Analysis Report: Re-identification in a Single Feed

Objective:

This project's primary goal is to identify, track, and re-identify players in a single static video feed—more especially, in a sports setting. The objective is to extract useful statistics like movement trajectories, speed, presence heatmaps, and possible ball passes in addition to preserving consistent player identities across frames. Sports analytics, tactical assessment, and player performance measurement from monocular video sources all depend on this kind of analysis.

Methodology:

A specially trained YOLOv8 model, which recognizes players and the ball in each frame, is used for object detection at the start of the pipeline. To preserve consistent object identities, these detections are fed into the Deep SORT tracking algorithm. A ResNet18 model pretrained on ImageNet is used to generate appearance-based embeddings, which further guarantee robustness against ID switches and occlusions. To compare these embeddings and, if necessary, re-identify players, cosine similarity is employed. Additionally, the code incorporates utility functions for proximity-based ball pass detection, speed estimation, and visualization outputs like trajectory plots, heatmaps, and annotated videos.

Techniques Tried:

Among the important strategies used are:

1. For real-time object detection, use YOLOv8.
2. Deep SORT for tracking multiple objects.
3. Visual embeddings for appearance-based re-identification based on ResNet18.
4. Cosine similarity thresholding for player identity matching.
5. displacement computations from frame to frame for estimating speed.
6. Heuristics based on ball proximity for pass detection. For heatmap and trajectory

visualization, use Seaborn and OpenCV.
To create a reliable single-camera re-identification and analysis system, each component was integrated one after the other.

Results:

In a 15-second test video, the system was able to identify and follow several players. Consistent ID assignment, cropped player image extraction, spatial presence heatmap generation, speed estimation in pixels per second, and ball pass detection based on distance thresholds were all capabilities it possessed. All pertinent overlays, including IDs, speeds, and passes, were visible in the output video (output.mp4). According to the analysis, this technique is a promising tool for sports video analytics since it can accurately capture player movement and spatial density.

Challenges:

During development, the following difficulties were faced:
1. ID switches happened when players were moving quickly or overlapping in crowded scenes.
2. Sometimes the object was too small or occluded to detect the ball.
3. The lack of sports-specific data during training may have limited the embedding accuracy of the ResNet18 model.
4. The fixed distance thresholds used for pass detection may not apply to other camera configurations or sports.
These difficulties highlight areas that require more flexible reasoning or better models for wider applicability.

Future Work:

Future enhancements to the system might consist of:
1. Utilizing a labeled sports dataset to refine the appearance embedding model for improved re-identification.
2. To avoid sudden ID switches, use temporal smoothing methods or sequence models

to mitigate this issue.

3. Using a trained event detection classifier in place of heuristic-based pass detection.

4. Incorporating multi-view video support for sophisticated analytics and 3D positioning.

5. Incorporating databases or interactive dashboards to compare and manage multi-match analytics.

For more thorough and sophisticated video-based sports analysis solutions, this system provides a strong basis.