Object Detection and Pose Estimation for Golf Shot Analysis Using YOLOv8

Introduction

Object detection and pose estimation play a significant role in sports analytics, where movement precision and posture directly influence performance. In golf, swing mechanics are critical for accuracy, power, and consistency. In this project, an object detection and pose estimation system was developed to analyze golf shot movements using YOLOv8 implemented through the Ultralytics framework. The objective of this assignment was to construct a custom golf shot dataset from video clips, train an object detection model to detect golfers, evaluate model performance using standard metrics, and apply pose estimation to analyze swing posture and body alignment. The project demonstrates an end-to-end deep learning workflow applied to sports motion analysis.

Dataset

The dataset was created from multiple golf shot video clips capturing different swing phases. From these clips, frames were extracted and manually annotated using bounding boxes in YOLO format. Each annotation file contains the class ID along with normalized bounding box parameters representing the object's center coordinates, width, and height. The dataset includes one primary class labeled as "golfer."

The images were divided into training and validation sets using an approximate 80:20 split to ensure proper evaluation. A YAML configuration file was created to define dataset paths, training and validation directories, number of classes (nc = 1), and class names. The dataset link is provided separately as required in the submission. Although the dataset is relatively compact, it effectively represents different phases of the golf swing, including backswing, downswing, and follow-through.

Methodology

Object Detection Model

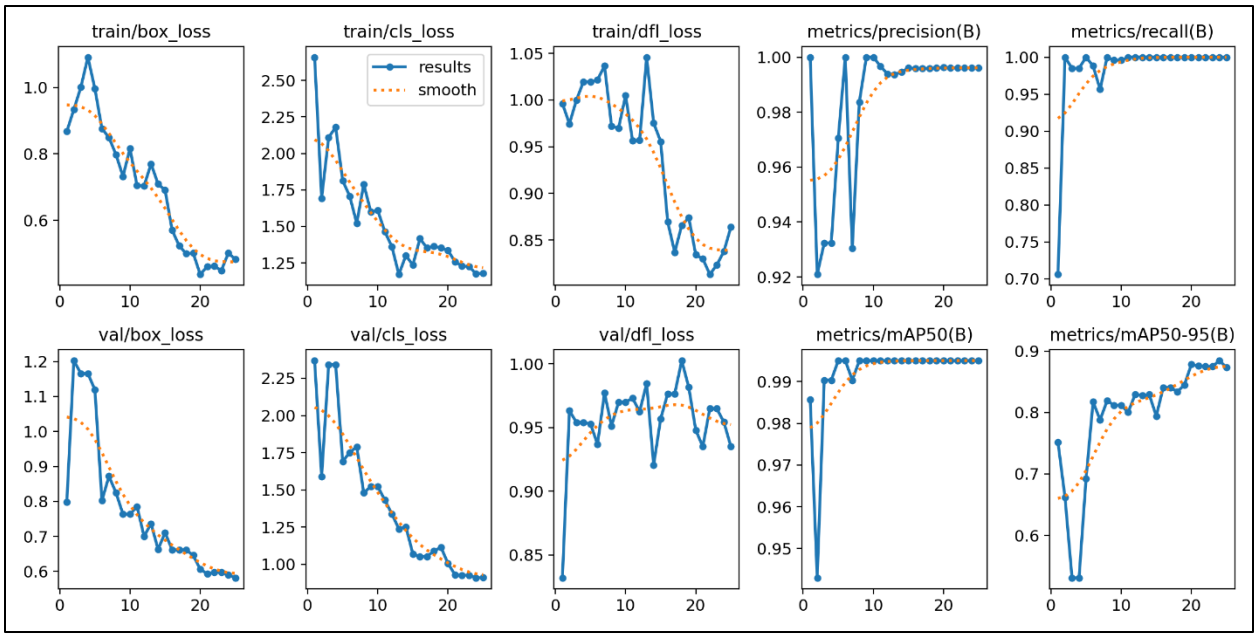
The object detection model was implemented using the YOLOv8 nano (YOLOv8n) architecture due to its efficiency and lightweight design. The model was trained for multiple epochs with an input image resolution of 640 × 640 pixels. During training, the optimizer and learning parameters were automatically selected by the framework to ensure stable convergence.

The training process monitored key loss components, including box loss for bounding box regression accuracy, classification loss for identifying the golfer class, and distribution focal loss for improved localization precision. The loss curves demonstrated a consistent downward trend across epochs, indicating effective learning and convergence of the model. Sample validation outputs confirmed that the trained model successfully detected the golfer in various swing positions.

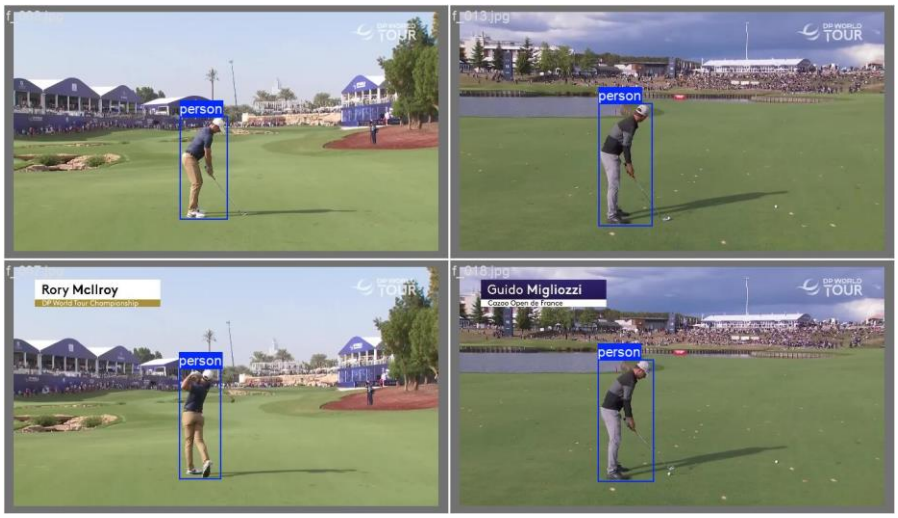
Pose Estimation

Following object detection, pose estimation was applied using a pretrained YOLOv8 pose model. The model detects key human joints, including shoulders, elbows, wrists, hips, knees, and ankles. These keypoints are particularly important for analyzing golf swing biomechanics. The pose model was applied to the golf shot videos, generating output sequences with skeleton overlays that visually represent body posture during different swing phases. The detected keypoints allow for analysis of body rotation, arm extension, knee flexion, and overall posture alignment. The results demonstrate stable and accurate keypoint tracking throughout the swing motion, highlighting the potential application of deep learning in automated sports performance evaluation.

Performance Evaluation



Bounding Box detection



Keypoint detection

Conclusion

In conclusion, this project successfully implemented a deep learning–based object detection and pose estimation pipeline for golf shot analysis using YOLOv8. The custom dataset was manually annotated, and the model was trained and evaluated using standard detection metrics. The results demonstrate effective golfer detection and stable keypoint tracking during swing movements. Despite dataset limitations, the project highlights the practical application of modern object detection and pose estimation techniques in sports analytics and performance evaluation.