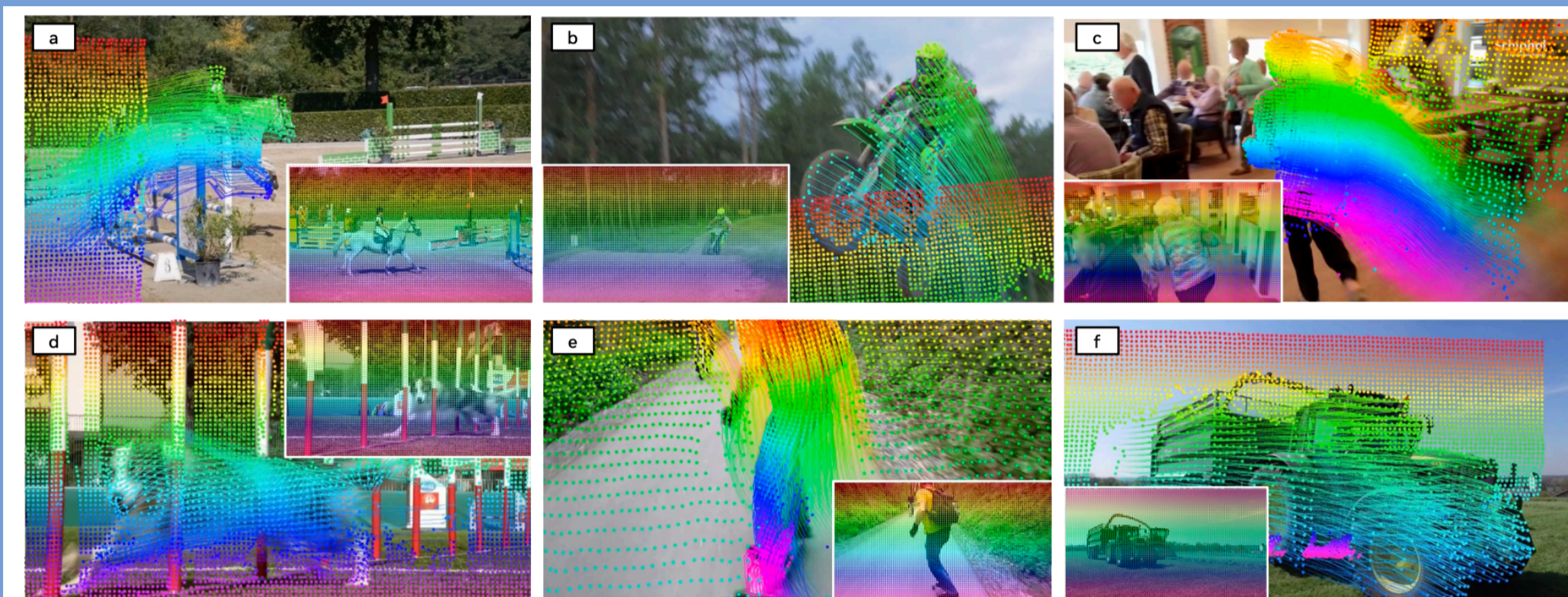
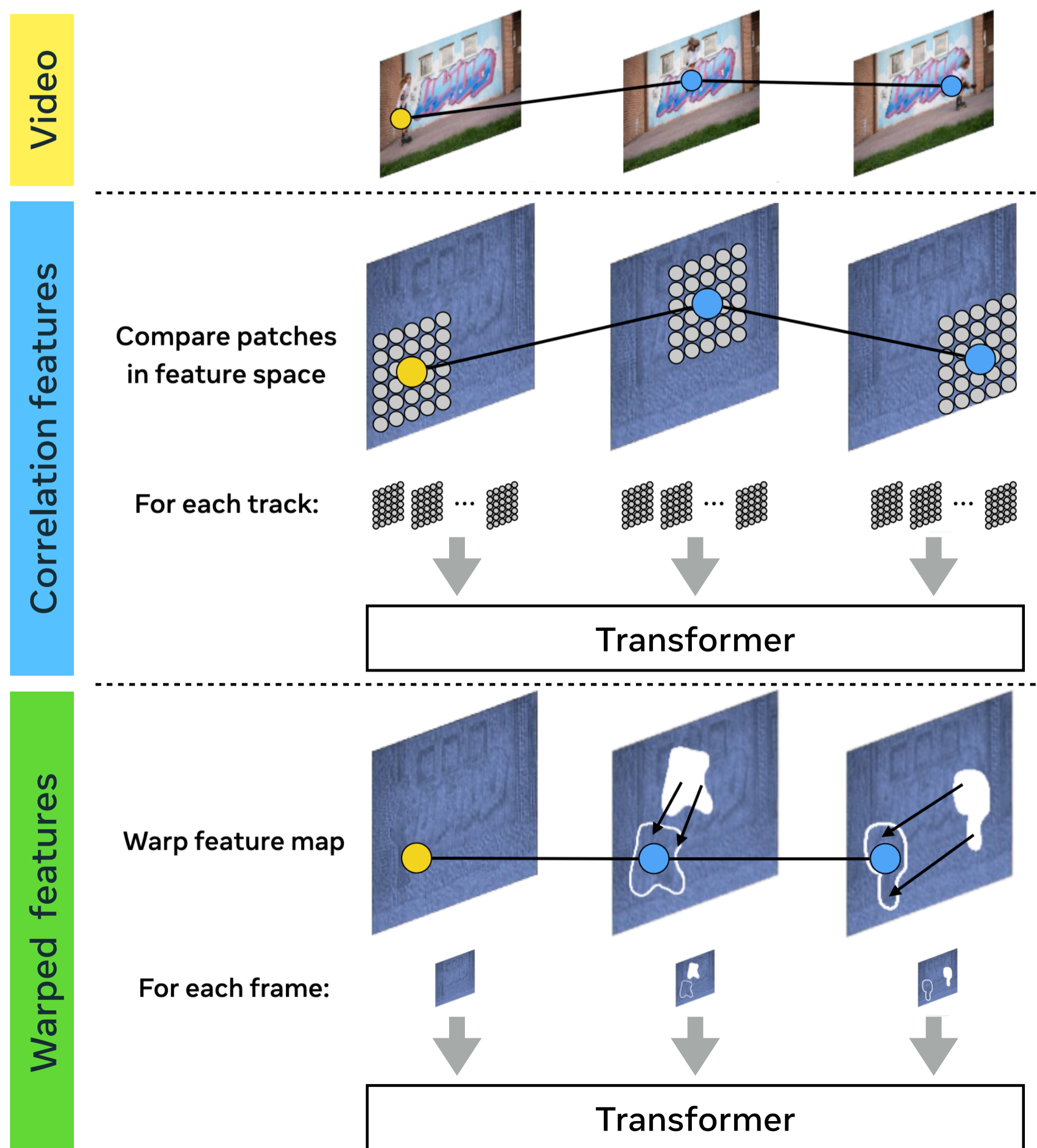


CoWTracker Results

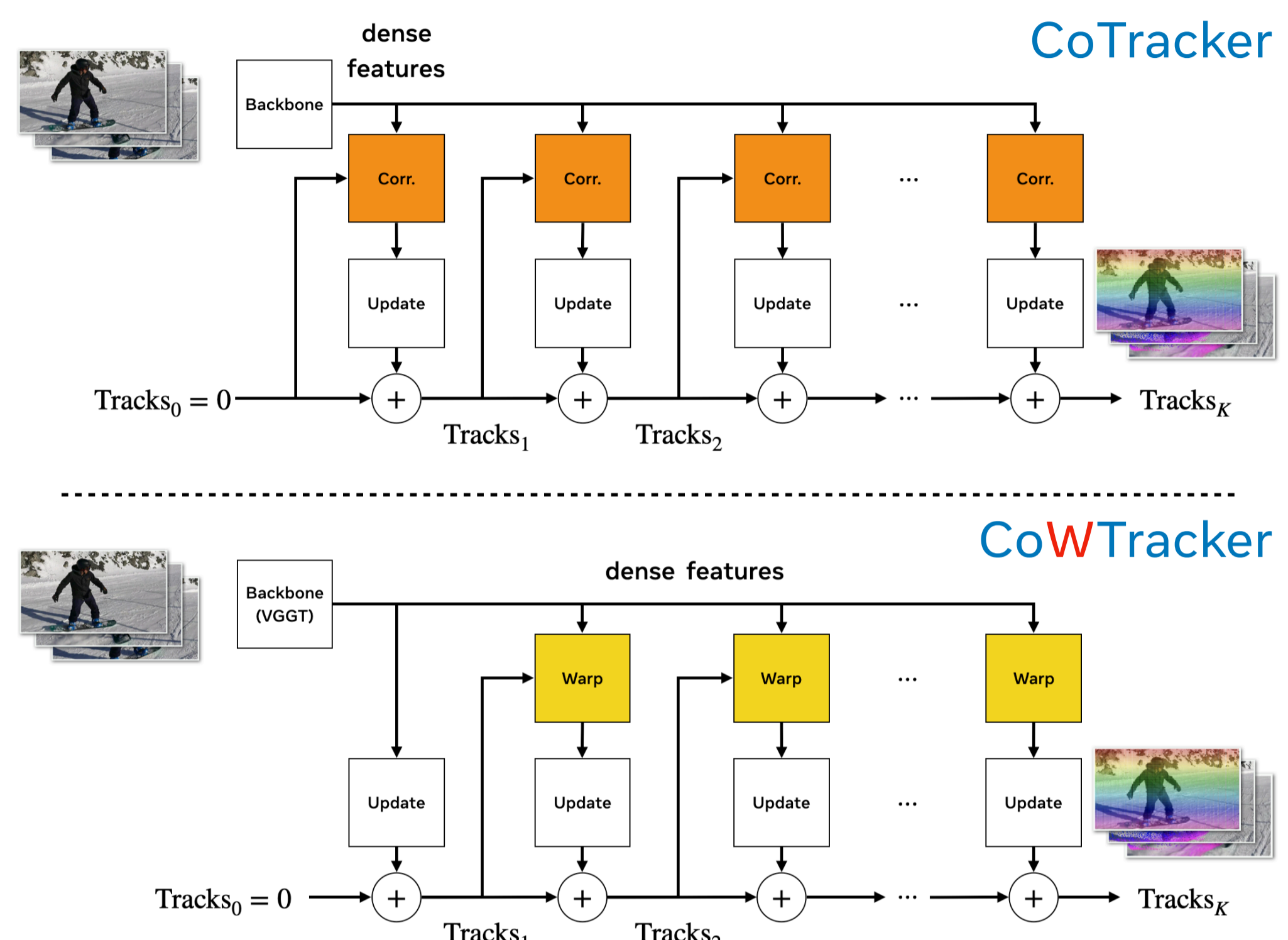


CoWTracker takes a video as input and outputs **dense per-pixel trajectories** over time.

From Correlation to Warping

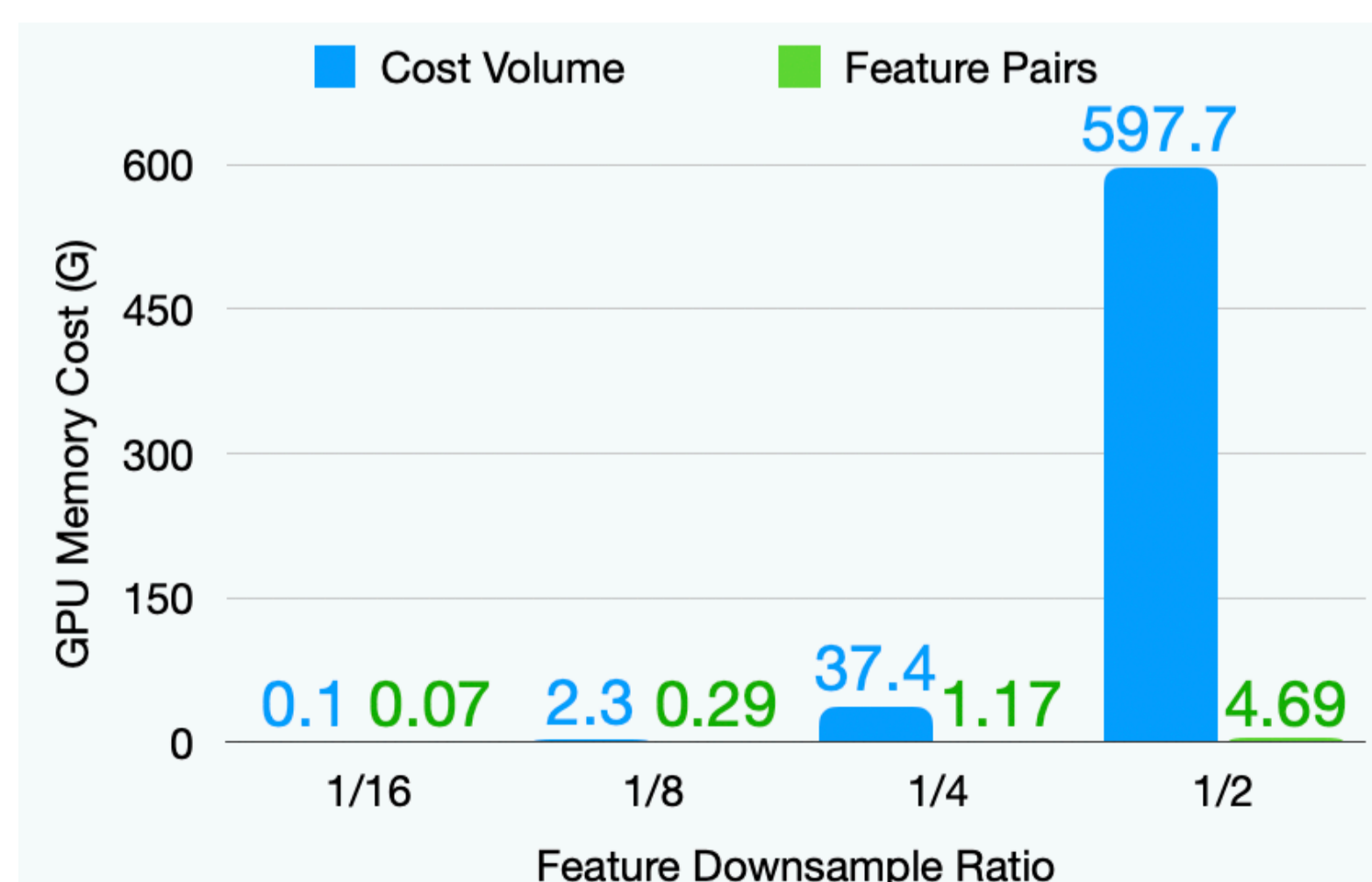


From CoTracker to CoWTracker



CoWTracker **replaces Correlation with Iterative Warping**:
 Simpler, Memory-efficient, and Accurate

Warping is Memory Efficient



The main advantage is to **dramatically reduce the GPU memory required**, particularly for high-resolution features.

Point Tracking Results

TAP-Vid & RoboTAP Benchmarks (Average Jaccard ↑)

Method	DAVIS	RGB-S	RoboTAP	Kinetics	Mean AJ
Sparse Trackers					
TAPIR	56.2	55.5	49.6	49.6	52.7
CoTracker3	64.4	74.3	54.7	54.7	62.0
BootsTAPIR	61.4	70.8	54.6	54.6	60.4
Dense Trackers					
DELTA	60.8	72.2	60.2	51.0	61.1
AllTracker	63.3	81.1	71.9	59.1	68.9
CoWTracker	65.5	85.4	73.2	60.9	71.3

+2.4
 Mean AJ vs. AllTracker

+9.3
 Mean AJ vs. CoTracker3

+10.9
 Mean AJ vs. BootsTAPIR

Zero-Shot Optical Flow

Same model, no flow-specific training. EPE (↓ better):

Method	Sintel Clean	Sintel Final	KITTI EPE	KITTI Fl-all
RAFT	1.15	1.86	1.53	7.81
SEA-RAFT	0.97	1.96	1.60	8.26
WAFT	0.94	2.09	1.15	5.29
CoWTracker	0.78	1.48	1.04	4.87

17%–20% relative EPE reduction on Sintel vs. best specialized flow model, despite zero flow-specific training.

Zero-Shot Optical Flow Qualitative Results



Optical flow on Sintel using the same tracking model. Predicted flows closely match ground truth even under large motion, occlusions, and clutter.

Point Tracking Qualitative Comparison



CoWTracker maintains accurate tracks along the front segment and rails despite occlusions and boundaries.

Conclusions

- CoWTracker replaces cost volumes with **iterative warping** — simpler, more memory-efficient, and more accurate.
- Achieves **SOTA on 4 tracking benchmarks** and competitive **zero-shot optical flow** on 3 benchmarks.
- Demonstrates that **warping-based architectures can unify** dense point tracking and optical flow estimation.
- The design is a stepping stone for future research bridging tracking, optical flow, and other correspondence tasks.

