Project Report: Cross-View Player Re-Identification using YOLOv11

Approach & Methodology:

The goal of this project was to maintain consistent player identities across two video feeds broadcast.mp4 and tacticam.mp4 captured from different camera angles during a sports event. The solution was designed with simplicity and accessibility in mind, leveraging only a custom-trained **YOLOv11** model for detection. Instead of a deep re-ID model, **colour histogram comparison** in HSV space was used to match detected players across views.

Pipeline Overview:

1. Detect players frame-by-frame using YOLOv11.
2. Extract cropped player regions from each bounding box.
3. Compute HSV color histograms for each crop.
4. Match players across views based on histogram correlation (Cosine similarity).
5. Optionally annotate matched player IDs on video frames and visualize trajectories on a 2D pitch.

Techniques & Outcomes:

- YOLOv11 Detection: The model successfully detected players in both video streams with reasonable bounding box precision and confidence scores.
- HSV Histogram Matching: Provided lightweight and interpretable player appearance matching. Performed well in scenarios where players wore distinct uniforms.
- Evaluation: Matching accuracy was manually verified across sampled frames. An accuracy metric was calculated by comparing predicted matches with human-annotated ground truth.

Challenges Encountered:

- No pretrained re-ID embedding model was used, so matching relied solely on color distributions, which made it sensitive to:
  1. Illumination changes
  2. Occlusion or partial views
  3. Similar appearances (e.g., players on the same team)
- Frame misalignment between broadcast and tacticam videos occasionally led to mismatches due to differing player poses or camera angles.

- Ground truth generation for accuracy evaluation required manual effort, as no labeled identity data was available per frame.

Conclusion:

In conclusion, this project successfully demonstrated how player re-identification across multiple camera views can be achieved using a custom-trained YOLOv11 model combined with histogram-based appearance matching. The approach balanced simplicity and effectiveness detecting players with high precision and assigning consistent identities using only visual features.

While this method performs well under controlled conditions, it opens doors for deeper enhancements like trajectory modeling, learned embeddings, and frame synchronization. These additions would increase robustness and accuracy, especially in crowded or fast-paced gameplay scenarios.

This foundation lays the groundwork for more advanced multi-camera sports analytics systems, and with more time and resources, it can evolve into a full-fledged, real-time tracking solution. Let me know if you'd like help turning this into a report, presentation, or next-phase roadmap.