

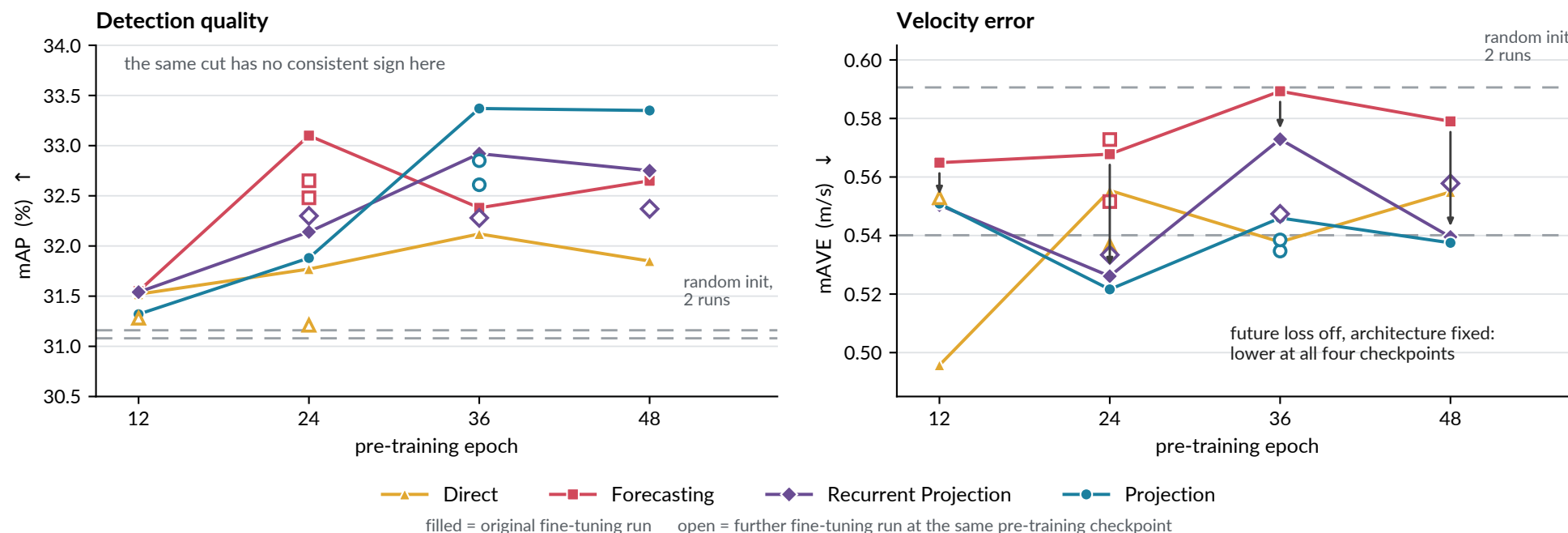
# What Does the Future Buy? – supplement

Full pre-training trajectories, additional fine-tuning executions, and the ONCE detection result.

Luca Brembilla · Politecnico di Milano

luca.brembilla@polimi.it

## Full four-branch trajectories, nuScenes D160



**What the poster shows instead.** After the main Projection–Forecasting result, the poster carries one compact control row: Recurrent Projection against Forecasting on NDS versus cumulative pre-training compute (nuScenes) and on AP@50 across pre-training doses (ONCE), then Projection – Direct on  $\Delta$  mAP. The  $\Delta$  NDS series for Projection – Direct is not on the sheet; its values are  $-0.85$ ,  $+0.11$ ,  $+1.35$ ,  $+1.43$  pp at epochs 12/24/36/48, the same pattern as  $\Delta$  mAP. These trajectories are the full, per-execution version of all of it.

**Replication status.** One pre-training run per branch. The four checkpoints are four points on a single trajectory, not four independent pre-training replications. Every open marker above is a further fine-tuning execution at a fixed pre-training checkpoint, not a further pre-training run, and they are never pooled as independent samples.

**Protocol.** Pre-training on nuScenes; the sparse-convolutional encoder transfers to TransFusion-L. Fine-tuning on D160 – every 160th training sample, 176 labelled samples before balanced resampling – with global batch 4 on a  $16\times$  schedule, evaluated at fine-tuning epoch 320. Random initialisation under the same recipe gives the two dashed references.

**Readings.** On detection the matched future-loss cut has no consistent sign across the four checkpoints, and the branches trade places. At later checkpoints the branches with an added transformation are above Direct. On velocity error the matched cut is lower at all four evaluated checkpoints; the arrows mark it.

**Exclusion.** The Projection sibling scoring 33.13 mAP at pre-training epoch 36 is excluded everywhere: it shared fine-tuning with the epoch-36 original through epoch 280 and is not an independent execution.

Every plotted value — original fine-tuning execution at each checkpoint

| pre-training epoch   | mAP (%)      |       |       |       | NDS (%)      |       |       |       | mAVE (m/s)     |        |        |        |
|----------------------|--------------|-------|-------|-------|--------------|-------|-------|-------|----------------|--------|--------|--------|
|                      | 12           | 24    | 36    | 48    | 12           | 24    | 36    | 48    | 12             | 24     | 36     | 48     |
| Forecasting          | 31.55        | 33.10 | 32.38 | 32.65 | 45.06        | 46.32 | 45.80 | 45.50 | 0.5649         | 0.5678 | 0.5893 | 0.5790 |
| Recurrent Projection | 31.54        | 32.14 | 32.92 | 32.75 | 45.38        | 45.97 | 46.13 | 46.65 | 0.5504         | 0.5261 | 0.5729 | 0.5395 |
| Projection           | 31.32        | 31.88 | 33.37 | 33.35 | 45.28        | 46.15 | 47.09 | 46.71 | 0.5509         | 0.5216 | 0.5460 | 0.5375 |
| Direct               | 31.52        | 31.77 | 32.12 | 31.85 | 46.13        | 46.04 | 45.74 | 45.28 | 0.4957         | 0.5554 | 0.5378 | 0.5549 |
| random init, 2 runs  | 31.16, 31.08 |       |       |       | 45.11, 44.49 |       |       |       | 0.5401, 0.5906 |        |        |        |

Registry: TREND/ablations/results\_registry/nuscenes\_d160\_results.csv. For Forecasting and Direct the original execution is the seedless row; for Recurrent Projection and Projection it is fine-tuning seed 666.

### Additional fine-tuning executions

| branch               | pre-training epoch | mAP (%) | NDS (%) | mAVE (m/s) |
|----------------------|--------------------|---------|---------|------------|
| Direct               | 12                 | 31.28   | 45.23   | 0.5529     |
| Direct               | 24                 | 31.21   | 45.56   | 0.5365     |
| Forecasting          | 24                 | 32.65   | 45.86   | 0.5728     |
| Forecasting          | 24                 | 32.48   | 46.23   | 0.5516     |
| Recurrent Projection | 24                 | 32.30   | 46.23   | 0.5334     |
| Recurrent Projection | 36                 | 32.28   | 45.90   | 0.5474     |
| Recurrent Projection | 48                 | 32.37   | 46.26   | 0.5578     |
| Projection           | 36                 | 32.85   | 46.57   | 0.5347     |
| Projection           | 36                 | 32.61   | 46.43   | 0.5385     |

Same pre-training checkpoints as the table above, different fine-tuning executions (seed 42, and same-seed reruns). Open markers in the trajectory plot.

### The two left-hand control panels, value by value

|                                          | pre-training epoch   |                      |                      |                      |
|------------------------------------------|----------------------|----------------------|----------------------|----------------------|
|                                          | 12                   | 24                   | 36                   | 48                   |
| Forecasting – NDS (%)                    | 45.06 <sup>n=1</sup> | 46.09 <sup>n=2</sup> | 45.80 <sup>n=1</sup> | 45.50 <sup>n=1</sup> |
| Recurrent Projection – NDS (%)           | 45.38 <sup>n=1</sup> | 46.10 <sup>n=2</sup> | 46.02 <sup>n=2</sup> | 46.45 <sup>n=2</sup> |
| Forecasting – GPU-h, measured            | 44                   | 89                   | 137                  | 186                  |
| Recurrent Projection – GPU-h, calibrated | 30                   | 60                   | 89                   | 119                  |
| Direct – mAP (%)                         | 31.40 <sup>n=2</sup> | 31.49 <sup>n=2</sup> | 32.12 <sup>n=1</sup> | 31.85 <sup>n=1</sup> |
| Projection – mAP (%)                     | 31.32 <sup>n=1</sup> | 31.88 <sup>n=1</sup> | 32.99 <sup>n=2</sup> | 33.35 <sup>n=1</sup> |

### ONCE D5 – CenterPoint, AP@50 (%), three fine-tuning seeds per cell

| branch               | PT ep 12     | PT ep 24     | PT ep 36               |
|----------------------|--------------|--------------|------------------------|
| Forecasting          | 58.30 ± 0.84 | 58.24 ± 0.42 | 57.91 <sup>1 run</sup> |
| Recurrent Projection | 58.48 ± 0.53 | 58.25 ± 0.02 | 58.18 ± 0.41           |
| Projection           | 58.09 ± 0.28 | 58.21 ± 0.54 | 57.76 ± 0.49           |
| random init, 1 run   | 58.43        |              |                        |

Seeds 666 / 42 / 2026; mean ± sample SD over the three seeds. At epoch 36 the 2026-08-01 jobs named `seed42` / `seed2026` were mislabelled – the launcher ignored `RANDOM_SEED` – so they are same-seed 666 executions and were collapsed to their mean before averaging across seeds; only the 2026-08-02 `true_seed` jobs are genuinely distinct seeds. Random initialisation is a single run at each fine-tuning length: 58.13 at 200 epochs, 58.43 at 320, 57.82 at 400. **No branch separates from the others or from random initialisation.** Overlapping spreads are not an equivalence test.

**The baseline drawn on the poster.** The poster’s ONCE panel uses the random-init baseline *as published for this subset*,  $57.29 \pm 0.29$  over 5 seeds (paper Table 7). Read it with care: `TREND/ablations/MAIN.md` finds that published baseline under-converged – our own scratch on the same split reaches 58.43 at 320 fine-tuning epochs, about 1.1 above it, and the analysis there attributes the published pre-training gain on this subset to that shortfall rather than to pre-training. The gap between our arms and the published line should therefore not be read as a pre-training gain. ONCE AP@50 carries no velocity term, so this result speaks only to detection and says nothing about the nuScenes mAVE finding – the central result is a single-dataset nuScenes finding. Registry: `once_d5_ft_results.csv`.

NDS is the mean over the fine-tuning executions at that pre-training checkpoint, same-seed reruns excluded; <sup>n</sup> gives how many executions each mean is over. GPU-h provenance is below. On ONCE the poster plots the cell means from the table above; every cell is a three-seed mean except Forecasting at epoch 36, which is a single run (57.91) — the poster draws it like the others, and the per-cell seed count is recorded here.

## Compute provenance — how each GPU-hour figure on the poster was obtained

The poster prints compute figures without distinguishing measurement from estimate, to keep the comparison readable. The distinction is here.

**Measured** — per-iteration timestamps, 4×A100-64GB, global batch 24, 1114 iterations per epoch, one uninterrupted job each: *Forecasting* 186.2 GPU-h to pre-training epoch 48 (44.2 / 88.6 / 137.0 / 186.2 at epochs 12/24/36/48); *Direct* 91.9 GPU-h; the width-64 projector variant 94.8 GPU-h.

**Calibrated estimate** — hours = epochs × (iters × s/iter + 285 s)/3600, validated to ≈ 2% low on stationary arms: the *Projection* objective 98 GPU-h to epoch 48, and the *Recurrent Projection* objective 119 GPU-h (30 / 60 / 89 / 119 at epochs 12/24/36/48).

**Why the estimates are needed.** The launched Projection and Recurrent Projection runs both declared three input frames and rendered a future target at loss weight zero. The zero weight means no gradient, but the wall clock is real, so those jobs did not realise the saving their objective implies and their own runtimes do not cost their objectives. The estimates above are for the corrected runtimes of the same objectives (pt16 and pt17), which were never trained — so the compute axis in the poster's left-hand control panel pairs measured accuracy with a calibrated cost on the Recurrent Projection side. Sources: TREND/ablations/timing/nuscenes\_pretrain\_checkpoints.csv and TREND/ablations/timing/README.md.