

CS181: Computer Vision

Object Detection (classical)

Tuesday, September 29th, 2026



Md Alimoor Reza

Associate Professor of Computer Science

What type of recognition problem?



Object Recognition

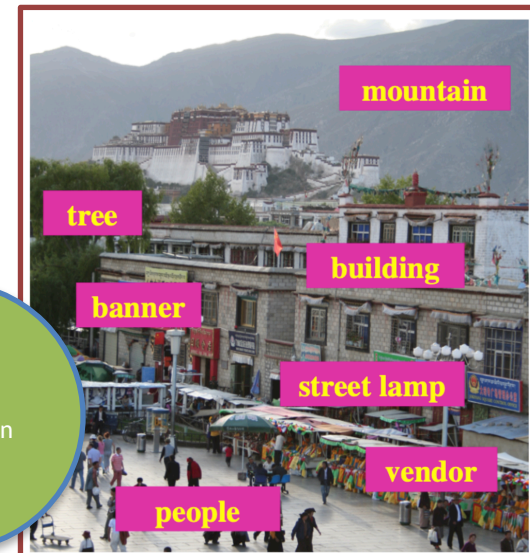


What type of recognition problem?



Single label classification

What type of recognition problem?



Multi-label classification

What type of recognition problem?



Scene and context classification

Types of Recognition Tasks

- Image classification
- Object detection
 - Specific objects categories
 - Generic objects categories
- Semantic segmentation

Image Classification: what are the objects in this image?



<u>Label</u>		<u>Probability</u>
Person	+	0.99
Building		0.01
Cat		0.00
Dog		0.01
...		...
...		...
...		...
...		...

Input:

1. an image



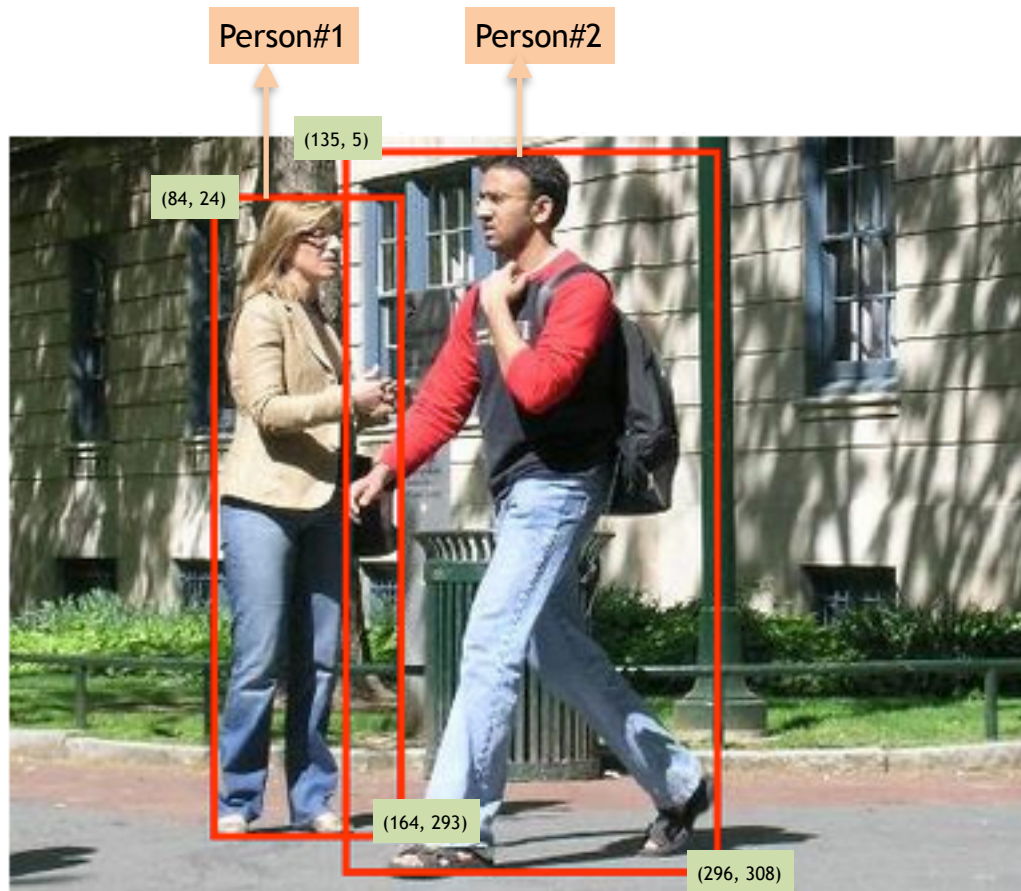
Classification model



Output:

1. Labels of prediction
2. Probability of labels

Object Detection: what are the locations of objects?



[xmin, ymin, xmax, ymax]

<u>Label</u>	<u>Probs.</u>	<u>Bounding Box Coordinates</u>
Person#1	0.99	[84, 24, 164, 293]
Person#2	0.99	[139, 5, 296, 308]
...	...	...
...	...	...
...	...	...
...	...	...
...	...	...
...	...	...

Input:
1. an image



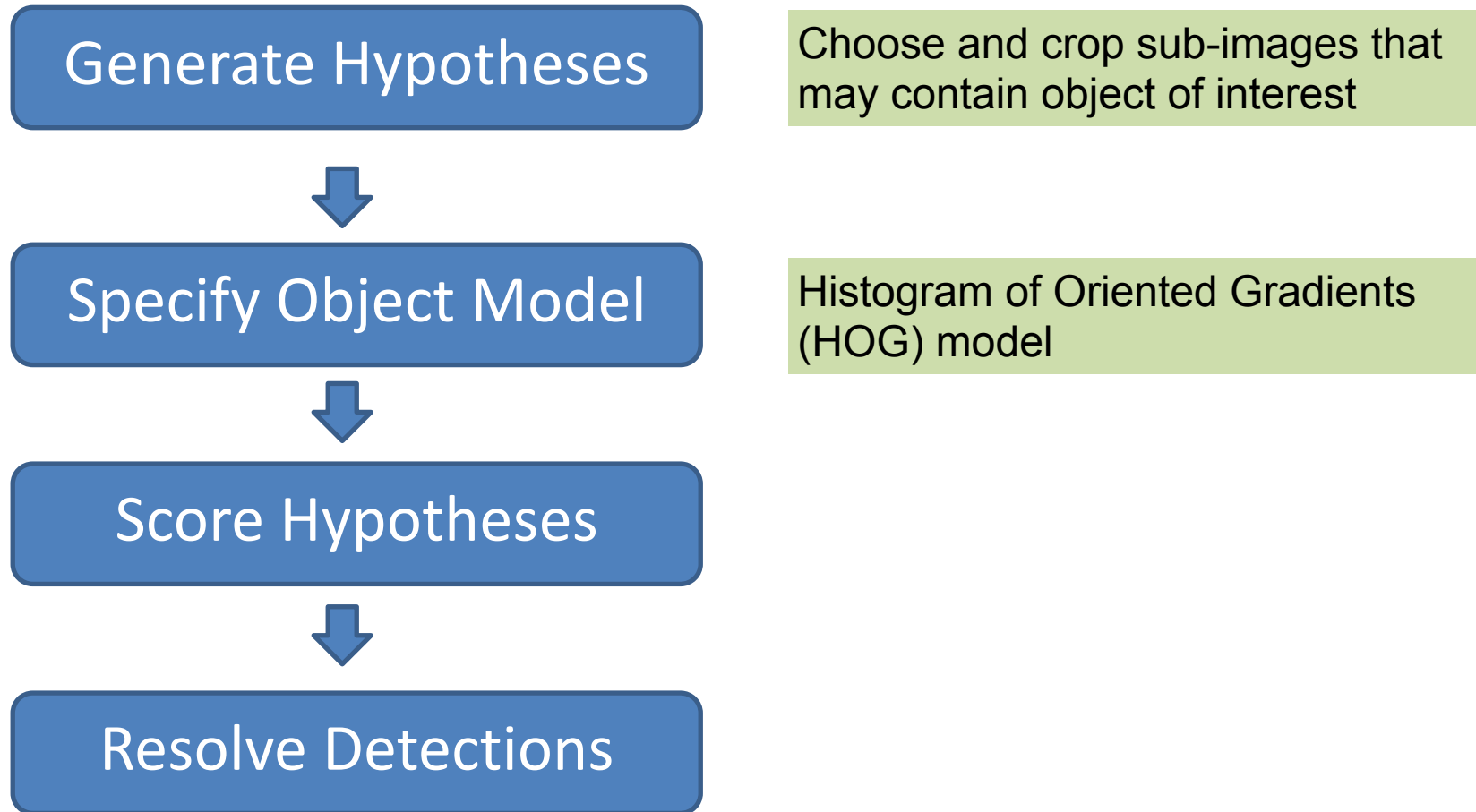
Detection model



Output:

1. Labels of prediction
2. Probabilities of labels
3. Locations of rectangular bounding boxes associated with each label

General Process of Object Detection



Slide credit: Derek Hoiem

Generating hypotheses

One option: Sliding window

- Test patch at each location and scale



Scale#1: Retain original resolution of the image

Test the image at each location of its original resolution: 400x300 pixels

Slide credit: Derek Hoiem

Generating hypotheses

One option: Sliding window

- Test patch at each location and scale



Scale#2: Resize the image 0.5 times of its original resolution.

For example, if the original resolution of the image is 400x300 pixels, after scaling by 0.5 factor it becomes 200x150

Test the image at each location of this scaled image 200x150 pixels

Slide credit: Derek Hoiem

Sliding window: a simple hypothesis generation technique

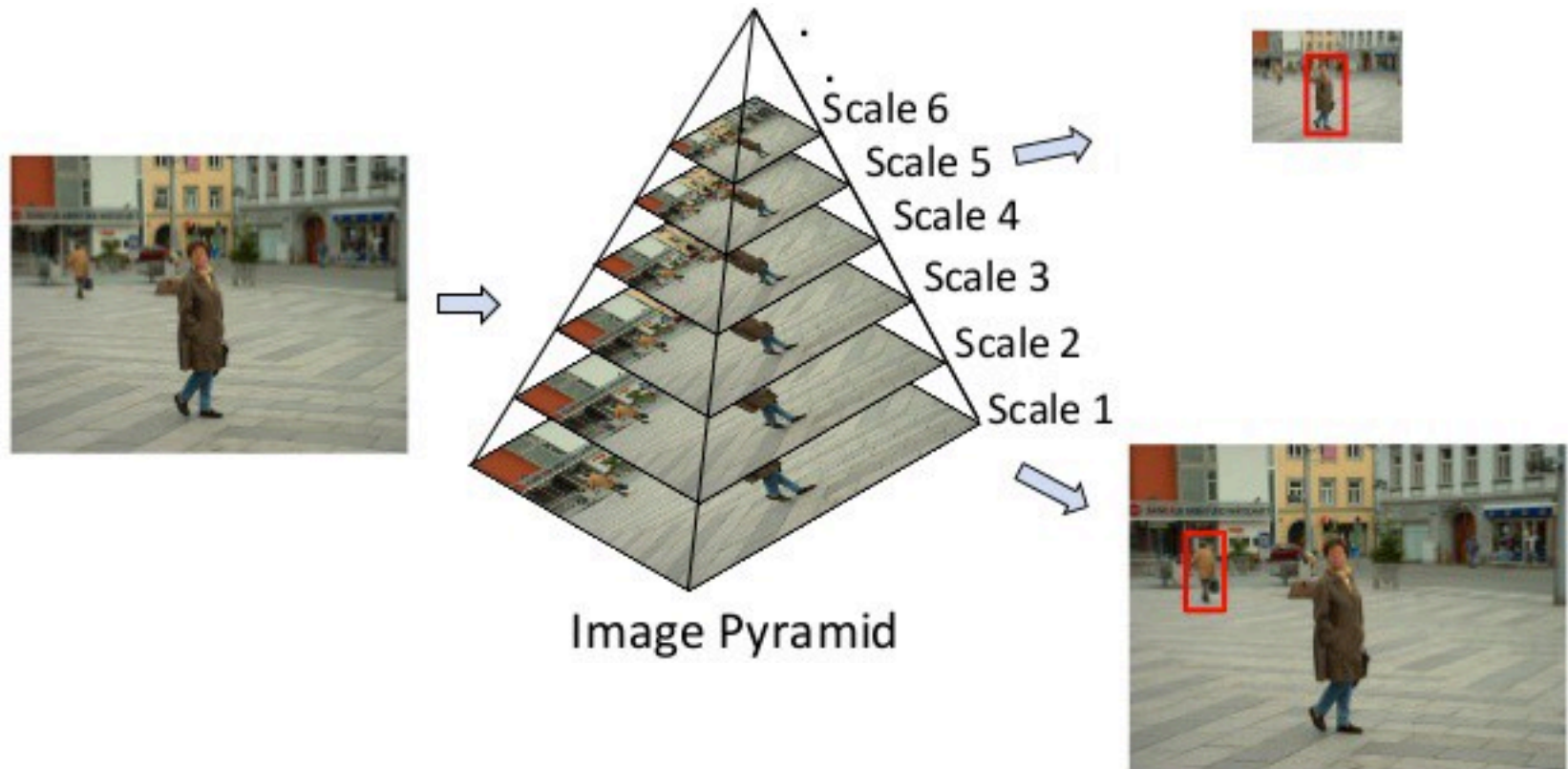


Scale#1
400x300 pixels



Scale#2
200x150 pixels

Sliding window: a simple hypothesis generation technique



Generated hypothesis samples at various locations and scales



Model: HOG Detector

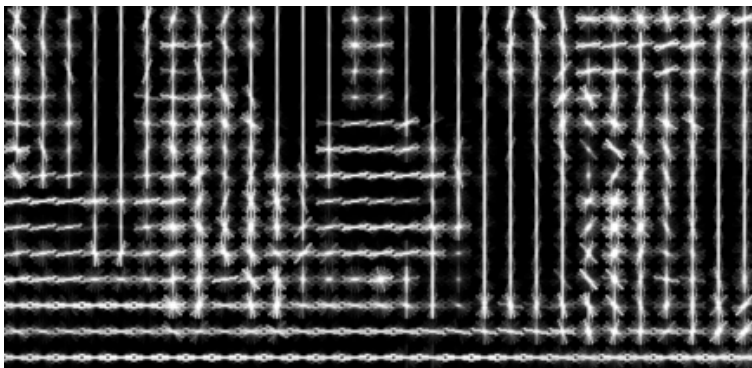


1. Extract fixed-sized (64x128 pixel) window at each position and scale
2. Compute HOG (histogram of gradient) features within each window
3. Score the window with a linear SVM classifier (or your any other favorite classifier)
4. Perform non-maxima suppression to remove overlapping detections with lower scores

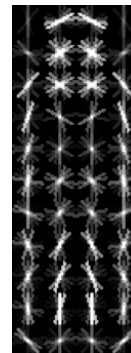
N. Dalal and B. Triggs, [Histograms of Oriented Gradients for Human Detection](#), CVPR 2005

Pedestrian Detection (Binary Problem)

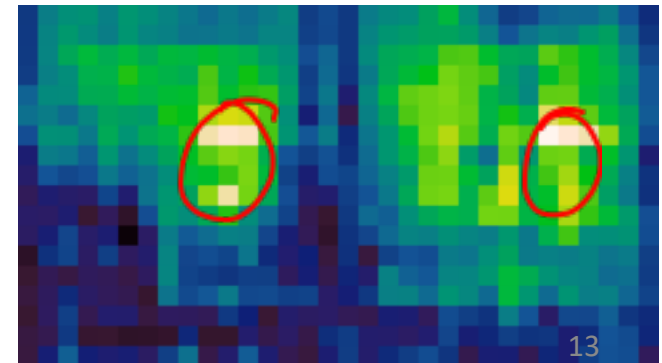
- Features: Histograms of oriented gradients (HOG)
 - Partition image into 8x8 pixel blocks and compute histogram of gradient orientations in each block
- Learn a pedestrian template using a linear SVM-classifier
 - At test time, convolve feature map with template



HOG feature map

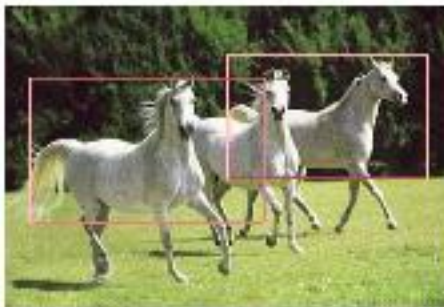


Template



Detector response map

Other Applications of Object Detection



Motorbike Detection: Examples

Correct Detections

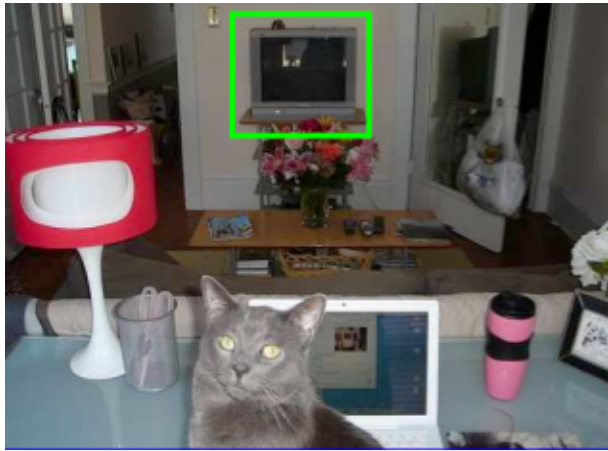


False Positives

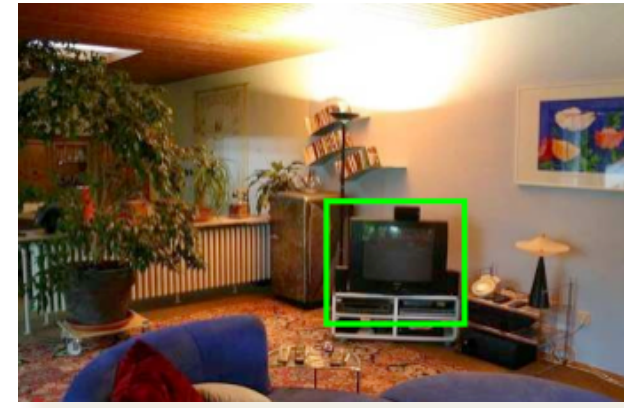
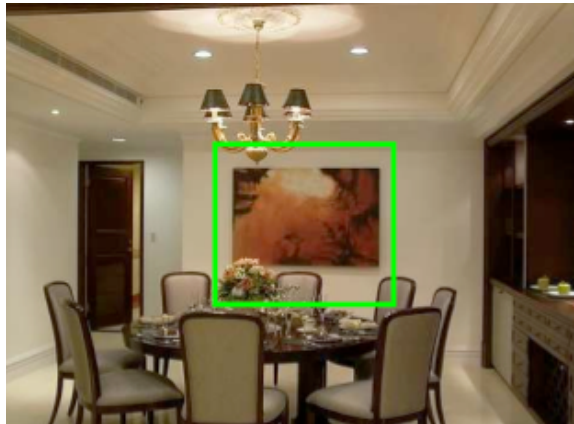


TV/Monitor Detections: Examples

Correct Detections



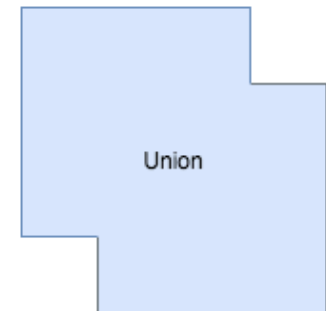
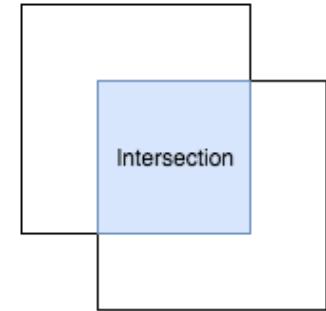
False Positives



How do you evaluate your detection?

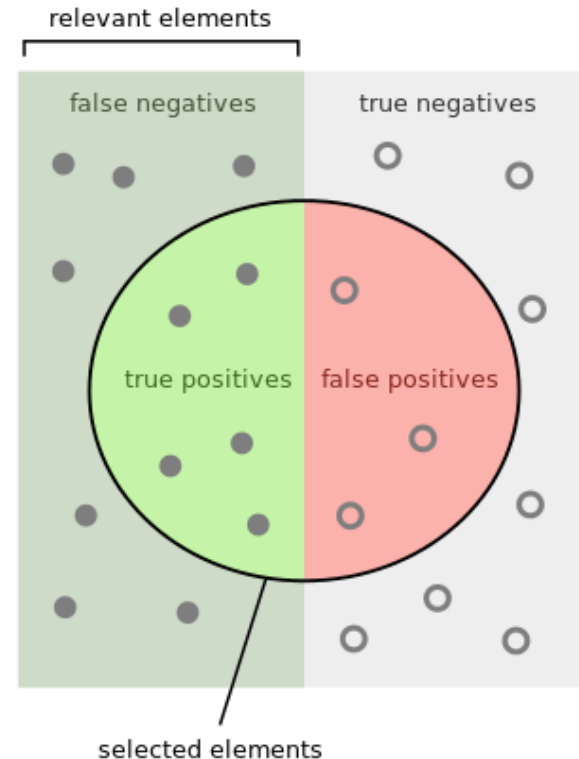
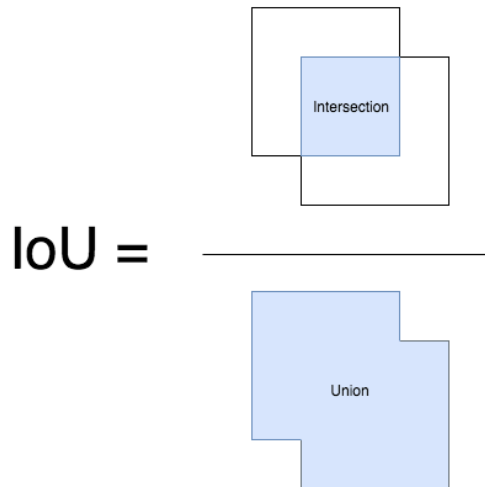
- Intersection over Union (IoU):
 - evaluation metric to assess the detection performance. It requires computation of two entities
 - **Numerator:** Intersection between the predicted bounding box and the ground-truth bounding box
 - **Denominator:** Union between the predicted bounding box and the ground-truth bounding box

$\text{IoU} =$



How do you evaluate your detection?

- $\text{IoU} > 0.5$ is True Positive
- $\text{IoU} < 0.5$ is False Positive



How many selected items are relevant?

Precision = $\frac{\text{true positives}}{\text{true positives} + \text{false positives}}$

How many relevant items are selected?

Recall = $\frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$