3D Multi-Camera Reasoning for Movement Turn Counts at Intersections

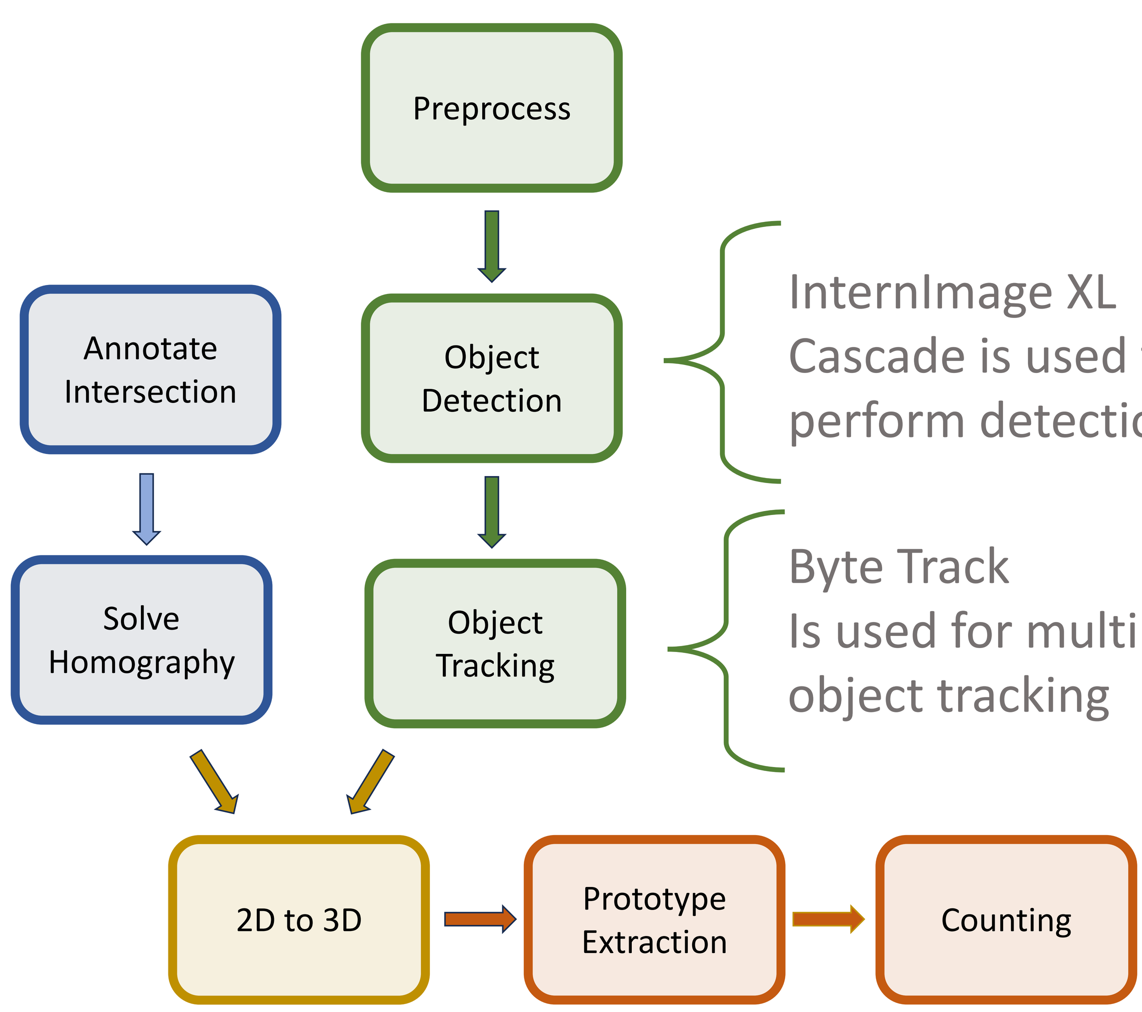
Sajjad Pakdamansavoji, Kumar Vaibhav Jha, James H. Elder

Abstract

This poster tackles Movement Turn Count (MTC) through the lens of 3D reasoning, illustrating its accuracy and robustness in applications like surveillance and urban planning. By merging 3D tracking with geo-spatial elements integration, it improves MTC classification and its scalability and practicality across varied sensory environments, highlighting a significant leap forward in real-world applicability.

Methodology

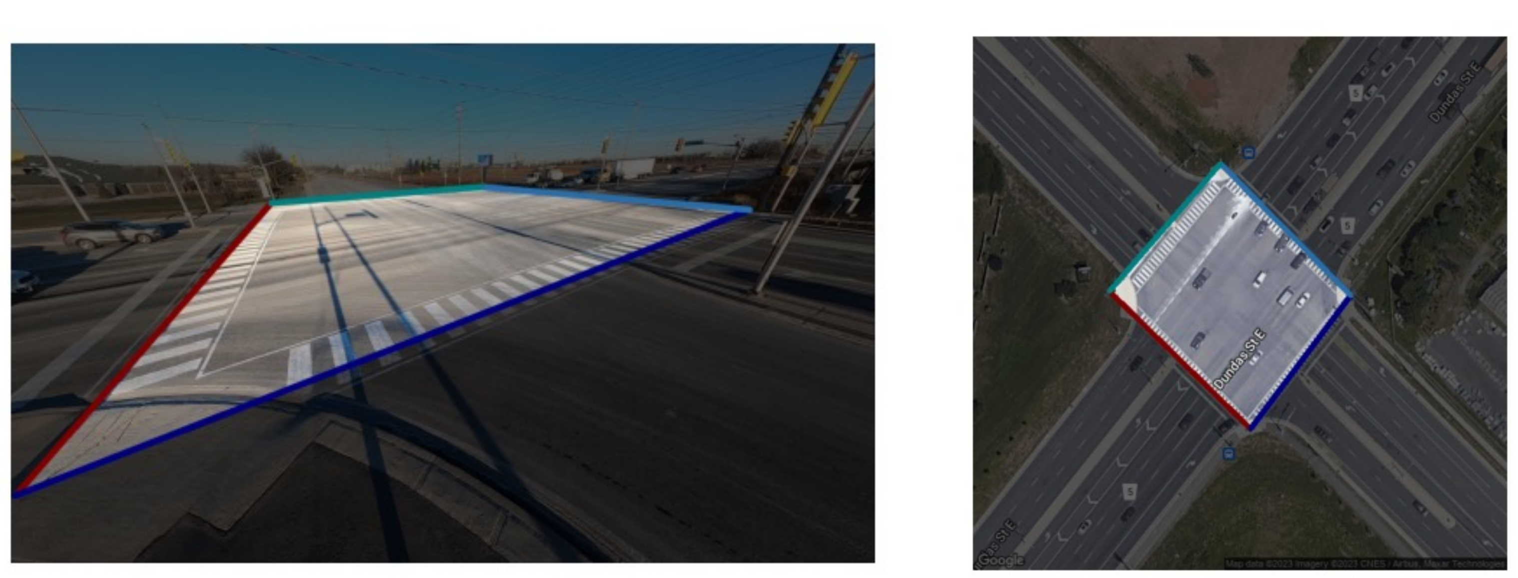
Our MTC method follows a pipeline consisting of six main blocks: preprocessing, detection and tracking, reprojection to 3D, and counting.



InternImage XL Cascade is used to perform detections

Byte Track Is used for multiple object tracking

Firstly one needs to select the intersection boundary.

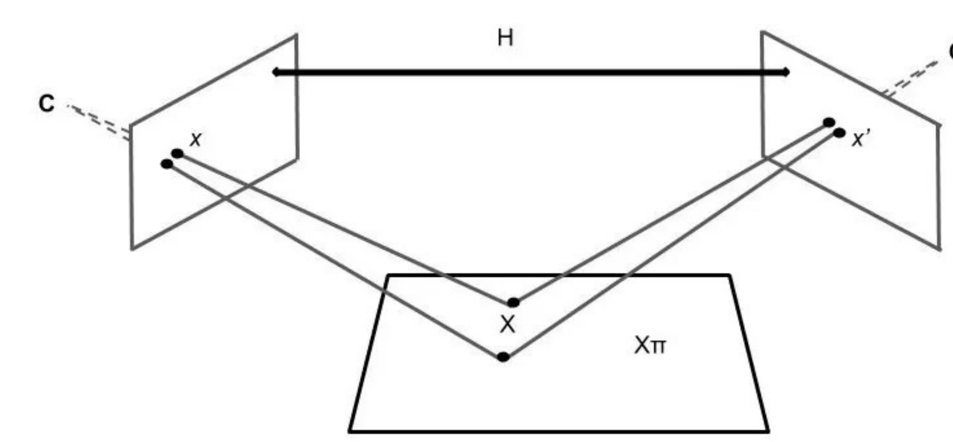


a) Camera view

b) Aerial view

2D to 3D Transition

The transition is based on projective geometry and the fact that road users are approximately on a 2D plane.



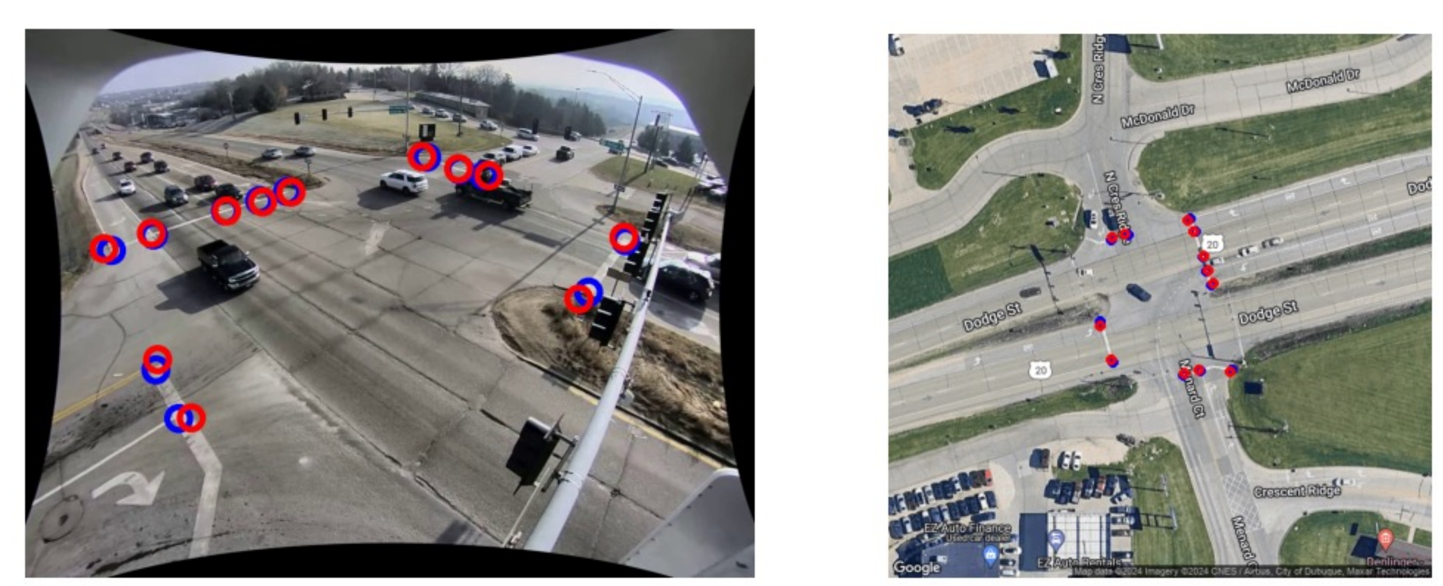
$$x = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \qquad x' = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

$$c \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

$$A = \begin{pmatrix} -x & -y & -1 & 0 & 0 & 0 & ux & uy & u \\ 0 & 0 & 0 & -x & -y & -1 & vx & vy & v \end{pmatrix}$$

$$\min_h \|Ah\|_2 \quad \text{subject to} \quad \operatorname{rank}(A) = 3$$

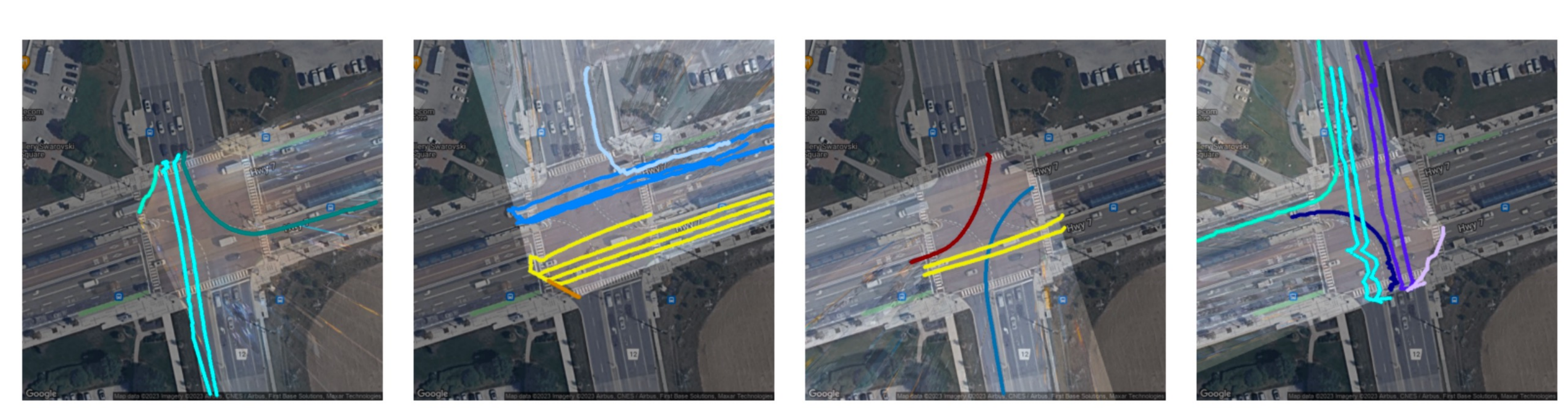
To solve the optimization, we take a LMS approach requiring selection of pair points.



Selected pair points and their back-projections

Multi-camera Prototypes

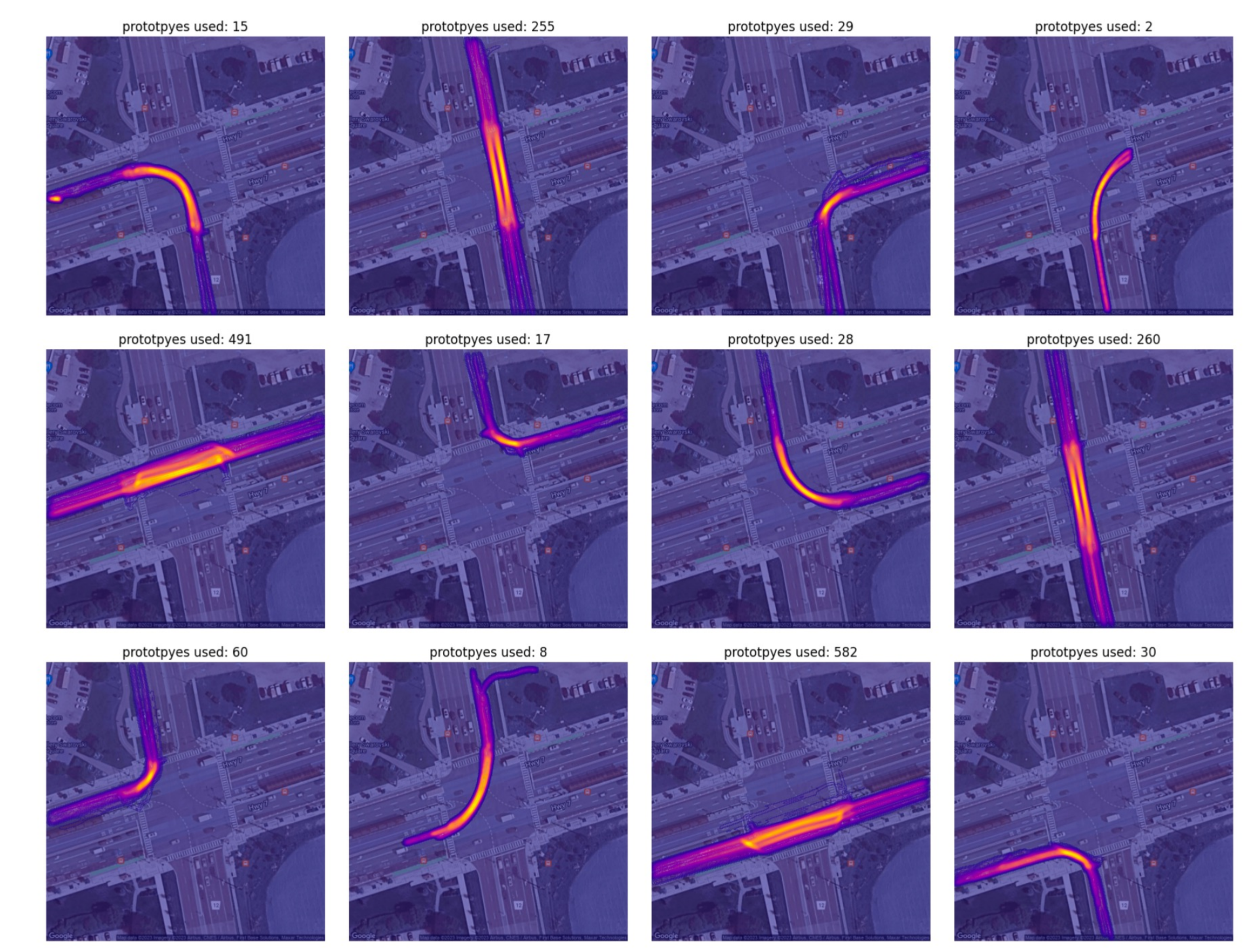
To extract prototypes, we use all cameras and merge tracks on the reference 3D coordinates.



Selected prototypes from each camera view.

Turn Counts

Given prototypes of each movement, kernel density estimation is used to estimate probability distributions used for the generative Bayesian classifier.



Heatmap of estimated densities for each movement

Results

Split	MC Integrated	kDE	Bias
Train	×	11.16	5.53
Train	✓	6.15	0.83
Valid	×	10.53	5.57
Valid	✓	5.39	0.063

Classification percent error on Region of York Dataset.

Conclusion

Utilizing 3D reasoning enhances distance measurement accuracy and facilitates the integration of multi-sensory data through conversion into a standardized real-world 3D coordinate system.

References

K. Bhargava, S. Pakdamansavoji, B. Abdulhai, J. H. Elder, and S. L. Waslander, "Intersection traffic dataset," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2024.