

=== GPU Information ===

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

=== Checking if vision_ccs.py exists ===

-rw-r----- 1 mdemirev mdemirev 17K Oct 22 02:39 vision_ccs.py

=== Running vision_ccs.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.

=====

VISIONCCS PIPELINE

LLaVA + CCS (Original Methodology)

=====

Configuration:

Model: llava-hf/llava-1.5-7b-hf

Samples per category:

- object_detection: 1323
- attribute_recognition: 3410
- spatial_recognition: 1030

Batch size: 40

Cache enabled: False

Categories: object_detection, attribute_recognition, spatial_recognition

CCS Training:

Epochs per trial: 1000

Random restarts: 10

Learning rate: 0.001

Weight decay: 0.01

#####

CATEGORY: OBJECT_DETECTION

#####

LOADING DATA for category: 'object_detection'

Using 1323 samples from '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 LLAVA MODEL: llava-hf/llava-1.5-7b-hf

=====

Device: cuda

Loading checkpoint shards: 0%| | 0/3 [00:00<?, ?it/s]
 Loading checkpoint shards: 33%|███| 1/3 [00:02<00:04, 2.42s/it]
 Loading checkpoint shards: 67%|██████| 2/3 [00:04<00:02, 2.33s/it]
 Loading checkpoint shards: 100%|██████████| 3/3 [00:06<00:00, 2.16s/it]
 Loading checkpoint shards: 100%|██████████| 3/3 [00:06<00:00, 2.21s/it]
 ✓ Model loaded successfully

Batches: 0%| | 0/34 [00:00<?, ?it/s]
 Batches: 3%|██| 1/34 [00:13<07:27, 13.56s/it]
 Batches: 6%|███| 2/34 [00:26<07:02, 13.20s/it]
 Batches: 9%|████| 3/34 [00:37<06:21, 12.31s/it]

Batches: 12%  | 4/34 [00:49<06:00, 12.02s/it]
Batches: 15%  | 5/34 [01:00<05:40, 11.72s/it]
Batches: 18%  | 6/34 [01:13<05:36, 12.00s/it]
Batches: 21%  | 7/34 [01:25<05:23, 11.98s/it]
Batches: 24%  | 8/34 [01:35<05:02, 11.62s/it]
Batches: 26%  | 9/34 [01:47<04:49, 11.59s/it]
Batches: 29%  | 10/34 [01:59<04:42, 11.79s/it]
Batches: 32%  | 11/34 [02:11<04:31, 11.82s/it]
Batches: 35%  | 12/34 [02:23<04:18, 11.75s/it]
Batches: 38%  | 13/34 [02:32<03:50, 11.00s/it]
Batches: 41%  | 14/34 [02:44<03:47, 11.36s/it]
Batches: 44%  | 15/34 [02:55<03:31, 11.12s/it]
Batches: 47%  | 16/34 [03:07<03:27, 11.53s/it]
Batches: 50%  | 17/34 [03:18<03:12, 11.33s/it]
Batches: 53%  | 18/34 [03:30<03:05, 11.62s/it]
Batches: 56%  | 19/34 [03:44<03:01, 12.13s/it]
Batches: 59%  | 20/34 [03:55<02:47, 11.97s/it]
Batches: 62%  | 21/34 [04:07<02:35, 11.95s/it]
Batches: 65%  | 22/34 [04:18<02:20, 11.71s/it]
Batches: 68%  | 23/34 [04:30<02:09, 11.75s/it]
Batches: 71%  | 24/34 [04:43<02:02, 12.21s/it]
Batches: 74%  | 25/34 [04:56<01:50, 12.23s/it]
Batches: 76%  | 26/34 [05:07<01:35, 11.96s/it]
Batches: 79%  | 27/34 [05:20<01:25, 12.16s/it]
Batches: 82%  | 28/34 [05:30<01:10, 11.69s/it]
Batches: 85%  | 29/34 [05:42<00:58, 11.74s/it]
Batches: 88%  | 30/34 [05:53<00:46, 11.58s/it]
Batches: 91%  | 31/34 [06:05<00:34, 11.57s/it]
Batches: 94%  | 32/34 [06:15<00:22, 11.28s/it]
Batches: 97%  | 33/34 [06:27<00:11, 11.27s/it]
Batches: 100%  | 34/34 [06:28<00:00, 8.22s/it]
Batches: 100%  | 34/34 [06:28<00:00, 11.42s/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.

=====

EXTRACTION COMPLETE

=====

✓ 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/cache_object_detection_1323_llava.npz

=====

TRAINING CCS PROBE

=====

Dataset split (Stratified):

Train: 797 samples (403 pos, 394 neg)

Test: 343 samples (173 pos, 170 neg)
Hidden dim: 4096

Probe architecture:
Input: 4096
Hidden: 256 → 128
Output: 1 (probability)

Training config:
Epochs per trial: 1000
Number of trials: 10
Learning rate: 0.001
Weight decay: 0.01

=====

TRAINING WITH MULTIPLE RANDOM RESTARTS

=====

Trial 1/10: Loss = 0.008419
✓ New best probe found!
Trial 2/10: Loss = 0.006876
✓ New best probe found!
Trial 3/10: Loss = 0.007835
Trial 4/10: Loss = 0.007059
Trial 5/10: Loss = 0.007925
Trial 6/10: Loss = 0.007450
Trial 7/10: Loss = 0.006797
✓ New best probe found!
Trial 8/10: Loss = 0.007042
Trial 9/10: Loss = 0.006951
Trial 10/10: Loss = 0.007232

=====

EVALUATION WITH BEST PROBE

=====

Best loss: 0.006797

Test Results:

Overall Accuracy: 79.3% (272/343)
Positive samples: 85.5% (173 samples)
Negative samples: 72.9% (170 samples)

✓ COMPLETE: object_detection → 79.3%

CATEGORY: ATTRIBUTE_RECOGNITION

LOADING DATA for category: 'attribute_recognition'
Using 3410 samples from '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 LLAVA MODEL: llava-hf/llava-1.5-7b-hf

=====

Device: cuda

Loading checkpoint shards: 0%| | 0/3 [00:00<?, ?it/s]
Loading checkpoint shards: 33%|███ | 1/3 [00:02<00:04, 2.47s/it]
Loading checkpoint shards: 67%|██████ | 2/3 [00:04<00:02, 2.36s/it]
Loading checkpoint shards: 100%|██████████| 3/3 [00:06<00:00, 2.16s/it]
Loading checkpoint shards: 100%|██████████| 3/3 [00:06<00:00, 2.23s/it]
✓ Model loaded successfully

Batches:	0%		0/86 [00:00<?, ?it/s]
Batches:	1%		1/86 [00:12<17:44, 12.53s/it]
Batches:	2%		2/86 [00:26<18:28, 13.20s/it]
Batches:	3%		3/86 [00:39<18:07, 13.10s/it]
Batches:	5%		4/86 [00:50<17:05, 12.51s/it]
Batches:	6%		5/86 [01:04<17:15, 12.78s/it]
Batches:	7%		6/86 [01:15<16:19, 12.24s/it]
Batches:	8%		7/86 [01:27<16:16, 12.36s/it]
Batches:	9%		8/86 [01:36<14:36, 11.24s/it]
Batches:	10%		9/86 [01:47<14:25, 11.24s/it]
Batches:	12%		10/86 [01:58<13:51, 10.94s/it]
Batches:	13%		11/86 [02:10<14:07, 11.30s/it]
Batches:	14%		12/86 [02:21<13:52, 11.25s/it]
Batches:	15%		13/86 [02:32<13:33, 11.15s/it]
Batches:	16%		14/86 [02:43<13:24, 11.17s/it]
Batches:	17%		15/86 [02:55<13:36, 11.50s/it]
Batches:	19%		16/86 [03:09<14:02, 12.03s/it]
Batches:	20%		17/86 [03:21<14:01, 12.19s/it]
Batches:	21%		18/86 [03:31<13:09, 11.61s/it]
Batches:	22%		19/86 [03:43<12:58, 11.61s/it]
Batches:	23%		20/86 [03:56<13:04, 11.89s/it]
Batches:	24%		21/86 [04:08<13:06, 12.10s/it]
Batches:	26%		22/86 [04:21<13:13, 12.40s/it]
Batches:	27%		23/86 [04:34<13:05, 12.47s/it]
Batches:	28%		24/86 [04:45<12:29, 12.09s/it]
Batches:	29%		25/86 [04:58<12:38, 12.44s/it]
Batches:	30%		26/86 [05:11<12:29, 12.49s/it]
Batches:	31%		27/86 [05:22<11:55, 12.13s/it]
Batches:	33%		28/86 [05:33<11:21, 11.75s/it]
Batches:	34%		29/86 [05:45<11:19, 11.91s/it]
Batches:	35%		30/86 [05:56<10:50, 11.61s/it]
Batches:	36%		31/86 [06:08<10:45, 11.73s/it]
Batches:	37%		32/86 [06:22<11:05, 12.32s/it]
Batches:	38%		33/86 [06:33<10:36, 12.01s/it]
Batches:	40%		34/86 [06:44<10:02, 11.58s/it]
Batches:	41%		35/86 [06:57<10:16, 12.08s/it]
Batches:	42%		36/86 [07:10<10:22, 12.45s/it]
Batches:	43%		37/86 [07:22<10:03, 12.31s/it]
Batches:	44%		38/86 [07:34<09:45, 12.21s/it]
Batches:	45%		39/86 [07:45<09:15, 11.82s/it]
Batches:	47%		40/86 [07:57<09:04, 11.84s/it]
Batches:	48%		41/86 [08:09<08:59, 11.98s/it]
Batches:	49%		42/86 [08:20<08:24, 11.46s/it]
Batches:	50%		43/86 [08:32<08:18, 11.59s/it]
Batches:	51%		44/86 [08:43<08:06, 11.58s/it]
Batches:	52%		45/86 [08:55<07:59, 11.68s/it]
Batches:	53%		46/86 [09:07<07:48, 11.72s/it]
Batches:	55%		47/86 [09:19<07:43, 11.88s/it]
Batches:	56%		48/86 [09:32<07:43, 12.19s/it]
Batches:	57%		49/86 [09:43<07:20, 11.90s/it]
Batches:	58%		50/86 [09:55<07:05, 11.82s/it]
Batches:	59%		51/86 [10:08<07:01, 12.04s/it]
Batches:	60%		52/86 [10:19<06:42, 11.84s/it]
Batches:	62%		53/86 [10:32<06:42, 12.20s/it]

Batches: 63% ██████████ | 54/86 [10:45<06:37, 12.42s/it]
Batches: 64% ██████████ | 55/86 [10:57<06:20, 12.26s/it]
Batches: 65% ██████████ | 56/86 [11:09<06:05, 12.17s/it]
Batches: 66% ██████████ | 57/86 [11:22<06:02, 12.50s/it]
Batches: 67% ██████████ | 58/86 [11:35<05:56, 12.74s/it]
Batches: 69% ██████████ | 59/86 [11:49<05:51, 13.01s/it]
Batches: 70% ██████████ | 60/86 [12:01<05:27, 12.59s/it]
Batches: 71% ██████████ | 61/86 [12:13<05:16, 12.68s/it]
Batches: 72% ██████████ | 62/86 [12:26<05:01, 12.58s/it]
Batches: 73% ██████████ | 63/86 [12:37<04:40, 12.20s/it]
Batches: 74% ██████████ | 64/86 [12:50<04:32, 12.40s/it]
Batches: 76% ██████████ | 65/86 [13:02<04:15, 12.15s/it]
Batches: 77% ██████████ | 66/86 [13:14<04:07, 12.37s/it]
Batches: 78% ██████████ | 67/86 [13:25<03:47, 11.96s/it]
Batches: 79% ██████████ | 68/86 [13:35<03:24, 11.34s/it]
Batches: 80% ██████████ | 69/86 [13:47<03:12, 11.30s/it]
Batches: 81% ██████████ | 70/86 [13:59<03:05, 11.59s/it]
Batches: 83% ██████████ | 71/86 [14:11<02:57, 11.81s/it]
Batches: 84% ██████████ | 72/86 [14:23<02:44, 11.74s/it]
Batches: 85% ██████████ | 73/86 [14:34<02:29, 11.50s/it]
Batches: 86% ██████████ | 74/86 [14:47<02:23, 11.93s/it]
Batches: 87% ██████████ | 75/86 [14:58<02:11, 11.93s/it]
Batches: 88% ██████████ | 76/86 [15:11<02:00, 12.05s/it]
Batches: 90% ██████████ | 77/86 [15:22<01:46, 11.82s/it]
Batches: 91% ██████████ | 78/86 [15:35<01:37, 12.16s/it]
Batches: 92% ██████████ | 79/86 [15:46<01:22, 11.78s/it]
Batches: 93% ██████████ | 80/86 [15:58<01:11, 11.85s/it]
Batches: 94% ██████████ | 81/86 [16:10<01:00, 12.06s/it]
Batches: 95% ██████████ | 82/86 [16:21<00:46, 11.62s/it]
Batches: 97% ██████████ | 83/86 [16:34<00:36, 12.14s/it]
Batches: 98% ██████████ | 84/86 [16:46<00:23, 11.96s/it]
Batches: 99% ██████████ | 85/86 [16:59<00:12, 12.27s/it]
Batches: 100% ██████████ | 86/86 [17:02<00:00, 9.63s/it]
Batches: 100% ██████████ | 86/86 [17:02<00:00, 11.89s/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.

=====

EXTRACTION COMPLETE

=====

✓ 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/cache_attribute_recognition_3410_llava.npz

=====

TRAINING CCS PROBE

=====

Dataset split (Stratified):

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

Probe architecture:

Input: 4096
Hidden: 256 → 128
Output: 1 (probability)

Training config:

Epochs per trial: 1000
Number of trials: 10
Learning rate: 0.001
Weight decay: 0.01

=====

TRAINING WITH MULTIPLE RANDOM RESTARTS

=====

Trial 1/10: Loss = 0.011204
✓ New best probe found!
Trial 2/10: Loss = 0.011921
Trial 3/10: Loss = 0.012789
Trial 4/10: Loss = 0.012990
Trial 5/10: Loss = 0.011565
Trial 6/10: Loss = 0.011913
Trial 7/10: Loss = 0.012190
Trial 8/10: Loss = 0.011544
Trial 9/10: Loss = 0.011536
Trial 10/10: Loss = 0.011247

=====

EVALUATION WITH BEST PROBE

=====

Best loss: 0.011204

Test Results:

Overall Accuracy: 75.9% (684/901)
Positive samples: 77.9% (456 samples)
Negative samples: 73.9% (445 samples)

✓ COMPLETE: attribute_recognition → 75.9%

#####

CATEGORY: SPATIAL_RECOGNITION

#####

LOADING DATA for category: 'spatial_recognition'

Using 1030 samples from '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 LLAVA MODEL: llava-hf/llava-1.5-7b-hf

=====

Device: cuda

Loading checkpoint shards: 0%| | 0/3 [00:00<?, ?it/s]
Loading checkpoint shards: 33%|██| 1/3 [00:02<00:04, 2.41s/it]
Loading checkpoint shards: 67%|██████| 2/3 [00:04<00:02, 2.31s/it]
Loading checkpoint shards: 100%|██████████| 3/3 [00:06<00:00, 2.13s/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 [00:12<05:21, 12.86s/it]
Batches:	8%		2/26 [00:26<05:13, 13.08s/it]
Batches:	12%		3/26 [00:35<04:23, 11.47s/it]
Batches:	15%		4/26 [00:47<04:16, 11.67s/it]
Batches:	19%		5/26 [00:59<04:07, 11.78s/it]
Batches:	23%		6/26 [01:11<03:58, 11.91s/it]
Batches:	27%		7/26 [01:24<03:48, 12.04s/it]
Batches:	31%		8/26 [01:35<03:35, 12.00s/it]
Batches:	35%		9/26 [01:46<03:15, 11.52s/it]
Batches:	38%		10/26 [01:58<03:06, 11.64s/it]
Batches:	42%		11/26 [02:10<02:57, 11.81s/it]
Batches:	46%		12/26 [02:22<02:44, 11.74s/it]
Batches:	50%		13/26 [02:32<02:29, 11.47s/it]
Batches:	54%		14/26 [02:42<02:11, 10.98s/it]
Batches:	58%		15/26 [02:54<02:01, 11.08s/it]
Batches:	62%		16/26 [03:05<01:51, 11.12s/it]
Batches:	65%		17/26 [03:16<01:40, 11.16s/it]
Batches:	69%		18/26 [03:27<01:29, 11.20s/it]
Batches:	73%		19/26 [03:40<01:20, 11.53s/it]
Batches:	77%		20/26 [03:52<01:11, 11.84s/it]
Batches:	81%		21/26 [04:03<00:57, 11.46s/it]
Batches:	85%		22/26 [04:15<00:46, 11.70s/it]
Batches:	88%		23/26 [04:26<00:34, 11.57s/it]
Batches:	92%		24/26 [04:39<00:23, 11.80s/it]
Batches:	96%		25/26 [04:50<00:11, 11.62s/it]
Batches:	100%		26/26 [04:59<00:00, 10.90s/it]
Batches:	100%		26/26 [04:59<00:00, 11.52s/it]

=====

EXTRACTION COMPLETE

=====

✓ 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/cache_spatial_recognition_1030_llava.npz

=====

TRAINING CCS PROBE

=====

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

Probe architecture:
Input: 4096
Hidden: 256 → 128
Output: 1 (probability)

Training config:

Epochs per trial: 1000

Number of trials: 10

Learning rate: 0.001

Weight decay: 0.01

=====

TRAINING WITH MULTIPLE RANDOM RESTARTS

=====

Trial 1/10: Loss = 0.007229

✓ New best probe found!

Trial 2/10: Loss = 0.009063

Trial 3/10: Loss = 0.008533

Trial 4/10: Loss = 0.007940

Trial 5/10: Loss = 0.007532

Trial 6/10: Loss = 0.007640

Trial 7/10: Loss = 0.007630

Trial 8/10: Loss = 0.007813

Trial 9/10: Loss = 0.007305

Trial 10/10: Loss = 0.007739

=====

EVALUATION WITH BEST PROBE

=====

Best loss: 0.007229

Test Results:

Overall Accuracy: 71.7% (190/265)

Positive samples: 75.4% (126 samples)

Negative samples: 68.3% (139 samples)

✓ COMPLETE: spatial_recognition → 71.7%

=====

ALL EXPERIMENTS FINISHED

=====

Final Results:

object_detection : 79.3%

attribute_recognition : 75.9%

spatial_recognition : 71.7%

Average : 75.6%

=====

=== Job finished at Wed Oct 22 03:35:33 CEST 2025 with exit code: 0 ===