Player Re-Identification in Sports Footage

Objective

The goal of this project is to implement a player re-identification system that can assign consistent IDs to players across frames of a sports video, even if they momentarily go out of view. The system should track each player using a YOLOv11 detection model and re-identify them using appearance-based features.

Approach & Methodology

- **Player Detection**: A fine-tuned YOLOv11 model (best.pt) is used to detect only players (class ID = 0).

- **Feature Extraction**: After detecting players, their cropped bounding boxes are processed to extract **HSV color histograms**, which serve as lightweight appearance features.

- **ID Assignment**: Re-identification is done by computing the **cosine distance** between feature vectors from the current and previous frames. The **Hungarian algorithm** is used for optimal matching.

- **ID Consistency**: If a feature match is below a defined threshold, the same ID is retained; otherwise, a new ID is assigned.

Techniques Tried

- **YOLOv11 (Ultralytics)**: Used for efficient and accurate player detection.

- **HSV Histogram Matching**: Color-based feature extraction that works well in short-term and when player jerseys differ.

- **Hungarian Algorithm**: Ensures optimal one-to-one matching between players across frames.

Challenges Faced

- **Similar Uniforms**: Players with identical color schemes (e.g., same team) can confuse the histogram matching.
- **Occlusion and Crowd Density**: When players overlap, the detection and matching may falter.
- **Short Video**: Limited testing on only 15 seconds of footage restricted evaluation of long-term re-ID accuracy.

Output

- Annotated video frames saved in tracking_output/

- Logs saved in tracking_logs/player_tracking_log.csv, including:

  - Frame index

  - Bounding box coordinates

  - Player ID

  - Timestamp

Future Improvements (If Given More Time)

- Integrate **Deep Re-ID Models** (e.g., OSNet, FastReID) to improve appearance matching.

- Add **temporal smoothing** or **Kalman filtering** to handle occlusion and jitter.

- Extend the system for **multi-camera** setups with homography and cross-view ID mapping.