NavOCR: A Dataset Generator for Navigation-Relevant Text Detection in Mobile Robots

KC Machine Learning Lab

NeurIPS 2026 Evaluations & Datasets Track

NEURAL INFORMATION PROCESSING SYSTEMS

KC ML2

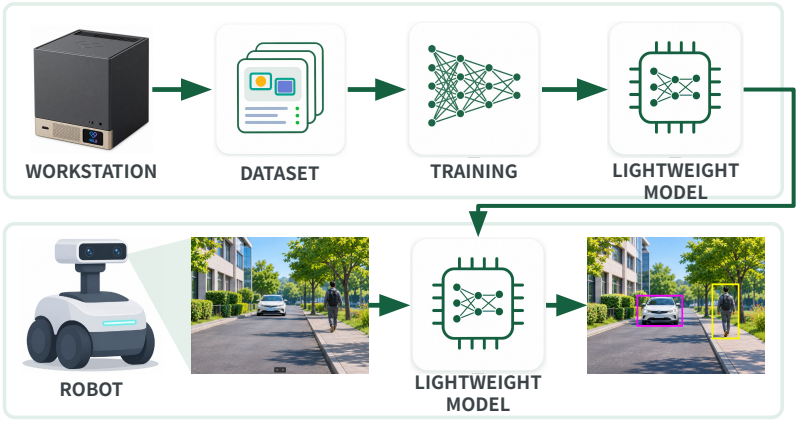
Why a Dataset Generator in the LLM Era?

Modern VLMs/LLMs can extract almost any semantic information.
However, they are still too heavy for mobile robots.

We propose a **workstation-to-robot pipeline**:

1. **Workstation**: Dataset generation & Training
2. **Robot**: Lightweight inference

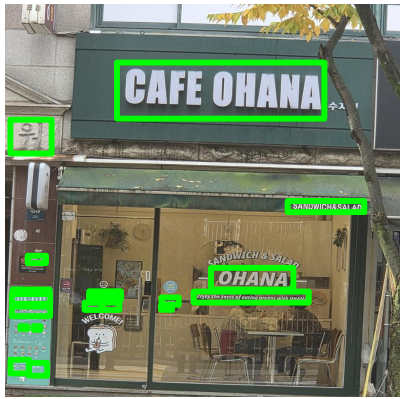
We first focus on **text as a landmark** for semantic navigation.



Key Idea

Conventional OCR detects **all** visible text.
(ex. ads, price tags, menus)

However, mobile robots need **only navigation-relevant text**.
(ex. signboard, directional guides, room numbers)



OCR
(PP-OCRv5)



NavOCR



OCR
(PP-OCRv5)



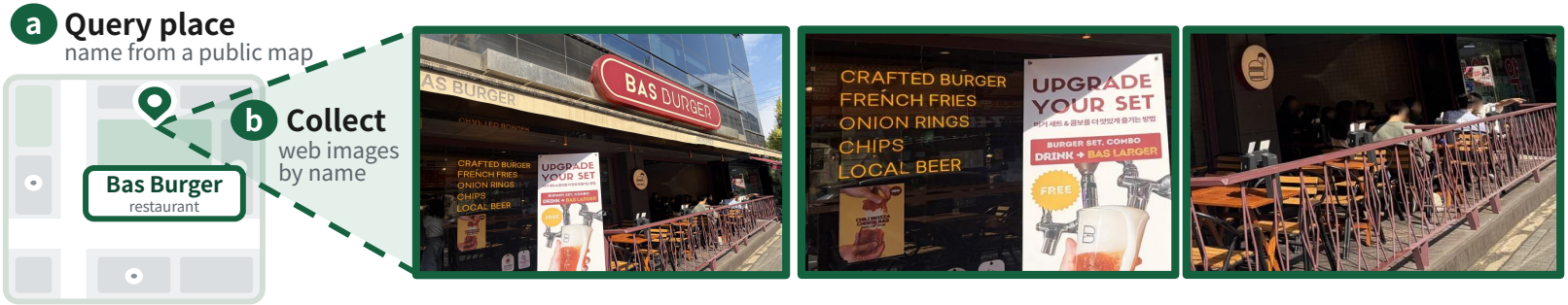
NavOCR

NavOCR

NavOCR: Our dataset generator for **navigation-relevant text** detection

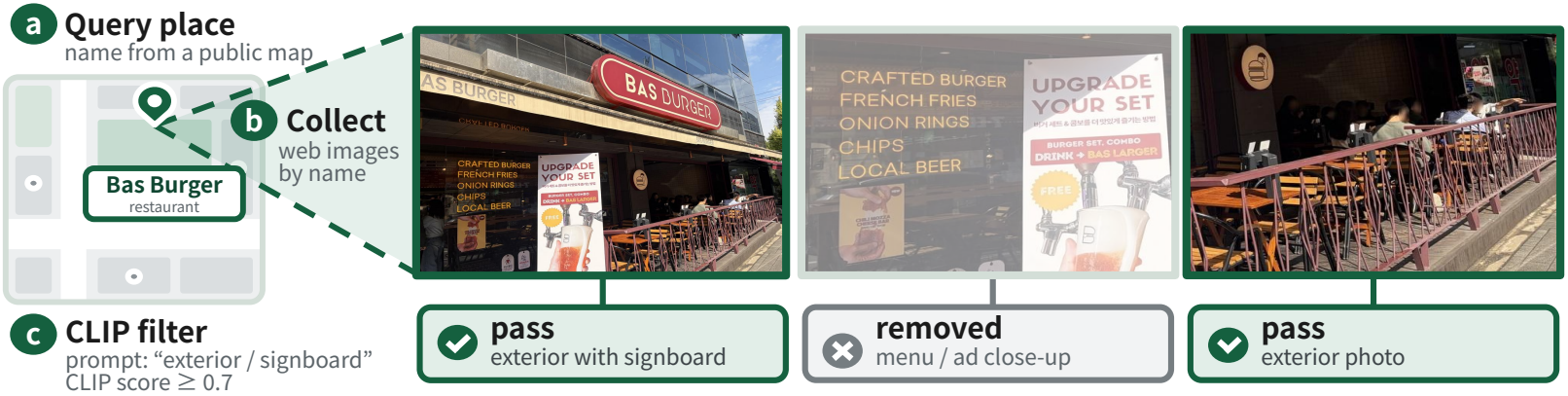
NavOCR

NavOCR: Our dataset generator for **navigation-relevant text** detection



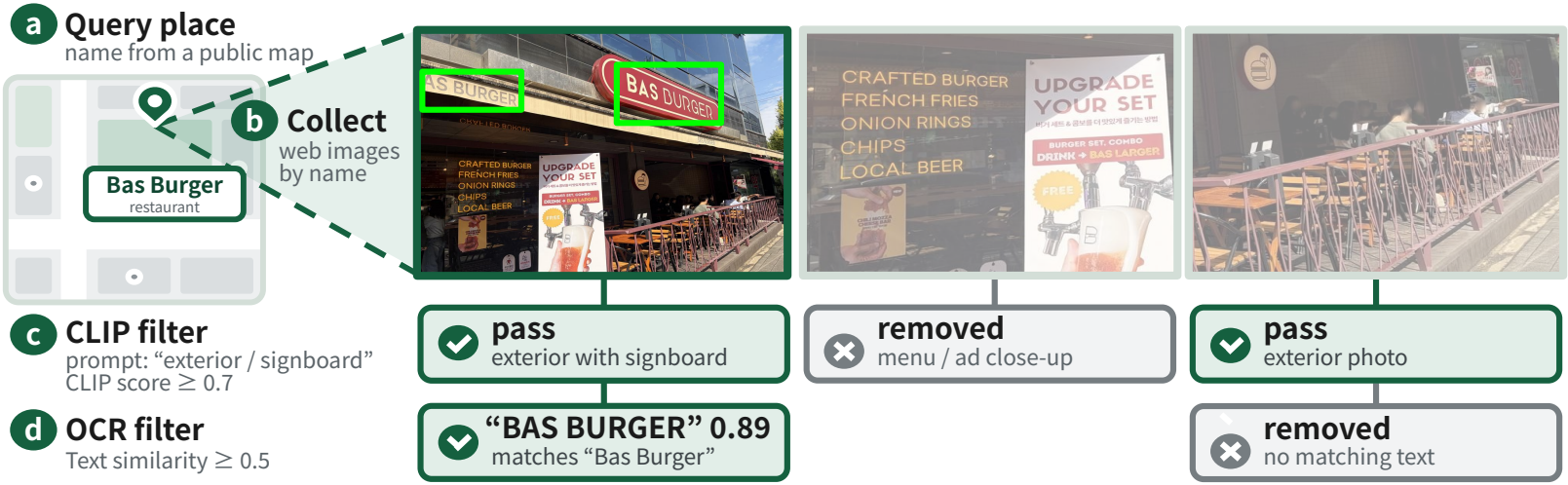
NavOCR

NavOCR: Our dataset generator for **navigation-relevant text** detection



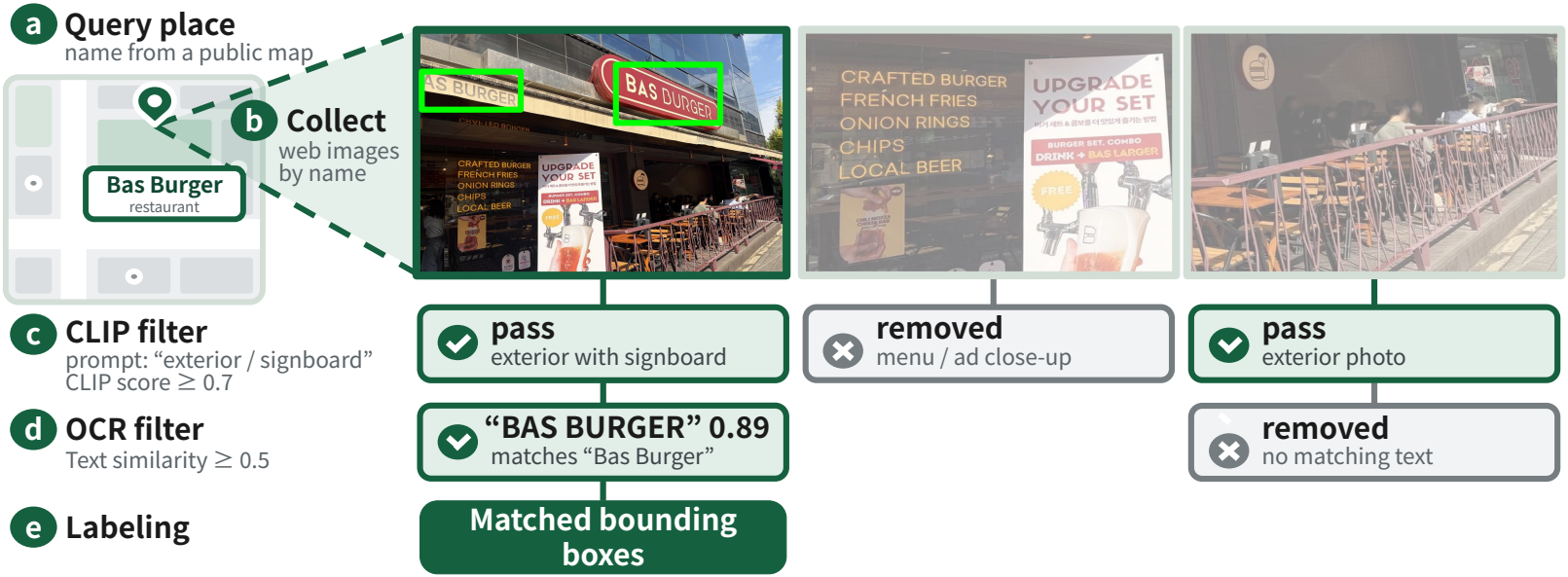
NavOCR

NavOCR: Our dataset generator for **navigation-relevant text** detection



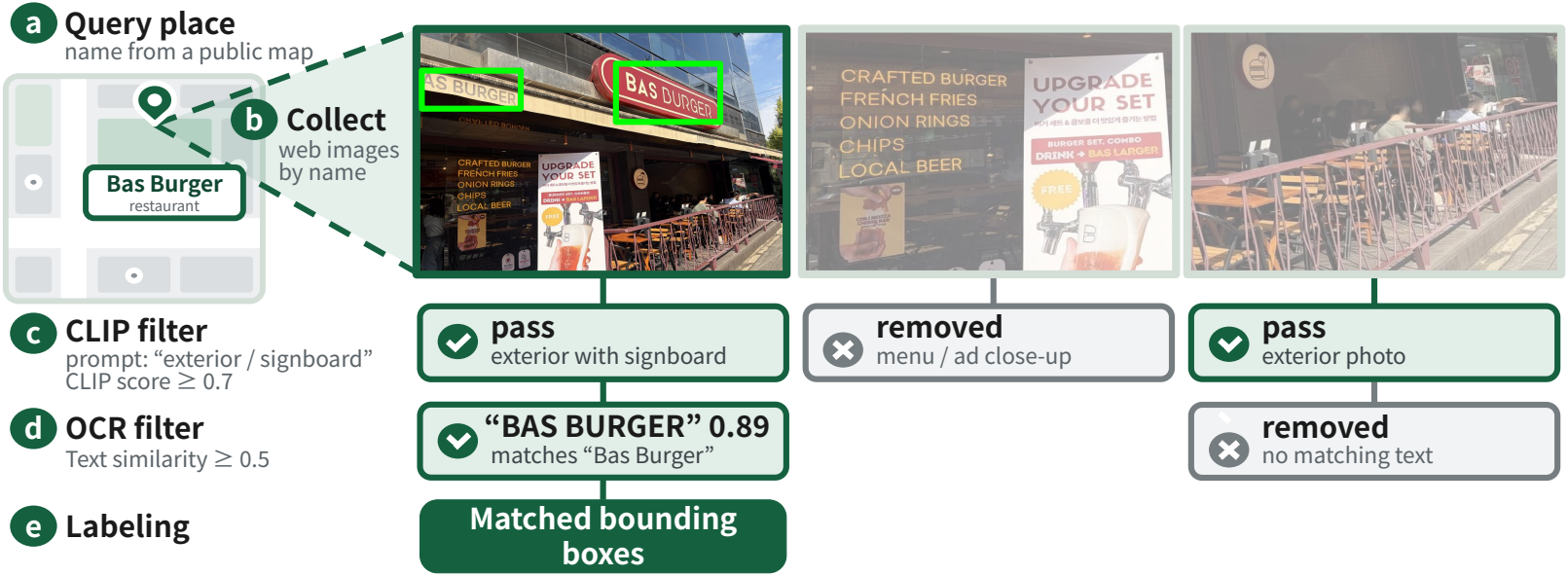
NavOCR

NavOCR: Our dataset generator for **navigation-relevant text** detection



NavOCR

NavOCR: Our dataset generator for **navigation-relevant text** detection



Contribution: Our text-specific pipeline generates **higher-quality datasets than LLM**-based approaches.

Experiment Result

Generated Dataset Quality

NavOCR generator
Given the place name
from map



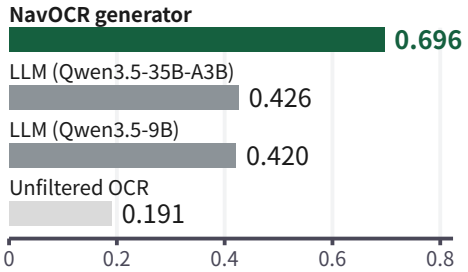
GROUNDB

LLM generator
(Qwen3.5-9B)
Given no place name



GROUNDB, SANDWICH, WELC,
BEER&WINE, LEA, ONE,

correct wrong



F1 score vs. human labels

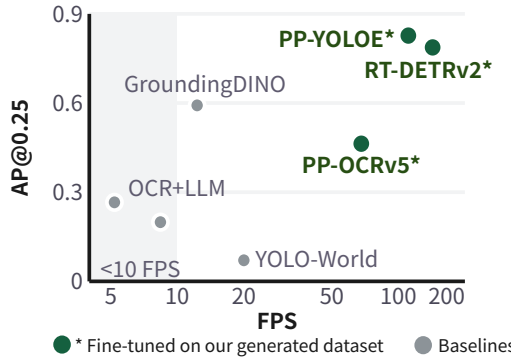
Detection Benchmark

NavOCR



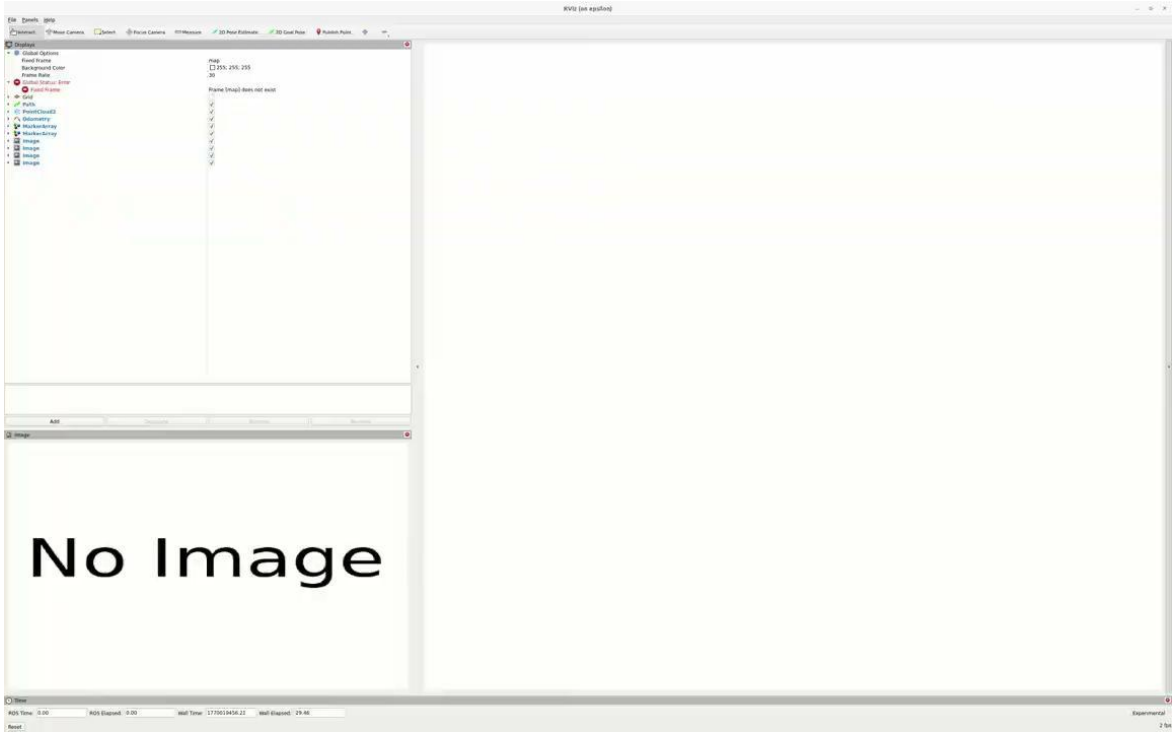
Grounding DINO





AP@0.25 vs. FPS (on RTX 4090)

Downstream Tasks



Beyond Text: Any User-Defined Landmark

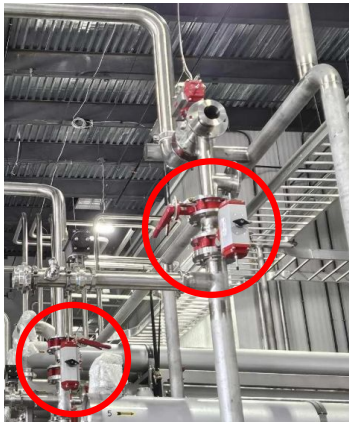
Our **workstation-to-robot pipeline** enables site-specific, lightweight robotic inference.

Any **user-defined** visual feature can be a landmark.

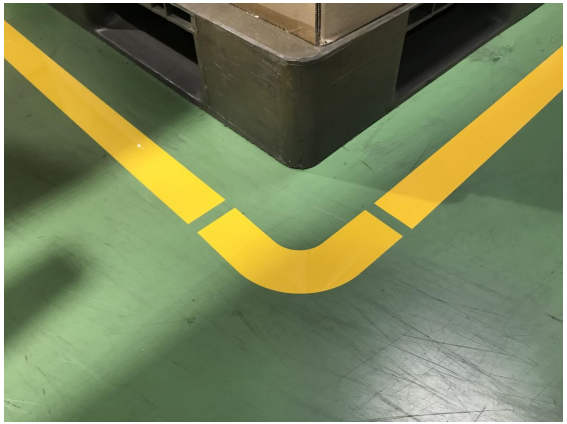
Generate on workstations. Infer on robots.



Loading docks



Pipe components



Floor tape junctions

Thank for listening!

<Our poster>

Date: Tuesday, December 8, 2026
Time: 10:00 AM – 1:00 PM
Session: Sydney Poster Session 1
Location: Hall 1–4

code

KC ML2

Text mapping example