PPE Detection on Construction Sites

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

Executive Summary

This project implements an automated Personal Protective Equipment (PPE) detection system for construction sites using YOLOv8n (Nano). The model detects 10 classes — hardhats, masks, safety vests (and their absence), people, safety cones, machinery, and vehicles — across two iterative versions. Version 1 established a strong baseline on 2,605 images, while Version 2 applied a "Clean Seed" audit strategy on 612 curated images to maximise data integrity and safety-critical reliability. Both models achieve real-time inference speeds suitable for live CCTV integration.

Version 1: Baseline Performance

Trained for 30 epochs on 2,605 auto-labelled images (Kaggle/Roboflow), the baseline model achieved **75.4% mAP@50** and **87.6% precision** with an inference speed of **1.9 ms/image** (~500 FPS on RTX 4070). Detecting PPE *presence* proved substantially easier (Hardhat 84%, Mask 86%, Vest 81%) than detecting its *absence* (NO-Hardhat 67%, NO-Mask 56%). Vehicle detection was the weakest class at 47.4% mAP@50 due to high visual diversity and occlusion.

Metric	Value
Precision	0.876
Recall	0.677
F1 Score	0.76

Metric	Value
mAP@50	0.754
mAP@50-95	0.441
Inference	1.9 ms

Version 2: Audited Model

A 100% manual audit of 204 high-priority images corrected bounding box inaccuracies in Hardhat and Safety Vest classes. With a 3x augmentation multiplier producing 612 training images, the audited model achieved notably higher reliability in safety-critical classes. Vehicle detection reached **1.00 recall** (zero missed heavy machinery), Hardhat detection improved to **0.677 mAP@50**, and Safety Vest detection achieved **0.79 recall**. Inference speed remained at **3.0 ms/image** (~333 FPS on Tesla T4).

Class	mAP@50	Recall	Strategic Impact
Vehicle	0.663	1.00	Zero-miss machinery detection
Hardhat	0.677	0.82	Stable head-protection monitoring
Safety Vest	0.535	0.79	Torso-level PPE verification
Person	0.498	0.82	Consistent worker identification

Error Analysis Summary

False Positives: Equipment colour confusion (yellow buckets detected as Hardhats), specular reflections flagged as Safety Vests, rounded objects misidentified as helmets, and a 26% background-to-person leakage rate from vertical site structures.

False Negatives: Small-object detection gap at >10m distances, partial occlusion (>40% coverage) causing missed PPE, low-light/shadow contrast deficiency, and a 9% Hardhat-Mask cross-class confusion due to spatial proximity in the head region.

Prioritised improvements: (1) Tiling preprocessing for small-object resolution, (2) Hard negative mining with 150+ empty-site images, (3) Contrast-robust augmentations for harsh lighting, (4) Head-region differentiation via higher resolution or class weighting.

Limitations & Future Work

The model is optimised for daytime, outdoor conditions; performance may degrade in low-light, night, or indoor environments without retraining. It should not be used as the sole compliance mechanism, for worker identification, or as legal evidence without human verification. Production deployment should incorporate multi-frame temporal smoothing to reduce false alerts from momentary occlusions. The Safety-First Bias principle recommends tuning the model to favour recall over precision in deployment — it is better to over-alert than to miss a genuine safety violation.

Governance & Licensing

No personally identifiable information or biometric data is processed; the model detects PPE objects only. All training data originates from public Kaggle/Roboflow datasets. The project code is released under the MIT License; the YOLOv8 framework is AGPL-3.0 (commercial deployment requires review of Ultralytics licensing). For on-site deployment, workers must be informed of computer vision monitoring per local labour regulations.

Full results, training curves, error analysis, and governance documentation: github.com/CarloCogni/computer-vision-with-YOLO