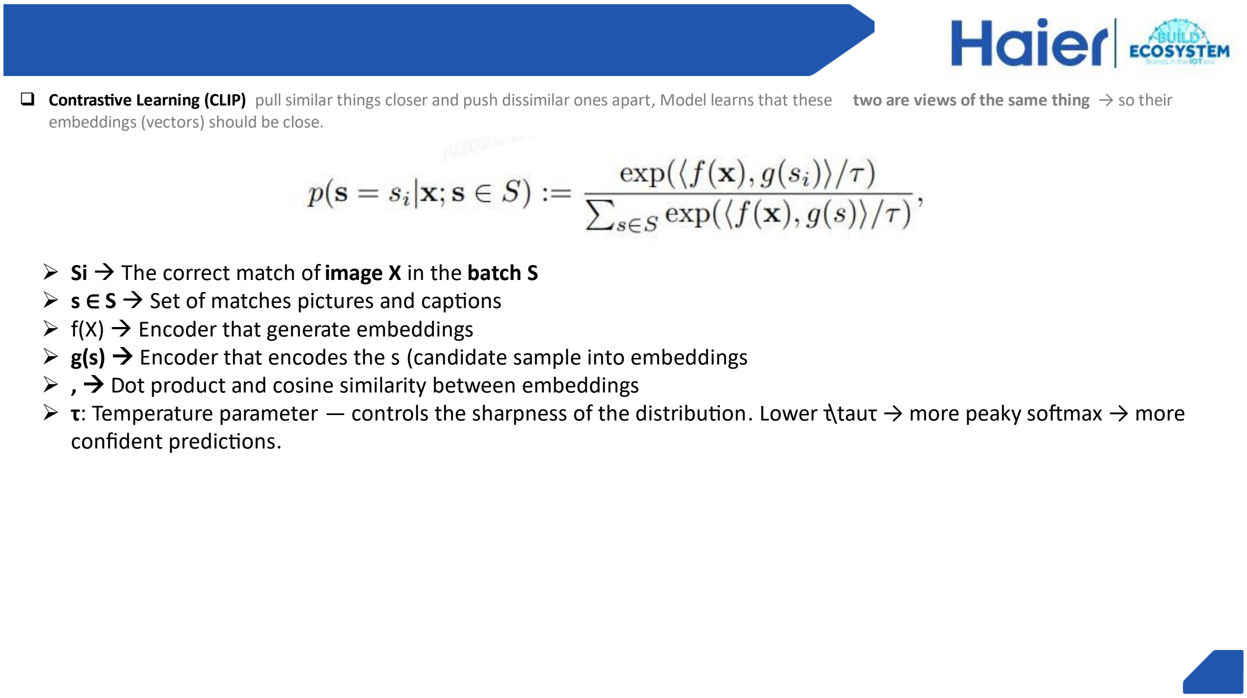
Deep Learning Assignment 02

Firstly, I didn't include the evaluation metrics and the graphs during the assignment. Then I trained again to include the metrics in it like precision, accuracy, f1-score etc.

How contrastive Learning Works:



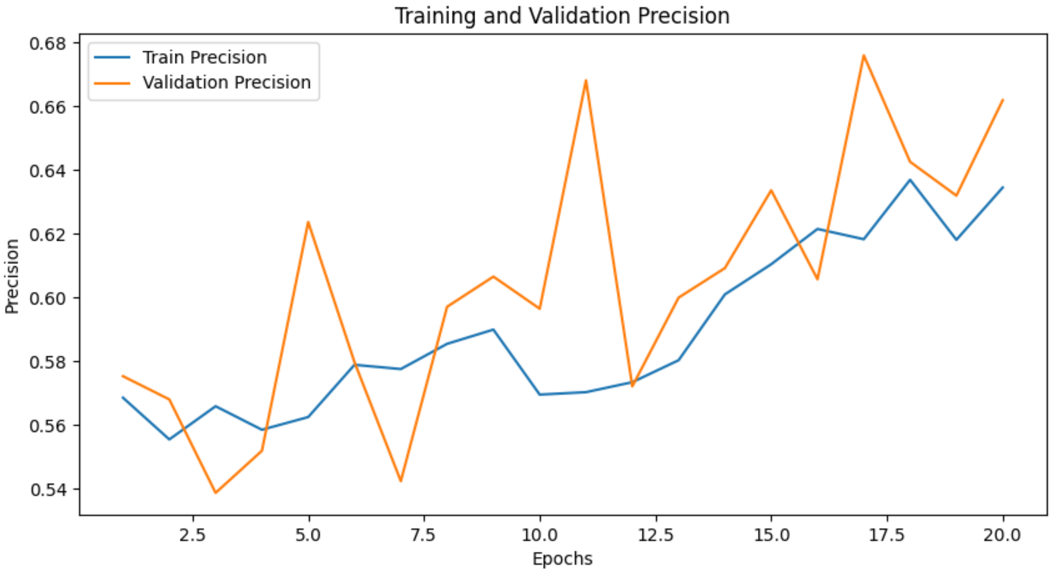
❑ **Contrastive Learning (CLIP)** pull similar things closer and push dissimilar ones apart, Model learns that these **two are views of the same thing** → so their embeddings (vectors) should be close.

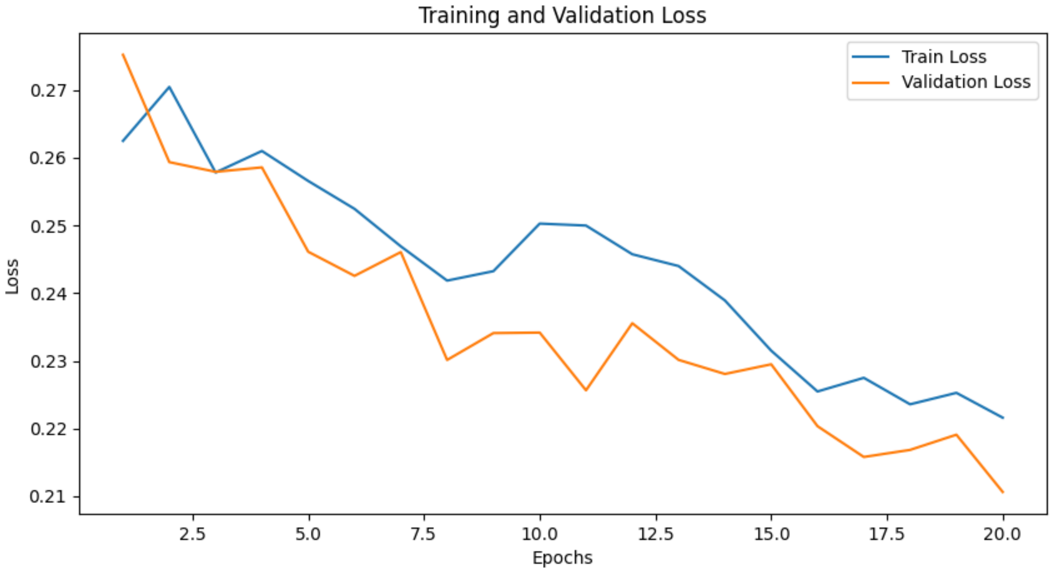
$$p(\mathbf{s}=s_i|\mathbf{x};\mathbf{s}\in S) := \frac{\exp(\langle f(\mathbf{x}), g(s_i)\rangle/\tau)}{\sum_{s\in S}\exp(\langle f(\mathbf{x}), g(s)\rangle/\tau)},$$

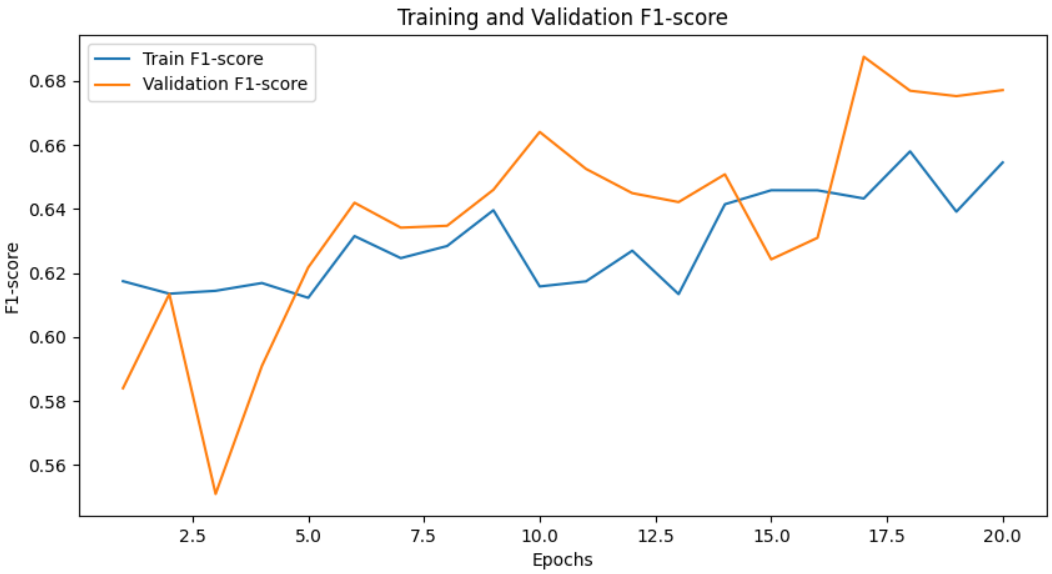
- **Si** → The correct match of **image X** in the **batch S**
- **s ∈ S** → Set of matches pictures and captions
- f(X) → Encoder that generate embeddings
- **g(s)** → Encoder that encodes the s (candidate sample into embeddings
- **,** → Dot product and cosine similarity between embeddings
- **τ**: Temperature parameter — controls the sharpness of the distribution. Lower τ\tauτ → more peaky softmax → more confident predictions.

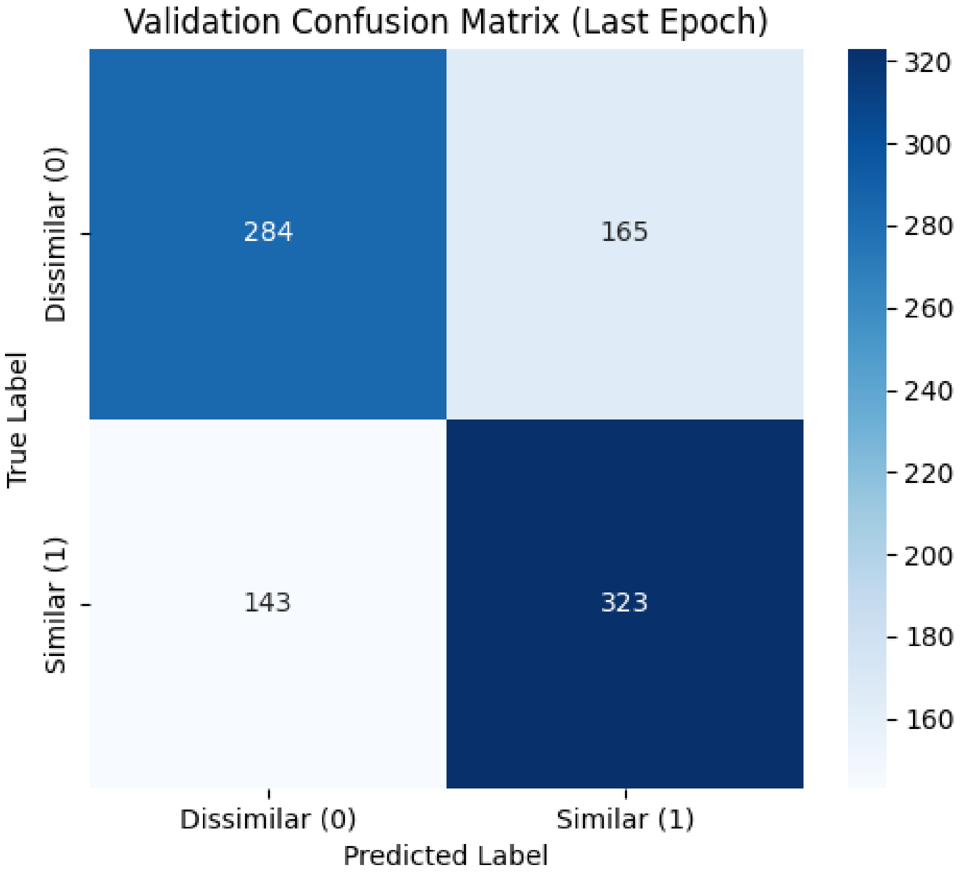
Below are the graphs:

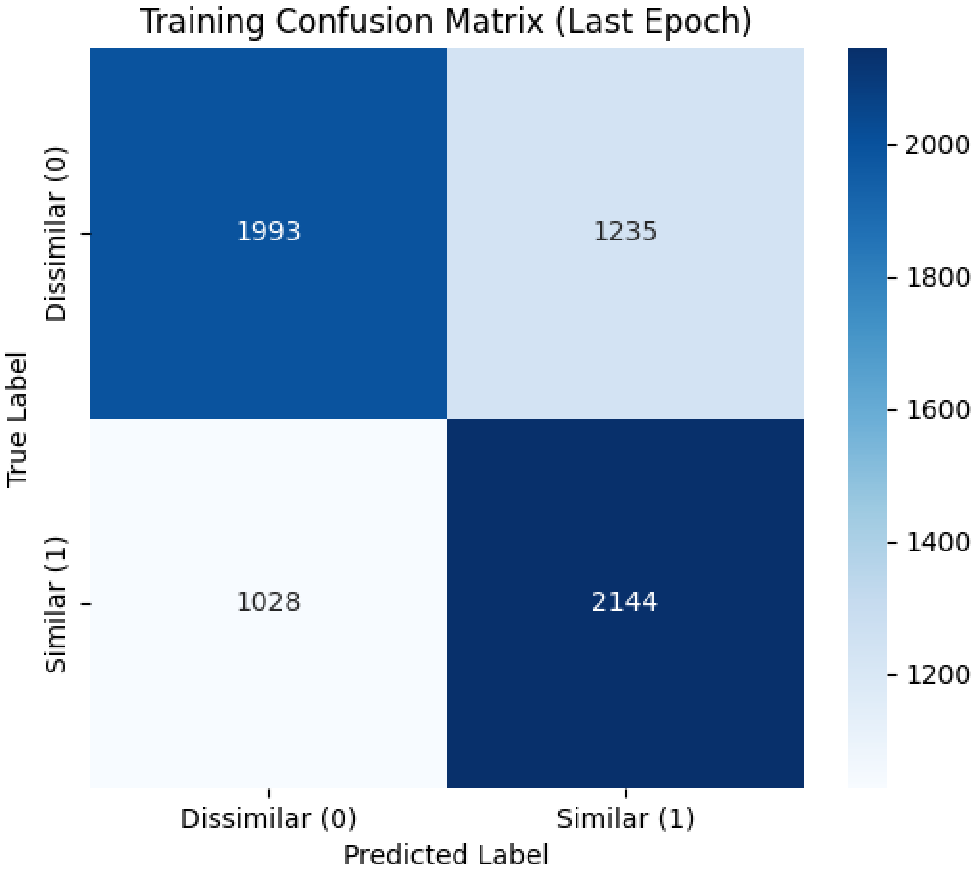
Training and Validation Graphs:

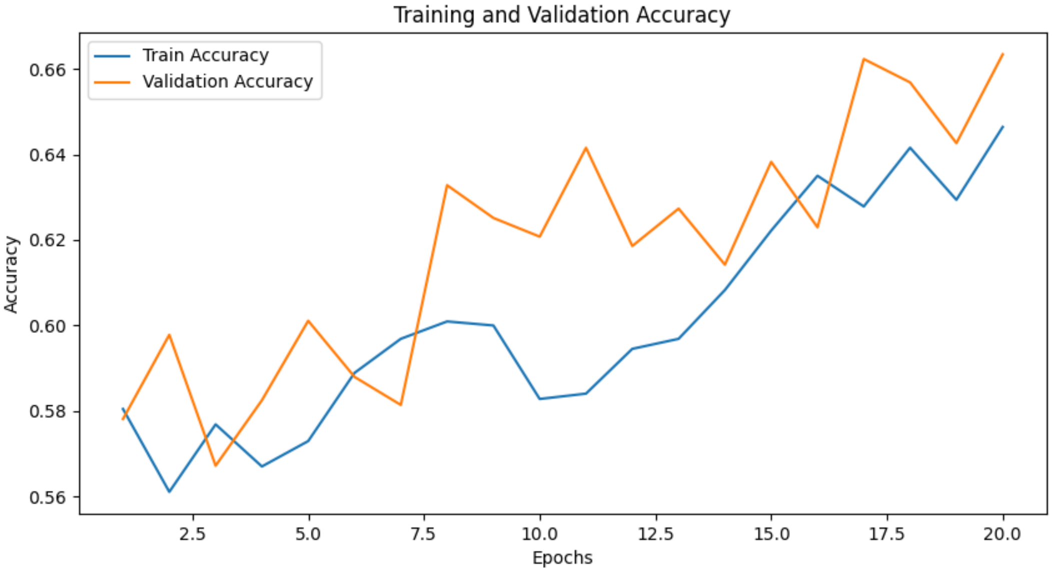


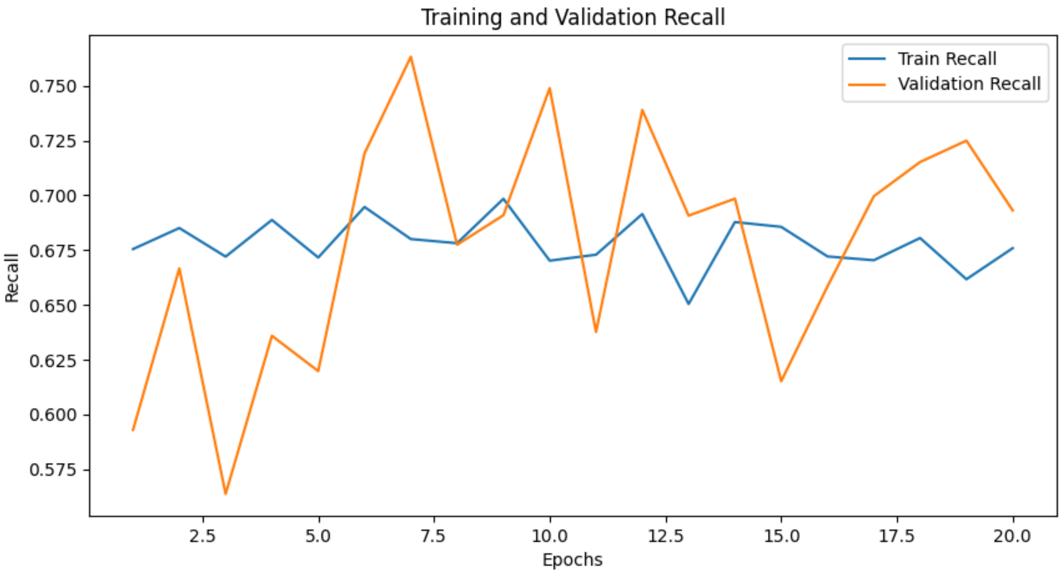



Training and Validation F1-score
Train F1-score
Validation F1-score
0.68
0.66
0.64
0.62
0.60
0.58
0.56
F1-score
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Epochs

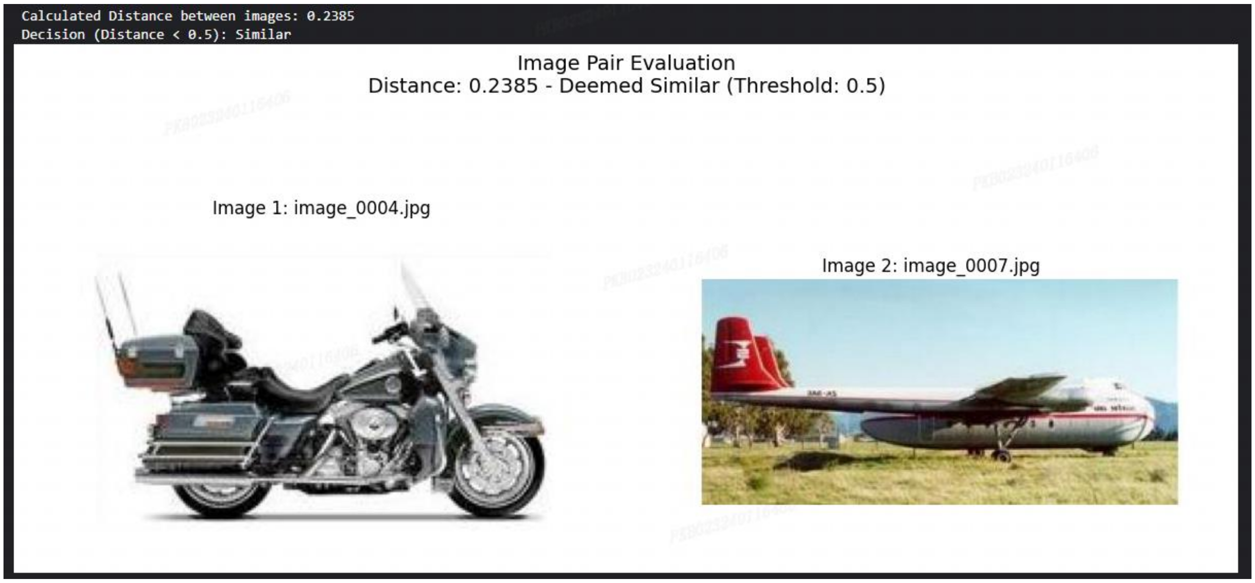

Validation Confusion Matrix (Last Epoch)
True Label
Dissimilar (0)
Similar (1)
284
165
143
323
Dissimilar (0)
Similar (1)
Predicted Label
320
300
280
260
240
220
200
180
160


Training Confusion Matrix (Last Epoch)
True Label
Dissimilar (0)
Similar (1)
1993
1235
1028
2144
Dissimilar (0)
Similar (1)
Predicted Label
2000
1800
1600
1400
1200


Training and Validation Accuracy
Train Accuracy
Validation Accuracy
Accuracy
0.66
0.64
0.62
0.60
0.58
0.56
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Epochs



Testing the model with the same images and the different images

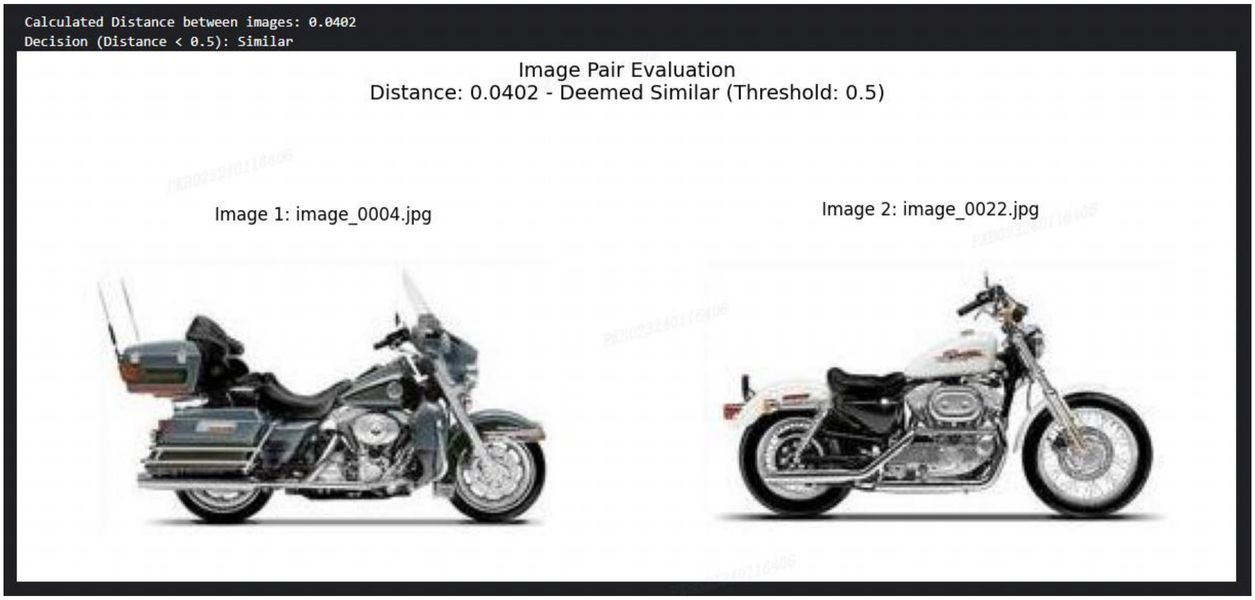


```
Calculated Distance between images: 0.2385
Decision (Distance < 0.5): Similar
```

Image Pair Evaluation
Distance: 0.2385 - Deemed Similar (Threshold: 0.5)

Image 1: image_0004.jpg

Image 2: image_0007.jpg



```
Calculated Distance between images: 0.0402
Decision (Distance < 0.5): Similar
```

Image Pair Evaluation
Distance: 0.0402 - Deemed Similar (Threshold: 0.5)

Image 1: image_0004.jpg

Image 2: image_0022.jpg