

Finding the Right Entrance for Autonomous Driving: Building Entrance Detection and Door-to-Building Association from Drive-by Imagery

Author: Sia Xu¹ Supervisor: Neil Oliver¹ Industry partner: HERE Technologies

¹The London School of Economics and Political Science

Background and Motivation

Autonomous vehicles navigate to an address — but passengers need to be dropped off at the right **entrance**, and entrances are rarely recorded in standard map data. Existing work largely treats drop-off planning, entrance detection and building-to-map matching separately. This project focuses on the entrance-level link: which mapped building does each observed entrance belong to? With HERE Technologies, this work addresses that gap.

Data and Methods

Data. 5,634 labelled street-level images and 10 continuous drive sequences covering 543 physical entrances, together with camera poses and OpenStreetMap building footprints. The task is split into two measurable stages:

- **Stage 1 — Entrance detection.** A YOLOv9c detector combines full-image context with high-resolution crops to preserve small and distant entrances. Repeated observations of the same doorway are grouped into physical entrance events.



Why event-level evaluation? The same physical entrance is observed repeatedly along a drive.

- **Stage 2 — Door-to-building association.** Each detected entrance is projected from the camera onto mapped building footprints (Figure 1). Experiments identified **camera-position error** as the main geometric bottleneck, so the pose is refined using facade evidence and route consistency.
- **Scene-conditioned routing.** No single correction works everywhere, so observable camera and street conditions determine which association path is used (Figure 2).



Figure 1: From a detected entrance to its mapped building.

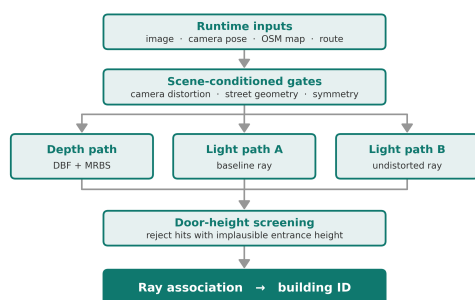


Figure 2: Runtime scene conditions select the association path.

Key Findings

- **Detection.** On continuous drives, frame recall is 50.4%, but repeated observations raise physical-entrance discovery to 70.7%. Small scale and occlusion remain key challenges (Figure 3).

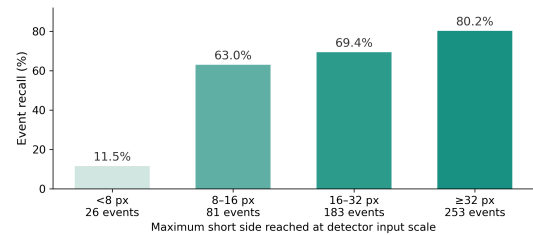


Figure 3: Repeated visibility improves entrance discovery.

- **Association.** The adaptive pipeline improves accuracy from 52.3% to 58.7% on the development sequences, and from 18.4% to 48.5% on the held-out stress test (Figure 4).

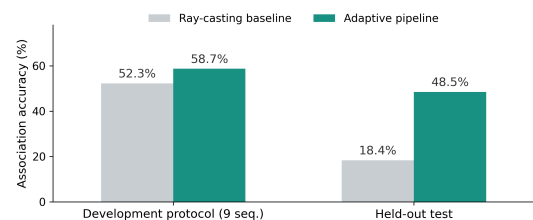


Figure 4: Association accuracy improves on development and held-out sequences.

- **End-to-end.** Replacing annotated boxes with detector predictions changes only 1 of 667 matched building associations. The interface itself is not the main bottleneck; the remaining losses come from missed detections and association geometry.

Value of the research

This work turns "stopping at the right door" into two measurable, independently diagnosable stages, providing the entrance-level evidence layer that standard maps lack. It shows that deployment should be judged at the drive level — repeated observations are an asset — and that geometric corrections must be scene-conditioned: the same fix helps one device or street type and harms another.

The pipeline could support enrichment of digital map layers with door-to-building associations, potentially benefiting drop-off planning, accessibility routing and curbside applications. Its diagnostics also identify the main priorities for future work: improving drive-domain detection and resolving ambiguous building candidates (Figure 5).

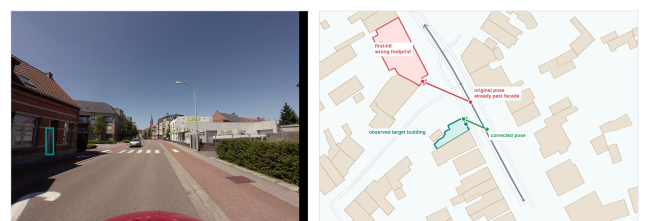


Figure 5: Pose correction restores the correct building hit.