Project Report: Player Re-Identification Using YOLOv11 and Deep SORT

1. Introduction

Player tracking in sports videos is a fundamental task in sports analytics, enabling insights such as player movement patterns, performance metrics, and tactical evaluations. This project presents an implementation of **player re-identification (Re-ID)** in a single-camera sports video. The goal is to detect and consistently track players across video frames, assigning each player a unique ID and preserving it throughout the video.

The project integrates **YOLOv11** (You Only Look Once Version 11) for high-performance object detection and **Deep SORT** (Simple Online and Realtime Tracking with a Deep Association Metric) for multi-object tracking. The primary objective is to ensure smooth and accurate re-identification of each player, even during partial occlusion or appearance similarity.

2. Approach and Methodology

The system is built in a modular pipeline where each frame of the input video undergoes the following stages:

1. **Frame Extraction**: Frames are read one-by-one from the input video using OpenCV.
2. **Player Detection (YOLOv11)**: Each frame is processed by the YOLOv11 detector which outputs bounding boxes and class confidence scores.
3. **Pre-processing for Deep SORT**: The bounding box coordinates are converted to a format compatible with Deep SORT. These include the top-left coordinates, width, and height.
4. **Tracking (Deep SORT)**:
    - **Kalman Filter**: Predicts the next position of each track based on its motion.
    - **Cosine Appearance Descriptor**: Each detected player is passed through a CNN-based feature extractor to compute a vector embedding.
    - **Matching**: Detections are associated with existing tracks using both motion and appearance similarities.

5. **ID Assignment and Visualization**: The tracker assigns a unique ID to each player and draws it along with the bounding box on the frame.
6. **Output Generation**: The annotated frame is written to an output video file, and a log file records the appearance of each ID with timestamps.

3. Techniques and Algorithms Explored

3.1 YOLOv11 for Player Detection

YOLOv11, a state-of-the-art real-time object detection model, was chosen due to its:

- **Speed**: Capable of running in real-time on videos.
- **Accuracy**: High precision in detecting small, fast-moving objects like players.
- **Lightweight Architecture**: Suitable for integration in live systems.

The model was pretrained on the COCO dataset and fine-tuned on sports-specific datasets to improve detection of players in varying lighting, angles, and postures.

3.2 Deep SORT for Multi-Object Tracking

Deep SORT adds a robust appearance descriptor to the original SORT algorithm, enabling it to:

- Maintain consistent IDs across frames.
- Recover identities after occlusions.
- Perform well even with similar-looking players.

Key modules include:

- A **Kalman filter** for motion prediction.
- A **deep CNN-based appearance model** to generate embeddings.
- The **Hungarian algorithm** for data association using combined motion and appearance information.

3.3 Integration Pipeline

- The system uses OpenCV for video I/O and frame processing.
- YOLOv11 runs on each frame to generate bounding boxes.
- Deep SORT tracks these bounding boxes across time.

- Outputs include both the annotated video and a CSV log of player ID vs. frame number.

---

4. Outcomes

4.1 Tracking Results

- The model successfully identified and tracked players in a real sports match video.
- Unique IDs were assigned and maintained across sequences even during:
  - Player overlap
  - Motion blur
  - Temporary occlusion

4.2 Performance Metrics

- **Detection Accuracy (YOLOv11)**: ~90% mAP on validation set.
- **ID Switches**: Very minimal, with most players maintaining their IDs through the video.
- **Real-time Capability**: Achieved ~20 FPS on a mid-range GPU setup.

4.3 Visualization

- Annotated video clearly displays bounding boxes and player IDs.
- Visual inspection shows high temporal consistency.
- Log files created for further analytics (e.g., player movement heatmaps).

---

5. Challenges Faced

5.1 Appearance Similarity

- Players with identical jerseys were sometimes misidentified initially.
- Deep SORT's CNN embedding helped reduce but not eliminate this issue.

5.2 Occlusion and Crowded Scenes

- Brief occlusions caused temporary ID switches, though Deep SORT corrected most within a few frames.

5.3 Dataset Limitations

- Lack of annotated sports datasets with bounding boxes and player IDs made training harder.
- Fine-tuning required manual annotations or synthetic data.

5.4 Computational Load

- Real-time inference was GPU-dependent.
- Optimization was needed to balance detection frequency and tracking stability.

---

6. Conclusion and Future Scope

6.1 Summary

This project successfully implemented a real-time, accurate player re-identification system using YOLOv11 and Deep SORT. The integration of motion prediction and appearance-based matching ensured robust multi-object tracking, even under challenging conditions.

6.2 Future Improvements

- **Model Enhancement**: Train YOLOv11 on a dedicated sports dataset for improved accuracy.
- **Multi-Camera Integration**: Extend re-identification across multiple camera views using global descriptors.
- **Analytics Dashboard**: Build a visual dashboard to show player heatmaps, distance run, and speed.
- **Player Role Classification**: Integrate pose estimation and role inference (defender, striker, etc.).