

Abstract

Object detection using deep learning faces significant challenges in real-world scenarios with occluded objects. The YOLOv8n architecture, popular for real-time detection, can potentially be enhanced using attention modules. This research evaluates the effect of the insertion location of two attention modules, CBAM (Convolutional Block Attention Module) and CoordAtt (Coordinate Attention), on YOLOv8n's performance in detecting occluded objects. The evaluation focuses on a modified KITTI dataset representing varying occlusion levels across six combined object classes (vehicles and pedestrians). The core problem is how insertion location (Neck, Backbone, or Both) and attention type (CBAM vs. CoordAtt) affect YOLOv8n's ability to handle occlusion. Experimental results consistently show that inserting attention modules into the Neck section yields the most significant performance improvement. The CBAM-Neck configuration achieved the highest mAP@0.5 (0.683) and mAP@0.5:0.95 (0.476), a substantial improvement over the baseline (0.659 and 0.449). Qualitative analysis confirms enhanced detection of moderately occluded objects with Neck attention, although severe occlusion remains challenging. Backbone or Both insertions were less effective.

Keywords: YOLOv8, CBAM, CoordAtt, object detection, occlusion, attention, KITTI.