

=== 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 16K Oct 21 18:35 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: 1000

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 1000 samples from 'object\_detection'

=====

EXTRACTING HIDDEN STATES: OBJECT\_DETECTION

=====

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

Processing 1000 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, 2.39s/it]  
 Loading checkpoint shards: 67%| | 2/3 [00:04<00:02, 2.25s/it]  
 Loading checkpoint shards: 100%| | 3/3 [00:06<00:00, 2.06s/it]  
 Loading checkpoint shards: 100%| | 3/3 [00:06<00:00, 2.13s/it]  
 ✓ Model loaded successfully

Batches: 0%| | 0/25 [00:00<?, ?it/s]  
 Batches: 4%| | 1/25 [00:13<05:22, 13.44s/it]  
 Batches: 8%| | 2/25 [00:26<05:00, 13.08s/it]  
 Batches: 12%| | 3/25 [00:37<04:27, 12.16s/it]  
 Batches: 16%| | 4/25 [00:48<04:09, 11.86s/it]  
 Batches: 20%| | 5/25 [00:59<03:51, 11.58s/it]

Batches: 24%  | 6/25 [01:12<03:45, 11.86s/it]  
 Batches: 28%  | 7/25 [01:24<03:33, 11.84s/it]  
 Batches: 32%  | 8/25 [01:34<03:15, 11.49s/it]  
 Batches: 36%  | 9/25 [01:46<03:03, 11.47s/it]  
 Batches: 40%  | 10/25 [01:58<02:54, 11.66s/it]  
 Batches: 44%  | 11/25 [02:10<02:43, 11.69s/it]  
 Batches: 48%  | 12/25 [02:21<02:31, 11.63s/it]  
 Batches: 52%  | 13/25 [02:30<02:10, 10.90s/it]  
 Batches: 56%  | 14/25 [02:42<02:03, 11.27s/it]  
 Batches: 60%  | 15/25 [02:53<01:50, 11.02s/it]  
 Batches: 64%  | 16/25 [03:05<01:42, 11.43s/it]  
 Batches: 68%  | 17/25 [03:16<01:29, 11.22s/it]  
 Batches: 72%  | 18/25 [03:28<01:20, 11.51s/it]  
 Batches: 76%  | 19/25 [03:41<01:12, 12.01s/it]  
 Batches: 80%  | 20/25 [03:53<00:59, 11.85s/it]  
 Batches: 84%  | 21/25 [04:05<00:47, 11.83s/it]  
 Batches: 88%  | 22/25 [04:16<00:34, 11.59s/it]  
 Batches: 92%  | 23/25 [04:27<00:23, 11.63s/it]  
 Batches: 96%  | 24/25 [04:40<00:12, 12.09s/it]  
 Batches: 100%  | 25/25 [04:53<00:00, 12.10s/it]  
 Batches: 100%  | 25/25 [04:53<00:00, 11.72s/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: 870/1000

X Skipped (missing/error): 130/1000

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: (870, 4096)

Negative: (870, 4096)

Labels: (870,)



Cached to: hidden\_states\_cache/cache\_object\_detection\_1000\_llava.npz

=====

TRAINING CCS PROBE

=====

Dataset split:

Train: 609 samples

Test: 261 samples

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 = 0.200740

✓ New best probe found!

Trial 2/10: Loss = 0.200833

Trial 3/10: Loss = 0.200778

Trial 4/10: Loss = 0.200491

✓ New best probe found!

Trial 5/10: Loss = 0.200768

Trial 6/10: Loss = 0.200835

Trial 7/10: Loss = 0.200960

Trial 8/10: Loss = 0.200619

Trial 9/10: Loss = 0.200452

✓ New best probe found!

Trial 10/10: Loss = 0.200616

=====

## EVALUATION WITH BEST PROBE

=====

Best loss: 0.200452

## Test Results:

Overall Accuracy: 54.4% (142/261)

Positive samples: 85.3% (143 samples)

Negative samples: 83.1% (118 samples)

✓ COMPLETE: object\_detection → 54.4%

#####

# CATEGORY: ATTRIBUTE\_RECOGNITION

#####

LOADING DATA for category: 'attribute\_recognition'

Using 1000 samples from 'attribute\_recognition'

=====

EXTRACTING HIDDEN STATES: ATTRIBUTE\_RECOGNITION

=====

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

Processing 1000 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&lt;?, ?it/s]

Loading checkpoint shards: 33%| | 1/3 [00:02&lt;00:04, 2.39s/it]

Loading checkpoint shards: 67%| | 2/3 [00:04&lt;00:02, 2.27s/it]

Loading checkpoint shards: 100%| | 3/3 [00:06&lt;00:00, 2.08s/it]

Loading checkpoint shards: 100%| | 3/3 [00:06&lt;00:00, 2.14s/it]

✓ Model loaded successfully

Batches: 0%| | 0/25 [00:00&lt;?, ?it/s]

Batches: 4%| | 1/25 [00:12&lt;04:57, 12.39s/it]

Batches: 8%| | 2/25 [00:25&lt;05:00, 13.06s/it]

Batches: 12%| | 3/25 [00:38&lt;04:45, 12.98s/it]

Batches: 16%| | 4/25 [00:50&lt;04:19, 12.38s/it]


















Batches: 20%| | 5/25 [01:03&lt;04:13, 12.65s/it]

Batches: 24%| | 6/25 [01:14&lt;03:50, 12.11s/it]

Batches: 28%| | 7/25 [01:26&lt;03:40, 12.23s/it]

Batches: 32%| | 8/25 [01:35&lt;03:09, 11.13s/it]

Batches: 36%| | 9/25 [01:46&lt;02:58, 11.13s/it]

Batches: 40%  | 10/25 [01:56<02:42, 10.82s/it]  
 Batches: 44%  | 11/25 [02:08<02:36, 11.19s/it]  
 Batches: 48%  | 12/25 [02:20<02:24, 11.14s/it]  
 Batches: 52%  | 13/25 [02:30<02:12, 11.05s/it]  
 Batches: 56%  | 14/25 [02:41<02:01, 11.07s/it]  
 Batches: 60%  | 15/25 [02:54<01:53, 11.39s/it]  
 Batches: 64%  | 16/25 [03:07<01:47, 11.93s/it]  
 Batches: 68%  | 17/25 [03:19<01:36, 12.08s/it]  
 Batches: 72%  | 18/25 [03:29<01:20, 11.50s/it]  
 Batches: 76%  | 19/25 [03:41<01:08, 11.50s/it]  
 Batches: 80%  | 20/25 [03:53<00:58, 11.77s/it]  
 Batches: 84%  | 21/25 [04:06<00:47, 12.00s/it]  
 Batches: 88%  | 22/25 [04:19<00:36, 12.29s/it]  
 Batches: 92%  | 23/25 [04:31<00:24, 12.35s/it]  
 Batches: 96%  | 24/25 [04:42<00:11, 11.98s/it]  
 Batches: 100%  | 25/25 [04:56<00:00, 12.34s/it]  
 Batches: 100%  | 25/25 [04:56<00:00, 11.84s/it]

Special tokens have been added in the vocabulary, make sure the associated word embeddings are fine-tuned or trained.

=====


EXTRACTION COMPLETE

=====

✓ Successfully processed: 879/1000  
 X Skipped (missing/error): 121/1000

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: (879, 4096)  
 Negative: (879, 4096)  
 Labels: (879,)

 Cached to: hidden\_states\_cache/cache\_attribute\_recognition\_1000\_llava.npz

=====

TRAINING CCS PROBE

=====

Dataset split:  
 Train: 615 samples  
 Test: 264 samples  
 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 = 0.200807  
 ✓ New best probe found!  
 Trial 2/10: Loss = 0.200837  
 Trial 3/10: Loss = 0.200868  
 Trial 4/10: Loss = 0.200640  
 ✓ New best probe found!

Trial 5/10: Loss = 0.200641  
Trial 6/10: Loss = 0.200724  
Trial 7/10: Loss = 0.200748  
Trial 8/10: Loss = 0.200426  
✓ New best probe found!  
Trial 9/10: Loss = 0.200719  
Trial 10/10: Loss = 0.200582

=====

EVALUATION WITH BEST PROBE

=====

Best loss: 0.200426

Test Results:

Overall Accuracy: 58.0% (153/264)  
Positive samples: 80.6% (129 samples)  
Negative samples: 63.7% (135 samples)

✓ COMPLETE: attribute\_recognition → 58.0%

#####  
# CATEGORY: SPATIAL\_RECOGNITION  
#####  
LOADING DATA for category: 'spatial\_recognition'  
Using 1000 samples from 'spatial\_recognition'

=====

EXTRACTING HIDDEN STATES: SPATIAL\_RECOGNITION

=====

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

Processing 1000 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, 2.38s/it]  
Loading checkpoint shards: 67%|██████ | 2/3 [00:04<00:02, 2.26s/it]  
Loading checkpoint shards: 100%|██████████| 3/3 [00:06<00:00, 2.08s/it]  
Loading checkpoint shards: 100%|██████████| 3/3 [00:06<00:00, 2.14s/it]

✓ Model loaded successfully

Batches: 0%| | 0/25 [00:00<?, ?it/s]  
Batches: 4%|█ | 1/25 [00:12<05:04, 12.69s/it]  
Batches: 8%|██ | 2/25 [00:25<04:57, 12.92s/it]  
Batches: 12%|███ | 3/25 [00:35<04:09, 11.34s/it]  
Batches: 16%|████ | 4/25 [00:47<04:02, 11.54s/it]  
Batches: 20%|█████ | 5/25 [00:58<03:52, 11.64s/it]  
Batches: 24%|██████ | 6/25 [01:10<03:43, 11.78s/it]  
Batches: 28%|███████ | 7/25 [01:23<03:34, 11.91s/it]  
Batches: 32%|████████ | 8/25 [01:34<03:21, 11.88s/it]  
Batches: 36%|█████████ | 9/25 [01:45<03:02, 11.41s/it]  
Batches: 40%|██████████ | 10/25 [01:57<02:52, 11.52s/it]  
Batches: 44%|███████████| 11/25 [02:09<02:43, 11.70s/it]  
Batches: 48%|███████████| 12/25 [02:20<02:31, 11.63s/it]  
Batches: 52%|███████████| 13/25 [02:31<02:16, 11.37s/it]  
Batches: 56%|███████████| 14/25 [02:41<01:59, 10.88s/it]  
Batches: 60%|███████████| 15/25 [02:52<01:49, 10.96s/it]  
Batches: 64%|███████████| 16/25 [03:03<01:39, 11.00s/it]  
Batches: 68%|███████████| 17/25 [03:14<01:28, 11.05s/it]

Batches: 72% ██████████ | 18/25 [03:25<01:17, 11.09s/it]  
Batches: 76% ██████████ | 19/25 [03:37<01:08, 11.41s/it]  
Batches: 80% ██████████ | 20/25 [03:50<00:58, 11.72s/it]  
Batches: 84% ██████████ | 21/25 [04:00<00:45, 11.34s/it]  
Batches: 88% ██████████ | 22/25 [04:12<00:34, 11.59s/it]  
Batches: 92% ██████████ | 23/25 [04:24<00:22, 11.46s/it]  
Batches: 96% ██████████ | 24/25 [04:36<00:11, 11.68s/it]  
Batches: 100% ██████████ | 25/25 [04:47<00:00, 11.52s/it]  
Batches: 100% ██████████ | 25/25 [04:47<00:00, 11.50s/it]

=====


EXTRACTION COMPLETE

=====

✓ Successfully processed: 853/1000  
X Skipped (missing/error): 147/1000

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: (853, 4096)  
Negative: (853, 4096)  
Labels: (853,)

 Cached to: hidden\_states\_cache/cache\_spatial\_recognition\_1000\_llava.npz

=====

TRAINING CCS PROBE

=====

Dataset split:  
Train: 597 samples  
Test: 256 samples  
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 = 0.201162  
✓ New best probe found!  
Trial 2/10: Loss = 0.200755  
✓ New best probe found!  
Trial 3/10: Loss = 0.200603  
✓ New best probe found!  
Trial 4/10: Loss = 0.200714  
Trial 5/10: Loss = 0.200718  
Trial 6/10: Loss = 0.200759  
Trial 7/10: Loss = 0.200732  
Trial 8/10: Loss = 0.200611  
Trial 9/10: Loss = 0.200898  
Trial 10/10: Loss = 0.200520  
✓ New best probe found!

=====

EVALUATION WITH BEST PROBE

=====  
Best loss: 0.200520

Test Results:

Overall Accuracy: 53.5% (137/256)  
Positive samples: 82.5% (126 samples)  
Negative samples: 74.6% (130 samples)

✓ COMPLETE: spatial\_recognition → 53.5%

=====  
ALL EXPERIMENTS FINISHED  
=====

Final Results:

object\_detection : 54.4%  
attribute\_recognition : 58.0%  
spatial\_recognition : 53.5%

Average : 55.3%

=====  
=== Job finished at Tue Oct 21 19:08:09 CEST 2025 with exit code: 0 ===