Technical Report: Cricket Player Detection, Tracking, and Bird's-Eye Projection

1. Detection Model Architecture

The system utilizes the **YOLOv11 Nano (n)** architecture for player detection.

- **Selection Rationale:** The Nano variant was selected due to its optimal balance between inference speed and accuracy. In a sports broadcasting context, maintaining high FPS (Frames Per Second) is critical.
- **Fine-Tuning:** To address specific challenges such as motion blur and varying player poses (e.g., a crouching batsman or a diving fielder), the model was fine-tuned on a custom-curated dataset. The training utilized **Mosaic Augmentation** and **Multi-Scale training** to ensure the model remains robust when players appear at different distances from the camera.

2. Multi-Object Tracking (MOT) Algorithm

The **ByteTrack** algorithm was implemented to handle player tracking and ID assignment.

- **Mechanism:** Unlike standard trackers that only process high-confidence detection boxes, ByteTrack leverages a "tracklet" association method that includes low-confidence boxes.
- **Application in Cricket:** This is particularly effective during high-action sequences where players may be partially occluded by other players or blurred during rapid movement, ensuring that a player does not lose their unique ID when detection confidence temporarily dips.

3. ID Consistency Maintenance

Consistency is maintained across frames through a combination of spatial and temporal logic:

- **Kalman Filtering:** The system predicts a player's location in the next frame based on their current velocity. If a detection is missed, the tracker "searches" in the predicted area.
- **Patience Threshold:** A "missing buffer" of 30 frames is implemented. This allows the system to re-identify a player with the same ID even if they temporarily leave the frame or are fully occluded behind another person for up to one second.

4. Challenges Encountered & Solutions

- **False Positives (Shadows & Pitch Lines):** Initial tests showed the model occasionally misidentified pitch markings or dark shadows as players.
    - **Solution:** A **Data Cleaning Filter** was applied post-inference to check the aspect ratio of detections. Objects with a height-to-width ratio below 0.5 (representing flat ground objects) were automatically discarded.
- **Perspective Distortion:** Projecting coordinates from a tilted camera angle to a flat map is mathematically complex.
    - **Solution:** A **Homography-based coordinate transformation** was calibrated using 4-point correspondence. This allows the 3D world coordinates to be mapped accurately onto a 2D top-view representation.

5. Limitations

- **Static Homography:** The current system uses a static transformation matrix. While highly accurate for fixed-camera shots, significant camera panning or zooming requires the matrix to be re-calibrated dynamically to maintain projection precision.
- **Identity Swaps:** In cases of extreme occlusion (three or more players in identical uniforms overlapping for several seconds), there remains a small risk of ID switching.

6. Future Improvements

- **Dynamic Homography:** Integrating feature-matching (ORB or SIFT) to update the Homography matrix in real-time as the camera moves.
- **Pose Estimation:** Integrating Keypoint Detection to analyze player technique (e.g., bowling action or batting stance) in addition to simple location tracking.
- **Automatic Role Labeling:** Using the player's proximity to the pitch and their movement patterns to automatically label them as "Batsman," "Bowler," or "Wicketkeeper."