

Abstract — “You Only Look Once: Real-Time Object Detection with Neural Networks”

Real-time object detection must localize and classify multiple objects fast enough for on-device uses such as assistive robotics, AR, and safety monitoring. This poster explains how YOLO achieves high throughput by predicting boxes, objectness, and classes in a single forward pass (Backbone \rightarrow Neck \rightarrow Head) followed by non-maximum suppression. We summarize the box-regression objective (IoU-family losses) and analyze the resolution trade-off—higher input size typically improves AP but reduces FPS—using literature data from recent YOLO variants. We complement this with a practical playbook: where YOLO excels (clear, near objects) and where it fails (tiny, occluded, low-light, motion-blurred scenes), plus quick fixes without retraining (input-size tweaks, threshold/NMS tuning, lightweight tracking). A small, reproducible demo with pretrained weights illustrates these ideas on consumer hardware. Overall, the poster provides an intuitive map from architecture to deployment choices when balancing accuracy and latency in real-time detection.