Computer Vision System for Traffic Counts at Intersections

Sajjad Pakdamansavoji, James H. Elder

YORK UNIVERSITÉ UNIVERSITY · VISTA VISION: SCIENCE TO APPLICATIONS YORK UNIVERSITY · Centre for Vision Research

Abstract

Real-time sensing of the vehicle traffic at an intersection can support many important smart city tasks, including real-time optimization of signaling and longer-term urban planning. While conventional inductive loop and radar technologies have been proven in the field, video-based computer vision systems are emerging as a potential lower-cost and more broadly informative competitor. Still, challenges remain, including the extent of manual calibration required and the effects of occlusion. Here we report a prototype one-camera system that addresses these challenges.

Our pipeline involves a lightweight calibration procedure, detection, tracking, clustering, and counting modules. The product is a time-series of turn counts. Evaluation of the system on two quite different camera installations in the Greater Toronto Area reveals a turn-count error of ~7% in daylight conditions.

Method

Our Traffic Analytics Software follows a pipeline consisting of six main blocks: preprocessing, detection and tracking, reprojection, and analysis as depicted in Figure 1.

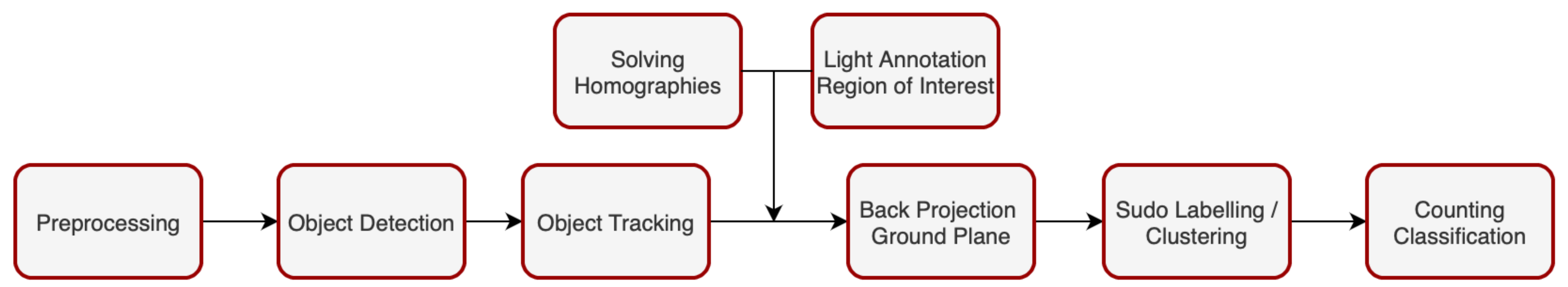

Fig 1. The six main blocks of our traffic analytics pipeline

- **Preprocessing** given camera intrinsic non-linear distortion is removed
- **Obj Detection** bounding boxes of vehicles are estimated with detectron2 Resnet50 + FPN + detection and segmentation heads
- **Obj Tracking** unique ids are assigned using StrongSort + osnet_x1_0 re-ID
- **Back Project** with homographies project bounding box bottom to ground
- **Clustering** use RoI and clustering to extract prototypes for each MoI
- **Counting** classify based on distributions estimated with KDE

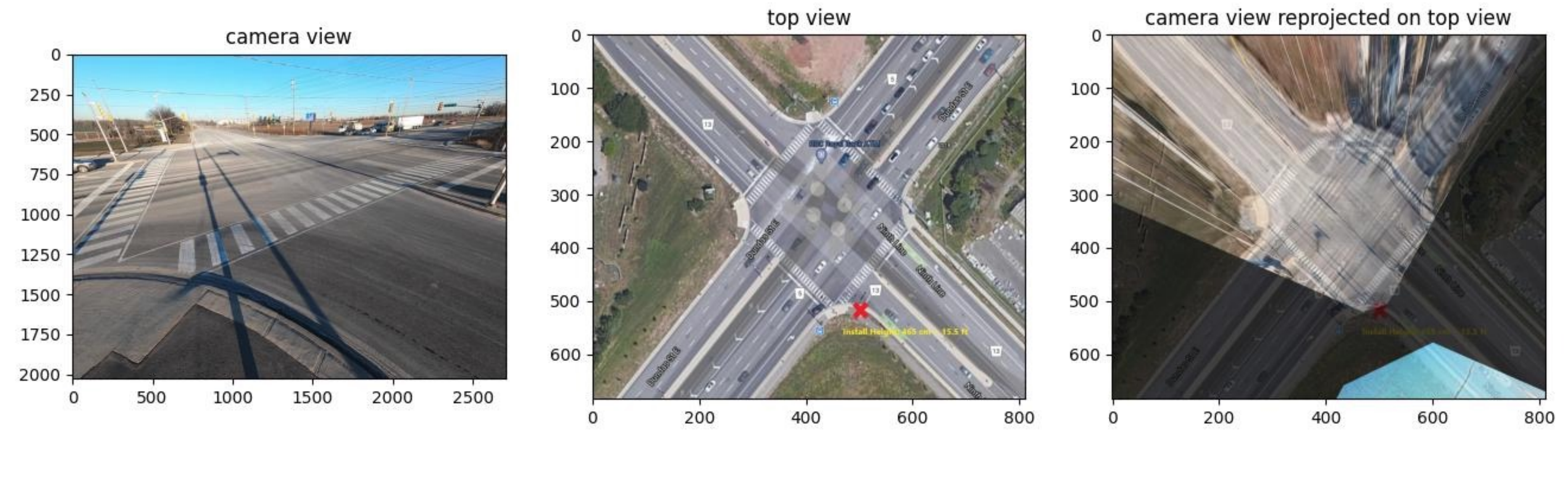

Fig 2. Projection of points from camera view to top view

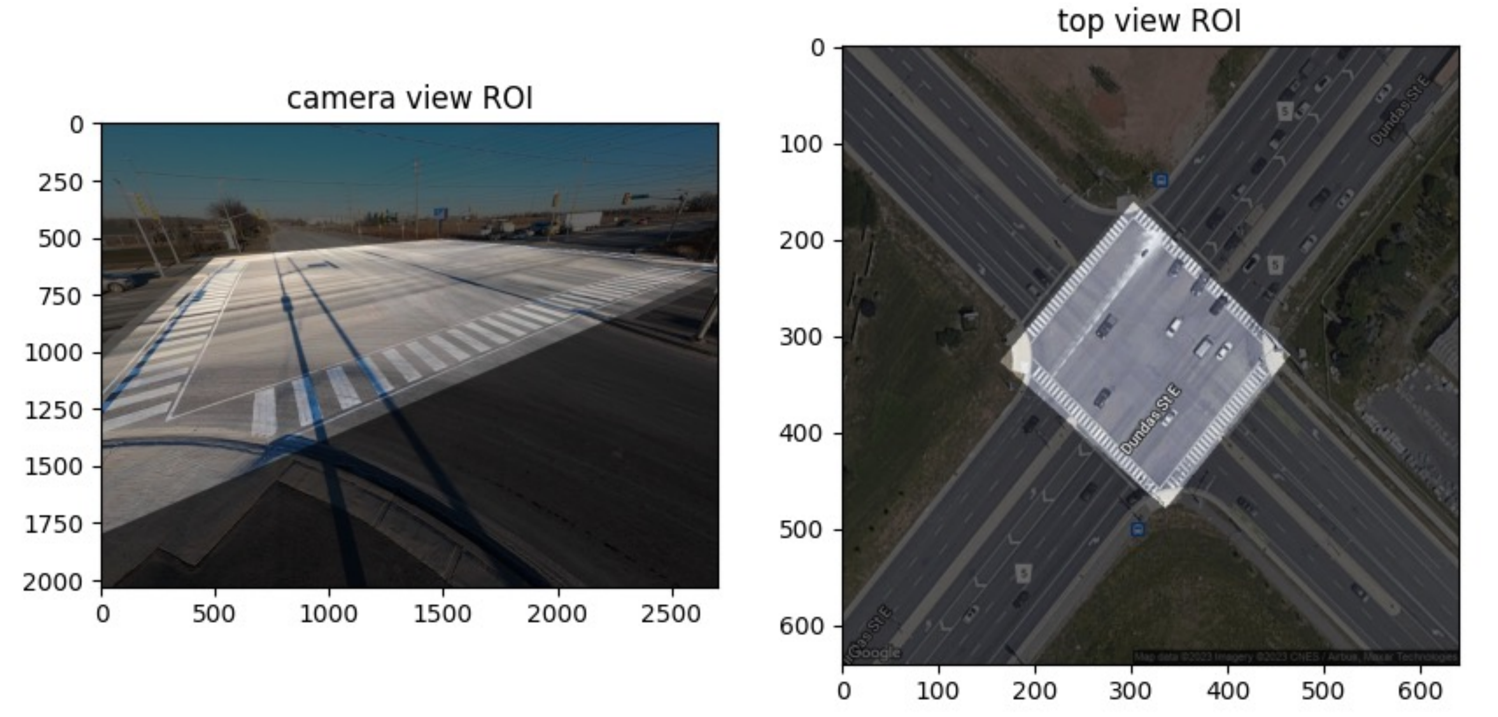

Fig 3. Region of Interest(ROI) camera view(left) and ROI on ground (right)

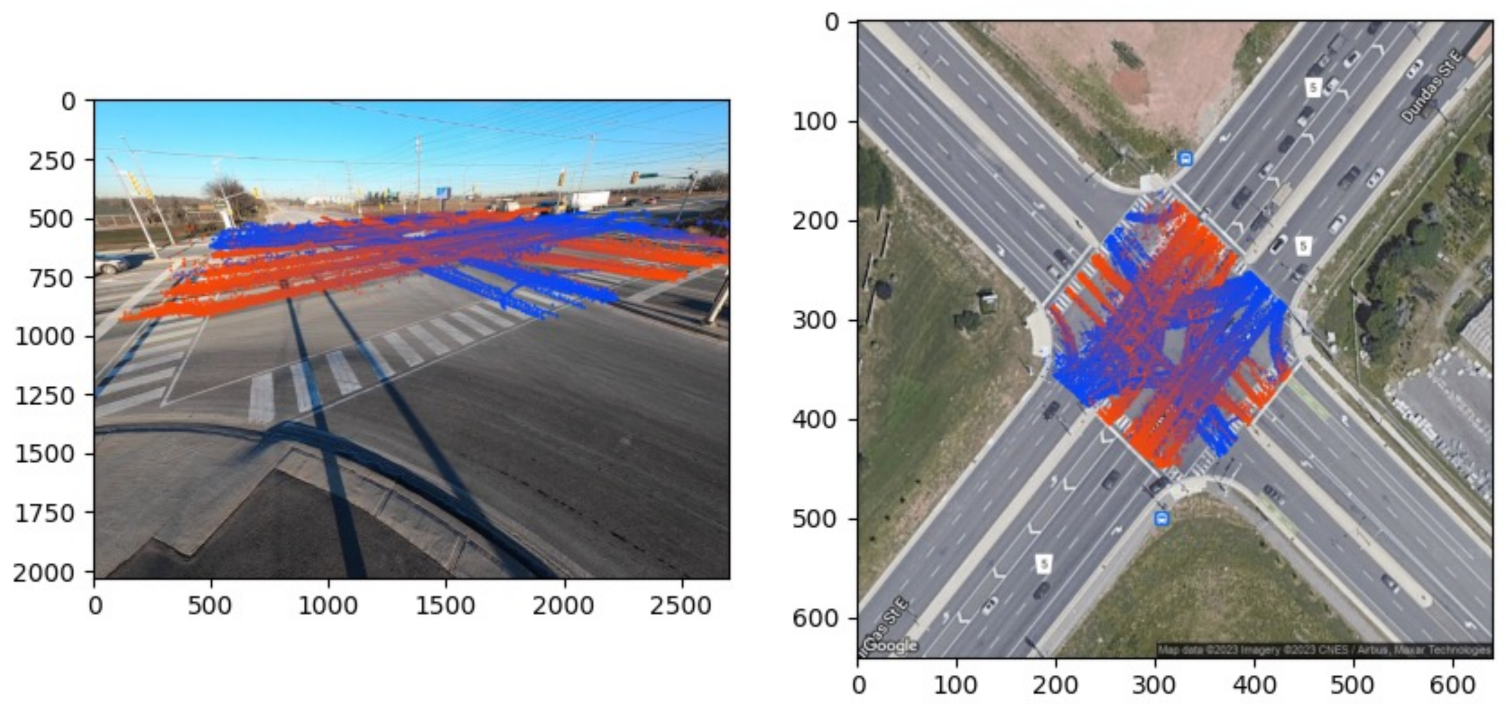

Fig 4. Visualization of tracks of camera view(left) and their corresponding trajectories on the ground plane(right)

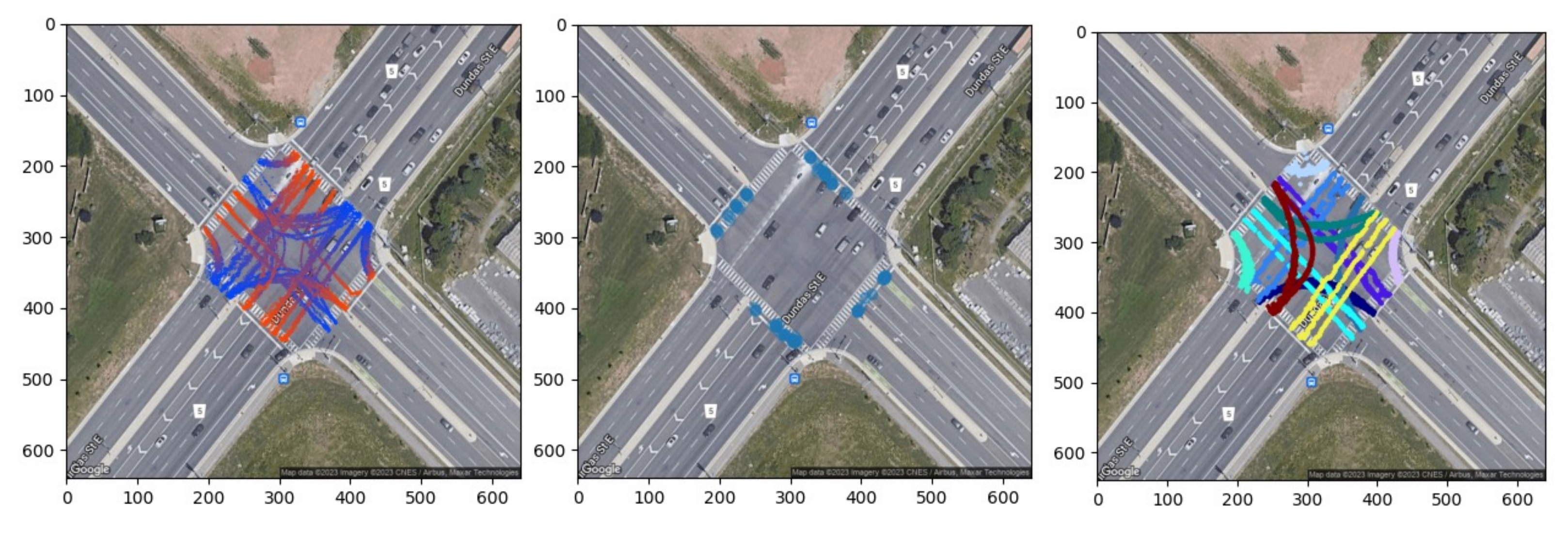

Fig 5. Tracks used for clustering(left), clusters of starting points(middle), automatically selected prototypes for each MoI(right)

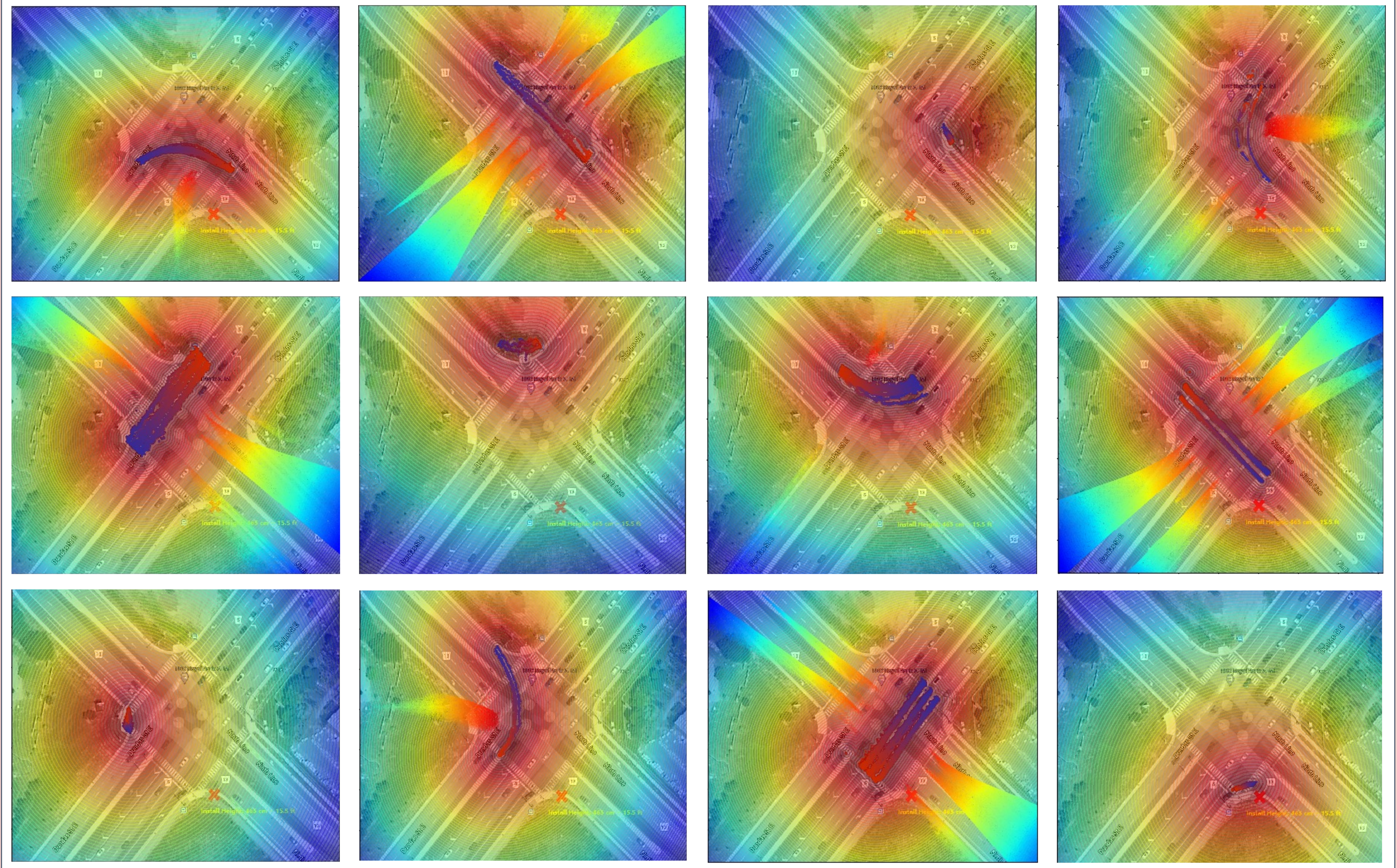

Fig 6. Log densities estimated with KDE on complete tracks

Results

Deriving accurate numerical results for the presented pipeline requires a vast dataset with known calibrated camera. Since these requirements are not commonly satisfied for available datasets, we are capturing a novel dataset that includes two distinct traffic surveillance scenarios in the GTA. The first dataset, created in collaboration with our industry partner TransPlan, was captured by a GoPro camera temporarily mounted on a light pole at the corner of a number of intersections. The second dataset, created in collaboration with the University of Toronto and the Region of York, features synchronized video from four cameras. As such here we only present the numerical result for one intersection located at Dandas St. and 9th Line.

MoI	GT	DT	Diff	Err%	MoI	GT	DT	Diff	Err%
1	18	21	3	0.166	1	13	13	0	0
2	20	21	1	0.05	2	30	32	2	0.066
3	20	13	7	0.35	3	6	6	0	0
4	10	9	1	0.1	4	8	8	0	0
5	127	135	8	0.062	5	87	90	3	0.034
6	9	10	1	0.111	6	12	11	1	0.083
7	29	31	2	0.068	7	18	18	0	0
8	40	41	1	0.025	8	30	30	0	0
9	14	13	1	0.071	9	22	21	1	0.045
10	18	19	1	0.055	10	14	14	0	0
11	176	179	3	0.017	11	85	88	3	0.035
12	28	20	8	0.285	12	0	0	0	0
All	509	512	37	**0.072**	All	325	331	10	**0.030**

Table 1. Numerical results of the pipeline. Dandas(left) and HW7(right)

Conclusion

The proposed pipeline provides a way of analyzing traffic flow at intersections which is independent of the camera angle and other physical constrains. This works has yet to be improved in terms of numerical results after being provided with our novel dataset.

References

[1]: Design and Perceptual Validation of Performance Measures for Salient Object Segmentation Movahedi, V. & Elder, J.H.

[2]: Detectron2, Yuxin Wu and Alexander Kirillov and Francisco Massa and Wan-Yen Lo and Ross Girshick

[3]: GIAOTracker: A Comprehensive Framework for MCMOT With Global Information and Optimizing Strategies in VisDrone, Du, Yunhao and Wan.

[4]: Torchreid: A Library for Deep Learning Person Re-Identification in Pytorch, Zhou, Kaiyang and Xiang.

NSERC CRSNG