```
=== Running vision_linear.py ===
The image processor of type `Qwen2VLImageProcessor` is now loaded as a fast processor by default,
even if the model checkpoint was saved with a slow processor. This is a breaking change and may
produce slightly different outputs. To continue using the slow processor, instantiate this class
with `use_fast=False`. Note that this behavior will be extended to all models in a future release.
`torch_dtype` is deprecated! Use `dtype` instead!
Qwen/Qwen2-VL-7B-Instruct + Contrast Pairs + Supervised Methods

############################################################################
# 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: qwen2
Device: cuda

Loading checkpoint shards:   0%|              | 0/5 [00:00<?, ?it/s]
Loading checkpoint shards:  20%|              | 1/5 [00:04<00:17,  4.42s/it]
Loading checkpoint shards:  40%|              | 2/5 [00:08<00:13,  4.34s/it]
Loading checkpoint shards:  60%|              | 3/5 [00:13<00:08,  4.33s/it]
Loading checkpoint shards:  80%|              | 4/5 [00:17<00:04,  4.32s/it]
Loading checkpoint shards: 100%|              | 5/5 [00:18<00:00,  3.19s/it]
Loading checkpoint shards: 100%|              | 5/5 [00:18<00:00,  3.70s/it]
✓ Model loaded successfully


Batches:   0%|              | 0/34 [00:00<?, ?it/s]
Batches:   3%|              | 1/34 [00:07<04:06,  7.47s/it]
Batches:   6%|              | 2/34 [00:14<03:43,  6.98s/it]
Batches:   9%|              | 3/34 [00:19<03:11,  6.18s/it]
Batches:  12%|              | 4/34 [00:24<02:56,  5.88s/it]
Batches:  15%|              | 5/34 [00:30<02:47,  5.78s/it]
Batches:  18%|              | 6/34 [00:36<02:45,  5.92s/it]
Batches:  21%|              | 7/34 [00:42<02:38,  5.88s/it]
Batches:  24%|              | 8/34 [00:47<02:27,  5.69s/it]
Batches:  26%|              | 9/34 [00:53<02:22,  5.68s/it]
Batches:  29%|              | 10/34 [00:59<02:18,  5.78s/it]
Batches:  32%|              | 11/34 [01:05<02:13,  5.80s/it]
Batches:  35%|              | 12/34 [01:10<02:07,  5.80s/it]
Batches:  38%|              | 13/34 [01:15<01:55,  5.49s/it]
Batches:  41%|              | 14/34 [01:21<01:52,  5.62s/it]
Batches:  44%|              | 15/34 [01:26<01:44,  5.50s/it]
Batches:  47%|              | 16/34 [01:32<01:42,  5.69s/it]
Batches:  50%|              | 17/34 [01:38<01:35,  5.59s/it]
Batches:  53%|              | 18/34 [01:44<01:31,  5.73s/it]
Batches:  56%|              | 19/34 [01:50<01:29,  5.96s/it]
Batches:  59%|              | 20/34 [01:56<01:21,  5.84s/it]
Batches:  62%|              | 21/34 [02:02<01:16,  5.88s/it]
Batches:  65%|              | 22/34 [02:07<01:08,  5.75s/it]
Batches:  68%|              | 23/34 [02:13<01:03,  5.81s/it]
Batches:  71%|              | 24/34 [02:20<01:00,  6.06s/it]
Batches:  74%|              | 25/34 [02:26<00:54,  6.01s/it]
Batches:  76%|              | 26/34 [02:32<00:47,  5.93s/it]
Batches:  79%|              | 27/34 [02:38<00:42,  6.05s/it]
Batches:  82%|              | 28/34 [02:43<00:34,  5.77s/it]
Batches:  85%|              | 29/34 [02:49<00:29,  5.80s/it]
Batches:  88%|              | 30/34 [02:54<00:22,  5.69s/it]
Batches:  91%|              | 31/34 [03:00<00:17,  5.73s/it]
```

```
Batches:  94%|██████████   | 32/34 [03:05<00:11,  5.51s/it]
Batches:  97%|███████████  | 33/34 [03:11<00:05,  5.52s/it]
Batches: 100%|████████████ | 34/34 [03:11<00:00,  4.08s/it]
Batches: 100%|████████████ | 34/34 [03:11<00:00,  5.65s/it]
Successfully processed: 1140/1323
Skipped (missing/error): 183/1323

There are skipped images: 000000262227.jpg, 000000262440.jpg, 000000262440.jpg, 000000262682.jpg,
000000262682.jpg, 000000262682.jpg, 000000139684.jpg, 000000000632.jpg, 000000000632.jpg,
000000000632.jpg

Extracted shapes:
  Positive hidden states: (1140, 3584)
  Negative hidden states: (1140, 3584)
  Labels: (1140,)

Cached to: hidden_states_cache/cache_object_detection_1323_supervised_contrast_qwen2.npz


======================================================================
SUPERVISED LOGISTIC REGRESSION
======================================================================

Dataset split:
  Train: 570 samples (273 pos, 297 neg)
  Test:  570 samples (303 pos, 267 neg)
  Hidden dim: 3584

Logistic regression accuracy: 83.2%


======================================================================
SUPERVISED LINEAR PROBE
======================================================================

Results:
  Train Accuracy: 90.4%
  Test Accuracy:  85.3%
  Positive samples: 84.8% (303 samples)
  Negative samples: 85.8% (267 samples)


======================================================================
COMPARISON SUMMARY
======================================================================
  Logistic Regression:     83.2%
  Supervised Linear Probe: 85.3%


✓ COMPLETE: object_detection

######################################################################
# 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: qwen2
Device: cuda

Loading checkpoint shards:   0%|          | 0/5 [00:00<?, ?it/s]
Loading checkpoint shards:  20%|██        | 1/5 [00:01<00:07,  1.84s/it]
Loading checkpoint shards:  40%|████      | 2/5 [00:03<00:05,  1.73s/it]
Loading checkpoint shards:  60%|██████    | 3/5 [00:05<00:03,  1.72s/it]
Loading checkpoint shards:  80%|████████  | 4/5 [00:06<00:01,  1.71s/it]
Loading checkpoint shards: 100%|██████████| 5/5 [00:07<00:00,  1.26s/it]
Loading checkpoint shards: 100%|██████████| 5/5 [00:07<00:00,  1.47s/it]
✓ Model loaded successfully
```

```
Batches:    0%|                    | 0/86 [00:00<?, ?it/s]
Batches:    1%|                    | 1/86 [00:06<08:49,  6.23s/it]
Batches:    2%||                   | 2/86 [00:13<09:18,  6.65s/it]
Batches:    3%||                   | 3/86 [00:19<09:03,  6.55s/it]
Batches:    5%||                   | 4/86 [00:25<08:23,  6.14s/it]
Batches:    6%||                   | 5/86 [00:31<08:25,  6.25s/it]
Batches:    7%|■                   | 6/86 [00:36<07:55,  5.94s/it]
Batches:    8%|■                   | 7/86 [00:42<07:47,  5.92s/it]
Batches:    9%|■                   | 8/86 [00:47<07:06,  5.47s/it]
Batches:   10%|■                   | 9/86 [00:52<06:57,  5.42s/it]
Batches:   12%|■                   | 10/86 [00:57<06:37,  5.24s/it]
Batches:   13%|■                   | 11/86 [01:03<06:52,  5.51s/it]
Batches:   14%|■                   | 12/86 [01:09<06:46,  5.49s/it]
Batches:   15%|■                   | 13/86 [01:14<06:40,  5.49s/it]
Batches:   16%|■■                  | 14/86 [01:19<06:34,  5.48s/it]
Batches:   17%|■■                  | 15/86 [01:25<06:38,  5.62s/it]
Batches:   19%|■■                  | 16/86 [01:32<06:53,  5.91s/it]
Batches:   20%|■■                  | 17/86 [01:38<06:50,  5.95s/it]
Batches:   21%|■■                  | 18/86 [01:43<06:22,  5.62s/it]
Batches:   22%|■■                  | 19/86 [01:48<06:14,  5.59s/it]
Batches:   23%|■■                  | 20/86 [01:55<06:21,  5.78s/it]
Batches:   24%|■■                  | 21/86 [02:01<06:24,  5.91s/it]
Batches:   26%|■■                  | 22/86 [02:07<06:23,  6.00s/it]
Batches:   27%|■■                  | 23/86 [02:13<06:21,  6.06s/it]
Batches:   28%|■■■                 | 24/86 [02:19<06:05,  5.89s/it]
Batches:   29%|■■■                 | 25/86 [02:25<06:09,  6.06s/it]
Batches:   30%|■■■                 | 26/86 [02:32<06:08,  6.14s/it]
Batches:   31%|■■■                 | 27/86 [02:37<05:53,  5.98s/it]
Batches:   33%|■■■                 | 28/86 [02:43<05:37,  5.82s/it]
Batches:   34%|■■■                 | 29/86 [02:49<05:34,  5.87s/it]
Batches:   35%|■■■                 | 30/86 [02:54<05:19,  5.71s/it]
Batches:   36%|■■■                 | 31/86 [03:00<05:17,  5.77s/it]
Batches:   37%|■■■                 | 32/86 [03:07<05:30,  6.13s/it]
Batches:   38%|■■■                 | 33/86 [03:12<05:14,  5.94s/it]
Batches:   40%|■■■■                | 34/86 [03:18<05:01,  5.80s/it]
Batches:   41%|■■■■                | 35/86 [03:24<05:09,  6.06s/it]
Batches:   42%|■■■■                | 36/86 [03:31<05:09,  6.19s/it]
Batches:   43%|■■■■                | 37/86 [03:37<04:58,  6.10s/it]
Batches:   44%|■■■■                | 38/86 [03:43<04:47,  6.00s/it]
Batches:   45%|■■■■                | 39/86 [03:48<04:32,  5.81s/it]
Batches:   47%|■■■■                | 40/86 [03:54<04:26,  5.80s/it]
Batches:   48%|■■■■                | 41/86 [03:59<04:19,  5.77s/it]
Batches:   49%|■■■■                | 42/86 [04:04<04:02,  5.51s/it]
Batches:   50%|■■■■■               | 43/86 [04:10<04:02,  5.65s/it]
Batches:   51%|■■■■■               | 44/86 [04:16<03:52,  5.55s/it]
Batches:   52%|■■■■■               | 45/86 [04:22<03:52,  5.67s/it]
Batches:   53%|■■■■■               | 46/86 [04:28<03:51,  5.78s/it]
Batches:   55%|■■■■■               | 47/86 [04:33<03:46,  5.81s/it]
Batches:   56%|■■■■■               | 48/86 [04:40<03:47,  6.00s/it]
Batches:   57%|■■■■■               | 49/86 [04:45<03:35,  5.84s/it]
Batches:   58%|■■■■■               | 50/86 [04:51<03:28,  5.78s/it]
Batches:   59%|■■■■■               | 51/86 [04:57<03:26,  5.91s/it]
Batches:   60%|■■■■■■              | 52/86 [05:03<03:17,  5.81s/it]
Batches:   62%|■■■■■■              | 53/86 [05:09<03:15,  5.94s/it]
Batches:   63%|■■■■■■              | 54/86 [05:15<03:13,  6.04s/it]
Batches:   64%|■■■■■■              | 55/86 [05:21<03:06,  6.01s/it]
```

```
Batches:  65%|██████    |  56/86 [05:27<02:59,  5.97s/it]
Batches:  66%|██████▌   |  57/86 [05:34<02:57,  6.12s/it]
Batches:  67%|██████▋   |  58/86 [05:40<02:54,  6.23s/it]
Batches:  69%|██████▊   |  59/86 [05:47<02:51,  6.35s/it]
Batches:  70%|███████   |  60/86 [05:52<02:40,  6.15s/it]
Batches:  71%|███████   |  61/86 [05:59<02:35,  6.20s/it]
Batches:  72%|███████▏  |  62/86 [06:05<02:27,  6.17s/it]
Batches:  73%|███████▎  |  63/86 [06:10<02:16,  5.95s/it]
Batches:  74%|███████▍  |  64/86 [06:17<02:13,  6.06s/it]
Batches:  76%|███████▌  |  65/86 [06:22<02:06,  6.01s/it]
Batches:  77%|███████▋  |  66/86 [06:29<02:00,  6.02s/it]
Batches:  78%|███████▊  |  67/86 [06:34<01:53,  5.97s/it]
Batches:  79%|███████▉  |  68/86 [06:39<01:41,  5.66s/it]
Batches:  80%|████████  |  69/86 [06:45<01:36,  5.68s/it]
Batches:  81%|████████  |  70/86 [06:51<01:32,  5.75s/it]
Batches:  83%|████████▎ |  71/86 [06:57<01:28,  5.89s/it]
Batches:  84%|████████▍ |  72/86 [07:03<01:20,  5.77s/it]
Batches:  85%|████████▍ |  73/86 [07:08<01:13,  5.67s/it]
Batches:  86%|████████▌ |  74/86 [07:15<01:10,  5.91s/it]
Batches:  87%|████████▋ |  75/86 [07:21<01:05,  5.94s/it]
Batches:  88%|████████▊ |  76/86 [07:27<01:00,  6.01s/it]
Batches:  90%|████████▉ |  77/86 [07:32<00:52,  5.88s/it]
Batches:  91%|█████████ |  78/86 [07:39<00:48,  6.08s/it]
Batches:  92%|█████████▏|  79/86 [07:44<00:40,  5.85s/it]
Batches:  93%|█████████▎|  80/86 [07:50<00:35,  5.85s/it]
Batches:  94%|█████████▍|  81/86 [07:56<00:29,  5.96s/it]
Batches:  95%|█████████▌|  82/86 [08:01<00:22,  5.65s/it]
Batches:  97%|█████████▋|  83/86 [08:08<00:17,  5.93s/it]
Batches:  98%|█████████▊|  84/86 [08:13<00:11,  5.87s/it]
Batches:  99%|█████████▉|  85/86 [08:20<00:06,  6.00s/it]
Batches: 100%|██████████|  86/86 [08:22<00:00,  4.74s/it]
Batches: 100%|██████████|  86/86 [08:22<00:00,  5.84s/it]
/gpfs/home6/mdemirev/snellius/venv/lib/python3.11/site-
packages/sklearn/linear_model/_logistic.py:473: ConvergenceWarning: lbfgs failed to converge after
100 iteration(s) (status=1):
STOP: TOTAL NO. OF ITERATIONS REACHED LIMIT

Increase the number of iterations to improve the convergence (max_iter=100).
You might also want to scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
Successfully processed: 3002/3410
Skipped (missing/error): 408/3410

There are skipped images: 000000393282.jpg, 000000393282.jpg, 000000393282.jpg, 000000393469.jpg,
000000000285.jpg, 000000262440.jpg, 000000262440.jpg, 000000262440.jpg, 000000262440.jpg,
000000131386.jpg

Extracted shapes:
  Positive hidden states: (3002, 3584)
  Negative hidden states: (3002, 3584)
  Labels: (3002,)

Cached to: hidden_states_cache/cache_attribute_recognition_3410_supervised_contrast_qwen2.npz

========================================================================
SUPERVISED LOGISTIC REGRESSION
========================================================================

Dataset split:
```

```
   Train: 1501 samples (745 pos, 756 neg)
   Test:  1501 samples (773 pos, 728 neg)
   Hidden dim: 3584

 Logistic regression accuracy: 77.3%


   ======================================================================
   SUPERVISED LINEAR PROBE
   ======================================================================

 Results:
   Train Accuracy: 84.7%
   Test Accuracy:  82.9%
   Positive samples: 80.6% (773 samples)
   Negative samples: 85.4% (728 samples)


   ======================================================================
   COMPARISON SUMMARY
   ======================================================================
   Logistic Regression:     77.3%
   Supervised Linear Probe: 82.9%


 ✓ COMPLETE: attribute_recognition

 ######################################################################
 # 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: qwen2
 Device: cuda

 Loading checkpoint shards:   0%|           | 0/5 [00:00<?, ?it/s]
 Loading checkpoint shards:  20%|█          | 1/5 [00:01<00:07,  1.86s/it]
 Loading checkpoint shards:  40%|██         | 2/5 [00:03<00:05,  1.75s/it]
 Loading checkpoint shards:  60%|███        | 3/5 [00:05<00:03,  1.74s/it]
 Loading checkpoint shards:  80%|████       | 4/5 [00:06<00:01,  1.73s/it]
 Loading checkpoint shards: 100%|█████      | 5/5 [00:07<00:00,  1.27s/it]
 Loading checkpoint shards: 100%|█████      | 5/5 [00:07<00:00,  1.49s/it]
 ✓ Model loaded successfully


 Batches:   0%|           | 0/26 [00:00<?, ?it/s]
 Batches:   4%|█          | 1/26 [00:06<02:39,  6.37s/it]
 Batches:   8%|█          | 2/26 [00:12<02:35,  6.49s/it]
 Batches:  12%|█          | 3/26 [00:17<02:09,  5.62s/it]
 Batches:  15%|█          | 4/26 [00:23<02:06,  5.76s/it]
 Batches:  19%|██         | 5/26 [00:29<02:02,  5.85s/it]
 Batches:  23%|██         | 6/26 [00:35<01:57,  5.86s/it]
 Batches:  27%|███        | 7/26 [00:41<01:51,  5.85s/it]
 Batches:  31%|███        | 8/26 [00:47<01:45,  5.83s/it]
 Batches:  35%|███        | 9/26 [00:52<01:35,  5.62s/it]
 Batches:  38%|███        | 10/26 [00:58<01:31,  5.70s/it]
 Batches:  42%|████       | 11/26 [01:03<01:26,  5.76s/it]
 Batches:  46%|████       | 12/26 [01:09<01:20,  5.77s/it]
 Batches:  50%|████       | 13/26 [01:15<01:13,  5.64s/it]
 Batches:  54%|████       | 14/26 [01:19<01:04,  5.41s/it]
 Batches:  58%|█████      | 15/26 [01:25<01:00,  5.48s/it]
 Batches:  62%|█████      | 16/26 [01:31<00:54,  5.48s/it]
 Batches:  65%|█████      | 17/26 [01:36<00:49,  5.51s/it]
 Batches:  69%|██████     | 18/26 [01:42<00:44,  5.52s/it]
```

```
Batches:  73%|████████   | 19/26 [01:48<00:39,  5.63s/it]
Batches:  77%|████████▌  | 20/26 [01:54<00:34,  5.75s/it]
Batches:  81%|█████████  | 21/26 [01:59<00:27,  5.57s/it]
Batches:  85%|█████████▍ | 22/26 [02:05<00:22,  5.72s/it]
Batches:  88%|█████████▊ | 23/26 [02:10<00:17,  5.70s/it]
Batches:  92%|██████████▏| 24/26 [02:17<00:11,  5.83s/it]
Batches:  96%|██████████▌| 25/26 [02:22<00:05,  5.68s/it]
Batches: 100%|███████████| 26/26 [02:26<00:00,  5.27s/it]
Batches: 100%|███████████| 26/26 [02:26<00:00,  5.65s/it]
Successfully processed: 880/1030
Skipped (missing/error): 150/1030
```

There are skipped images: 000000393282.jpg, 000000000285.jpg, 000000262682.jpg, 000000000632.jpg, 000000262895.jpg, 000000043816.jpg, 000000043816.jpg, 000000043816.jpg, 000000043816.jpg, 000000000785.jpg

```
Extracted shapes:
  Positive hidden states: (880, 3584)
  Negative hidden states: (880, 3584)
  Labels: (880,)
```

Cached to: hidden_states_cache/cache_spatial_recognition_1030_supervised_contrast_qwen2.npz

```
========================================================================
SUPERVISED LOGISTIC REGRESSION
========================================================================

Dataset split:
  Train: 440 samples (214 pos, 226 neg)
  Test:  440 samples (206 pos, 234 neg)
  Hidden dim: 3584

Logistic regression accuracy: 74.1%

========================================================================
SUPERVISED LINEAR PROBE
========================================================================

Results:
  Train Accuracy: 86.1%
  Test Accuracy:  78.4%
  Positive samples: 76.7% (206 samples)
  Negative samples: 79.9% (234 samples)

========================================================================
COMPARISON SUMMARY
========================================================================
  Logistic Regression:     74.1%
  Supervised Linear Probe: 78.4%
```

✓ COMPLETE: spatial_recognition

```
========================================================================
FINAL RESULTS SUMMARY
========================================================================
```

Category	LogReg	Linear Probe
object_detection	83.2%	85.3%
attribute_recognition	77.3%	82.9%
spatial_recognition	74.1%	78.4%
Average	78.2%	82.2%

```
========================================================================

=== Job finished at Fri Oct 24 20:27:53 CEST 2025 with exit code: 0 ===
```