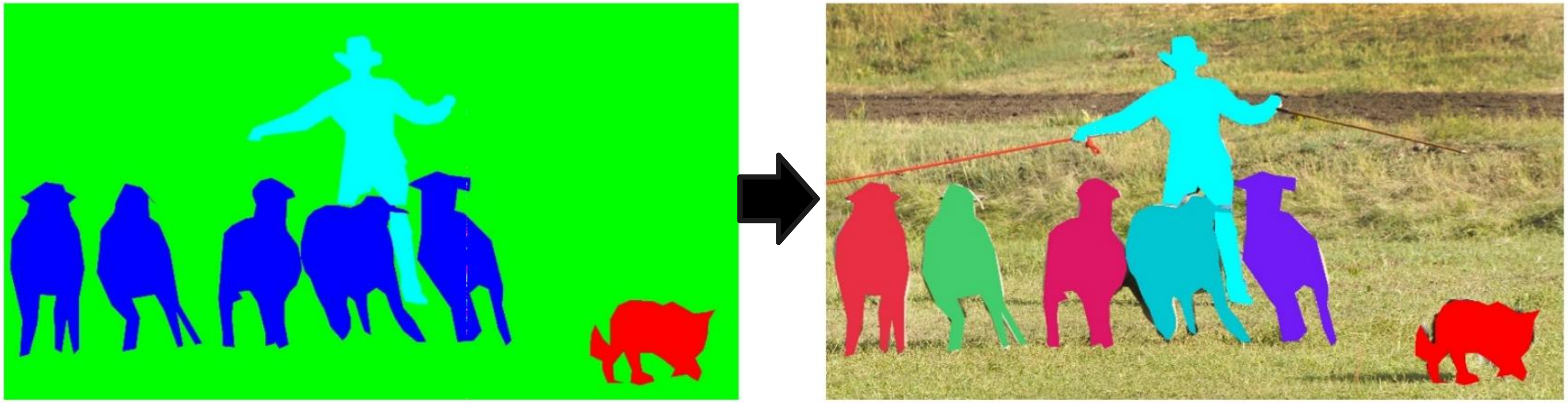


Semantic segmentation Instance segmentation

Instance Segmentation

for istance segmentation, here is a starter---

find out instances in a class-agnostic way, or object proposals



Semantic segmentation　　　　Instance segmentation

Instance Segmentation

for istance segmentation, here is a starter---

find out instances in a class-agnostic way, or object proposals

how to find the individual instances in the picture?

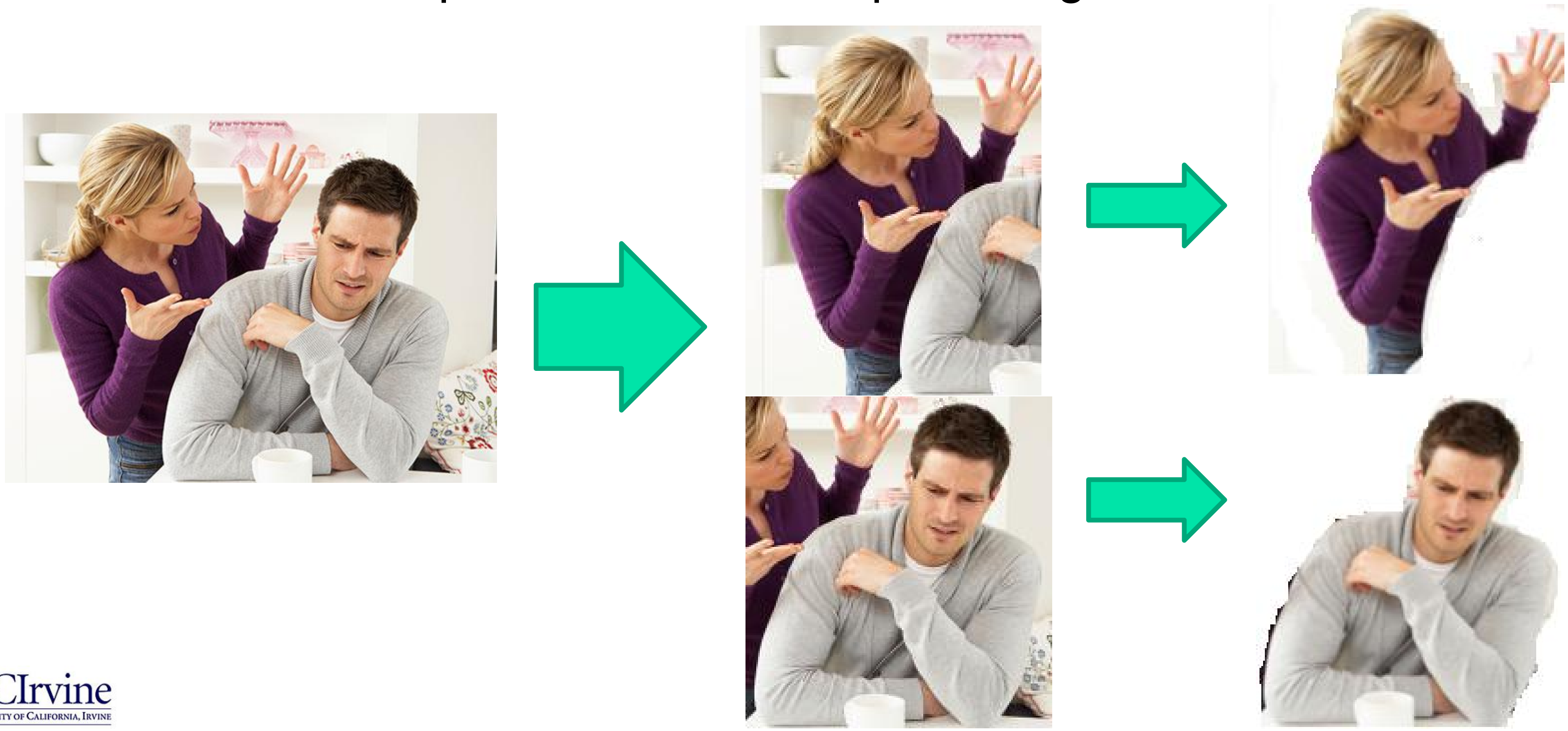
Instance Segmentation

for istance segmentation, here is a starter---

find out instances in a class-agnostic way, or object proposals

how to find the individual instances in the picture?

philosophy -- crop image (sliding window?), highlight the instance centered in the crop, and zero out the pixels/regions outside the instance



Instance Segmentation

Methods --

1. implement the idea described above

Instance Segmentation

Methods --

1. implement the idea described above
2. fancier output for instance **inference**

upper boundary

left boundary

right boundary

so on and so forth............

Proposal based Instance Segmentation

1. problem definition
2. object proposal based instance refinement
3. FCN architecture with smarter label
4. others
5. Conclusion

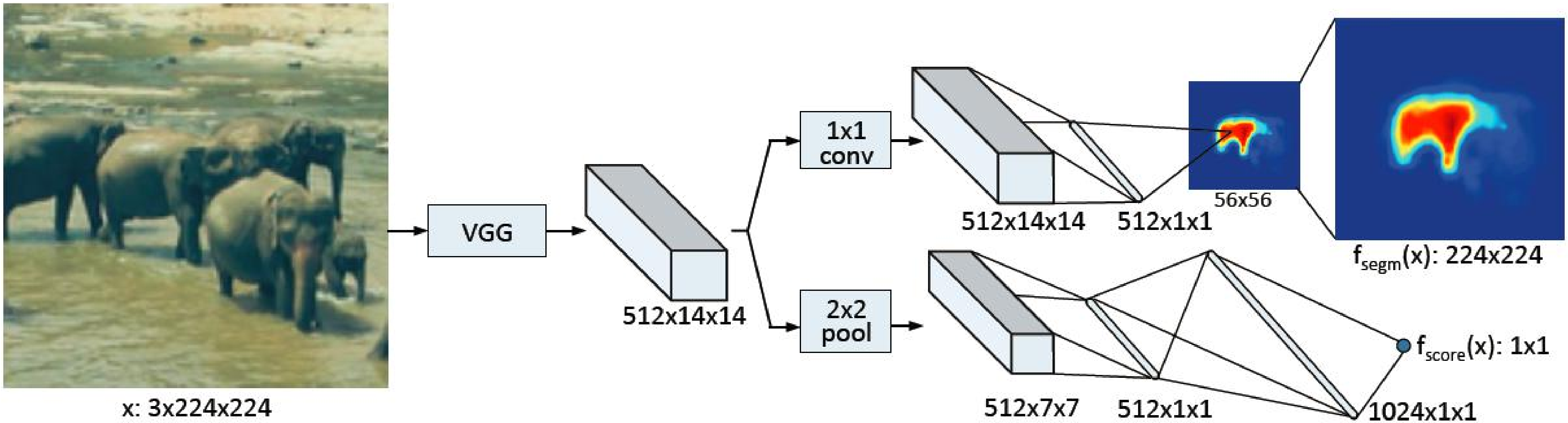
Proposal based Instance Segmentation

crop image (sliding window?), highlight the instance centered in the crop, and zero out the pixels/regions outside the instance

Pinheiro et al., "Learning to Segment Object Candidates", NIPS, 2015

Proposal based Instance Segmentation

architecture



top branch -- predicting the mask for the instance centered at the patch

bottom branch -- predicting a score to indicate whether there is a “valid” instance in the patch

Pinheiro et al., "Learning to Segment Object Candidates", NIPS, 2015

Proposal based Instance Segmentation

samping data for training -- triplet input (input image, mask, score)

Pinheiro et al., "Learning to Segment Object Candidates", NIPS, 2015

Proposal based Instance Segmentation

samping data for training -- triplet input (input image, mask, score)

input image -- reshaped into 224x224x3

Pinheiro et al., "Learning to Segment Object Candidates", NIPS, 2015

Proposal based Instance Segmentation

samping data for training -- triplet input (input image, mask, score)

input image -- reshaped into 224x224x3

mask -- binary map of size 224x224

Pinheiro et al., "Learning to Segment Object Candidates", NIPS, 2015

Proposal based Instance Segmentation

samping data for training -- triplet input (input image, mask, score)

input image -- reshaped into 224x224x3

mask -- binary map of size 224x224

score -- binary label, 1 for valid patch (**green**), -1 for invalid patch (**red**)

UCIrvine
University of California, Irvine

Pinheiro et al., "Learning to Segment Object Candidates", NIPS, 2015

Proposal based Instance Segmentation

samping data for training -- triplet input (input image, mask, score)

input image -- reshaped into 224x224x3

mask -- binary map of size 224x224

score -- binary label, 1 for valid patch (**green**), -1 for invalid patch (**red**)

Constraints --

1. the patch contains an object roughly centered in the patch
2. the object is fully contained in the patch and in a given scale range

Pinheiro et al., "Learning to Segment Object Candidates", NIPS, 2015

Proposal based Instance Segmentation

objective function -- a sum of binary logistic regression losses

$$\mathcal{L}(\theta) = \sum_k \left(\frac{1+y_k}{2w^o h^o} \sum_{ij} \log(1 + e^{-m_k^{ij} f_{segm}^{ij}(x_k)}) + \lambda \log(1 + e^{-y_k f_{score}(x_k)}) \right)$$

x_k the k-th patch

m_k its mask

y_k its objectness score

i,j the pixel location

$$\lambda = \frac{1}{32}$$

the output of the classification layer to be $h^o \times w^o$

Pinheiro et al., "Learning to Segment Object Candidates", NIPS, 2015

Proposal based Instance Segmentation

objective function -- a sum of binary logistic regression losses

$$\mathcal{L}(\theta) = \sum_k \left(\frac{1+y_k}{2w^o h^o} \sum_{ij} \log(1 + e^{-m_k^{ij} f_{segm}^{ij}(x_k)}) + \lambda \log(1 + e^{-y_k f_{score}(x_k)}) \right)$$

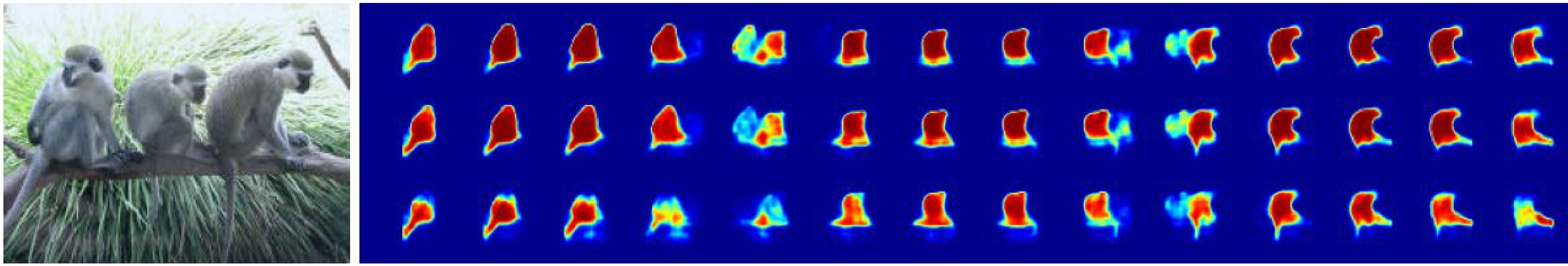
remarks --

1. negative samples do not contribute segmentation loss (**critical**)
2. alternating backpropogating the two branches
3. for scoring branch, sampling data with equal number of positive&negative
4. can be deployed in a fully convolutional manner
5. sampling data includes translation shift, scale deformation, horizontal flip
6. non-trainable upsampling layer (bilinear upsampling)

Pinheiro et al., "Learning to Segment Object Candidates", NIPS, 2015

Proposal based Instance Segmentation

qualitative results -- pretty visualization on model generalization

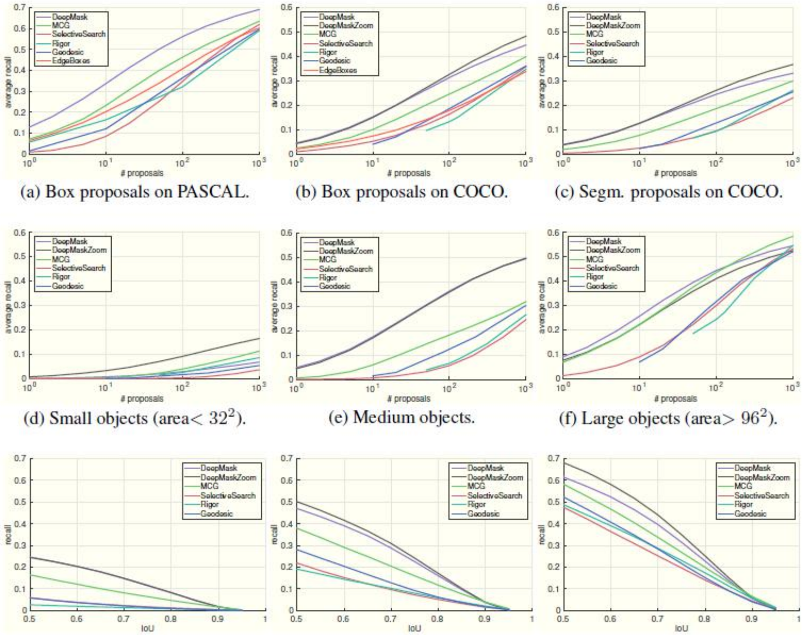


Pinheiro et al., "Learning to Segment Object Candidates", NIPS, 2015

Proposal based Instance Segmentation

quantitative results -- seems awesome

metrics -- Intersection over Union (IoU), Average Recall (AR) btwn IoU 0.5~1.0



(a) Box proposals on PASCAL. (b) Box proposals on COCO. (c) Segm. proposals on COCO.
(d) Small objects (area< 32^2). (e) Medium objects. (f) Large objects (area> 96^2).
(g) Recall with 10 proposals. (h) Recall with 100 proposals. (i) Recall with 1000 proposals.

UCIrvine
UNIVERSITY OF CALIFORNIA, IRVINE

Pinheiro et al., "Learning to Segment Object Candidates", NIPS, 2015

Proposal based Instance Segmentation

quantitative results -- seems awesome

metrics -- Intersection over Union (IoU), Average Recall (AR) btwn IoU 0.5~1.0

	Box Proposals				Segmentation Proposals						
	AR@10	AR@100	AR@1000	AUC	AR@10	AR@100	AR@1000	AUC^S	AUC^M	AUC^L	AUC
EdgeBoxes [34]	.074	.178	.338	.139	-	-	-	-	-	-	-
Geodesic [16]	.040	.180	.359	.126	.023	.123	.253	.013	.086	.205	.085
Rigor [14]	-	.133	.337	.101	-	.094	.253	.022	.060	.178	.074
SelectiveSearch [31]	.052	.163	.357	.126	.025	.095	.230	.006	.055	.214	.074
MCG [24]	.101	.246	.398	.180	.077	.186	.299	.031	.129	.324	.137
DeepMask20	.139	.286	.431	.217	.109	.215	.314	.020	.227	.317	.164
DeepMask20*	.152	.306	.432	.228	.123	.233	.314	.020	.257	.321	.175
DeepMaskZoom	.150	**.326**	**.482**	**.242**	**.127**	**.261**	**.366**	**.068**	**.263**	.308	**.194**
DeepMaskFull	.149	.310	.442	.231	.118	.235	.323	.020	.244	**.342**	.176
DeepMask	**.153**	.313	.446	.233	.126	.245	.331	.023	.266	.336	.183

~1.5s per image

Pinheiro et al., "Learning to Segment Object Candidates", NIPS, 2015

FCN with Fancier Label

1. problem definition
2. object proposal based instance refinement
3. FCN architecture with smarter label
4. others
5. Conclusion

FCN with Fancier Label

fancier output for instance **inference**

upper boundary

left boundary

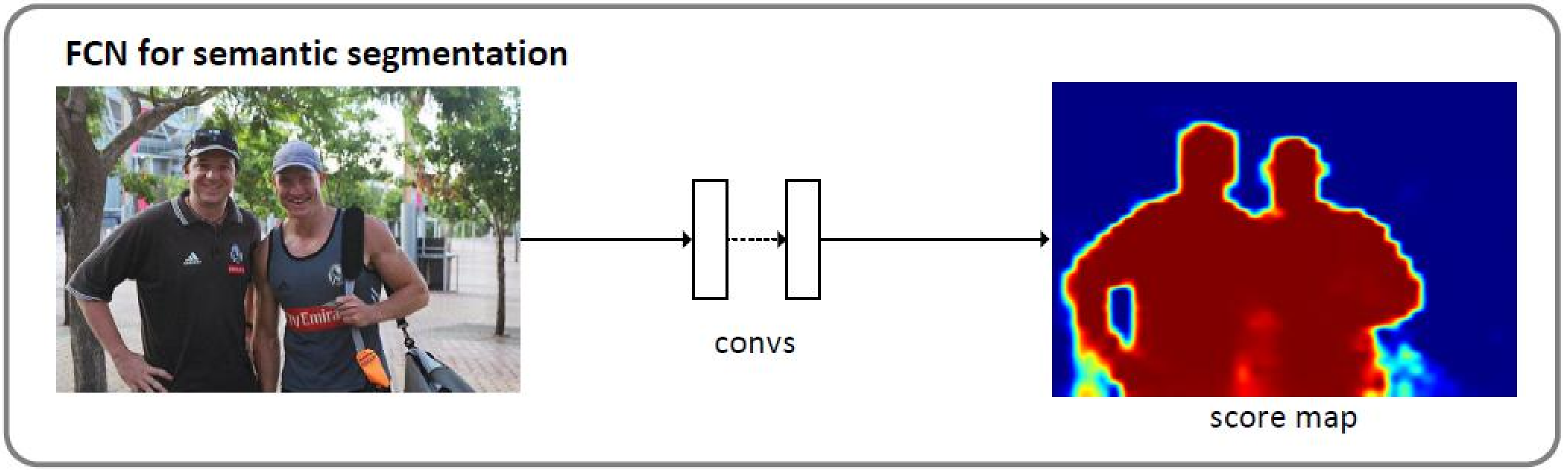
right boundary

so on and so forth............

Dai et al., "Instance-sensitive Fully Convolutional Networks", ECCV, 2016

FCN with Fancier Label

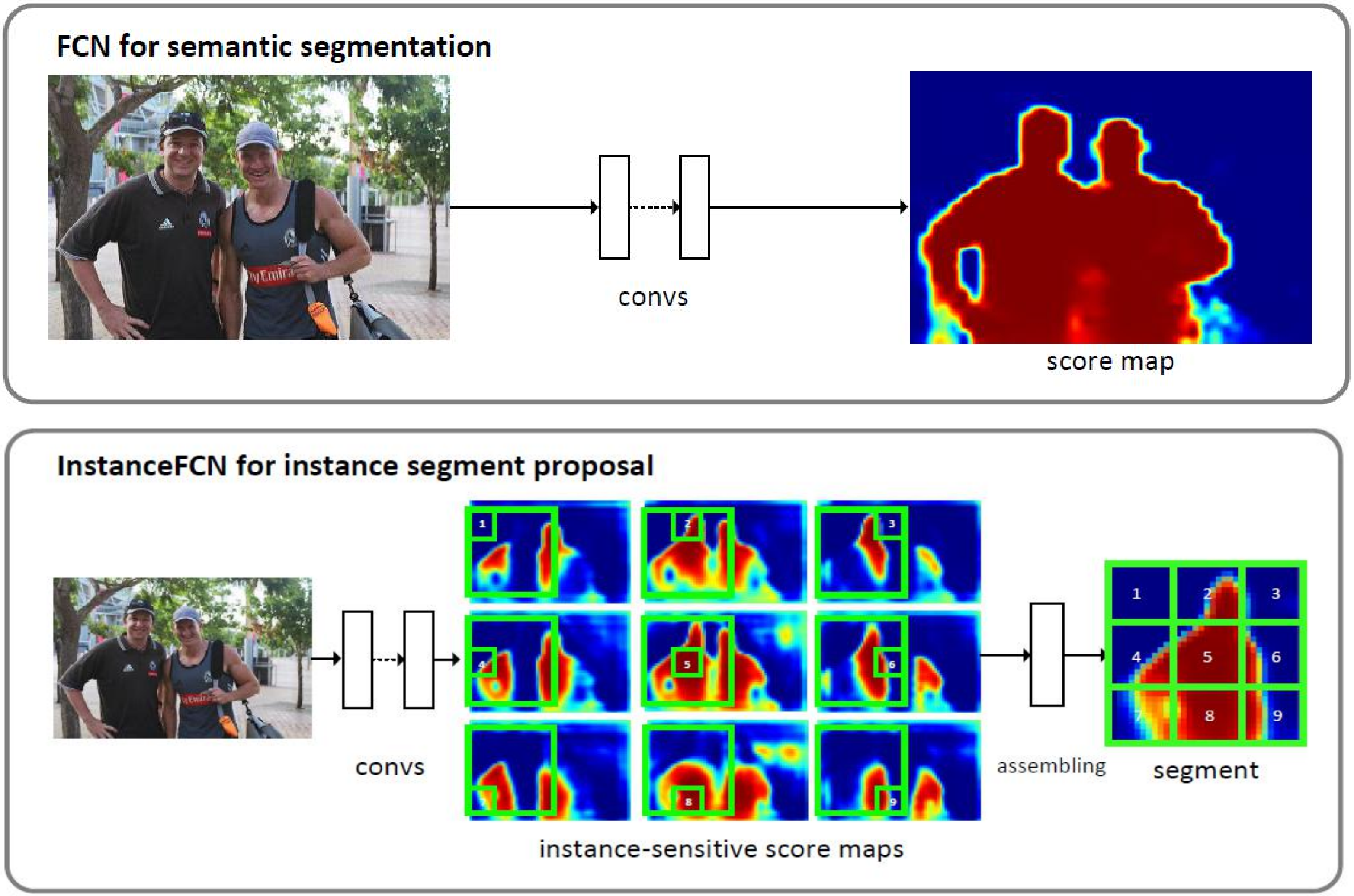
from FCN to InstanceFCN



Dai et al., "Instance-sensitive Fully Convolutional Networks", ECCV, 2016

FCN with Fancier Label

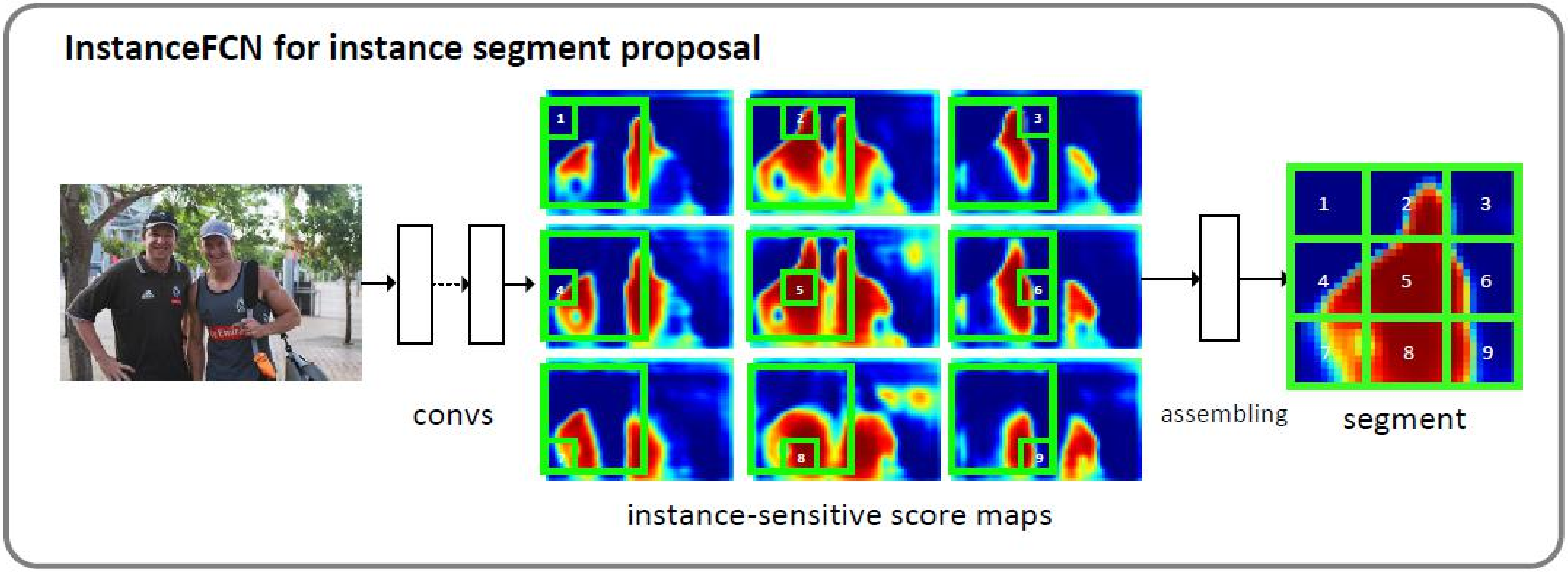
from FCN to InstanceFCN



Dai et al., "Instance-sensitive Fully Convolutional Networks", ECCV, 2016

FCN with Fancier Label

InstanceFCN -- differentiate left from right regions

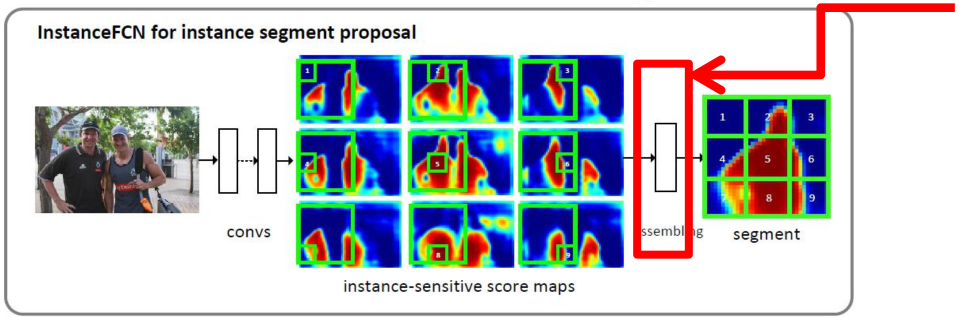


train a model to produce instance-sensitive score maps with relative position of instances

Dai et al., "Instance-sensitive Fully Convolutional Networks", ECCV, 2016

FCN with Fancier Label

Instance Assembling Module - producing instance based on maps



say, 9 output maps, mosaic them w.r.t relative positions,

similar to **mosaic upsampling**

Dai et al., "Instance-sensitive Fully Convolutional Networks", ECCV, 2016

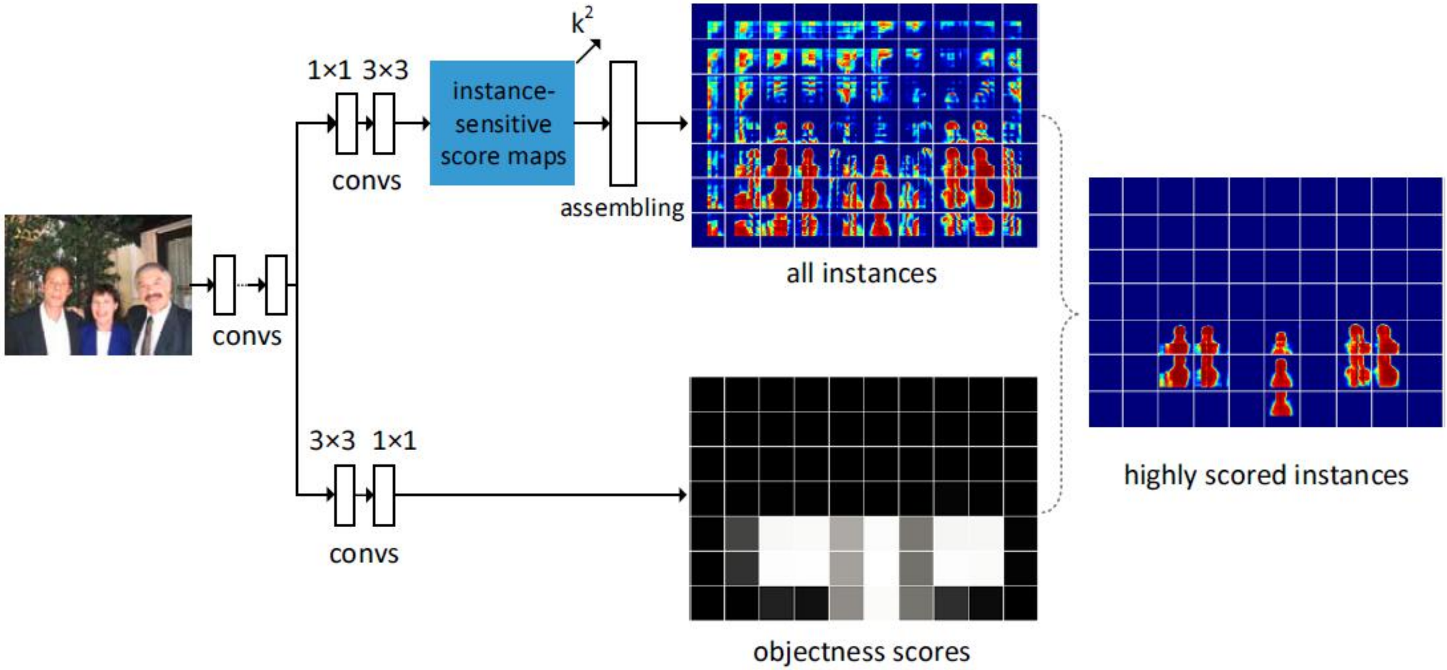
FCN with Fancier Label

training

vgg16 as base model

modify it with reduced stride at pool4, “hole algorithm” at conv5_1 and conv5_3

two fc branches for segmentation and scoring objectness



Dai et al., "Instance-sensitive Fully Convolutional Networks", ECCV, 2016

FCN with Fancier Label

training $$\sum_i \left(\mathcal{L}(p_i, p_i^*) + \sum_j \mathcal{L}(S_{i,j}, S_{i,j}^*)\right)$$

sampling for training $$600 \times 1.5^{\{-4,-3,-2,-1,0,1\}}$$

8-GPU, each for one image with 256 sampled windows -- batch-8

~1.5s for testing one image

NMS (0.8) for final set of proposals

Dai et al., "Instance-sensitive Fully Convolutional Networks", ECCV, 2016

FCN with Fancier Label

Quantitative Results

Table 2. Ablation comparisons between ~DeepMask and our method on the PASCAL VOC 2012 validation set. "~DeepMask" is our implementation based on controlled settings (see more descriptions in the main text).

method	train	test	AR@10 (%)	AR@100 (%)	AR@1000 (%)
~DeepMask	crop 224×224	sliding *fc*	31.2	42.9	47.0
ours	crop 224×224	fully conv.	37.4	48.4	51.4
	fully conv.	fully conv.	**38.9**	**49.7**	**52.6**

Table 3. Comparisons with state-of-the-art segment proposal methods on the PASCAL VOC 2012 validation set. The results of SS [6] and MCG [12] are from the publicly available code, and the results of MNC [20] is provided by the authors of [20].

method	AR@10 (%)	AR@100 (%)	AR@1000 (%)
SS [6]	7.0	23.5	43.3
MCG [12]	18.9	36.8	49.5
~DeepMask	31.2	42.9	47.0
MNC [20]	<u>33.4</u>	<u>48.5</u>	**53.8**
ours	**38.9**	**49.7**	<u>52.6</u>

UCIrvine
University of California, Irvine

Dai et al., "Instance-sensitive Fully Convolutional Networks", ECCV, 2016

FCN with Fancier Label

Quantitative Results

Table 5. Comparisons of instance segment proposals on the first 5k images [8] from the MS COCO validation set. DeepMask's results are from [8].

segment proposals	AR@10 (%)	AR@100 (%)	AR@1000 (%)
GOP [29]	2.3	12.3	25.3
Rigor [30]	-	9.4	25.3
SS [6]	2.5	9.5	23.0
MCG [7]	7.7	18.6	29.9
DeepMask [8]	12.6	24.5	33.1
DeepMaskZoom [8]	12.7	26.1	36.6
ours	**16.6**	**31.7**	**39.2**

Dai et al., "Instance-sensitive Fully Convolutional Networks", ECCV, 2016

FCN with Fancier Label

Qualitative Results

Dai et al., "Instance-sensitive Fully Convolutional Networks", ECCV, 2016

Conclusion

NO conclusion.

Thank you

Question & Answer

UCIrvine
UNIVERSITY OF CALIFORNIA, IRVINE

leaving blank

Content after this page is not suitable for people to watch!