INFERRING VEHICLE SPEEDS, MAKES, AND MODELS FROM VIDEO CAMERA FOOTAGE: READY OPPORTUNITIES FOR TRANSPORT DATA ANALYSIS

Mohammadreza Mirzaei  
Department of Civil Engineering  
Sharif University of Technology, Tehran, Iran  
m.mirzaei77@student.sharif.edu

Mohaddese Salchian Khaled Abady  
Department of Civil Engineering  
Sharif University of Technology, Tehran, Iran  
M.salehian99@student.sharif.edu

Alizer Khowaja  
Undergraduate Research Assistant  
Department of Civil, Architectural and Environmental Engineering  
The University of Texas at Austin  
alizerkhowaja@utexas.edu

Kara M. Kockelman, Ph.D., P.E.  
Dewitt Greer Centennial Professor of Transportation Engineering  
Department of Civil, Architectural and Environmental Engineering  
The University of Texas at Austin – 6.9 E. Cockrell Jr. Hall 14 Austin, TX 78712-1076  
kkockelm@mail.utexas.edu

1 Corresponding Author
ABSTRACT

This paper shows how to quickly infer the body type, make, model and speed of vehicles using a single intersection camera in Austin, Texas. To develop a tuned Vehicle Make and Model Recognition model capable of object tracking, YOLOv4, DeepSort, and TensorFlow were used in conjunction with vehicle images from the ResNet-152 Cars Dataset. Regression results suggest motorcycles, utility vehicles, and vans average 1.5 to 0.7 km/h higher speeds than passenger cars, and 2.4 to 1.6 km/h higher speeds than pickup trucks.

Keywords: Object Tracking, Vehicle Recognition, Speed Inference, YOLOv4, DeepSORT

INTRODUCTION

Video surveillance cameras have been leveraged by municipalities globally to remotely enforce traffic laws, monitor system performance, and maintain roadway safety with reduced bias (“Europe’s Noise Capital Tries to Turn Down the Volume,” 2022; Frangoul, n.d.; Melecki, 2013; Paybarah, 2019; Tewolde, 2012; Wilson et al., 2010). These devices can be programmed to provide vehicle/speed tracking, identify red-light violations, read license plates, apply tolls, and analyze pedestrian-vehicle interactions (Al-Smadi et al., 2016; Jain et al., 2016; Makino et al., 2018; Manlises et al., 2015; Neuhold et al., 2019; Pesti et al., 2008; Skogan, 2019). Pedestrian crashes in particular have been rising in the US over the past decade with vehicles becoming larger and traveling faster (Ballesteros et al., 2004; Bernhardt and Kockelman, 2021; Billah et al., 2021; Rosenthal et al., 2022; Zuniga-Garcia et al., 2022). The spatial distribution between pedestrian accidents (n=3,576 from January 2012 to February 2022) and city-managed traffic cameras (n=369) was analyzed to identify high-risk intersections in Austin (“CRIS Query,” n.d.). Austin’s Mobility Management Center then shared 12 hours of footage from 3 high-crash sites, recorded between 1 and 3pm on Tuesday, April 5, 2022, as an example of what those videos look like.

METHODS

30 minutes of video from the Loop 360 at Courtyard Drive intersection (shown in Figure 1) were used to demonstrate application opportunities. A Vehicle Make and Model Recognition (VMMR) model was fine-tuned using the ResNet-152 Cars Dataset, containing 16,185 images of the 196 vehicles. Vehicles observed heading in the southbound direction were identified by Make, Model, and Year, (e.g., a 2012, Tesla, Model S) (Krause et al., 2013). Image data was split into 50% training and 50% testing sets. Object tracking features were also implemented in this VMMR model, using YOLOv4, DeepSORT, and TensorFlow. Figure 1’s dark blue region detects vehicles upstream of the signal light, and the light green region defines the vehicle recognition area as shown in Figure 1. Vehicle speeds were estimated by comparing video frame rate (30 frames per second) with time taken by vehicles moving through the intersection to cross the red region (from Fig 1’s Speed 1 area to Speed 2 area).
Figure 1. Areas for Vehicle Detection and Speed Inference

Figure 2. Loop 360 and Courtyard Drive Intersection Location

FINDINGS
These speed values were used as the dependent variable \( Y = \text{vehicle speed} \) in the ordinary least squares (OLS) model, with vehicle body types, lane position, signal light status, and other AI-inferred variables used as covariates to ascertain the relationship between speed and body type. Tables 1 and 2 provide summary statistics and OLS results of analyzed variables. Results underscore how sport and crossover utility vehicles (SUVs/CUVs) and motorcycles generally travel faster than cars and trucks at this intersection. The average speed of SUVs/CUVs was 1.4 km/h faster than cars, 2.3 km/h faster than trucks, 0.7 km/h faster than vans, and 0.1 km/h slower than motorcycles. Those owning SUVs and CUVs may notice ground speed less and may be sensation seeking. Motorcyclists have better peripheral vision than vehicle drivers on intersection approaches, which may be why they also move faster through intersections.

Table 1. Summary Statistics of Variables Inferred \((n = 683)\)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>Median</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed</td>
<td>37.727</td>
<td>40.5</td>
<td>11.008</td>
<td>3.393</td>
<td>72</td>
</tr>
<tr>
<td>Motorcycle</td>
<td>0.007</td>
<td>0</td>
<td>0.085</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>0.075</td>
<td>0</td>
<td>0.263</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car</td>
<td>0.521</td>
<td>1</td>
<td>0.500</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Truck</td>
<td>0.078</td>
<td>0</td>
<td>0.268</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUV/CUV</td>
<td>0.187</td>
<td>0</td>
<td>0.391</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Van/Minivan</td>
<td>0.132</td>
<td>0</td>
<td>0.338</td>
<td></td>
<td></td>
</tr>
<tr>
<td>light_state_green</td>
<td>0.876</td>
<td>1</td>
<td>0.330</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lane_1</td>
<td>0.589</td>
<td>1</td>
<td>0.492</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Car class includes Convertibles, Coupes, Sedans, and Hatchbacks.
Table 2. Results of OLS Regression Model to Predict Y = Speed (in km/hr)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Const</td>
<td>20.255</td>
<td>1.138</td>
<td>1.780</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Motorcycle</td>
<td>4.122</td>
<td>3.638</td>
<td>0.628</td>
<td>1.13</td>
<td>0.258</td>
</tr>
<tr>
<td>Other</td>
<td>4.401</td>
<td>1.320</td>
<td>0.628</td>
<td>3.34</td>
<td>0.001</td>
</tr>
<tr>
<td>Car</td>
<td>2.652</td>
<td>0.820</td>
<td>0.312</td>
<td>3.24</td>
<td>0.001</td>
</tr>
<tr>
<td>Truck</td>
<td>1.729</td>
<td>1.298</td>
<td>0.359</td>
<td>1.33</td>
<td>0.183</td>
</tr>
<tr>
<td>SUV/CUV</td>
<td>4.010</td>
<td>0.991</td>
<td>0.490</td>
<td>4.05</td>
<td>0.000</td>
</tr>
<tr>
<td>Van/Minivan</td>
<td>3.342</td>
<td>1.098</td>
<td>0.476</td>
<td>3.04</td>
<td>0.002</td>
</tr>
<tr>
<td>light_state_green</td>
<td>17.279</td>
<td>1.092</td>
<td>1.634</td>
<td>15.83</td>
<td>0.000</td>
</tr>
<tr>
<td>lane_1</td>
<td>-1.229</td>
<td>0.738</td>
<td>-2.021</td>
<td>-1.67</td>
<td>0.096</td>
</tr>
</tbody>
</table>

Sample Size: 683 vehicles
R-squared: 0.277
Adj. R-squared: 0.269
Dep. Variable: Y= Speed (estimated in km/hr)
Standard Error (eY): +/- 9.41 km/hr
REFERENCES


