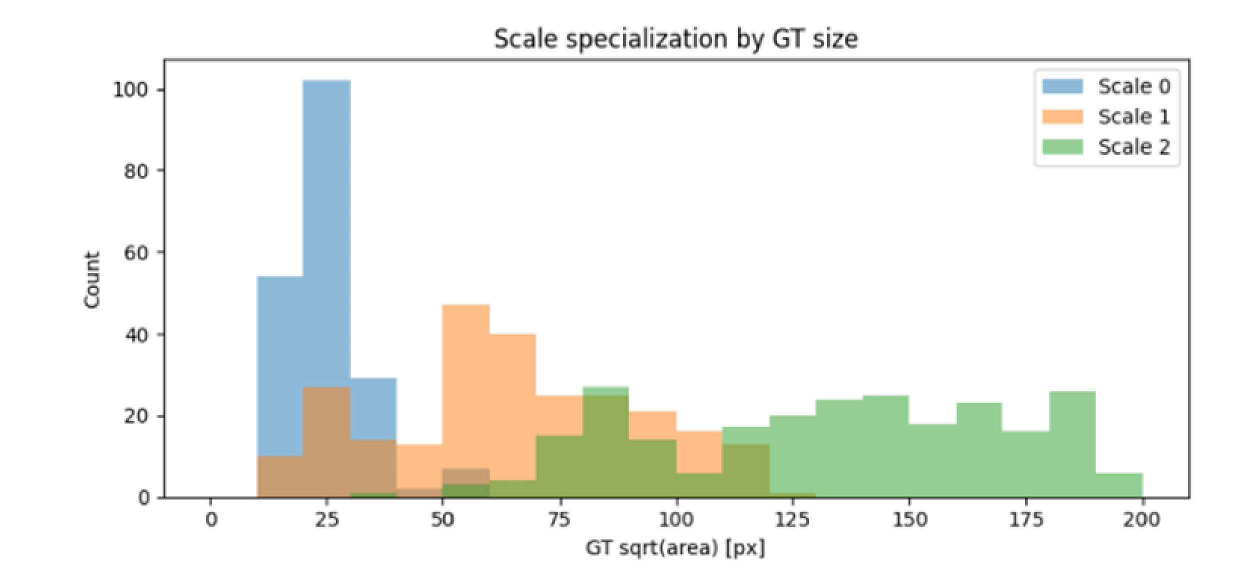
# EE641 Homework 1

Vivin Thiyagarajan

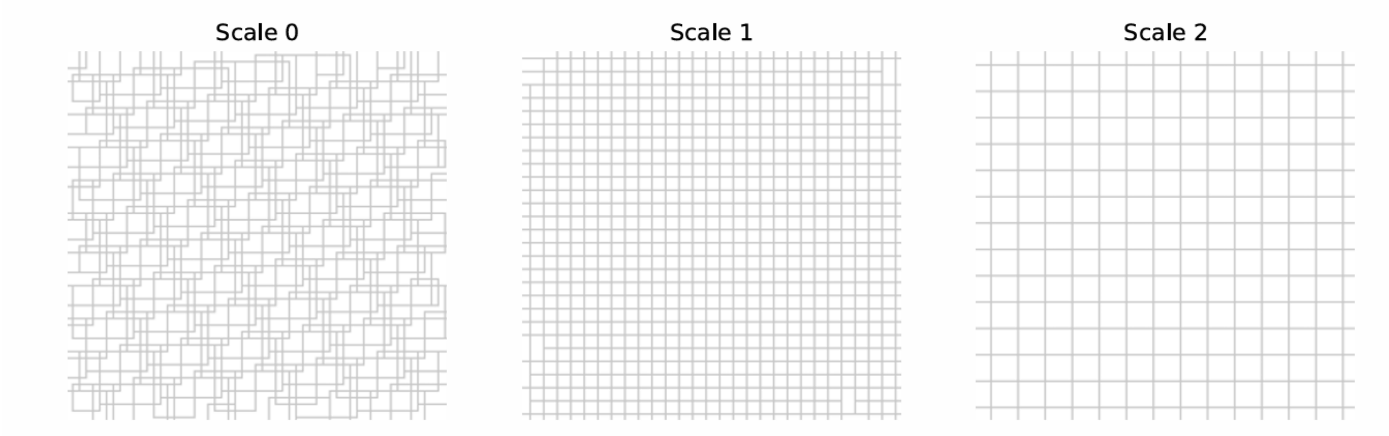
**Problem 1 Analysis:**

a) **How different scales specialize for different object sizes?**



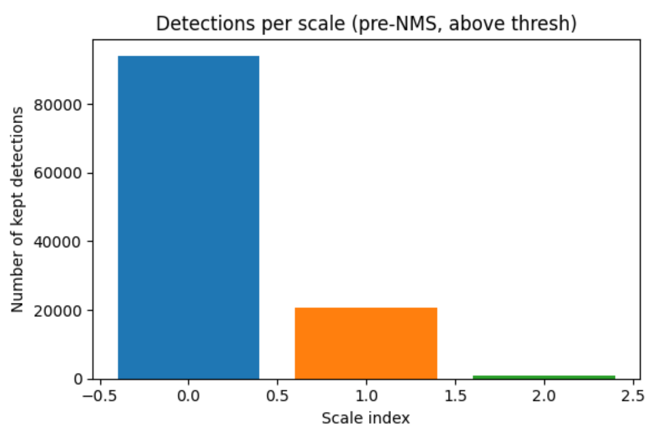

- The **first scale (blue)** mainly picks up very small objects, mostly below 40 px, and doesn't really contribute much once objects get bigger.
- The **second scale (orange)** takes over in the middle range, around 50–100 px, where it has the most detections.
- The **third scale (green)** is clearly focused on large objects, above 100 px, and becomes the dominant contributor there.
- Together, the three scales divide the work: one for small, one for medium, and one for large objects.

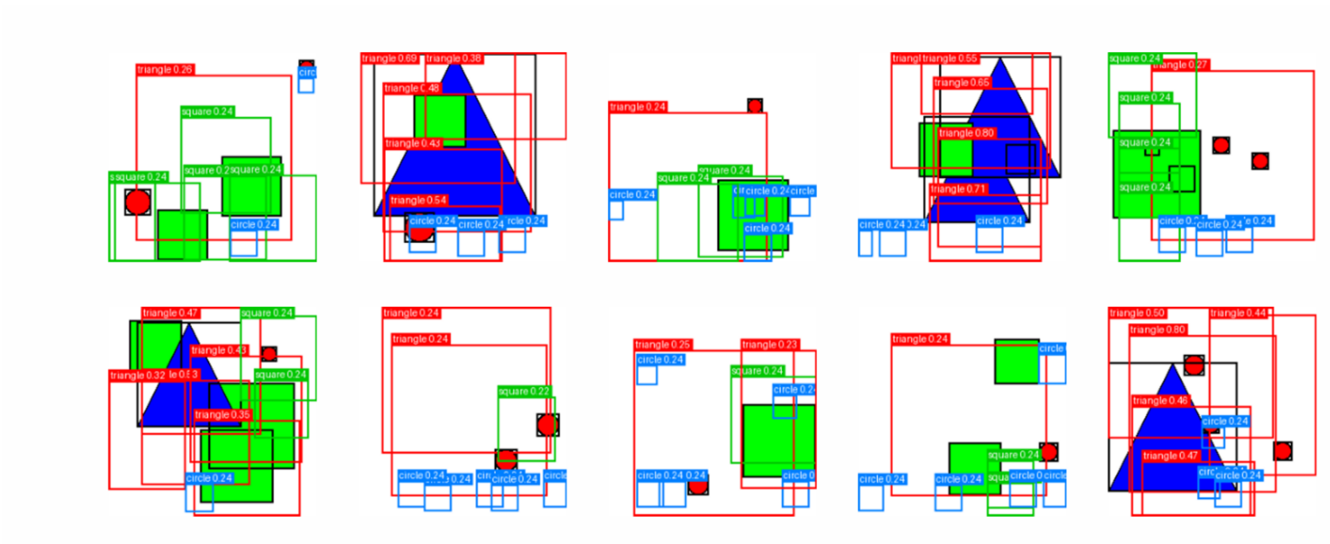
## b) The effect of anchor scales on detection performance




- Matching anchor scales to object sizes improves recall and makes box regression easier.
- Mismatched anchors force large adjustments, leading to poor localization and missed detections.
- Multi-scale anchors [16,24,32] -> [48,64,96] -> [96,128,192] ensure coverage across small, medium, and large objects.
- Fine scales give dense coverage for small objects, while coarse scales use larger, sparser anchors for big ones.

## c) Visualization of the learned features at each scale

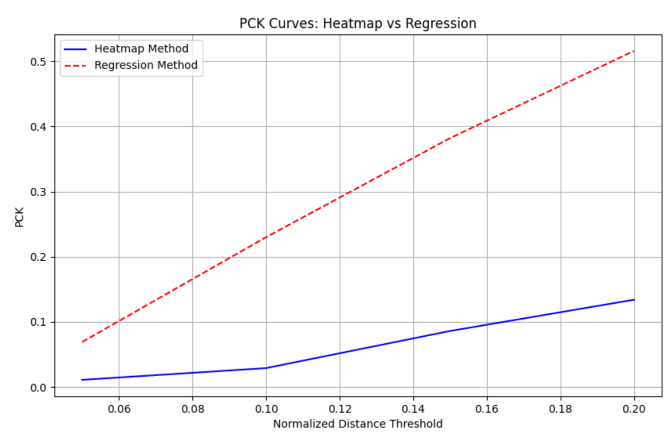




- The chart shows tons of detections at scale 0.5 (around 8000), dropping to 2000 at 1.0, with almost none after.
- The second image has bounding boxes with confidence scores, spotting shapes like triangles and squares across scales.
- Features pop up more at smaller scales, with clear shape variety in the boxes.
- Confidence scores (e.g., 0.71 for a triangle) show the model's pretty solid at picking out features.

**Problem 2 Analysis:**

**a) PCK Curves**




- Heatmap method starts low (0.011, 0.029, 0.086, 0.134) and rises steadily.
- Regression method outperforms, jumping from 0.069 to 0.516, nearly doubling at each step.
- The chart reflects this trend, with regression (red) climbing faster than heatmap (blue).

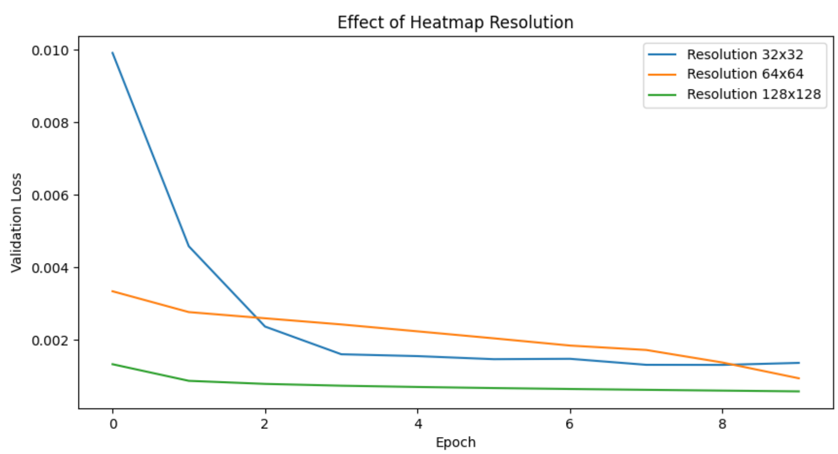
- Regression seems more accurate at higher thresholds, hitting over 0.5 at 0.2.

**b) Analysis of why heatmap approach underperforms**

- The simple dataset with sparse key points does not suit heatmap strengths.
- Down sampling to 32×32 or 64×64 reduces accuracy for locating points.
- Limited data leads to blurry, off-target heatmap predictions.
- Regression works better by directly predicting coordinates, avoiding heatmap errors.

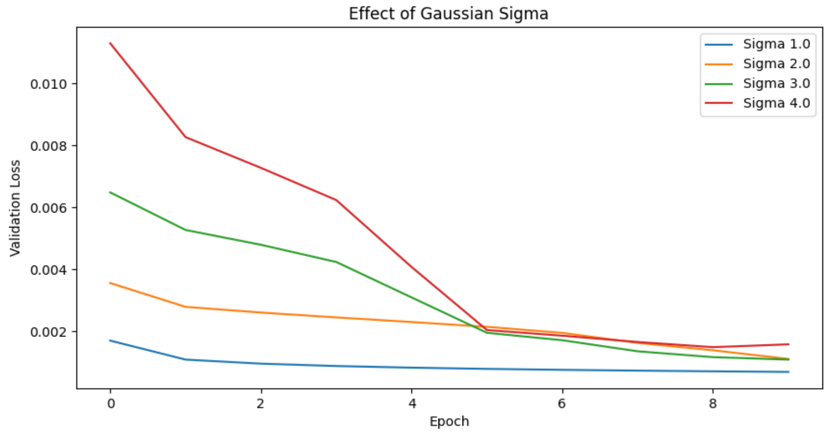
**c) Ablation Study**

**1. Effect of heatmap resolution**

- 


- Higher resolutions (32x32 to 128x128) reduce validation loss, with 128x128 achieving the best at 0.000575 (epoch 9).
- Increased resolution improves keypoint localization accuracy by minimizing discretization errors.
- The 128x128 resolution offers a 4x improvement over 32x32 (0.009912 at epoch 0 vs. 0.000575) due to better spatial detail preservation.

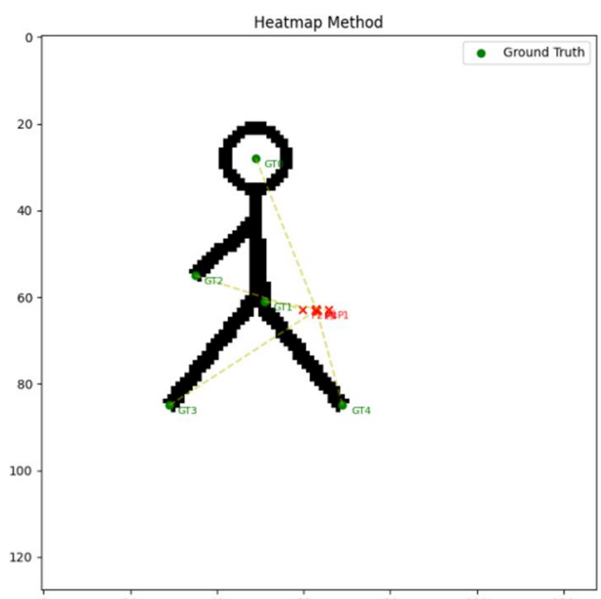
## 2. Effect of sigma resolution



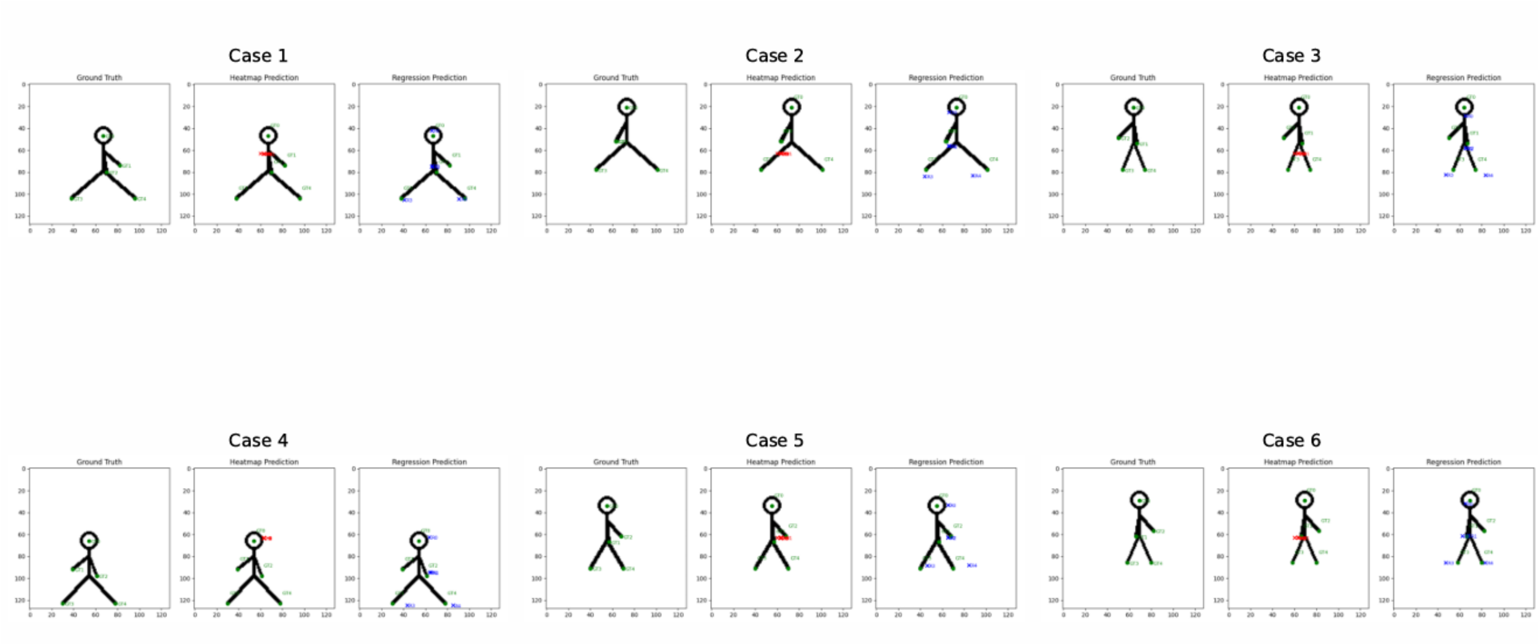

- Sigma 1.0 yields the lowest validation loss at 0.000688 (epoch 9), creating sharp targets.
- Sigma 2.0-3.0 balances precision and gradient stability, with losses around 0.001108 and 0.001086, respectively.
- Sigma 4.0 results in the highest loss at 0.001578 (epoch 9), reducing precision with blurry targets.
- Sigma 1.0 is optimal with ample training data and high-resolution heatmaps.

## d) Visualization of heatmaps and failure cases

### -> Learned Heatmaps: Sample Predictions:



**-> Failure Cases Analysis:**




- Across all six cases, the heatmap predictions often show diffuse or misplaced keypoints (e.g., Case 4 GT0, Case 5 GT3), struggling to pinpoint exact locations compared to ground truth.
- Regression predictions generally align closer to ground truth (e.g., Case 1 GT1, Case 3 GT4) but still deviate in challenging poses or when keypoints are close together (e.g., Case 6 GT1, GT4).
- Both methods consistently falter with occluded or overlapping keypoints (e.g., Case 5 GT2, Case 6 GT3), resulting in scattered or shifted predictions.