

How Hard is Snow? A Paired Domain Adaptation Dataset for Clear and Snowy Weather: CADC+

Mei Qi Tang, Chengjie Huang, Sean Sedwards, Krzysztof Czarnecki



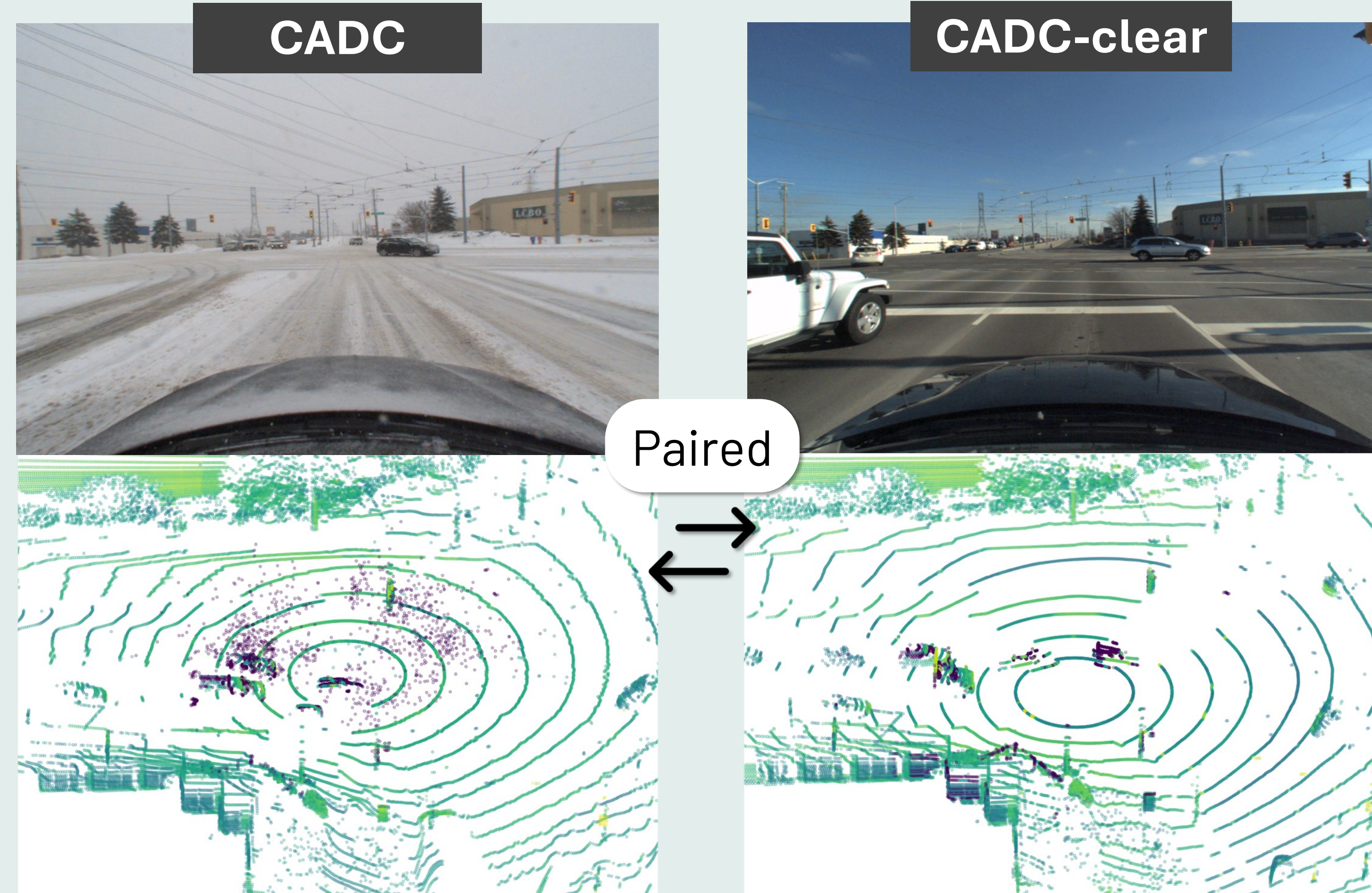
CADC+ Dataset

Motivation

- Underexplored **impact of snow on 3D object detections**
- Lack of** existing **paired weather datasets**
- Synthetic datasets (e.g., created using de-snowing) often **lack realism** and introduce **confounding factors** that hinder accurate evaluation

Contributions

- First** multi-modal **paired** domain adaptation dataset in clear and snowy weather
- 74 sequence pairs** recorded in the **same season**, on **similar roads**, using the **same vehicle platform**
- 3D bounding box** annotations for 3D object detection



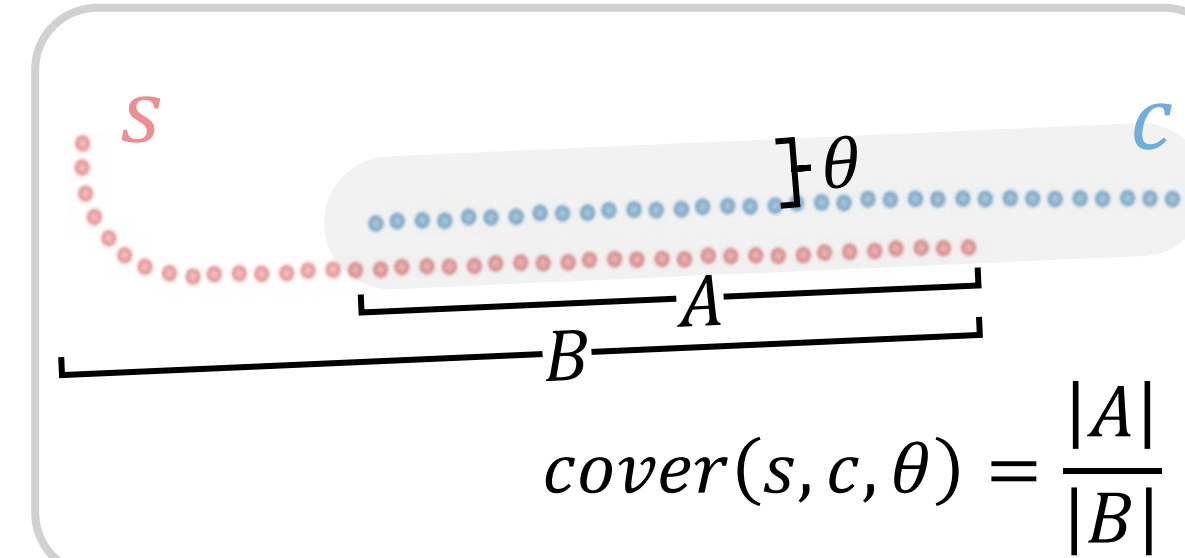
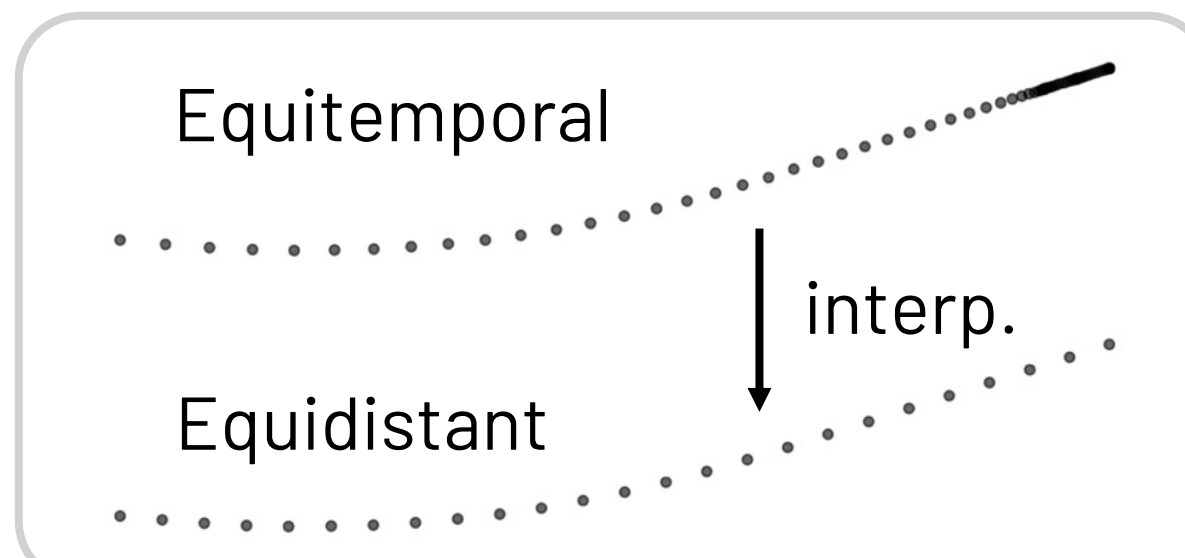
Sequence Matching

Goal: minimize spatial and temporal differences between two sequences

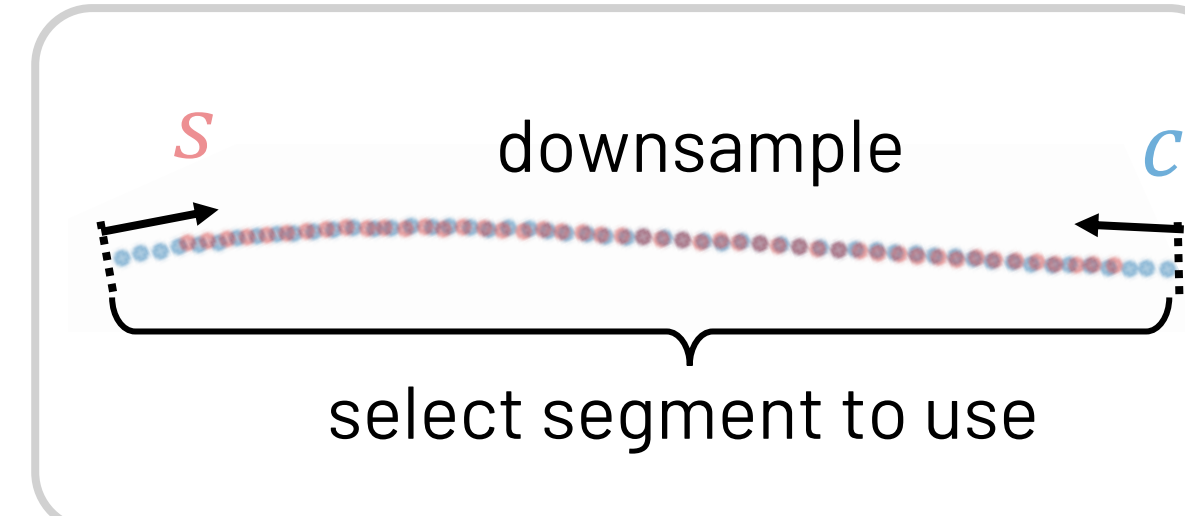
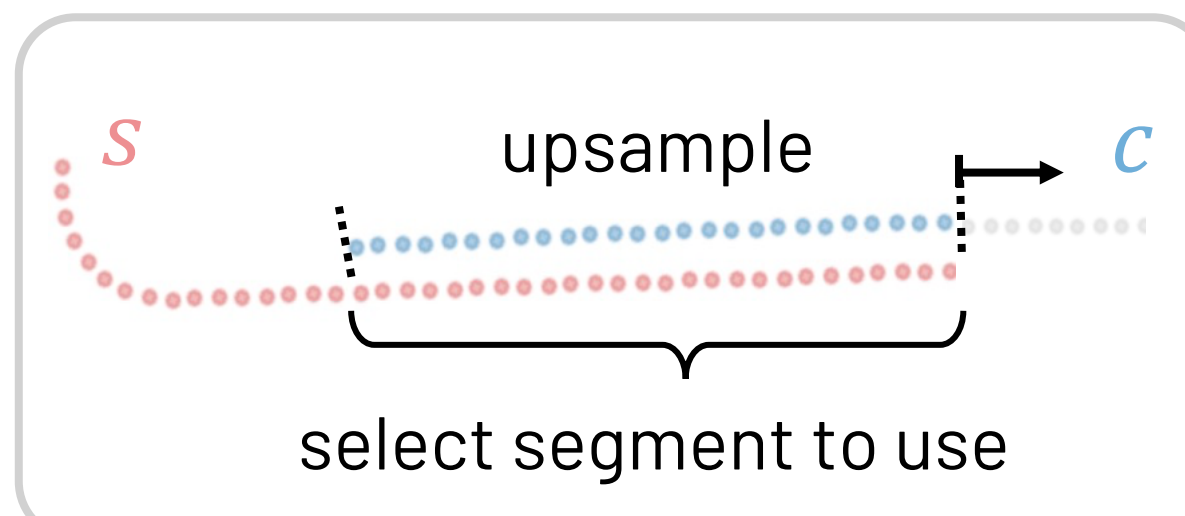
Matching preferences:

- Same lane
- Different lane but same direction
- Any lane on the same road
- Roads with partial coverage
- Similar road type and traffic density

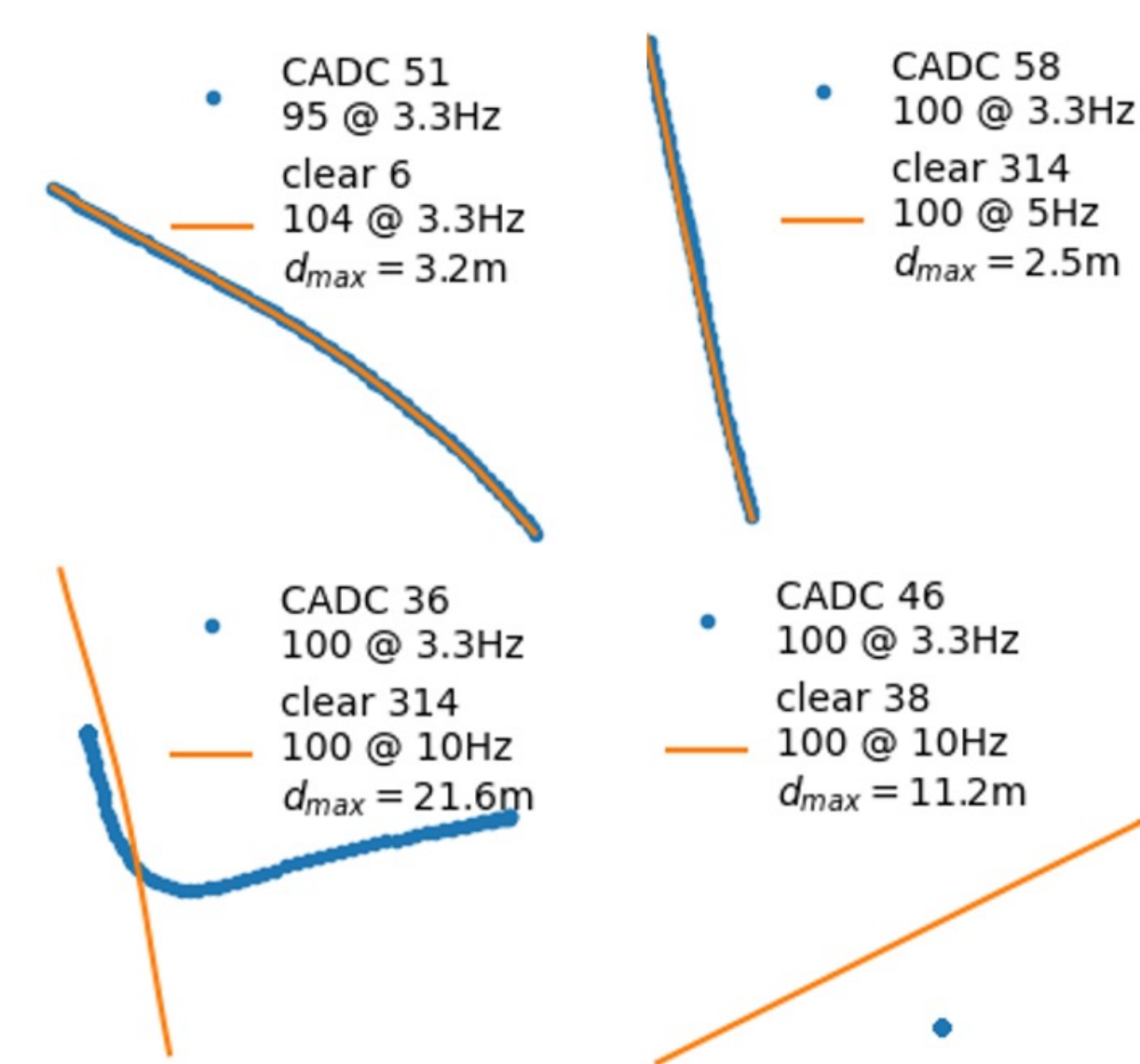
- Interpolate spatial models
- Compute coverage ratios



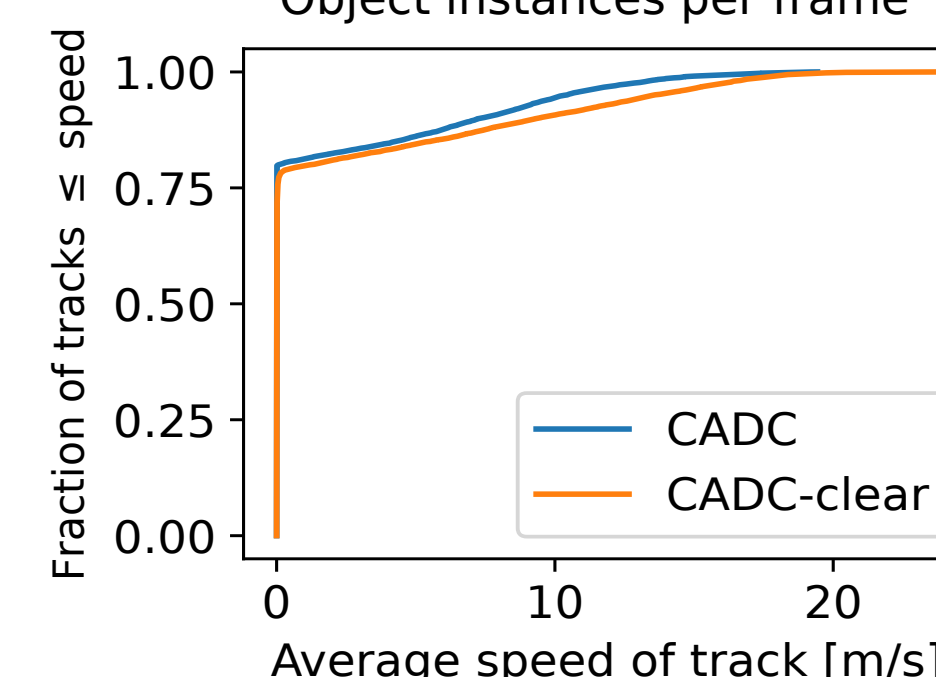
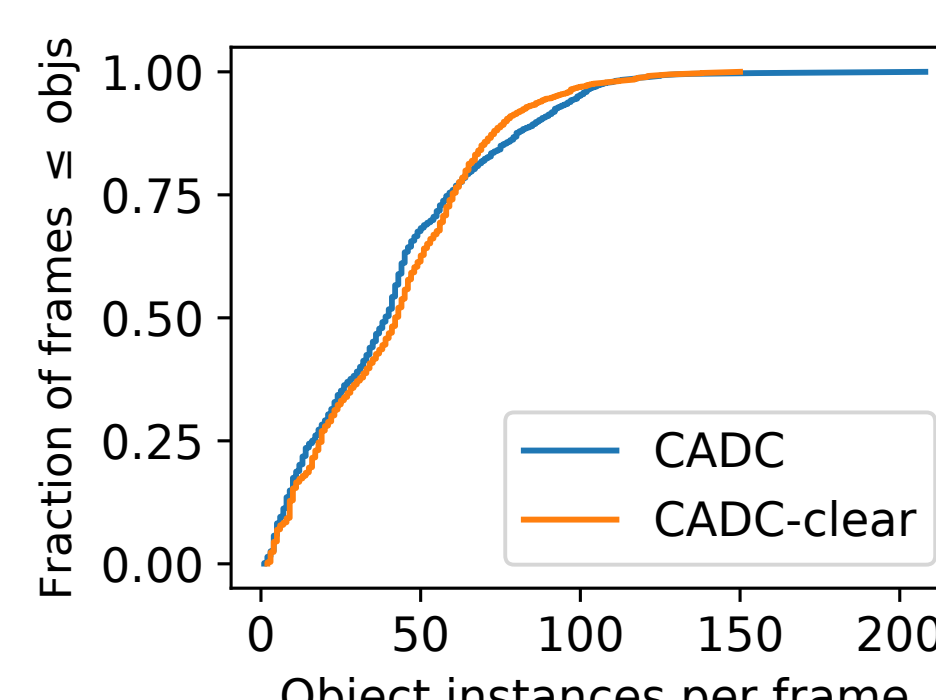
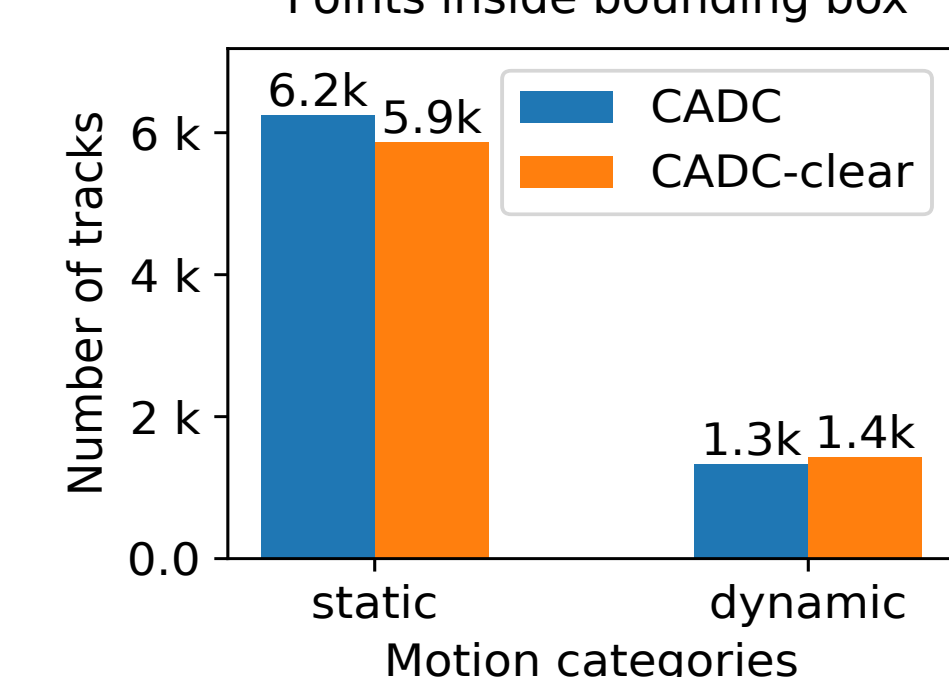
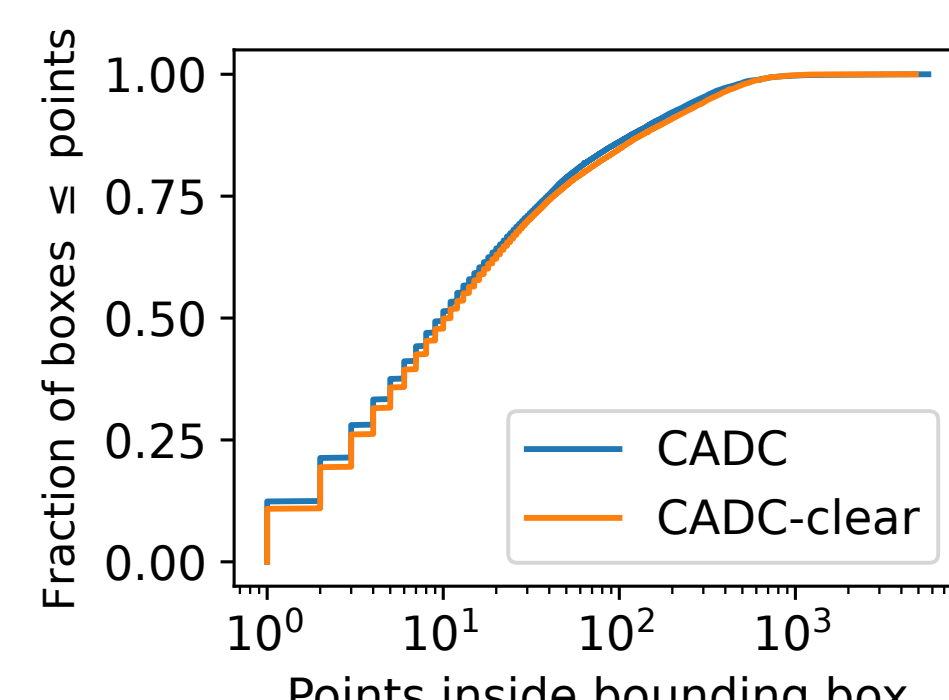
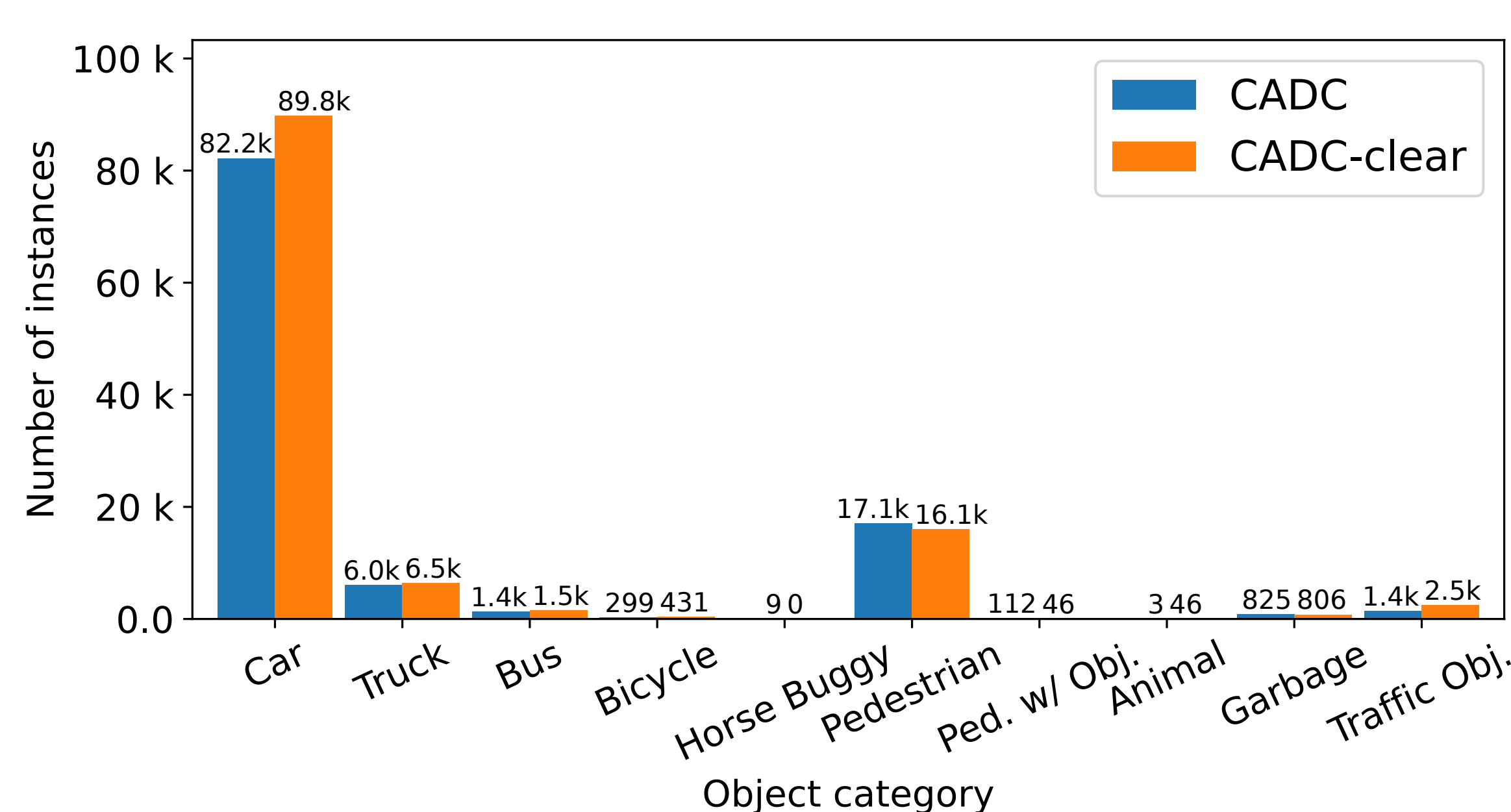
- Choose sampling frequency and endpoints



Examples



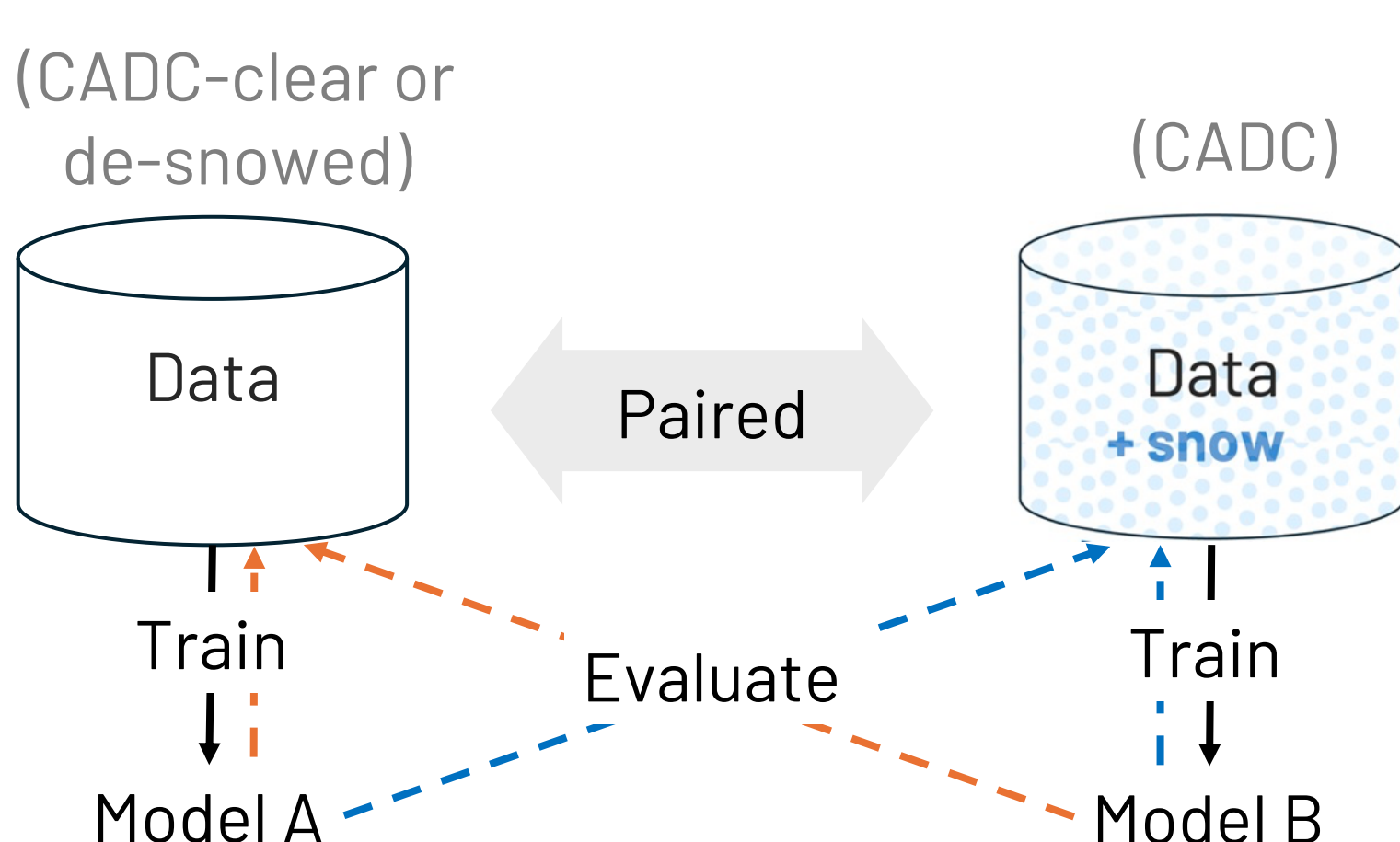
Statistics



- Sequences are **≥ 100** frames
- 53 / 74 sequence pairs overlap spatially
- Training sequences are labelled on every 10th frame, but the **validation set is fully labelled** to ensure accurate evaluation

Experiments: How Hard is Snow?

Setup

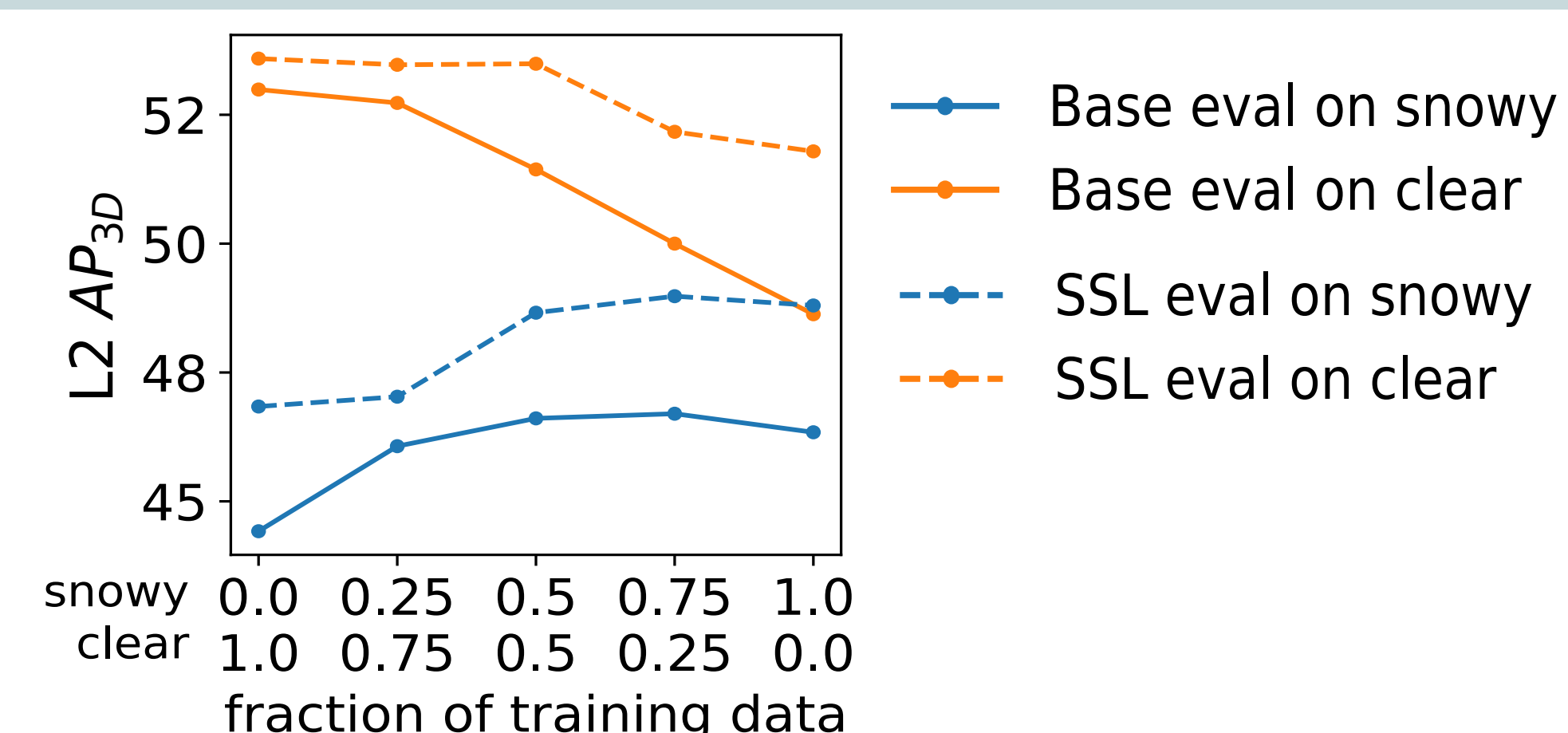


- Downsample CADC labels to match CADC-clear
- Create data mixtures of varying amount of snowy data
- Train Base and SSL models using sparse and pseudo labels¹

Impact of Snow on 3D Object Detection

- Snowy objects are more **difficult to detect**
- Snowy objects are a **distinct domain**
- Training with **balanced** snowy and clear data performs best

¹Models use the VoxelNeXt architecture. The Base model trains only on sparse labels. The SSL model trains using high quality pseudo labels that Base model produces, via semi-supervised learning.



Effectiveness of Synthetic Clear Weather Data Generation via De-snowing

Training data	Base model				SSL model			
	Eval on snowy		Eval on clear		Eval on snowy		Eval on clear	
clear	44.42	Δ%	52.99	Δ%	46.84	Δ%	53.59	Δ%
de-snowed _{DROR}	42.93	-1.49	46.26	-6.73	47.44	0.60	49.74	-3.85
de-snowed _{LiSnowNet}	44.44	0.02	48.13	-4.86	48.36	1.51	51.33	-2.27
snowy	46.34	1.92	48.63	-4.36	48.80	1.96	51.79	-1.80

Δ% indicates the difference w.r.t. the clear model. Evaluation results show Waymo OD's L2 AP_{3D} (%) metric.

- Low quality de-snowing is more detrimental than helpful
- De-snowing is **not a realistic** way of generating clear weather data