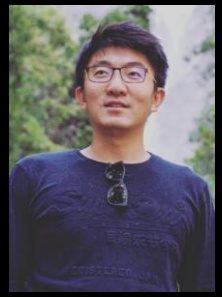
Traffic4D: Single View Reconstruction of Repetitious Activity Using Longitudinal Self-Supervision

Fangyu Li, N Dinesh Reddy, Xudong Chen, Srinivasa G. Narasimhan

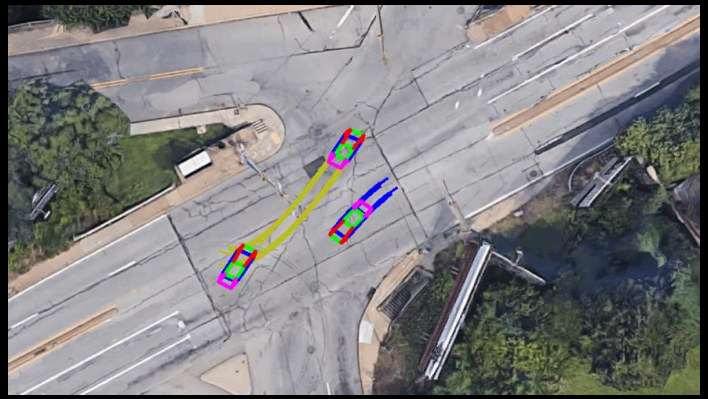




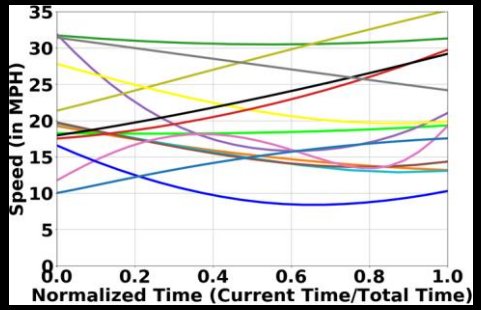

Carnegie Mellon University

NVIDIA

Motivation: Traffic in 4D



Traffic Activity
4D = 3D Real World + Time



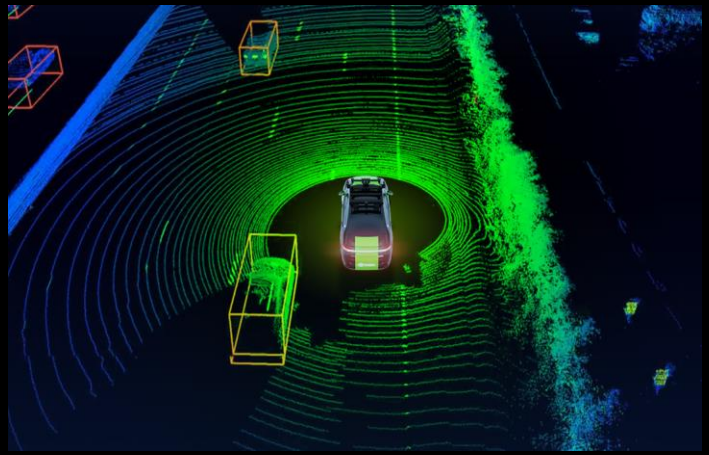
Real World Velocity

Locating Traffic Anomaly

Driving Assistance

Traffic Monitoring

Challenge 1: Sensors at City Scale



Depth Sensors

- Direct 3D information
- Impractical for large scale

Traffic Surveillance Cameras

- Everywhere, anytime
- Need to reconstruct 3D from 2D

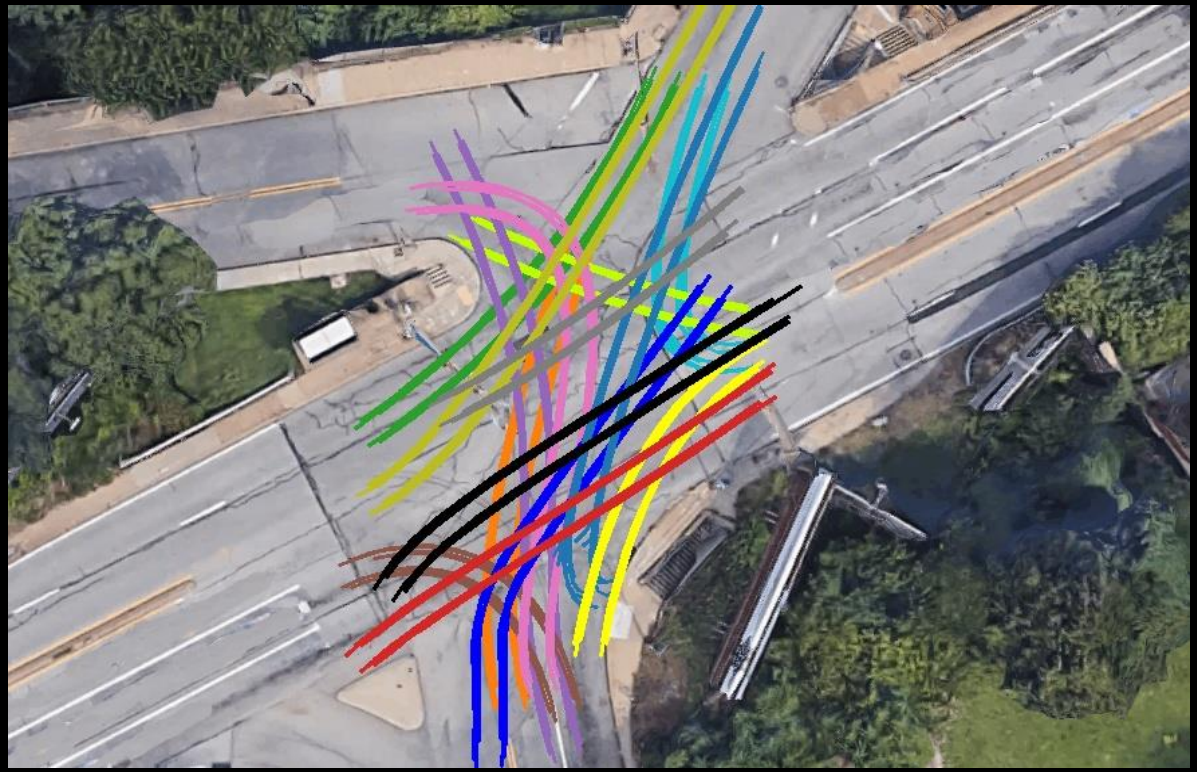
Challenge 2: Stable & Accurate Algorithms

Failed detection

Spatially inaccurate, temporally inconsistent reconstruction

Longitudinal Information in Repetitious Activity

Longitudinal Information in Repetitious Activity



2D & 3D Improvement With Longitudinal Information

Initial Keypoints before Longitudinal Self-Supervision

Initial 2D Keypoints

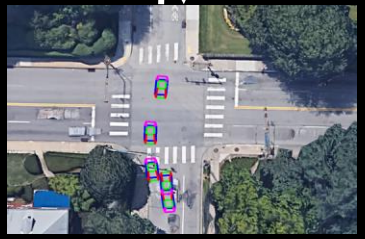
Keypoints: HRNet*

* J. Wang, K. Sun, T. Cheng, B. Jiang, C. Deng, Y. Zhao, D. Liu, Y. Mu,M. Tan, X. Wang,et al., “Deep high-resolution representation learning for visual recognition,” TPAMI, 2020

Initial Keypoints before Longitudinal Self-Supervision

Initial 2D Keypoints

Reproject **Reconstruct**



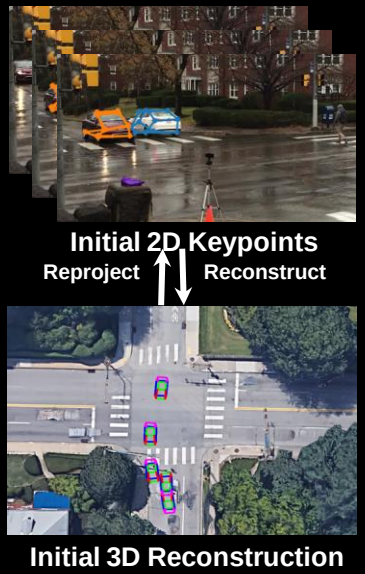
Initial 3D Reconstruction

To solve:
3D Shape and Pose
for *all* vehicles

→

Based on:
Long term geometric
constraints

Initial Keypoints before Longitudinal Self-Supervision



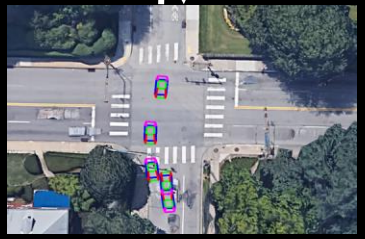
Enforce rigidity over frames:
Multi-Object Tracker V-IOU*

* E. Bochinski, T. Senst, and T. Sikora, “Extending iou based multi-object tracking by visual information,” in IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2018.

Initial Keypoints before Longitudinal Self-Supervision

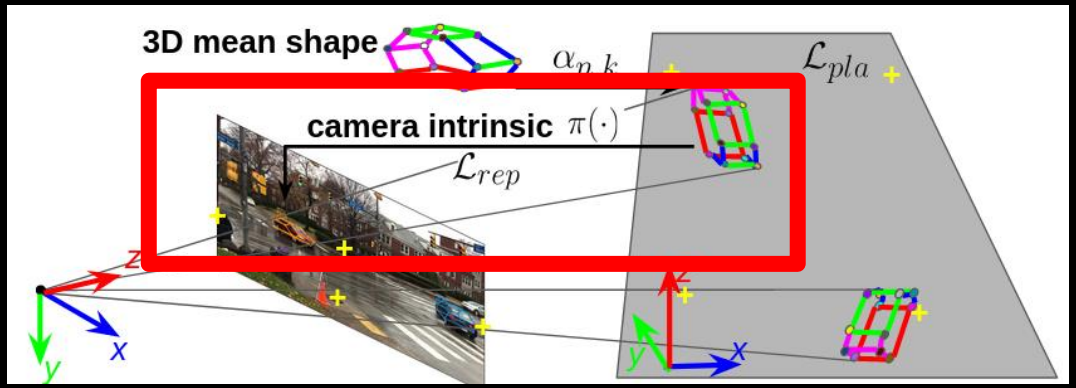
Initial 2D Keypoints

Reproject ↑↓ Reconstruct



Initial 3D Reconstruction

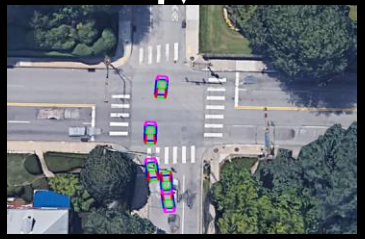
Reprojection Loss:
Project 3D keypoints -> 2D detected keypoints on image



Initial Keypoints before Longitudinal Self-Supervision

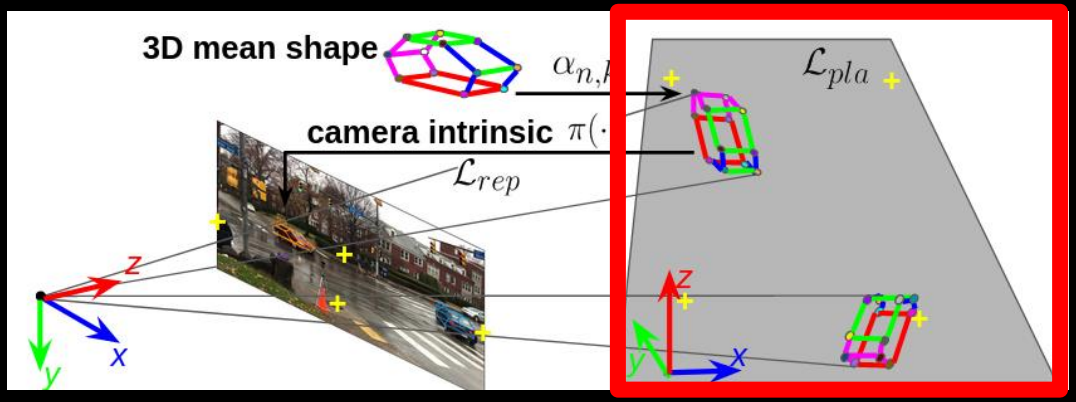
Initial 2D Keypoints

Reproject ↑↓ Reconstruct

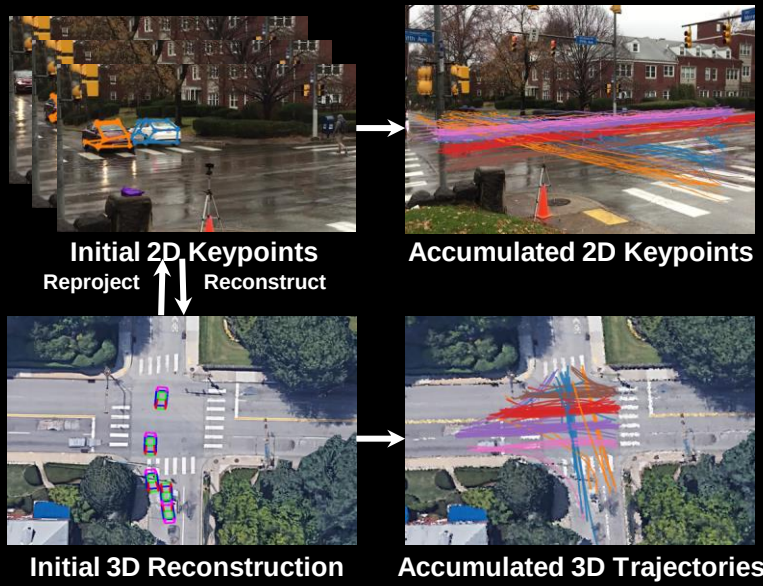


Initial 3D Reconstruction

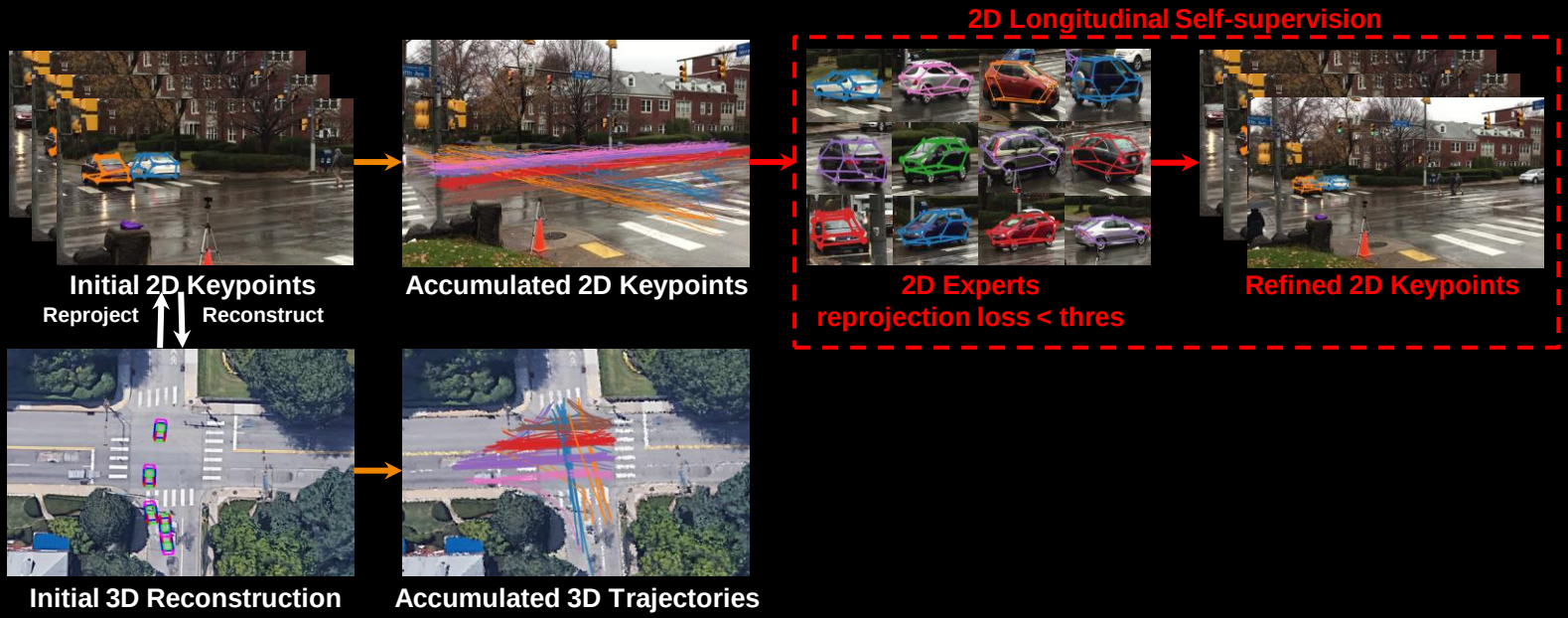
Joint Planar Loss:
All vehicles in the *long term* video
as close as possible to a plane



2D Experts to Improve Detection



2D Experts to Improve Detection



2D Experts to Improve Detection

Initial Erroneous Detection

2D Experts to Improve Detection

Initial Erroneous Detection

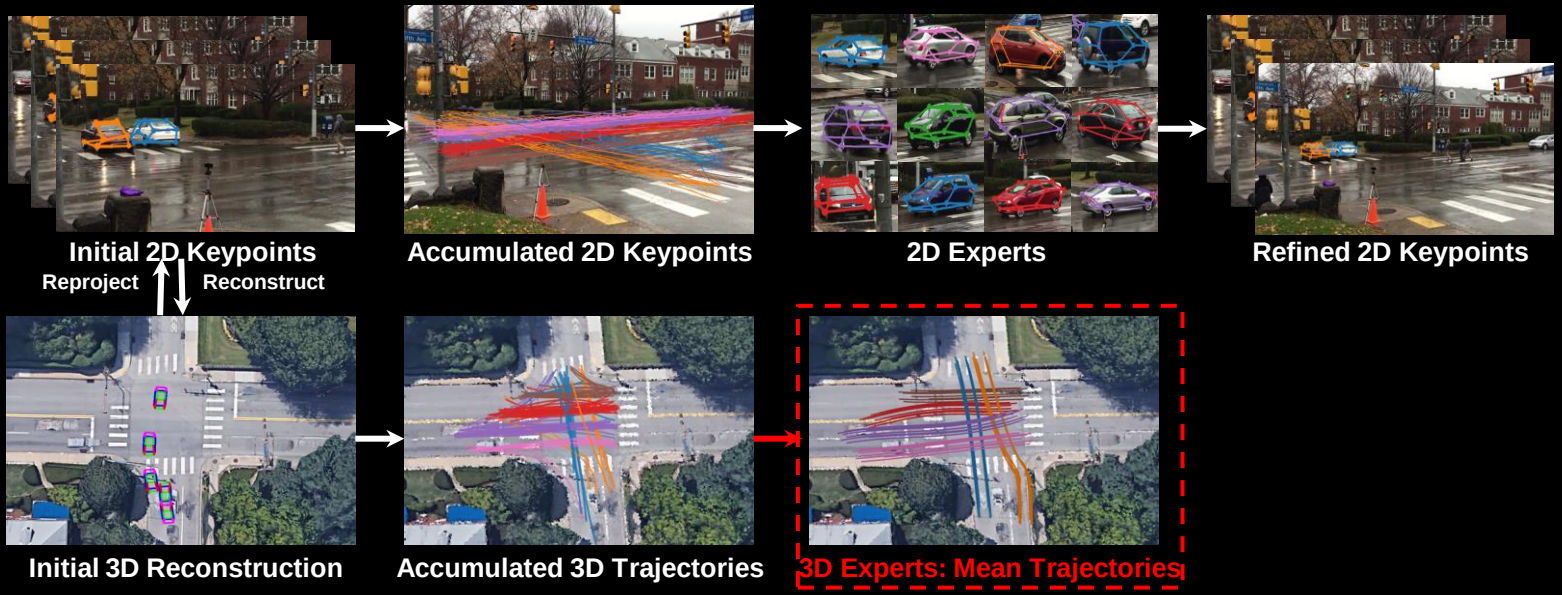
Nearest Neighbor with Bounding Boxes

Refined Detection

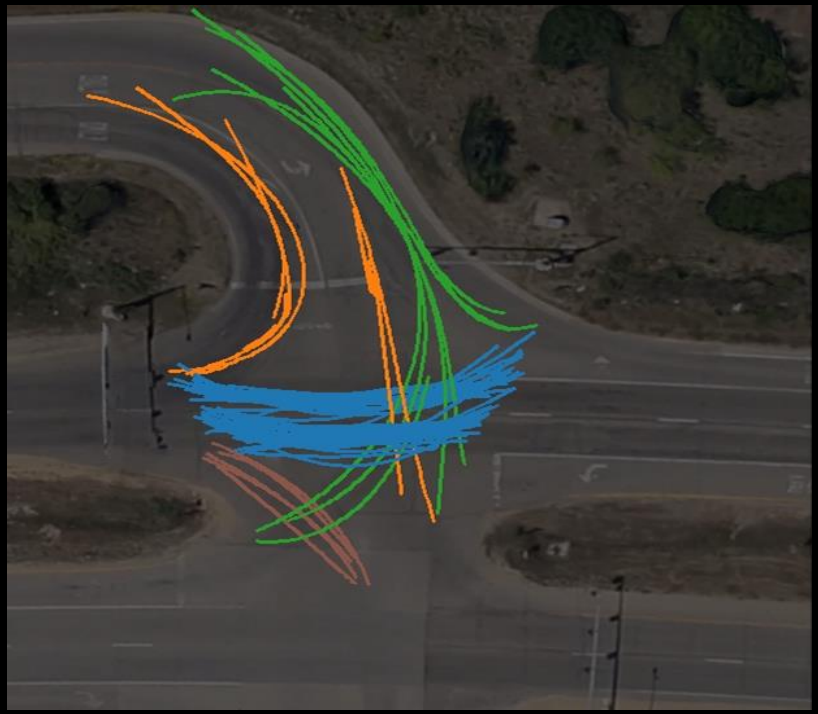
2D Experts to Improve Detection

Refined Detection

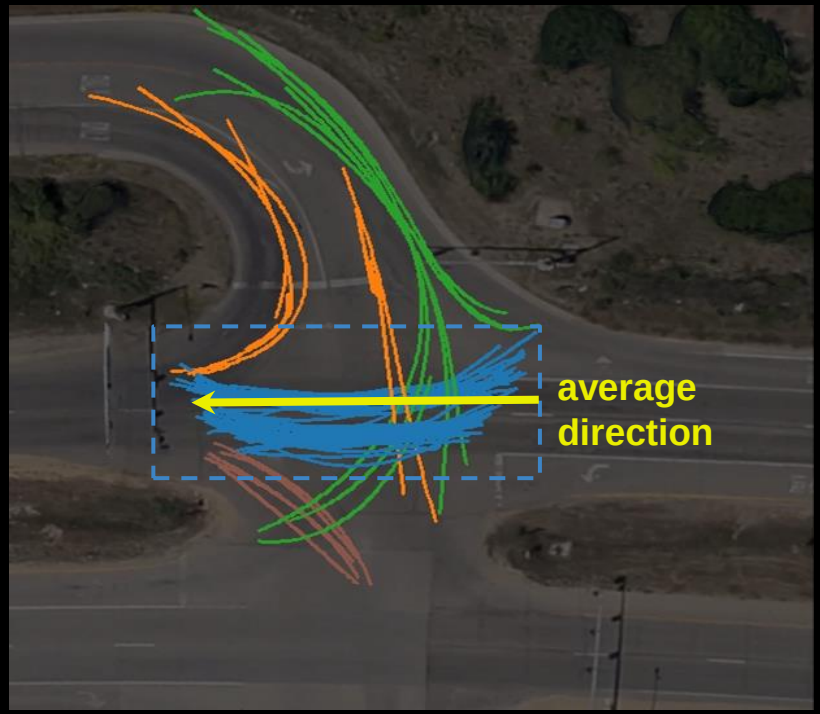
Learning 3D Experts: Hierarchical Clustering



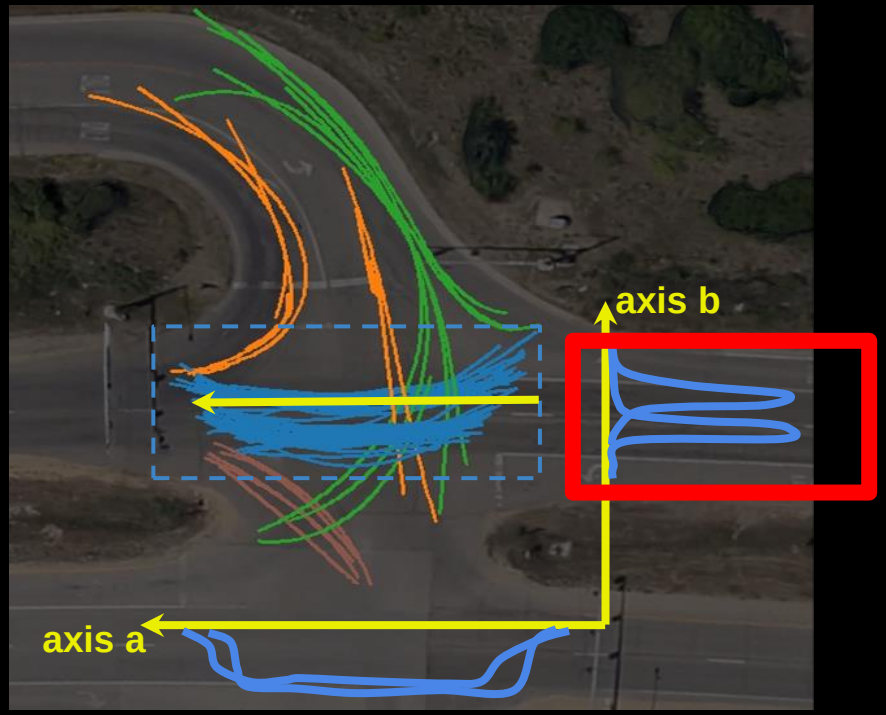
Learning 3D Experts: Hierarchical Clustering



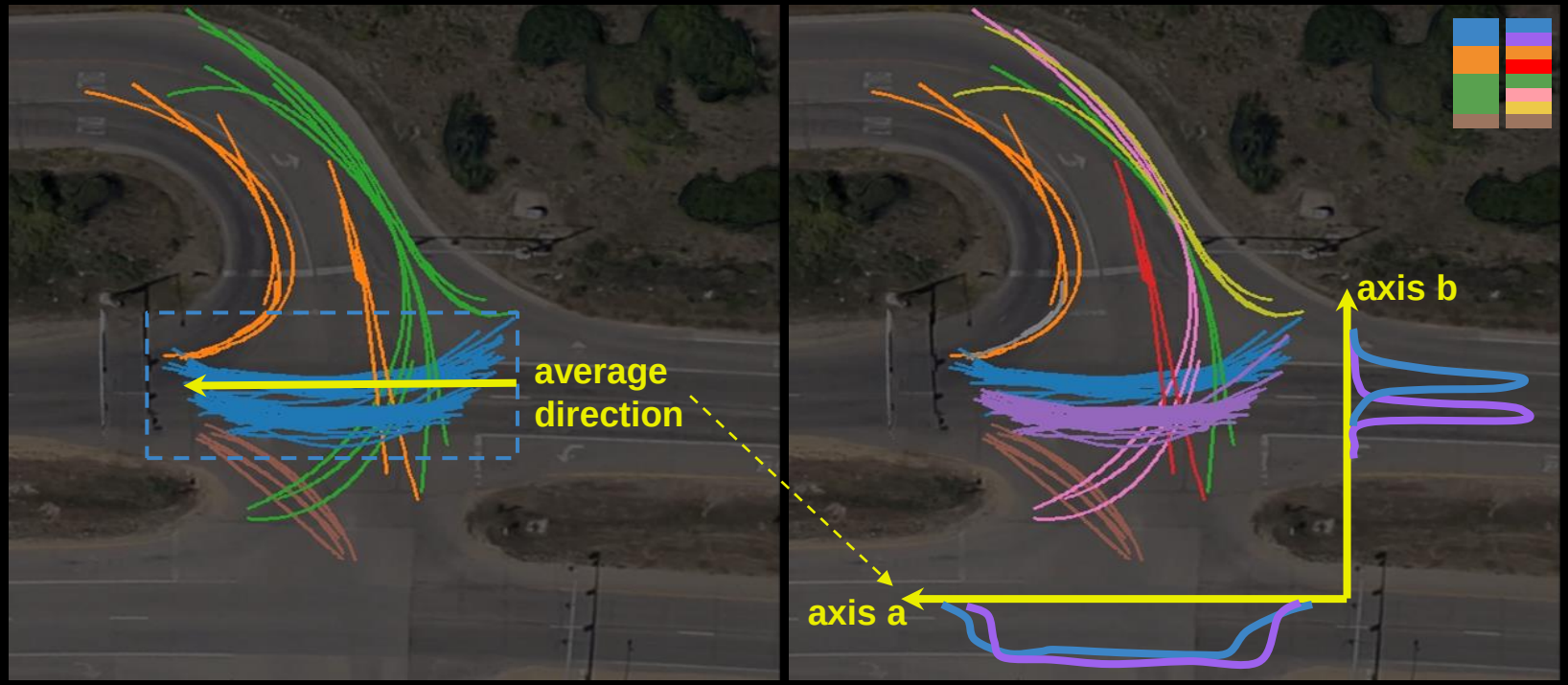
Learning 3D Experts: Hierarchical Clustering



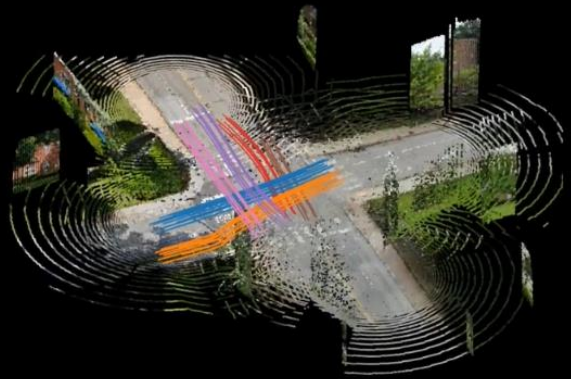
Learning 3D Experts: Hierarchical Clustering



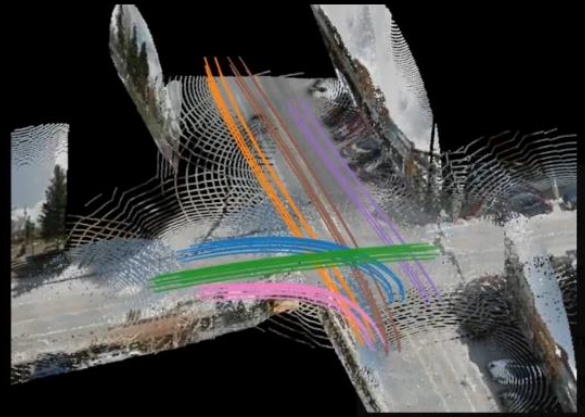
Learning 3D Experts: Hierarchical Clustering



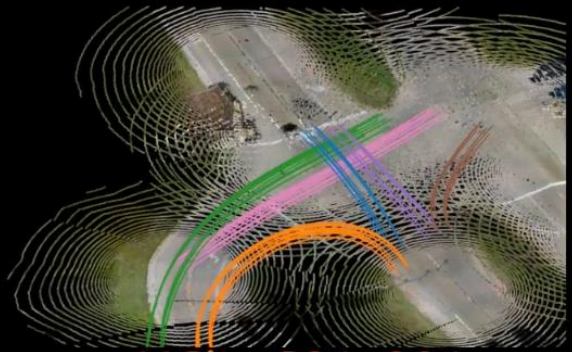
Results: 3D Experts Clustering



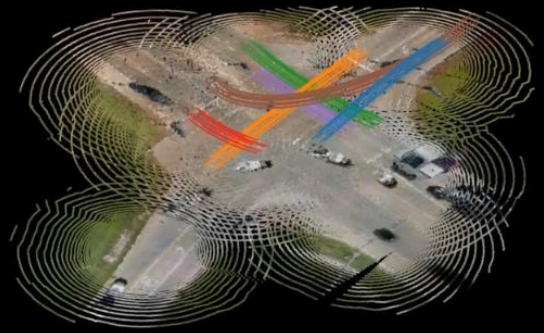
TRAFFIC4D-001



TRAFFIC4D-009

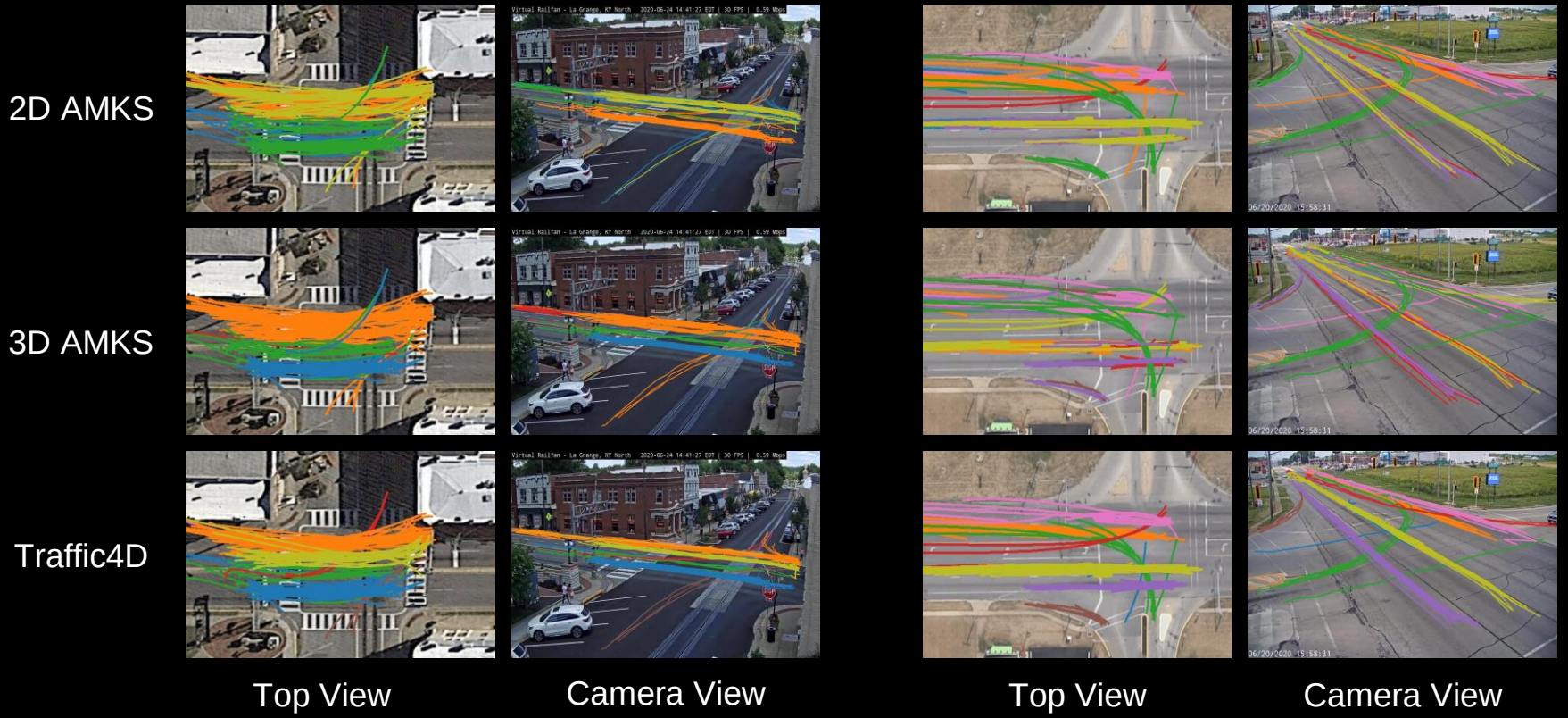


AICity S01_001



AICity S02_c007

Results: 3D Experts Clustering



AMKS: H. Xu, Y. Zhou, W. Lin, and H. Zha, “Unsupervised trajectory clustering via adaptive multi-kernel-based shrinkage,” in ICCV, 2015

3D Experts to Improve Reconstruction

Initial 2D Keypoints

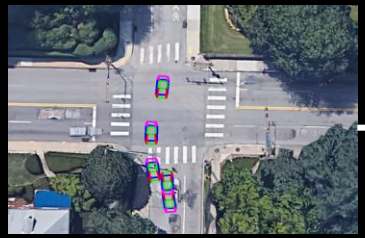
Reproject Reconstruct

Accumulated 2D Keypoints

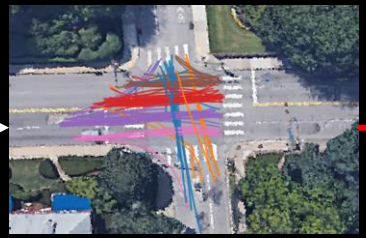
2D Experts

Refined 2D Keypoints

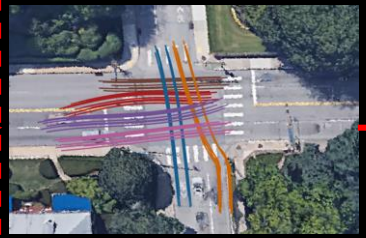
Reproject Reconstruct



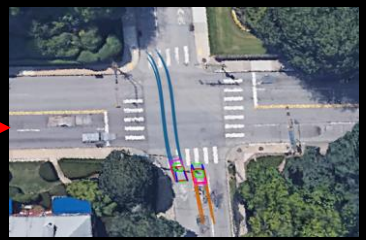
Initial 3D Reconstruction



Accumulated 3D Trajectories



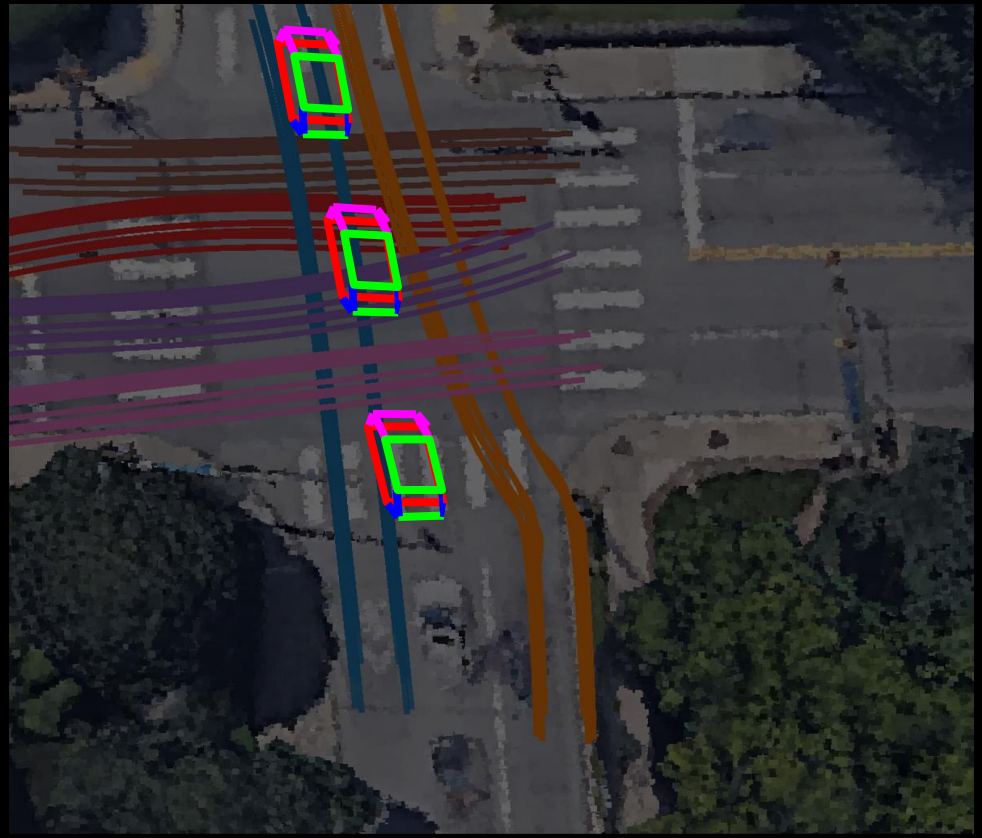
3D Experts: Mean Trajectories



Refined 3D Reconstruction

3D Longitudinal Self-supervision

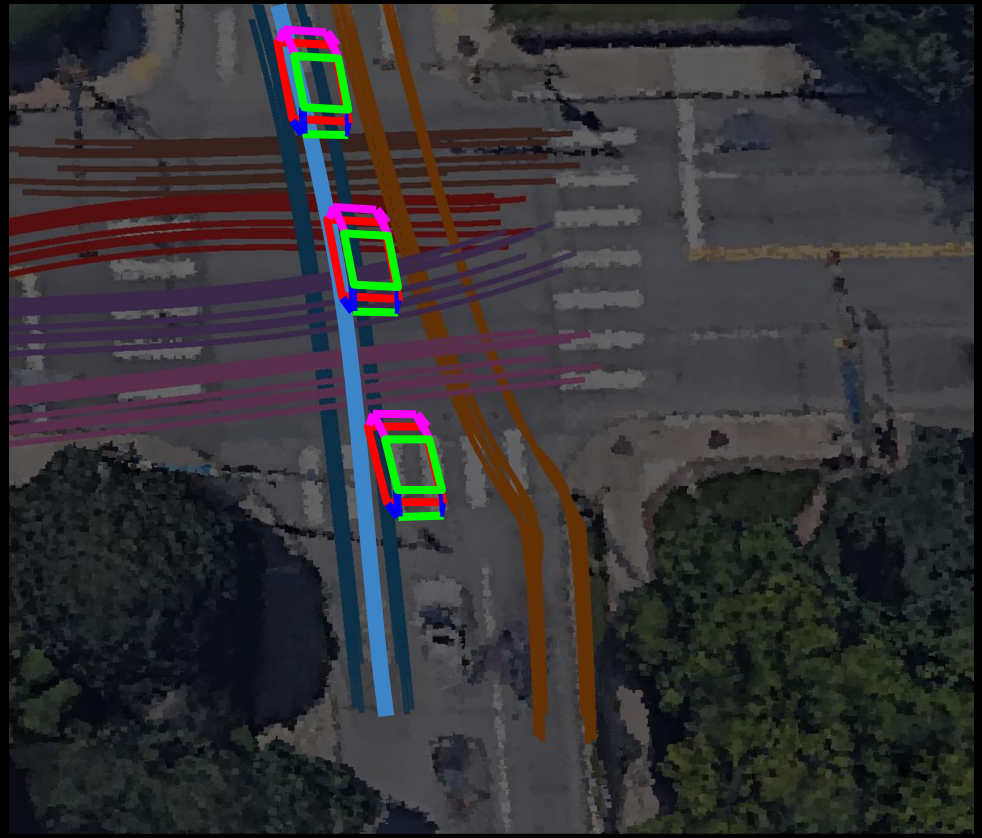
3D Experts to Improve Reconstruction



For each vehicle in consecutive frames...

- Find the corresponding expert with Chamfer distance

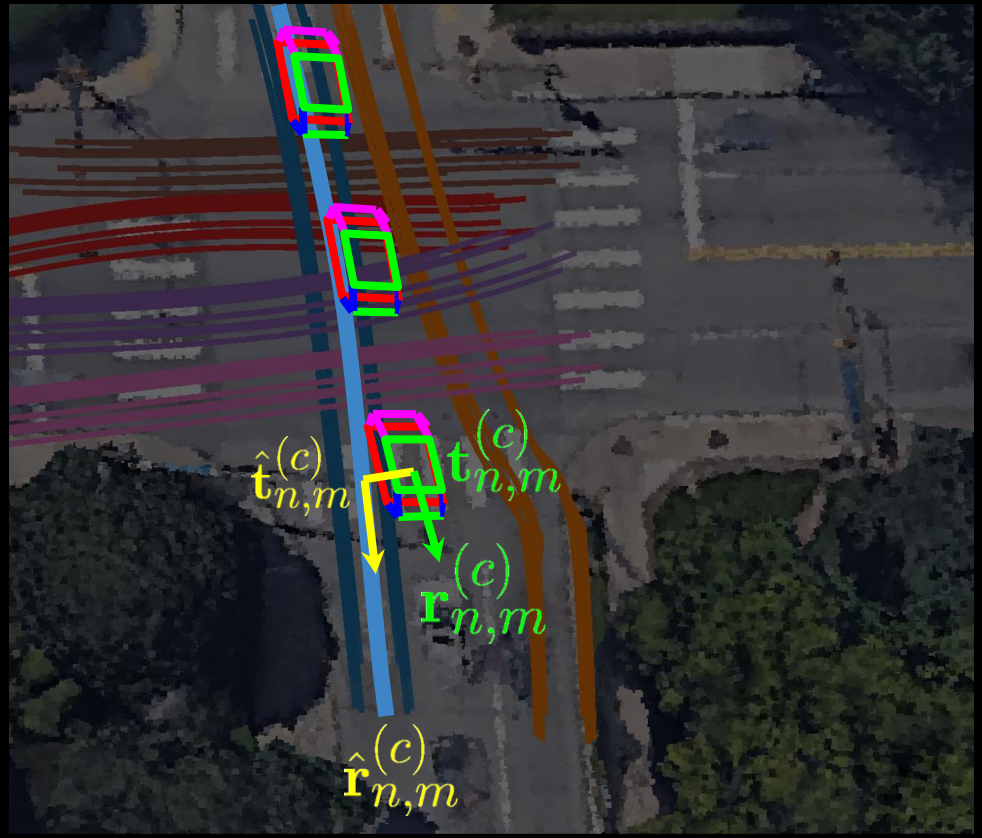
3D Experts to Improve Reconstruction



For each vehicle in consecutive frames...

- Find the corresponding expert with Chamfer distance

3D Experts to Improve Reconstruction



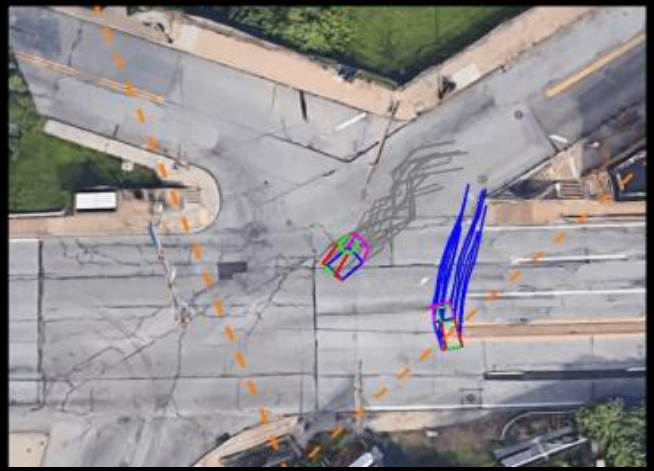
For each vehicle in consecutive frames...

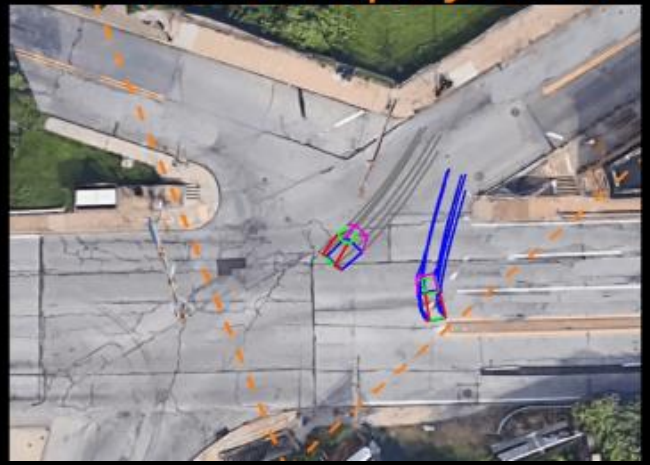
- Find the corresponding expert with Chamfer distance
- Refind 3D poses using correspondence: Longitudinal Loss

Results: Detection & Reconstruction Improvement with Longitudinal Self-Supervision

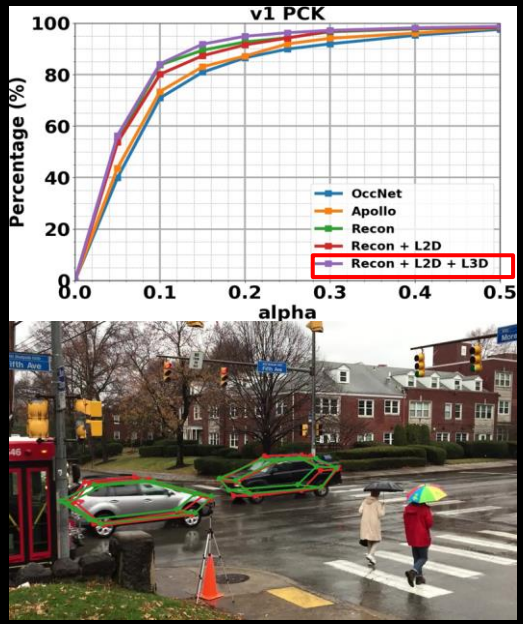
Initial Detections

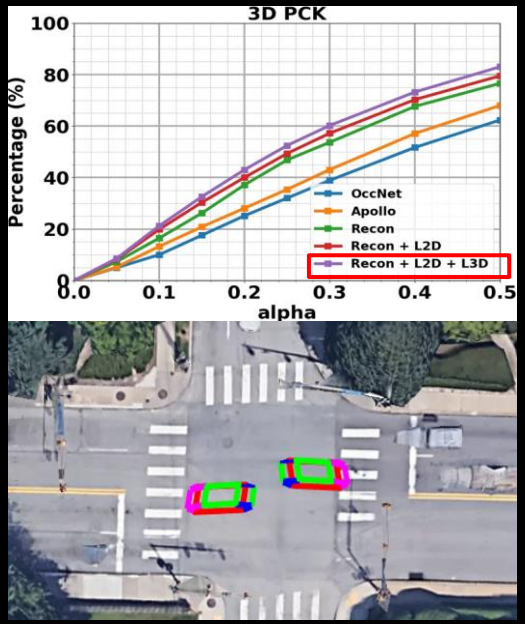
Traffic4D Reprojection

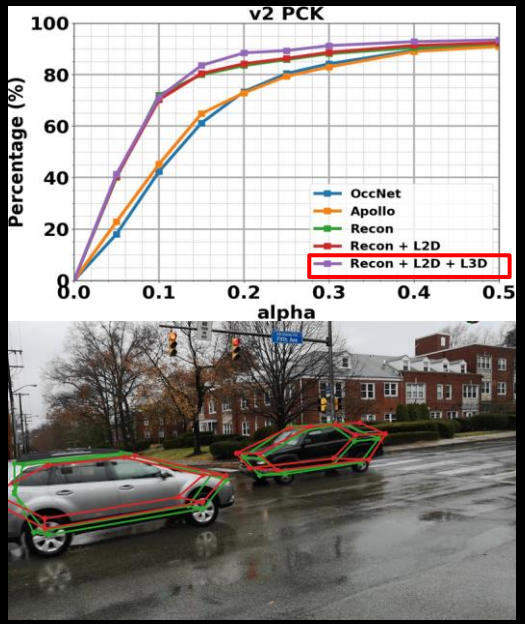




Results: Ablation Study of Keypoint Accuracy with Longitudinal Self-Supervision







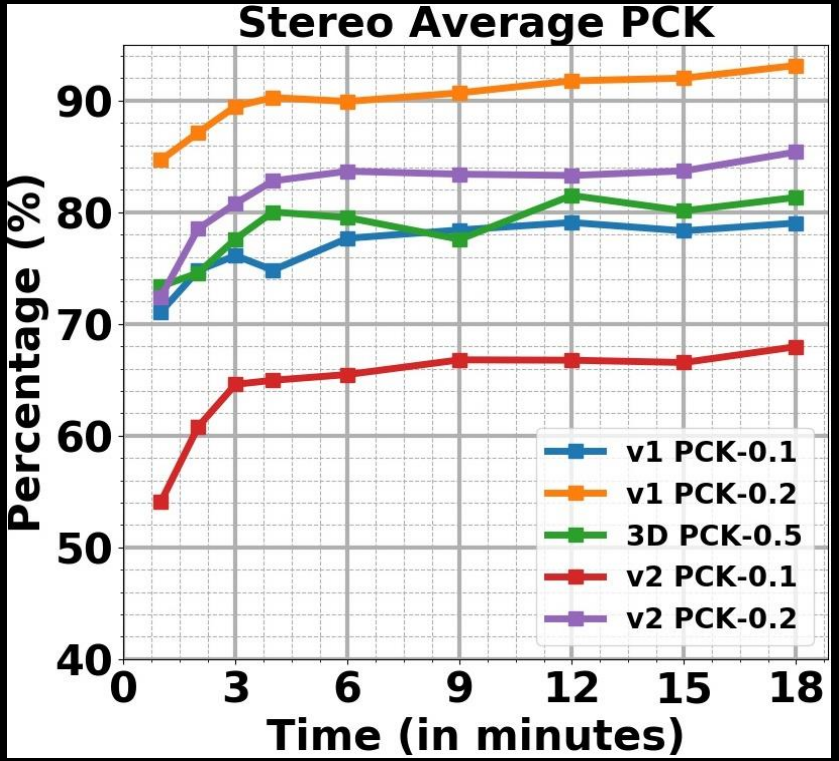
2D Keypoints on Input Frame

3D Keypoints Reconstructed

2D Keypoints Multi-View Consistency

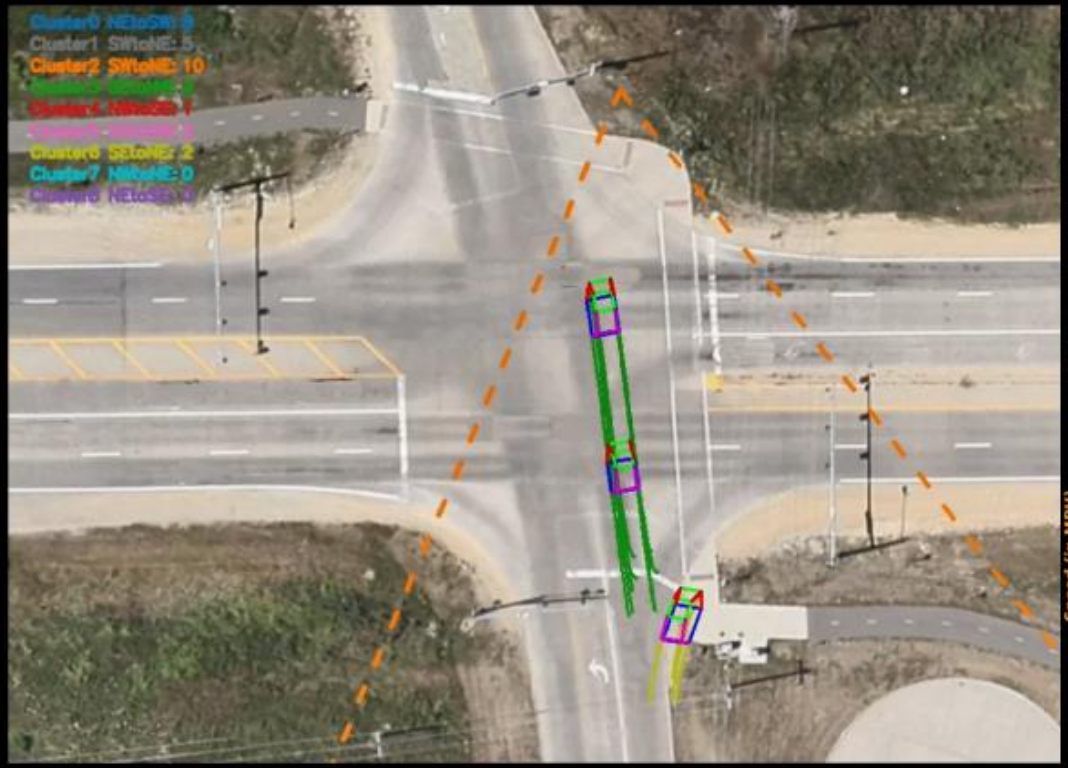
OccNet: N. D. Reddy, M. Vo, and S. G. Narasimhan, "Occlusion-net: 2d/3d occluded keypoint localization using graph networks," in CVPR, 2019
Apollo: X. Song, P. Wang, D. Zhou, R. Zhu, C. Guan, Y. Dai, H. Su, H. Li,and R. Yang, "Apollocar3d: A large 3d car instance understanding benchmark for autonomous driving," in CVPR, 2019.

Results: Keypoint Accuracy Improving Overtime with Longitudinal Self-Supervision

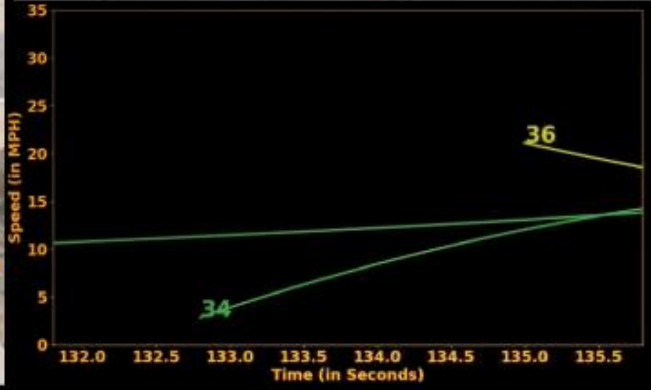


Application: Velocity Estimation & Vehicle Counting

Speed 3X

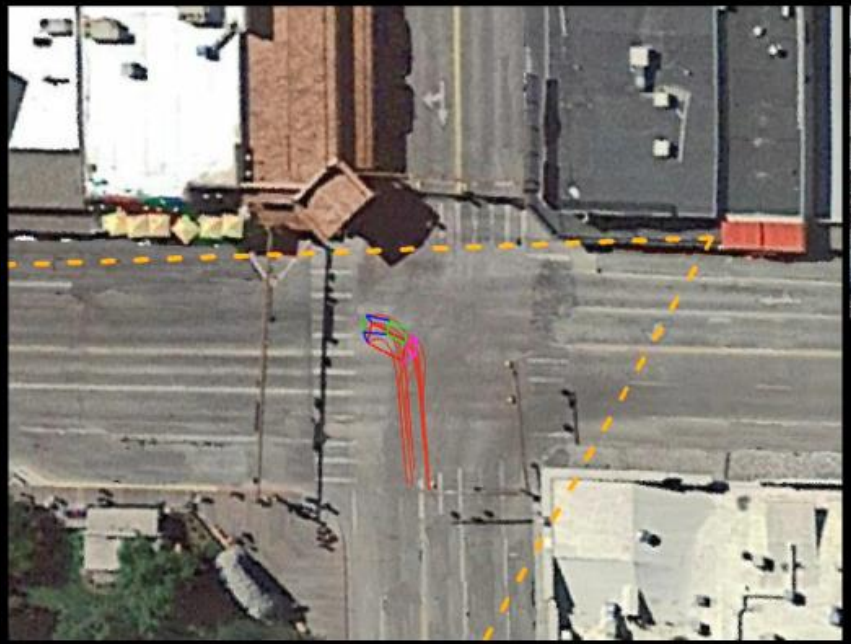


Reprojection



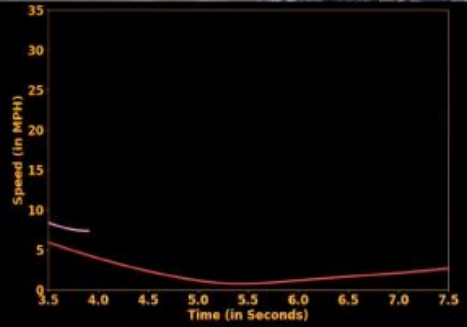
Application: Anomaly detection

Speed 3X



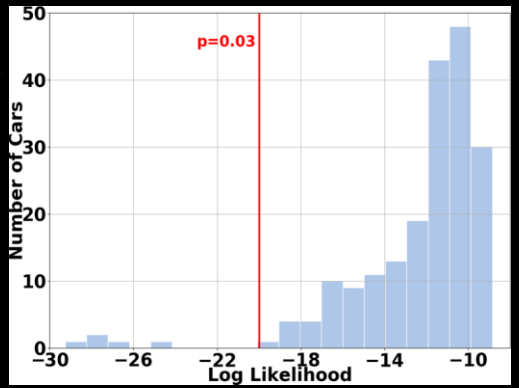
4D Reconstruction

Reprojection



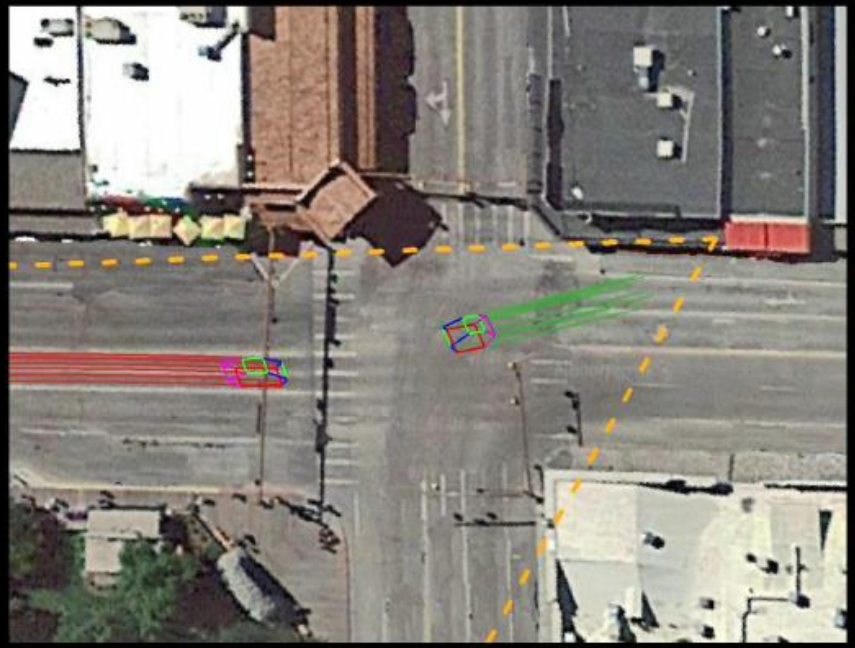
Velocity Profile

No Left Turn



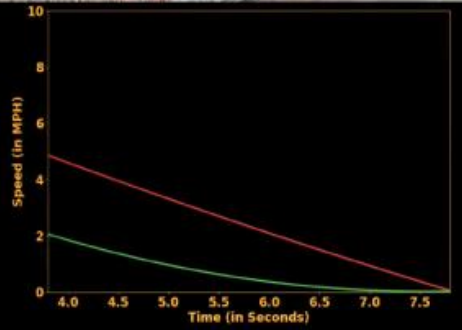
Application: Anomaly detection

Speed 3X



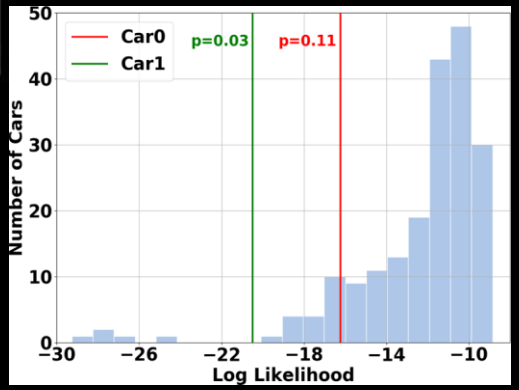
4D Reconstruction

Reprojection



Velocity Profile

Near Miss

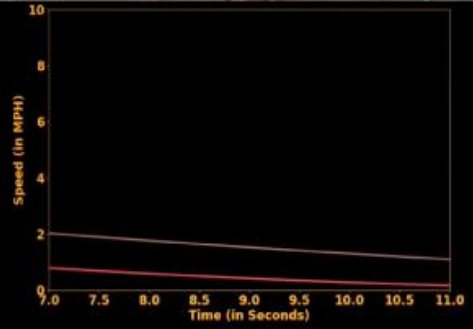


Application: Anomaly detection

Speed 3X

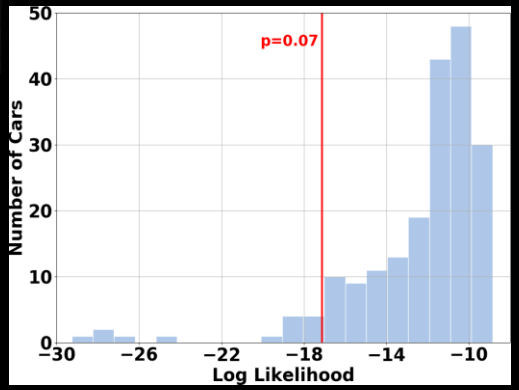
4D Reconstruction

Reprojection



Velocity Profile

Accident



Traffic4D: Single View Reconstruction of Repetitious Activity Using Longitudinal Self-Supervision

Thank You

For More Results Visit our Website at:

http://www.cs.cmu.edu/~ILIM/projects/IM/TRAFFIC4D/