Problem Formulation

Course 3, Module 4, Lesson 1

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

Learning Objectives

- In this module…
  - Define the 2D object detection problem
  - Apply ConvNets to 2D object detection
  - Challenges of object detection
  - 2D object tracking problem

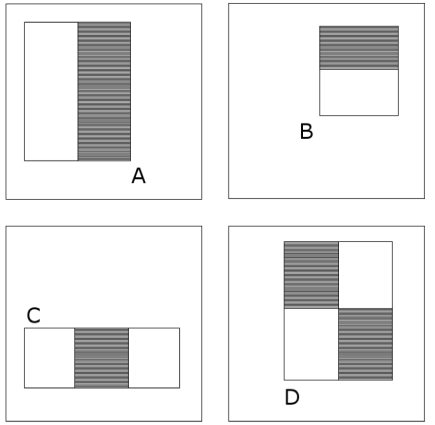
Learning Objectives

- Learn the 2D object detection task problem formulation.

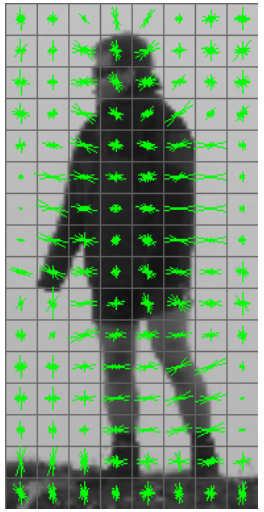
- Learn to determine how good a 2D object detector is through evaluating performance measures.

Brief History of Object Detection

- 2001 – Viola, Jones – Viola Jones Object Detection Framework



- 2005 – Dalal, Triggs – Histogram of Oriented Gradients

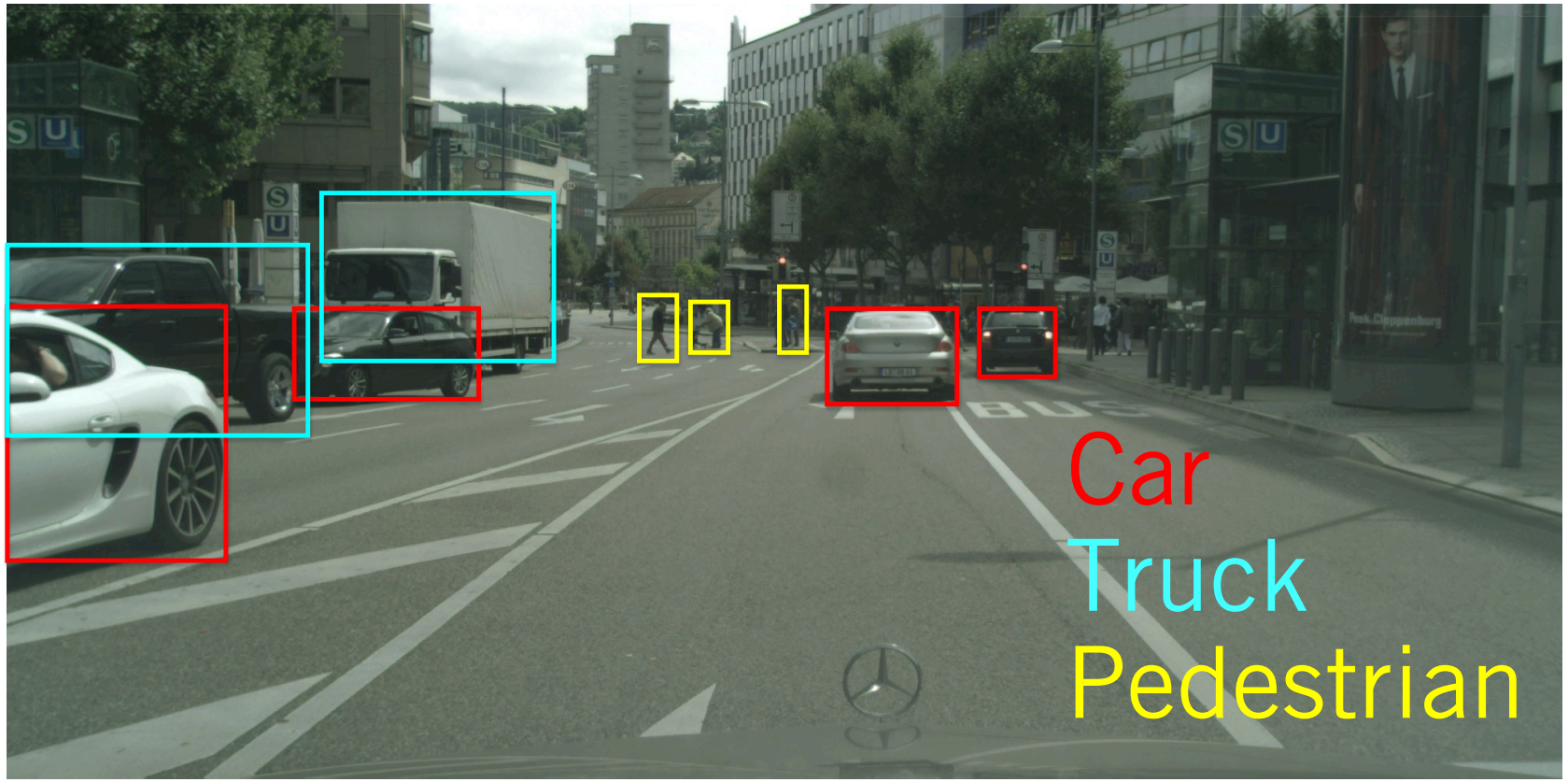


- 2012 – Krizhevsky, Sutskever, Hinton - Alexnet

Imagene

The Object Detection Problem

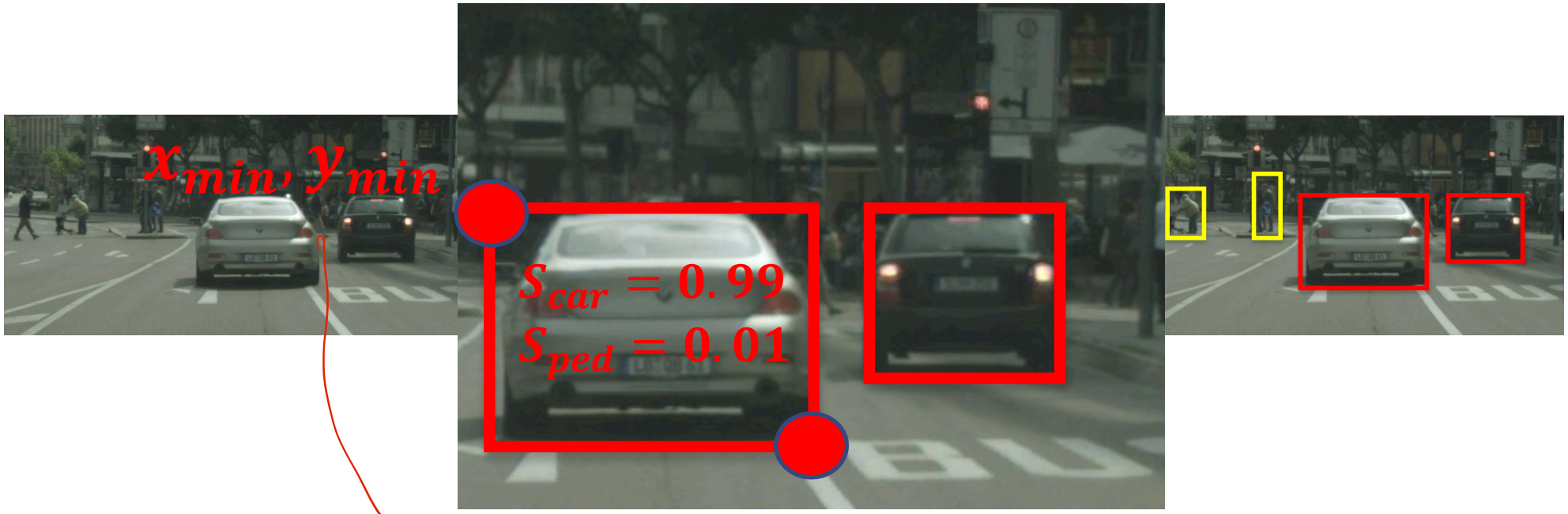
location + class



Object Detection Is Not Trivial !

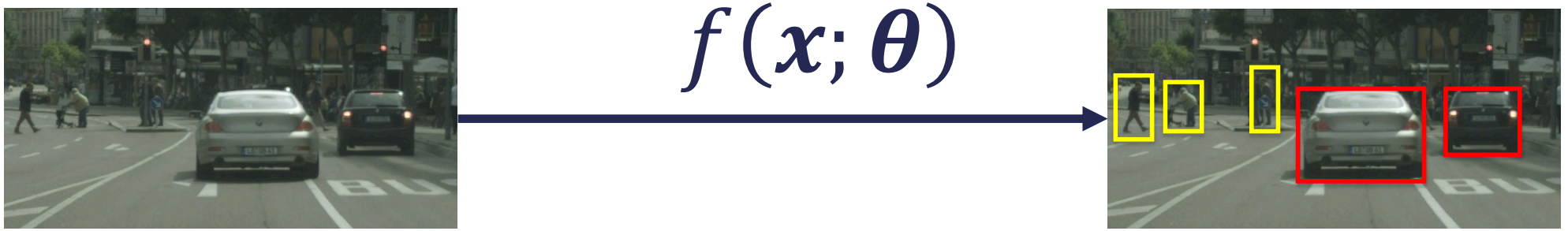
- **Extent of objects is not fully observed!**
  - **Occlusion:** Background objects covered by foreground objects
  - **Truncation:** Objects are out of image boundaries
- **Scale:** Object size gets smaller as the object moves farther away
- Illumination Changes:
  - **Too bright**
  - **Too dark**

Mathematical Problem Formulation



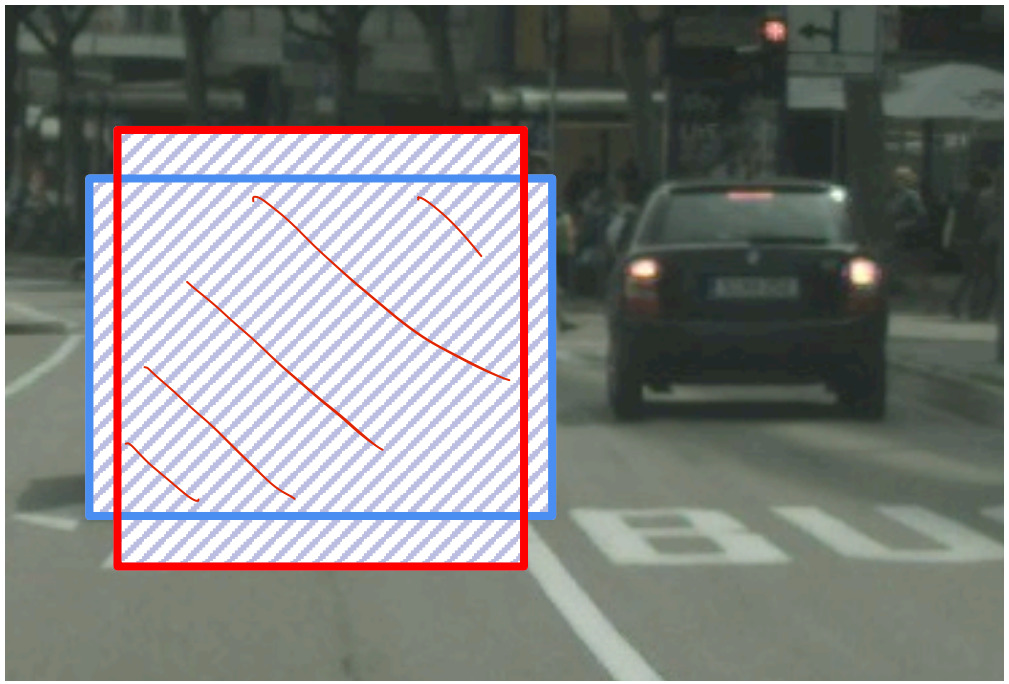
$$f(x;\theta) = [x_{min}, y_{min}, x_{max}, y_{max}, S_{class_1} \dots, S_{class_k}]$$

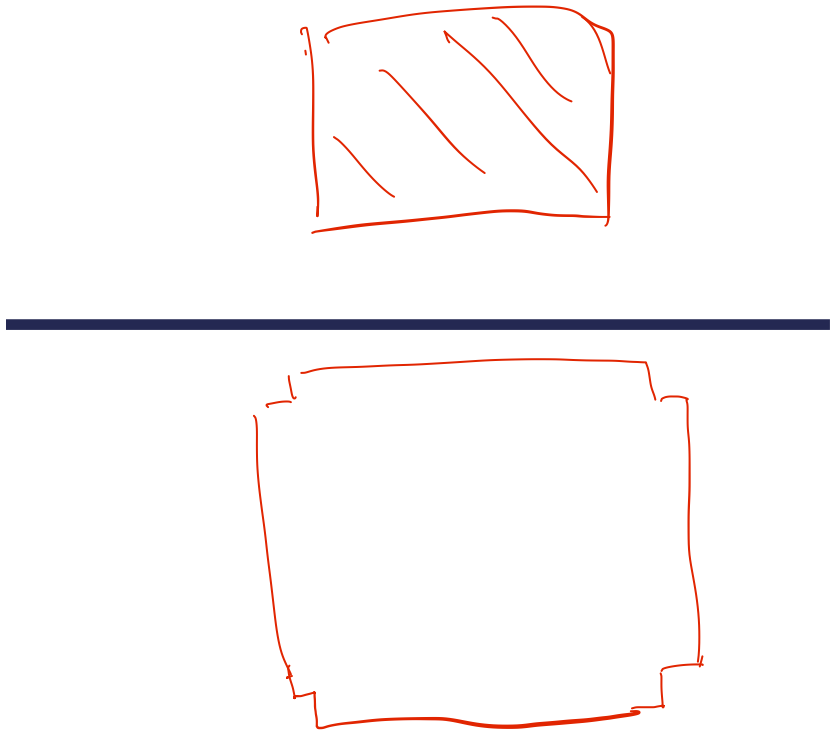
ConvNets For Object Detection



Evaluation Metrics

- **Intersection-Over-Union (IOU):** area of intersection of **predicted box** with a **ground truth box**, divided by the area of their union





Evaluation Metrics

- **True Positive (TP):** Object class score > score threshold, and IOU > IOU threshold
- **False Positive (FP):** Object class score > score threshold, and IOU < IOU threshold
- **False Negative (FN):** Number of ground truth objects not detected by the algorithm
- **Precision:** TP/ (TP + FP)
- **Recall:** TP / (TP + FN)
- **Precision Recall Curve (PR-Curve):** Use multiple object class score thresholds to compute precision and recall. Plot the values with precision on y-axis, and recall on x-axis
- **Average Precision (AP):** Area under PR-Curve for a single class. Usually approximated using **11** recall points

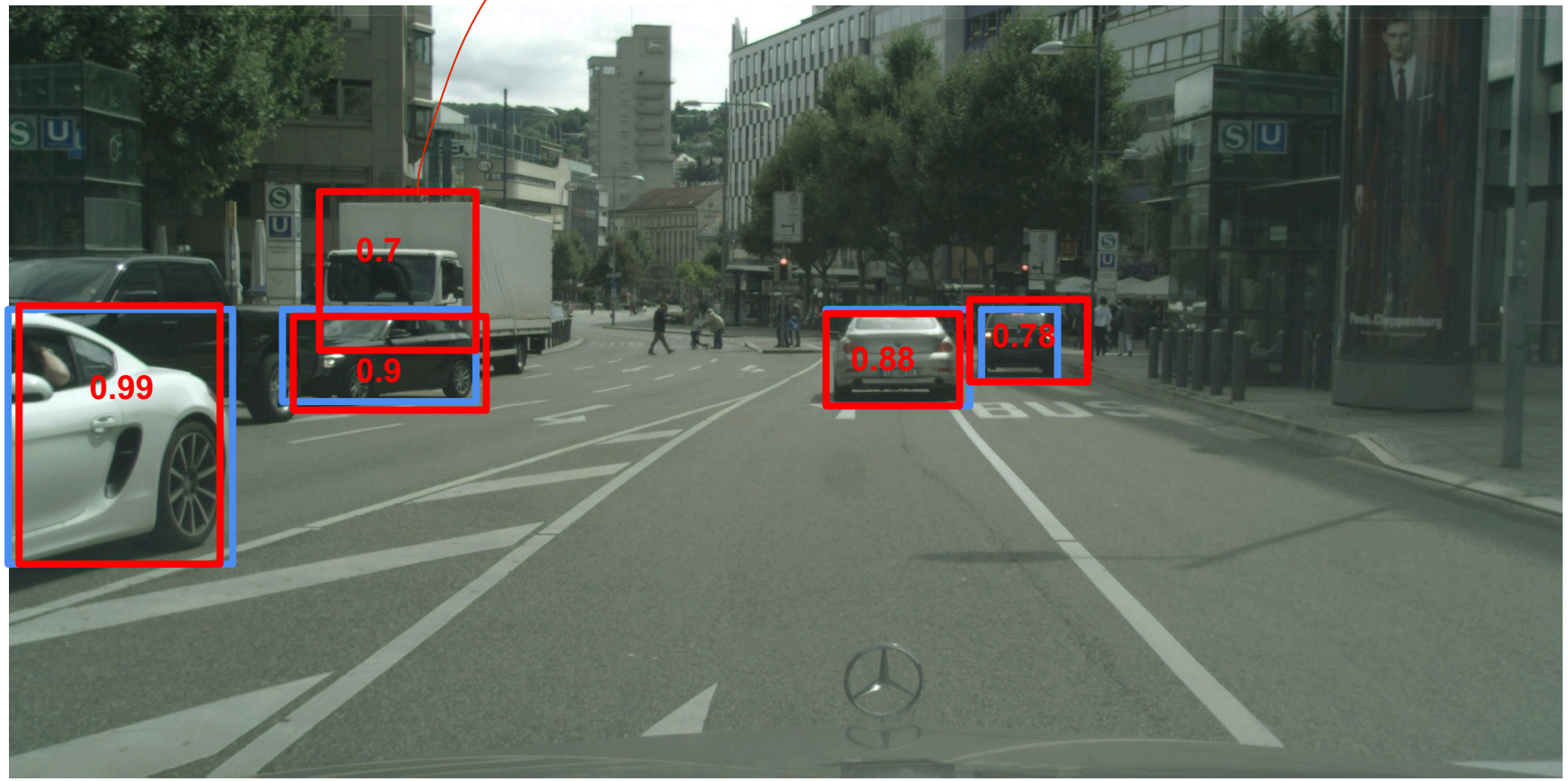
TP 3

FP 0

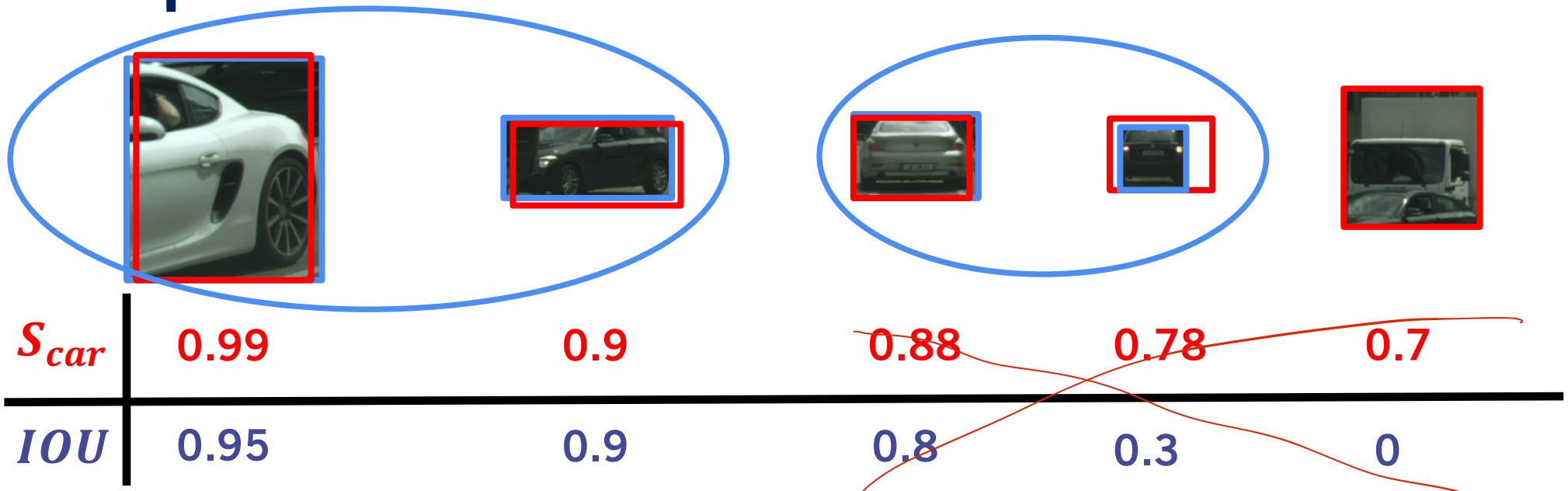
FN 2

$\frac{3}{3+2}$

Example



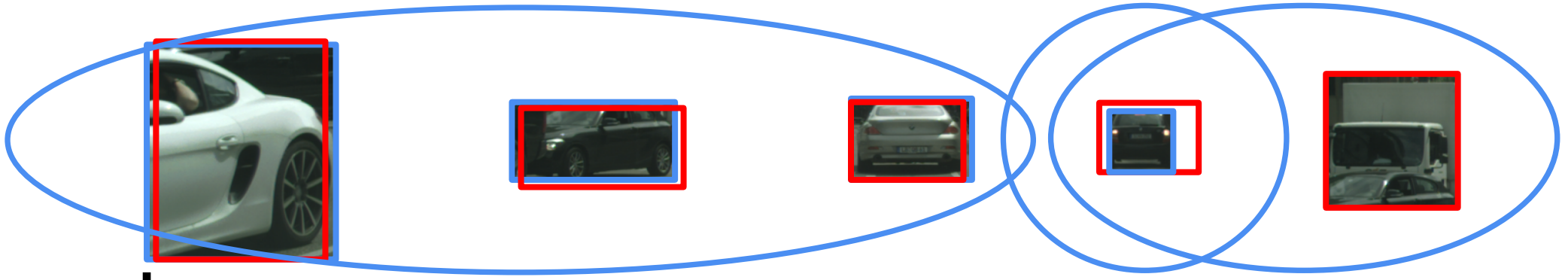
Example



- Score Threshold: 0.9
- IOU Threshold: 0.7

- TP = 2
- FP = 0
- FN = 2
- Precision = 2/2 = 1
- Recall = 2/4 = 0.5

Example



S_{car}	0.99	0.9	0.88	0.78	0.7
IOU	0.95	0.9	0.8	0.3	0

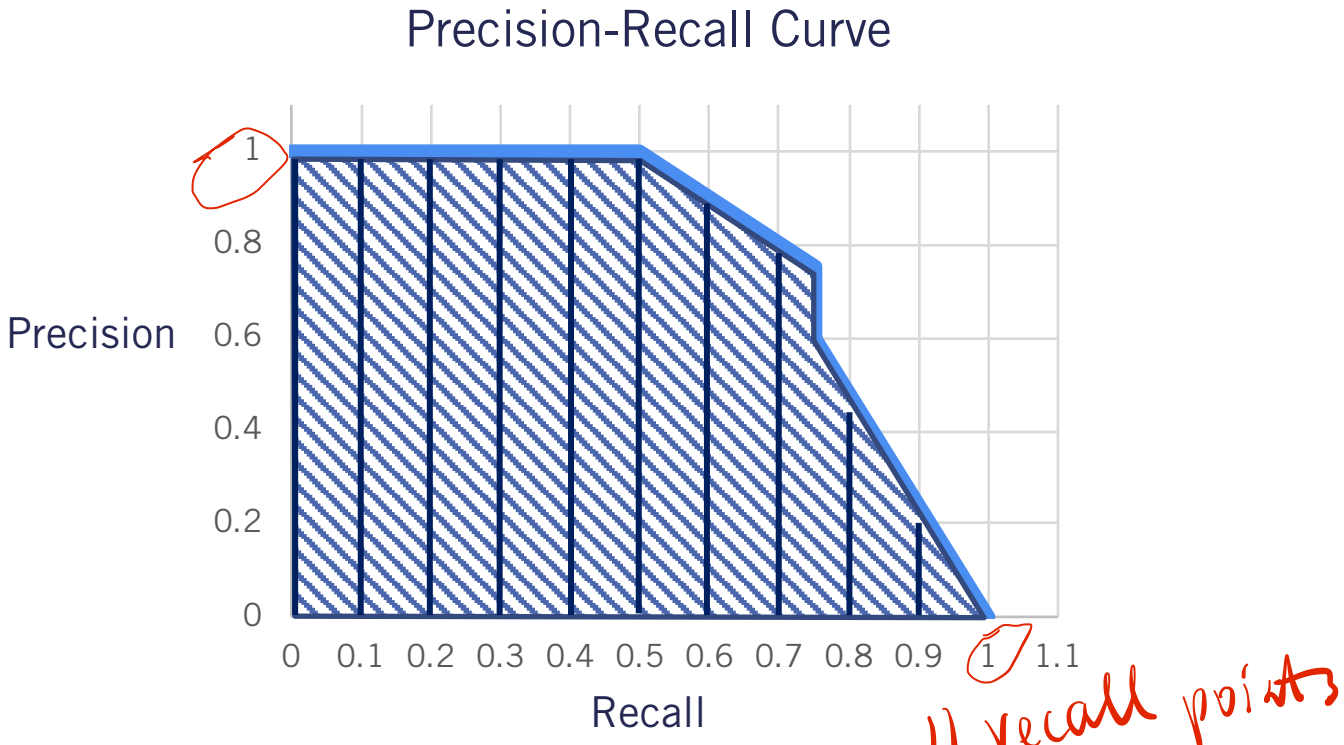
- Score Threshold: 0.7
- IOU Threshold: 0.7

Lowering
- less accurate detection results
- but detecting more objects

- TP = 3
- FP = 2
- FN = 1
- Precision = 3/5 = 0.6
- Recall = 3/4 = 0.75

Example

Score threshold	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
Precision	1	0.75	0.6	0.6	0.6	0.6	0.6	0.6	0.6
Recall	0.5	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75



$$AP = \frac{1}{11}\sum_{r=0}^{11} p_r \approx 0.75$$

Summary

- 2D object detection comprises of localizing an object and determining what the object is
- 2D object detectors are evaluated using the Average Precision metric, at a specific IOU threshold

- **Next: 2D object detection using ConvNets**