Player Re-Identification Assignment Report

1. Approach and Methodology

The system performs real-time player re-identification from video using a combination of object detection (YOLOv11) and deep feature extraction (ResNet50). YOLO detects players in each frame, while ResNet50 extracts appearance features. These features are compared across frames using cosine similarity and IoU to assign consistent IDs. Bounding boxes and IDs are visualized and written to an output video.

2. Techniques Tried and Their Outcomes

- Used YOLOv11 for object detection with confidence filtering to ensure only high-quality detections.
- Leveraged ResNet50 (pre-trained on ImageNet) with final FC layer removed for dense embeddings.
- Applied cosine similarity and IoU combined (weighted scoring) to match identities between frames.
- Used a simple ID tracking mechanism that updates based on similarity thresholds.
- Resulted in consistent re-identification across frames in short clips with stable detections.

3. Challenges Encountered

- Lack of labeled data for evaluating true accuracy made it difficult to benchmark quantitatively.
- Appearance variation and partial occlusions occasionally affected feature quality.
- Limited GPU memory required optimization of frame processing pipeline.
- False positives in YOLO occasionally led to ID confusion in crowded frames.

4. Future Improvements

- Integrate motion tracking (e.g., Kalman filter or SORT) to improve robustness.
- Add temporal consistency checks to refine ID assignment.
- Train a custom feature extractor for Re-ID tasks using domain-specific datasets.
- Explore vision-language models (e.g., CLIP) to improve feature embedding quality.