CSE 598: Action and Perception

Homework 3

Yu-I Chang, yuic, 55241226

October 6, 2025

1 Predict ImageNet-100 labels using pretrained ResNet-50 CLIP.

1.1 Precision/Recall per Class with Histogram Overlay

First, the confusion matrix is computed using the true and predicted class labels. Then, all classes are ranked in descending order based on the number of predictions per class, treated as the ranked class index. These counts are visualized as a gray histogram to represent the distribution of predicted samples across classes. Precision (blue) and Recall (orange) for each ranked class are plotted on top of the histogram, as shown in Figure 1.

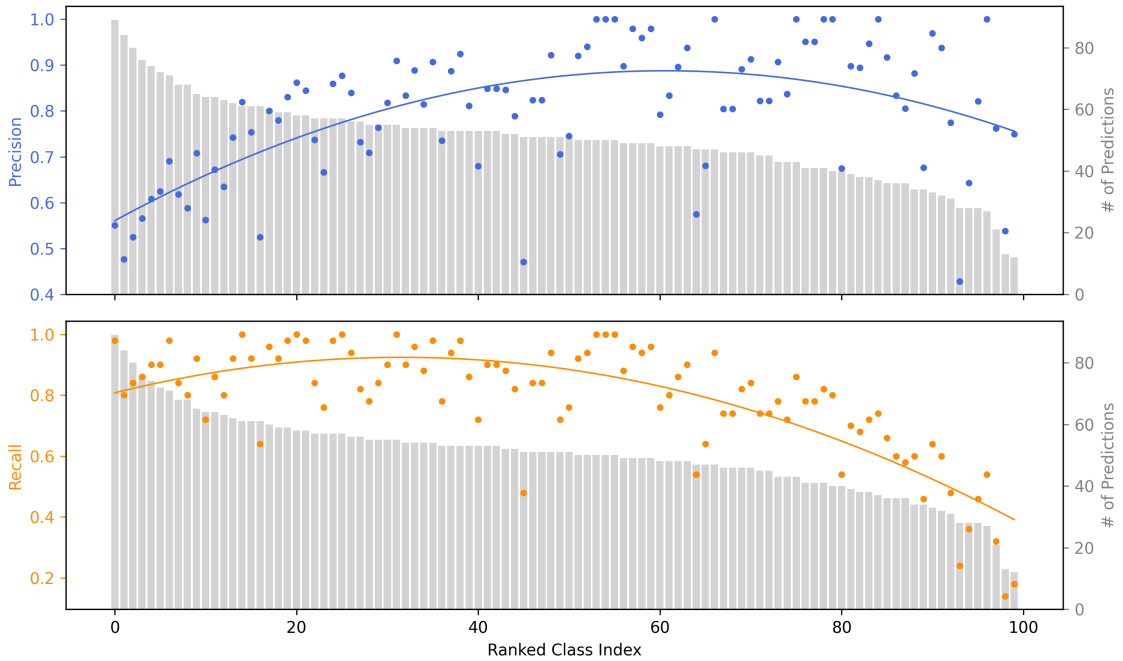

Figure 1: Precision/Recall per Class with Histogram Overlay

1.2 Confusion Matrix

The `confusion_matrix` function from the `scikit-learn` package is used to compute the confusion matrix, where the entry at the i-th row and j-th column represents the number of samples with true label i and predicted label j. After computing the matrix, all classes are sorted in descending order based on the total number of predictions per class. Therefore, in the reordered matrix, shown in Figure 2, the first column corresponds to the most frequently predicted class, while the last column corresponds to the least predicted class.

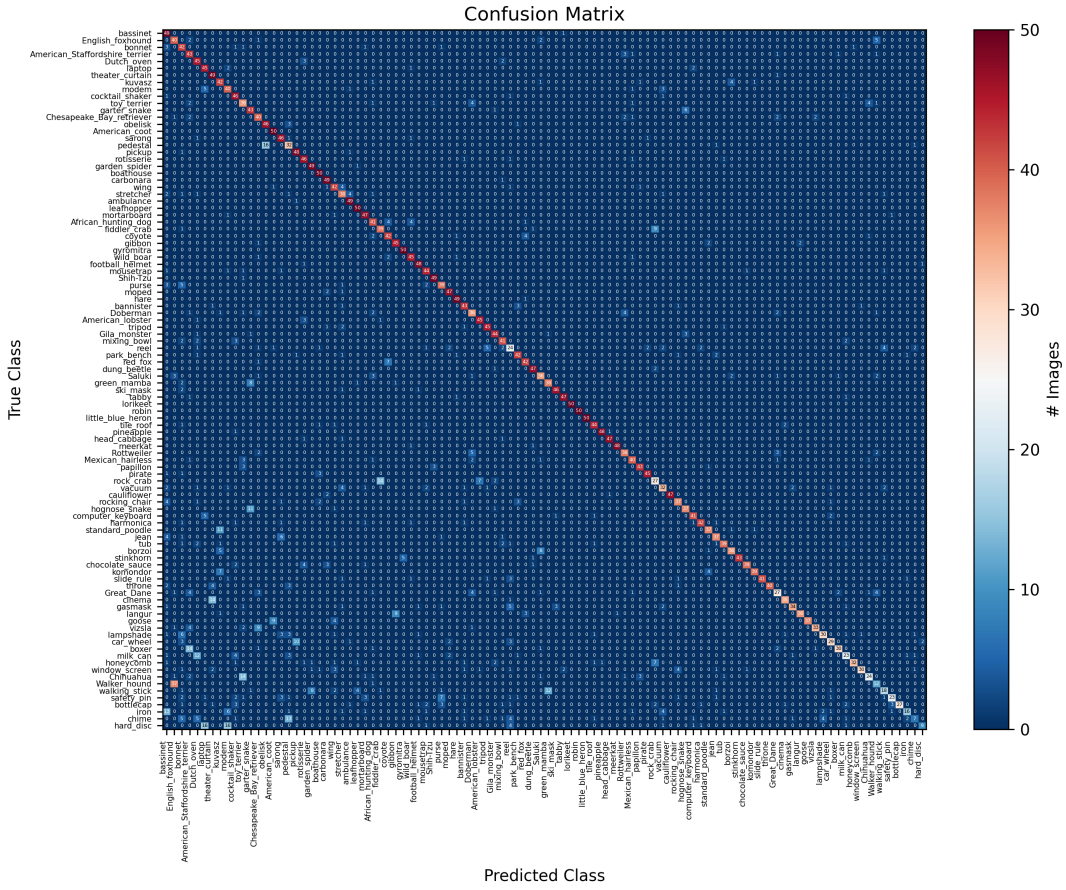

Figure 2: Confusion Matrix

1.3 Logit Centroid and Similarity Analysis

Most predicted class

- Name: **bassinet** (index 46)
- Top-10 Most Similar Classes and Corresponding Cosine Similarity Vector:

Rank	Class	Cosine Similarity
1	bonnet	0.9956
2	tub	0.9956
3	rocking_chair	0.9946
4	throne	0.9937
5	pedestal	0.9927
6	jean	0.9927
7	lampshade	0.9927
8	bannister	0.9927
9	purse	0.9927
10	milk_can	0.9927

Least predicted class

- Name: **hard_disc** (index 58)
- Top-10 Most Similar Classes and Corresponding Cosine Similarity Vector:

Rank	Class	Cosine Similarity
1	modem	0.9961
2	car_wheel	0.9951
3	reel	0.9941
4	laptop	0.9941
5	vacuum	0.9937
6	iron	0.9927
7	computer_keyboard	0.9922
8	bottlecap	0.9917
9	mousetrap	0.9917
10	slide_rule	0.9902

2 Evaluate debiased fine-tuned ResNet-50 CLIP (zero-shot) on ImageNet-100.

2.1 Precision/Recall per Class with Histogram Overlay

To keep the ranked class index order consistent with Question 1, the ranked indices were stored as a NumPy array. The same ranking was later loaded to plot Figure 3, following the procedure described in Section 1.1.

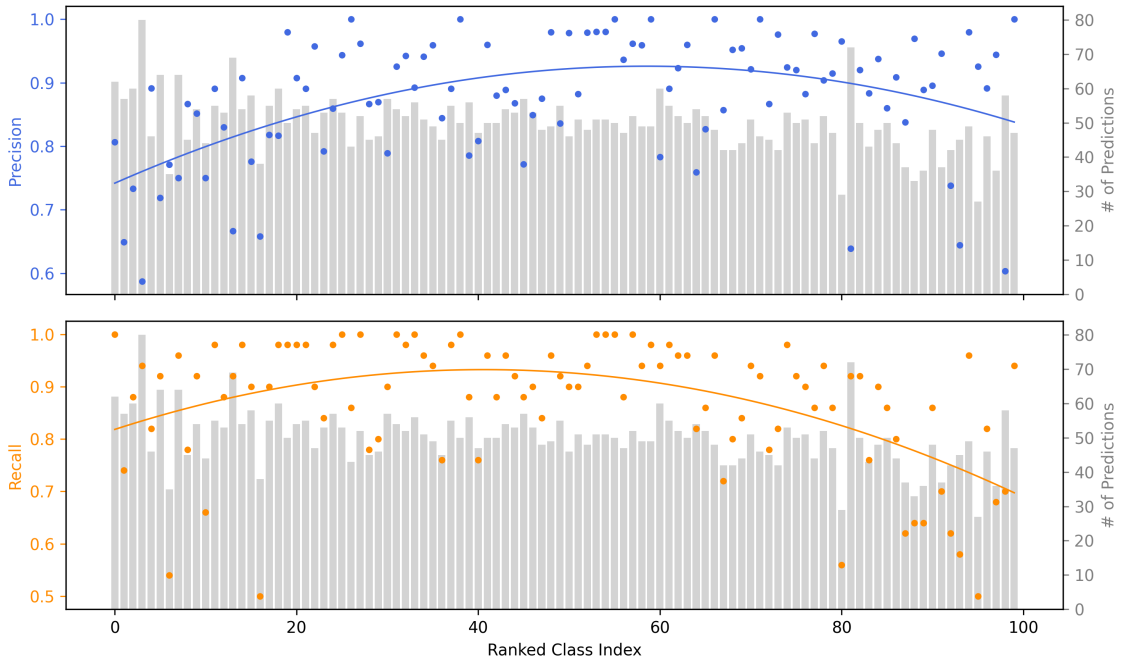

Figure 3: Precision/Recall per Class with Histogram Overlay

2.2 Confusion Matrix

Using the stored ranked class index, the confusion matrix in Figure 4 is plotted by following the same procedure described in Section 1.2.

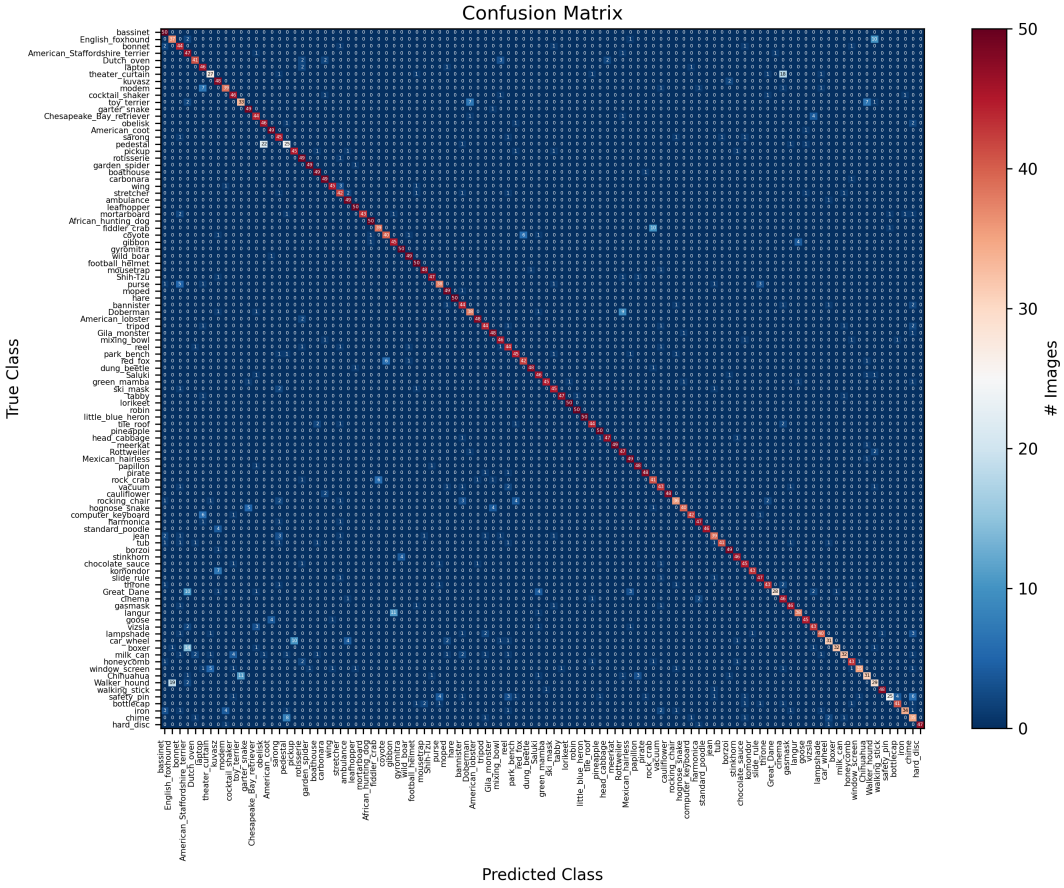

Figure 4: Confusion Matrix

2.3 Logit Centroid and Similarity Analysis

Most predicted class (CLIP)

- Name: **bassinet** (index 46)
- Top-10 Most Similar Classes and Corresponding Cosine Similarity Vector:

Rank	Class	Cosine Similarity
1	rocking_chair	0.6814
2	bonnet	0.6510
3	purse	0.5437
4	tub	0.5427
5	vacuum	0.5084
6	throne	0.4853
7	iron	0.4715
8	tabby	0.4608
9	bannister	0.4535
10	sarong	0.4402

Least predicted class (CLIP)

- Name: **hard_disc** (index 58)
- Top-10 Most Similar Classes and Corresponding Cosine Similarity Vector:

Rank	Class	Cosine Similarity
1	modem	0.7375
2	laptop	0.6984
3	slide_rule	0.5543
4	computer_keyboard	0.5381
5	mousetrap	0.4844
6	iron	0.4778
7	harmonica	0.4619
8	vacuum	0.4332
9	safety_pin	0.3920
10	purse	0.3909

Most predicted class (ZSL)

- Name: **American_Staffordshire_terrier** (index 21)
- Top-10 Most Similar Classes and Corresponding Cosine Similarity Vector:

Rank	Class	Cosine Similarity
1	boxer	0.8947
2	Great_Dane	0.8208
3	vizsla	0.8009
4	Walker_hound	0.7326
5	Doberman	0.7141
6	English_foxhound	0.7064
7	Mexican_hairless	0.6932
8	Chesapeake_Bay_retriever	0.6710
9	Rottweiler	0.6391
10	Saluki	0.5263

Least predicted class (ZSL)

- Name: **safety_pin** (index 80)
- Top-10 Most Similar Classes and Corresponding Cosine Similarity Vector:

Rank	Class	Cosine Similarity
1	purse	0.6270
2	mousetrap	0.6106
3	bottlecap	0.5512
4	cocktail_shaker	0.4920
5	harmonica	0.4796
6	reel	0.4744
7	slide_rule	0.4530
8	jean	0.4363
9	iron	0.4216
10	chime	0.4105