

## Assignment: Computer Vision Engineer

### Objective:

We are building an exciting team and are looking for an ingenious Computer Vision engineer to join us. The selected candidate will be responsible for analyzing two videos where the user is performing a specific football drill and dynamically tracking the objects of interest (footballs) across frames. As part of the same, here's a small assignment to understand your multiple-object tracking skills better. You are free to use any open-source computer vision models, and if required train the same for better accuracy. Feel free to connect with the team for any other clarification.

### Model Suggestions:

- 1) YOLO - for Object Detection
- 2) Mediapipe - for player's pose estimation (Optional - emphasis is on your object tracking skills)

### Video Analysis Requirements:

Using the provided [drive link](#), dynamically display the following data overlaid on the two test videos. Use the reference sample\_output video to get a better understanding of what is expected in terms of tracking objects of interest.

### Deliverables:

#### 1. Dynamic Display:

- Dynamic and overlaid bounding boxes on stationary balls and a tracking point on the action ball across all frames for both videos (test1 and test2)
- Action Ball: Ball being used by player to perform the drill; Stationary Ball: Balls being used as regional markers by the player

#### 2. Code:

- The code written by you in order to achieve the required functionality in the form of a .ipynb or a .py file.

#### 3. Report:

- A short and concise report covering your methodology and the models used by you.
- Emphasis on the object tracking techniques used by you and how you managed to maintain ID consistency for objects of interest across frames.

### Evaluation Criteria:

- Accuracy of your tracker in tracking objects of interest.
- ID consistency and uniqueness of your tracker in occlusion/overlap cases across frames. (Refer to sample\_output for clarity and avoid ID swaps in case of overlapping objects)
- Efficiency of the code written by you.
- Quality and clarity of your overlay annotated videos.

## FAQs:

### 1. What is the difference between object detection and object tracking?

Object detection is the process of identifying and localizing objects of interest in individual frames of a video. It typically outputs bounding boxes and class labels per frame. Models like YOLO are widely used for this purpose.

Object tracking on the other hand involves associating detections across consecutive frames to maintain a consistent identity (ID) for each object over time. Tracking ensures that we know which object in frame  $t$  corresponds to which object in frame  $t+1$ , even during occlusion, overlap, or appearance changes.

In this assignment:

Detection helps locate all footballs in a frame.

Tracking helps assign unique and persistent IDs to these balls across all frames.

### 2. What is the difference between stationary balls and action balls?

The stationary balls remain fixed in position across the video, acting as reference points for the drill. The action ball is the one actively used by the player and moves significantly across frames.

### 3. What should the overlay video outputs look like?

Refer to the provided sample\_output video. Your output video must:

- Show **bounding boxes** on **stationary balls**, with consistent IDs.
- Show a **clearly marked tracking point or trail** on the **action ball**.
- Ensure overlays are visible and non-obstructive.
- Maintain a smooth visual experience with no frame flickers or ID jumps.