#DLUPC

Day 2 Lecture 4

Instance Segmentation

http://bit.ly/dlcv2018

Míriam Bellver
miriam.bellver@bsc.edu

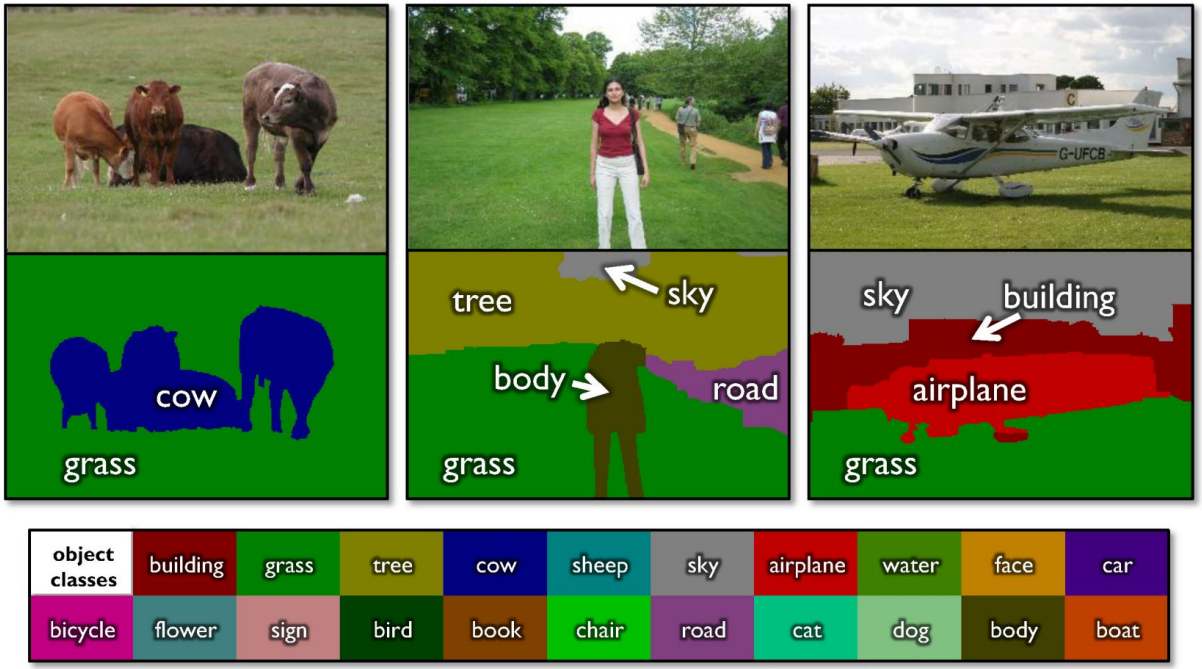
PhD Candidate
Barcelona Supercomputing Center

Semantic Segmentation

Label every pixel!

Don’t differentiate instances (cows)

Classic computer vision problem



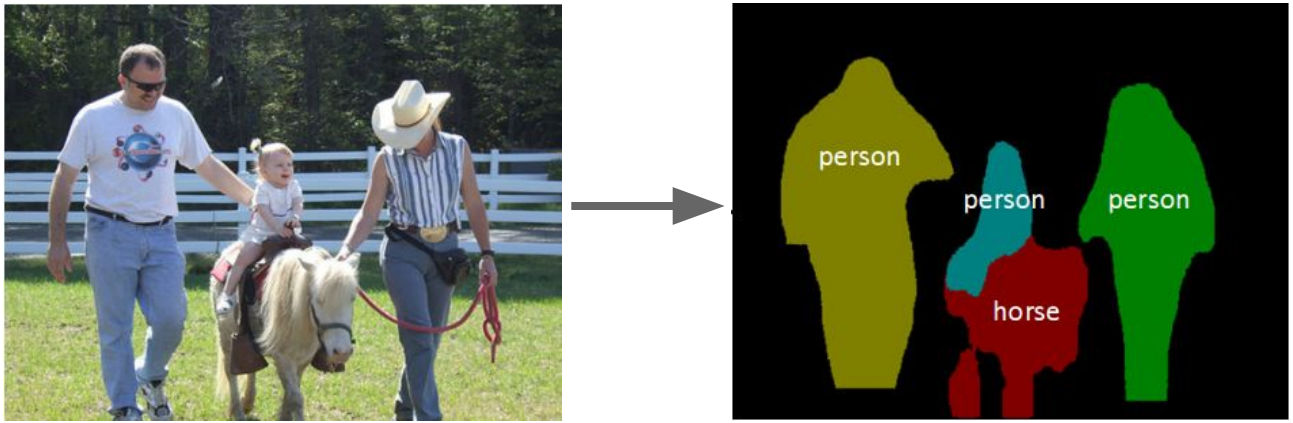
Slide Credit: CS231n

Instance Segmentation

Detect instances, give category, label pixels

“simultaneous detection and segmentation” (SDS)

Label are class-aware and instance-aware



Slide Credit: CS231n

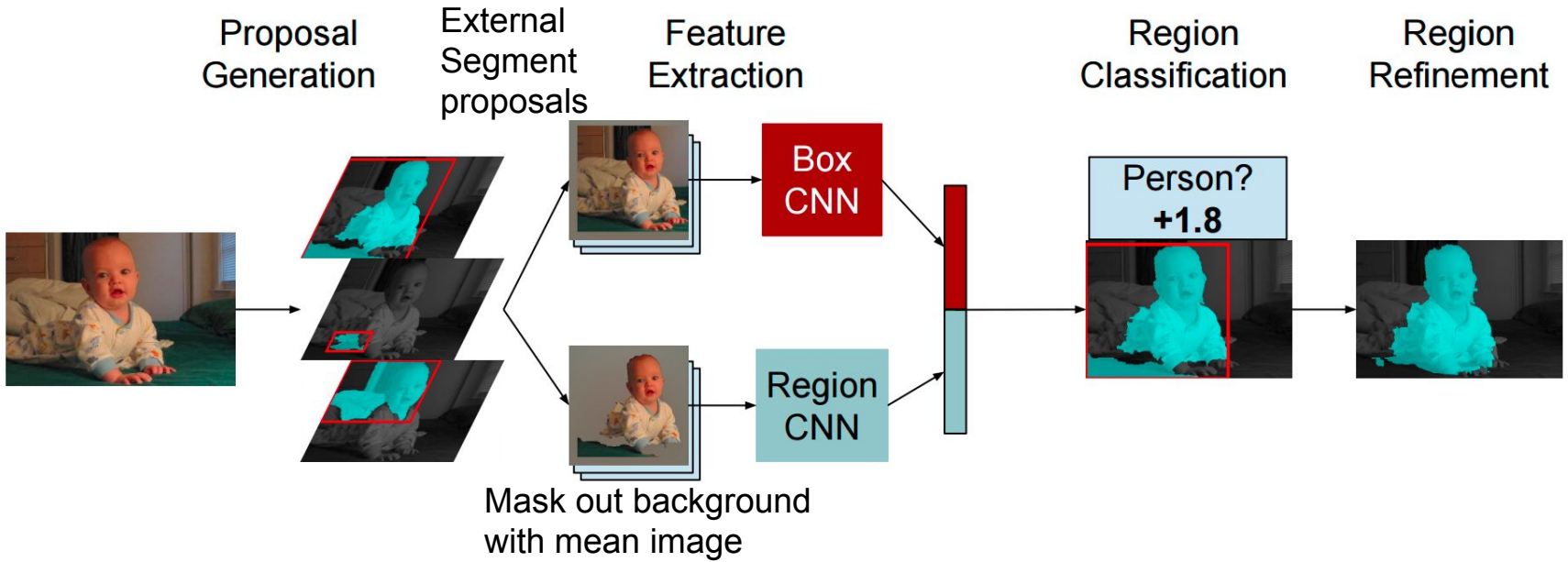
Outline

Instance Segmentation Methods

- Proposal-Based
- Recurrent
- Instance Embedding

Proposal-based

Similar to R-CNN, but with segment proposals

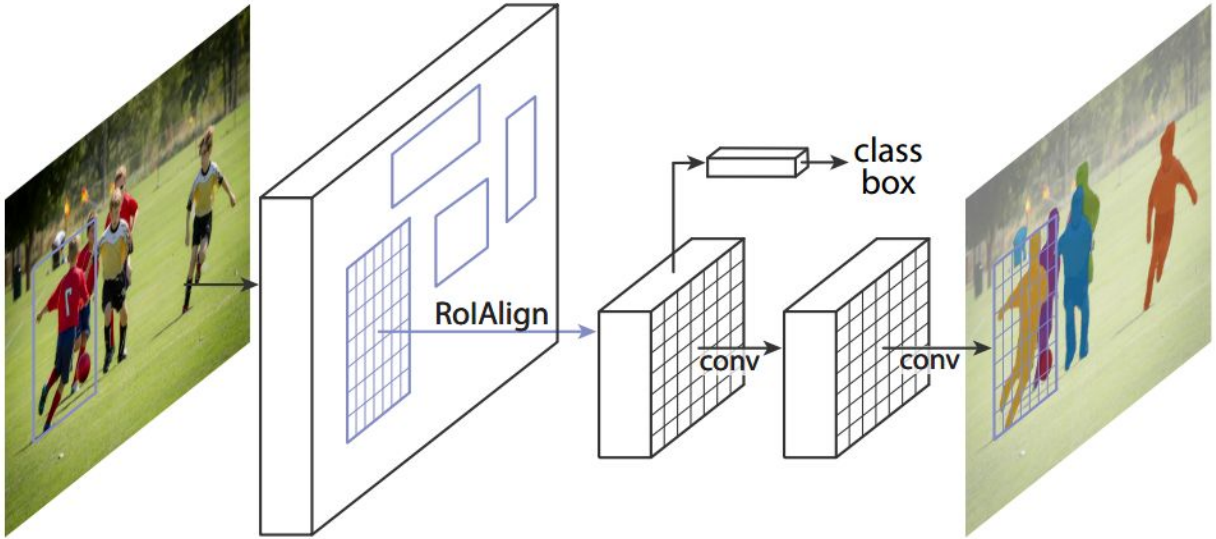


Hariharan et al. Simultaneous Detection and Segmentation. ECCV 2014

Slide Credit: CS231n

Proposal-based Instance Segmentation: Mask R-CNN

Faster R-CNN for Pixel Level Segmentation as a **parallel prediction of masks and class labels**

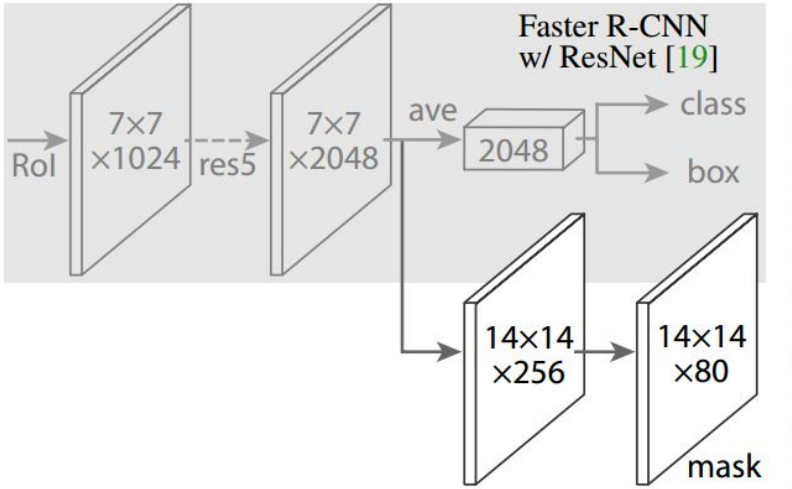


He et al. Mask R-CNN. ICCV 2017

Mask R-CNN

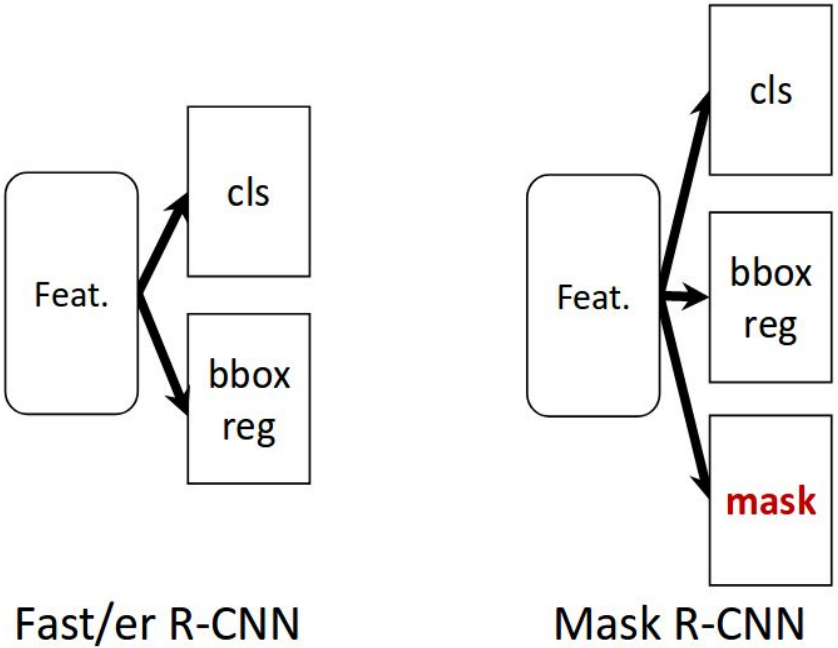
- Classification & box detection losses are identical to those in Faster R-CNN
- Addition of a new loss term for mask prediction:

The network outputs a *K x m x m* volume for mask prediction, where *K* is the number of categories and *m* is the size of the mask (square)



He et al. Mask R-CNN. ICCV 2017

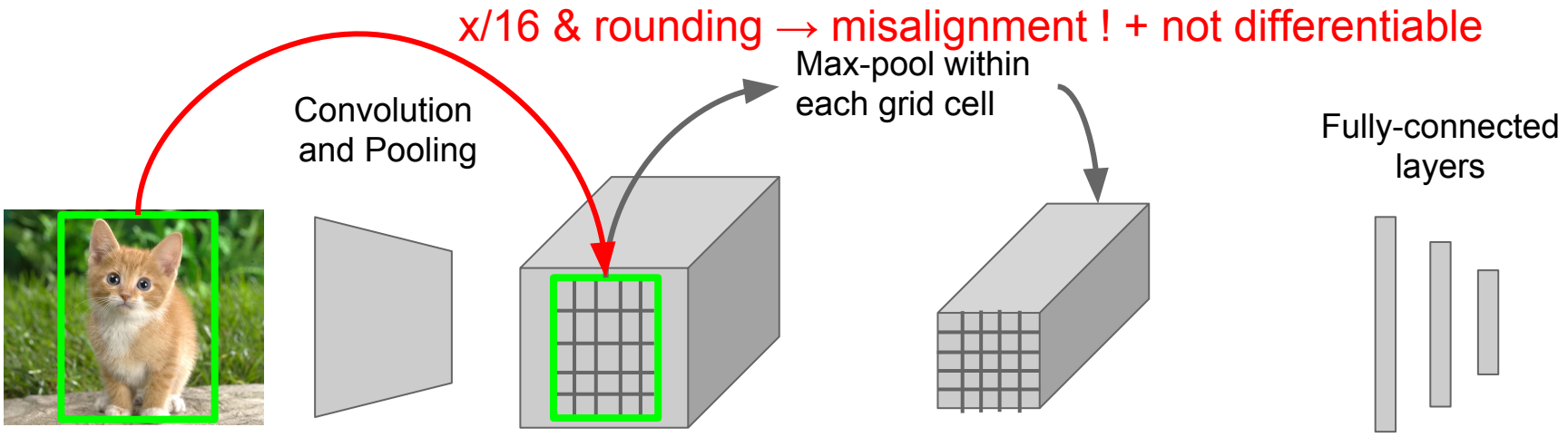
Mask R-CNN



He et al. Mask R-CNN. ICCV 2017

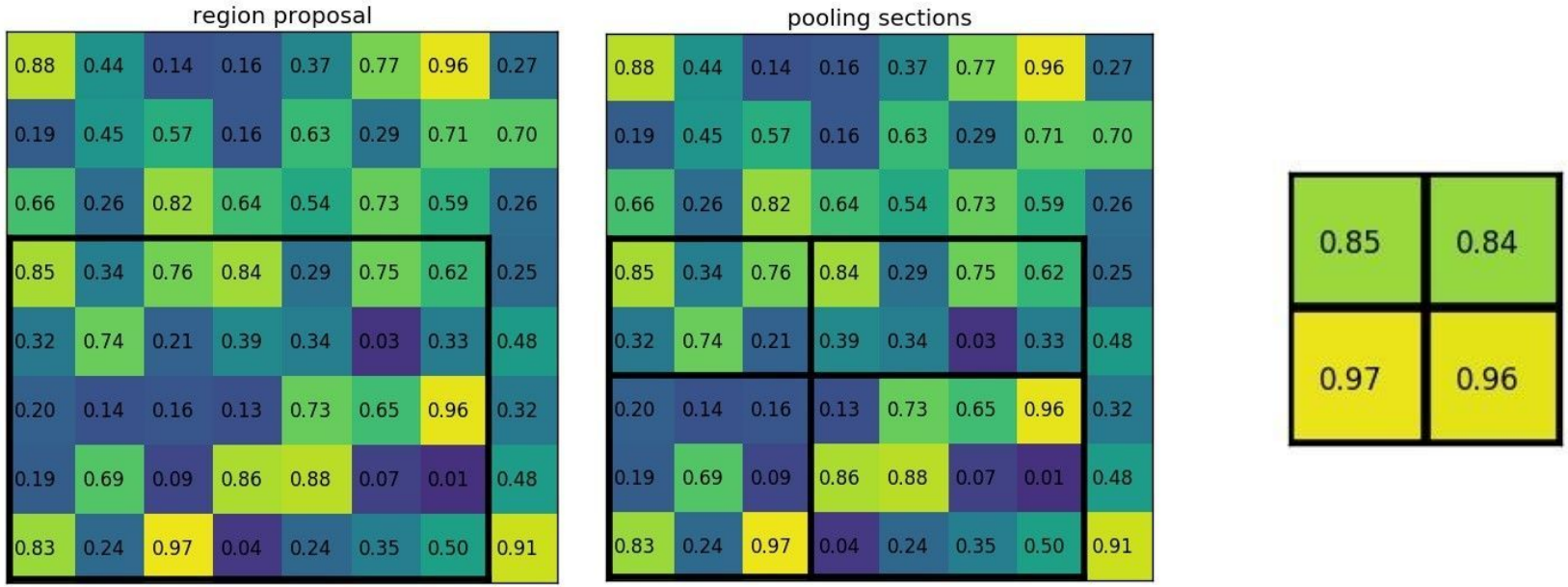
Mask R-CNN: RoI Align

Reminder: RoI Pool from Fast R-CNN



He et al. Mask R-CNN. ICCV 2017

Mask R-CNN: RoI Align

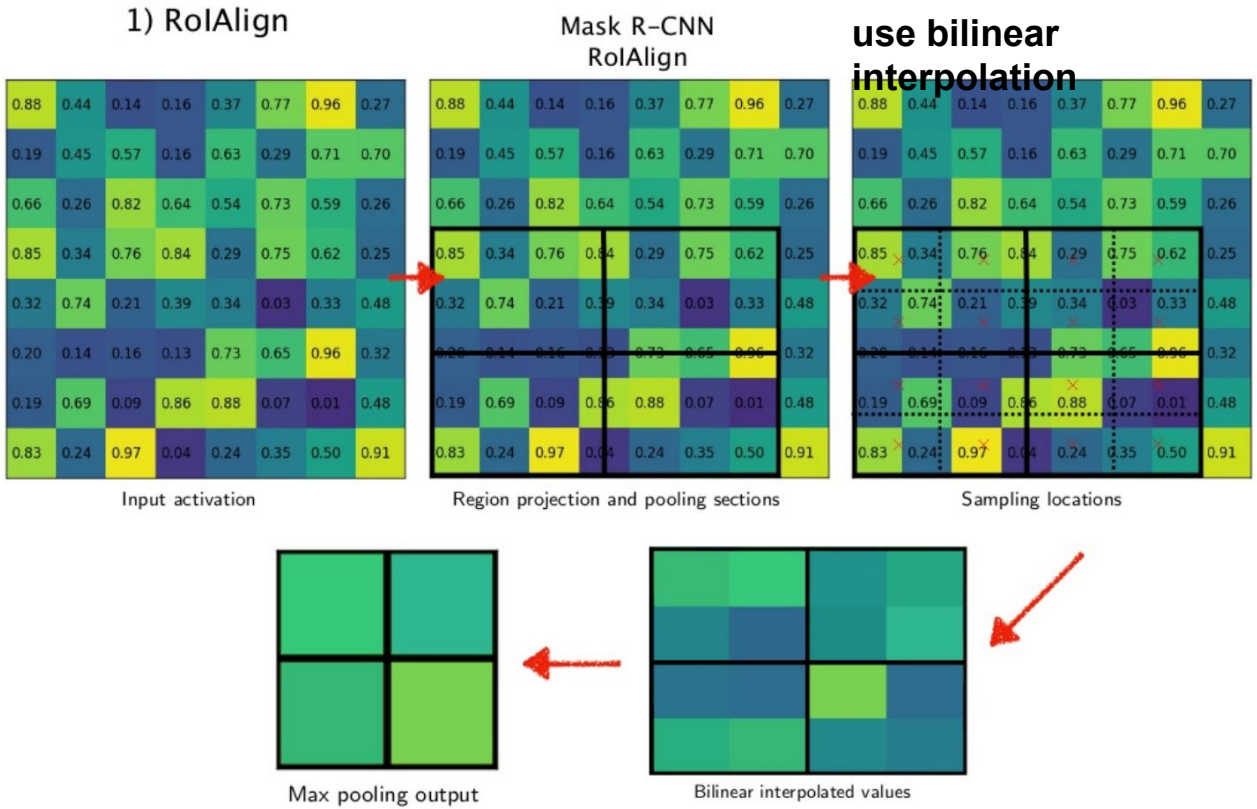


Roi Pooling example

He et al. Mask R-CNN. ICCV 2017

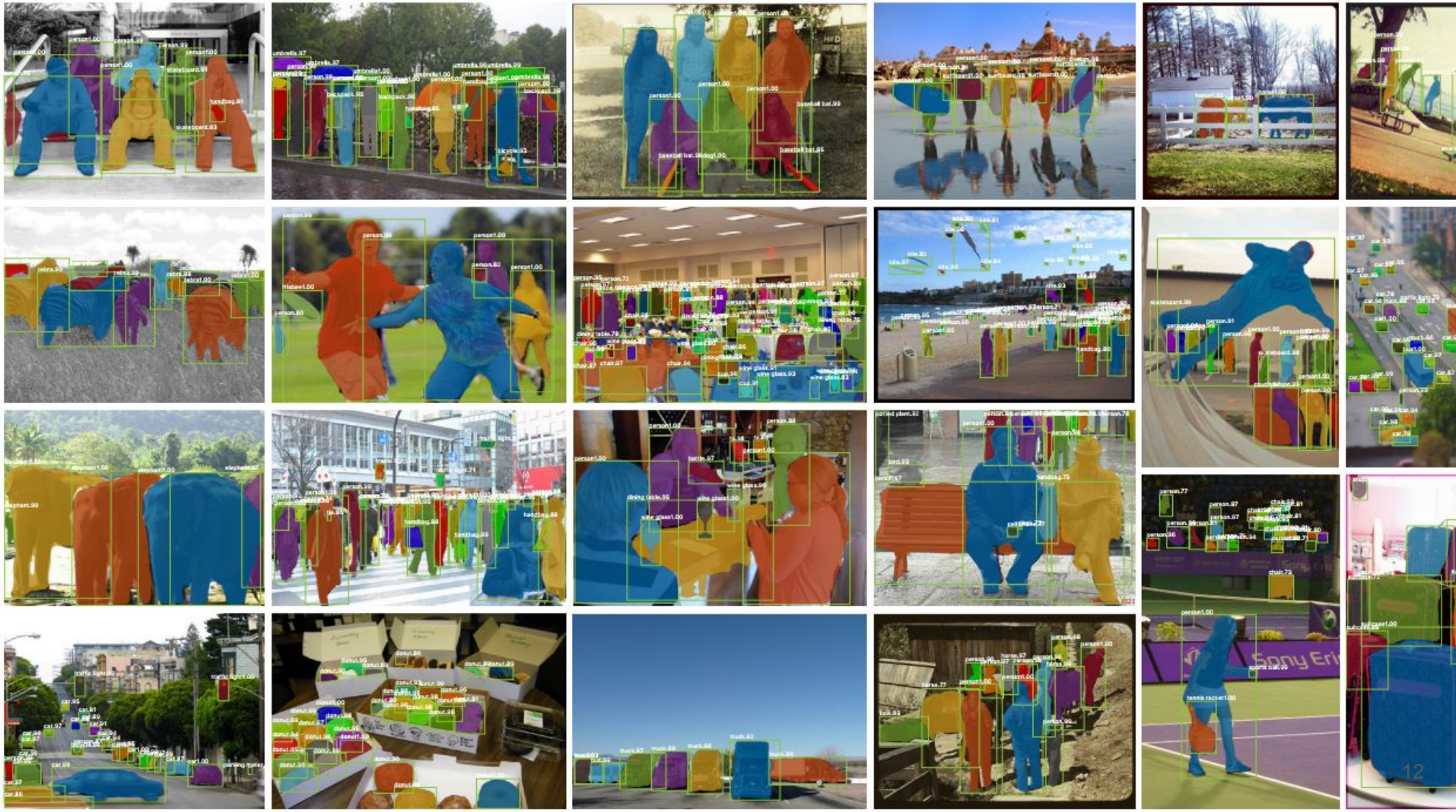
Figure Credit: Deepsense.ai

Mask R-CNN: RoI Align



He et al. Mask R-CNN. ICCV 2017

Figure Credit: Silvio Galesso

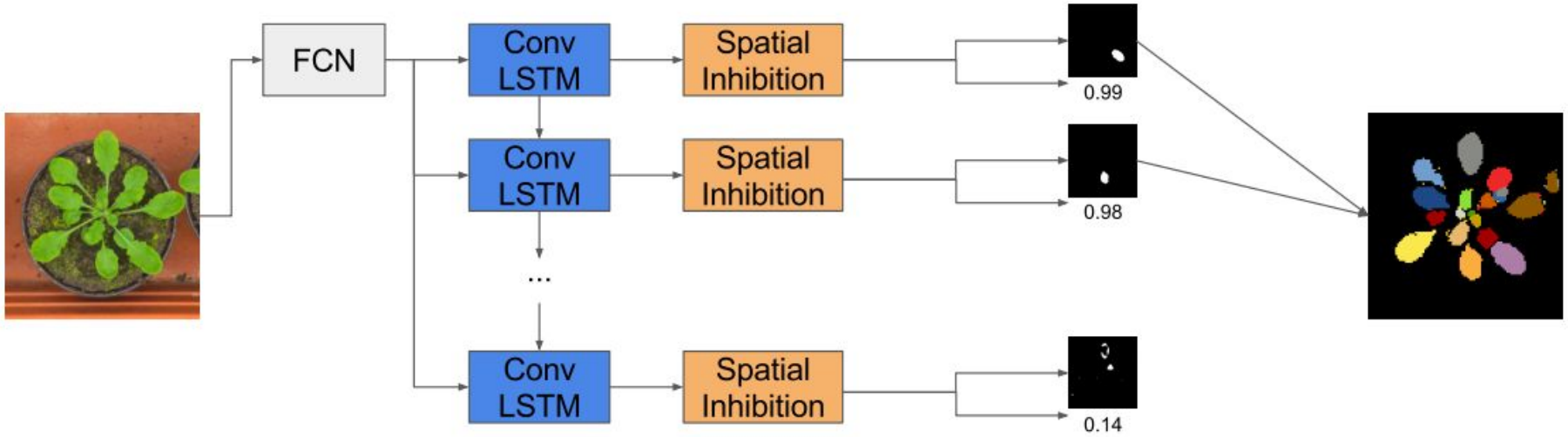


Limitations of Proposal-based models

1. Two objects might share the same bounding box: Only one will be kept after NMS step.
2. Choice of NMS threshold is application dependant
3. Same pixel can be assigned to multiple instances
4. Number of predictions is limited by the number of proposals.

Recurrent Instance Segmentation

Sequential mask generation

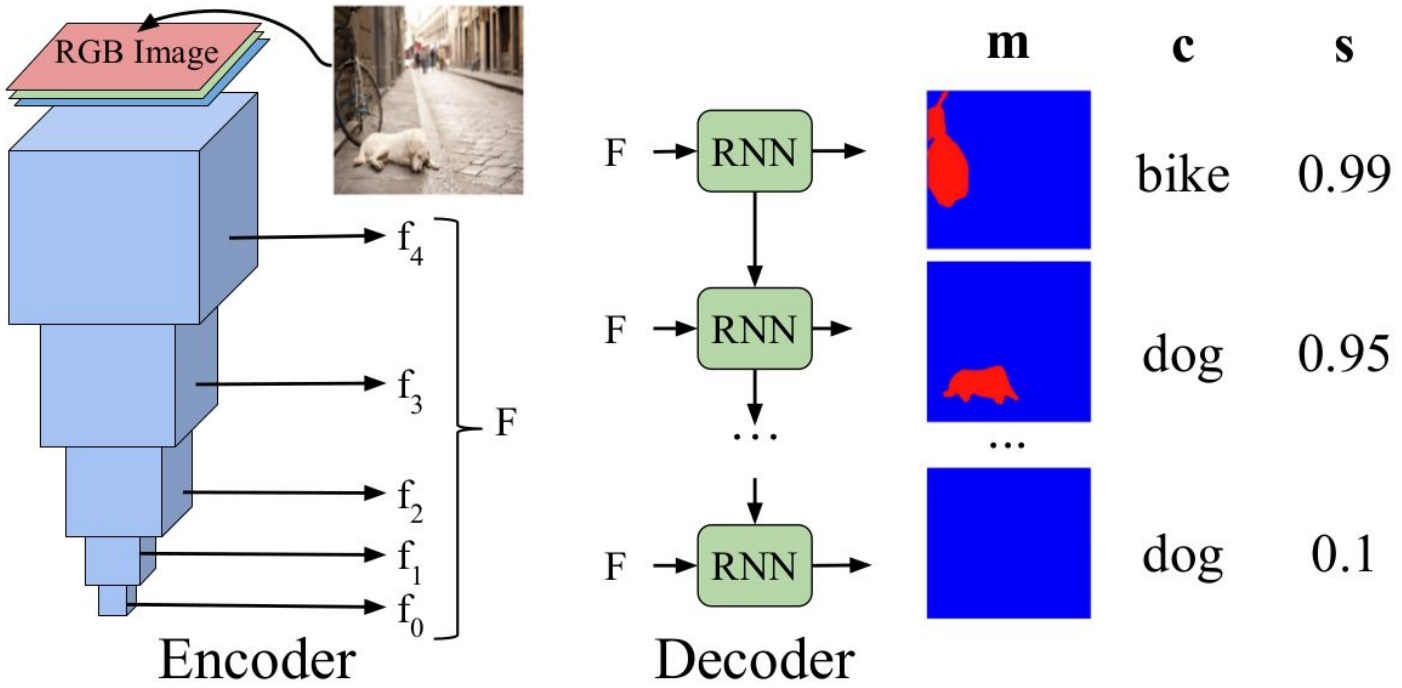


Romera-Paredes & H.S. Torr. Recurrent Instance Segmentation ECCV 2016

Recurrent Instance Segmentation

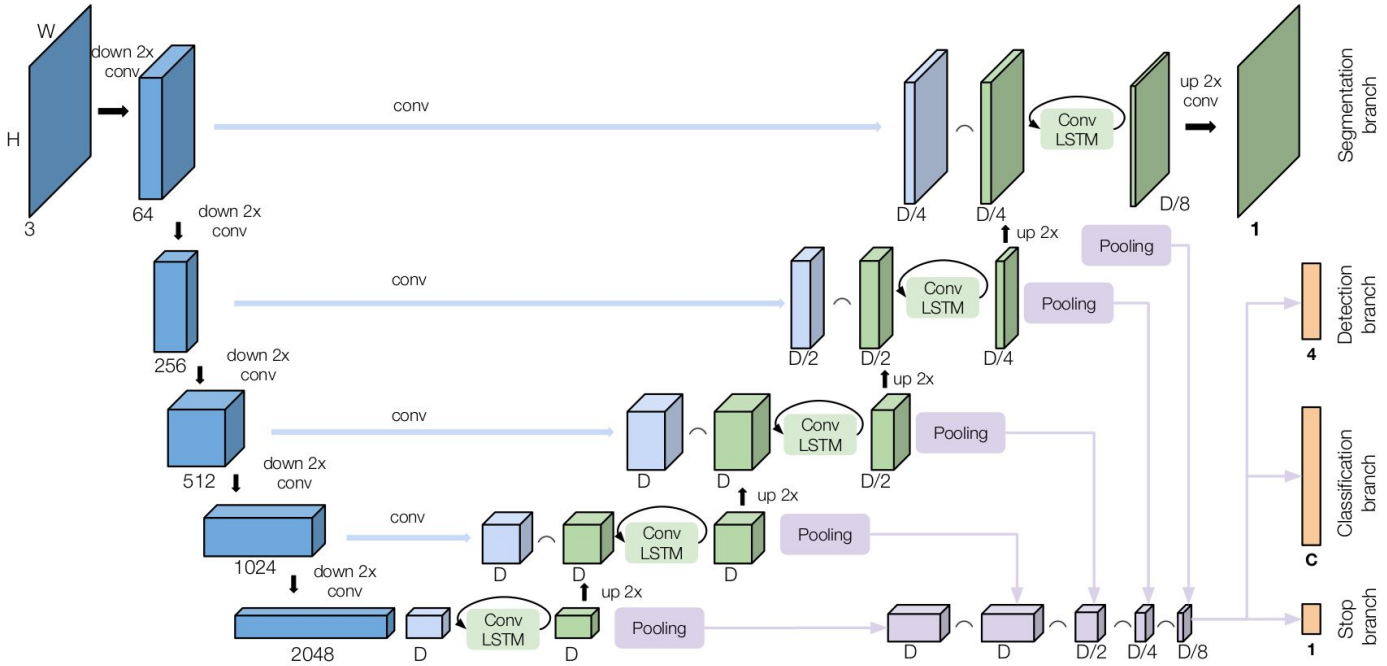
Romera-Paredes & H.S. Torr. Recurrent Instance Segmentation ECCV 2016

Recurrent Semantic Instance Segmentation



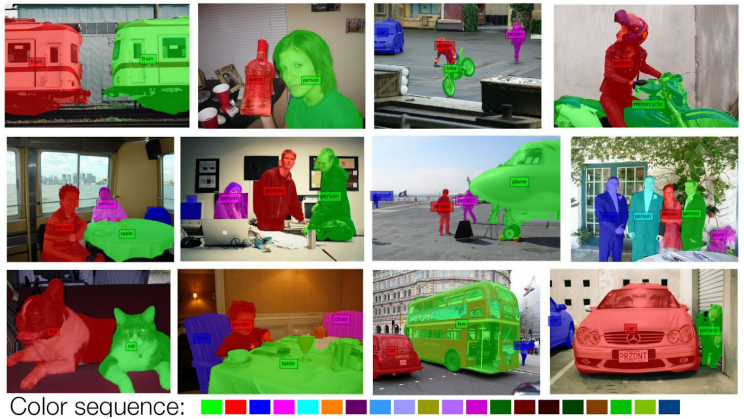
Salvador, A., Bellver, Campos. V, M., Baradad, M., Marqués, F., Torres, J., & Giro-i-Nieto, X. (2018) From Pixels to Object Sequences: Recurrent Semantic Instance Segmentation.

Recurrent Semantic Instance Segmentation

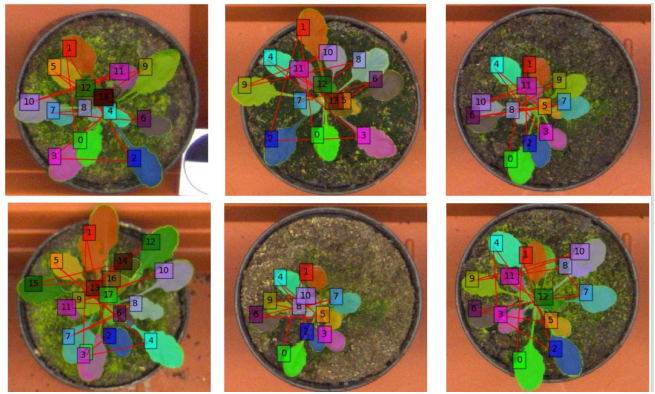


Salvador, A., Bellver, Campos. V, M., Baradad, M., Marqués, F., Torres, J., & Giro-i-Nieto, X. (2018) From Pixels to Object Sequences: Recurrent Semantic Instance Segmentation.

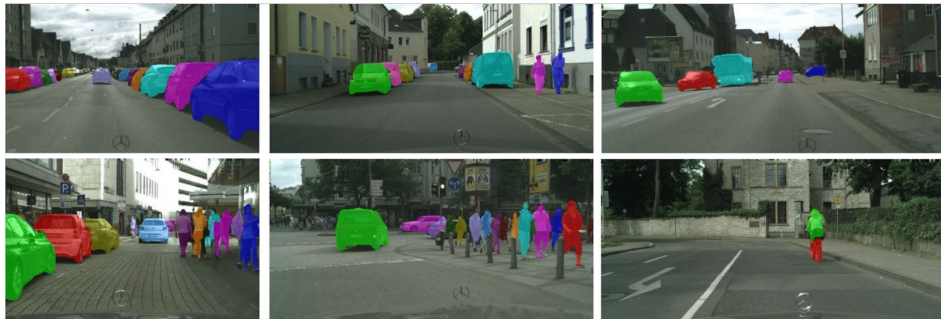
Recurrent Semantic Instance Segmentation



Pascal VOC



CVPPP



Cityscapes

Salvador, A., Bellver, Campos. V, M., Baradad, M., Marqués, F., Torres, J., & Giro-i-Nieto, X. (2018) From Pixels to Object Sequences: Recurrent Semantic Instance Segmentation.

Recurrent Semantic Instance Segmentation

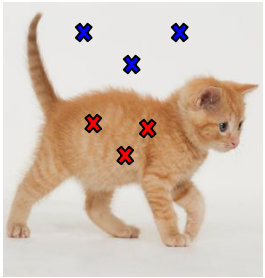
Object discovery patterns

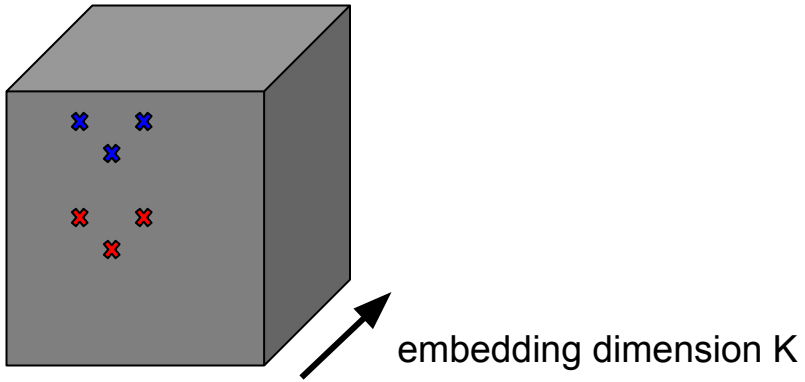
Salvador, A., Bellver, Campos. V, M., Baradad, M., Marqués, F., Torres, J., & Giro-i-Nieto, X. (2018) From Pixels to Object Sequences: Recurrent Semantic Instance Segmentation.

Instance Embedding

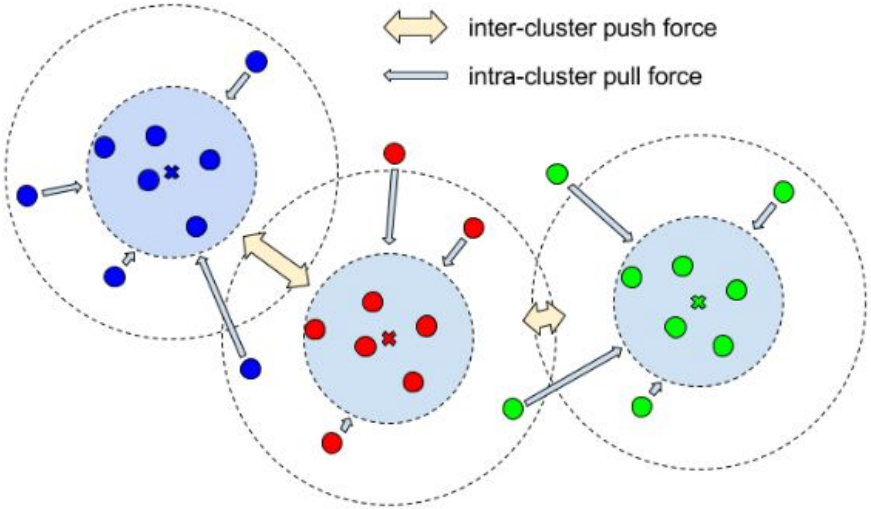
Basis of Instance Embedding Segmentation

- Each pixel in the output of the network is a point in the embedding space
- Pixel belonging to same object are close in the embedding space, and pixels belonging to different objects, are distant
- Parsing the image embeddings involves some **clustering**





Instance Embedding

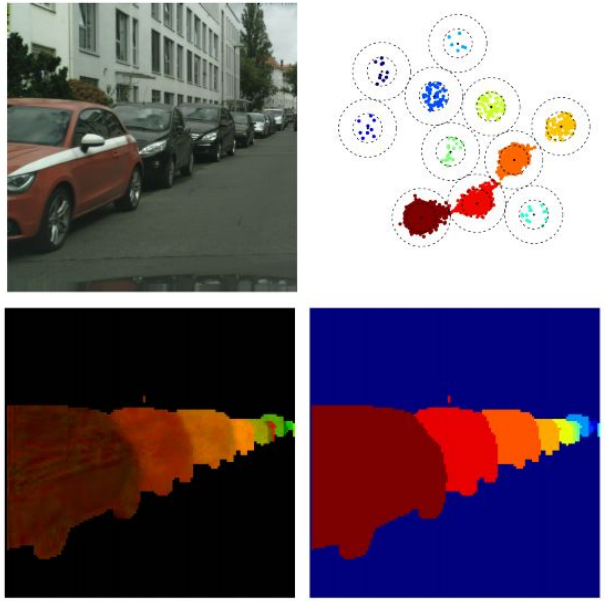


1. **variance term**: an intra-cluster pull-force that draws embeddings towards the mean embedding, i.e. the cluster center.
2. **distance term**: an inter-cluster push-force that pushes clusters away from each other, increasing the distance between the cluster centers.
3. **regularization term**: a small pull-force that draws all clusters towards the origin, to keep the activations bounded.

Brabandere et al. Semantic Instance Segmentation with a Discriminative Loss Function. CVPRW 2017

Instance Embedding

Mapping pixels to a N-dimensional space where pixels belonging to the same object are close to each other.



	AP	AP0.5	AP100m	AP50m
R-CNN+MCG	4.6	12.9	7.7	10.3
FCN+Depth	8.9	21.1	15.3	16.7
JGD	9.8	23.2	16.8	20.3
InstanceCut	13.0	27.9	22.1	26.1
Boundary-aware	17.4	36.7	29.3	34.0
DWT	19.4	35.3	31.4	36.8
Pixelwise DIN	20.0	38.8	32.6	37.6
Mask R-CNN	26.2	49.9	37.6	40.1
Ours	17.5	35.9	27.8	31.0

Results on Cityscapes

Brabandere et al. Semantic Instance Segmentation with a Discriminative Loss Function. CVPRW 2017

Outline

Instance Segmentation Methods

- Proposal-Based
- Recurrent
- Instance Embedding

Questions?