1. Load the dataset and preprocess it.
2. Write the code for the yolov5 model and train the dataset for 100 epochs.
3. Initial test the weights on the input soccer match video.
4. Write the placeholder to read the video frames and write the video frames for later use.
5. Implement tracking: For every frame the players have the bounding boxes and for next frame the player position changes and so the position of bounding boxes also changes.

We are going to use **ByteTrack**: a multi-object tracking algorithm, uses these bounding boxes, along with their associated confidence scores, to track the movements and identities of objects across multiple frames.

1. We created the tracker using supervision library, we convert the goalkeeper to normal player, just for consistency.
2. We will save the tracks into a stub (pkl file).
3. Once we have all the tracker in place we can create the annotation, annotation means we will have an ellipse for players and referees, for ball we have triangle.