1. Raw Visual Feed
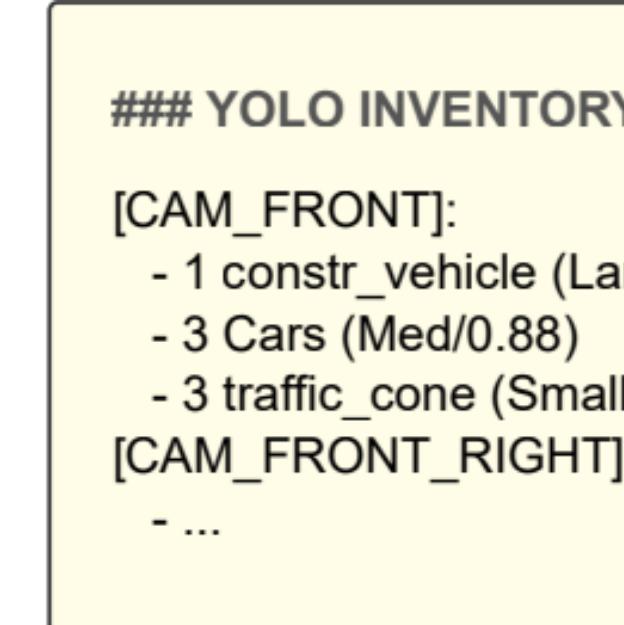*(Front Camera)*
(Front Right Camera)
(Front Left Camera)

2. Open-Vocabulary Segmentation
(YOLOE-11 / WOD Taxonomy)

YOLO INVENTORY

[CAM_FRONT]:
   - 1 constr_vehicle (Large/0.94)
   - 3 Cars (Med/0.88)
   - 3 traffic_cone (Small/0.45)
[CAM_FRONT_RIGHT]:
   - ...

3. Prompt Injection
(Structured Text)

High Recall Policy
Low conf (0.45) retained
for VLM verification