

CSE4261: Neural Network and Deep Learning

Lecture: 26.06.2025



Sangeeta Biswas, Ph.D.
Associate Professor,
University of Rajshahi, Rajshahi-6205, Bangladesh

Some Videos on Application of Object Detection

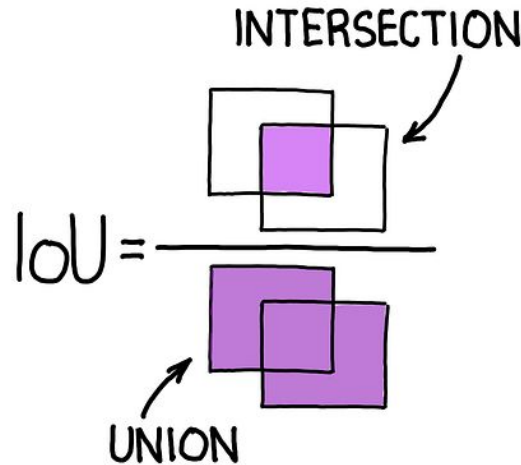
- <https://www.youtube.com/watch?v=GkiJoR5uWxl>
- <https://www.youtube.com/watch?v=srER2HKMkl0>
- <https://www.youtube.com/watch?v=BboT85CjG9o>
- <https://www.youtube.com/watch?v=lekdUXv82xQ>
- <https://www.youtube.com/watch?v=1DPU4F-0Gvo>
- <https://www.youtube.com/watch?v=RZsX6aoy8wo>
- <https://www.youtube.com/watch?v=yqvMuw-uedU>
- <https://www.youtube.com/watch?v=y8xq2wgCcNQ>

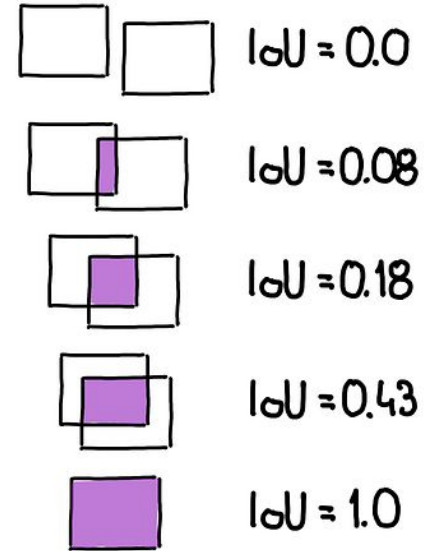
Recommended Andrew Ng's Lectures on Object Detection

- **C4W3L01 Object Localization:**
https://www.youtube.com/watch?v=GSwYGkTfOKk&list=PL_IHmaMAvkVxdDOBRg2CbcJBq9SY7ZUvs
- **C4W3L03 Object Detection:**
https://www.youtube.com/watch?v=5e5pjeojznk&list=PL_IHmaMAvkVxdDOBRg2CbcJBq9SY7ZUvs&index=3
- **C4W3L07 Nonmax Suppression:**
https://www.youtube.com/watch?v=VAo84c1hQX8&list=PL_IHmaMAvkVxdDOBRg2CbcJBq9SY7ZUvs&index=6
- **C4W3L08 Anchor Boxes:**
https://www.youtube.com/watch?v=RTlwl2bv0Tg&list=PL_IHmaMAvkVxdDOBRg2CbcJBq9SY7ZUvs&index=7
- **C4W3L09 YOLO Algorithm:**
https://www.youtube.com/watch?v=9s_FpMpdYW8&list=PL_IHmaMAvkVxdDOBRg2CbcJBq9SY7ZUvs&index=8
-

Intersection over Union (IoU)

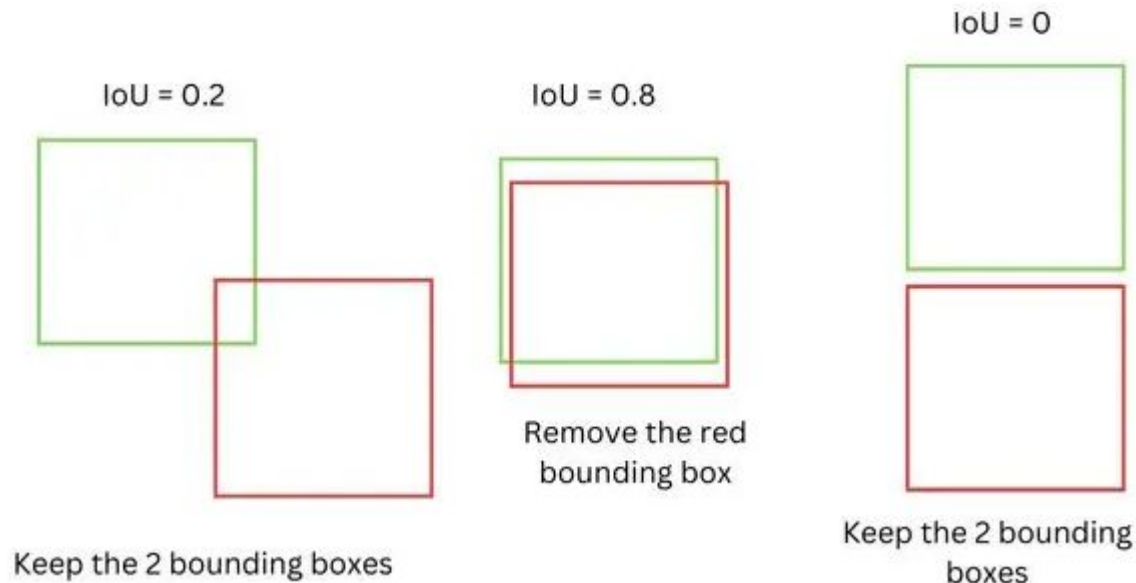
- It is a metric.
- It quantifies the overlap between a predicted bounding box or segmentation area and the corresponding ground truth (actual) object.
- A higher IoU value indicates a better overlap, signifying a more accurate prediction.

$$\text{IoU} = \frac{\text{INTERSECTION}}{\text{UNION}}$$




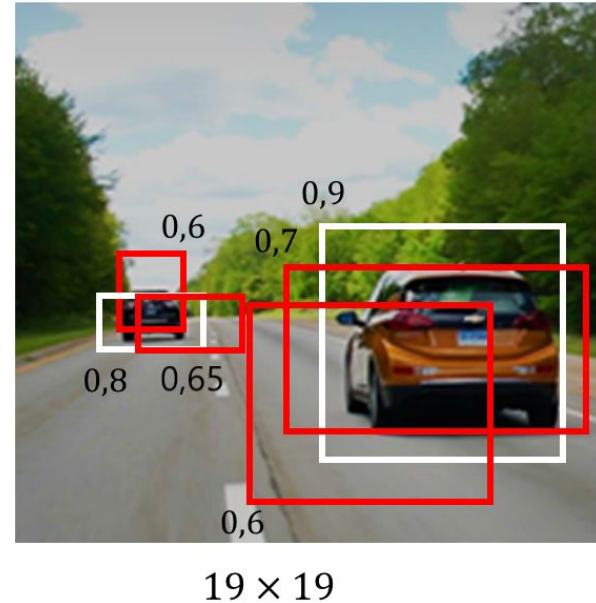
Non-Max Suppression (NMS)

- It is a post-processing technique
- It identifies and removes redundant or overlapping bounding boxes.
- It ensures that each object is represented by a single, precise bounding box.



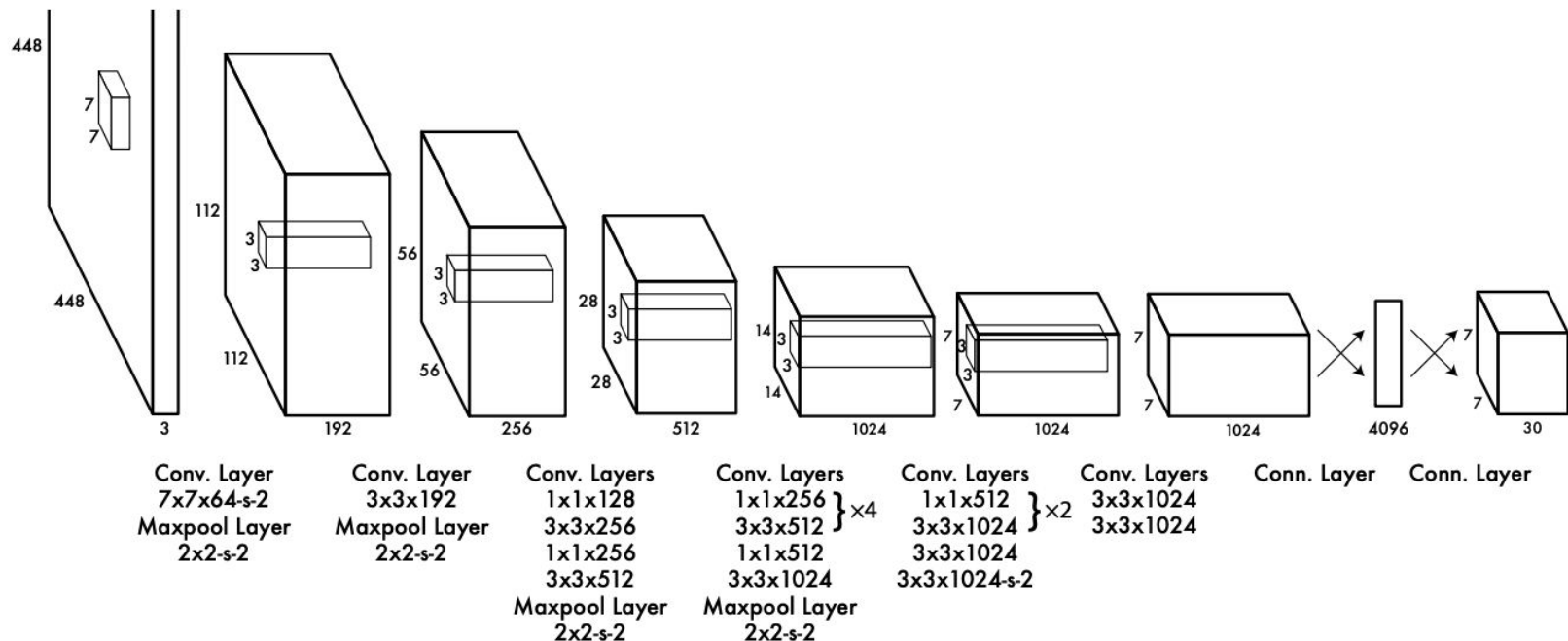
Steps of NMS

- Sort bounding boxes in descending order based on their confidence scores.
- Select the bounding box with the highest confidence score as a detection.
- Suppress or remove all other bounding boxes that have a significant overlap (usually measured by Intersection over Union or IoU) with the selected box.



Architecture of YOLOv1

- 24 Convolution Layers + 4 MaxPooling Layers + 2 Fully Connected Layers



Lab Work

WIDER Face Dataset:

- **32,203** images and label **393,703** faces
- Images are splitted into **61** event classes.
- For each event class, 40%, 10%, and 50% data were selected to prepare training, validation and testing sets respectively.
- Labels are not available for the test set
- <http://shuoyang1213.me/WIDERFACE/>

Lab Work: Prepare a face detector by applying transfer learning on Yolov8, YOLOv11 and YOLOv12 object detector pre-trained by the images of 80 classes of the COCO dataset.