I see you: A Vehicle-Pedestrian Interaction Dataset from Traffic Surveillance Cameras.

Hanan Quispe[1] Jorshinno Sumire [1], Patricia Condori[2] , Edwin Alvarez[3], Harley Vera[1].

[1]Universidad Nacional de San Antonio Abad del Cusco, [3]Pontificia Universidad Católica del Perú, [2]Business on Engineering and Technology - BE Tech

{163819,160347,harley.vera}@unsaac.edu.pe, edwin.alvarez@pucp.edu.pe, patricia.condoriobregon@betech.net.pe

1. Abstract

The development of autonomous vehicles arises new challenges in urban traffic scenarios where vehicle-pedestrian interactions are frequent e.g. vehicle yields to pedestrians, pedestrian slows down due approaching to the vehicle. Over the last years, several datasets have been developed to model these interactions. However, available datasets do not cover near-accident scenarios that our dataset covers. We introduce *I see you*, a new vehicle-pedestrian interaction dataset that tackles the lack of trajectory data in near-accident scenarios using YOLOv5 and camera calibration methods. *I see you* consist of 170 near-accident occurrences in seven intersections in Cusco-Peru. This new dataset and pipeline code are available on (**https://github.com/hvzzzz/Vehicle_trajectory_Dataset**).

2. Introduction

The problem of predicting pedestrian paths is crucial for the development of autonomous vehicles (AVs) that have to handle complex interactions with pedestrians in urban traffic environments e.g. pedestrians recklessly crossing the road, and vehicles speeding up to overtake a pedestrian.

Figure 1: Vehicle-Pedestrian interactions captured by our dataset.

I see you captures vehicle-pedestrian avoidance behaviors in dangerous situations and scenarios where vehicle-pedestrian are very close but do not represent a dangerous situation (Figure 1) as this would provide valuable negative feedback. For each case, we provide processed vehicle and pedestrian trajectories in GPS coordinates.

6. Statistics

Speed is preferred to train pedestrian trajectory prediction models rather than absolute position. This is because it does not depend on the reference system and therefore it can be used as a more general way to describe motion. In figure speed distributions were calculated for vehicles and pedestrians. Both distributions show peaks at $0\frac{km}{h}$ this is because vehicles and pedestrians are stopped by traffic lights at signalized intersections. Also, each distribution has a second peak that shows the mean speed of vehicles and pedestrians during the near-accident scenarios.

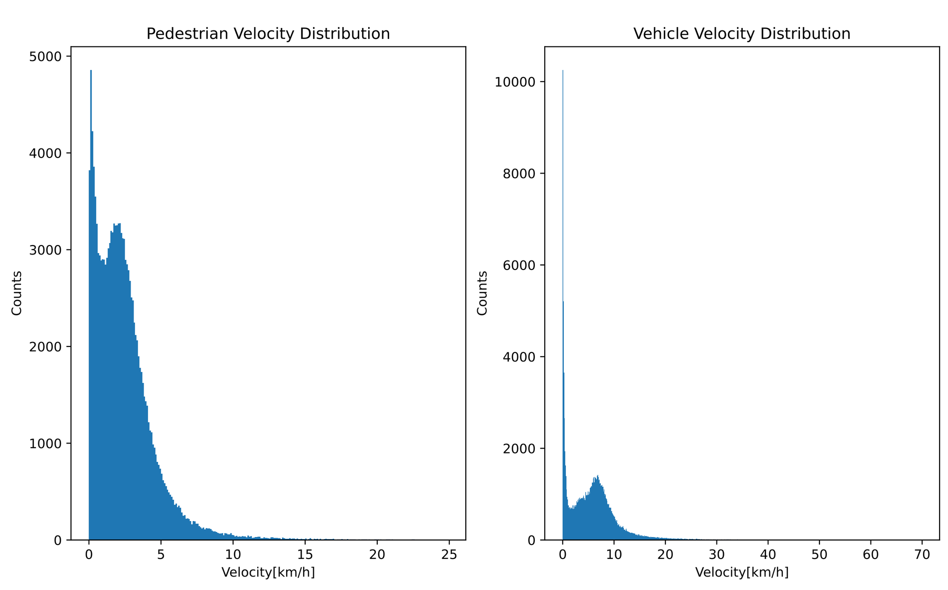

Figure 3: Velocity distributions for vehicle and pedestrian categories in I see you.

3. I see you Dataset

We developed a pipeline to collect *I see you* (Figure 2). The object detection task was performed using a YOLOv5, for tracking we used StrongSORT then we used a Linear Kalman filter to remove noise from the trajectories and finally for transformation to GPS coordinates we used Perspective-n-Point.

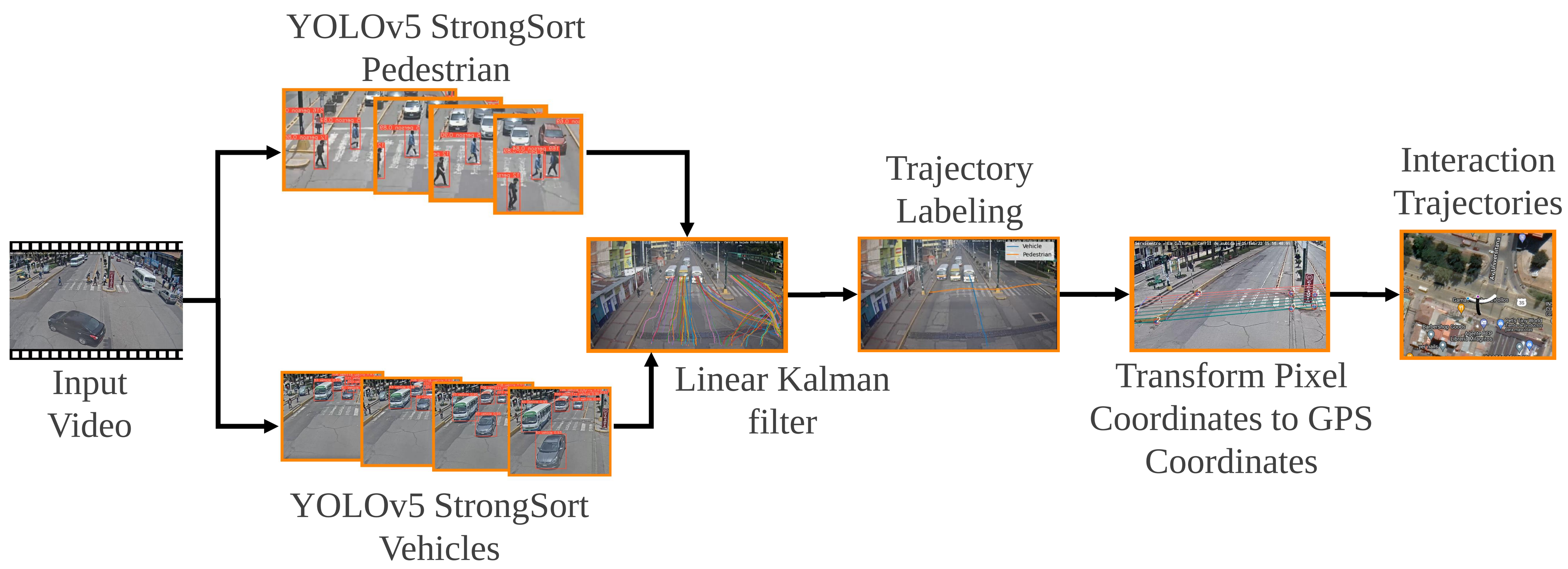

Figure 2: Pipeline for the proposed system.

Video footage was collected from publicly available traffic surveillance videos at seven signalized intersections in Cusco-Peru . The seven intersections were selected near schools, colleges, churches, and hospitals to ensure a diverse distribution in pedestrian characteristics (genre, age). To capture all illumination variations we sampled the 18-hour videos saving a frame every 6 minutes resulting in 4246 images in total. Employing a 90% train, 10% test distribution, we obtained the results described for object detection in the test subset for each category.

Category	P	R	mAP@0.5	mAP@0.5:0.95
Vehicle	0.934	0.916	0.974	0.76
Pedestrian	0.864	0.853	0.91	0.595

The filtered trajectories were manually labeled with the id generated by the tracker of the vehicle and pedestrian involved in the near-accident scenario. The labeled trajectories were converted from pixel coordinates to GPS coordinates using semi-automatic camera calibration based on Perspective-n-Point (PnP). Finally, we compare our dataset with other trajectory datasets used for pedestrian trajectory prediction in various cases (Table 2).

5. Comparison with Other Datasets

Dataset	Scenarios	Interaction Type	Agents	FPS	Amount of Trajectories
DUT [?]	Campus	Vehicle-Crowd in shared spaces	Car, Pedestrian	23.98	1793
I see you	Signalized Intersections	Vehicle-Pedestrian in near-accident scenarios	Car, Pedestrian	30	340
ETH [?]	Campus, Urban Street	Pedestrians in busy scenarios	Pedestrian	2.5	650
UCY [?]	Urban Street, Campus, Park	multi-human interaction scenarios	Pedestrian	2.5	909

7. Conclusions

In this work, we addressed the need for trajectory data in near-accident scenarios for which we have developed a pipeline for collecting the trajectories of pedestrians and vehicles. During this process, we have assessed the limitations of our pipeline in which manual labeling is a process that can be automated using clusters to select the trajectories that correspond to the interactions between pedestrians and vehicles in near-accident scenarios. Also, the clustering approach could be used to select vehicle-pedestrian interactions in other scenarios.