Football Match Analysis system

Ankur Kumar
Department of Computer Applications
PES University, Bangalore, India
Supervised by Dr. S.visalakshi

***Abstract*—Football is a game that relies heavily on strategy, skill, and precise decision-making. However, manual analysis of football matches is time-consuming and often lacks accuracy, making it difficult for coaches and analysts to obtain quick and reliable insights. This project presents an automated football match analysis system using computer vision and deep learning techniques to address these challenges. The system uses the YOLO model to detect players, the ball, and the referee from match video footage. It tracks player movements across frames, identifies team possession using jersey colour analysis, and computes important performance metrics such as realtime average player speed, as well as distance covered. By automating tasks like detection, tracking, and performance evaluation, the system reduces the dependency on manual analysis and provides accurate, real-time match insights. This AI-powered approach helps coaches and analysts better understand player performance, team behaviour, and match dynamics, enabling improved tactical planning and decision-making.**

***Index Terms*—Football analytics, computer vision, deep learning, YOLOv5, OpenCV, AI in sports,player tracking, tactical insights, sports technology.**

I. Introduction

Football is not merely defined by goals and final scores; it is a dynamic sport driven by strategy, continuous movement, and rapid on-field decision-making. Coaches and performance analysts depend heavily on match analysis to understand player behavior, team coordination, and tactical effectiveness. Traditionally, such analysis has been conducted through manual review of match recordings, which is time-intensive, prone to human error, and incapable of delivering timely insights during live matches. As the pace and competitiveness of modern football continue to increase, conventional methods fall short in meeting the demand for accurate and real-time performance evaluation.

Recent advances in artificial intelligence and computer vision have opened new possibilities for automating football match analysis. AI-driven systems enable automatic detection and tracking of players and the ball, extraction of movement patterns, and computation of performance metrics that are difficult to capture through manual observation. These technologies make it possible to convert raw video data into structured and meaningful information that supports objective analysis and data-driven decision-making.

This work presents an automated football match analysis framework based on deep learning and computer vision techniques. The proposed system employs YOLO-based object detection and OpenCV-based tracking to identify and follow players, the ball, and referees throughout match footage. Team affiliation is determined using jersey color analysis, allowing estimation of ball possession and team dominance. Furthermore, the system computes essential performance indicators such as player speed, distance covered, and spatial movement patterns. Camera motion is compensated using motion estimation and perspective transformation to improve spatial accuracy and reliability of measurements. The extracted information is visualized through heatmaps, movement trajectories, and statistical summaries to provide interpretable insights for tactical assessment.

By automating the analysis pipeline, the proposed approach significantly reduces manual effort while improving accuracy and consistency in performance evaluation. The system is designed to support coaches in tactical planning, assist clubs in player development, and enable near real-time insights during matches. Overall, this research demonstrates how AI-powered automation can enhance the efficiency, objectivity, and effectiveness of modern football analytics, paving the way for more intelligent and data-driven sports performance analysis systems.

II. System Overview

The proposed Automated Football Match Analysis System follows a structured, end-to-end processing pipeline to transform raw match videos into meaningful performance analytics. As illustrated in the process flow diagram, the system begins by taking recorded football match footage as input, which is first converted into individual frames for frame-wise analysis. These frames are processed using a YOLO-based object detection model to identify key entities on the field, including players, the ball, and the referee.

To enable consistent analysis over time, the detected objects are tracked across consecutive frames using unique identity assignment, allowing the system to follow individual player movements throughout the match. Team affiliation is then determined using K-means clustering on jersey color features, which facilitates accurate team classification and supports ball possession estimation. Based on the spatial proximity between players and the ball, the system estimates ball possession for each team across different match phases.

Since broadcast football videos often involve continuous camera motion, the framework incorporates camera motion estimation using optical flow to compensate for background movement. This ensures that player trajectories are analyzed in a stabilized reference frame. Furthermore, pixel-level coordinates are transformed into real-world measurements through camera transformation techniques, enabling conversion from pixel distances to metric units. Using these calibrated trajectories, the system computes key performance indicators such as player speed and total distance covered.

Finally, the extracted analytics are presented through visual overlays, movement trajectories, heatmaps, and statistical summaries, forming the final output of the system. This systematic process flow enables automated, accurate, and interpretable football match analysis, supporting data-driven tactical evaluation and performance assessment for coaches and analysts.

III. Methodology

A. Data Input and Preprocessing

The system begins by loading the football match video using the OpenCV library. The input video is converted into individual frames to enable frame-wise processing. Each frame is resized to a fixed resolution to maintain uniform input dimensions for the detection model. Since OpenCV reads images in BGR format, frames are converted to RGB format to match the input requirements of the deep learning model. The preprocessed frames are normalized and prepared for inference, ensuring stable and consistent detection performance across varying video qualities.

B. Object Detection using YOLO

A YOLO-based deep learning model is employed to detect players, the ball, and the referee in each video frame. The model outputs bounding boxes with confidence scores. Redundant detections are removed using non-maximum suppression (NMS), defined as:

$$\mathrm{IoU}(B_i, B_j) = \frac{|B_i \cap B_j|}{|B_i \cup B_j|} \tag{1}$$

Bounding boxes with high overlap and lower confidence are suppressed. Low-confidence detections are filtered to reduce false positives.

C. Object Tracking using ByteTrack / Deep SORT

Multi-object tracking assigns a unique identity to each detected object across frames. Let the position of object k at time t be $\mathbf{p}_k^t = (x_k^t, y_k^t)$. The predicted position in the next frame is estimated using a motion model:

$$\hat{\mathbf{p}}_k^{t+1} = \mathbf{p}_k^t + \mathbf{v}_k^t \Delta t \tag{2}$$

where $\mathbf{v}_k^t$ denotes the estimated velocity. This prediction helps maintain identity continuity during temporary occlusions.

D. Player Feature Extraction

For each tracked player, spatial features are extracted from bounding boxes. The player position is represented by the bounding box center:

$$(x_c, y_c) = \left(x + \frac{w}{2},\ y + \frac{h}{2}\right) \tag{3}$$

where (x, y) denotes the top-left corner and (w, h) represent width and height. Jersey color features are extracted from the region of interest for team classification.

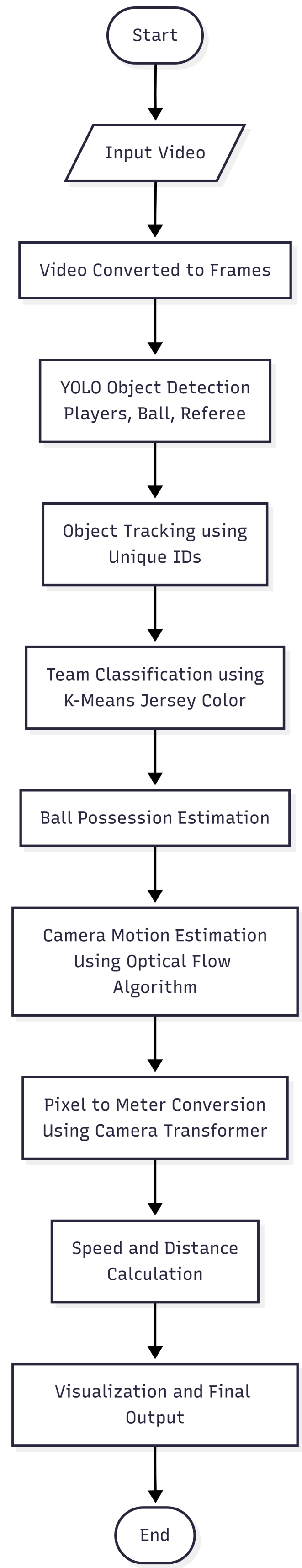

Fig. 1. Process Flow Diagram

E. Team Classification using K-Means Clustering

Jersey color features are clustered into two groups using K-means clustering:

$$\min_{\{\mu_1,\mu_2\}} \sum_{i=1}^{N} \min_{k\in\{1,2\}} \|\mathbf{c}_i - \mu_k\|^2 \tag{4}$$

where $\mathbf{c}_i$ represents the color feature vector of player i, and μ_k denotes the cluster centroids corresponding to the two teams.

F. Ball Tracking and Possession Estimation

Ball possession is estimated based on the distance between the ball and nearby players:

$$d_i^t = \|\mathbf{p}_{\text{ball}}^t - \mathbf{p}_i^t\| \tag{5}$$

The player with the minimum distance to the ball within a predefined threshold is assigned possession. Team possession percentage over time T is computed as:

$$\text{Possession}_{\text{team}} = \frac{1}{T}\sum_{t=1}^{T} \mathbb{I}(\text{team holds ball at } t) \tag{6}$$

G. Camera Motion Estimation

Camera motion is estimated using optical flow. Given a pixel location (x, y) and its displacement (u, v) between frames:

$$I_x u + I_y v + I_t = 0 \tag{7}$$

where I_x, I_y, and I_t denote image gradients. The Lucas–Kanade method is used to solve this equation for stable field features.

H. Camera Movement Compensation

Estimated camera displacement $(\Delta x, \Delta y)$ is subtracted from object positions:

$$\tilde{\mathbf{p}}_k^t = \mathbf{p}_k^t - (\Delta x^t, \Delta y^t) \tag{8}$$

This compensation stabilizes trajectories and isolates true player movement.

I. Perspective Transformation and Pixel-to-Meter Conversion

The mapping from image coordinates $\mathbf{x}$ to field coordinates $\mathbf{x}'$ is given by the homography matrix H:

$$\mathbf{x}' \sim H\mathbf{x} \tag{9}$$

Pixel distances are converted to metric units using a scaling factor s:

$$D_{\text{meters}} = s \cdot D_{\text{pixels}} \tag{10}$$

J. Speed and Distance Metric Computation

Player displacement between frames is computed as:

$$\Delta D_k^t = \|\mathbf{p}_k^t - \mathbf{p}_k^{t-1}\| \tag{11}$$

Speed is calculated as:

$$v_k^t = \frac{\Delta D_k^t}{\Delta t} \tag{12}$$

The total distance covered by a player over T frames is:

$$D_k = \sum_{t=1}^{T} \Delta D_k^t \tag{13}$$

K. Output Rendering and Visualization

The final results are visualized by overlaying player identities, trajectories, and heatmaps on the original video frames. Statistical summaries are exported in CSV and JSON formats to support further analysis and reporting. The annotated video provides an interpretable visual representation of match dynamics and player performance.

IV. Results and Discussion

The performance of the proposed automated football match analysis system is evaluated through qualitative visual results and quantitative observations obtained from processed football match videos. The system outputs intermediate visualizations at different stages of the pipeline, enabling verification of detection accuracy, tracking stability, team classification, and performance metric computation.

A. Object Detection Results (YOLO)

The YOLO-based detection module successfully identifies players, the ball, and referees across video frames. The detected bounding boxes are overlaid on the input frames, demonstrating the ability of the model to localize multiple objects in dynamic match scenes. Detection performance remains stable in well-lit and open-field conditions, while minor degradation is observed during occlusions and rapid camera motion.

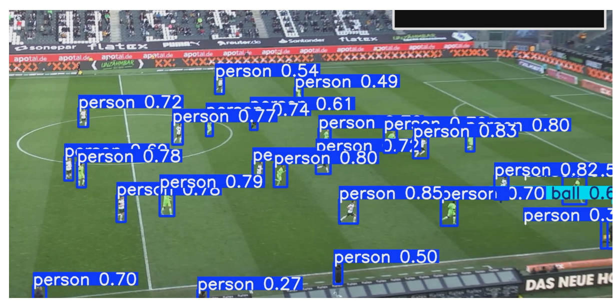

Fig. 2. YOLO-based detection of players, ball, and referee in a sample frame.

B. Team Classification using Jersey Color (K-Means)

The team classification module clusters player jersey colors into two dominant groups using K-means clustering. The extracted jersey regions enable effective differentiation between opposing teams, which is visually validated through color-coded player bounding boxes. The clustering approach performs reliably when team jersey colors are visually distinct, although minor misclassification may occur when color patterns are similar or lighting conditions vary significantly.

C. Ball Tracking and Possession Estimation

The ball tracking module maintains a consistent identity for the ball across frames and estimates possession based on proximity between players and the ball. The possession assignment is visualized by highlighting the player currently closest to the ball. Aggregated possession statistics reflect realistic match dynamics, providing an intuitive understanding of team dominance during different phases of play.

Fig. 3. Team classification using K-means clustering on extracted jersey color features.

D. Camera Motion Compensation and Trajectory Stabilization

Camera motion estimation using optical flow effectively separates background motion from actual player movement. After compensation, player trajectories appear stable relative to the field, allowing accurate spatial analysis. This stabilization improves the reliability of movement patterns, particularly in broadcast videos involving continuous camera panning and zooming.

E. Final Visualization and Performance Metrics

The final output of the system includes annotated videos with player identities, team colors, ball trajectories, and motion statistics. Heatmaps and trajectory overlays provide an interpretable summary of spatial behavior, while numerical metrics such as speed and distance covered offer quantitative insights into player workload and team activity. These outputs demonstrate the practical applicability of the proposed framework for tactical analysis and performance evaluation.

F. Discussion

The experimental results indicate that the proposed system can effectively transform raw match videos into meaningful analytical insights. The integration of detection, tracking, team classification, camera motion compensation, and metric computation enables comprehensive performance analysis. While the system performs well in standard broadcast conditions, limitations are observed under heavy occlusion, fast ball motion, and visually similar team jerseys. These challenges highlight potential directions for future improvements, such as incorporating appearance-based re-identification models and multi-view analysis to enhance robustness.

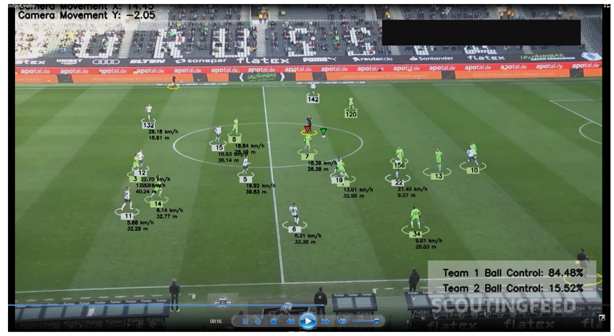

Fig. 4. Final visualization showing player trajectories, heatmaps, and analytical overlays.

V. Conclusion and Future Work

VI. Conclusion and Future Work

This work presents a practical and interpretable framework for automated football match analysis using computer vision and deep learning techniques. By integrating object detection, multi-object tracking, team classification, camera motion compensation, and performance metric computation into a unified pipeline, the proposed system demonstrates how raw football match videos can be transformed into meaningful analytical insights. The framework enables objective evaluation of player movement patterns, team dynamics, and ball possession trends, thereby supporting data-driven tactical analysis for coaches and performance analysts.

Despite achieving reliable performance under standard broadcast conditions, certain challenges such as heavy occlusion, rapid camera zoom, and visually similar team jerseys can affect detection and tracking accuracy. These limitations highlight opportunities for further improvement. Future work may focus on incorporating pose estimation techniques to capture fine-grained player actions such as passing, tackling, and shooting, enabling deeper behavioral analysis. In addition, integrating real-time processing pipelines and model optimization strategies can facilitate deployment in live match scenarios, providing near real-time tactical feedback. Furthermore, extending the framework with predictive analytics and learning-based tactical modeling can enable anticipation of play patterns and strategic recommendations. Such enhancements would further improve the robustness, scalability, and practical utility of automated football analytics systems in professional sports environments.

References

[1] J. Eichner, J. Nowak, B. Grzelak, T. Górecki, T. Piłka, and K. Dyczkowski, "Advanced Vision Techniques in Soccer Match Analysis: From Detection to Classification," in *Proc. VISIGRAPP/VISAPP Conf.*, 2025. [Online]. Available: https://www.scitepress.org/Papers/2025/133771/133771.pdf

[2] A. Temptation, W. Kusakunniran, and W. Limroongreungrat, "Deep Learning-Based Detection of Players and Teams in Soccer Videos with Positioning Heatmap Generation," *Applied Computing and Informatics*, 2025. [Online]. Available: https://doi.org/10.1108/ACI-07-2024-0257

[3] S. Akan and S. Varlı, "Use of Deep Learning in Soccer Videos Analysis: A Survey," *Multimedia Systems*, Springer, 2023. [Online]. Available: https://doi.org/10.1007/s00530-022-01027-0

[4] A. Karimi, R. Toosi, and M. A. Akhaee, "Soccer Event Detection Using Deep Learning," arXiv preprint, 2021. [Online]. Available: https://arxiv.org/abs/2102.04331

[5] O. A. N. Rongved *et al.*, "Automated Event Detection and Classification in Soccer: Potential of Using Multiple Modalities," *Multimodal Technologies and Interaction*, MDPI, vol. 3, no. 4, 2019. [Online]. Available: https://www.mdpi.com/2504-4990/3/4/51

[6] F. Vidal-Codina *et al.*, "Automatic Event Detection in Football Using Tracking Data," *Sports Engineering*, Springer, 2022. [Online]. Available: https://link.springer.com/article/10.1007/s12283-022-00381-6

[7] S. Rangappa, B. Li, and R. Qian, "Tracking and Identification for Football Video Analysis Using Deep Learning," in *Proc. SPIE Machine Vision Conf.*, 2021. [Online]. Available: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11605/2586798

[8] A. K. Valdani, "Deep Learning for Automated Football Goal Detection: Comparing 2D and 3D Residual Networks," M.S. thesis, KTH School of Electrical Engineering and Computer Science, 2024. [Online]. Available: https://www.diva-portal.org/smash/get/diva2:1912371/FULLTEXT01.pdf