Player Tracking Using YOLO and OCR-Based Jersey Recognition

1. Approach and Methodology

The goal of this project was to develop a system that can track individual players in sports videos by assigning consistent IDs across frames. Our approach integrates three key components:

- **Object Detection using YOLOv11**: We used a pre-trained or custom YOLOv11 model to detect players in each frame of the input video. This model produces bounding boxes and confidence scores for detected players.
- **Feature Extraction**: For each detected player, we extracted two sets of features:
  - **Color Histogram** (HSV-based) to capture clothing and appearance.
  - **Deep Visual Features** using a custom CNN backbone to encode shape, posture, and texture.
- **Jersey Number Recognition**: We employed EasyOCR to read numeric jersey numbers from detected players. This acts as an auxiliary feature to improve tracking accuracy.

These features are combined to match newly detected players to previously tracked individuals by measuring cosine similarity and comparing positional proximity. If a match is not found, a new ID is assigned.

2. Techniques Tried and Their Outcomes

☑ Feature-Based Tracking (Current Approach)

We used color histograms and deep visual embeddings for matching detections across frames.

Deep visual embeddings consist of the method of CNN layers to extract players features using zero shot method.

This worked surprisingly well in cases where:

- Players were moving slowly or moderately fast.
- Lighting and contrast conditions were consistent.
- The camera had minimal zoom or pan movement.

✅ Jersey Number OCR

OCR provided an excellent bonus signal. When EasyOCR correctly identified the jersey number, it greatly improved player ID consistency across frames. We added a small score boost in the matching function if two players shared the same jersey number.

⚠️ Matching Thresholds

We empirically tested various similarity thresholds. We found:

- A cosine similarity threshold of **0.65** worked best.
- Too high a threshold caused frequent ID switching.
- Too low a threshold caused unrelated players to get the same ID.

3. Challenges Encountered

🎯 OCR Accuracy

One of the biggest challenges was jersey number recognition:

- OCR struggled with blurry frames, occlusions, and players turning their backs.
- Some numbers were either missed or misread (e.g., "11" read as "H" or "1I").
- The system doesn’t use the OCR confidence score much beyond accepting the best candidate.

📦 Bounding Box Noise

- YOLO sometimes produced false positives or incomplete boxes for players partially out of frame.
- Very small players in distant parts of the frame were either not detected or were ignored due to size constraints.

📍 Player Position Drift

- Since we don't use a motion model, tracking relies on visual features and proximity alone.
- In fast-motion scenes (like counterattacks or basketball sprints), players moved so quickly that the visual similarity alone wasn’t enough for reliable tracking.

👥 Similar Appearance

- Players on the same team wearing identical uniforms caused confusion, especially in team sports where many players are clustered together.
- This sometimes led to frequent ID switching when two similar-looking players crossed paths.