Convolutional Neural Networks For Semantic Segmentation

Course 3, Module 5, Lesson 2

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

Learning Objectives

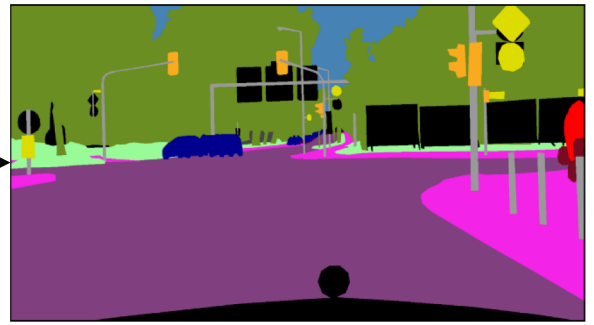
- Learn how to use convolutional neural networks to perform the semantic segmentation task
- Learn the different layers required for the good performance of semantic segmentation models

The Semantic Segmentation Problem

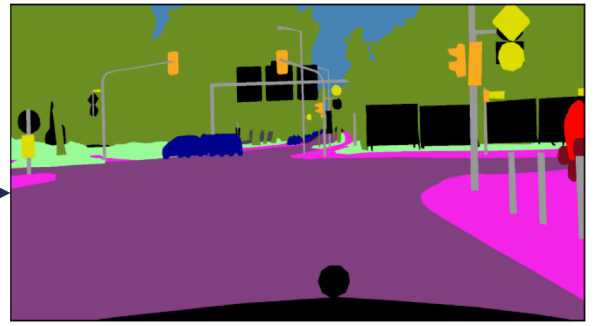
- Road
- Sidewalk
- Pole
- Traffic Light
- Traffic Signs
- Vegetation
- Terrain
- Sky
- Car
- Background

ConvNets For Semantic Segmentation

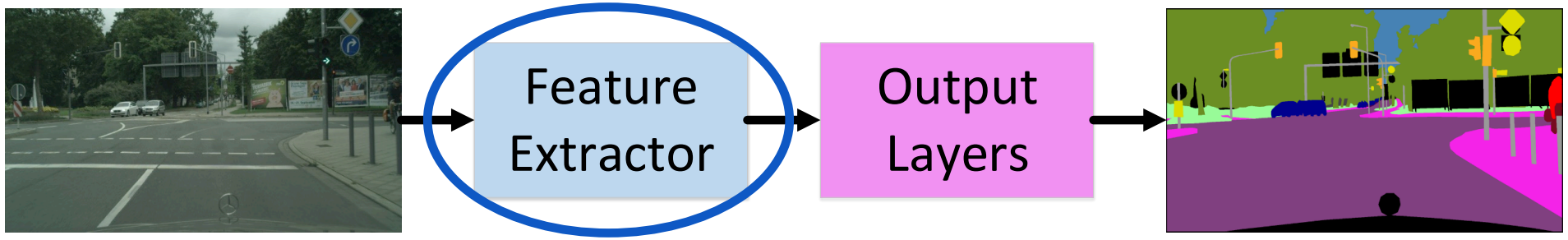
$$f(x;\theta)$$



ConvNet



ConvNets For Semantic Segmentation



The Feature Extractor

VGG

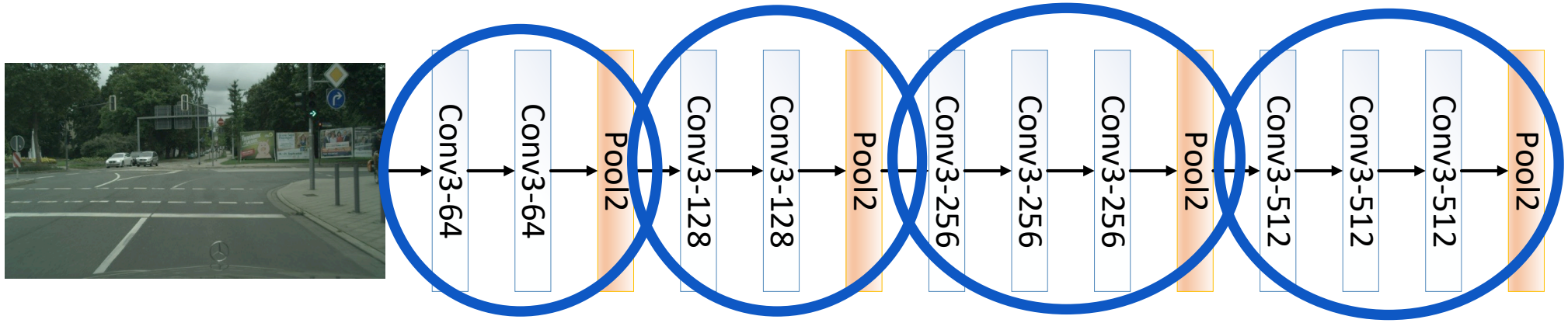
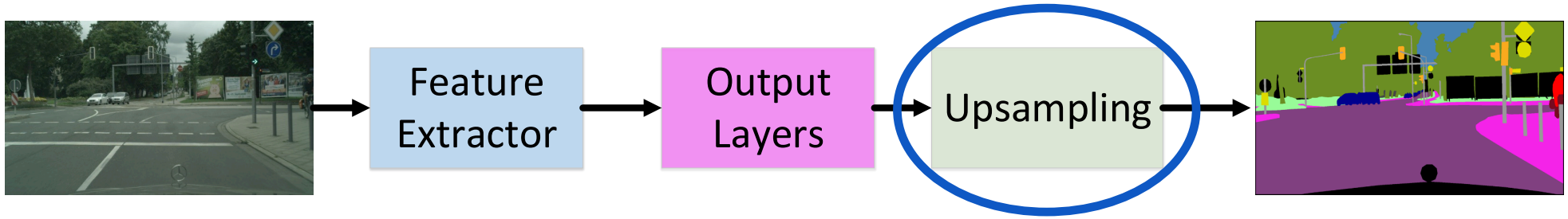


	Image	Conv1	Conv2	Conv3	Conv4
Width	M	M/2	M/4	M/8	M/16
Height	N	N/2	N/4	N/8	N/16
Depth	3	64	128	256	512

Shrink!

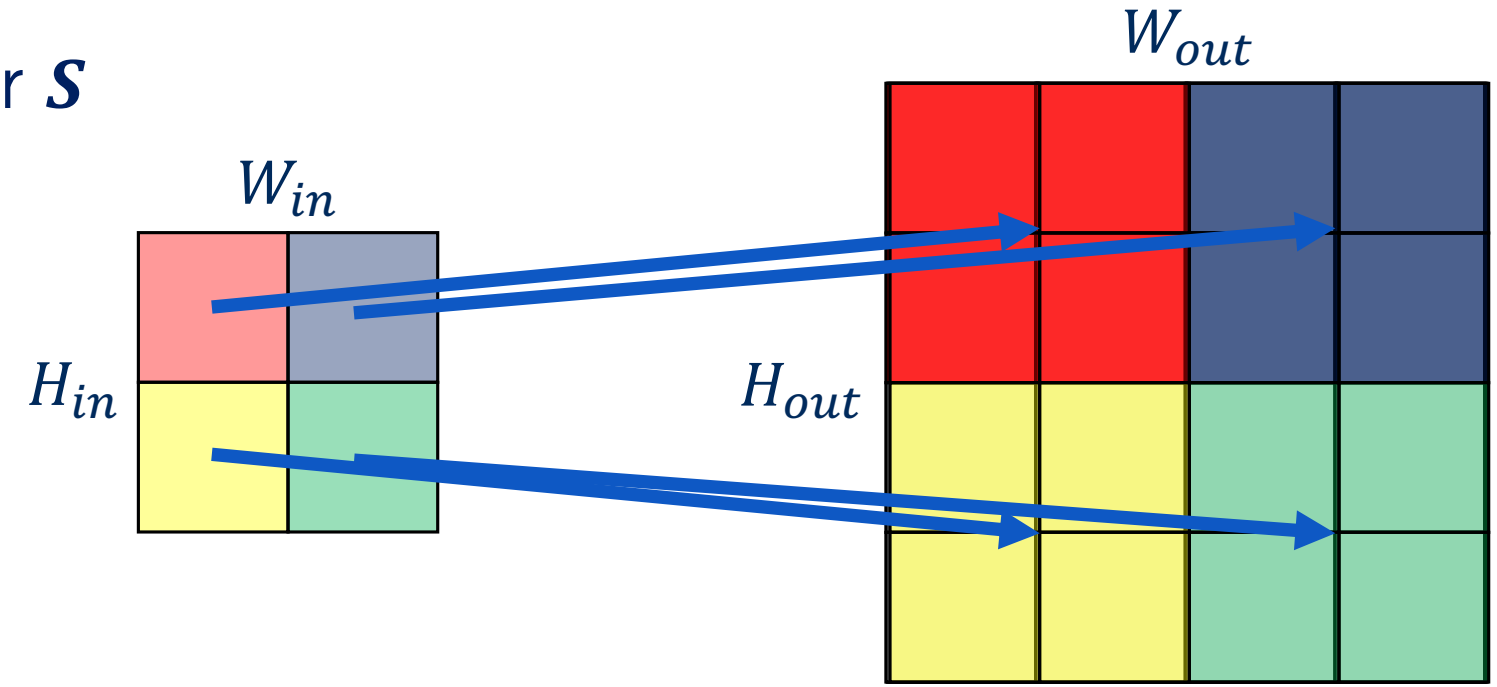
Upsampling the Output



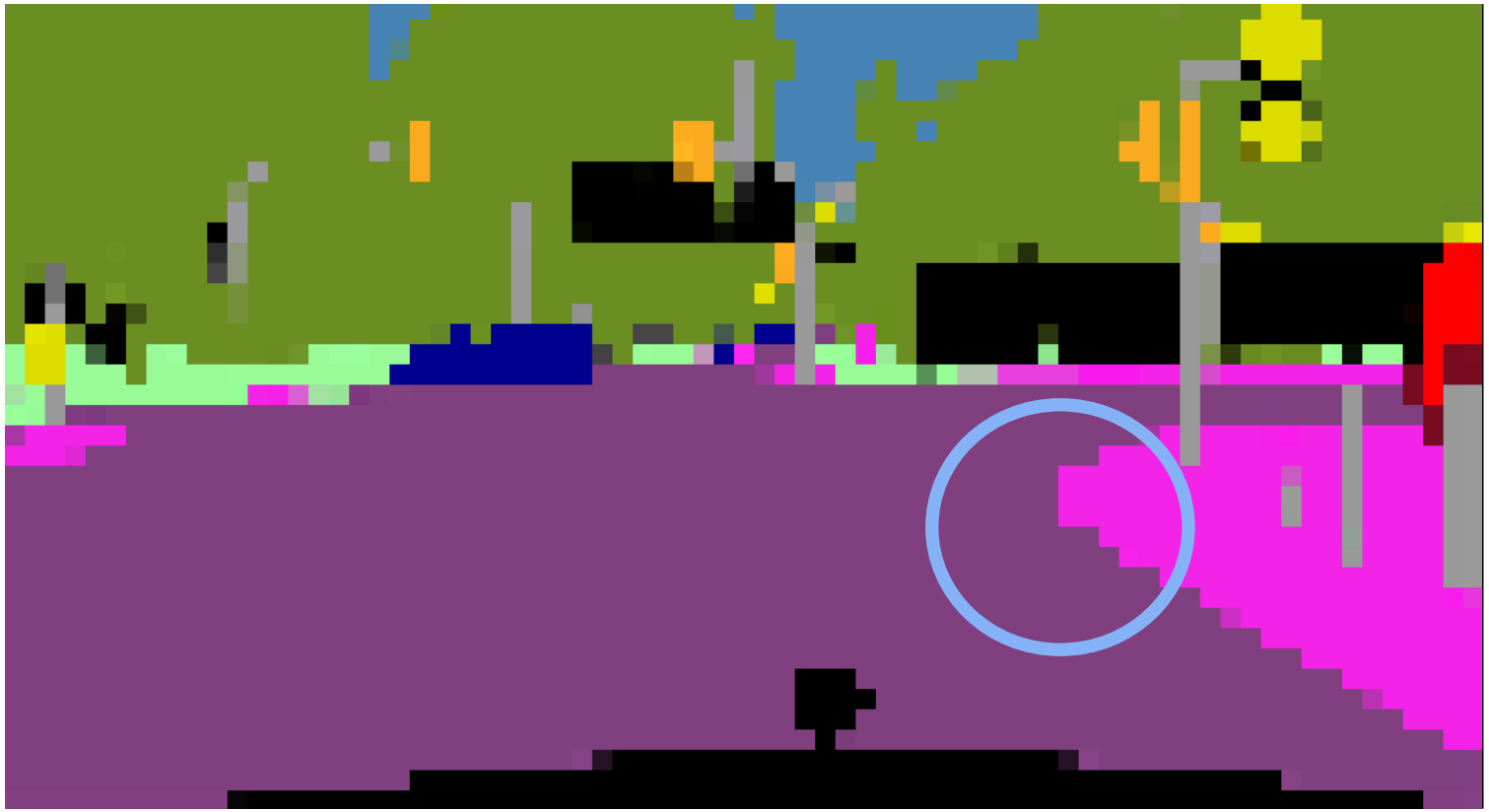
Upsampling Layer

- Upsampling Multiplier $\boldsymbol{S}$

$$\boldsymbol{W_{out}} = S \times \boldsymbol{W_{in}}$$
$$\boldsymbol{H_{out}} = S \times \boldsymbol{H_{in}}$$
$$\boldsymbol{D_{out}} = \boldsymbol{D_{in}}$$



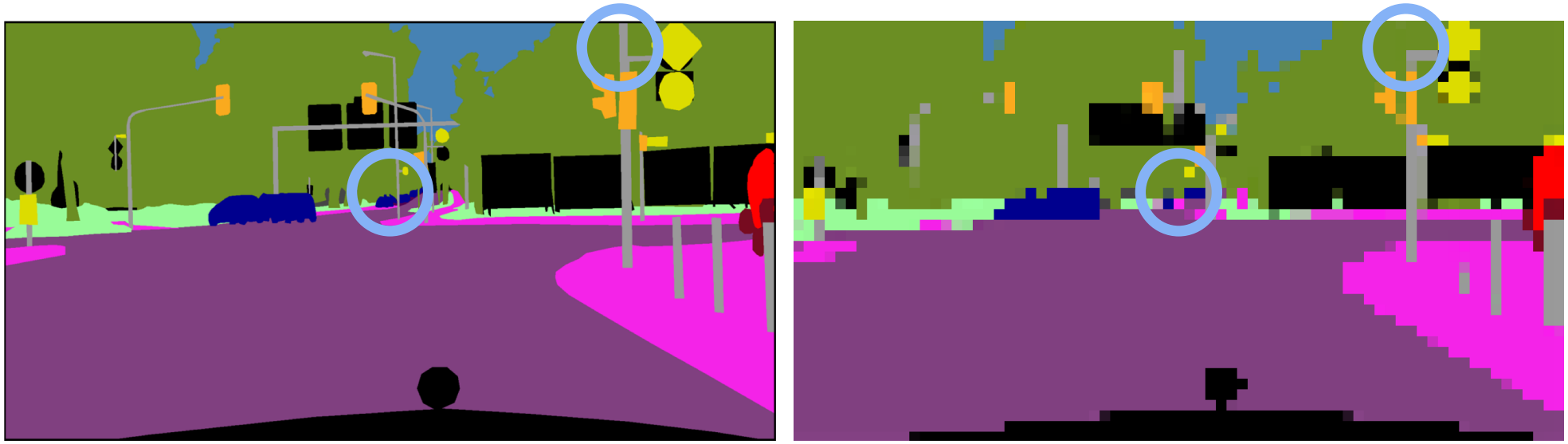
Upsampling The Output



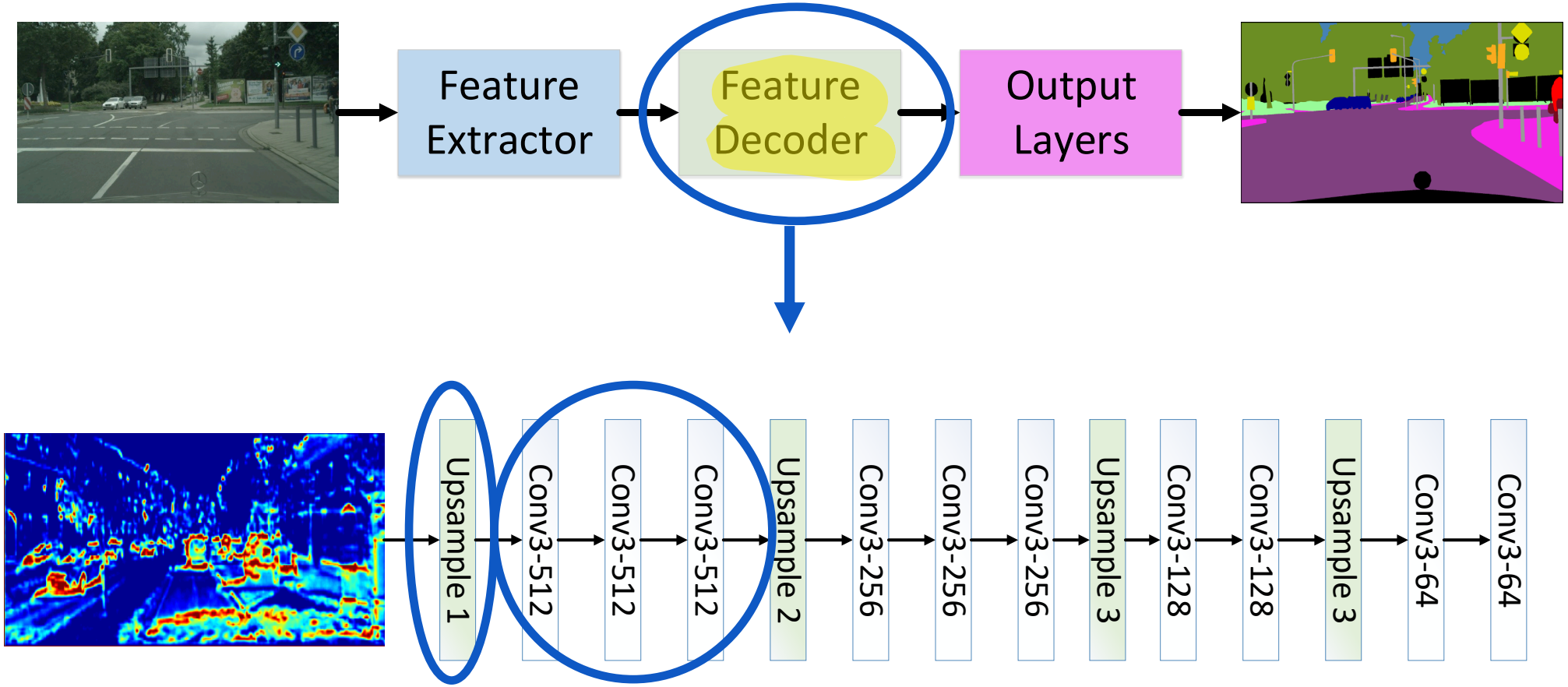
Coarse boundary.

Upsampling The Output

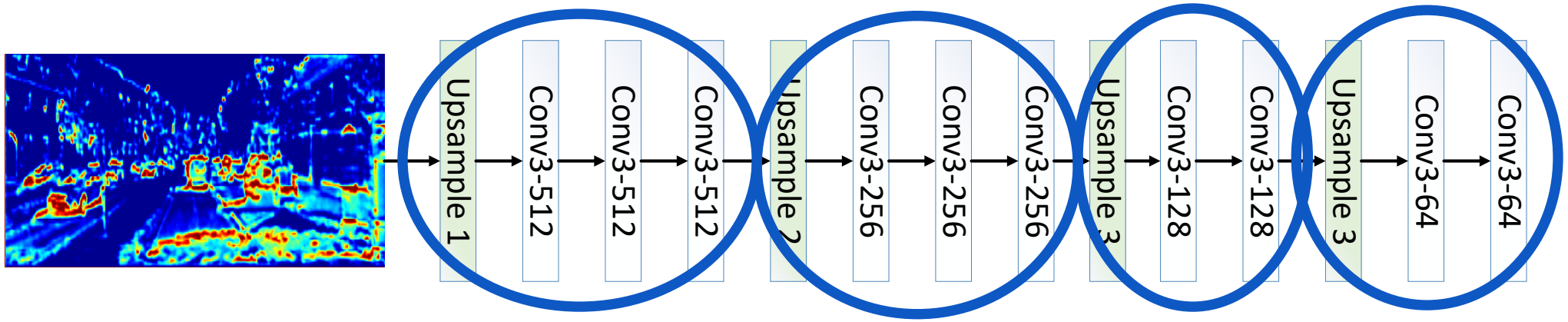
Object less than 16 pixels in width or height fully disappears.



Learning Same Resolution Feature Maps

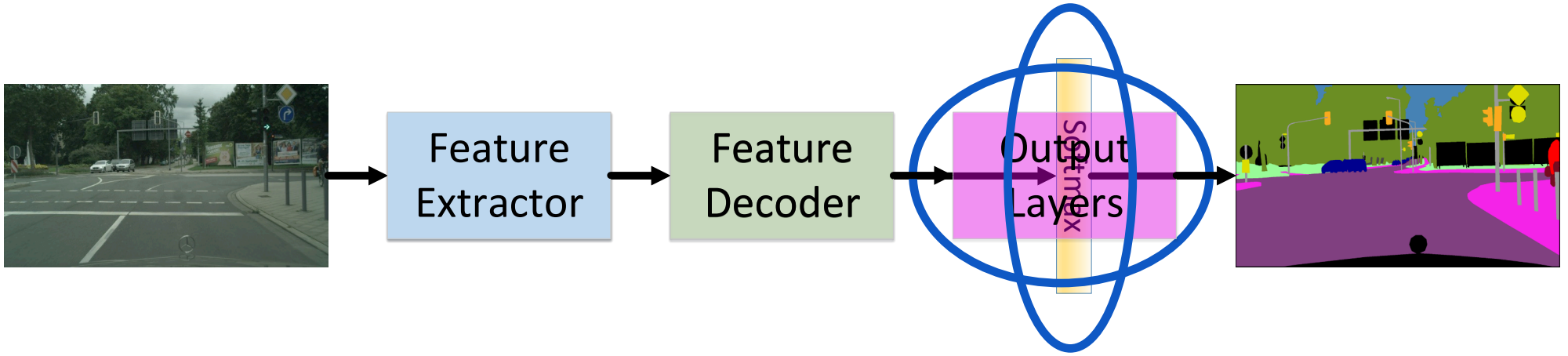


The Feature Decoder

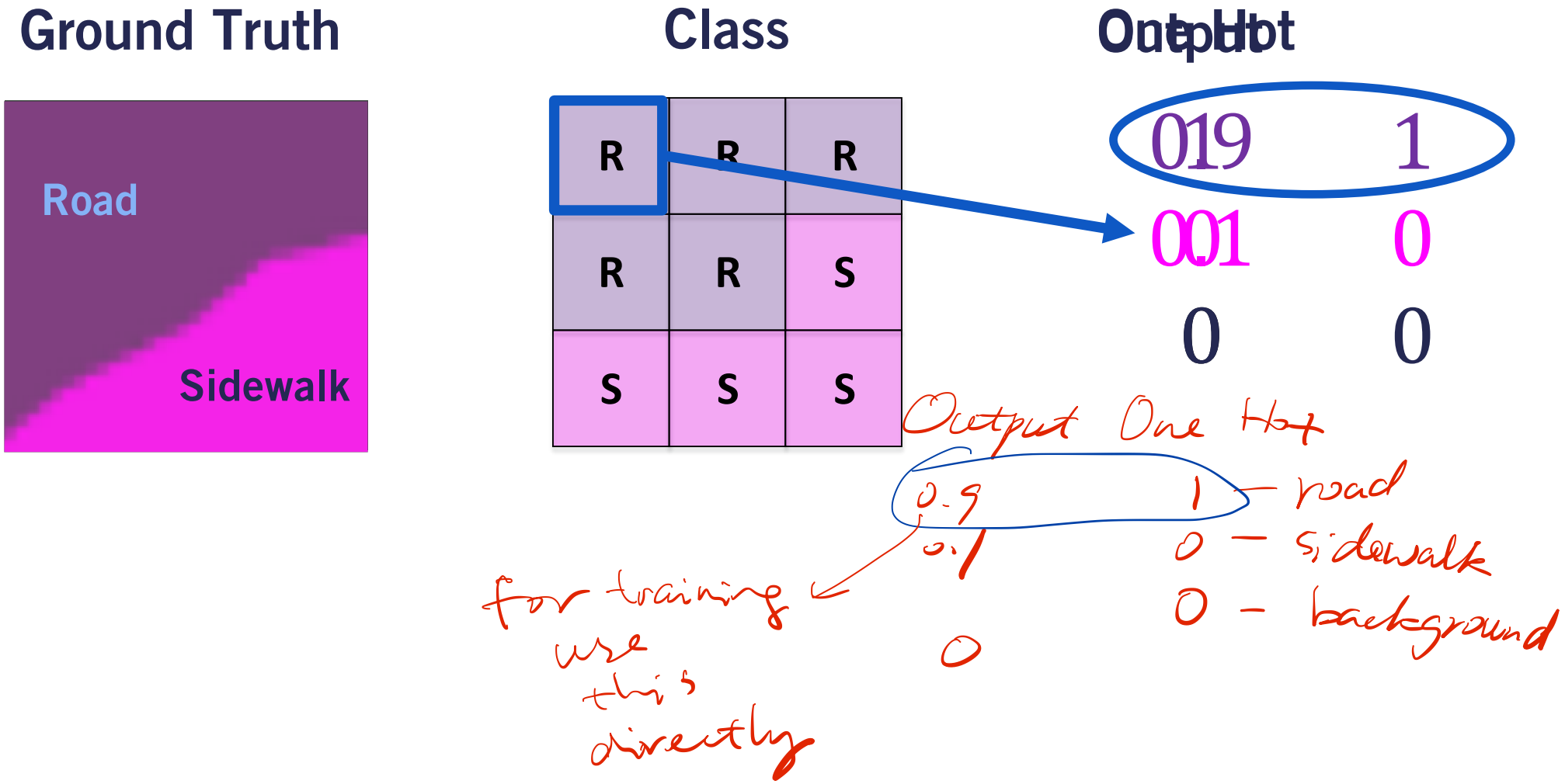


	Feature Map	Deconv1	Deconv2	Deconv3	Deconv4
Width	M/16	M/8	M/4	M/2	M
Height	N/16	N/8	N/4	N/2	N
Depth	512	512	256	128	64

Learning Same Resolution Feature Maps



Output Representation

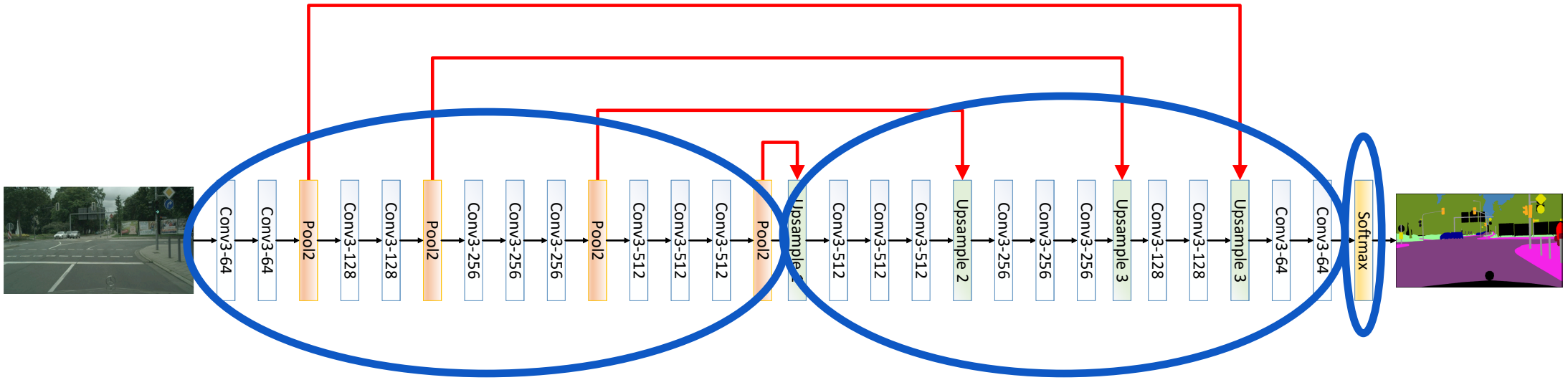


Classification Loss

$$L_{cls} = \frac{1}{N_{total}} \sum_i CrossEntropy(s_i^*, s_i)$$

- N_{total} is the number of pixels in all images of our **minibatch** 8 – 16 images
- s_i is the output of the neural network
- s_i^* is the ground truth classification

ConvNets For Semantic Segmentation



Summary

- Convolutional Neural Networks can be used to solve the semantic segmentation problem
- In a feature extractor and a feature decoder are required to provide the final output of semantic segmentation models

- **Next: Semantic Segmentation For Autonomous Driving**