

Real-Time Object Detection using YOLO

CCTV cameras record everything and understand nothing. Operators cannot watch dozens of feeds round the clock, so incidents surface only after the fact. This project builds a real-time object detection system using YOLOv8 that turns any ordinary camera into an intelligent sensor: it watches the live stream, recognises people, vehicles and everyday objects, and raises an alert the moment something enters a user-defined intrusion zone.

YOLO (You Only Look Once) frames detection as a single regression problem. Instead of proposing regions and then classifying them, it divides the frame into a grid and predicts bounding boxes, class probabilities and confidence scores in one forward pass of a convolutional backbone. YOLOv8's anchor-free decoupled head keeps this single-pass speed while pushing COCO accuracy higher — roughly 25–30 FPS on a modest GPU and 8–12 FPS on a laptop CPU.

The pipeline is built to demo well. OpenCV grabs frames from a USB webcam, Raspberry Pi camera or IP stream; each frame is resized to 640×640 and passed through YOLOv8; non-maximum suppression removes duplicate boxes; the annotated frame — boxes, labels, confidence scores — streams to the VisionGuard dashboard over Flask MJPEG. Zones are drawn once on a reference frame; every detection centroid is tested against the polygon, and a breach logs an event, captures a snapshot and fires an alert. The nano variant (YOLOv8n) runs on a Raspberry Pi 4 at usable frame rates, so the same code scales from laptop to edge hardware.

The project ships with pretrained COCO weights and a fine-tuning pipeline for custom classes — helmets, number plates, lab equipment. Evaluation uses precision, recall and mAP@0.5 on a held-out set, giving the report real numbers and the viva a solid technical narrative.