```
=== GPU Information ===
NVIDIA A100-SXM4-40GB, 40960 MiB, 580.95.05
=== Checking if vision_ccs_corrected.py exists ===
-rw-r-----. 1 mdemirev mdemirev 18K Oct 23 14:31 vision_ccs_corrected.py
=== Running vision_ccs_corrected.py ===
/home/mdemirev/.local/lib/python3.11/site-packages/huggingface_hub/file_download.py:945:
FutureWarning: `resume_download` is deprecated and will be removed in version 1.0.0. Downloads
always resume when possible. If you want to force a new download, use `force_download=True`.
  warnings.warn(
Special tokens have been added in the vocabulary, make sure the associated word embeddings are
fine-tuned or trained.
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 [00:07<00:14,  7.18s/it]
Loading checkpoint shards:  67%|██████▋   | 2/3 [00:15<00:07,  7.59s/it]
Loading checkpoint shards: 100%|██████████| 3/3 [00:21<00:00,  7.20s/it]
Loading checkpoint shards: 100%|██████████| 3/3 [00:21<00:00,  7.27s/it]
✓ Model loaded successfully


Batches:   0%|          | 0/34 [00:00<?, ?it/s]
Batches:   3%|▎         | 1/34 [00:07<03:55,  7.15s/it]
Batches:   6%|▌         | 2/34 [00:13<03:38,  6.83s/it]
Batches:   9%|▉         | 3/34 [00:19<03:15,  6.32s/it]
Batches:  12%|█▏        | 4/34 [00:25<03:04,  6.14s/it]
Batches:  15%|█▍        | 5/34 [00:31<02:53,  5.99s/it]
Batches:  18%|█▊        | 6/34 [00:37<02:51,  6.13s/it]
Batches:  21%|██        | 7/34 [00:43<02:45,  6.11s/it]
Batches:  24%|██▎       | 8/34 [00:49<02:34,  5.94s/it]
Batches:  26%|██▋       | 9/34 [00:54<02:28,  5.92s/it]
Batches:  29%|██▉       | 10/34 [01:01<02:24,  6.03s/it]
Batches:  32%|███▏      | 11/34 [01:07<02:18,  6.03s/it]
Batches:  35%|███▌      | 12/34 [01:13<02:12,  6.00s/it]
Batches:  38%|███▊      | 13/34 [01:17<01:58,  5.62s/it]
Batches:  41%|████      | 14/34 [01:24<01:55,  5.80s/it]
Batches:  44%|████▍     | 15/34 [01:29<01:47,  5.67s/it]
Batches:  47%|████▋     | 16/34 [01:35<01:46,  5.89s/it]
Batches:  50%|█████     | 17/34 [01:41<01:38,  5.79s/it]
Batches:  53%|█████▎    | 18/34 [01:47<01:35,  5.94s/it]
Batches:  56%|█████▌    | 19/34 [01:54<01:32,  6.19s/it]
Batches:  59%|█████▉    | 20/34 [02:00<01:25,  6.11s/it]
Batches:  62%|██████▏   | 21/34 [02:06<01:19,  6.10s/it]
Batches:  65%|██████▍   | 22/34 [02:12<01:11,  5.98s/it]
Batches:  68%|██████▊   | 23/34 [02:18<01:06,  6.01s/it]
Batches:  71%|███████   | 24/34 [02:25<01:02,  6.24s/it]
Batches:  74%|███████▎  | 25/34 [02:31<00:56,  6.24s/it]
```

```
Batches:  76%|███████▌  | 26/34 [02:37<00:48,  6.11s/it]
Batches:  79%|███████▉  | 27/34 [02:43<00:43,  6.20s/it]
Batches:  82%|████████▏ | 28/34 [02:49<00:35,  5.96s/it]
Batches:  85%|████████▍ | 29/34 [02:55<00:29,  5.99s/it]
Batches:  88%|████████▊ | 30/34 [03:00<00:23,  5.91s/it]
Batches:  91%|█████████ | 31/34 [03:06<00:17,  5.90s/it]
Batches:  94%|█████████▍| 32/34 [03:12<00:11,  5.76s/it]
Batches:  97%|█████████▋| 33/34 [03:17<00:05,  5.75s/it]
Batches: 100%|██████████| 34/34 [03:18<00:00,  4.21s/it]
Batches: 100%|██████████| 34/34 [03:18<00:00,  5.84s/it]
/home/mdemirev/.local/lib/python3.11/site-packages/huggingface_hub/file_download.py:945:
FutureWarning: `resume_download` is deprecated and will be removed in version 1.0.0. Downloads
always resume when possible. If you want to force a new download, use `force_download=True`.
  warnings.warn(
Special tokens have been added in the vocabulary, make sure the associated word embeddings are
fine-tuned or trained.

======================================================================
✓ Successfully processed: 1140/1323
✗ 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 = 0.001015
  Trial  2/10: Loss = 0.000915
  Trial  3/10: Loss = 0.001956
  Trial  4/10: Loss = 0.001456
  Trial  5/10: Loss = 0.001009
  Trial  6/10: Loss = 0.001258
  Trial  7/10: Loss = 0.000903
  Trial  8/10: Loss = 0.000975
  Trial  9/10: Loss = 0.000891
  Trial 10/10: Loss = 0.001275

======================================================================
EVALUATION WITH BEST PROBE
======================================================================
Best loss: 0.000891

Test Results:
  Overall Accuracy: 80.5% (276/343)
  Positive samples: 82.1% (173 samples)
  Negative samples: 78.8% (170 samples)

✓ COMPLETE: object_detection → 80.5%
```

```
######################################################################
# 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 [00:02<00:04,  2.43s/it]
Loading checkpoint shards:  67%|██████▋   | 2/3 [00:04<00:02,  2.31s/it]
Loading checkpoint shards: 100%|██████████| 3/3 [00:06<00:00,  2.11s/it]
Loading checkpoint shards: 100%|██████████| 3/3 [00:06<00:00,  2.17s/it]
✓ Model loaded successfully


Batches:   0%|          | 0/86 [00:00<?, ?it/s]
Batches:   1%|          | 1/86 [00:06<09:04,  6.41s/it]
Batches:   2%|▏         | 2/86 [00:13<09:26,  6.74s/it]
Batches:   3%|▎         | 3/86 [00:19<09:14,  6.68s/it]
Batches:   5%|▍         | 4/86 [00:25<08:42,  6.37s/it]
Batches:   6%|▌         | 5/86 [00:32<08:47,  6.51s/it]
Batches:   7%|▋         | 6/86 [00:38<08:17,  6.22s/it]
Batches:   8%|▊         | 7/86 [00:44<08:15,  6.28s/it]
Batches:   9%|▉         | 8/86 [00:49<07:26,  5.73s/it]
Batches:  10%|▉         | 9/86 [00:54<07:21,  5.73s/it]
Batches:  12%|█         | 10/86 [01:00<07:03,  5.58s/it]
Batches:  13%|█▏        | 11/86 [01:06<07:13,  5.78s/it]
Batches:  14%|█▎        | 12/86 [01:12<07:05,  5.76s/it]
Batches:  15%|█▍        | 13/86 [01:17<06:56,  5.70s/it]
Batches:  16%|█▌        | 14/86 [01:23<06:50,  5.71s/it]
Batches:  17%|█▋        | 15/86 [01:29<06:56,  5.87s/it]
Batches:  19%|█▊        | 16/86 [01:36<07:09,  6.14s/it]
Batches:  20%|█▉        | 17/86 [01:42<07:08,  6.21s/it]
Batches:  21%|██        | 18/86 [01:48<06:42,  5.92s/it]
Batches:  22%|██▏       | 19/86 [01:54<06:36,  5.92s/it]
Batches:  23%|██▎       | 20/86 [02:00<06:40,  6.07s/it]
Batches:  24%|██▍       | 21/86 [02:06<06:41,  6.18s/it]
Batches:  26%|██▌       | 22/86 [02:13<06:45,  6.33s/it]
Batches:  27%|██▋       | 23/86 [02:19<06:40,  6.36s/it]
Batches:  28%|██▊       | 24/86 [02:25<06:22,  6.17s/it]
Batches:  29%|██▉       | 25/86 [02:32<06:27,  6.35s/it]
Batches:  30%|███       | 26/86 [02:38<06:22,  6.38s/it]
Batches:  31%|███▏      | 27/86 [02:44<06:05,  6.19s/it]
Batches:  33%|███▎      | 28/86 [02:50<05:48,  6.01s/it]
Batches:  34%|███▎      | 29/86 [02:56<05:47,  6.10s/it]
Batches:  35%|███▍      | 30/86 [03:02<05:32,  5.95s/it]
Batches:  36%|███▌      | 31/86 [03:08<05:30,  6.01s/it]
Batches:  37%|███▋      | 32/86 [03:15<05:40,  6.30s/it]
Batches:  38%|███▊      | 33/86 [03:21<05:25,  6.14s/it]
Batches:  40%|███▉      | 34/86 [03:26<05:08,  5.93s/it]
Batches:  41%|████      | 35/86 [03:33<05:15,  6.18s/it]
Batches:  42%|████▏     | 36/86 [03:40<05:18,  6.37s/it]
```

```
Batches:  43%|████▎     | 37/86 [03:46<05:08,  6.30s/it]
Batches:  44%|████▍     | 38/86 [03:52<04:59,  6.24s/it]
Batches:  45%|████▌     | 39/86 [03:57<04:43,  6.04s/it]
Batches:  47%|████▋     | 40/86 [04:03<04:38,  6.06s/it]
Batches:  48%|████▊     | 41/86 [04:10<04:35,  6.12s/it]
Batches:  49%|████▉     | 42/86 [04:15<04:17,  5.85s/it]
Batches:  50%|█████     | 43/86 [04:21<04:14,  5.92s/it]
Batches:  51%|█████     | 44/86 [04:27<04:08,  5.91s/it]
Batches:  52%|█████▏    | 45/86 [04:33<04:04,  5.96s/it]
Batches:  53%|█████▎    | 46/86 [04:39<03:59,  6.00s/it]
Batches:  55%|█████▍    | 47/86 [04:45<03:56,  6.07s/it]
Batches:  56%|█████▌    | 48/86 [04:52<03:56,  6.23s/it]
Batches:  57%|█████▋    | 49/86 [04:58<03:45,  6.08s/it]
Batches:  58%|█████▊    | 50/86 [05:04<03:37,  6.03s/it]
Batches:  59%|█████▉    | 51/86 [05:10<03:35,  6.15s/it]
Batches:  60%|██████    | 52/86 [05:16<03:24,  6.03s/it]
Batches:  62%|██████▏   | 53/86 [05:22<03:24,  6.21s/it]
Batches:  63%|██████▎   | 54/86 [05:29<03:22,  6.32s/it]
Batches:  64%|██████▍   | 55/86 [05:35<03:13,  6.25s/it]
Batches:  65%|██████▌   | 56/86 [05:41<03:06,  6.21s/it]
Batches:  66%|██████▋   | 57/86 [05:48<03:04,  6.38s/it]
Batches:  67%|██████▋   | 58/86 [05:55<03:01,  6.50s/it]
Batches:  69%|██████▊   | 59/86 [06:02<02:59,  6.63s/it]
Batches:  70%|██████▉   | 60/86 [06:08<02:46,  6.42s/it]
Batches:  71%|███████   | 61/86 [06:14<02:41,  6.47s/it]
Batches:  72%|███████▏  | 62/86 [06:20<02:34,  6.42s/it]
Batches:  73%|███████▎  | 63/86 [06:26<02:23,  6.23s/it]
Batches:  74%|███████▍  | 64/86 [06:33<02:19,  6.33s/it]
Batches:  76%|███████▌  | 65/86 [06:39<02:10,  6.21s/it]
Batches:  77%|███████▋  | 66/86 [06:45<02:06,  6.31s/it]
Batches:  78%|███████▊  | 67/86 [06:51<01:55,  6.10s/it]
Batches:  79%|███████▉  | 68/86 [06:56<01:44,  5.79s/it]
Batches:  80%|████████  | 69/86 [07:02<01:38,  5.77s/it]
Batches:  81%|████████  | 70/86 [07:08<01:34,  5.92s/it]
Batches:  83%|████████▎ | 71/86 [07:14<01:30,  6.03s/it]
Batches:  84%|████████▎ | 72/86 [07:20<01:23,  5.99s/it]
Batches:  85%|████████▍ | 73/86 [07:26<01:16,  5.87s/it]
Batches:  86%|████████▌ | 74/86 [07:32<01:13,  6.09s/it]
Batches:  87%|████████▋ | 75/86 [07:38<01:06,  6.09s/it]
Batches:  88%|████████▊ | 76/86 [07:45<01:01,  6.15s/it]
Batches:  90%|████████▉ | 77/86 [07:50<00:54,  6.03s/it]
Batches:  91%|█████████ | 78/86 [07:57<00:49,  6.21s/it]
Batches:  92%|█████████▏| 79/86 [08:03<00:42,  6.02s/it]
Batches:  93%|█████████▎| 80/86 [08:09<00:36,  6.05s/it]
Batches:  94%|█████████▍| 81/86 [08:15<00:30,  6.16s/it]
Batches:  95%|█████████▌| 82/86 [08:21<00:23,  5.93s/it]
Batches:  97%|█████████▋| 83/86 [08:27<00:18,  6.19s/it]
Batches:  98%|█████████▊| 84/86 [08:33<00:12,  6.10s/it]
Batches:  99%|█████████▉| 85/86 [08:40<00:06,  6.25s/it]
Batches: 100%|██████████| 86/86 [08:42<00:00,  4.92s/it]
Batches: 100%|██████████| 86/86 [08:42<00:00,  6.07s/it]
/home/mdemirev/.local/lib/python3.11/site-packages/huggingface_hub/file_download.py:945:
FutureWarning: `resume_download` is deprecated and will be removed in version 1.0.0. Downloads
always resume when possible. If you want to force a new download, use `force_download=True`.
  warnings.warn(
Special tokens have been added in the vocabulary, make sure the associated word embeddings are
fine-tuned or trained.
```

```
======================================================================
✓ Successfully processed: 3002/3410
✗ 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 = 0.005310
  Trial  2/10: Loss = 0.006096
  Trial  3/10: Loss = 0.005737
  Trial  4/10: Loss = 0.004670
  Trial  5/10: Loss = 0.006275
  Trial  6/10: Loss = 0.005877
  Trial  7/10: Loss = 0.006481
  Trial  8/10: Loss = 0.006128
  Trial  9/10: Loss = 0.006220
  Trial 10/10: Loss = 0.005875

======================================================================
EVALUATION WITH BEST PROBE
======================================================================
Best loss: 0.004670

Test Results:
  Overall Accuracy: 76.8% (692/901)
  Positive samples: 75.4% (456 samples)
  Negative samples: 78.2% (445 samples)

✓ COMPLETE: attribute_recognition → 76.8%

################################################################################
# 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 [00:02<00:04,  2.46s/it]
Loading checkpoint shards:  67%|██████▋   | 2/3 [00:04<00:02,  2.33s/it]
```

```
Loading checkpoint shards: 100%|██████████| 3/3 [00:06<00:00,  2.13s/it]
Loading checkpoint shards: 100%|██████████| 3/3 [00:06<00:00,  2.19s/it]
✓ Model loaded successfully


Batches:   0%|          | 0/26 [00:00<?, ?it/s]
Batches:   4%|▍         | 1/26 [00:06<02:43,  6.54s/it]
Batches:   8%|▊         | 2/26 [00:13<02:39,  6.65s/it]
Batches:  12%|█▏        | 3/26 [00:18<02:14,  5.85s/it]
Batches:  15%|█▌        | 4/26 [00:24<02:11,  5.96s/it]
Batches:  19%|█▉        | 5/26 [00:30<02:06,  6.01s/it]
Batches:  23%|██▎       | 6/26 [00:36<02:01,  6.08s/it]
Batches:  27%|██▋       | 7/26 [00:42<01:56,  6.14s/it]
Batches:  31%|███       | 8/26 [00:48<01:50,  6.12s/it]
Batches:  35%|███▍      | 9/26 [00:54<01:40,  5.89s/it]
Batches:  38%|███▊      | 10/26 [01:00<01:35,  5.94s/it]
Batches:  42%|████▏     | 11/26 [01:06<01:30,  6.03s/it]
Batches:  46%|████▌     | 12/26 [01:12<01:24,  6.00s/it]
Batches:  50%|█████     | 13/26 [01:18<01:16,  5.87s/it]
Batches:  54%|█████▍    | 14/26 [01:23<01:07,  5.62s/it]
Batches:  58%|█████▊    | 15/26 [01:28<01:02,  5.66s/it]
Batches:  62%|██████▏   | 16/26 [01:34<00:56,  5.68s/it]
Batches:  65%|██████▌   | 17/26 [01:40<00:51,  5.71s/it]
Batches:  69%|██████▉   | 18/26 [01:46<00:45,  5.73s/it]
Batches:  73%|███████▎  | 19/26 [01:52<00:41,  5.88s/it]
Batches:  77%|███████▋  | 20/26 [01:58<00:36,  6.04s/it]
Batches:  81%|████████  | 21/26 [02:04<00:29,  5.86s/it]
Batches:  85%|████████▍ | 22/26 [02:10<00:23,  5.99s/it]
Batches:  88%|████████▊ | 23/26 [02:16<00:17,  5.93s/it]
Batches:  92%|█████████▏| 24/26 [02:22<00:12,  6.04s/it]
Batches:  96%|█████████▌| 25/26 [02:28<00:05,  5.94s/it]
Batches: 100%|██████████| 26/26 [02:33<00:00,  5.57s/it]
Batches: 100%|██████████| 26/26 [02:33<00:00,  5.89s/it]

======================================================================
✓ Successfully processed: 880/1030
✗ 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,
000000000785.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 = 0.002678
  Trial  2/10: Loss = 0.001629
```

```
  Trial  3/10: Loss = 0.001632
  Trial  4/10: Loss = 0.001704
  Trial  5/10: Loss = 0.001500
  Trial  6/10: Loss = 0.001322
  Trial  7/10: Loss = 0.001643
  Trial  8/10: Loss = 0.001287
  Trial  9/10: Loss = 0.001858
  Trial 10/10: Loss = 0.002163

======================================================================
EVALUATION WITH BEST PROBE
======================================================================
Best loss: 0.001287

Test Results:
  Overall Accuracy: 74.3% (197/265)
  Positive samples: 78.6% (126 samples)
  Negative samples: 70.5% (139 samples)

✓ COMPLETE: spatial_recognition → 74.3%

======================================================================

Final Results:
  object_detection         : 80.5%
  attribute_recognition    : 76.8%
  spatial_recognition      : 74.3%

  Average                  : 77.2%

======================================================================

=== Job finished at Thu Oct 23 14:51:31 CEST 2025 with exit code: 0 ===
```