

## === GPU Information ===

NVIDIA A100-SXM4-40GB, 40960 MiB, 580.95.05

## === Installing/upgrading dependencies in venv ===

Requirement already satisfied: pip in ./venv/lib/python3.11/site-packages (25.2)

Looking in indexes: https://download.pytorch.org/whl/cu121

Requirement already satisfied: torch==2.5.1 in ./venv/lib/python3.11/site-packages (2.5.1+cu121)

Requirement already satisfied: torchvision==0.20.1 in ./venv/lib/python3.11/site-packages (0.20.1+cu121)

Requirement already satisfied: filelock in ./venv/lib/python3.11/site-packages (from torch==2.5.1) (3.19.1)

Requirement already satisfied: typing-extensions&gt;=4.8.0 in ./venv/lib/python3.11/site-packages (from torch==2.5.1) (4.15.0)

Requirement already satisfied: networkx in ./venv/lib/python3.11/site-packages (from torch==2.5.1) (3.5)

Requirement already satisfied: jinja2 in ./venv/lib/python3.11/site-packages (from torch==2.5.1) (3.1.6)

Requirement already satisfied: fsspec in ./venv/lib/python3.11/site-packages (from torch==2.5.1) (2025.9.0)

Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.1.105 in ./venv/lib/python3.11/site-packages (from torch==2.5.1) (12.1.105)

Requirement already satisfied: nvidia-cuda-runtime-cu12==12.1.105 in ./venv/lib/python3.11/site-packages (from torch==2.5.1) (12.1.105)

Requirement already satisfied: nvidia-cuda-cupti-cu12==12.1.105 in ./venv/lib/python3.11/site-packages (from torch==2.5.1) (12.1.105)

Requirement already satisfied: nvidia-cudnn-cu12==9.1.0.70 in ./venv/lib/python3.11/site-packages (from torch==2.5.1) (9.1.0.70)

Requirement already satisfied: nvidia-cublas-cu12==12.1.3.1 in ./venv/lib/python3.11/site-packages (from torch==2.5.1) (12.1.3.1)

Requirement already satisfied: nvidia-cufft-cu12==11.0.2.54 in ./venv/lib/python3.11/site-packages (from torch==2.5.1) (11.0.2.54)

Requirement already satisfied: nvidia-curand-cu12==10.3.2.106 in ./venv/lib/python3.11/site-packages (from torch==2.5.1) (10.3.2.106)

Requirement already satisfied: nvidia-cusolver-cu12==11.4.5.107 in ./venv/lib/python3.11/site-packages (from torch==2.5.1) (11.4.5.107)

Requirement already satisfied: nvidia-cusparse-cu12==12.1.0.106 in ./venv/lib/python3.11/site-packages (from torch==2.5.1) (12.1.0.106)

Requirement already satisfied: nvidia-nccl-cu12==2.21.5 in ./venv/lib/python3.11/site-packages (from torch==2.5.1) (2.21.5)

Requirement already satisfied: nvidia-nvtx-cu12==12.1.105 in ./venv/lib/python3.11/site-packages (from torch==2.5.1) (12.1.105)

Requirement already satisfied: triton==3.1.0 in ./venv/lib/python3.11/site-packages (from torch==2.5.1) (3.1.0)

Requirement already satisfied: sympy==1.13.1 in ./venv/lib/python3.11/site-packages (from torch==2.5.1) (1.13.1)

Requirement already satisfied: numpy in ./venv/lib/python3.11/site-packages (from torchvision==0.20.1) (1.26.4)

Requirement already satisfied: pillow!=8.3.\*,&gt;=5.3.0 in ./venv/lib/python3.11/site-packages (from torchvision==0.20.1) (11.3.0)

Requirement already satisfied: nvidia-nvjitlink-cu12 in ./venv/lib/python3.11/site-packages (from nvidia-cusolver-cu12==11.4.5.107-&gt;torch==2.5.1) (12.9.86)

Requirement already satisfied: mpmath&lt;1.4,&gt;=1.1.0 in ./venv/lib/python3.11/site-packages (from sympy==1.13.1-&gt;torch==2.5.1) (1.3.0)

Requirement already satisfied: MarkupSafe&gt;=2.0 in ./venv/lib/python3.11/site-packages (from jinja2-&gt;torch==2.5.1) (2.1.5)

Requirement already satisfied: transformers in ./venv/lib/python3.11/site-packages (4.57.1)

Requirement already satisfied: accelerate in ./venv/lib/python3.11/site-packages (1.11.0)

Requirement already satisfied: pillow in ./venv/lib/python3.11/site-packages (11.3.0)

Requirement already satisfied: tqdm in ./venv/lib/python3.11/site-packages (4.67.1)

Requirement already satisfied: scikit-learn in ./venv/lib/python3.11/site-packages (1.7.2)

Requirement already satisfied: numpy&lt;2 in ./venv/lib/python3.11/site-packages (1.26.4)

Requirement already satisfied: qwen-vl-utils in ./venv/lib/python3.11/site-packages (0.0.14)

Requirement already satisfied: filelock in ./venv/lib/python3.11/site-packages (from transformers) (3.19.1)

Requirement already satisfied: huggingface-hub&lt;1.0,&gt;=0.34.0 in ./venv/lib/python3.11/site-packages (from transformers) (0.36.0)

Requirement already satisfied: packaging&gt;=20.0 in ./venv/lib/python3.11/site-packages (from transformers) (25.0)

Requirement already satisfied: pyyaml&gt;=5.1 in ./venv/lib/python3.11/site-packages (from transformers) (6.0.3)

Requirement already satisfied: regex!=2019.12.17 in ./venv/lib/python3.11/site-packages (from transformers) (2025.10.23)

Requirement already satisfied: requests in ./venv/lib/python3.11/site-packages (from transformers) (2.32.5)

Requirement already satisfied: tokenizers<=0.23.0,>=0.22.0 in ./venv/lib/python3.11/site-packages (from transformers) (0.22.1)

Requirement already satisfied: safetensors>=0.4.3 in ./venv/lib/python3.11/site-packages (from transformers) (0.6.2)

Requirement already satisfied: fsspec>=2023.5.0 in ./venv/lib/python3.11/site-packages (from huggingface-hub<1.0,>=0.34.0->transformers) (2025.9.0)

Requirement already satisfied: typing-extensions>=3.7.4.3 in ./venv/lib/python3.11/site-packages (from huggingface-hub<1.0,>=0.34.0->transformers) (4.15.0)

Requirement already satisfied: hf-xet<2.0.0,>=1.1.3 in ./venv/lib/python3.11/site-packages (from huggingface-hub<1.0,>=0.34.0->transformers) (1.1.10)

Requirement already satisfied: psutil in ./venv/lib/python3.11/site-packages (from accelerate) (7.1.1)

Requirement already satisfied: torch>=2.0.0 in ./venv/lib/python3.11/site-packages (from accelerate) (2.5.1+cu121)

Requirement already satisfied: scipy>=1.8.0 in ./venv/lib/python3.11/site-packages (from scikit-learn) (1.16.2)

Requirement already satisfied: joblib>=1.2.0 in ./venv/lib/python3.11/site-packages (from scikit-learn) (1.5.2)

Requirement already satisfied: threadpoolctl>=3.1.0 in ./venv/lib/python3.11/site-packages (from scikit-learn) (3.6.0)

Requirement already satisfied: av in ./venv/lib/python3.11/site-packages (from qwen-vl-utils) (16.0.1)

Requirement already satisfied: networkx in ./venv/lib/python3.11/site-packages (from torch>=2.0.0->accelerate) (3.5)

Requirement already satisfied: jinja2 in ./venv/lib/python3.11/site-packages (from torch>=2.0.0->accelerate) (3.1.6)

Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.1.105 in ./venv/lib/python3.11/site-packages (from torch>=2.0.0->accelerate) (12.1.105)

Requirement already satisfied: nvidia-cuda-runtime-cu12==12.1.105 in ./venv/lib/python3.11/site-packages (from torch>=2.0.0->accelerate) (12.1.105)

Requirement already satisfied: nvidia-cuda-cupti-cu12==12.1.105 in ./venv/lib/python3.11/site-packages (from torch>=2.0.0->accelerate) (12.1.105)

Requirement already satisfied: nvidia-cudnn-cu12==9.1.0.70 in ./venv/lib/python3.11/site-packages (from torch>=2.0.0->accelerate) (9.1.0.70)

Requirement already satisfied: nvidia-cublas-cu12==12.1.3.1 in ./venv/lib/python3.11/site-packages (from torch>=2.0.0->accelerate) (12.1.3.1)

Requirement already satisfied: nvidia-cufft-cu12==11.0.2.54 in ./venv/lib/python3.11/site-packages (from torch>=2.0.0->accelerate) (11.0.2.54)

Requirement already satisfied: nvidia-curand-cu12==10.3.2.106 in ./venv/lib/python3.11/site-packages (from torch>=2.0.0->accelerate) (10.3.2.106)

Requirement already satisfied: nvidia-cusolver-cu12==11.4.5.107 in ./venv/lib/python3.11/site-packages (from torch>=2.0.0->accelerate) (11.4.5.107)

Requirement already satisfied: nvidia-cuspars-cu12==12.1.0.106 in ./venv/lib/python3.11/site-packages (from torch>=2.0.0->accelerate) (12.1.0.106)

Requirement already satisfied: nvidia-nccl-cu12==2.21.5 in ./venv/lib/python3.11/site-packages (from torch>=2.0.0->accelerate) (2.21.5)

Requirement already satisfied: nvidia-nvtx-cu12==12.1.105 in ./venv/lib/python3.11/site-packages (from torch>=2.0.0->accelerate) (12.1.105)

Requirement already satisfied: triton==3.1.0 in ./venv/lib/python3.11/site-packages (from torch>=2.0.0->accelerate) (3.1.0)

Requirement already satisfied: sympy==1.13.1 in ./venv/lib/python3.11/site-packages (from torch>=2.0.0->accelerate) (1.13.1)

Requirement already satisfied: nvidia-nvjitlink-cu12 in ./venv/lib/python3.11/site-packages (from nvidia-cusolver-cu12==11.4.5.107->torch>=2.0.0->accelerate) (12.9.86)

Requirement already satisfied: mpmath<1.4,>=1.1.0 in ./venv/lib/python3.11/site-packages (from sympy==1.13.1->torch>=2.0.0->accelerate) (1.3.0)

Requirement already satisfied: MarkupSafe>=2.0 in ./venv/lib/python3.11/site-packages (from jinja2->torch>=2.0.0->accelerate) (2.1.5)

Requirement already satisfied: charset\_normalizer<4,>=2 in ./venv/lib/python3.11/site-packages (from requests->transformers) (3.4.4)

Requirement already satisfied: idna<4,>=2.5 in ./venv/lib/python3.11/site-packages (from requests->transformers) (3.11)

Requirement already satisfied: urllib3<3,>=1.21.1 in ./venv/lib/python3.11/site-packages (from requests->transformers) (2.5.0)

Requirement already satisfied: certifi>=2017.4.17 in ./venv/lib/python3.11/site-packages (from requests->transformers) (2025.10.23)

```
requests->transformers) (2025.10.5)
=== Checking if vision_ccs_corrected.py exists ===
-rw-r----- 1 mdemirev mdemirev 20K Oct 24 02:15 vision_ccs_corrected.py
=== Running vision_ccs_corrected.py ===
`torch_dtype` is deprecated! Use `dtype` instead!
Model: llava-hf/llava-1.5-7b-hf
```

```
#####
# CATEGORY: OBJECT_DETECTION
#####
```

```
=====
EXTRACTING HIDDEN STATES: OBJECT_DETECTION
=====
⚠ Cache disabled (use_cache=False). Extracting new...
```

Processing 1323 samples in batches of 40  
Searching in 2 image directories  
LOADING MODEL: llava  
Device: cuda

```
Loading checkpoint shards: 0%|          | 0/3 [00:00<?, ?it/s]
Loading checkpoint shards: 33%|███      | 1/3 [00:06<00:12, 6.45s/it]
Loading checkpoint shards: 67%|██████   | 2/3 [00:12<00:06, 6.18s/it]
Loading checkpoint shards: 100%|████████| 3/3 [00:17<00:00, 5.59s/it]
Loading checkpoint shards: 100%|████████| 3/3 [00:17<00:00, 5.78s/it]
✓ Model loaded successfully
```

```
Batches: 0%|          | 0/34 [00:00<?, ?it/s]
Batches: 3%|█         | 1/34 [00:06<03:35, 6.54s/it]
Batches: 6%|██        | 2/34 [00:12<03:14, 6.09s/it]
Batches: 9%|███       | 3/34 [00:17<02:53, 5.58s/it]
Batches: 12%|████      | 4/34 [00:22<02:42, 5.40s/it]
Batches: 15%|█████     | 5/34 [00:27<02:32, 5.27s/it]
Batches: 18%|██████    | 6/34 [00:33<02:31, 5.39s/it]
Batches: 21%|███████   | 7/34 [00:38<02:24, 5.36s/it]
Batches: 24%|████████  | 8/34 [00:43<02:15, 5.20s/it]
Batches: 26%|█████████ | 9/34 [00:48<02:09, 5.19s/it]
Batches: 29%|█████████ | 10/34 [00:53<02:07, 5.29s/it]
Batches: 32%|█████████ | 11/34 [00:59<02:02, 5.31s/it]
Batches: 35%|█████████ | 12/34 [01:04<01:56, 5.28s/it]
Batches: 38%|█████████ | 13/34 [01:08<01:43, 4.94s/it]
Batches: 41%|█████████ | 14/34 [01:14<01:41, 5.09s/it]
Batches: 44%|█████████ | 15/34 [01:18<01:34, 4.98s/it]
Batches: 47%|█████████ | 16/34 [01:24<01:33, 5.17s/it]
Batches: 50%|█████████ | 17/34 [01:29<01:26, 5.09s/it]
Batches: 53%|█████████ | 18/34 [01:34<01:23, 5.21s/it]
Batches: 56%|█████████ | 19/34 [01:40<01:21, 5.44s/it]
Batches: 59%|█████████ | 20/34 [01:46<01:15, 5.38s/it]
Batches: 62%|█████████ | 21/34 [01:51<01:10, 5.44s/it]
Batches: 65%|█████████ | 22/34 [01:56<01:05, 5.42s/it]
Batches: 68%|█████████ | 23/34 [02:03<01:02, 5.65s/it]
Batches: 71%|█████████ | 24/34 [02:11<01:05, 6.51s/it]
Batches: 74%|█████████ | 25/34 [02:19<01:02, 6.97s/it]
Batches: 76%|█████████ | 26/34 [02:27<00:56, 7.09s/it]
Batches: 79%|█████████ | 27/34 [02:34<00:51, 7.32s/it]
Batches: 82%|█████████ | 28/34 [02:42<00:43, 7.28s/it]
Batches: 85%|█████████ | 29/34 [02:50<00:37, 7.56s/it]
Batches: 88%|█████████ | 30/34 [02:57<00:30, 7.55s/it]
Batches: 91%|█████████ | 31/34 [03:05<00:22, 7.55s/it]
```

Batches: 94% ██████████ | 32/34 [03:12<00:14, 7.35s/it]  
Batches: 97% ██████████ | 33/34 [03:19<00:07, 7.38s/it]  
Batches: 100% ██████████ | 34/34 [03:20<00:00, 5.40s/it]  
Batches: 100% ██████████ | 34/34 [03:20<00:00, 5.90s/it]

=====

✓ Successfully processed: 1140/1323  
X Skipped (missing/error): 183/1323

First 10 skipped: 000000262227.jpg, 000000262440.jpg, 000000262440.jpg, 000000262682.jpg,  
000000262682.jpg, 000000262682.jpg, 000000139684.jpg, 000000000632.jpg, 000000000632.jpg,  
000000000632.jpg...

Extracted shapes:  
Positive: (1140, 4096)  
Negative: (1140, 4096)  
Labels: (1140,)

Cached to: hidden\_states\_cache\_corrected/cache\_object\_detection\_1323\_llava.npz

=====

TRAINING CCS PROBE

=====

Dataset split:  
Train: 797 samples (403 pos, 394 neg)  
Test: 343 samples (173 pos, 170 neg)  
Hidden dim: 4096

=====

TRAINING WITH MULTIPLE RANDOM RESTARTS

=====

Trial 1/10: Loss = 0.000870  
Trial 2/10: Loss = 0.001189  
Trial 3/10: Loss = 0.000930  
Trial 4/10: Loss = 0.001145  
Trial 5/10: Loss = 0.001945  
Trial 6/10: Loss = 0.001364  
Trial 7/10: Loss = 0.001339  
Trial 8/10: Loss = 0.000959  
Trial 9/10: Loss = 0.001259  
Trial 10/10: Loss = 0.001249

=====

EVALUATION WITH BEST PROBE

=====

Best loss: 0.000870

Test Results:  
Overall Accuracy: 79.0% (271/343)  
Positive samples: 80.9% (173 samples)  
Negative samples: 77.1% (170 samples)

✓ COMPLETE: object\_detection → 79.0%

#####  
# CATEGORY: ATTRIBUTE\_RECOGNITION  
#####

=====

EXTRACTING HIDDEN STATES: ATTRIBUTE\_RECOGNITION

=====

⚠ Cache disabled (use\_cache=False). Extracting new...

Processing 3410 samples in batches of 40  
Searching in 2 image directories  
LOADING MODEL: llava  
Device: cuda

```
Loading checkpoint shards: 0%|          | 0/3 [00:00<?, ?it/s]
Loading checkpoint shards: 33%|██████    | 1/3 [00:06<00:12, 6.28s/it]
Loading checkpoint shards: 67%|██████████| 2/3 [00:12<00:06, 6.12s/it]
Loading checkpoint shards: 100%|██████████| 3/3 [00:17<00:00, 5.79s/it]
Loading checkpoint shards: 100%|██████████| 3/3 [00:17<00:00, 5.89s/it]
✓ Model loaded successfully
```

```
Batches: 0%|          | 0/86 [00:00<?, ?it/s]
Batches: 1%|          | 1/86 [00:07<10:44, 7.58s/it]
Batches: 2%||         | 2/86 [00:15<10:57, 7.83s/it]
Batches: 3%||         | 3/86 [00:23<10:41, 7.73s/it]
Batches: 5%||         | 4/86 [00:30<10:13, 7.48s/it]
Batches: 6%||         | 5/86 [00:38<10:16, 7.61s/it]
Batches: 7%|█         | 6/86 [00:45<09:48, 7.36s/it]
Batches: 8%|█         | 7/86 [00:52<09:49, 7.47s/it]
Batches: 9%|█         | 8/86 [00:58<09:03, 6.97s/it]
Batches: 10%|█         | 9/86 [01:05<08:55, 6.95s/it]
Batches: 12%|██        | 10/86 [01:12<08:44, 6.91s/it]
Batches: 13%|██        | 11/86 [01:20<09:07, 7.30s/it]
Batches: 14%|██        | 12/86 [01:27<09:00, 7.31s/it]
Batches: 15%|██        | 13/86 [01:35<08:59, 7.40s/it]
Batches: 16%|███       | 14/86 [01:46<10:19, 8.60s/it]
Batches: 17%|███       | 15/86 [02:00<11:57, 10.11s/it]
Batches: 19%|███       | 16/86 [02:08<10:59, 9.42s/it]
Batches: 20%|███       | 17/86 [02:15<10:08, 8.82s/it]
Batches: 21%|███       | 18/86 [02:22<09:17, 8.20s/it]
Batches: 22%|███       | 19/86 [02:29<08:50, 7.92s/it]
Batches: 23%|███       | 20/86 [02:37<08:41, 7.90s/it]
Batches: 24%|███       | 21/86 [02:45<08:28, 7.83s/it]
Batches: 26%|███       | 22/86 [02:53<08:21, 7.84s/it]
Batches: 27%|███       | 23/86 [03:00<08:09, 7.76s/it]
Batches: 28%|███       | 24/86 [03:07<07:51, 7.60s/it]
Batches: 29%|███       | 25/86 [03:15<07:48, 7.68s/it]
Batches: 30%|███       | 26/86 [03:23<07:41, 7.69s/it]
Batches: 31%|███       | 27/86 [03:30<07:26, 7.57s/it]
Batches: 33%|███       | 28/86 [03:37<07:09, 7.41s/it]
Batches: 34%|███       | 29/86 [03:45<07:03, 7.43s/it]
Batches: 35%|███       | 30/86 [03:51<06:44, 7.22s/it]
Batches: 36%|███       | 31/86 [03:59<06:38, 7.25s/it]
Batches: 37%|███       | 32/86 [04:07<06:44, 7.50s/it]
Batches: 38%|███       | 33/86 [04:14<06:31, 7.39s/it]
Batches: 40%|███       | 34/86 [04:21<06:10, 7.13s/it]
Batches: 41%|███       | 35/86 [04:29<06:17, 7.39s/it]
Batches: 42%|███       | 36/86 [04:36<06:16, 7.52s/it]
Batches: 43%|███       | 37/86 [04:44<06:07, 7.50s/it]
Batches: 44%|███       | 38/86 [04:51<05:58, 7.46s/it]
Batches: 45%|███       | 39/86 [04:58<05:43, 7.30s/it]
Batches: 47%|███       | 40/86 [05:05<05:36, 7.31s/it]
Batches: 48%|███       | 41/86 [05:13<05:32, 7.38s/it]
Batches: 49%|███       | 42/86 [05:20<05:18, 7.23s/it]
Batches: 50%|███       | 43/86 [05:27<05:13, 7.28s/it]
Batches: 51%|███       | 44/86 [05:35<05:06, 7.29s/it]
Batches: 52%|███       | 45/86 [05:42<04:58, 7.28s/it]
Batches: 53%|███       | 46/86 [05:49<04:52, 7.32s/it]
Batches: 55%|███       | 47/86 [05:57<04:47, 7.36s/it]
```

Batches: 56%	<div><div></div></div>	48/86 [06:05<04:45, 7.50s/it]
Batches: 57%	<div><div></div></div>	49/86 [06:12<04:35, 7.45s/it]
Batches: 58%	<div><div></div></div>	50/86 [06:19<04:26, 7.41s/it]
Batches: 59%	<div><div></div></div>	51/86 [06:27<04:20, 7.45s/it]
Batches: 60%	<div><div></div></div>	52/86 [06:34<04:08, 7.32s/it]
Batches: 62%	<div><div></div></div>	53/86 [06:42<04:06, 7.48s/it]
Batches: 63%	<div><div></div></div>	54/86 [06:49<04:01, 7.54s/it]
Batches: 64%	<div><div></div></div>	55/86 [06:57<03:50, 7.45s/it]
Batches: 65%	<div><div></div></div>	56/86 [07:04<03:40, 7.36s/it]
Batches: 66%	<div><div></div></div>	57/86 [07:12<03:37, 7.51s/it]
Batches: 67%	<div><div></div></div>	58/86 [07:19<03:31, 7.56s/it]
Batches: 69%	<div><div></div></div>	59/86 [07:27<03:27, 7.68s/it]
Batches: 70%	<div><div></div></div>	60/86 [07:35<03:17, 7.58s/it]
Batches: 71%	<div><div></div></div>	61/86 [07:42<03:10, 7.63s/it]
Batches: 72%	<div><div></div></div>	62/86 [07:50<03:01, 7.58s/it]
Batches: 73%	<div><div></div></div>	63/86 [07:57<02:50, 7.40s/it]
Batches: 74%	<div><div></div></div>	64/86 [08:05<02:46, 7.55s/it]
Batches: 76%	<div><div></div></div>	65/86 [08:12<02:37, 7.48s/it]
Batches: 77%	<div><div></div></div>	66/86 [08:20<02:31, 7.57s/it]
Batches: 78%	<div><div></div></div>	67/86 [08:26<02:18, 7.29s/it]
Batches: 79%	<div><div></div></div>	68/86 [08:32<02:05, 6.95s/it]
Batches: 80%	<div><div></div></div>	69/86 [08:40<01:58, 7.00s/it]
Batches: 81%	<div><div></div></div>	70/86 [08:47<01:54, 7.19s/it]
Batches: 83%	<div><div></div></div>	71/86 [08:55<01:50, 7.37s/it]
Batches: 84%	<div><div></div></div>	72/86 [09:02<01:42, 7.33s/it]
Batches: 85%	<div><div></div></div>	73/86 [09:09<01:34, 7.25s/it]
Batches: 86%	<div><div></div></div>	74/86 [09:17<01:28, 7.37s/it]
Batches: 87%	<div><div></div></div>	75/86 [09:24<01:20, 7.35s/it]
Batches: 88%	<div><div></div></div>	76/86 [09:32<01:13, 7.37s/it]
Batches: 90%	<div><div></div></div>	77/86 [09:39<01:04, 7.21s/it]
Batches: 91%	<div><div></div></div>	78/86 [09:46<00:59, 7.39s/it]
Batches: 92%	<div><div></div></div>	79/86 [09:53<00:50, 7.26s/it]
Batches: 93%	<div><div></div></div>	80/86 [10:00<00:43, 7.23s/it]
Batches: 94%	<div><div></div></div>	81/86 [10:09<00:37, 7.54s/it]
Batches: 95%	<div><div></div></div>	82/86 [10:16<00:29, 7.39s/it]
Batches: 97%	<div><div></div></div>	83/86 [10:29<00:27, 9.17s/it]
Batches: 98%	<div><div></div></div>	84/86 [10:37<00:17, 8.65s/it]
Batches: 99%	<div><div></div></div>	85/86 [10:44<00:08, 8.38s/it]
Batches: 100%	<div><div></div></div>	86/86 [10:46<00:00, 6.50s/it]
Batches: 100%	<div><div></div></div>	86/86 [10:46<00:00, 7.52s/it]

=====

✓ Successfully processed: 3002/3410  
X Skipped (missing/error): 408/3410

First 10 skipped: 000000393282.jpg, 000000393282.jpg, 000000393282.jpg, 000000393469.jpg, 000000000285.jpg, 000000262440.jpg, 000000262440.jpg, 000000262440.jpg, 000000262440.jpg, 000000131386.jpg...

Extracted shapes:  
Positive: (3002, 4096)  
Negative: (3002, 4096)  
Labels: (3002,)

Cached to: hidden\_states\_cache\_corrected/cache\_attribute\_recognition\_3410\_llava.npz

=====

TRAINING CCS PROBE

=====

## Dataset split:

Train: 2101 samples (1062 pos, 1039 neg)  
Test: 901 samples (456 pos, 445 neg)  
Hidden dim: 4096

## =====

## TRAINING WITH MULTIPLE RANDOM RESTARTS

=====

Trial 1/10: Loss = 0.006287  
Trial 2/10: Loss = 0.005402  
Trial 3/10: Loss = 0.006598  
Trial 4/10: Loss = 0.005286  
Trial 5/10: Loss = 0.006353  
Trial 6/10: Loss = 0.008032  
Trial 7/10: Loss = 0.006607  
Trial 8/10: Loss = 0.006231  
Trial 9/10: Loss = 0.005843  
Trial 10/10: Loss = 0.006463

## =====

## EVALUATION WITH BEST PROBE

=====

Best loss: 0.005286

## Test Results:

Overall Accuracy: 77.0% (694/901)  
Positive samples: 78.1% (456 samples)  
Negative samples: 76.0% (445 samples)

✓ COMPLETE: attribute\_recognition → 77.0%

#####

# CATEGORY: SPATIAL\_RECOGNITION

#####

## =====

## EXTRACTING HIDDEN STATES: SPATIAL\_RECOGNITION

=====

⚠ Cache disabled (use\_cache=False). Extracting new...

Processing 1030 samples in batches of 40

Searching in 2 image directories

LOADING MODEL: llava

Device: cuda

Loading checkpoint shards: 0%| | 0/3 [00:00<?, ?it/s]

Loading checkpoint shards: 33%| | 1/3 [00:06<00:12, 6.33s/it]

Loading checkpoint shards: 67%| | 2/3 [00:12<00:06, 6.22s/it]

Loading checkpoint shards: 100%| | 3/3 [00:17<00:00, 5.75s/it]

Loading checkpoint shards: 100%| | 3/3 [00:17<00:00, 5.89s/it]

✓ Model loaded successfully

Batches: 0%| | 0/26 [00:00<?, ?it/s]

Batches: 4%| | 1/26 [00:07<03:04, 7.36s/it]

Batches: 8%| | 2/26 [00:14<02:59, 7.47s/it]

Batches: 12%| | 3/26 [00:21<02:38, 6.90s/it]

Batches: 15%| | 4/26 [00:28<02:34, 7.04s/it]

Batches: 19%| | 5/26 [00:35<02:29, 7.14s/it]

Batches: 23%| | 6/26 [00:43<02:24, 7.22s/it]

Batches: 27%| | 7/26 [00:50<02:17, 7.23s/it]

Batches: 31%| | 8/26 [00:57<02:10, 7.23s/it]

Batches: 35%| | 9/26 [01:04<02:00, 7.09s/it]

Batches: 38%| | 10/26 [01:11<01:54, 7.16s/it]

```

Batches: 42% ██████████ | 11/26 [01:19<01:48, 7.23s/it]
Batches: 46% ██████████ | 12/26 [01:26<01:40, 7.21s/it]
Batches: 50% ██████████ | 13/26 [01:33<01:32, 7.09s/it]
Batches: 54% ██████████ | 14/26 [01:39<01:22, 6.91s/it]
Batches: 58% ██████████ | 15/26 [01:46<01:15, 6.90s/it]
Batches: 62% ██████████ | 16/26 [01:53<01:08, 6.88s/it]
Batches: 65% ██████████ | 17/26 [02:00<01:02, 6.92s/it]
Batches: 69% ██████████ | 18/26 [02:07<00:55, 6.96s/it]
Batches: 73% ██████████ | 19/26 [02:14<00:49, 7.08s/it]
Batches: 77% ██████████ | 20/26 [02:22<00:42, 7.16s/it]
Batches: 81% ██████████ | 21/26 [02:28<00:35, 7.06s/it]
Batches: 85% ██████████ | 22/26 [02:36<00:28, 7.17s/it]
Batches: 88% ██████████ | 23/26 [02:42<00:21, 7.04s/it]
Batches: 92% ██████████ | 24/26 [02:50<00:14, 7.11s/it]
Batches: 96% ██████████ | 25/26 [02:57<00:07, 7.03s/it]
Batches: 100% ██████████ | 26/26 [03:02<00:00, 6.56s/it]
Batches: 100% ██████████ | 26/26 [03:02<00:00, 7.02s/it]

```

```

=====
✓ Successfully processed: 880/1030
X Skipped (missing/error): 150/1030

```

First 10 skipped: 000000393282.jpg, 000000000285.jpg, 000000262682.jpg, 000000000632.jpg, 000000262895.jpg, 000000043816.jpg, 000000043816.jpg, 000000043816.jpg, 000000043816.jpg, 00000000785.jpg...

```

Extracted shapes:
  Positive: (880, 4096)
  Negative: (880, 4096)
  Labels: (880,)

```

Cached to: hidden\_states\_cache\_corrected/cache\_spatial\_recognition\_1030\_llava.npz

```

=====
TRAINING CCS PROBE
=====

```

```

Dataset split:
  Train: 615 samples (294 pos, 321 neg)
  Test: 265 samples (126 pos, 139 neg)
  Hidden dim: 4096

```

```

=====
TRAINING WITH MULTIPLE RANDOM RESTARTS
=====

```

```

Trial 1/10: Loss = 0.001677
Trial 2/10: Loss = 0.001346
Trial 3/10: Loss = 0.002143
Trial 4/10: Loss = 0.001893
Trial 5/10: Loss = 0.001935
Trial 6/10: Loss = 0.002284
Trial 7/10: Loss = 0.002975
Trial 8/10: Loss = 0.001574
Trial 9/10: Loss = 0.002096
Trial 10/10: Loss = 0.001778

```

```

=====
EVALUATION WITH BEST PROBE
=====

```

Best loss: 0.001346

```

Test Results:
  Overall Accuracy: 75.5% (200/265)
  Positive samples: 80.2% (126 samples)
  Negative samples: 71.2% (139 samples)

```



✓ COMPLETE: spatial\_recognition → 75.5%

=====

Final Results:

object_detection	: 79.0%
attribute_recognition	: 77.0%
spatial_recognition	: 75.5%

Average	: 77.2%
---------	---------

=====

=== Job finished at Fri Oct 24 02:35:27 CEST 2025 with exit code: 0 ===