

Real-Time PPE Detection for Construction Sites

M4U3 Assignment | MSc AI for Architecture & Construction
Carlo Cogni | February 2026

YOLOv8n | 10 Classes | Real-Time Inference

Problem & Approach

AECO Challenge

Manual PPE compliance monitoring is labour-intensive and error-prone on active construction sites.

Our Solution

Automated detection of 10 PPE classes using YOLOv8 Nano, optimised for real-time CCTV feeds (>30 FPS target).

Success Criteria

High precision (reduce alert fatigue) + sufficient recall to act as a reliable supplementary safety tool.

10 Detection Classes

- Person
- Hardhat / NO-Hardhat
- Mask / NO-Mask
- Safety Vest / NO-Safety Vest
- Safety Cone
- Machinery
- Vehicle

Version 1 — Baseline Results

2,605 images | 30 epochs | YOLOv8n

Metric	Value
Precision	0.876
Recall	0.677
mAP@50	0.754
mAP@50-95	0.441
Inference	1.9 ms (~500 FPS)

Key Takeaways

PPE presence detection is strong
(Hardhat 84%, Mask 86%, Vest 81%)

PPE absence detection is weaker
(NO-Mask 56%, NO-Hardhat 67%)

Vehicle class underperforms (47% mAP)
due to visual diversity & occlusion

Model is highly efficient at ~500 FPS,
suitable for live CCTV integration

Version 2 — Audited Model

Clean Seed Strategy | 612 audited images | 3x augmentation

Intervention:

100% manual review of 204 high-priority images
to correct bounding box inaccuracies.
3x augmentation multiplier applied.

Class	mAP@50	Recall	Impact
Vehicle	0.663	1.00	Zero-miss machinery
Hardhat	0.677	0.82	Stable head protection
Safety Vest	0.535	0.79	Torso verification
Person	0.498	0.82	Worker identification

Inference: 3.0 ms / image

(~333 FPS on Tesla T4)

Recommended deployment weights: best_v2.pt

Error Analysis

Key failure modes identified across V1 & V2

False Positives

- FP-1: Equipment colour confused as PPE
(yellow bucket flagged as Hardhat)
- FP-2: Reflective surfaces flagged as Vest
(glass cladding, metallic scaffolding)
- FP-3: Rounded objects detected as Hardhats
(buckets, lids carried by workers)
- FP-4: Background-to-Person leakage: 26%
(poles, scaffolding, cables)

False Negatives

- FN-1: Small objects missed at distance
(gloves < 10m resolution limit)
- FN-2: Partial occlusion (>40% covered)
(workers behind pallets, rebar)
- FN-3: Low-light / shadow conditions
(contrast deficiency in trenches)
- FN-4: Hardhat-Mask confusion: 9%
(spatial overlap in head region)

Prioritised Improvements

1. Tiling Preprocessing

Slice 4K images into 640x640 tiles to boost small-object detection (FN-1)

2. Hard Negative Mining

Add 150+ empty-site images to reduce 26% background-to-person leakage (FP-4)

3. Contrast Augmentations

Mosaic + random brightness to handle harsh lighting transitions (FN-3)

4. Head-Region Differentiation

Increase class weight or resolution to 1280 for Hardhat/Mask separation (FN-4)

Governance & Reproducibility

Governance

- No PII/biometric processing; PPE objects only
- Supplementary tool, not sole safety mechanism
- Safety-First Bias: recall prioritised over precision

Licensing

- Code: MIT License
- Dataset: Public (Kaggle/Roboflow), educational use
- YOLOv8: AGPL-3.0 (review for commercial use)

Reproducibility

- Colab notebooks run end-to-end from GitHub badges
- Auto-downloads dataset + weights; no local setup needed

GitHub Repository

[github.com/CarloCogni/
computer-vision-with-YOLO](https://github.com/CarloCogni/computer-vision-with-YOLO)