Ground Plane Projection for Improved Traffic Analytics at Intersections

Sajjad Pakdamansavoji[1] Kumar Vaibhav Jha[1] Baher Abdulhai[2]
James H. Elder[1]

[1]York University [2]University of Toronto

Count turns on the ground, not in the image: back-projecting tracks to the ground plane lowers turning-movement count error on all 3 datasets, and fusing 4 cameras nearly halves it again.

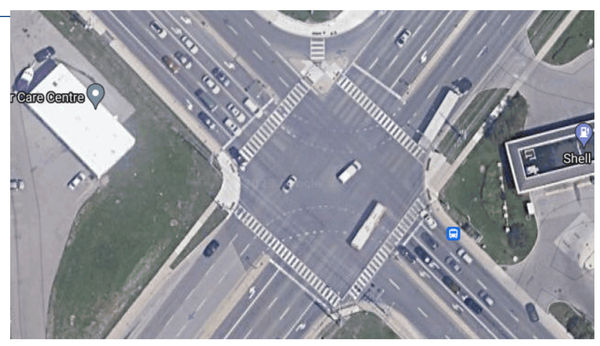
Motivation

- **Turning movement counts (TMC)** drive signal timing, congestion relief and road safety planning.
- Vision systems detect, track and then classify each trajectory into a turning class, usually in the **image plane**.
- Perspective projection **distorts and entangles** trajectories, especially for low, temporary survey cameras (<5 m).
- **Question:** is it better to classify trajectories on the **ground plane** in real-world coordinates?

Three Datasets



Camera view (Trans-Plan)



Orthophoto (ground plane)

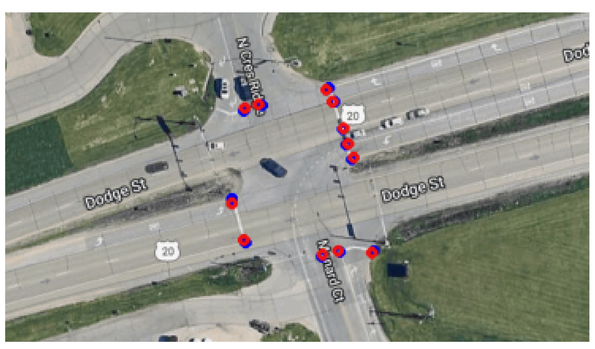
- **CityFlow V2**: 8 high-vantage public intersections, enhanced with turn counts and homographies.
- **Trans-Plan** (new): low-vantage 4K GoPro on a 5 m pole, 4 intersections.
- **Region of York** (new): **4 synchronized 4K cameras**, one per corner, 128 videos across weather and lighting.

Method

Detect → **track** → **classify**, with InternImage detection and ByteTrack tracking.

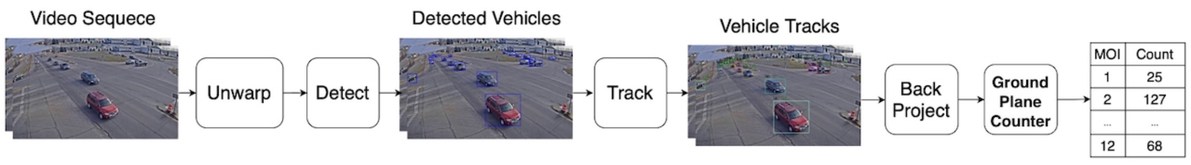
1. **Calibrate once per site**: undistort, then fit a camera-to-ground homography from keypoints matched to a Google Maps orthophoto (2.8–4.3 px error).





Matched keypoints (red) and reprojections (blue), camera vs. orthophoto.

2. **Learn without labels**: cluster tracks into prototypes per lane and turn class, then fit a dense **KDE likelihood map** per class.
3. **Infer**: back-project the bottom-centre of each box and pick the class with the **maximum likelihood** over the whole track.



Ground-plane inference pipeline.

Ground-plane prototypes

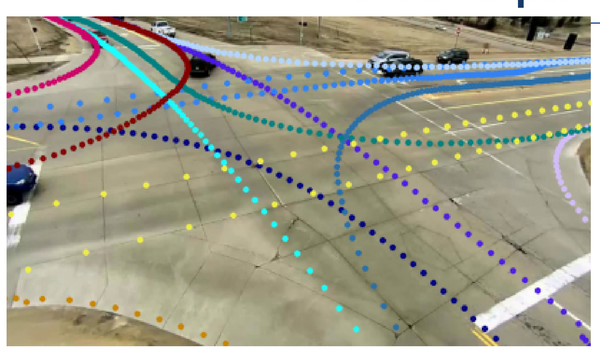
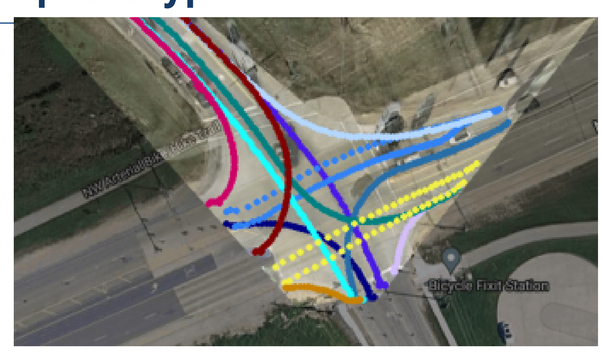


Image plane



Ground plane

One prototype track per lane and turn class; colours are classes. Classes separate cleanly on the ground.

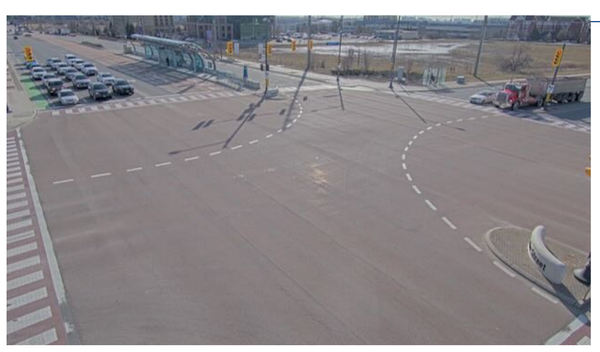
Result 1: ground plane beats image plane

TMC error (mean absolute error over classes, lower is better)

Dataset	Plane	EE	DIR	VOTE	ML
CityFlow	Camera	16.0%	6.6%	7.8%	5.1%
	Ground	14.0%	5.4%	7.9%	**4.0%**
Trans-Plan	Camera	17.7%	12.6%	14.2%	11.0%
	Ground	14.3%	10.5%	13.8%	**9.9%**
Region of York	Camera	–	20.0%	38.2%	19.5%
	Ground	–	14.9%	38.2%	**13.8%**

- The **ground plane wins** in all but two cases, for every classifier.
- Our **maximum-likelihood (ML)** classifier is best on all three datasets.
- Low cameras and narrow fields of view (Trans-Plan, York) are the hardest cases.

Result 2: weak multi-camera fusion



One of 4 York corner cameras



Same view on the ground plane

- Each York camera sees only part of the intersection, so each turn class is assigned to the camera with the lowest training error.

System	Error rate	Bias
Single camera	13.8%	5.2%
4 cameras, weak fusion	**7.3%**	**0.1%**

- Fusion **nearly halves the error** and removes the over-counting caused by tracks broken up by occlusion.

Conclusion

- Analyse traffic on the **ground plane**, not the image plane.
- **Unsupervised** likelihood models beat heuristic classifiers.
- **Multiple views** beat a single view, even with simple fusion.
- **Next:** terrain models, sway stabilization and early 3D fusion of detections.