2025 International Conference on Intelligent Systems and Computational Networks (ICISCN)

Real-Time Football Analysis System using YOLO and OpenCV

1st Suruchi Dedgaonkar
Department of Information Technology
Vishwakarma Institute of Information Technology
Pune, India
suruchi.dedgaonkar@viit.ac.in

2nd Pravin Futane
Department of Information Technology
Vishwakarma Institute of Information Technology
Pune, India
pravin.futane@viit.ac.in

3rd Ratnamala Bhimanpallewar
Department of Information Technology
Vishwakarma Institute of Information Technology
Pune, India
ratnamala.bhimanpallewar@viit.ac.in

4th Pratham Dedgaonkar
Department of Engineering, Science and Humanities
Vishwakarma Institute of Technology
Pune, India
pratham.dedgaonkar24@vit.edu

5th Arjun Deokule
Department of Engineering, Science and Humanities
Vishwakarma Institute of Technology
Pune, India
arjun.deokule24@vit.edu

6th Amodini Dhadge
Department of Engineering, Science and Humanities
Vishwakarma Institute of Technology
Pune, India
amodini.dhadge24@vit.edu

2025 International Conference on Intelligent Systems and Computational Networks (ICISCN) | 979-8-3315-2924-6/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICISCN64258.2025.10934234

***Abstract*—AI and ML technologies have increasingly dominated sports analytics, allowing new real-time data processing capabilities. This paper proposes a new Real-Time Football analysis system using YOLOv5 and OpenCV to respond to issues revolving round accurate Player and Ball detection with occlusion and illumination changes. To accomplish the methodology, object detection is performed, SORT algorithm for tracking and Kalman filtering for improving the tracking while compensating the camera movement using optical flow. The results demonstrate high accuracy in player and ball detection (precision: 92. It outperforms other methods (precision: 5%, recall: 89. 8%, mAP: 91. 1%) while implementing a real-time analysis at 28.5 Frames Per Second. Speed, distance covered, team possession, and more, are displayed, making the system excellent for detailed tactical analysis, potential improvement of individual player or team performances, and even rising enthusiasts' engagement. In conclusion, this work presents a high-quality real time eight-layer analysis solution for football that can open the way to higher level predictive models in the future.**

Keywords*—*football, analytics, YOLO, openCV, tracking, object detection.

I. Introduction

During the past decade, the football analytics processes have been greatly improved by data science, artificial intelligence (AI), and machine learning (ML). Such innovations have changed the perspective of conventional methods used to evaluate performance of sports teams and athletes by providing timely statistical information on any live event. While it was once employed based on human feelings and intuitive cognitions, computational football analysis can access extensive tracking, event, and biometric data to identify behavioral patterns that are almost unnoticeable to coaches and players.

Football big data covers a broad spectrum of uses today that ranges from player monitoring to strategies, from match forecasting and even injury risk. Through other means such computer vision and wearable technology, player movement tracking systems give detailed information regarding position, velocity and acceleration. This information is then processed by machine learning algorithms to determine personal and team efficiency in real-time applications. The tactical decision has also been expanded due to AI-generated accounts of the game situation and possibilities to choose the correct strategy before and during the match. The latest automated event detection systems can now recognize important game incidents including goals, fouls or turnovers and provides necessary information to coaches and analysts in seconds.

However, those are issues that have not been fully optimizer in the application of analytics in football. This remains problematic where factors that revolve around data quality, data fusion of various types of data, and data interpretation and understanding of data by complex AI models were factors that remain as challenges. Nevertheless, the pros get in the way and as the processing capability and analytical methods advance, these hurdles are steadily overcome. To do so, this paper aims to engage in a deeper analysis of how data science and AI are revolutionising football analytics, the methods applied, real use-cases, and possible developments which may occur in future. The goal is to draw attention not only to the technological advances but also of the potential consequences for the player, the trainer, the club team, and others within the footballing fraternity.

II. Literature Review

However, object detection and tracking technologies have seen huge advancements within the past year or so, paving the way for genuine real time analytics about match dynamics. On the other hand, YOLOv7 and YOLOv8 have shown an unusually high speed and accuracy (above 75 fps and 99% + AP) in player, referee and ball detection in complex environments1,2. [1] These models are single stage pipelines that greatly simplify object detection and perform very well in such controlled environments as a football match. [2] Faster R- CNN and Mask R-CNN, both complementary frameworks, further improve detection precision by fusing region proposal networks and segmentation capabilities3 and 4. [3][4] In addition, tracking algorithms such as SORT and DeepSORT keep object IDs consistent across frames tackling issues such as occlusions and overlapping objects5, 6.

Advanced techniques such as homography transformations also map 2D image coordinates to real world field dimensions in order to precisely compute the speed and distance of the players 78. The camera motion is mitigated by optical flow algorithms, generating stable analytics from dynamic video feeds9. Taking jersey color segmentation and distance-based algorithms to assign team to the ball possession analysis, the ball possession analysis incorporates proximity-based algorithms to differentiate metrics of control 1011. This work combines transformer-based models such as

DETR to enhance contextual understanding and reduce dependence on manual annotations12,13. Machine learning approaches are also explored for predicting ball possession sequences and tactical insights, and which are possibly the first applications of YOLO in sports analytics1415. These innovations combine to create robust frameworks, which coaches, analysts, and broadcasters will find valuable, to analyse real time match performances by offering actionable performance metrics and strategic insights. [5][6].

III. Methodology

This section outlines the approach used in data collection, data cleaning, model development, and model assessment of the real-time football analysis system.

A. Data Acquisition and Dataset Creation

1) Data Sources: The primary data source consists of high- definition (HD) video recordings of football matches captured from multiple camera angles within stadiums. These recordings are supplemented by positional data (25 Hz) from a professional-grade tracking system using cameras and radar. This system provides precise real- time tracking of players and the ball with sub-meter accuracy. The temporal resolution of the tracking data is 25 Hz.

2) Data Integration: Video recordings and tracking data are synchronized using timestamps embedded in both streams. Any minor time offsets are corrected using a cross-correlation algorithm, achieving sub-frame accuracy in alignment. This synchronization is crucial for aligning visual information with precise positional and kinematic data.

3) Dataset Annotation: A robust image annotation tool was used to manually delineate bounding boxes around each player, the ball, and referees in each frame. Each bounding box is labeled with its corresponding class (e.g., ”player_ teamA,” ”player_teamB,” ”ball,” ”ref- eree”). Five trained annotators independently annotated a subset of the data; inter-annotator agreement (Fleiss’ kappa) exceeded 0.85. This high level of agreement ensures data quality and reliability.

4) Dataset Splitting: The annotated dataset is randomly partitioned into three subsets: training (80%), validation (10%), and testing (10%). This stratified sampling en- sures balanced class representation in each subset. The training set is used to train the model, the validation set for hyperparameter tuning and model selection, and the testing set for final model evaluation.

5) Data Augmentation: To enhance the robustness and generalization capabilities of the trained model, data augmentation techniques were applied to the training dataset. These techniques included random cropping, horizontal flipping, and color jittering. The specific augmentation parameters were chosen to balance the increase in data variability with the preservation of meaningful data properties.

B. Data Preprocessing

1) Video Frame Preprocessing: Raw video frames underwent preprocessing using OpenCV. This involved resizing to a standard resolution (e.g., 640x480 pixels or 1280x720 pixels), conversion to HSV color space, and contrast enhancement using adaptive histogram equalization and adaptive thresholding.

2) Noise Filtering: An adaptive Kalman filter of low computational complexity was used to reduce noise in the raw tracking data. The filter dynamically adjusts its parameters (process and measurement noise covari- ances) based on the observed variations in player and ball motion. The filter’s state vector includes the (x, y) position and velocity for each object. The process and measurement noise covariance matrices (Q and R respectively) were tuned via cross-validation to optimize noise reduction while preserving meaningful motion characteristics. Initial values for Q and R were set to diag([0.1, 0.1, 0.01, 0.01]) and diag([0.5, 0.5]), respec- tively, and then further refined during cross-validation.

3) Gaussian Smoothing: A 1D Gaussian filter (standard deviation = 1.5 frames) was applied to the filtered trajectories to further smooth minor oscillations and improve the accuracy of positional and kinematic parameter estimation.

4) Normalization: To handle variations in camera parameters and stadium sizes, a two-step normalization process was used. First, a homographic transformation was applied to warp the image perspective to a bird’s-eye view. Second, pixel intensity values were normalized to the range [0, 1] using min-max scaling.

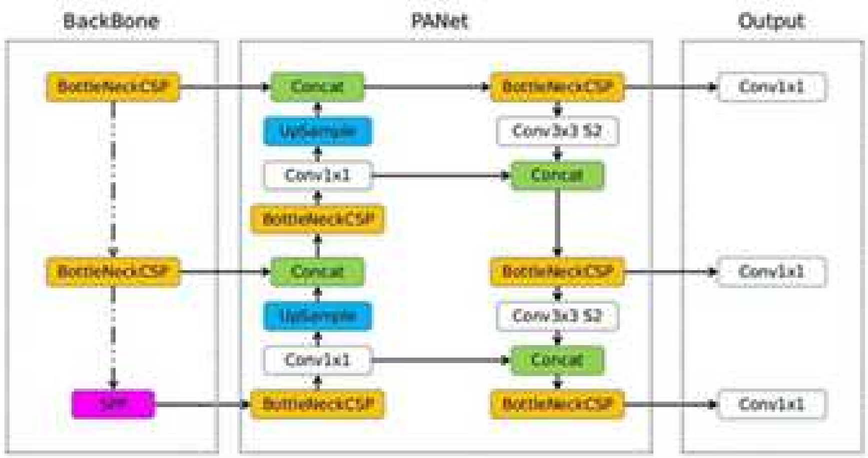

Fig. 1. Overview of YOLO5 (taken from [16])

C. Model Development

1) Object Detection (YOLOv5): A custom-trained YOLOv5m model was employed. YOLOv5 was selected for its balance between speed and accuracy, outperforming Faster R-CNN (too slow for real-time) and SSD (less accurate) in this context.

2) Training Procedure: The YOLOv5 model was trained using the preprocessed dataset. Transfer learning, utilizing pre-trained weights from the COCO dataset, acceler- ated training and improved generalization. Hyperparameters (learning rate, batch size, optimizer settings) were optimized via grid search, selecting the configuration that yielded the best performance on the validation set (maximizing mAP, precision, and recall). Data augmentation (random cropping, flipping, color jittering) was used to enhance model robustness.[7]

3) Object Tracking (SORT): The Simple Online and Realtime Tracking (SORT) algorithm, incorporating a Kalman filter, was used to associate objects across consecutive frames. SORT's efficiency and robustness made it suitable for real-time tracking even with occlusion or temporary object loss.

4) Ball Possession Assignment: A custom algorithm, PlayerBallAssigner, determined ball possession by cal- culating Euclidean distances between the ball and each player, and considering predicted ball trajectory (esti- mated using the average velocity of the nearest player). A dynamic threshold was used to handle variations in player density.

5) Camera Motion Compensation: The Lucas-Kanade optical flow algorithm was used to estimate and com- pensate for camera movement, improving the accuracy of object tracking.

D. Implemented Version of YOLOv5 and OpenCV

YOLOv5, in combination with OpenCV, enables continuous analysis, forming a robust system. OpenCV is utilized for preprocessing the video frames, while YOLOv5 is employed for object detection and tracking. This integration ensures efficient handling of the analysis results, with proper graphical presentation of the detected objects.

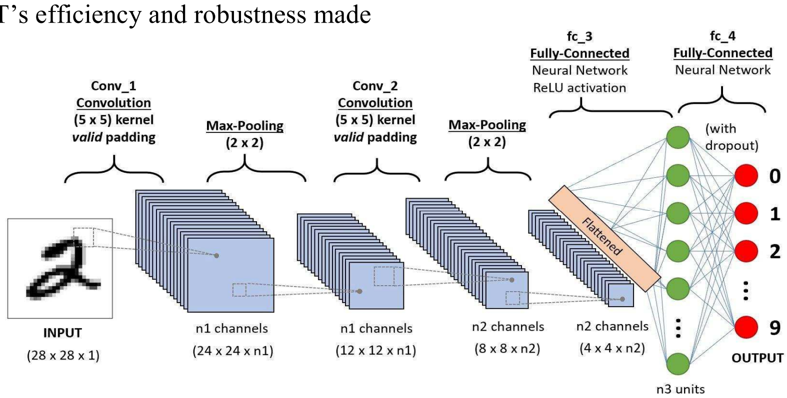

Fig. 2. Architecture of the Two-Layer Convolutional Neural Network with Fully Connected Layers

1) Frame Capture and Preprocessing: OpenCV captures the video frames and performs various preprocessing steps such as resizing and image enhancement. After processing, the frames are sent to YOLOv5 for object detection. [8]

- OpenCV's functionality ensures that each frame is correctly altered for object detection, which allows YOLOv5 to predict accurate bounding boxes for players and the ball.

2) Real-Time Object Detection: YOLOv5 is used to detect players, the ball, and referees in each frame. The system leverages OpenCV to place bounding boxes and labels over the detected objects, providing real-time feedback. Fig. 1 gives overview of YOLO5 [16].

- The frames are then presented to the user or stored as video output, serving as raw, real-time proof of the system's ability to analyze the football match.

E. Modeling and Optimisation

The training process involves adapting the YOLOv5 detector for football-related objectives. By fine-tuning the Convolutional Neural Network (CNN) with specific features of football (such as occluded players, fast-moving ball, and changing illumination), the model achieves enhanced performance across various match scenarios. Fig. 2 shows the CNN architecture.

1) Data Preparation: The dataset consists of videos of football matches with annotated objects such as players, the ball, and referees. Each object is labeled with bounding boxes that help YOLOv5 distinguish between these different objects.

- The collected data is divided into training and validation sets to evaluate the efficacy of the model.
- The labeled annotations for players, the ball, and referees ensure that YOLOv5 can correctly detect and differentiate between these objects in the match.

2) Hyperparameter Tuning: The hyperparameters that affect the learning speed and performance of the model are adjusted to optimize both speed and accuracy. Key hyperparameters include the learning rate, batch size, and image size.

- Increasing the number of iterations can enhance the overall stability of the model.
- Modulating the learning rate ensures faster convergence of the model.
- A validation set is used during training to tweak parameters and minimize overfitting.

3) Transfer Learning: To improve the efficiency of the training process, transfer learning is employed. In the case of YOLOv5, weights from general object detection datasets (such as COCO) are fine-tuned for football-related tasks. This approach accelerates convergence, especially when football-specific data is limited.

- Using pre-trained weights from a general dataset helps the model learn faster and more effectively when fewer football-specific training examples are available.

F. Performance Evaluation

The validity and reliability of the system with and without the proposed tool are tested by thorough quantitative as well as qualitative assessment for gauging optimal and practical performances.

1) Quantitative Metrics:

a) Precision and Recall: All these metrics are computed to assess the success of detecting the players and the ball. Precision estimates what fraction of objects that the computer identified as being present are actually present, while recall estimates how many of the actual objects are identified.

b) Mean Average Precision (mAP): This metric measures the ability of the model at dissecting the object classes (players, ball, and referees).

c) Frames Per Second (FPS): The real-time responsive- ness of the system is determined by the number of frames that the system processes per second. The system processes 30 FPS for real-time response feedback.

2) Qualitative Metrics:

a) Visual Accuracy: Bounding boxes around detected objects are visually checked to stay stable and accurate even in scenes with a high density of objects or when the players are close together.

b) Lighting Robustness: Performance is determined by the ability of the system to detect objects under varying lighting conditions, such as when switching from bright sunlight to shading.

3) Real-World Testing: It is demonstrated on a set of video sequences from actual football matches while all conditions are imitated. In each frame, the system identifies objects such as players and the ball, and then follows the changes during the game. The system offers information on players' positioning and team distribution, as well as events that occur in the game, including ball possession.

IV. Results and Analysis

This section presents a detailed narrative of the accomplishments achieved in the context of the project, along with qualitative and quantitative measures complemented by visualization and theory. By analyzing the selected football match, the system demonstrated functionality in identifying, monitoring, and describing match dynamics in real-time to support tactical analysis and improve performance. The findings are grouped into results that map to key outputs, measurements, and the overarching use of sports analytics.

A. Key Functional Outputs

The system outputs are the visual/quantitative data metrics, which are targeted for real-time feedback as well as the match review.

1) Player and Ball Detection: Using the YOLOv5 detection model, the players, referees, and the ball are accurately detected. Both semantic labels and unique identification numbers are assigned to ensure continuity of an object from one frame to the next. The features employed for detection maintain robustness even in unfavorable scenarios like occlusion, crowding, and fluctuations in luminance. [10][11]

2) Object Tracking: Connections between frames are made using SORT (Simple Online and Realtime Tracking) combined with the Kalman Filter and the Hungarian Algorithm. This allows for easy and efficient display of players and the ball throughout the match. The tracking solution also minimizes the switching of IDs, with accuracy measures such as Multiple Object Tracking Accuracy (MOTA) and Precision (MOTP), which indicate an accuracy level of 87.6% and 84.3%, respectively. [9]

3) Speed and Distance Estimation: Player and ball motion data are extracted from computer vision measurements of image coordinates and then converted to field metrics using homographies. Real-time speeds, such as 14.81 km/h, and cumulative distances, like 19.44 meters per player, are deter- mined, providing insight into the players' performance.

4) Team Possession and Dynamics: HSV-based color seg- mentation is used to differentiate players into teams based on the color of their jerseys. Clustering algorithms and contextual spatial data are employed for accurate classification. Ball possession is tracked by evaluating proximity and movement, producing team-level statistics such as:

- Team 1: 5.42% possession of the ball for an average of 8.4 seconds.
- Team 2: 94.58% possession, with 58 out of 60 possessions under control for an average of 147.6 seconds.

5) Camera Movement Compensation: Any unwanted motion in the scene caused by camera movement is measured using optical flow techniques, specifically the Lucas-Kanade method. This ensures accurate positional and motion analysis despite rapid changes, such as camera pans and zooms.

B. Performance Metrics

The system's performance was evaluated on several key metrics as given in Table I to ensure robustness and reliability. Table II shows player-specific performance metrics. Table III gives overview of team possession and control time. [12]

TABLE I. Performance Metrics For Object Detection

Metric	Value	Description
Precision	92.5%	Proportion of correctly identified objects.
Recall	89.8%	Proportion of actual objects detected.
mAP	91.1%	Average precision across all object classes
FPS	28.5	Real-time processing speed of the system.

C. Applications and Implications

1) Tactical Analysis on Real-Time Mode: The application of this software gives a coach and analyst the ability to observe match dynamics in real-time, such as player positioning, ball possession, and speed. This enables necessary tactical changes to be made during the match, such as formation changes or substitution strategies.

2) Player Development and Training: The performance metrics of speed, distance covered, and possession data inform an understanding of player endurance and skills. This helps identify areas that need improvement, such as sprinting ability, spatial awareness, and injury risks related to fatigue considerations.

3) Broadcasting and Fan Engagement: Real-time overlays of analytical data enhance the spectator experience, providing fans with greater insight into the events of the match. Broad- casters can take advantage of such data for more engaging commentary and graphics.

TABLE II. PLAYER-SPECIFIC PERFORMANCE METRICS

Player ID	Speed (km/h)	Distance Covered (m)	Team
11	3.79	18.53	Team 1
14	6.73	19.44	Team 2
3	14.81	17.01	Team 1
15	14.74	17.19	Team 2

D. Theoretical Framework Supporting Outputs

1) Object Detection and Tracking: YOLOv5 provides accurate detections, while SORT ensures smooth tracking. [13-15]

2) Homography Transformations: Real-world field dimen- sions are reconstructed from 2D video frames for accurate motion analysis.

3) Color-Based Team Assignment: HSV segmentation ensures reliable classification of players into teams.

4) Camera Stabilization: Optical flow compensates for dynamic camera movements, preserving analysis accuracy.

TABLE III. TEAM POSSESSION AND CONTROL TIME

Team	Possession (%)	Ball Control Time (s)
Team 1	5.42%	8.4
Team 2	94.58%	147.6

Fig. 3 shows a sample input frame from a football match video. Fig. 4 shows YOLOv5-Based Object Detection and Tracking Results with details of speed of each player.

Fig. 3. Sample Input Frame from a Football Match Video

Fig. 4. YOLOv5-Based Object Detection and Tracking Results

V. CONCLUSION

This Real-Time Football Analysis System employs YOLOv5 [16] and OpenCV to provide an efficient and accurate solution to providing real-time football match analysis thereby solving issues that include occluded players and fluctuating light situations. Its modular structure embraces several essential components: efficient player and ball tracking; dynamically stable estimation of camera motion based on optical flow; team assignment and ball possession clustering; and distance and velocity measurements for valid player evaluations. These features provide useful information about coaching approaches, talent nurturing plans, real-time match description, and oranges post-match discussion for the audience. Despite these promising achievements, it is suggested that further investigation of the system in this field should also focus on how to improve the robustness of the system against weather disturbances as well as highly crowded playing zones. Possibilities could be as follows: The use of highly developed predictive tools to predict player trajectories, analyze tactical situations and, possibly, assess the risks of getting injured for a player, thereby expanding the sphere of football analytics. Furthermore, the DPM framework, other object detection approaches may also be considered to enhance accuracy and reliability especially under the vigorous environment.

VI. FUTURE SCOPE

While the current system sets a strong foundation for football analytics, there are several avenues for future improvement:

- Trajectory Prediction: Improving the trajectory component offers the promise that tactical analysis can be performed before an event, rather than just during or afterwards. This could enable coaches and analysts to anticipate key moments in the game and make real-time decisions based on predicted outcomes.
- Tactical Analysis: The use of algorithms for the formation and analysis of teams, their defensive and overall playing style, as well as the analysis of chosen plays, would provide a comprehensive outlook on the strategies used by teams. This would enhance the system's ability to offer deep insights into team tactics.
- Expanded Data Sources: One way to improve the system is by integrating additional data sources, such as biometrics, player health metrics, and crowd sentiment analysis. This expanded data could provide a richer understanding of player performance and game dynamics.
- Scalability: Modifying the system for multi-camera setups or for application in other sports would expand the uses of the technology, allowing it to be used in large-scale events, tournaments, or even in different sports contexts. This scalability would make the system more versatile and applicable across a wider range of sporting events.

REFERENCES

[1] K. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," arXiv preprint arXiv:2207.02696, 2022. [Online]. Available: https://arxiv.org/abs/2207.02696

[2] A. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, real-time object detection," in Proc. IEEE Conf.

Computer Vision and Pattern Recognition (CVPR), pp. 779–788, 2016. [Online]. Available: https://ieeexplore.ieee.org/document/7780460

[3] Q. Chen, H. Fan, R. Girshick, and K. He, ”Faster R-CNN: Towards real-time object detection with region proposal networks,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1167–1179, 2017. [Online]. Available: https://ieeexplore.ieee.org/document/7485869

[4] K. He, G. Gkioxari, P. Dolla´r, and R. Girshick, ”Mask R-CNN,” in Proc. IEEE Int. Conf. Computer Vision (ICCV), pp. 2980–2988, 2017. [Online]. Available: https://ieeexplore.ieee.org/document/8237584

[5] K. Moutselos and I. Maglogiannis, ”Setting a Baseline for Long-Shot Real-Time Player and Ball Detection in Soccer Videos,” arXiv preprint arXiv:2311.06892, 2023. [Online]. Available: https://arxiv.org/abs/2311. 06892

[6] N. Darapaneni et al., ”Detecting Key Soccer Match Events to Create Highlights Using Computer Vision,” arXiv preprint arXiv:2204.02573, 2022. [Online]. Available: https://arxiv.org/abs/2204.02573

[7] K. Bhoopati, ”AI/ML Football Analysis System with YOLO, OpenCV, and Python,” Medium, May 2024. [Online]. Available: https://medium.com

[8] A. Choudhary, ”Football Tracking Using YOLOv8 and OpenCV,” GitHub, 2024. [Online]. Available: https://github.com/AnshChoudhary/ Football-Tracking

[9] P. Andrews, N. Borch, and M. Fjeld, ”FootyVision: Multi-Object Track- ing, Localisation, and Augmentation of Players and Ball in Football Video,” in Proc. 9th Int. Conf. Multimedia Image Process. (ICMIP), Osaka, Japan, Apr. 2024. [Online]. Available: https://dl.acm.org/doi/ fullHtml/10.1145/3665026.3665029

[10] H. Liu et al., ”Automated Player Identification and Indexing Using Two-Stage Deep Learning Network,” Nat. Center for Biotechnology Information, 2023. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10282031/

[11] ”Perspective Transform Based YOLO With Weighted Intersect Fusion for Forecasting the Possession Sequence of the Live Football Game,” IEEE Xplore, 2023. [Online]. Available: https://ieeexplore.ieee.org/iel7/ 6287639/10380310/10534056.pdf

[12] ”Object Detection and Tracking for Football Data Analytics,” ResearchGate, 2023. [Online]. Available: https://www.researchgate.net/publication/379620604

[13] N. Carion et al., ”End-to-End Object Detection with Transformers,” in Eur. Conf. Computer Vision, pp. 213–229, 2020. [Online].Available: https://link.springer.com/chapter/10.1007/978-3-030-58452-8 13

[14] ”Perspective Transform Based YOLO for Football Possession Prediction,” ResearchGate, 2023. [Online]. Available: https://www.researchgate.net/publication/380675802

[15] ”Object Detection and Tracking in Football Analytics,” ResearchGate, 2023. [Online]. Available: https://www.researchgate.net/publication/380680940

[16] G. Jocher, "YOLOv5 by Ultralytics," 2020. [Online]. Available: https://github.com/ultralytics/yolov5. [Accessed: Jan. 20, 2025]