Soccer Assignment Report

Introduction

The objective of this assignment was to consistently identify and map soccer players across two different video feeds—one from a broadcast camera and one from a tacticam—covering the same gameplay but from different angles. The challenge required building a pipeline for detection, tracking, and cross-camera player mapping using computer vision techniques.

Methodology

Detection & Tracking:
A YOLO-based model (provided) was used for player and ball detection in both videos. ByteTrack was employed for multi-object tracking, assigning unique IDs to each detected player within each video.

Frame Alignment:
The broadcast video is 5 seconds long, while the tacticam video is 8 seconds. By visual inspection, it was determined that the last 5 seconds of the tacticam video correspond to the broadcast video. Only these overlapping frames were used for mapping.

Normalization:
Player positions were normalized by the frame size (width and height) to enable spatial comparison between the two camera views.

Mapping:
For each frame, every detected player in the broadcast view was matched to the closest detected player in the tacticam view using Euclidean distance of normalized center coordinates. A distance threshold was applied to ensure only confident matches were assigned.

Global Player IDs:
Each matched pair of broadcast and tacticam players was assigned a unique global player ID, ensuring consistent labeling across both views.

Visualization:
For selected frames, both views were visualized side-by-side with bounding boxes and global IDs. Players not matched across views were labeled as "Unmatched".

Techniques Tried

- Spatial-only matching with various distance thresholds (0.15–0.3)
- Careful frame alignment to ensure only overlapping gameplay was mapped
- Visualization with global IDs for matched players and "Unmatched" for others

---

Challenges

- Different Camera Perspectives: The tacticam covers the whole field, while the broadcast view is often zoomed in. This made spatial matching challenging, especially for players far from the goal.
- Tracker Instability: Tracking IDs can change or reset if a player is occluded or missed by the detector.
- Partial Visibility: Some players are only visible in one camera view at a time, making consistent mapping impossible for those instances.

What Remains / Future Work

- Appearance Features: Incorporate color histograms or jersey number OCR to improve matching in crowded or ambiguous situations.
- Field Calibration: Calibrate both cameras to a common field coordinate system for more accurate spatial matching.
- Tracker Improvements: Explore more robust multi-object tracking algorithms to reduce ID switches.

Conclusion

The developed pipeline successfully detects, tracks, and maps players across two camera feeds during overlapping gameplay. The approach is robust to common real-world challenges, and the results are clearly visualized with consistent global IDs for matched players. With further enhancements such as appearance-based matching and field calibration, the system could achieve even higher accuracy and reliability.