AI-driven Hazard Detection for Workplace Safety in Construction Industry
Donia Gasmi

Professor :   Alcherio Martinoli
Assistant(s) :   Kagan Erünsal

The construction industry remains one of the most hazardous occupational sectors, with thousands of injuries and fatalities occurring annually despite existing safety protocols. While computer vision offers promise or automated safety monitoring, current approaches are hindered by fragmented datasets, poor generalization across real-world sites, and an inability to interpret contextual relationships, such as whether personal protective equipment (PPE) is being worn.

This internship project addresses these limitations through a unified, cluster-based vision pipeline for construction-site hazard detection. We introduce a novel dataset design that groups semantically related objects, such as PPE, vehicles, cranes, and safety structures, into scale, and context-homogeneous clusters. Critically, we demonstrate that training on partially labelled datasets induces catastrophic forgetting, motivating the creation of two large-scale, fully annotated cluster datasets (4,147 PPE images and 15,148 vehicle/machinery images) via a semi-automated pipeline combining Grounding DINO for open-vocabulary detection, the Segment Anything Model (SAM) for segmentation, and active learning for efficient human-in-the-loop refinement.

Among multiple architectures explored, including hybrid detection–classification pipelines and custom designs like GSO-YOLO and a DINOv3–YOLOv12 fusion. The YOLOv8x model achieved the strongest performance, with 99.3% mAP@50 on vehicles and 84.3% mAP@50 on PPE. To enable contextual safety reasoning, we developed a spatial predicate layer that fuses outputs from a YOLO-Pose model and the PPE detector to distinguish, for example, between a helmet being worn versus held in hand.

A lightweight User Interface (GUI) was implemented to demonstrate real-time inference, and a domain adaptation experiment confirmed that fine-tuning on just 300 site-specific images significantly improves performance in new environments.

By integrating rigorous data engineering, robust detection, and interpretable spatial reasoning, this work advances the feasibility of deploying intelligent, automated safety systems on real construction sites.

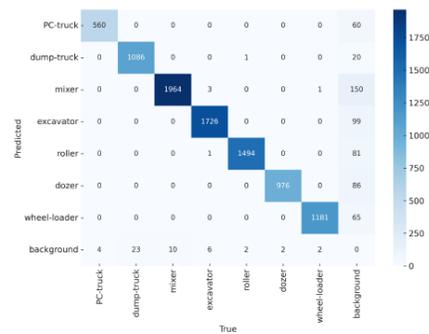*Figure 1 Overview of the five object clusters defined for construction-site hazard detection.*

Figure 2. Confusion matrix – Vehicles & Machinery. The model achieved near-perfect accuracy across all seven machinery categories, with minimal background confusion.
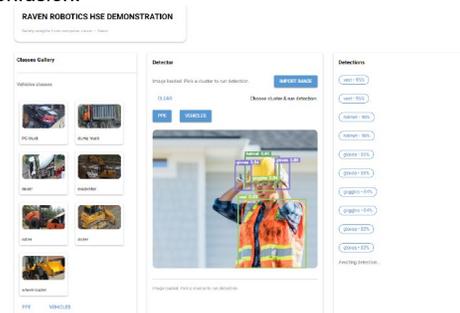

Figure 3. GUI demonstration for PPE model.