```
=== Running supervised_vision_ccs.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 + Logistic Regression

############################################################################
# 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,  4.41s/it]
Loading checkpoint shards:   40%|████      | 2/5 [00:08<00:12,  4.29s/it]
Loading checkpoint shards:   60%|██████    | 3/5 [00:12<00:08,  4.29s/it]
Loading checkpoint shards:   80%|████████  | 4/5 [00:17<00:04,  4.29s/it]
Loading checkpoint shards:  100%|██████████| 5/5 [00:18<00:00,  3.16s/it]
Loading checkpoint shards:  100%|██████████| 5/5 [00:18<00:00,  3.67s/it]
✓ Model loaded successfully


Batches:    0%|          | 0/34 [00:00<?, ?it/s]
Batches:    3%|▎         | 1/34 [00:07<04:00,  7.29s/it]
Batches:    6%|▌         | 2/34 [00:13<03:40,  6.89s/it]
Batches:    9%|▉         | 3/34 [00:19<03:09,  6.12s/it]
Batches:   12%|█▏        | 4/34 [00:24<02:54,  5.81s/it]
Batches:   15%|█▍        | 5/34 [00:29<02:45,  5.70s/it]
Batches:   18%|█▊        | 6/34 [00:36<02:43,  5.85s/it]
Batches:   21%|██        | 7/34 [00:41<02:37,  5.82s/it]
Batches:   24%|██▎       | 8/34 [00:47<02:26,  5.64s/it]
Batches:   26%|██▋       | 9/34 [00:52<02:20,  5.63s/it]
Batches:   29%|██▉       | 10/34 [00:58<02:16,  5.71s/it]
Batches:   32%|███▏      | 11/34 [01:04<02:11,  5.74s/it]
Batches:   35%|███▌      | 12/34 [01:10<02:06,  5.74s/it]
Batches:   38%|███▊      | 13/34 [01:14<01:54,  5.44s/it]
Batches:   41%|████      | 14/34 [01:20<01:51,  5.58s/it]
Batches:   44%|████▍     | 15/34 [01:25<01:43,  5.45s/it]
Batches:   47%|████▋     | 16/34 [01:32<01:41,  5.64s/it]
Batches:   50%|█████     | 17/34 [01:37<01:34,  5.54s/it]
Batches:   53%|█████▎    | 18/34 [01:43<01:30,  5.68s/it]
Batches:   56%|█████▌    | 19/34 [01:49<01:28,  5.91s/it]
Batches:   59%|█████▉    | 20/34 [01:55<01:21,  5.79s/it]
Batches:   62%|██████▏   | 21/34 [02:00<01:14,  5.77s/it]
Batches:   65%|██████▍   | 22/34 [02:06<01:07,  5.64s/it]
Batches:   68%|██████▊   | 23/34 [02:12<01:02,  5.71s/it]
Batches:   71%|███████   | 24/34 [02:18<00:59,  5.93s/it]
Batches:   74%|███████▎  | 25/34 [02:24<00:53,  5.90s/it]
Batches:   76%|███████▋  | 26/34 [02:30<00:46,  5.82s/it]
Batches:   79%|███████▉  | 27/34 [02:36<00:41,  5.94s/it]
Batches:   82%|████████▏ | 28/34 [02:41<00:34,  5.67s/it]
Batches:   85%|████████▌ | 29/34 [02:47<00:28,  5.72s/it]
Batches:   88%|████████▊ | 30/34 [02:52<00:22,  5.62s/it]
Batches:   91%|█████████ | 31/34 [02:58<00:16,  5.66s/it]
Batches:   94%|█████████▍| 32/34 [03:03<00:10,  5.45s/it]
```

```
Batches:  97%|███████  | 33/34 [03:08<00:05,  5.45s/it]
Batches: 100%|████████ | 34/34 [03:09<00:00,  4.03s/it]
Batches: 100%|████████ | 34/34 [03:09<00:00,  5.57s/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

TRAINING 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: 0.8315789473684211

✓ COMPLETE: object_detection → 83.2%

```
############################################################################
# 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,  1.86s/it]
Loading checkpoint shards:  40%|██       | 2/5 [00:03<00:05,  1.74s/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<08:46,  6.19s/it]
Batches:   2%||        | 2/86 [00:13<09:15,  6.61s/it]
Batches:   3%|▊        | 3/86 [00:19<09:01,  6.53s/it]
Batches:   5%|▊        | 4/86 [00:25<08:22,  6.12s/it]
Batches:   6%|▊        | 5/86 [00:31<08:25,  6.24s/it]
Batches:   7%|█        | 6/86 [00:36<07:54,  5.93s/it]
Batches:   8%|█        | 7/86 [00:42<07:47,  5.92s/it]
Batches:   9%|█        | 8/86 [00:47<07:05,  5.46s/it]
Batches:  10%|█        | 9/86 [00:52<06:56,  5.40s/it]
Batches:  12%|█        | 10/86 [00:57<06:36,  5.22s/it]
Batches:  13%|█        | 11/86 [01:03<06:51,  5.49s/it]
Batches:  14%|█▋       | 12/86 [01:08<06:43,  5.45s/it]
Batches:  15%|█▋       | 13/86 [01:14<06:37,  5.45s/it]
Batches:  16%|█▊       | 14/86 [01:19<06:31,  5.44s/it]
Batches:  17%|█▊       | 15/86 [01:25<06:35,  5.57s/it]
Batches:  19%|█▊       | 16/86 [01:31<06:49,  5.85s/it]
```

```
Batches:  20%|██          | 17/86 [01:38<06:47,  5.91s/it]
Batches:  21%|█           | 18/86 [01:42<06:19,  5.58s/it]
Batches:  22%|█           | 19/86 [01:48<06:11,  5.54s/it]
Batches:  23%|██          | 20/86 [01:54<06:17,  5.72s/it]
Batches:  24%|██          | 21/86 [02:00<06:20,  5.85s/it]
Batches:  26%|██          | 22/86 [02:06<06:19,  5.93s/it]
Batches:  27%|██          | 23/86 [02:12<06:16,  5.98s/it]
Batches:  28%|██          | 24/86 [02:18<05:59,  5.80s/it]
Batches:  29%|██          | 25/86 [02:24<06:04,  5.97s/it]
Batches:  30%|██          | 26/86 [02:30<06:03,  6.06s/it]
Batches:  31%|███         | 27/86 [02:36<05:48,  5.90s/it]
Batches:  33%|███         | 28/86 [02:41<05:33,  5.75s/it]
Batches:  34%|███         | 29/86 [02:47<05:30,  5.80s/it]
Batches:  35%|███         | 30/86 [02:52<05:15,  5.64s/it]
Batches:  36%|███         | 31/86 [02:58<05:13,  5.70s/it]
Batches:  37%|███         | 32/86 [03:05<05:26,  6.04s/it]
Batches:  38%|███         | 33/86 [03:11<05:10,  5.85s/it]
Batches:  40%|███         | 34/86 [03:16<04:57,  5.72s/it]
Batches:  41%|███         | 35/86 [03:23<05:05,  5.98s/it]
Batches:  42%|████        | 36/86 [03:29<05:05,  6.12s/it]
Batches:  43%|████        | 37/86 [03:35<04:55,  6.03s/it]
Batches:  44%|████        | 38/86 [03:40<04:44,  5.92s/it]
Batches:  45%|████        | 39/86 [03:46<04:30,  5.75s/it]
Batches:  47%|████        | 40/86 [03:51<04:23,  5.74s/it]
Batches:  48%|████        | 41/86 [03:57<04:16,  5.70s/it]
Batches:  49%|████        | 42/86 [04:02<03:59,  5.44s/it]
Batches:  50%|████        | 43/86 [04:08<03:59,  5.58s/it]
Batches:  51%|████        | 44/86 [04:13<03:49,  5.47s/it]
Batches:  52%|████        | 45/86 [04:19<03:49,  5.59s/it]
Batches:  53%|████        | 46/86 [04:25<03:48,  5.70s/it]
Batches:  55%|████        | 47/86 [04:31<03:43,  5.74s/it]
Batches:  56%|████        | 48/86 [04:37<03:44,  5.92s/it]
Batches:  57%|█████       | 49/86 [04:42<03:33,  5.76s/it]
Batches:  58%|█████       | 50/86 [04:48<03:25,  5.70s/it]
Batches:  59%|█████       | 51/86 [04:54<03:23,  5.82s/it]
Batches:  60%|█████       | 52/86 [05:00<03:14,  5.73s/it]
Batches:  62%|█████       | 53/86 [05:06<03:13,  5.86s/it]
Batches:  63%|█████       | 54/86 [05:12<03:10,  5.97s/it]
Batches:  64%|█████       | 55/86 [05:18<03:03,  5.93s/it]
Batches:  65%|█████       | 56/86 [05:24<02:56,  5.90s/it]
Batches:  66%|█████       | 57/86 [05:30<02:55,  6.04s/it]
Batches:  67%|█████       | 58/86 [05:36<02:52,  6.15s/it]
Batches:  69%|█████       | 59/86 [05:43<02:49,  6.27s/it]
Batches:  70%|██████      | 60/86 [05:49<02:37,  6.08s/it]
Batches:  71%|██████      | 61/86 [05:55<02:33,  6.12s/it]
Batches:  72%|██████      | 62/86 [06:01<02:26,  6.09s/it]
Batches:  73%|██████      | 63/86 [06:06<02:15,  5.89s/it]
Batches:  74%|██████      | 64/86 [06:12<02:11,  5.98s/it]
Batches:  76%|██████      | 65/86 [06:18<02:04,  5.94s/it]
Batches:  77%|██████      | 66/86 [06:24<01:58,  5.94s/it]
Batches:  78%|██████      | 67/86 [06:30<01:52,  5.90s/it]
Batches:  79%|██████      | 68/86 [06:35<01:40,  5.59s/it]
Batches:  80%|███████     | 69/86 [06:41<01:35,  5.60s/it]
Batches:  81%|███████     | 70/86 [06:46<01:30,  5.68s/it]
Batches:  83%|███████     | 71/86 [06:53<01:27,  5.81s/it]
```

```
Batches:  84%|████████     |  72/86 [06:58<01:19,  5.70s/it]
Batches:  85%|████████     |  73/86 [07:03<01:12,  5.60s/it]
Batches:  86%|████████     |  74/86 [07:10<01:10,  5.84s/it]
Batches:  87%|████████     |  75/86 [07:16<01:04,  5.87s/it]
Batches:  88%|████████     |  76/86 [07:22<00:59,  5.94s/it]
Batches:  90%|████████     |  77/86 [07:27<00:52,  5.80s/it]
Batches:  91%|████████     |  78/86 [07:34<00:47,  6.00s/it]
Batches:  92%|████████     |  79/86 [07:39<00:40,  5.77s/it]
Batches:  93%|████████     |  80/86 [07:45<00:34,  5.78s/it]
Batches:  94%|████████     |  81/86 [07:51<00:29,  5.89s/it]
Batches:  95%|████████     |  82/86 [07:56<00:22,  5.58s/it]
Batches:  97%|████████     |  83/86 [08:02<00:17,  5.86s/it]
Batches:  98%|████████     |  84/86 [08:08<00:11,  5.79s/it]
Batches:  99%|████████     |  85/86 [08:14<00:05,  5.92s/it]
Batches: 100%|████████     |  86/86 [08:16<00:00,  4.68s/it]
Batches: 100%|████████     |  86/86 [08:16<00:00,  5.77s/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

TRAINING 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: 0.7728181212524984

✓ COMPLETE: attribute_recognition → 77.3%

####################################################################
# 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,  1.83s/it]
Loading checkpoint shards:  40%|████      | 2/5 [00:03<00:05,  1.71s/it]
```
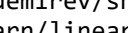
```
Loading checkpoint shards:  60%|██████     | 3/5 [00:05<00:03,  1.70s/it]
Loading checkpoint shards:  80%|███████    | 4/5 [00:06<00:01,  1.69s/it]
Loading checkpoint shards: 100%|█████████  | 5/5 [00:07<00:00,  1.24s/it]
Loading checkpoint shards: 100%|█████████  | 5/5 [00:07<00:00,  1.46s/it]
✓ Model loaded successfully


Batches:   0%|           | 0/26 [00:00<?, ?it/s]
Batches:   4%|▏          | 1/26 [00:06<02:35,  6.24s/it]
Batches:   8%|▎          | 2/26 [00:12<02:33,  6.38s/it]
Batches:  12%|▍          | 3/26 [00:17<02:07,  5.53s/it]
Batches:  15%|▌          | 4/26 [00:23<02:04,  5.67s/it]
Batches:  19%|▋          | 5/26 [00:29<02:01,  5.77s/it]
Batches:  23%|▊          | 6/26 [00:34<01:55,  5.78s/it]
Batches:  27%|██         | 7/26 [00:40<01:49,  5.76s/it]
Batches:  31%|██         | 8/26 [00:46<01:43,  5.75s/it]
Batches:  35%|███        | 9/26 [00:51<01:34,  5.53s/it]
Batches:  38%|███        | 10/26 [00:57<01:29,  5.62s/it]
Batches:  42%|████       | 11/26 [01:02<01:25,  5.67s/it]
Batches:  46%|████       | 12/26 [01:08<01:19,  5.68s/it]
Batches:  50%|█████      | 13/26 [01:13<01:12,  5.56s/it]
Batches:  54%|█████      | 14/26 [01:18<01:03,  5.33s/it]
Batches:  58%|██████     | 15/26 [01:24<00:59,  5.38s/it]
Batches:  62%|██████     | 16/26 [01:29<00:53,  5.38s/it]
Batches:  65%|███████    | 17/26 [01:35<00:48,  5.42s/it]
Batches:  69%|███████    | 18/26 [01:40<00:43,  5.44s/it]
Batches:  73%|███████    | 19/26 [01:46<00:38,  5.55s/it]
Batches:  77%|████████   | 20/26 [01:52<00:34,  5.67s/it]
Batches:  81%|████████   | 21/26 [01:57<00:27,  5.49s/it]
Batches:  85%|█████████  | 22/26 [02:03<00:22,  5.63s/it]
Batches:  88%|█████████  | 23/26 [02:09<00:16,  5.62s/it]
Batches:  92%|█████████  | 24/26 [02:15<00:11,  5.76s/it]
Batches:  96%|██████████ | 25/26 [02:20<00:05,  5.62s/it]
Batches: 100%|███████████| 26/26 [02:24<00:00,  5.21s/it]
Batches: 100%|███████████| 26/26 [02:24<00:00,  5.56s/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


TRAINING 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: 0.740909090909091

✓ COMPLETE: spatial_recognition → 74.1%

Final Results:
  object_detection        : 83.2%
  attribute_recognition   : 77.3%
```

```
    spatial_recognition        : 74.1%

    Average                     : 78.2%

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

  === Job finished at Fri Oct 24 12:34:14 CEST 2025 with exit code: 0 ===
```