The Semantic Segmentation Problem

Course 3, Module 5, Lesson 1

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

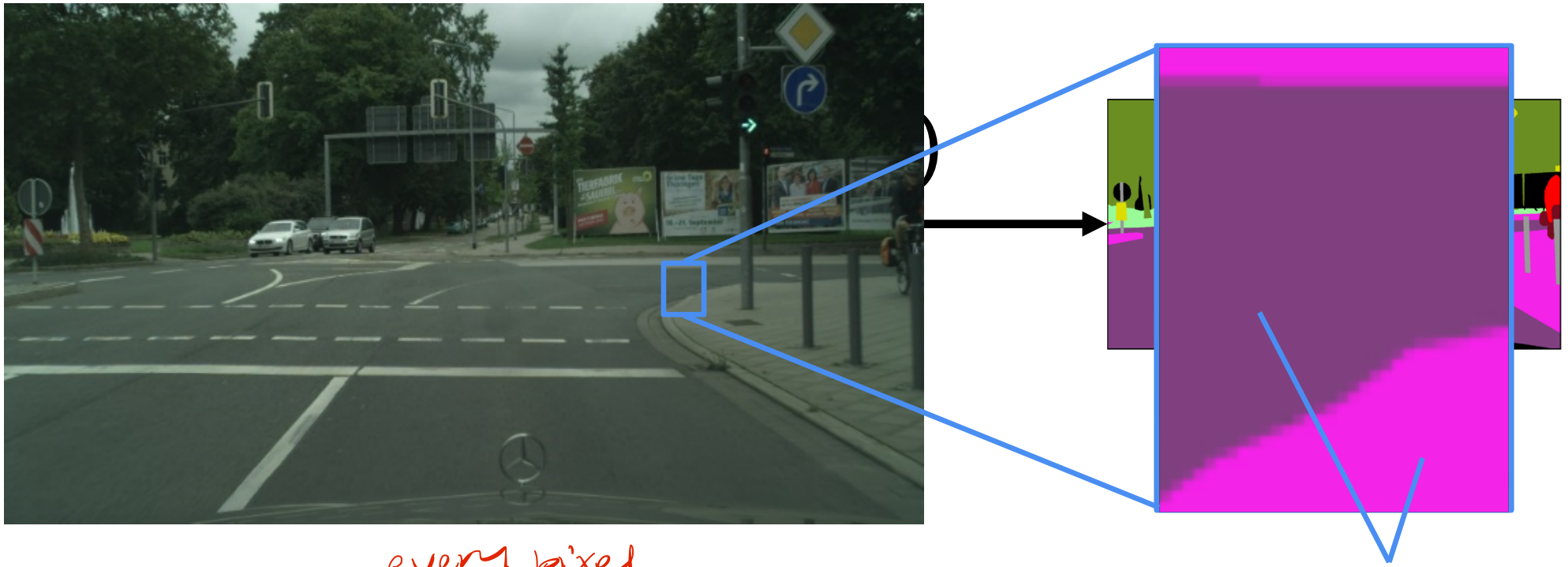
Learning Objectives

- Learn the semantic segmentation task problem formulation
- Learn to determine how well a semantic segmentation model is performing with task relevant performance measures

The Semantic Segmentation Problem

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

Mathematical Problem Formulation



every pixel

$$f(x;\theta) = [S_{class_1}, \cdots, S_{class_k}]$$

$S_{road} = 0.92$

$S_{sidewalk} = 0.02$

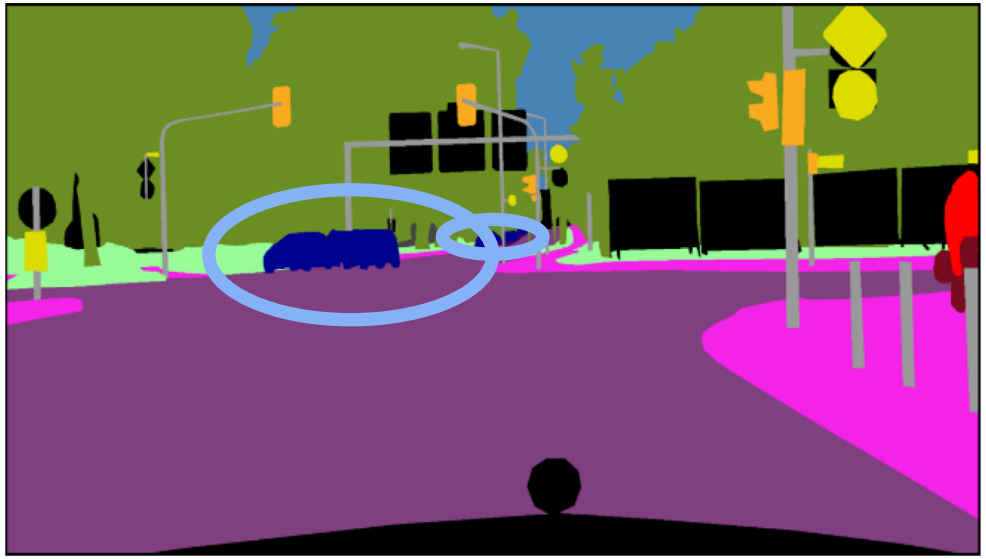
.

.

$S_{Background} = 0.02$

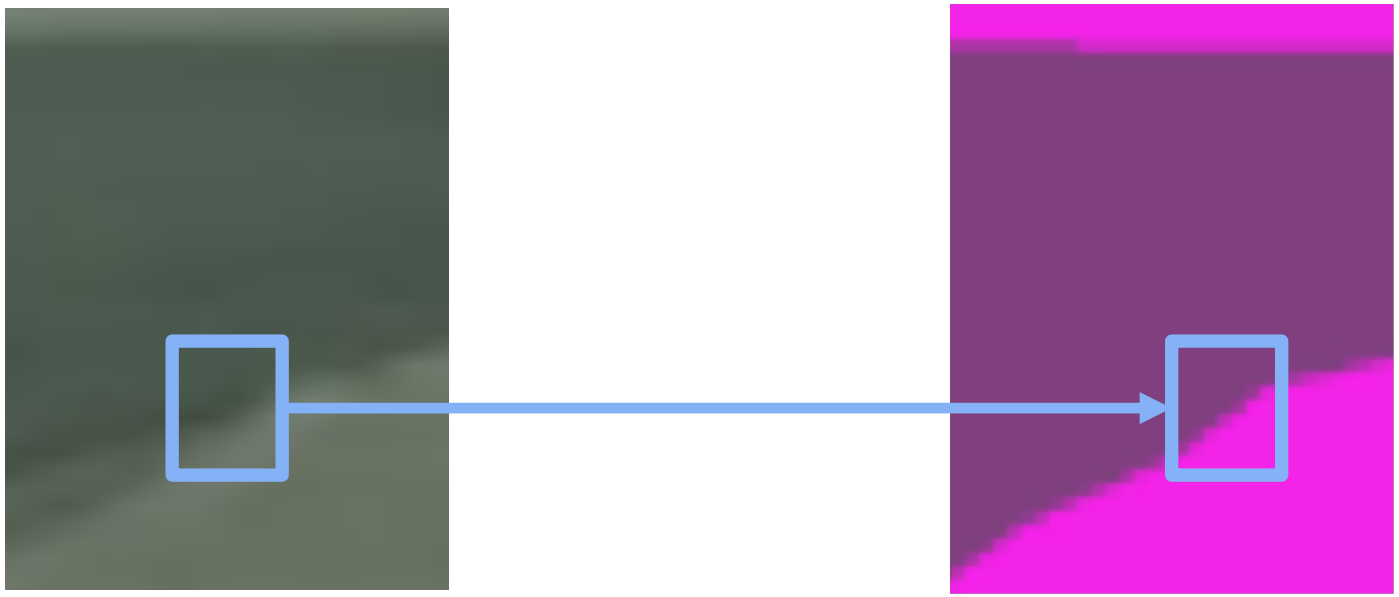
Semantic Segmentation is Not Trivial !

- Occlusion, truncation, scale, and illumination changes



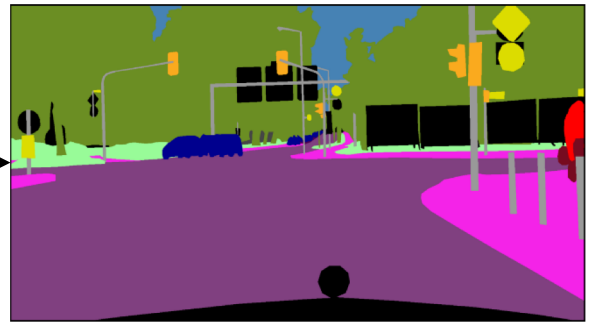
Semantic Segmentation is Not Trivial !

- Occlusion, truncation, scale, and illumination changes
- Smooth boundaries

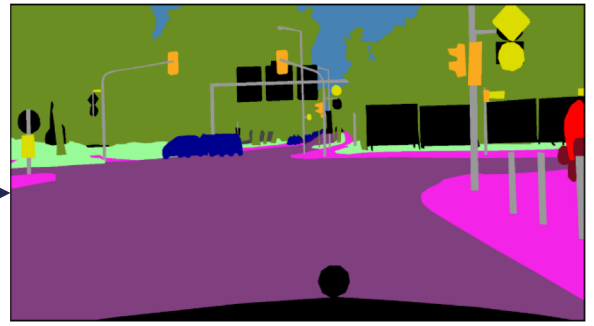


ConvNets For Semantic Segmentation

$$f(x;\theta)$$



ConvNet



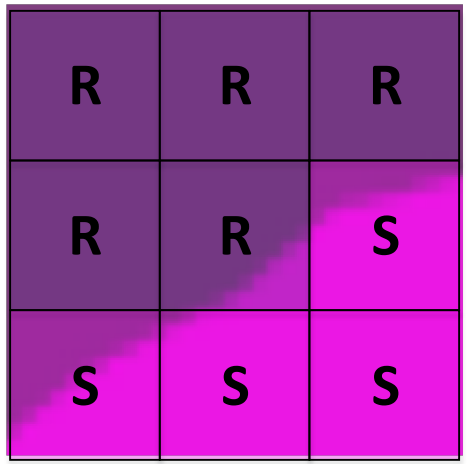
Evaluation Metrics

- **True Positive (TP):** The number of correctly classified pixels belonging to class X
- **False Positive (FP):** The number of pixels that **do not belong** to class X in ground truth but **are classified** that class by the algorithm
- **False Negative (FN):** The number of pixels that **do belong** to class X in ground truth, but **are not classified** as that class by the algorithm

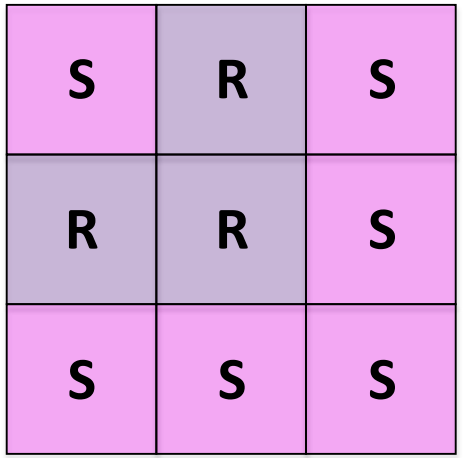
$$IOU_{class} = \frac{TP}{TP + FP + FN}$$

Evaluation Metrics

Ground Truth

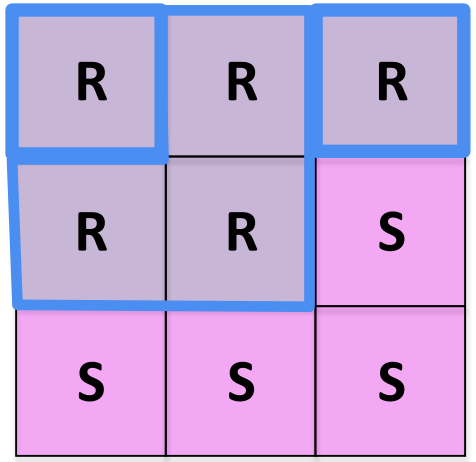


Prediction

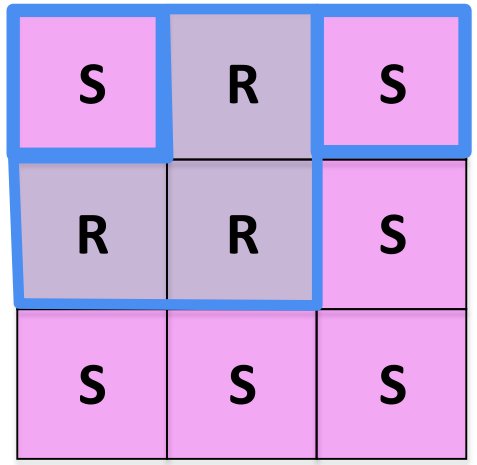


Evaluation Metrics

Ground Truth



Prediction



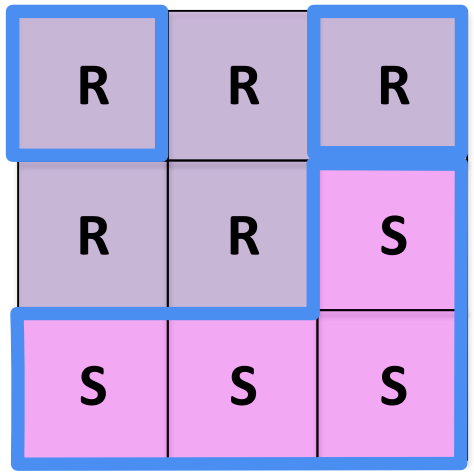
Class: Road

$$TP = 3$$
$$FP = 0$$
$$FN = 2$$

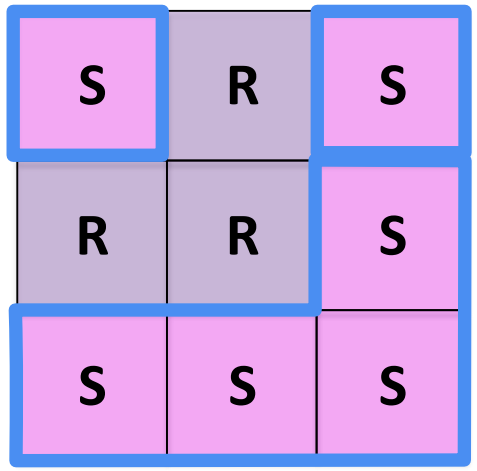
$$IOU_{Road} = \frac{3}{3+0+2} = \frac{3}{5}$$

Evaluation Metrics

Ground Truth



Prediction



Class: Sidewalk

$$TP = 4$$
$$FP = 2$$
$$FN = 0$$

$$IOU_{Road} = \frac{4}{4+2+0} = \frac{4}{6}$$

Evaluation Metrics

- **Class IOU** over all the data is calculated by computing the sum of TP, FP, FN for all images first
- Averaging the class IOU is usually not a very good idea!
- CityScapes Segmentation Dataset

benchmark

bias towards object instances that cover a large image area

Summary

- Semantic segmentation consists of providing a class label for every pixel in a 2D image
- Semantic segmentation models can be evaluated using class IOU

- **Next: ConvNets for Semantic Segmentation**