CNNs in Medicine: Beyond Classification

June 27, 2020

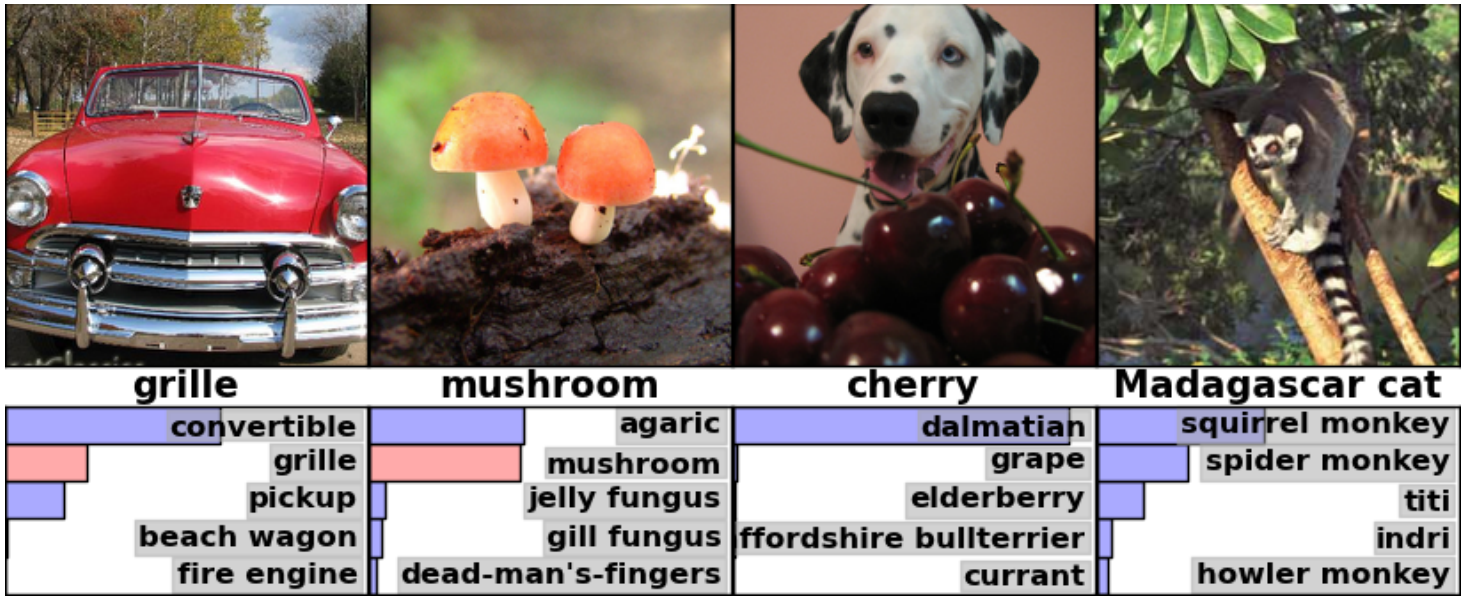
MMCi Applied Data Science

Matthew Engelhard

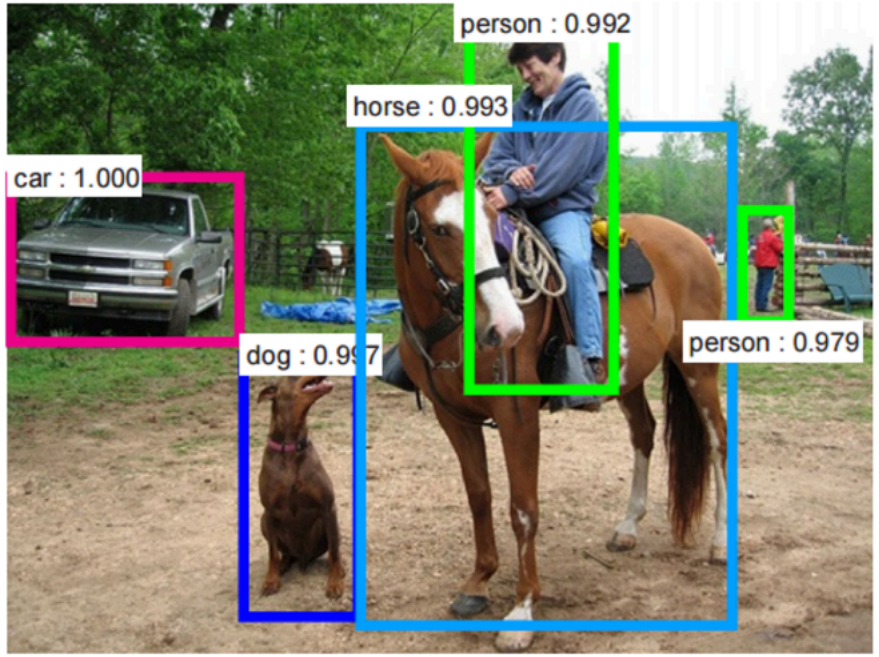
Find objects within images and identify them

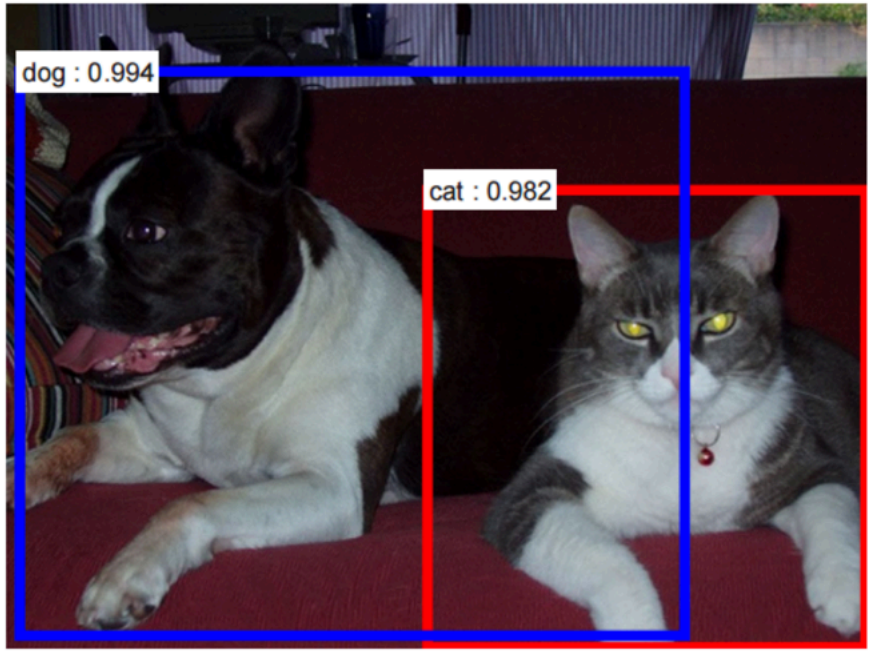
OBJECT DETECTION

Which label is best?

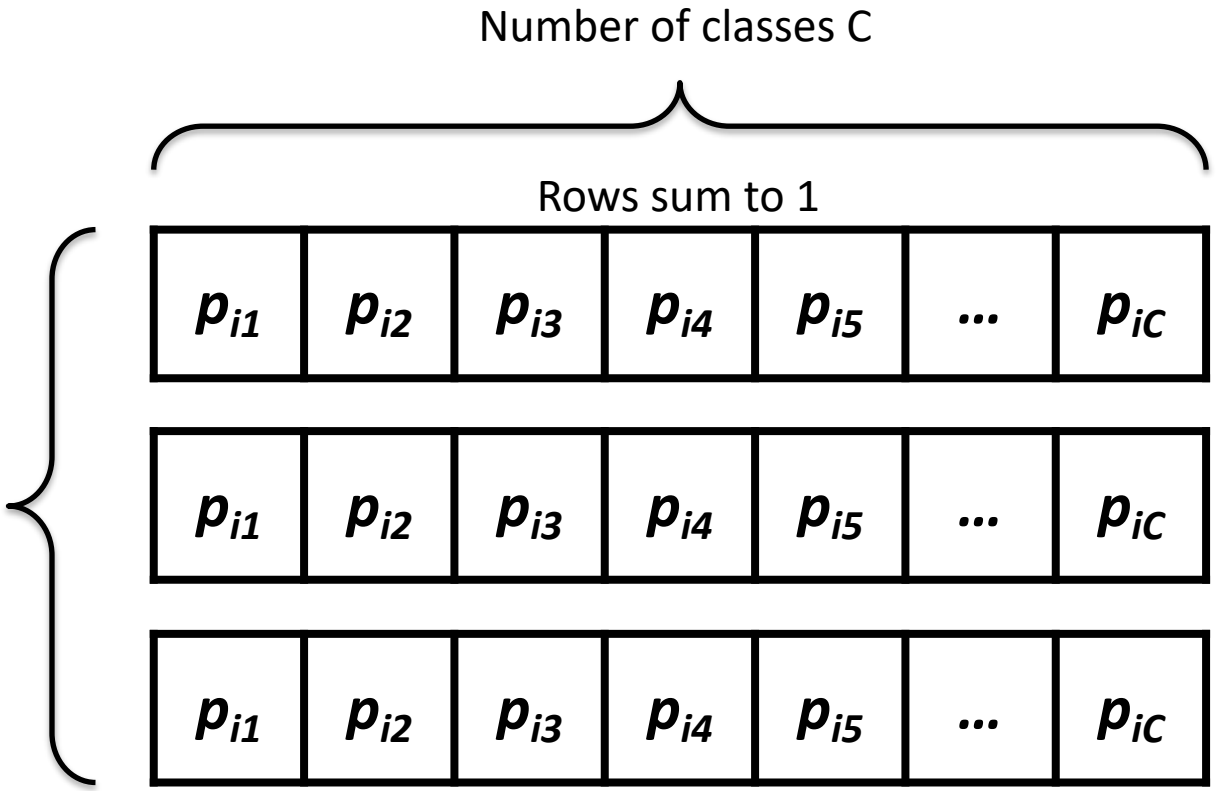


Detection: propose regions and predict their labels





Detection Output

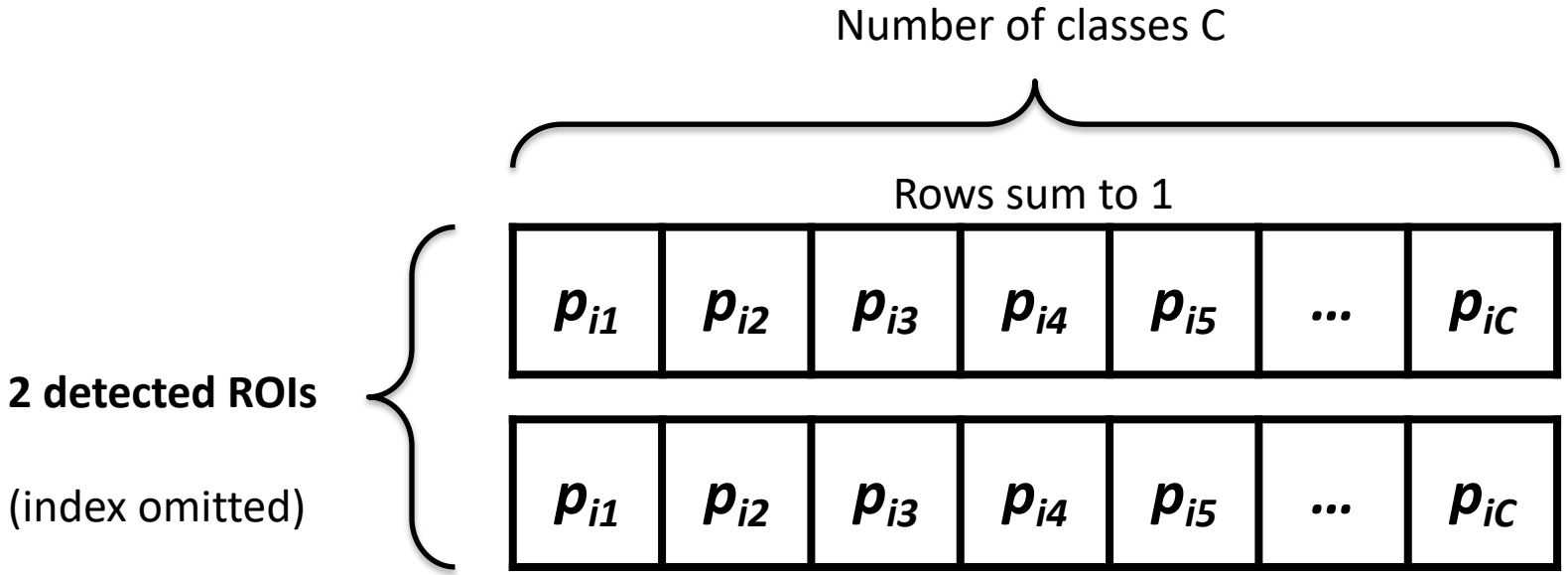


Variable Number of Detected Regions of Interest

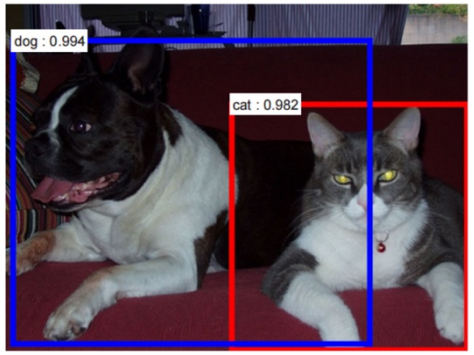
(index omitted)

$$p_{ij} = p(y_i = j|x_i)$$

Detection Output



$$p_{ij} = p(y_i = j|x_i)$$



Detection Output

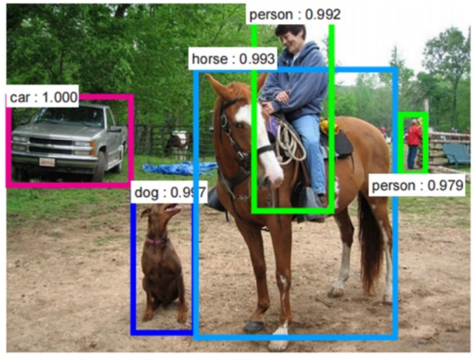
Number of classes C

Rows sum to 1

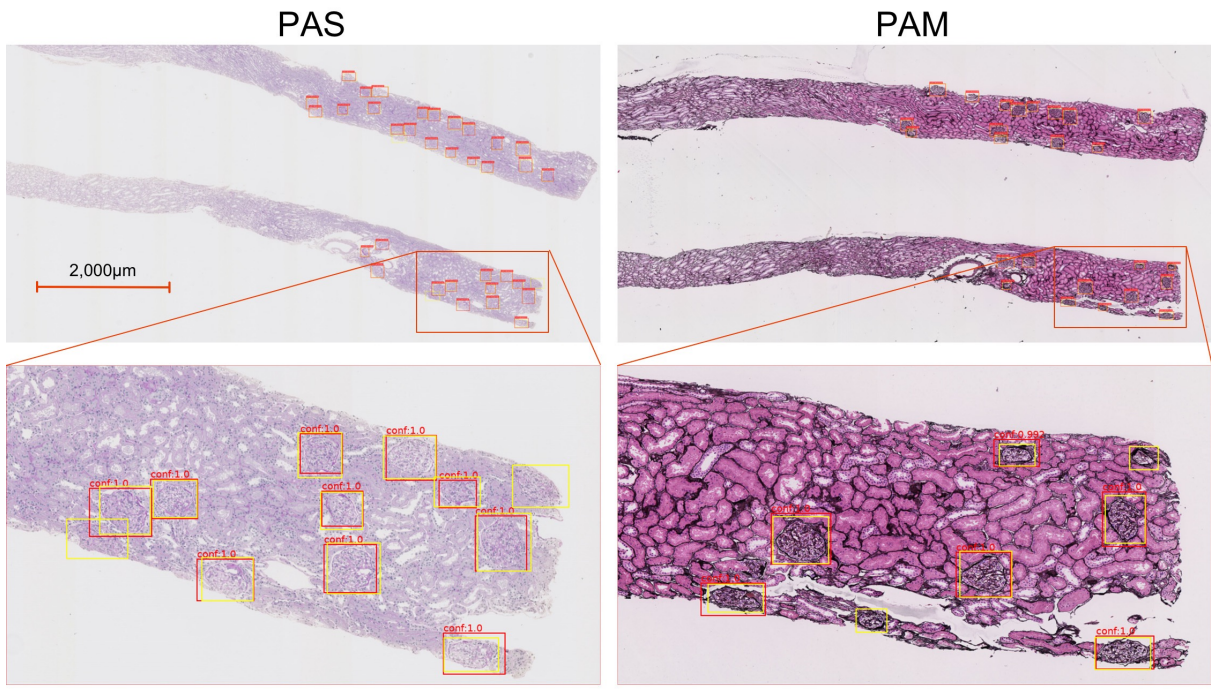
5 detected ROIs

(index omitted)

p_{i1}	p_{i2}	p_{i3}	p_{i4}	p_{i5}	…	p_{iC}
p_{i1}	p_{i2}	p_{i3}	p_{i4}	p_{i5}	…	p_{iC}
p_{i1}	p_{i2}	p_{i3}	p_{i4}	p_{i5}	…	p_{iC}
p_{i1}	p_{i2}	p_{i3}	p_{i4}	p_{i5}	…	p_{iC}
p_{i1}	p_{i2}	p_{i3}	p_{i4}	p_{i5}	…	p_{iC}

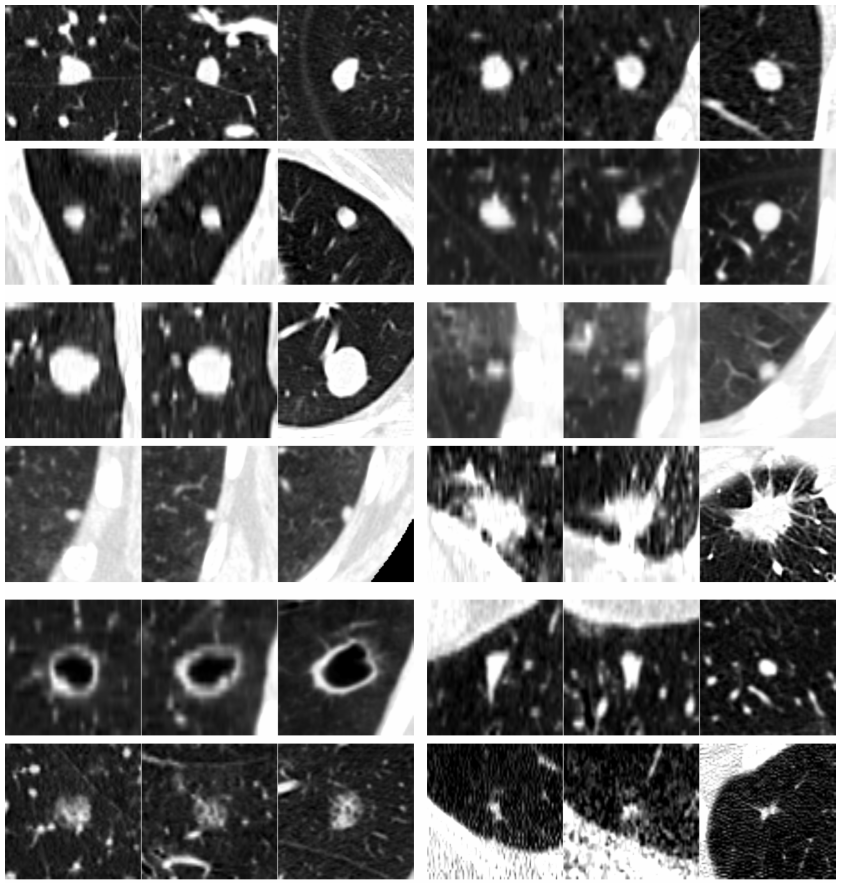


Detection in medicine



Glomerular Detection with Faster-RCNN

Kawazoe et al., *J. Imaging, 2018*



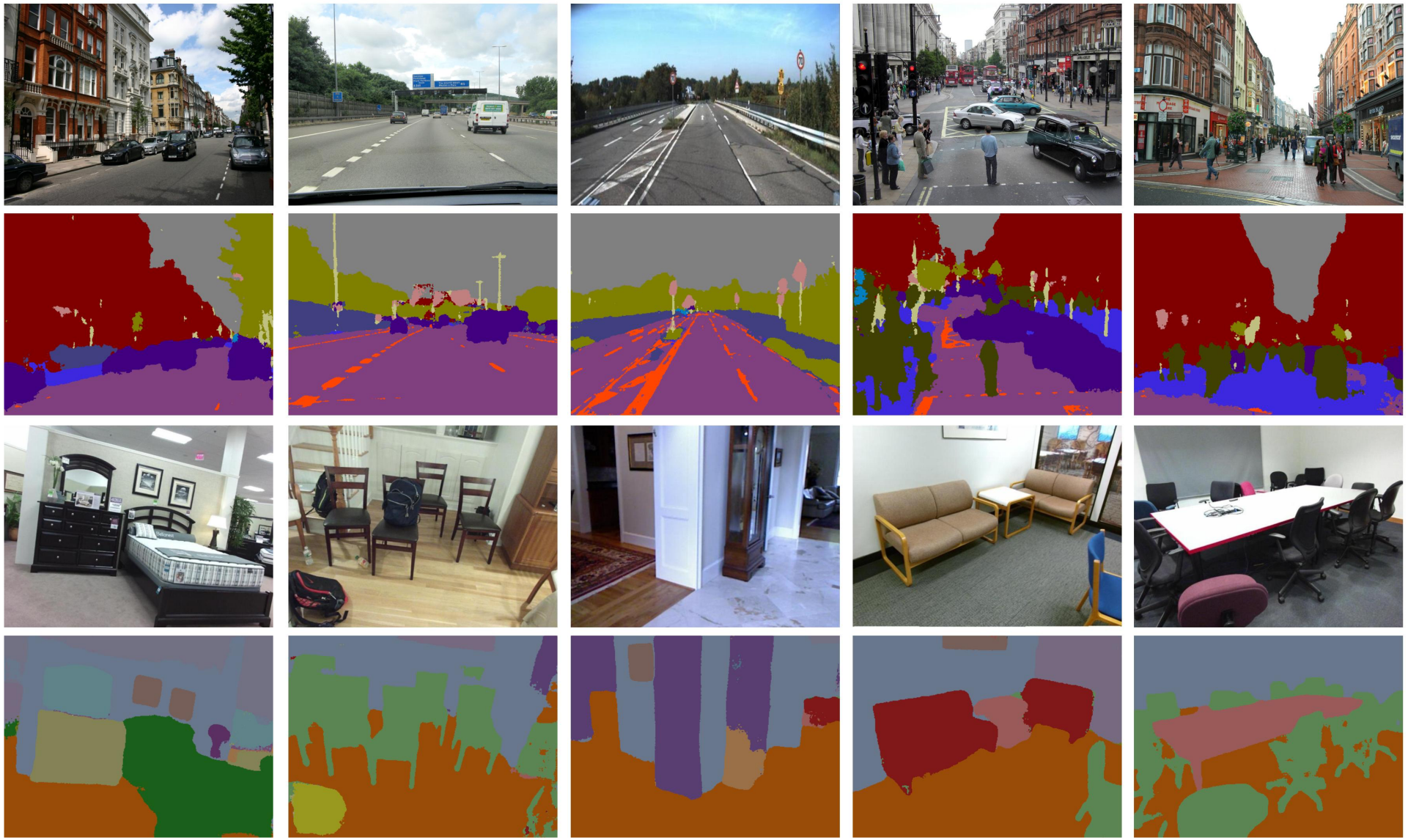
Pulmonary Nodule detection in CT

van Ginneken et al., *Biomedical Imaging,* 2015

Divide the image into meaningful regions

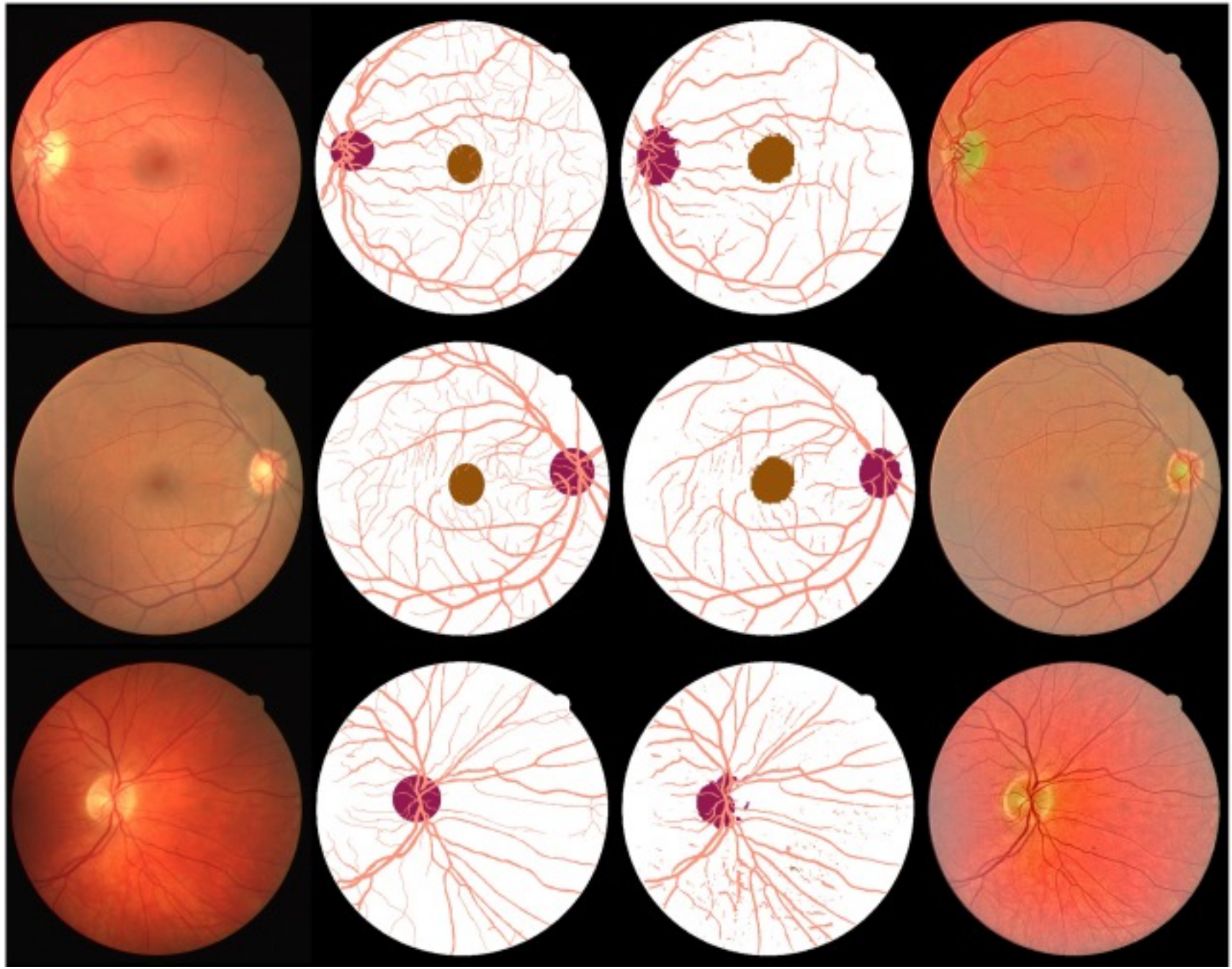
SEMANTIC SEGMENTATION

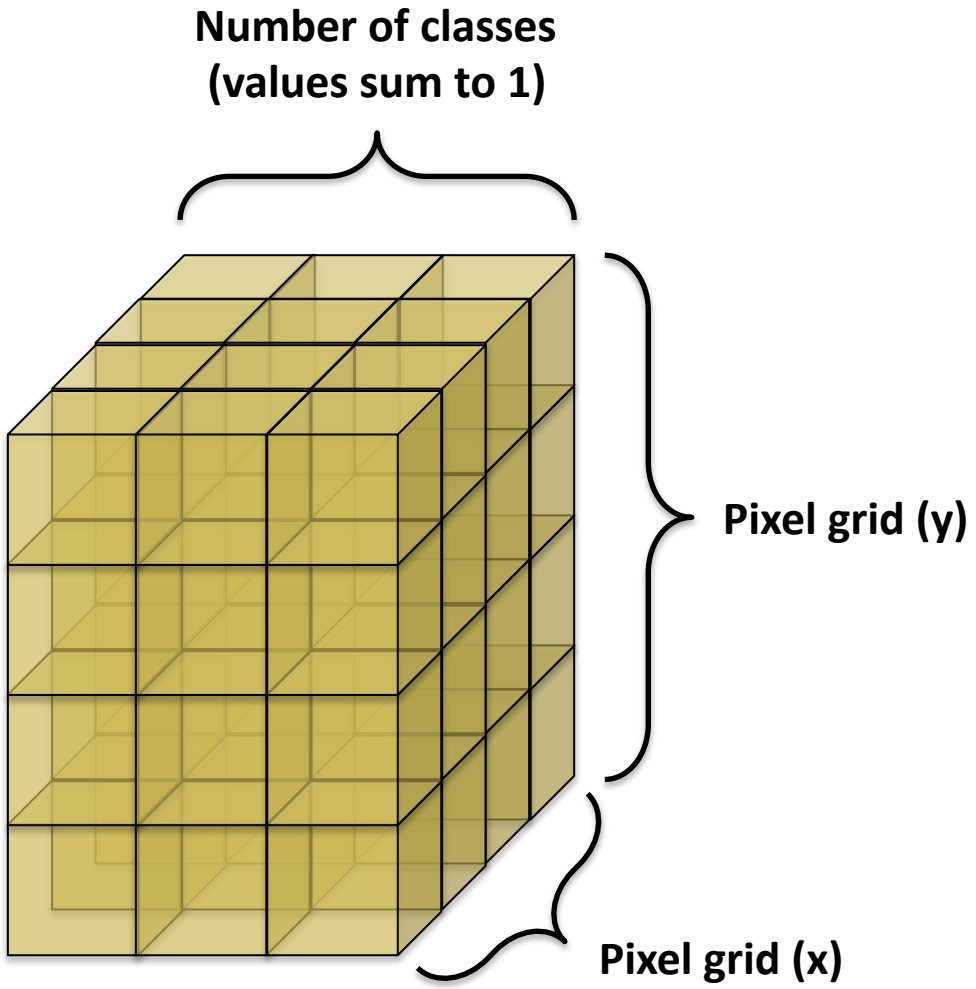
Segmentation: predict the label for each pixel



Segmentation of optic disc, fovea and retinal vasculature

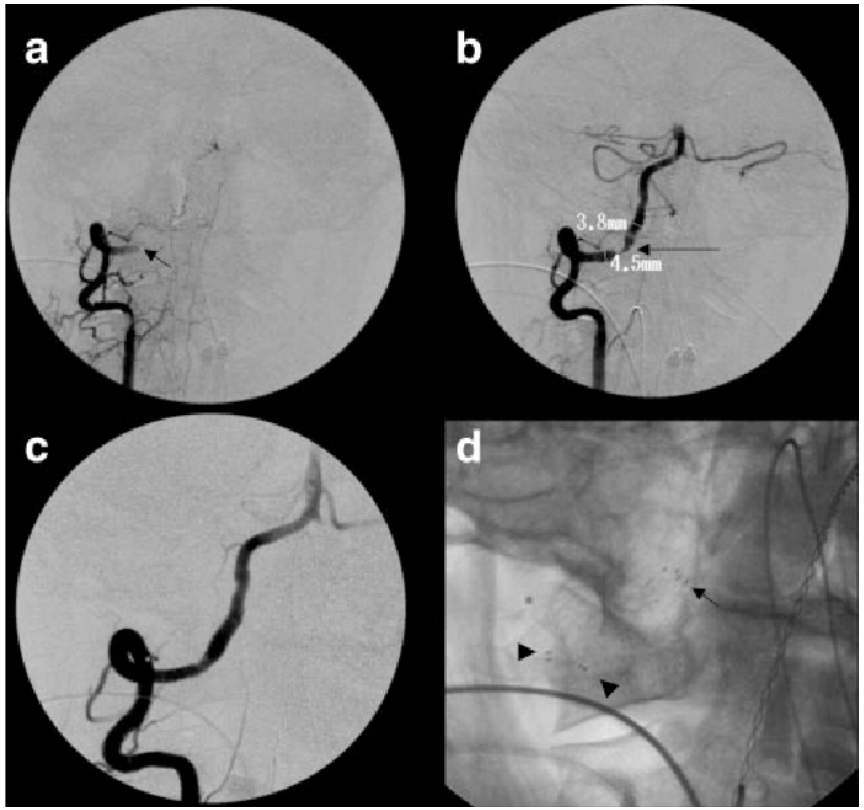
Journal of Computational Science, *20*, 70-79 (2017).



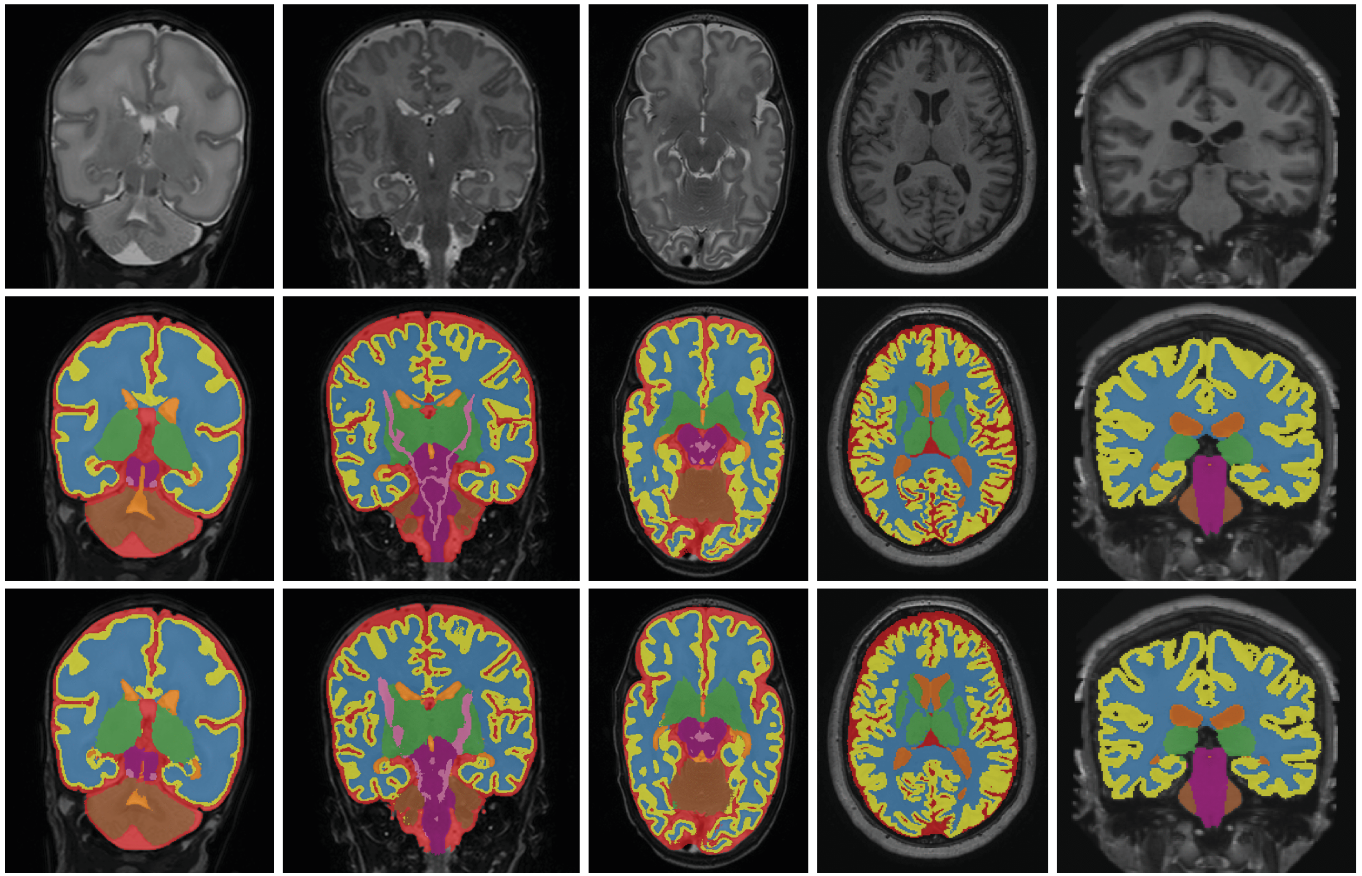


Segmentation Output: Distribution over Classes *for each pixel*

Precisely Identify Boundaries



Determine Areas or Volumes



Segmentation-based features when end-to-end classification is not feasible

Segment Bones → Calculate lengths and areas → Simple Classifier for Genetic or Growth Disorder

nature biomedical engineering

Article | Published: 10 October 2018

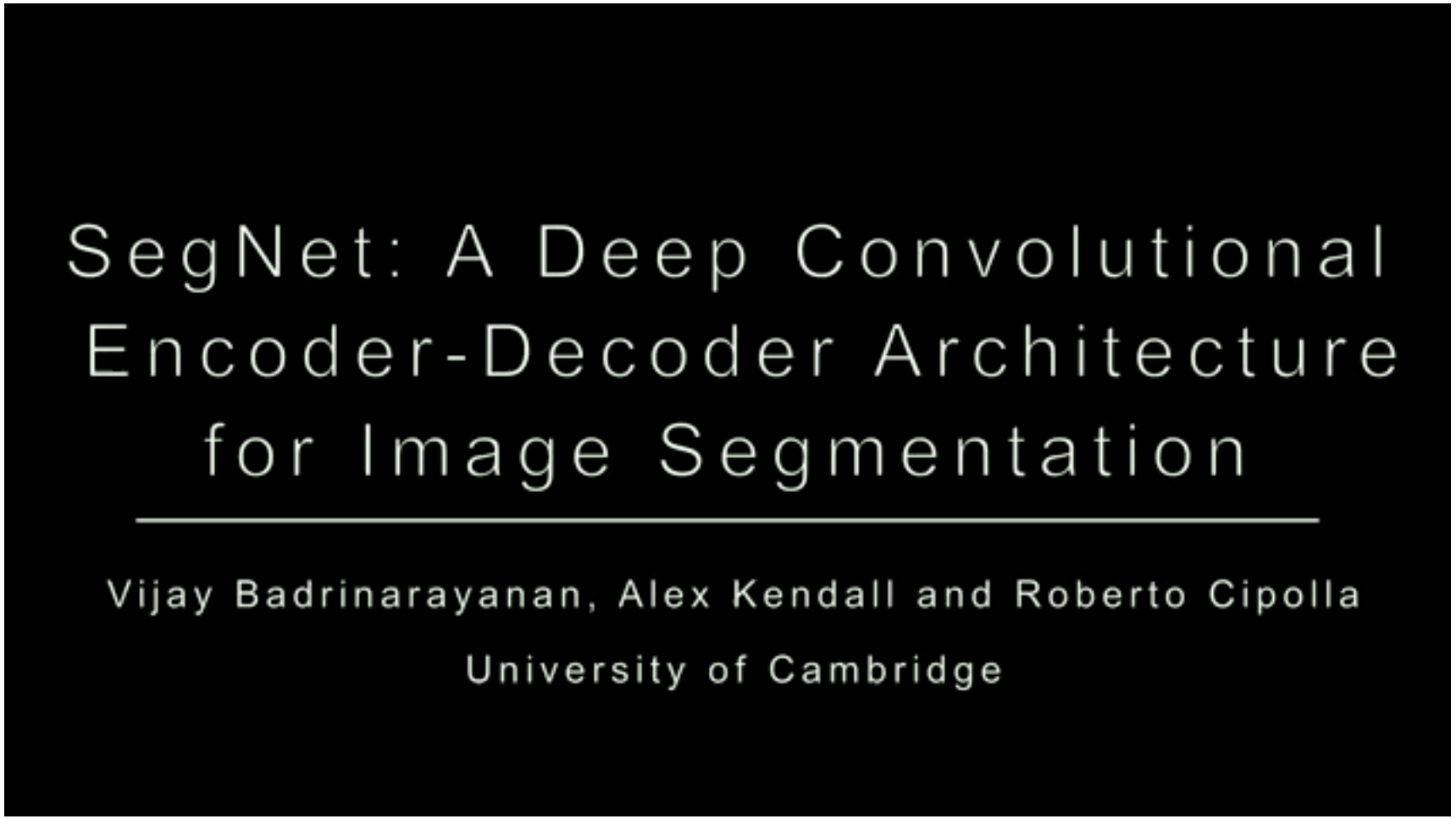
Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy

Pu Wang, Xiao Xiao, Jeremy R. Glissen Brown, Tyler M. Berzin, Mengtian Tu, Fei Xiong, Xiao Hu, Peixi Liu, Yan Song, Di Zhang, Xue Yang, Liangping Li, Jiong He, Xin Yi, Jingjia Liu & Xiaogang Liu

Nature Biomedical Engineering **2**, 741–748 (2018) | Download Citation

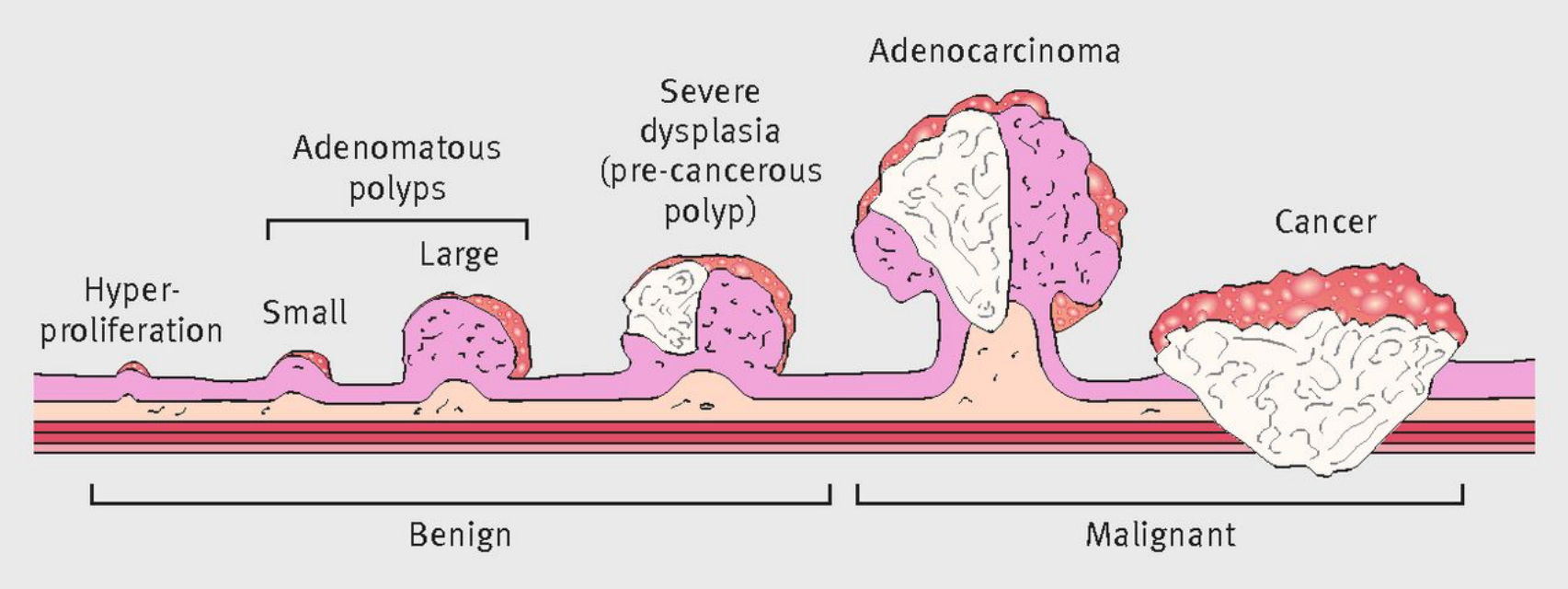
COLON POLYP SEGMENTATION

Approach: Start with SegNet (2015)



Retrain to segment polyps in real time

Focus on Adenomatous Polyps



BMJ 2016;354:i3590

Why are polyps missed?

1. Never in the visual field

2. In the field, but not recognized

-> Segment polyps in real time to act as an "extra set of eyes" during colonoscopy

4 validation datasets!!

A: new imageDB

Sichuan Provincial People's Hospital of China, 01-12/16

27,113 images from 1,138 consecutive patients who had polyps

20% of images had polyps

Confirmed histology

Excluded poor-quality images (by panel) and >2cm masses

B: public imageDB

Hospital Clinic of Barcelona (public CVC-ClinicDB)

612 image frames from 29 colonoscopy videos

C: polyp videos

Sichuan Provincial People's Hospital of China, 05-07/16

videos of 138 consecutively encountered polyps from 110 polyp patients

Confirmed histology

Used to assess **sensitivity**

D: complete videos

Sichuan Provincial People's Hospital, 01-02/18

54 unaltered full-length colonoscopy videos from 54 patients

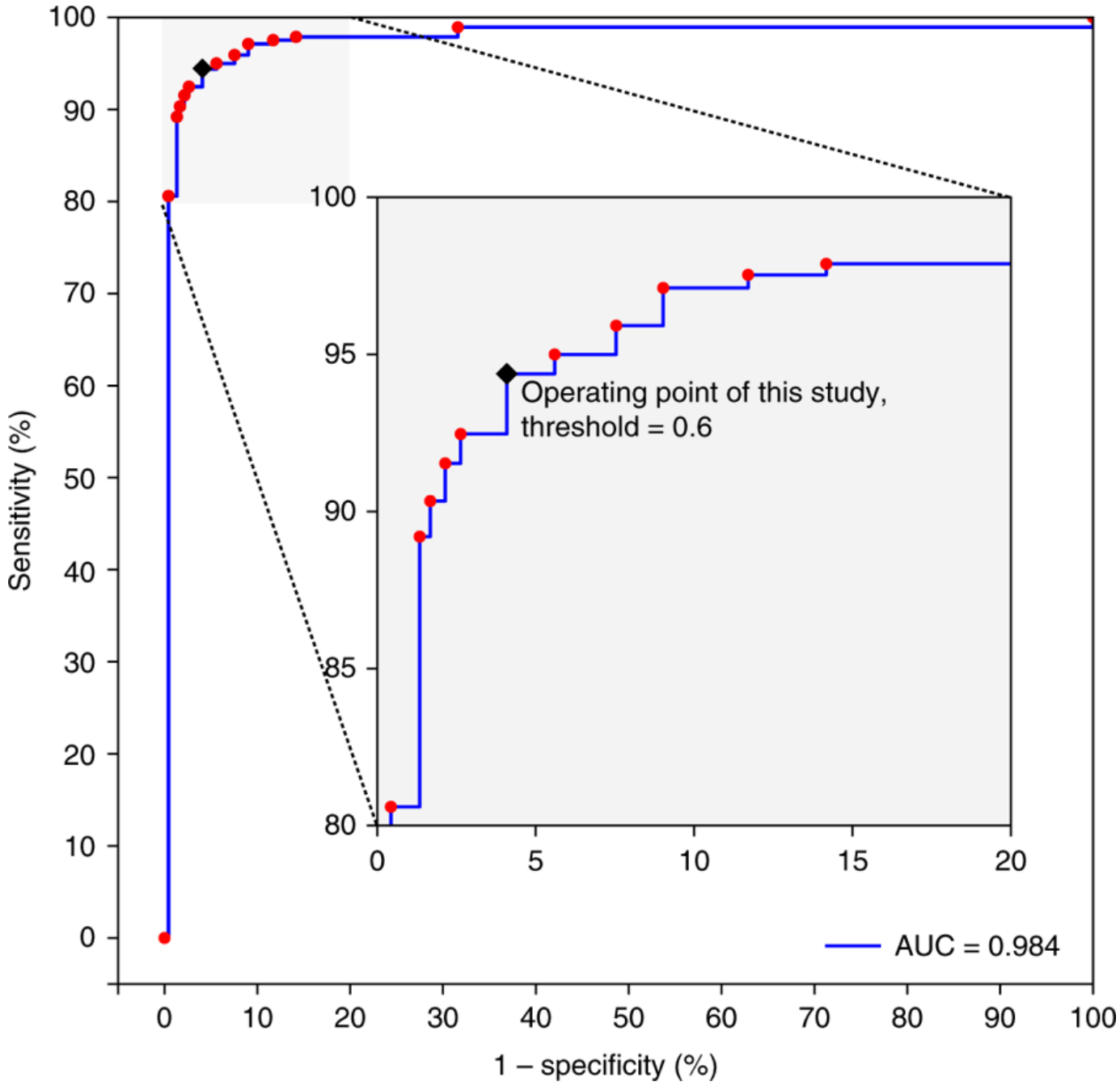
Used to assess **specificity**

Images

Video

Performance (dataset A)

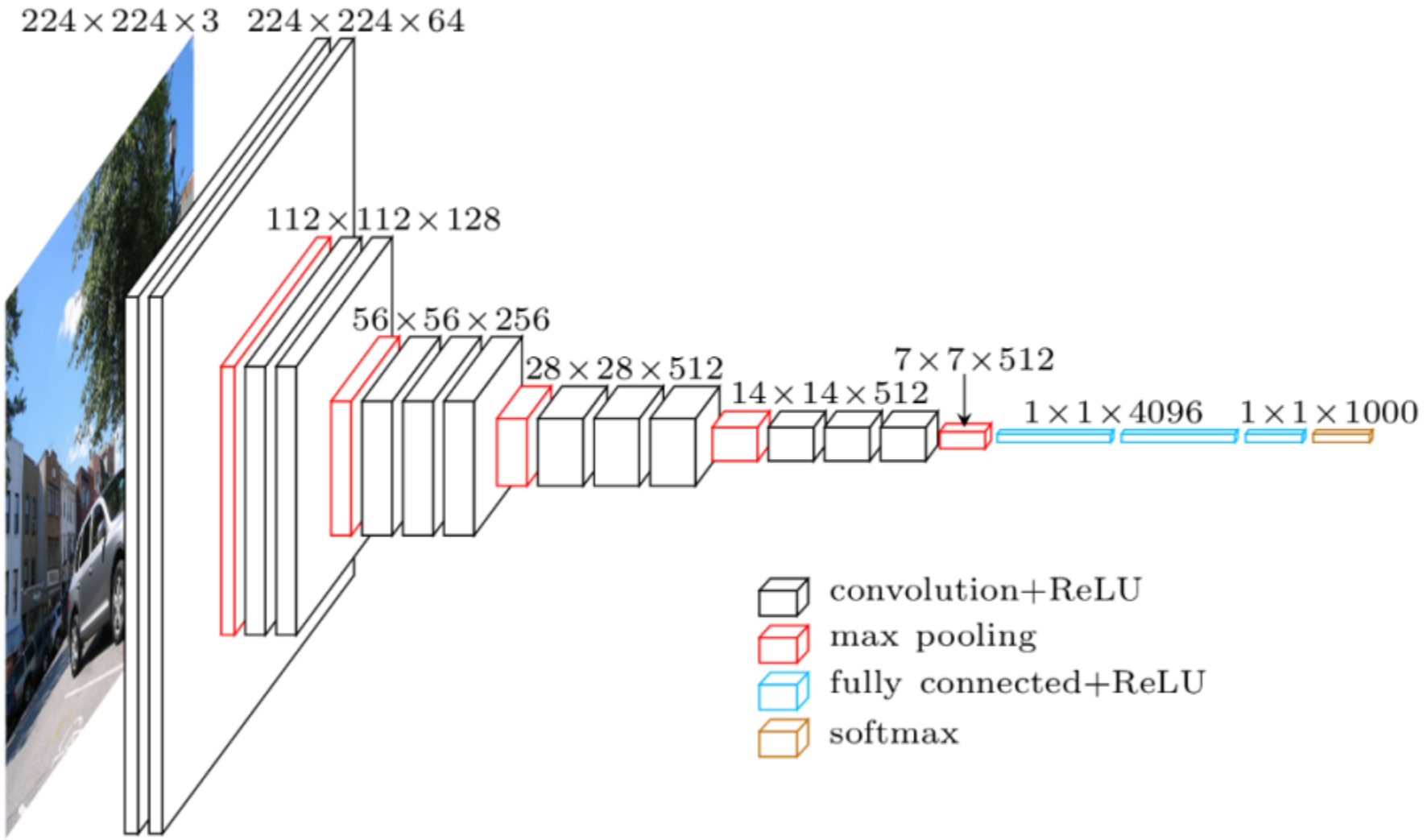
How can we get a meaningful comparison to expert performance?



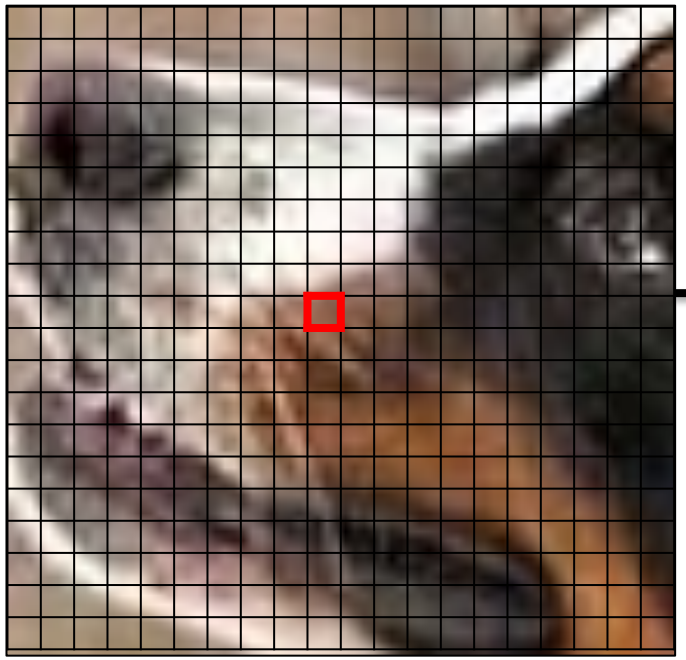
How does it work?

SEGMENTATION IN BRIEF

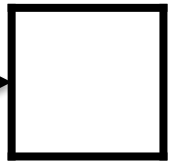
Classification Architecture (VGG)



A separate classifier for each pixel?



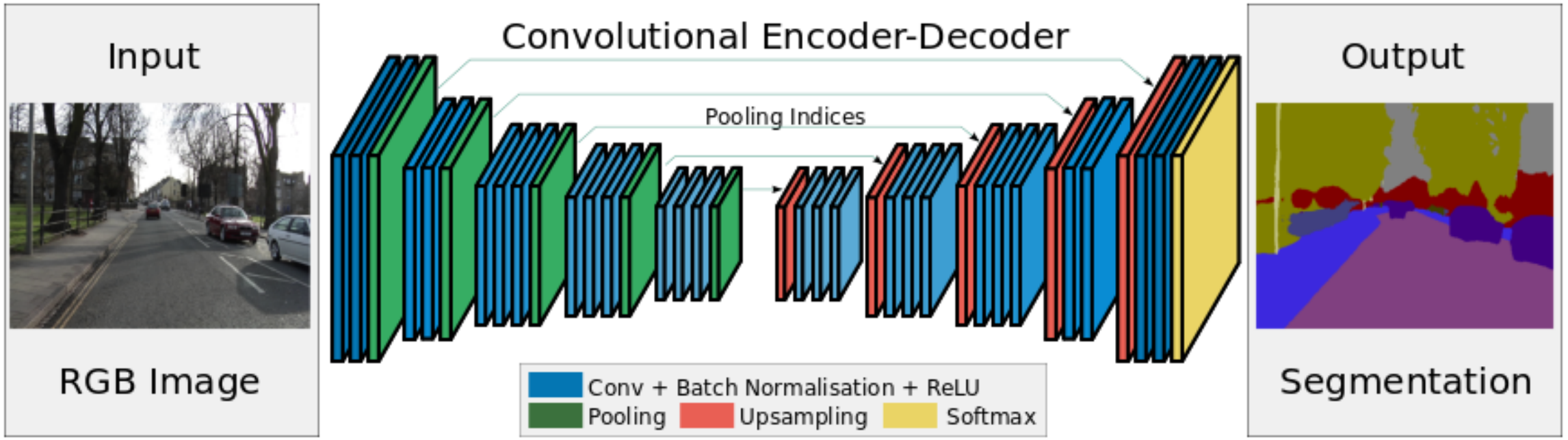
x, data/features



y, associated value or label

End goal: predict y from x

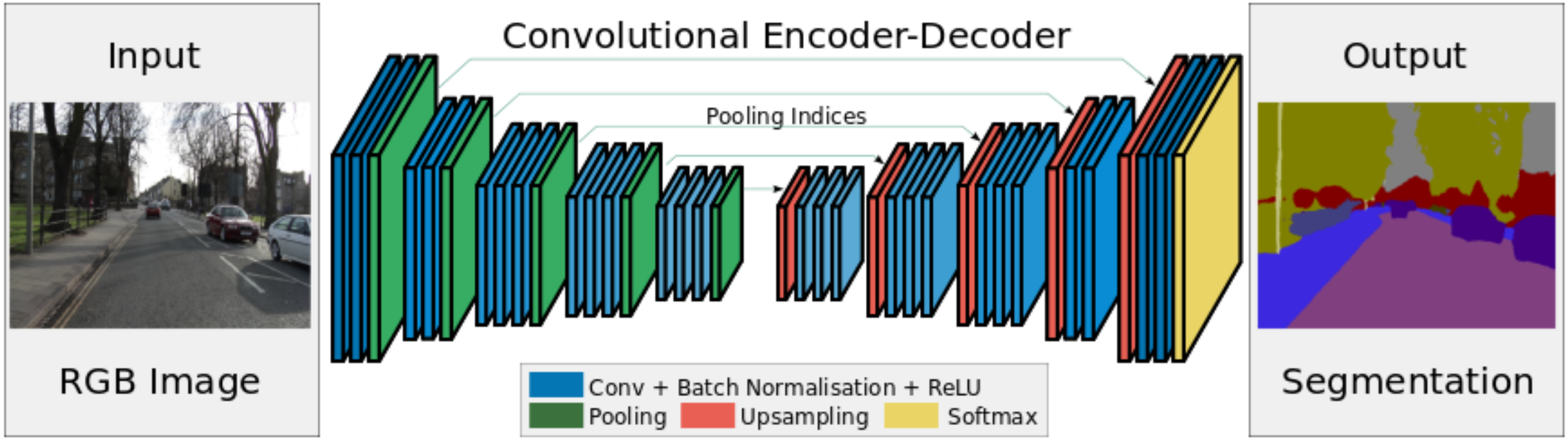
Segmentation Architecture



SegNet (2015): fully convolutional architecture

Stage 1: Encoder identifies features

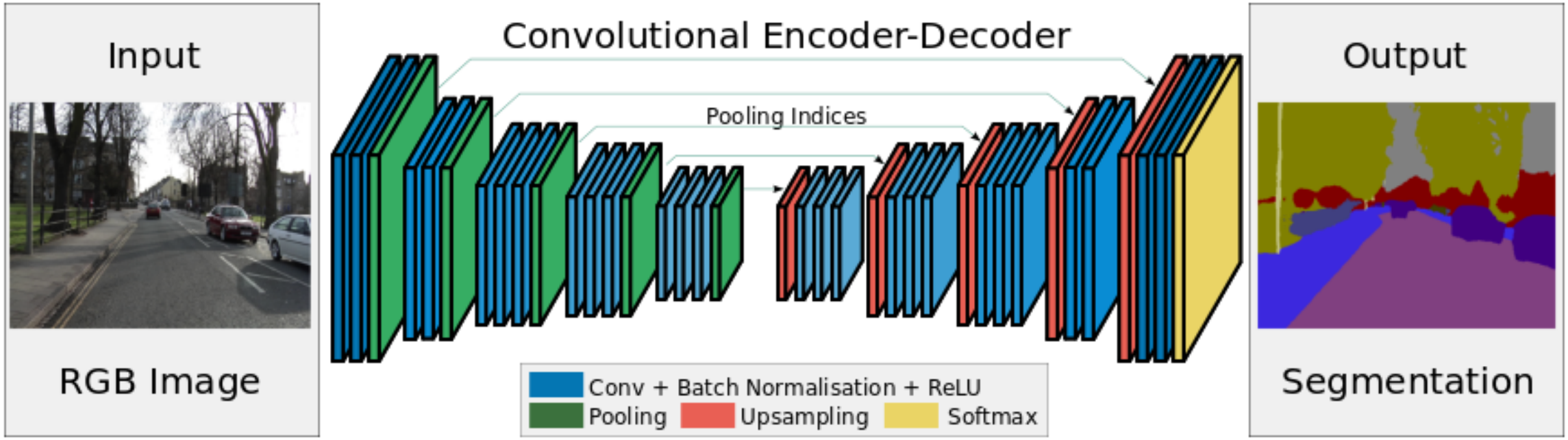
Stage 2: Decoder reconstructs the spatial map



Alternative architectures differ in their approach to the spatial reconstruction process

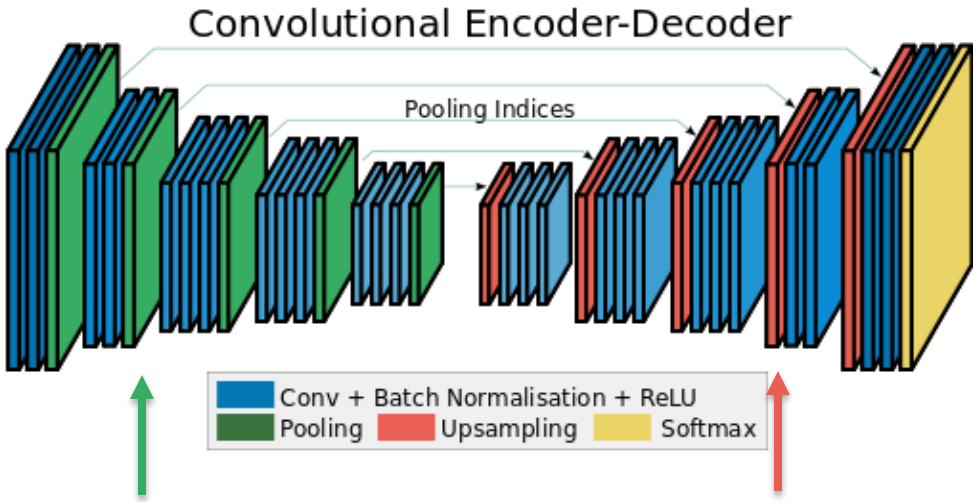
Stage 1: Encoder identifies features

Stage 2: Decoder reconstructs the spatial map



Segnet's approach to reconstruction: remember max-pooling indices

SegNet decoder overview



Remember max pooling indices

3	1	3	7
2	5	3	4
1	4	3	1
1	2	2	5

→

5	7
4	5

→ **Other Layers** →

3	4
8	2

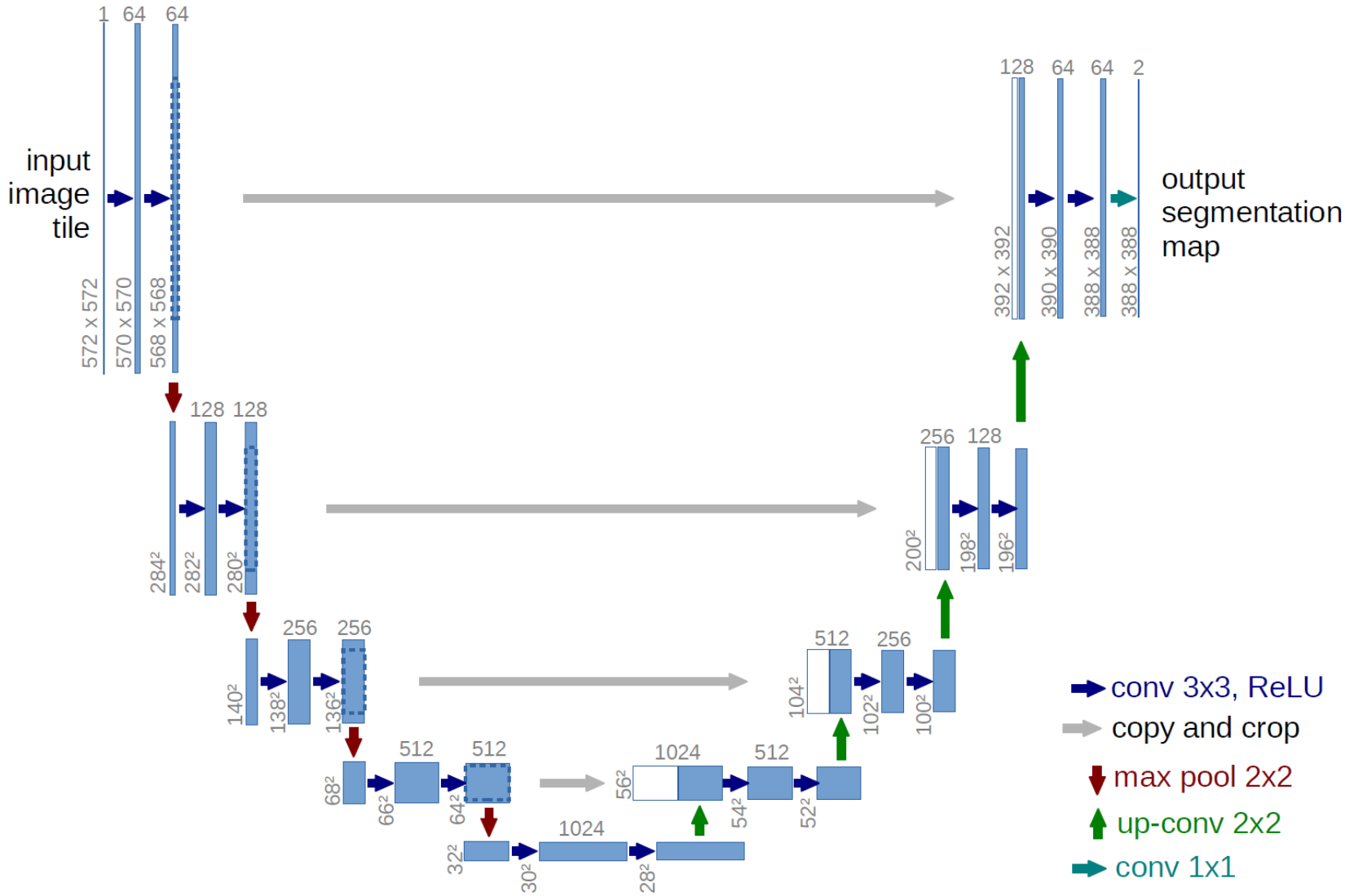
Use max pooling indices

→

0	0	0	4
0	3	0	0
0	8	0	0
0	0	0	2

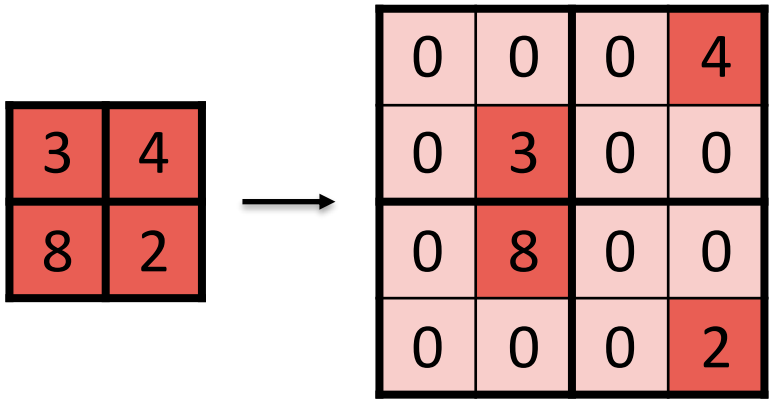
→ **Convolution with trainable decoder filters**

A Highly Successful Architecture: U-Net

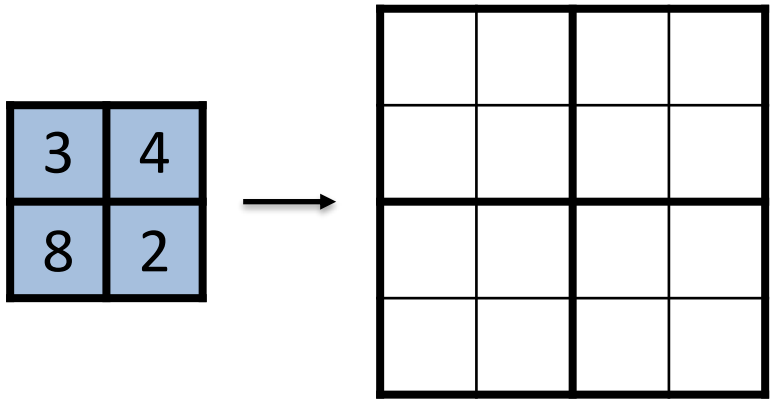


Concatenation of corresponding encoder feature map

SegNet: Upsampling w/ Stored Pooling Indices



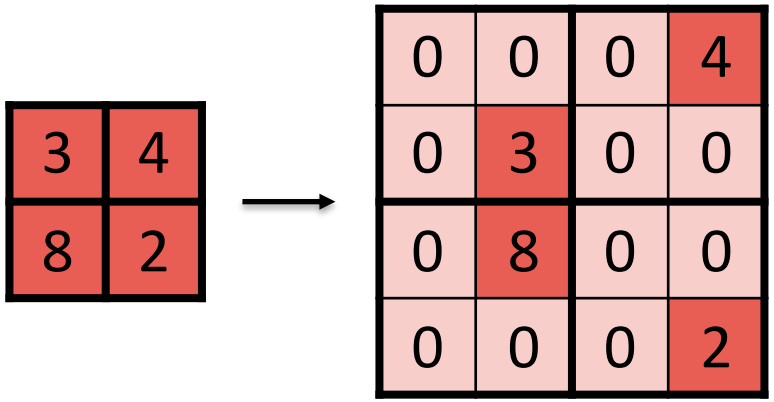
U-Net: 2x2 Upconvolution



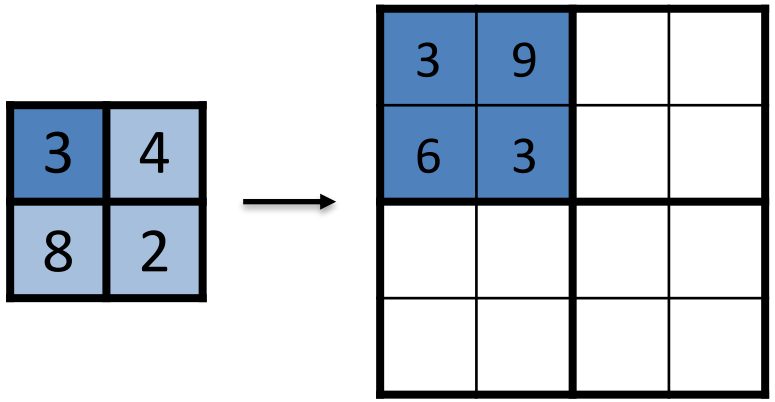
Learned filter

1	3
2	1

SegNet: Upsampling w/ Stored Pooling Indices



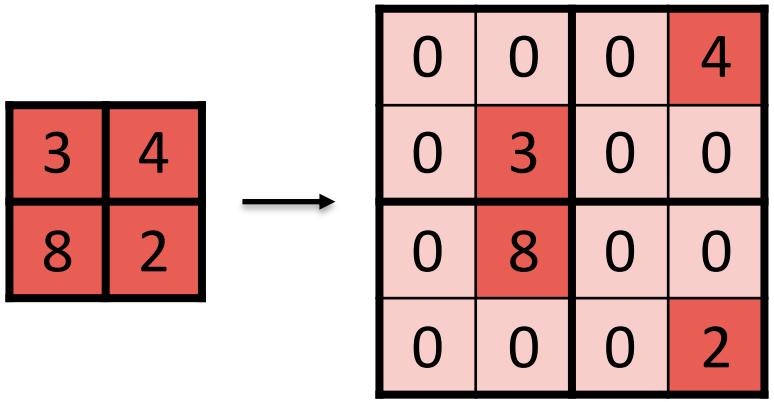
U-Net: 2x2 Upconvolution



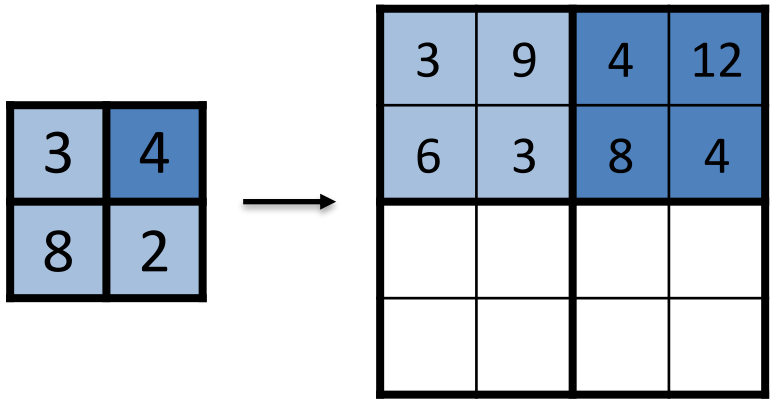
Learned filter

1	3
2	1

SegNet: Upsampling w/ Stored Pooling Indices



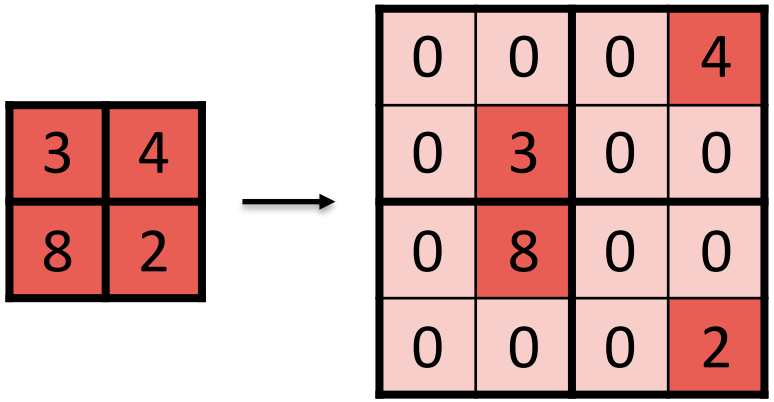
U-Net: 2x2 Upconvolution



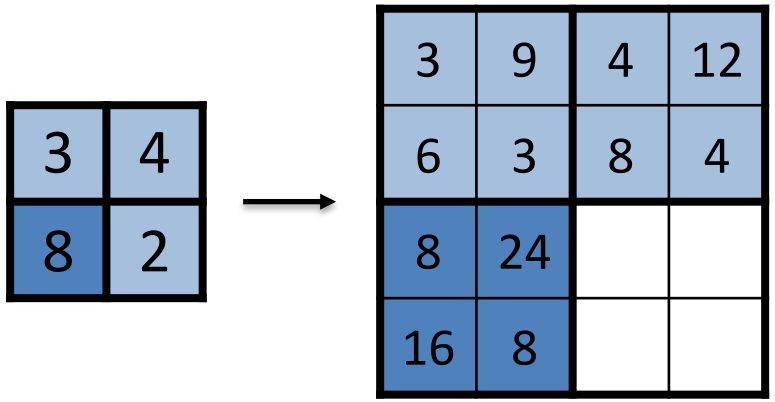
Learned filter

1	3
2	1

SegNet: Upsampling w/ Stored Pooling Indices



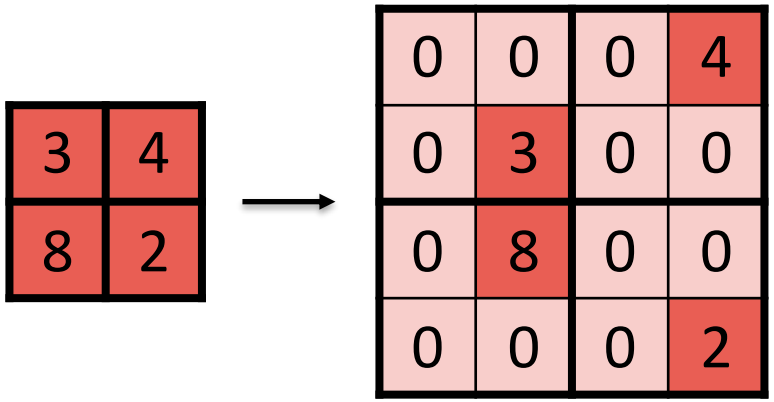
U-Net: 2x2 Upconvolution



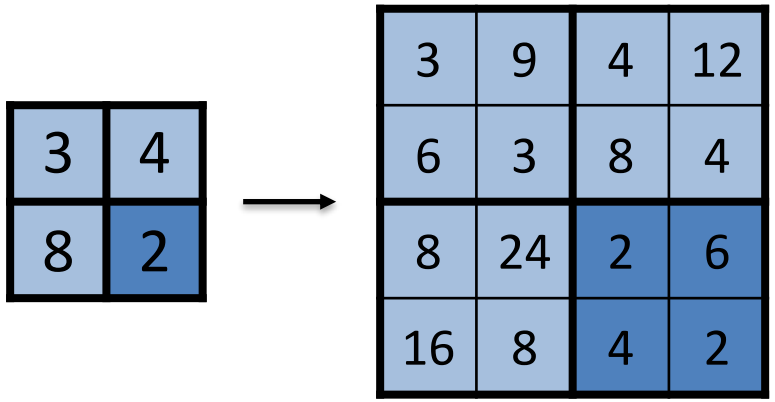
Learned filter

1	3
2	1

SegNet: Upsampling w/ Stored Pooling Indices



U-Net: 2x2 Upconvolution



Learned filter

1	3
2	1