

Practices in Visual Computing II
Spring 2023

Lab #2: Object Detection

Classification

Answers the question:
What is in this image?



CAT

Classification with localization

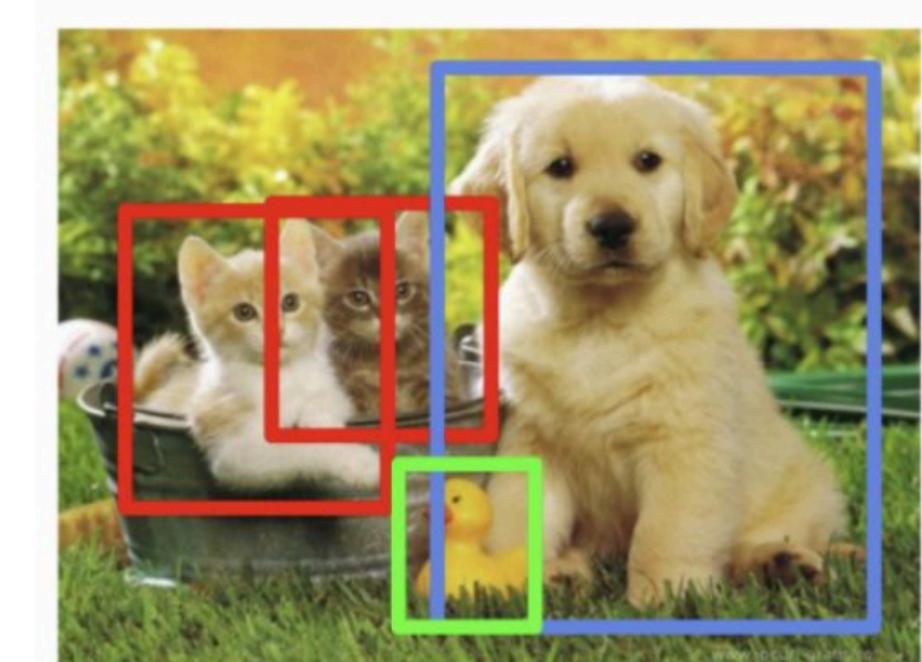
Answers the question:
*What is in this image
and where is it?*



CAT

Object Detection

Answers the question:
*What are all the objects
in this image and where
are they?*

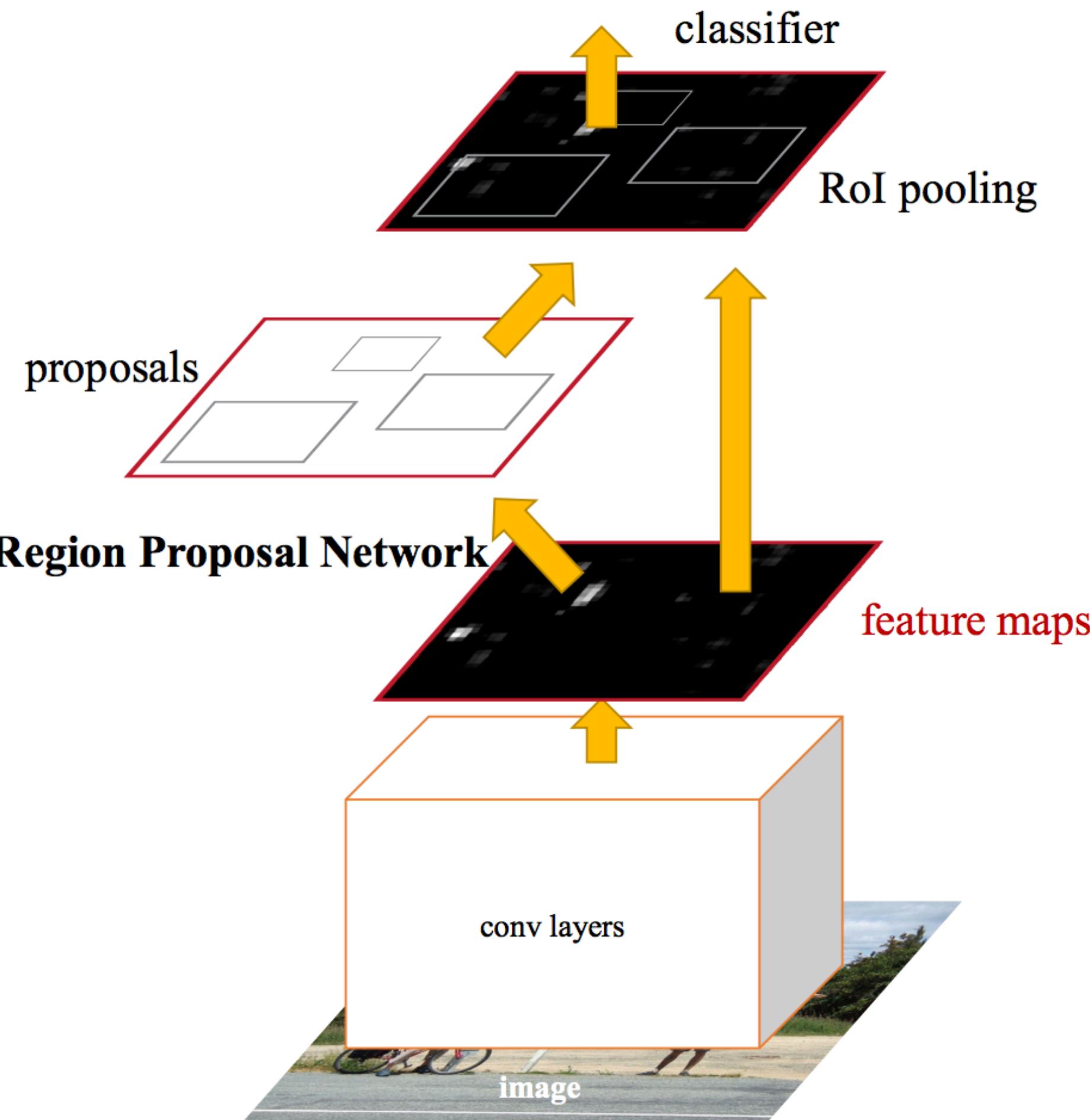


CAT, DOG, DUCK

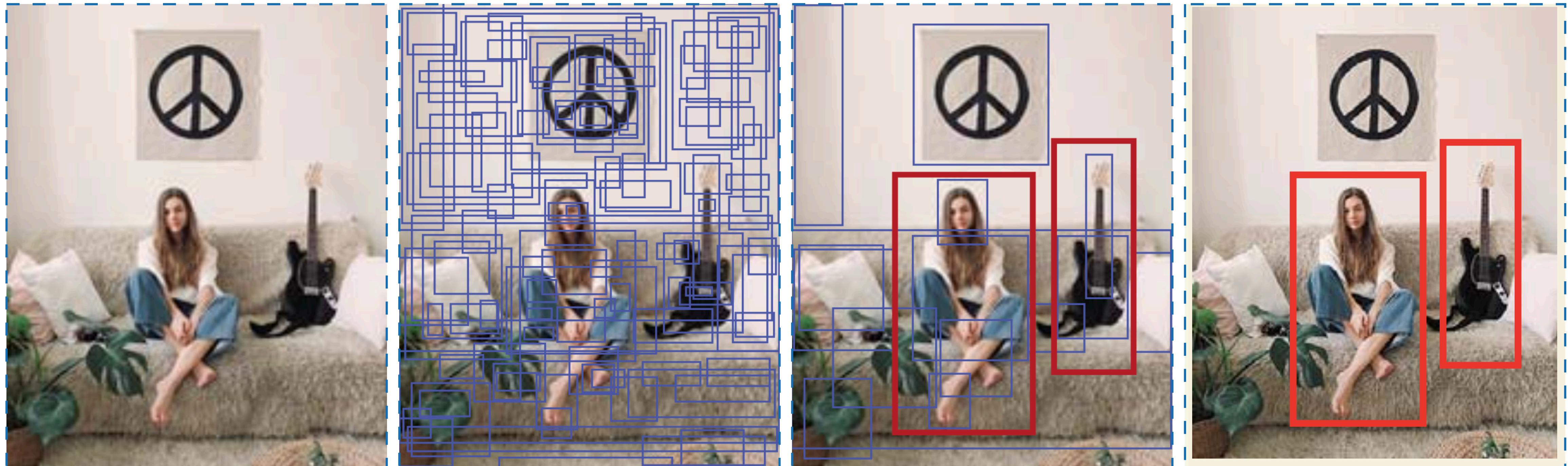
Sliding Windows



Region Proposal



Region Proposal



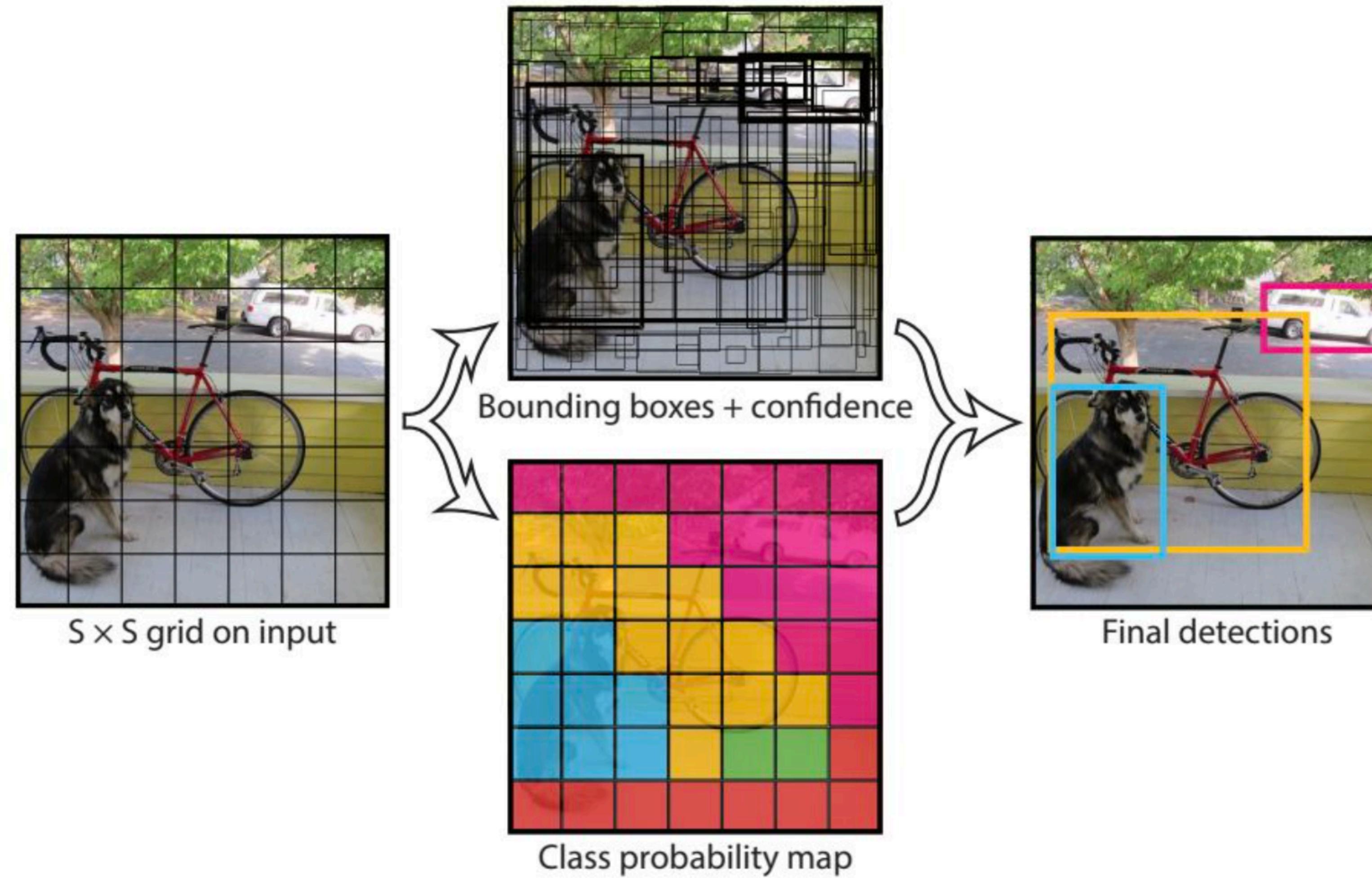
Single-Stage Methods

Why don't we replace the **ENTIRE** pipeline with a single convolutional network?

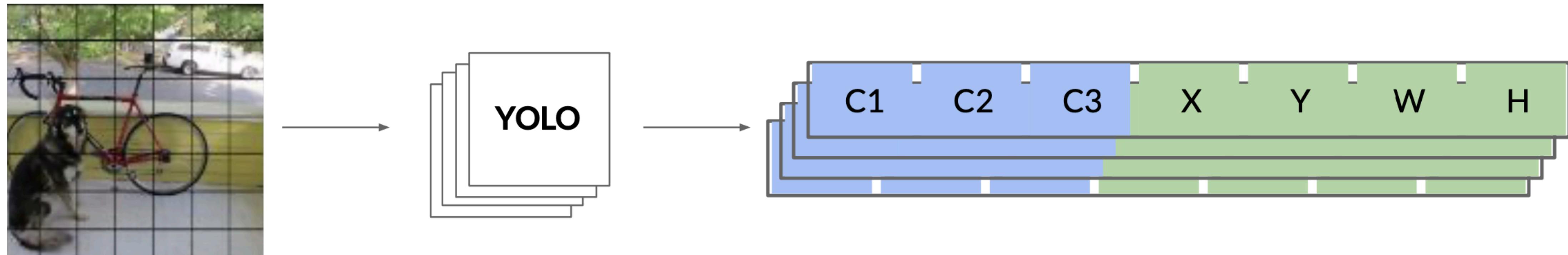
Task

Try `torchvision` multi-stage (**faster RCNN**) and single-stage (**SSD**) models.
Compare the inference speeds.

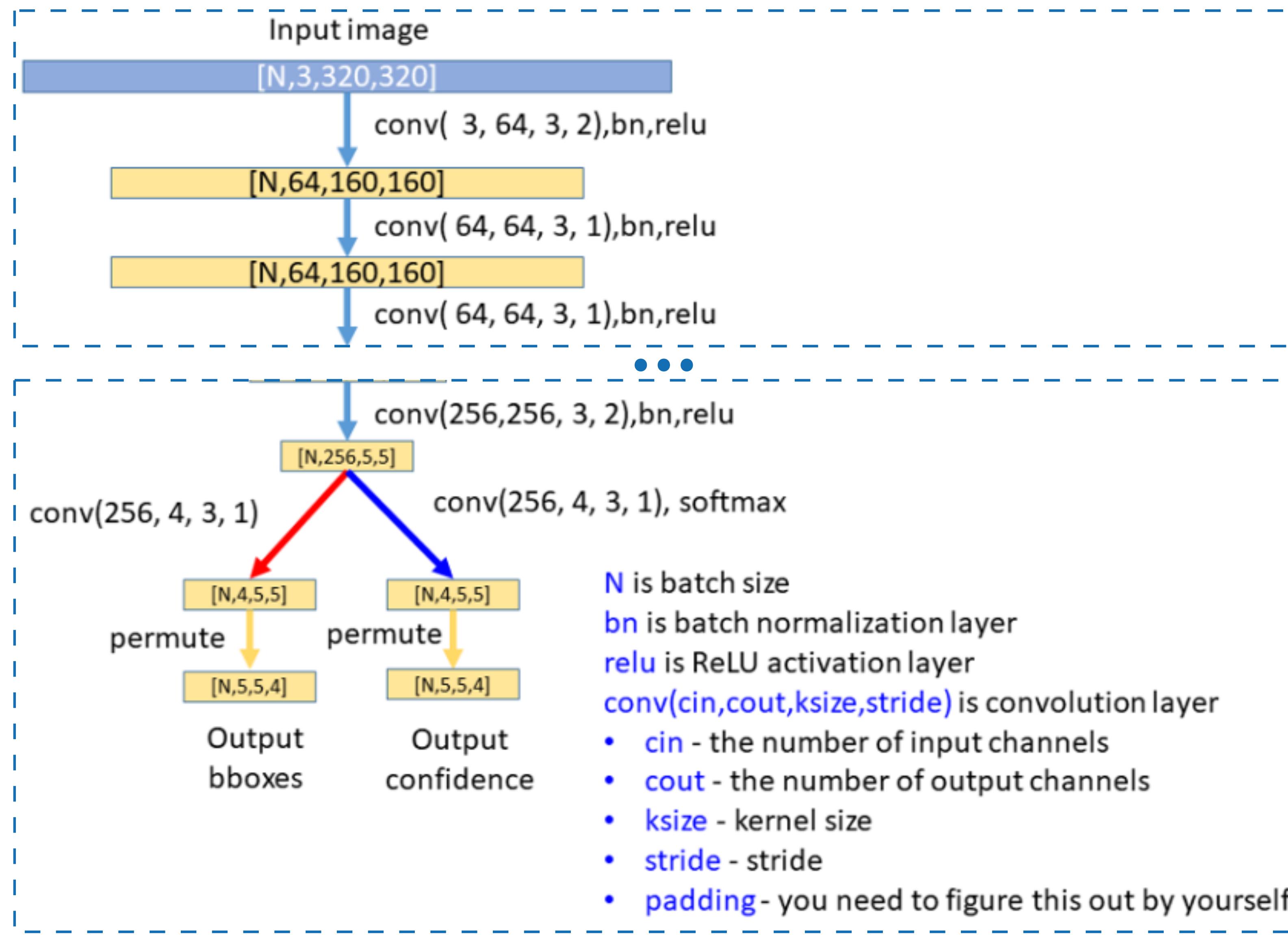
YOLO



YOLO



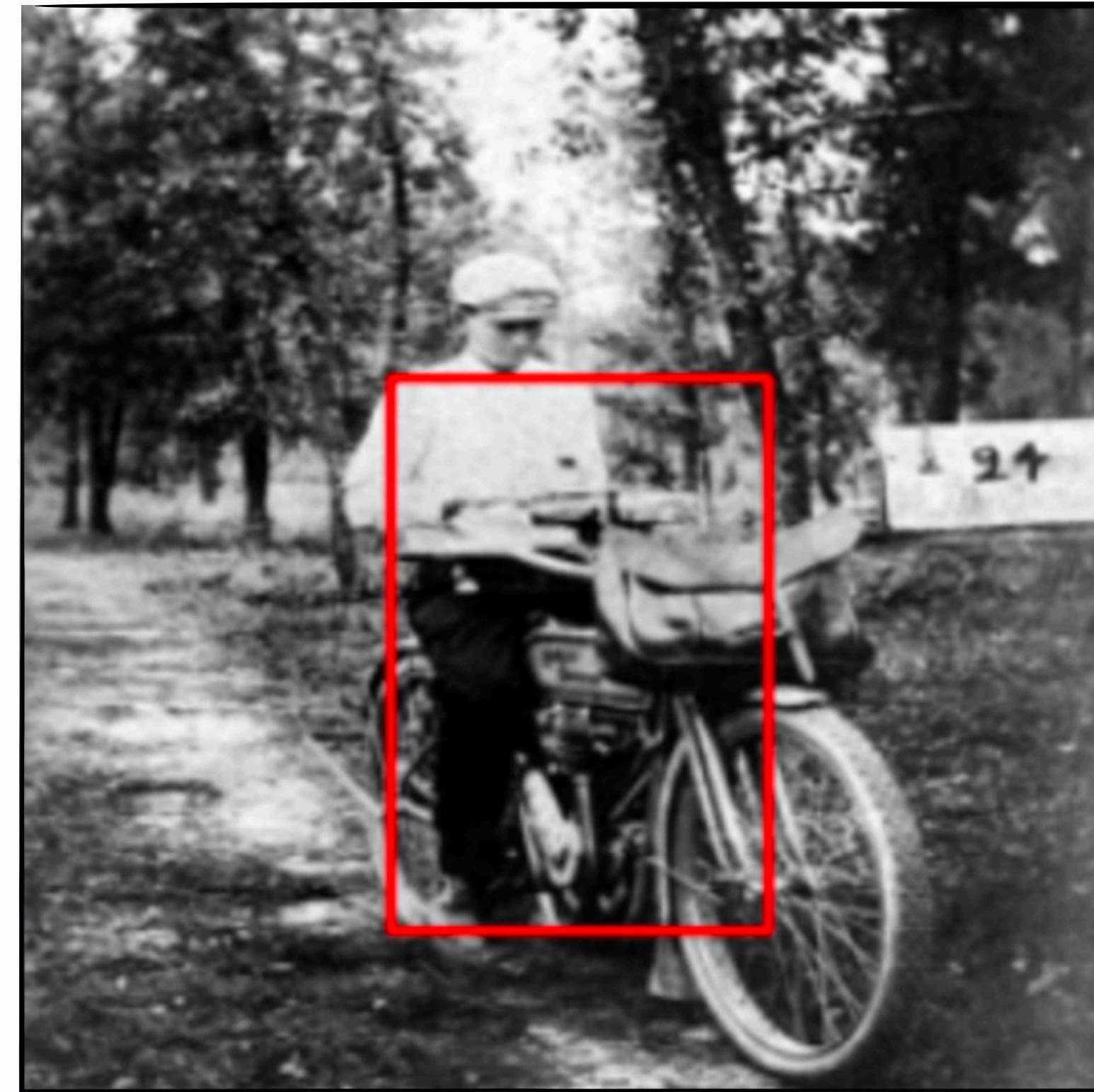
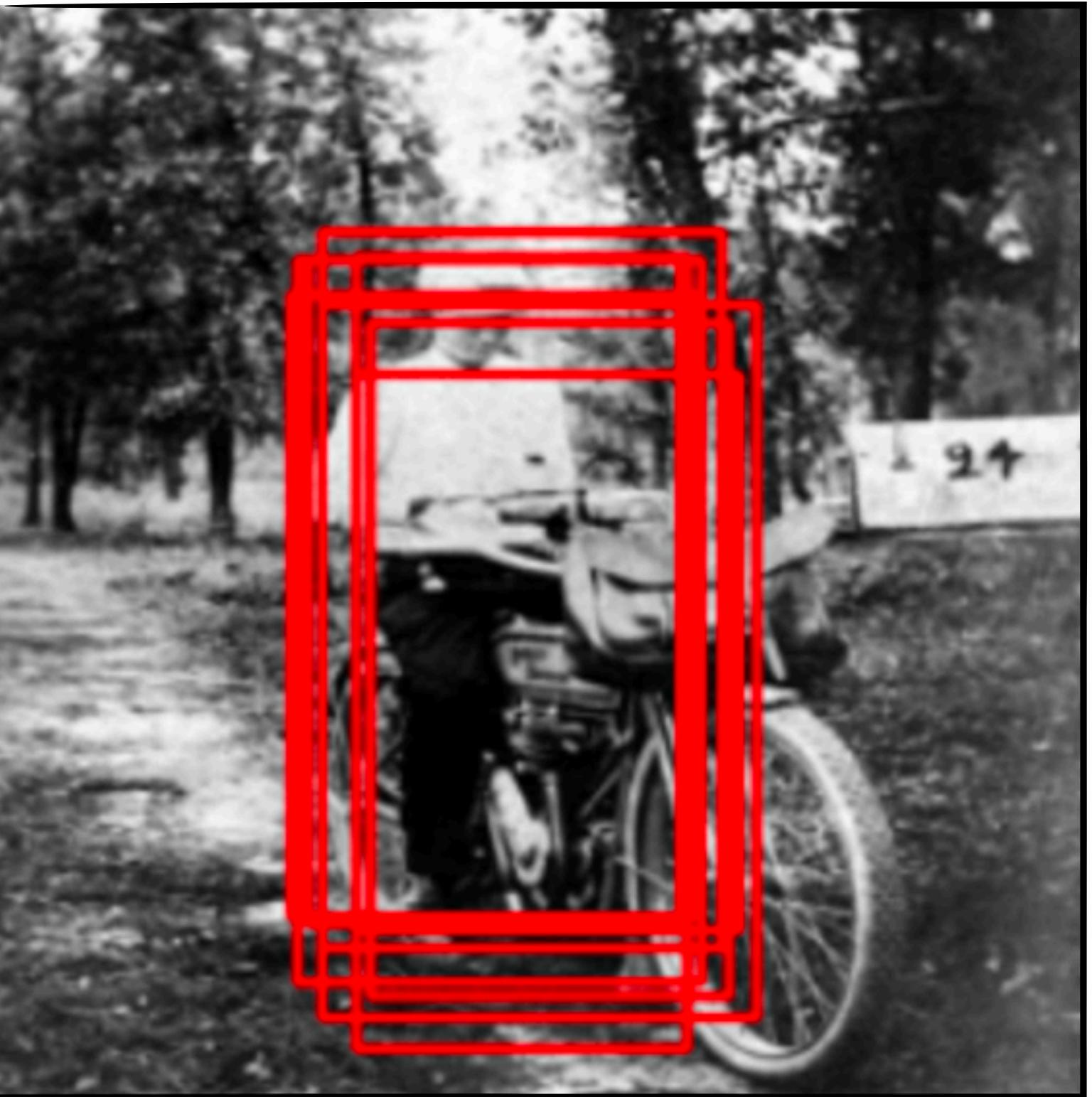
YOLO Architecture



Loss

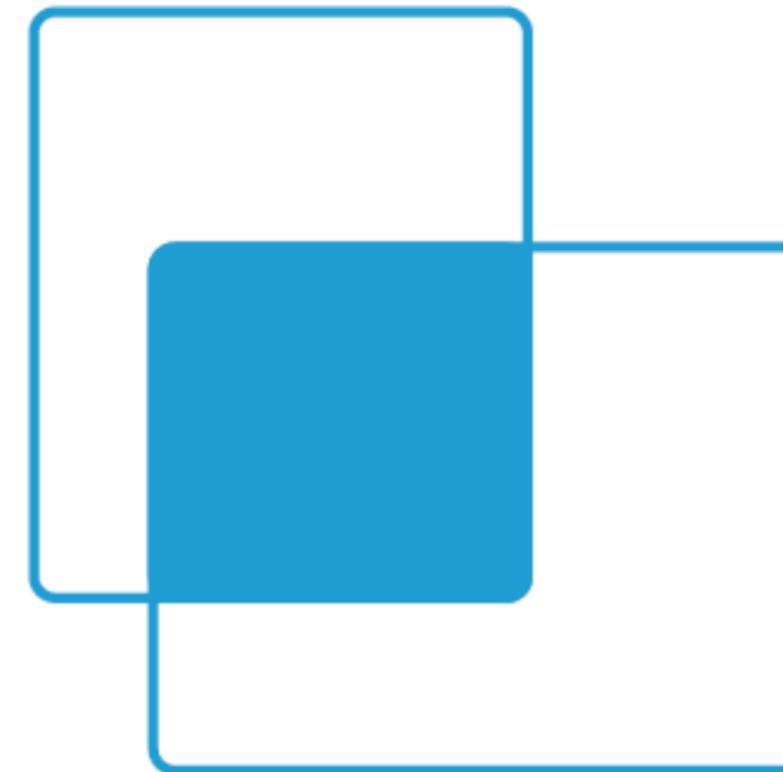
$$L_{cls} = \frac{1}{\sum_i x_i^{obj}} \sum_i x_i^{obj} \mathbf{crossentropy} \left(\text{conf}_i^{pred}, \text{conf}_i^{gt} \right) + 3 \cdot \frac{1}{\sum_i x_i^{noobj}} \sum_i x_i^{noobj} \mathbf{crossentropy} \left(\text{conf}_i^{pred}, \text{conf}_i^{gt} \right)$$
$$L_{box} = \frac{1}{\sum_i x_i^{obj}} \sum_i x_i^{obj} L1 \left(box_i^{pred}, box_i^{gt} \right)$$

Non-Maximum Suppression

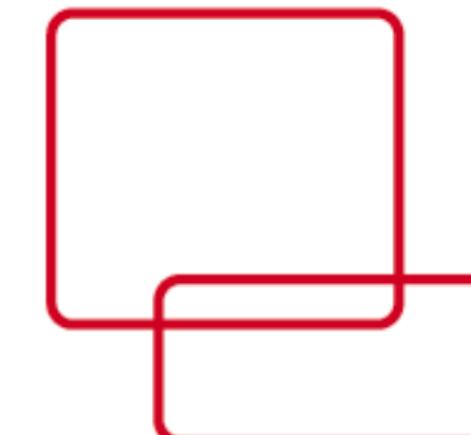


Intersection over Union (IOU)

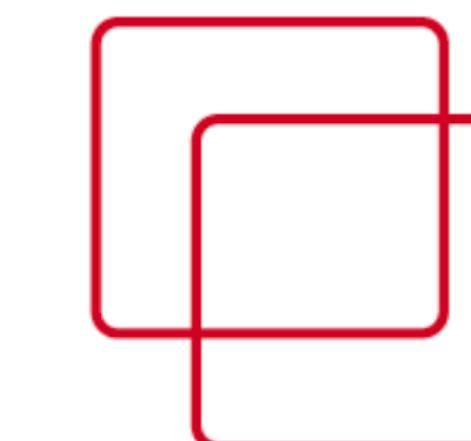
$$\text{IOU} = \frac{\text{Area of intersection}}{\text{Area of union}}$$



Examples



IOU = 0.1



IOU = 0.3



IOU = 0.6

Non-Maximum Suppression

```
def NMS(boxes, class_confidence, overlap, threshold):
    # Step 1: Filter out boxes with confidence < threshold
    filtered_boxes = []
    for box in boxes:
        if class_confidence[box] > threshold:
            filtered_boxes.append(box)
    # Step 2: Sort the boxes by confidence
    result_boxes = []
    for b1 in filtered_boxes:
        discard = False
        for b2 in filtered_boxes:
            if IOU(b1, b2) > overlap:
                if class_confidence[b2] > class_confidence[b1]:
                    discard = True
        if not discard:
            result_boxes.append(b1)
    return result_boxes
```