Model Architecture: YOLOv8

A sophisticated deep learning model called YOLOv8 was created for detecting objects in real time. Through the integration of a more streamlined architecture, it enhances speed and accuracy while building upon the efficiencies of its predecessors. YOLOv8 uses a deep convolutional neural network (CNN) to predict bounding boxes and class probabilities in a single forward pass by processing the full image during training and inference. This method is speedier and better suited for applications that need real-time processing because it differs greatly from region proposal-based approaches.

Training Methodology

Dataset Preparation: The Roboflow platform was used to annotate a custom dataset that was produced by gathering images from Google. To make sure the model could generalize to varied situations, this dataset contained a variety of object classifications.

Preprocessing: Images were resized to fit the 640x640 pixel input requirements of the YOLOv8 model and their pixel values were normalized. To improve the model's resilience to varying object orientations and scales, data augmentation techniques like rotation, scaling, and horizontal flipping were used.

Model Configuration: A batch size of 16 was used for training across 10 epochs, and a scheduler was used to modify the learning rate, which was initially set at 0.001, in response to the loss reduction plateau. The model used IoU (Intersection over Union) loss for bounding box prediction and cross-entropy loss for classification.

Training Environment: To take advantage of GPU acceleration and guarantee quick training times, the PyTorch backend with CUDA support was used.

Evaluation Metrics

The accuracy of the model and its capacity to reduce false negatives were assessed using precision and recall metrics.Precision quantifies how well optimistic predictions work.The ability to locate every pertinent instance in the dataset is indicated by recall.
Mean Average Precision (mAP): mAP was a critical performance metric that assessed the accuracy of the model across various IoU levels. For assessing overall performance in object detection tasks, this statistic is essential.

Results

- Precision and Recall: The model achieved a precision of 66.38% and a recall rate of 53.99%, proving its ability to recognize meaningful objects.
- Mean Average Precision (mAP): The model achieved a mAP of 58.09% at an IOU (Intersection Over Union) of 50%, suggesting strong detection skills.
- Performance Insights: The testing indicated areas for improvement, particularly in detecting tiny objects and in scenarios including overlapping objects.