Chapter　02. Image Localization 구현 및 성능 개선하기

Image Localization

Computer Vision Task
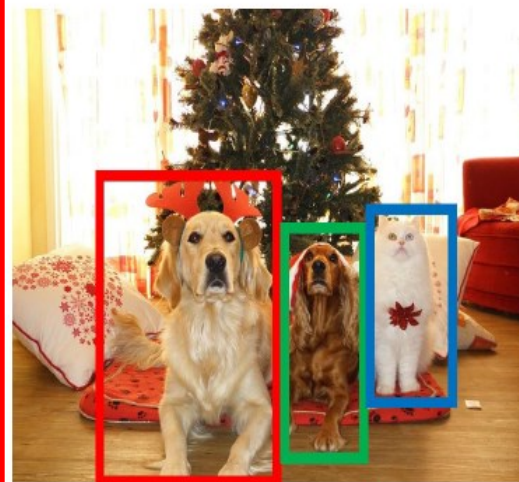
Classification & Localization

Classification: C classes
  Input: Image
  Output: Class label
  Evaluation metric: Accuracy

 → CAT

Localization:
  Input: Image
  Output: Box in the image (x, y, w, h)
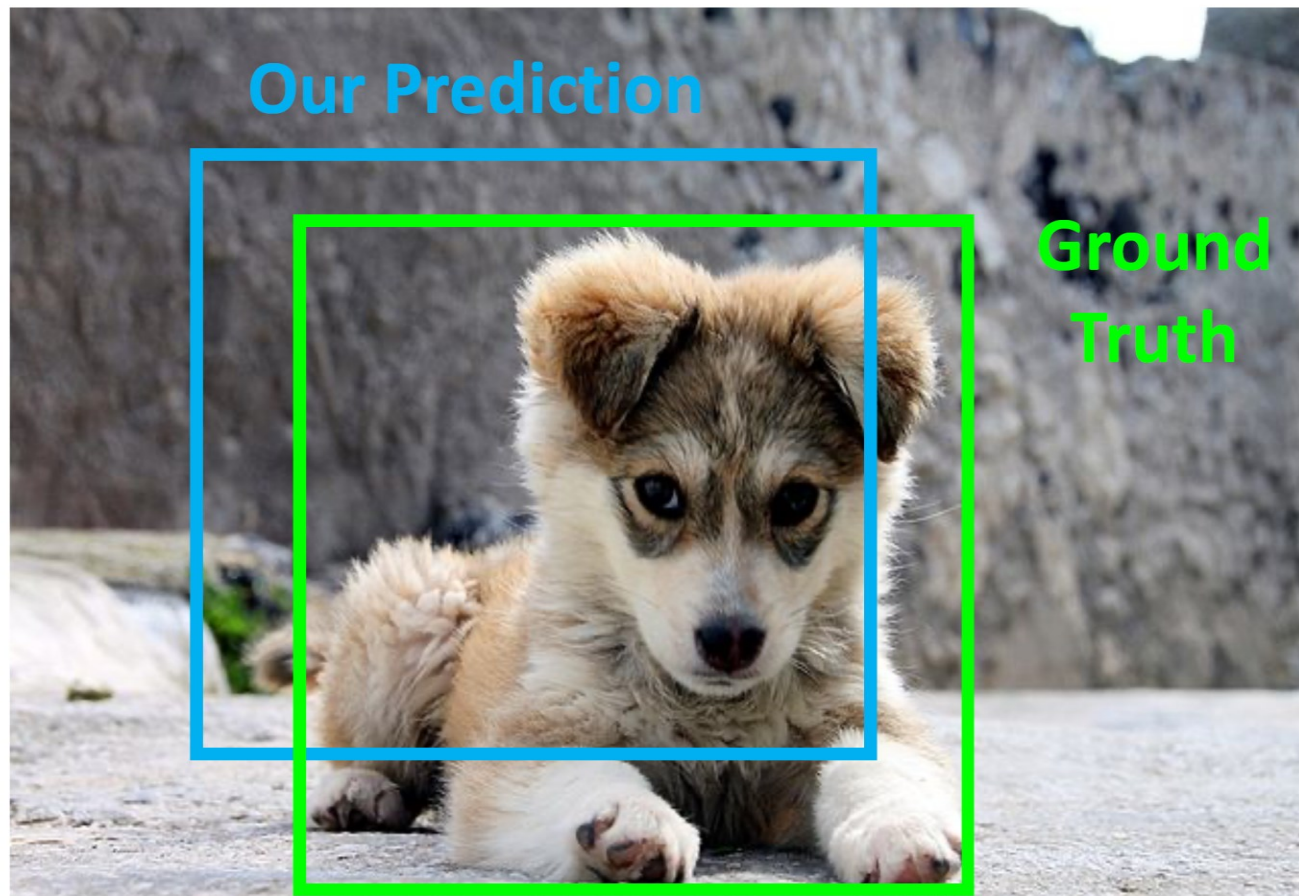 **Evaluation metric:** Intersection over Union

 → (x, y, w, h)

Classification + Localization: Do both

Comparing Boxes: Intersection over Union (IoU)

How can we compare our prediction to the ground-truth box?

Slide from Deep Learning for Computer Vision @Univ. of Michigan

Comparing Boxes: Intersection over Union (IoU)

How can we compare our prediction to the ground-truth box?

Intersection over Union (IoU) (Also called "Jaccard similarity" or "Jaccard index"):

$$\frac{Area\ of\ Intersection}{Area\ of\ Union}$$

Our Prediction

Ground Truth

Slide from Deep Learning for Computer Vision @Univ. of Michigan
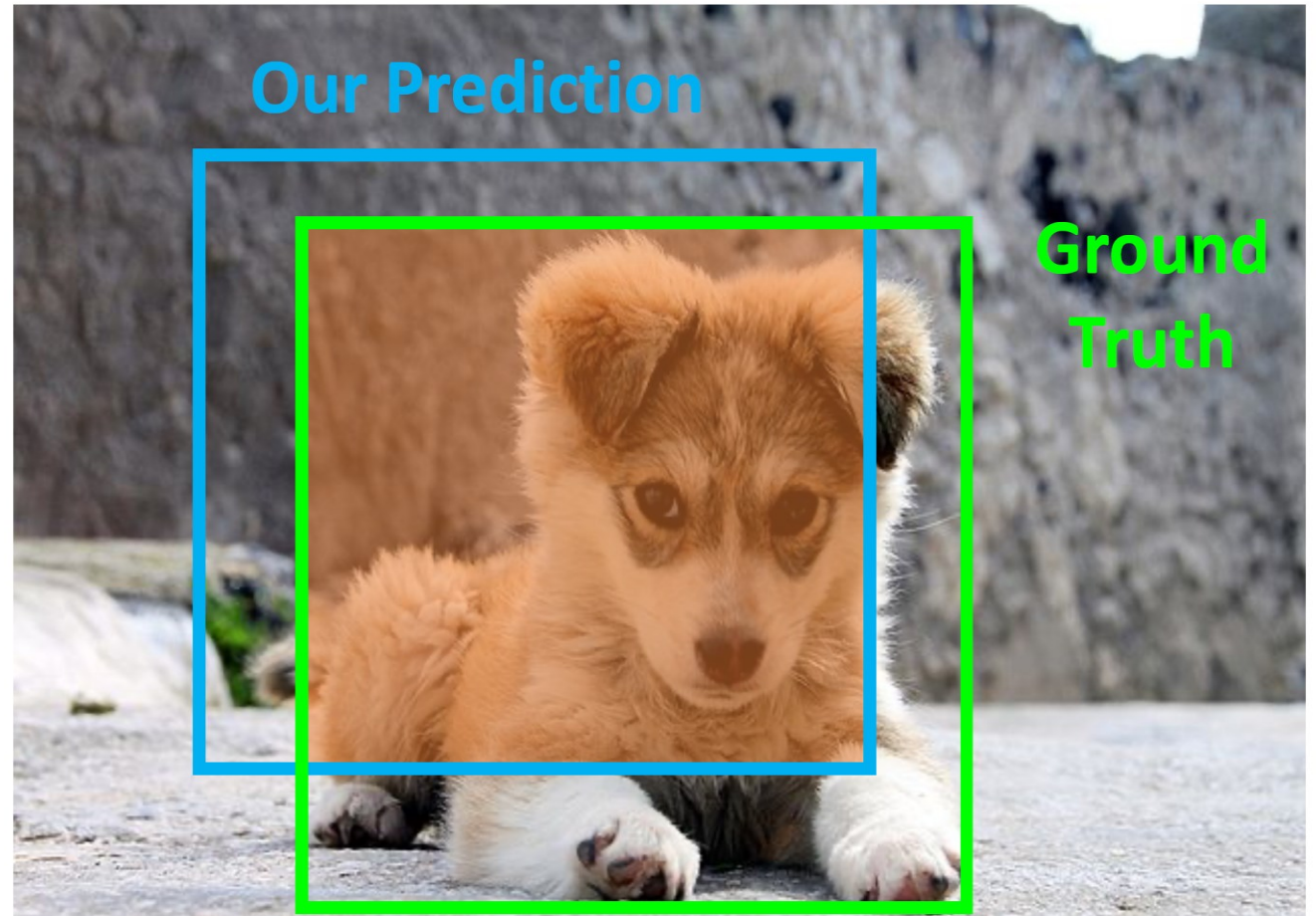
Comparing Boxes: Intersection over Union (IoU)

How can we compare our prediction to the ground-truth box?

Intersection over Union (IoU)
(Also called "Jaccard similarity" or "Jaccard index"):

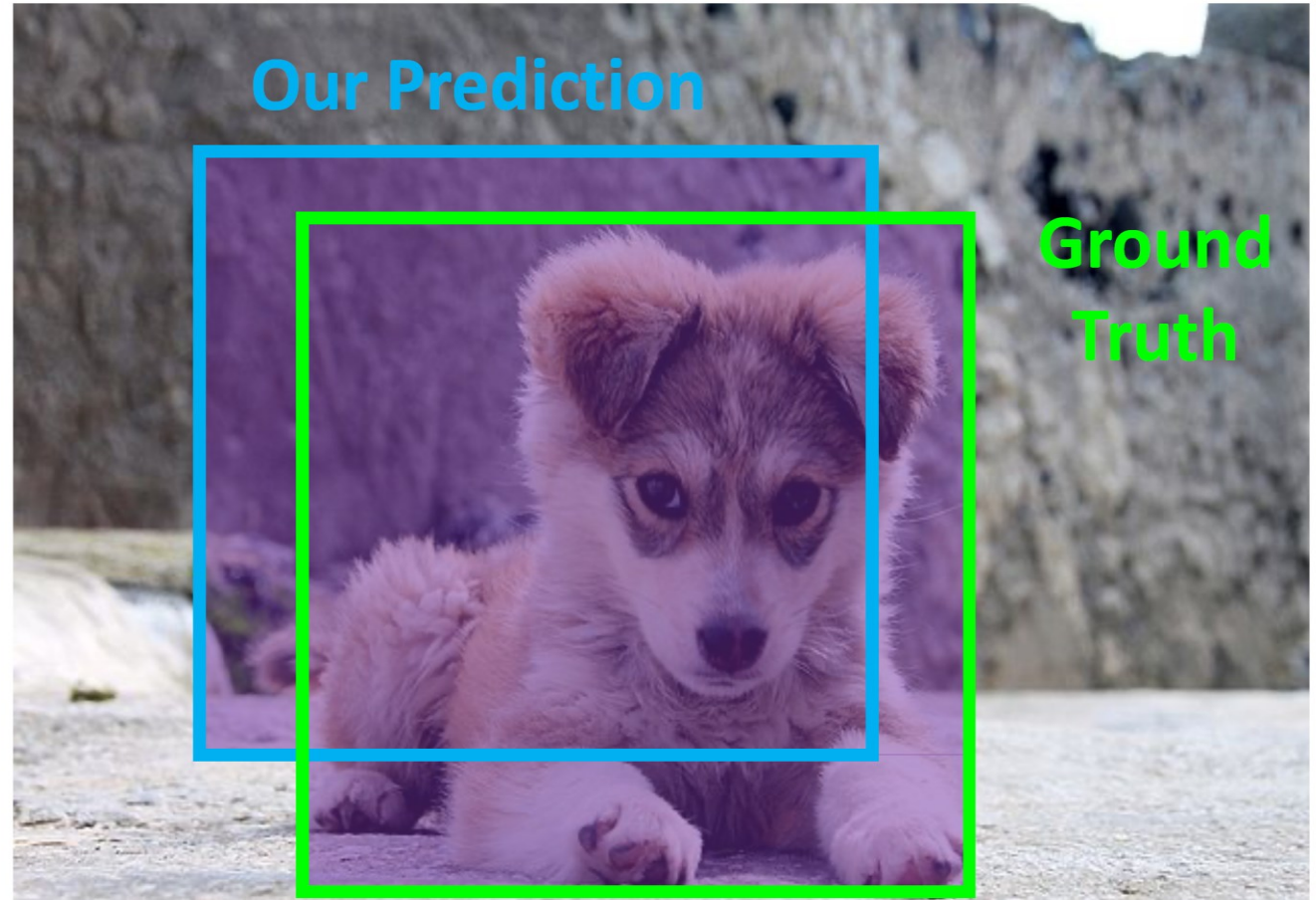$$\frac{\textit{Area of Intersection}}{\textit{Area of Union}}$$

Our Prediction

Ground Truth

Slide from Deep Learning for Computer Vision @Univ. of Michigan
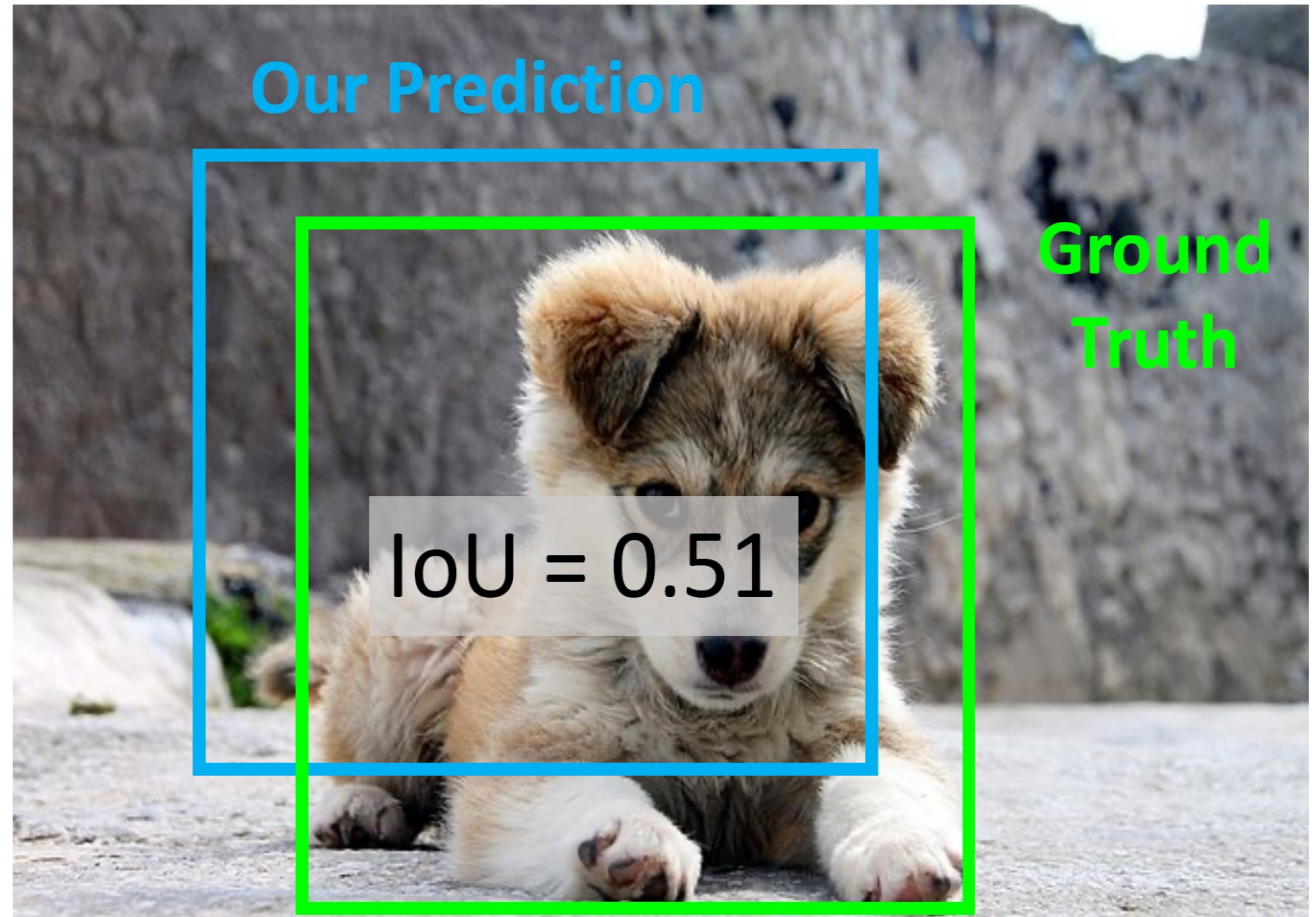
fast campus

Comparing Boxes: Intersection over Union (IoU)

How can we compare our prediction to the ground-truth box?

Intersection over Union (IoU)
(Also called "Jaccard similarity" or "Jaccard index"):

$$\frac{\textit{Area of Intersection}}{\textit{Area of Union}}$$

IoU > 0.5 is "decent"

Our Prediction

Ground Truth

IoU = 0.51

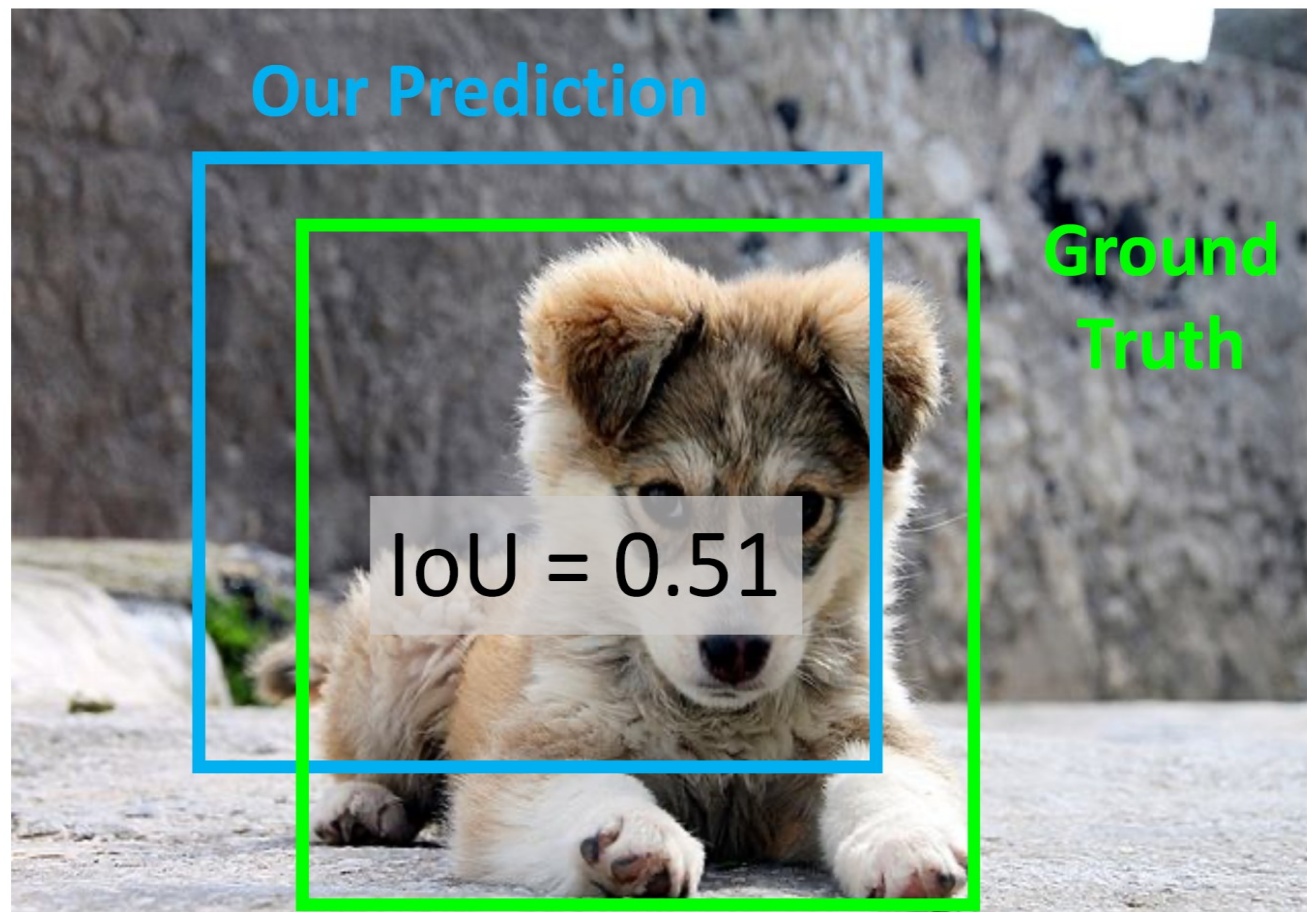Slide from Deep Learning for Computer Vision @Univ. of Michigan

Comparing Boxes: Intersection over Union (IoU)

How can we compare our prediction to the ground-truth box?

Intersection over Union (IoU)
(Also called "Jaccard similarity" or "Jaccard index"):

$$\frac{\textit{Area of Intersection}}{\textit{Area of Union}}$$

IoU > 0.5 is "decent"

Our Prediction

Ground Truth

IoU = 0.51

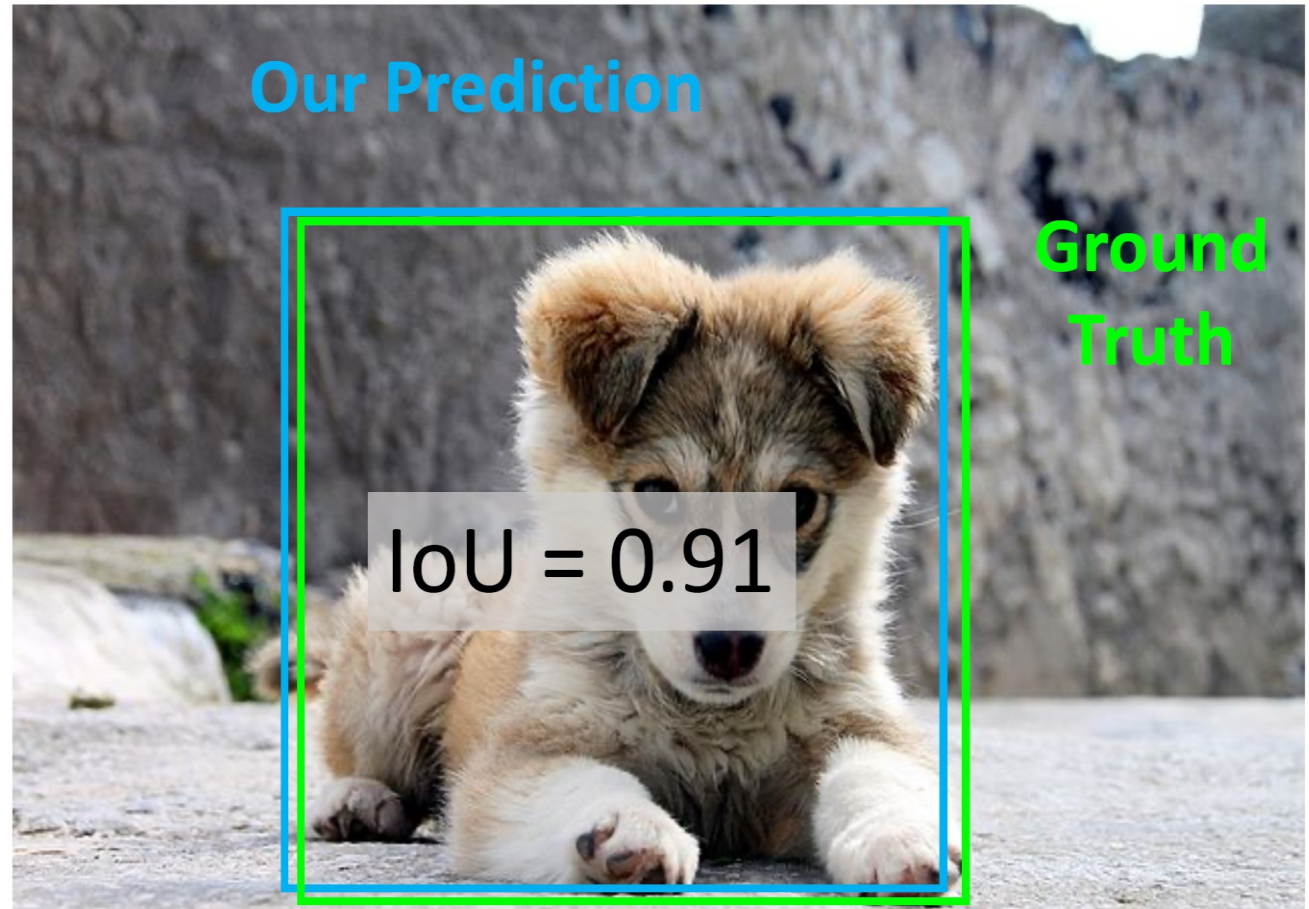Slide from Deep Learning for Computer Vision @Univ. of Michigan

Comparing Boxes: Intersection over Union (IoU)

How can we compare our prediction to the ground-truth box?

Intersection over Union (IoU)
(Also called "Jaccard similarity" or "Jaccard index"):

$$\frac{\textit{Area of Intersection}}{\textit{Area of Union}}$$

IoU > 0.5 is "decent",
IoU > 0.7 is "pretty good",
IoU > 0.9 is "almost perfect"

Our Prediction

Ground Truth

IoU = 0.91

Slide from Deep Learning for Computer Vision @Univ. of Michigan

fast campus

Localization

This image is CC0 public domain
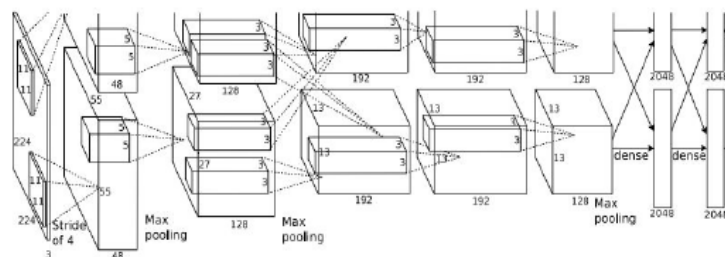
Vector: **Fully**
4096 **Connected**:
4096 to 4

Box
Coordinates ⟶ **L2 Loss**
(x, y, w, h)

Treat localization as a
regression problem!

Correct box:
(x', y', w', h')

fast
campus

Classification + Localization

Treat localization as a
regression problem!