A Proposed Vision-Based Football Analytics System for Automated Tactical Insights: Interim Report (Engineering Track)

Sarthak Gupta Sidhartha Garg Tanish Verma Asa Singh
Indraprastha Institute of Information Technology, Delhi
{sarthak22451, sidhartha22499, tanish22532, asa22113}@iiitd.ac.in

Abstract

This report presents interim progress on an automated, vision-based football analytics system (Engineering Track). Manual analysis is time-consuming and subjective. Our system aims to provide accessible, data-driven tactical insights using standard video footage. We employ YOLOv8 for object detection, which has been successfully fine-tuned and validated (mAP@50=0.912).

Current work is focused on integrating Object Detection and ByteTrack algorithm for multi-object tracking, as well as utilizing methods like K-means for reliable player-to-team assignment across frames. Future steps include implementing homography estimation for pitch mapping, developing the backend (FastAPI/PostgreSQL) and frontend interactive visualization (React/Fabric.js), full system integration, deployment, and refining analytical outputs. The goal is to empower coaches, analysts, broadcasters, and fans with objective positional data and tactical tools.

1. Problem Statement

Traditional football analysis relies on manual review, which is inefficient and subjective. This project aims to develop an automated system to process football match videos and output time-stamped positional data.

- **Input:** Standard football match videos.
- **Output:** Time-stamped positional data with Track ID, Frame, Pitch Coordinates, and Class Labels.
- **Target Users:** Coaches, analysts, broadcasters, scouts, and fans.
- **Interface:** Web application with interactive 2D maps and match footage synchronization.

2. User Interface & Progress

2.1. UI Design

The web application enables users to upload football footage for analysis through an LLM-powered pipeline that extracts performance metrics and provides insights. The system outputs a processed video with annotated insights, an interactive 2D pitch map (via Fabric.js) displaying tracked entities, and player-specific analytics, including heatmaps, movement patterns, and performance metrics.

The frontend is built with React.js and styled using Tailwind CSS, while the backend leverages FastAPI for processing. A conceptual UI mockup is planned (Fig. 2, Appendix).

2.2. Progress So Far

The development of core computer vision components is underway. The object detection module, based on a fine-tuned and validated YOLOv8 model, is complete (see Sections 4 and 5). ByteTrack has been selected for object tracking, with initial integration in progress. Player assignments done using K-Means clustering. The infrastructure stack, comprising Python, PyTorch, OpenCV, and NumPy, is in place.

Key pending tasks include refining object tracking, implementing homography estimation, developing backend APIs and database (FastAPI/PostgreSQL), integrating React with Fabric.js, and finalizing full system deployment. The development follows standard tools and best practices to ensure scalability and maintainability.

3. Related Work

We build on established CV techniques. **YOLOv8** [Ultralytics(2023)] provides fast, accurate object detection, fine-tuned for football. **ByteTrack** [Zhang et al.(2021)] offers robust MOT by associating high- and low-confidence detections, suitable for crowded sports scenes. Our work aligns with broader sports analysis research, e.g., **SoccerNet-v2** [Deliege et al.(2021)], providing context and baselines (YOLOv8 for detection, ByteTrack for tracking). The system described by the paper, [Jurca and Giosan(2022)], detects, tracks, and identifies players and staff, and maps their positions from broadcast images to their actual field positions. The methodology provided by [Skalski(2024)], leveraging YOLOv8 keypoint detection and homography, facilitates camera calibration in soccer footage through dataset labeling, keypoint detection model training, and homography application for perspective transformation.

4. Datasets & Evaluation Metrics

4.1. Dataset Utilized

The **"Football Players Detection"** dataset [Roboflow(2024)] from Roboflow Universe, licensed under **CC BY 4.0**, consists of **372 images** categorized into four classes: *Ball, Goalkeeper, Player, and Referee*. The validation split includes **49 images with 1,174 instances**, exhibiting a notable class imbalance—e.g., **973 Player instances versus only 45 Ball instances**—which may impact minority-class performance (Sec. 5). Preprocessing steps include **auto-orientation and resizing to 640×640**, specifically optimized for **YOLOv8 fine-tuning**. Additionally, variations in image quality and angles could affect detection accuracy.

Further, we will use the "Football Field Detection" dataset from Roboflow Universe, licensed under **CC BY 4.0**, comprises **317 images** annotated for keypoint detection of football fields. The dataset is partitioned into **training (255 images), validation (34 images), and test (28 images) sets.** Preprocessing steps include auto-orientation and resizing to **640×640 pixels**. No data augmentations were applied. The dataset is optimized for training keypoint detection models, such as YOLOv8

4.2. Evaluation Metrics

Performance is assessed using well-established metrics:

- **Object Detection (YOLOv8):** Evaluated using **Precision**, **Recall**, **mAP@50**, and **mAP@50:95** on the validation set. The **Precision-Recall (PR) curve** (Fig. 3, appendix) further visualizes performance across different confidence thresholds.
- **Multi-Object Tracking (ByteTrack):** Measured with:
  - **MOTA (Multiple Object Tracking Accuracy):** Quantifies overall tracking precision by penalizing false positives, false negatives, and ID switches.
  - **IDF1 (ID F-score):** Evaluates the consistency of object identity assignment throughout the tracking sequence.
  - **HOTA (Higher Order Tracking Accuracy):** Balances detection accuracy and association quality, providing a more comprehensive tracking performance metric.
- **System Performance:** Assessed in terms of **frames per second (FPS), inference latency, and computational resource utilization**. The objective is to maintain >**10–15 FPS** to ensure real-time processing efficiency across the full pipeline.

5. Analysis of Results

Focuses on the fine-tuned YOLOv8 detector's validation performance.

Detection Performance: Achieved strong overall detection: **mAP@50 = 0.912**, **mAP@50:95 = 0.681** (See validation summary and PR curve, Fig. 3, appendix).

Class-wise mAP@50 (Fig. 3, appendix) and the confusion matrix (Fig. 4, appendix) show high accuracy for Player (0.994), Goalkeeper (0.950), Referee (0.984). The 'Ball' class is weaker (mAP@50=0.718). The confusion matrix confirms good recall for major classes but lower recall (0.67) for the ball, often missed (predicted as background). This is likely due to size, motion blur, occlusion, and dataset imbalance. Qualitative results (Fig. 5, appendix) show visual examples.

Tracking Performance: ByteTrack integration ongoing. Qualitative tests show basic tracklets. Quantitative MOT evaluation pending. Goal: Robust tracking comparable to benchmarks.

Runtime: YOLOv8 inference speed is 41.4ms/image (24 FPS) on A6000 GPU (validation run), excluding pre/post-processing. Full pipeline target remains >10-15 FPS on deployment hardware.

6. Compute Requirements

- **Fine-tuning:** High-end GPU (>=24GB VRAM, e.g., A6000).
- **Inference/Deployment:** Dedicated GPU (>=8GB VRAM), multi-core CPU (>=4-8 cores), RAM (>=16GB).

Current resources (cloud/local) are adequate. Deployment hardware performance is key.

7. Individual Tasks

Primary ownership is distributed; collaboration is key.

Table 1. Individual Task Ownership (Primary Responsibility)

Member	Primary Task(s)
Sarthak Gupta	YOLOv8 Tuning, ByteTrack Integ
Sidhartha Garg	ByteTrack Integ, Player-Detector Eval
Tanish Verma	Dataset Mgmt and proc, Init 2D proj
Asa Singh	YOLOv8 Tuning, Init 2D proj
Team	*Integration, Testing, Docs, LLM (Stretch)*

8. Next Steps

Immediate work involves completing and integrating remaining modules:

1. **Homography Estimation:** Implement and test perspective transform. (Owner: Sidhartha Garg, Tanish Verma)
2. **Coordinate Mapping:** Develop pixel-to-pitch coordinate transformation. (Owner: Asa Singh, Sidhartha Garg)
3. **Backend Development:** Finalize DB schema; implement API endpoints. (Owners: Asa Singh, Sarthak Gupta)
4. **Frontend Development:** Build React UI; integrate Fabric.js map; implement playback controls. (Owner: Sarthak Gupta, Tanish Verma)
5. **Tracking Refinement:** Quantitatively evaluate MOT performance; tune/refine tracking. (Owner: Sidhartha Garg, Sarthak Gupta)
6. **System Integration & Testing:** Combine modules; test end-to-end pipeline. (Owner: Team)
7. **Deployment Strategy:** Plan and implement web app deployment. (Owner: Tanish Verma)
8. **(Stretch Goal) LLM Integration:** Explore feasibility of using multimodal LLM for analysis. (Owner: Team)

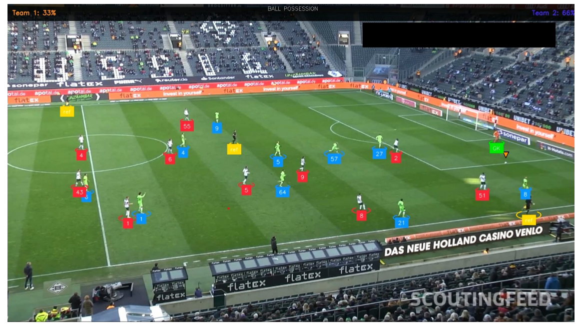

Figure 1. Fig 1: Conceptual illustration of detection (YOLOv8) and tracking (ByteTrack) pipeline identifying entities. Bounding boxes and Track IDs are overlaid.

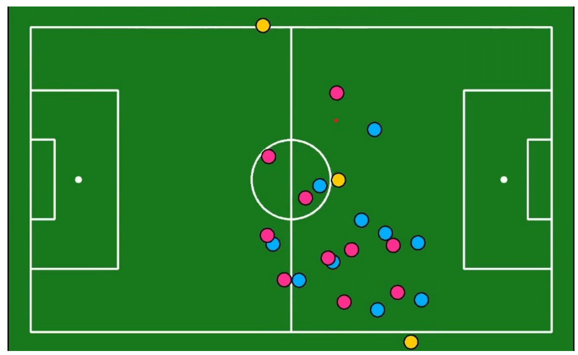

Figure 2. Fig 2: Design concept for the interactive 2D tactical map interface with tracked positions and playback controls.

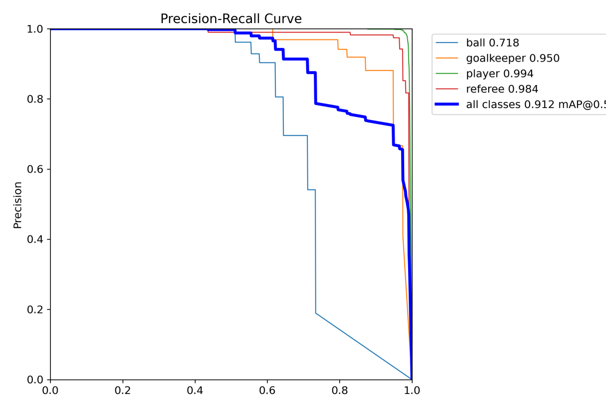

Figure 3. Fig 3: Precision-Recall Curve (YOLOv8 validation). Legend shows class mAP@0.5 and overall mAP@0.5 (0.912).

References

[Deliege et al.(2021)] Adrien Deliege, Anthony Cioppa, Silvio Giancola, Meisam J Seikavandi, Jacob V Duez, Congyan Zhou, Bernard Ghanem, and Marc Van Droogenbroeck. 2021. Soccernet-v2: A large-scale benchmark for football action spotting, tracking, and re-identification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 1290–1300. 1

[Jurca and Giosan(2022)] Mihnea Bogdan Jurca and Ion Giosan. 2022. A modern approach for positional football analysis using computer vision. In *2022 IEEE 18th International Conference on Intelligent Computer Communication and Processing (ICCP)*. 275–282. https://doi.org/10.1109/ICCP56966.2022.10053962 2

[Roboflow(2024)] Roboflow. 2024. Football Players Detection Dataset. https://universe.roboflow.com/roboflow-jvuqo/

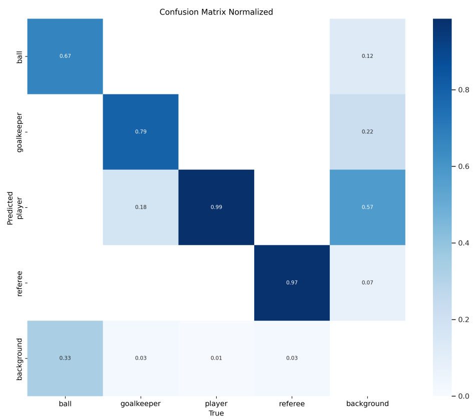

Figure 4. Fig 4: Normalized Confusion Matrix (YOLOv8 validation). Rows: Predicted, Columns: True. Diagonal shows recall.

Figure 5. Fig 5: Qualitative detection results (YOLOv8 validation) showing detected entities on sample images.

football-players-detection-3zvbc. 2

[Skalski(2024)] Piotr Skalski. 2024. Camera Calibration in Sports with Keypoints. *Roboflow Blog* (August 8 2024). https://blog.roboflow.com/camera-calibration-sports-computer-vision/ 2

[Ultralytics(2023)] Ultralytics. 2023. YOLOv8. https://github.com/ultralytics/ultralytics. 1

[Zhang et al.(2021)] Yifu Zhang, Peize Sun, Yi Jiang, Dongdong Yu, Zehuan Yuan, Ping Luo, Wenyu Liu, and Xinggang Wang. 2021. ByteTrack: Multi-Object Tracking by Associating Every Detection Box. *arXiv preprint arXiv:2110.06864* (2021). (To appear in ECCV 2022). 1