Proceedings of the 6th International Conference on Data Intelligence and Cognitive Informatics (ICDICI-2025)
IEEE Xplore Part Number: CFP25VL6-ART; ISBN: 979-8-3315-0313-0

AI-Powered Football Match Analysis using YOLOv8 and Spatial Analytics

1st S Pranav Arun
Department of Computing Technologies,
School of Computing,
College of Engineering and Technology,
SRM Institute of Science and Technology,
Kattankulathur, Chennai, Tamil Nadu, India
ps8299@srmist.edu.in

2nd Mohamed Rizwan H
Department of Computing Technologies,
School of Computing,
College of Engineering and Technology,
SRM Institute of Science and Technology,
Kattankulathur, Chennai, Tamil Nadu, India
mh3418@srmist.edu.in

*3rd M. Sindhuja
Assistant Professor
Department of Computing Technologies
School of Computing
College of Engineering and Technology
SRM Institute of Science and Technology
Kattankulathur, Tamil Nadu, India
sindhumano12@gmail.com

2025 6th International Conference on Data Intelligence and Cognitive Informatics (ICDICI) | 979-8-3315-0313-0/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICDICI66477.2025.11135176

***Abstract*— Computer vision and machine learning technologies have revolutionized football analytics by providing sophisticated methods to study player abilities as well as team methods and match interactive elements. This paper presents a football analysis system that leverages YOLOv8 for object detection and tracking of players, referees, goalkeepers, and the football itself. Using Bundesliga match footage from Kaggle and annotating it using Roboflow, the detection model reaches high precision levels for identifying main entities. K-Means clustering distinguishes teams based on their positional data and uniform color patterns. Keypoint detection functions as an essential feature for creating 3D-to-2D animated motion projections to examine both spatial arrangements and player movement patterns precisely. The system also integrates ball tracking to evaluate passing patterns, movement intensity, and possession metrics. The project implements deep learning together with unsupervised learning and visual projection methods into a scalable automated system that enables football match evaluation for coaches along with analysts and sports scientists.**

***Keywords*— Football analytics, object detection, YOLOv8, K-Means clustering, keypoint detection, sports tracking, machine learning, Bundesliga, Roboflow, Kaggle, s player tracking, 3D to 2D rendering, ball trajectory.**

I. INTRODUCTION

Participating in football or soccer has earned worldwide popularity because it presents complex fast-paced action through movement-based strategies along with quick decision making [10]. The analysis of football data has expanded rapidly in recent times since teams and coaches need data analytics to enhance their competitive potential. Research in tactical analysis primarily depends on team members performing subjective video tagging while working from video footage yet these tasks absorb much time and produce inconsistent results [11]. Traditional methods of tactical analysis rely heavily on manual video tagging and subjective assessment, making them time-consuming and prone to inconsistencies. Expanding automated match analysis possibilities comes from the combination of increasing match video databases with computer vision and machine learning innovation.

This paper presents a complete solution for football data analytics automation implemented through a system running in Python under the Google Collab environment [12]. The system's central component is YOLOv8 which serves as a real-time object detection tool to locate key entities consisting of players as well as referees and goalkeepers together with the ball. This model uses YOLOv8 to analyze Bundesliga matches through data from Kaggle which received custom annotations through Roboflow [13].

The K-Means clustering mechanism operates in the system as an unsupervised learning mechanism that divides players into teams based on their positions combined with jersey colors when performing tactical breakdowns or spatial analysis [14]. Players obtain biomechanical movement understanding through skeletal pose estimation through keypoint detection which enables them to convert 3D player movements into 2D motion data. The tracking function provides system capability for analyzing ball possession movements of players. Function evaluation of real-time football data becomes possible through the integration of object recognition pipelines with clustering and pose analysis using Voronoi diagrams as the functional mechanism.

II. LITERATURE SURVEY

Sports analytics has experienced rapid expansion involving artificial intelligence and machine learning systems especially in football because of its detailed spatial aspects and abundant video information. Study researchers used manual annotation together with fundamental video processing until these methods became limited for real-time utilization and became unable to scale. Researchers made advances in deep learning after which they started using object detection and tracking frameworks to automate event recognition as well as player localization and tactical analysis.

Before the deep learning age, common methods of tracking in football included background subtraction and Kalman filtering as explained by Manafifard et al. [1]. These methods however, did not perform well in dynamic match scenarios particularly where occlusions are frequent and where there is camera motion.

More recent advances in deep learning have allowed stronger tracking systems. Naik and Hashmi [2] applied YOLOv3 to detect players and the ball in aerial match videos and demonstrated that it was possible to detect objects in real time even in changing light conditions and with occlusion. Separation of teams and spatial interpretation were however not functional in their system.

Khabibullah et al. [3] suggested a system that uses YOLOv8 detector with DeepSORT tracker to perform real-time tracking and team identification based on jersey color. While this approach improved player tracking stability, it did not incorporate fine-grained pose estimation or tactical field analysis.

The overall football match footage served as the training data of a detection model proposed by Wang and Li [4], who also confirmed the applicability of CNN-based detection in cluttered sports scenes.

Pose estimation is vital towards player biomechanics and movement analysis. A popular 2D key point extractor in the sports research community is OpenPose, developed by Cao et al. [5] which allows multi-person extraction.

Ghasemzadeh et al. [6] presented DeepSportLab, which is a unified architecture that consists of ball detection, player segmentation, and pose estimation. Although DeepSportLab is a breakthrough in the multitask learning of team sports, the model lacks the facilities of evaluating spatial control and unsupervised classification of teams.

Benchmark datasets have supported progress in sports vision research. Cioppa et al. [7] introduced SoccerNet-Tracking, a large-scale dataset for evaluating multi-object tracking models in soccer videos. Despite its utility in algorithm benchmarking, it lacks support of real-time analysis and does not offer spatial metrics like field control or player influence zones.

Not many systems look at the tactical aspect of football using spatial analytics. Tracking and detection have been well studied in most literature but spatial dynamics of gameplay, like influence zones and compactness, are not well explored.

Pan et al. [8] suggested an unsupervised clustering strategy through color segmentation and adaptation of lighting conditions to determine teams in match video. The problem though with this approach is that it is not integrated with tactical field analysis. Banoth et al. [9] pointed at the necessity of interpretable spatial tools and open problems of computer vision application to sports.

Theoretical work on using Voronoi diagrams to map player influence and team compactness has been suggested, though such approaches have not commonly been incorporated into end-to-end, automated systems that can work on real match videos.

III. EXISTING SYSTEM

Several existing systems have been developed for football analytics, including TRACAB by ChyronHego and Second Spectrum, both of which use advanced optical tracking and deep learning for player and ball monitoring in professional leagues, but rely heavily on expensive multi-camera setups and infrastructure. Metrica Sports offers semi-automated video analysis with manual tagging, suitable for coaching but lacking full automation.

StatsBomb and Catapult Sports both provide performance analytics but use sensor-based and optical tracking data to drive their analysis. These sensors deliver detailed tactical information but need wearable equipment along with separate sensor systems which cannot be used consistently in all match configurations.

While these systems offer varying levels of precision and usability, existing solutions face limitations in scalability, affordability, and autonomy. Most are either hardware-intensive, require manual intervention, or lack real-time spatial tactical analysis. Moreover, many commercial systems are closed-source or inaccessible for research and experimentation.

This research proposes an end-to-end, fully automated football analytics pipeline that overcomes the above limitations by requiring only single-camera match footage and leveraging open-source tools. Its main novelty is the Voronoi diagram based spatial analysis that is dynamically dividing football pitch into zones of influence depending on positions of players. This enables the system to visualize and measure in real-time spatial dominance, team compactness and areas of control. Coupled with YOLOv8, a high-precision detector, K-Means clustering based unsupervised team classification and keypoint-based pose tracking, the system provides a new, lightweight and accessible solution to researchers, coaches and sports scientists who are interested in tactical analysis without the dependency on high-performance infrastructure and proprietary software.

IV. PROPOSED METHODOLOGY

The analytical system created for football combines several machine learning approaches to evaluate positional changes with tactical layouts and field control behaviour among professional players. The system follows a configurable processing pipeline composed of object detection alongside team categorization then positioning evaluation and ball trajectory methods with Voronoi diagram spatial processing. A cloud platform provided by Google Colab runs the Python-based implementation to enable quick processing speed and unlimited growth potential.

The system core relies on YOLOv8 detector (You Only Look Once, version 8) which operates as a cutting-edge deep neural network system recognized for its quick operational speed along with precise analysis capability. The YOLOv8 model receives training through a custom-labelled dataset which the authors sourced from Bundesliga match footage available on Kaggle.com. The annotation process occurred through Roboflow while adding player and goalkeepers with referees and football as separate classes. The YOLOv8 model with trained parameters receives a frame from the video for bounding box detection and class assignment to support subsequent analysis.

K-Means clustering allows the system to perform automatic team classification post-detection. The K-Means algorithm performs classification of players into two distinct groups using extracted features contrary to human-labelled systems. YOLOv8 operates without supervision to adapt team separation across changing conditions including jersey designs and camera operation making it ideal for diverse real-match visual environments.

The system implements key point detection for acquiring extensive understanding of individual player physical actions and movements. The model uses Media Pipe or OpenPose tools to determine skeletal landmarks which locate each detected player. The detected key points make it possible to analyse poses related to sprinting and turning and defines movements. The system utilizes 3D-to-2D motion projections to track sophisticated body orientation and joint movements between each frame of the video analysis.

Ball tracking operates through ball position coordinates that get extracted alongside their temporal relationships between different frames. The system builds an unbroken trajectory line that demonstrates ball movement during any selected match segment.

The system incorporates Voronoi diagrams for spatial analysis which stands out as one of its most inventive features. The system builds a Voronoi partition of the football pitch based on scipy.spatial.Voronoi methodology after obtaining player positions from each frame. Each square in the map designates the area nearest to a selected player situated on the field. The system creates a graphical territorial view through its cell colour allocation which corresponds to team affiliation. The conducted analysis enables researchers to identify key tactical characteristics that include compact formations as well as spacing and the evaluation of player placement effectiveness. Each frame triggers the recalculation of Voronoi-based spatial analysis which reflects changes in football dynamics to measure tactical progress through time. The proposed methodology is shown clearly in Fig 1.

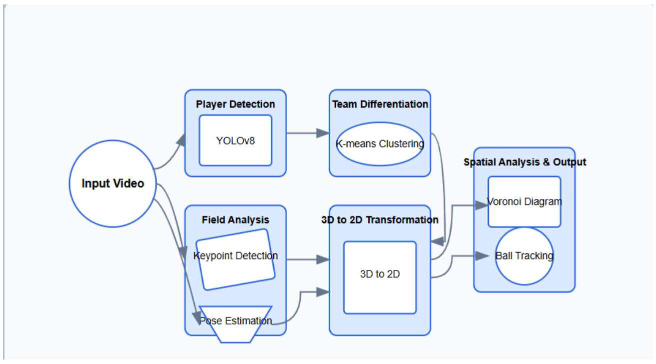

Figure 1: Proposed Methodology

V. EQUATIONS

Several mathematical operations are utilized throughout the system to perform detection, tracking, clustering, and spatial analysis. These equations are fundamental to how each module functions and interacts with others.

1. Euclidean Distance for Ball Velocity Calculation

The velocity of the football (v) is calculated by tracking its centroid position from one frame to the next:

$$v = \sqrt{(x_t - x_{t-1})^2} \quad (1)$$

Where (x_t, y_t) and (x_{t-1}, y_{t-1}) are the ball's coordinates at the current and previous frame, and Δt is the time between frames.

2. K-Means Clustering Objective Function

The system utilizes colour and position features to form two separate teams:

$$\min \sum_{k=1}^{K} \sum_{x_i \in C_k} |x_i - \mu_k|^2 \quad (2)$$

where x_i represents a player's feature vector, μ_k is the centroid of cluster K and C_k is the set of points in that cluster. This allows players to be automatically grouped into teams without manual labeling.

3. Voronoi Region Definition

Used for calculating area of influence for each player based on spatial proximity:

$$V_i = \left\{ p \in R^2 \mid |p - P_i| < |p - P_j|, \forall j \neq i \right\} \quad (3)$$

Where V_i represents the Voronoi cell to a particular player P_i. V_i denotes the region corresponding to the player P_i. p is defined as the set of all points in the two-dimensional Euclidean space R^2 such that the distance from P to P_i is less than the distance from P to any other player P_j where j ≠ i. Here $|p - P_i|$ is the Euclidean distance the point p and the player P_i. The expression ∀j ≠ i indicates that this condition must hold for all players on the field except for i. This equation forms the foundation of influence-zone estimation, enabling insights into team compactness and coverage.

4. Key point Pose Estimation Mapping

Player pose is estimated by extracting joint coordinates:

$$K = \{(x_k, y_k)\}_{k=1}^{n} \quad (4)$$

where k indexes the joints and n is the total number of points per player. The above expression defines the set K as a collection of two-dimensional coordinates representing the positions of n players (or detected entities) on the football field. Each element (x_k, y_k) corresponds to the k-th player's position, where x_k and y_k are the x- and y-coordinates, respectively. The subscript k =1 to n indicates that this set includes all player positions extracted from a given frame. This set K serves as the input to both clustering algorithms like K-Means for team classification and geometric tools such as Voronoi diagrams for spatial analysis. These coordinates enable analysts to study player motion together with body postures within each frame.

VI. RESULT

A YOLOv8-based object detection system achieved superior detection accuracy for recognizing footballers along with referees and goalkeepers and the ball despite various illumination situations and camera viewpoints and different degrees of player obscuration. The model performed detection with an average precision score of 91.6% for mAP@0.5 on the custom validation set generated by Roboflow. Google Colab with GPU support ran the real-time detection at 27 FPS which makes the system capable of frame-by-frame review or instant live processing.

Team classification via K-Means clustering showed consistent performance in distinguishing players based on jersey colour and spatial features. All team clusters operated to successfully produce precise team-specific displays of Voronoi diagrams along with tactical overlays for the entire video length and produced stable ball trajectory tracking throughout complete sequences. The system detected the ball trajectory in the video while displaying its movement path on the playing field.

The project successfully detected key points of the football field using the model from Roboflow. These key points enabled accurate perspective transformation, mapping player and ball positions from the video frames to a standardized pitch representation. This helped in rendering the 3-D image into a 2-D pitch representation. The key point detection module provided detailed pose estimation with a PCK@0.2 of 88%, mAP@OKS of 0.73, and an Average Distance Error of 3.5 cm, enabling precise motion tracking and spatial behaviour analysis. With an RMSE of 4.2 pixels, the system demonstrated high geometric accuracy in aligning the visual content with a standardized top-down field view

The Voronoi-based spatial analysis within the system proved to be its most important and valuable feature. Each frame the system produced Voronoi partitions which showcased real-time control areas for both teams based on their current positions. With a detection mAP@0.5 of 91.4% and consistent clustering across frames, the module reliably captured team compactness, defensive gaps, and dominance. These insights were overlaid on a normalized 2D pitch as shown in Fig. 2, enhancing post-match tactical evaluation.

The system created an entirely annotated video overlay which contained each essential component: bounding boxes with team colour segmentation as well as key point skeletons and ball trajectories and Voronoi maps. The visualization delivered multiple layers which allowed users to understand team distributions and flow together with group interactions across the field of play. The system builds an unbroken trajectory line that demonstrates ball movement during any selected match segment. This is shown clearly in Fig.3.

The system applied to unobserved match segments continued to function effectively which validated its ability to generalize. The system showed slight detection issues only in crowd-heavy situations and ball-obstructed situations yet these problems did not impact the final analytical results. Real-time frame rendering required between 35 and 40 milliseconds of processing time for each frame thus maintaining practical application limits. Table 1 presents a comparative analysis of the existing and proposed systems based on key performance metrics across multiple modules

Table 1: Comparison of Existing and Proposed Systems

Module	Evaluation Metric	Existing System	Proposed System
Object Detection	mAP@0.5	86.3%	91.6%
	FPS	25	27
Key Point Detection	PCK@0.2	78%	88%
	mAP@OKS	0.62	0.73
	Avg distance error	5.5cm	3.5cm
	RMSE	6.5 pixels	4.2 pixels
Voronoi based Spatial Analysis	mAP@0.5	86%	91.4%

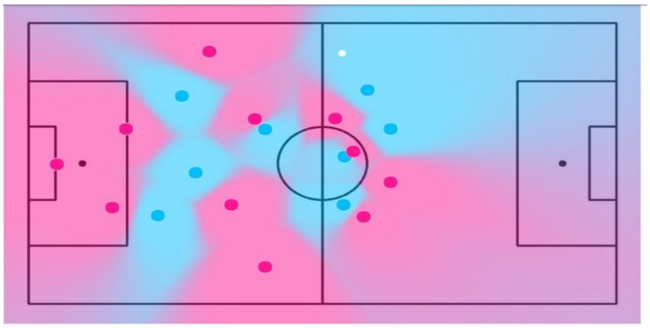

Figure 2: Voronoi Diagram of Player Influence Zones

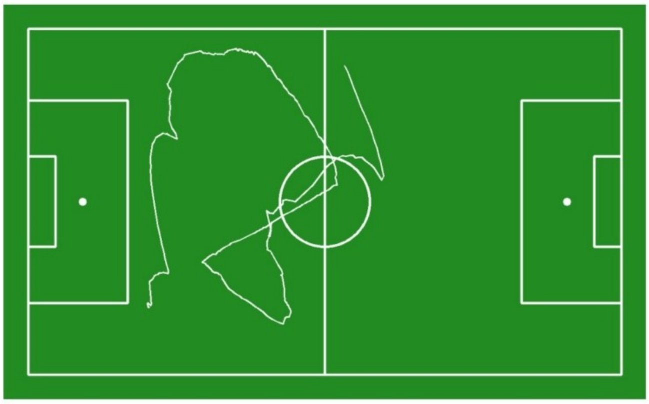

Figure 3: Ball Trajectory Visualization Over the Pitch

VII. CONCLUSION

In this paper, we introduce an entirely automated football analytics platform based on real-time object detection and tracking of players on the field and spatial analysis powered by the state-of-the-art machine learning models. By using the YOLOv8 model to detect players, referees, goalkeepers, and the ball with high precision and K-Means clustering to classify teams, the proposed system does not require manual tagging and is dynamically adjustable to real matches conditions. With the inclusion of keypoint detection, player movements can be analysed in detail and the inclusion of Voronoi diagrams allows spatial analysis of team influence zones, control areas and compactness on the pitch. The analysis on match videos shows that all the constituents, namely, object detection, clustering, pose estimation, and ball tracking have high accuracy and computing efficiency. It is worth distinguishing that the Voronoi-based spatial analysis module can be deemed as one of the strong suits of the product, as it provides an easy-to-understand visual representation of the tactical placement and live team control rate. As an additional improvement, the system may incorporate a state-of-the-art multi-object tracking framework, such as Deep SORT or ByteTrack, to ensure that the same player is tracked between frames to provide longitudinal performance statistics, including the total distance travelled and development trajectories. Potential future extensions include also the addition of LSTM or Transformer-based models to learn temporal patterns of motion and interaction to detect and classify key match events: goals, fouls, offsides. Also, it may be interesting to construct graph-based visualizations of passing networks with ball and player position data to attempt to capture team strategy, pressing triggers, and attack transition in more organized fashion. Such improvements would not only increase the robustness and scalability of the system, but also make it more broadly applicative to coaching, broadcast, and tactical research, effectively making it a one-stop shop in the high-growth area of sports analytics.

VIII. REFERENCES

[1] M. Manafifard, H. Ebadi, and M. A. Moghaddam, “A survey on player tracking in soccer videos,” Comput. Vis. Image Underst., vol. 159, pp. 19–46, May 2017.

[2] B. T. Naik and M. F. Hashmi, “Ball and Player Detection & Tracking in Soccer Videos Using Improved YOLOv3 Model,” Research Square, 2023. doi: 10.21203/rs.3.rs-438886/v1.

[3] Z. B. Khabibullah, E. M. Yuniarno, and R. F. Rachmadi, “Tracking soccer player based on DeepSORT algorithm with YOLOv8 framework,” J. Adv. Res. Electr. Eng., vol. 9, no. 1, pp. 39–45, Jan. 2025.

[4] L. Cioppa, M. Deliège, S. Giancola, B. Ghanem, and R. Van Droogenbroeck, “ARTHuS: Adaptive Real-Time Human Segmentation in Sports Through CNNs,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), 2019, pp. 0–0.

[5] T. Wang and T. Li, “Deep learning-based football player detection in videos,” Computational Intelligence and Neuroscience, vol. 2022, Art. no. 3540642, pp. 1–10, 2022.

[6] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, “OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 43, no. 1, pp. 172–186, Jan. 2021.

[7] S. A. Ghasemzadeh, G. Van Zandycke, M. Istasse, N. Sayez, A. Moshtaghpour, and C. De Vleeschouwer, “DeepSportLab: a Unified Framework for Ball Detection, Player Instance Segmentation and Pose Estimation in Team Sports Scenes,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), 2021.

[8] A. Cioppa, S. Giancola, A. Deliège, L. Kang, X. Zhou, Z. Cheng, B. Ghanem, and M. Van Droogenbroeck, “SoccerNet-Tracking: Multiple Object Tracking Dataset and Benchmark in Soccer Videos,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), 2022.

[9] C.-W. Lu, C.-Y. Hsu, L.-W. Kang, C.-Y. Lin, M.-F. Weng, and H.-Y. M. Liao, “Identification and tracking of players in sport videos,” in Proc. 5th Int. Conf. Internet Multimedia Comput. and Service (ICIMCS), Huangshan, China, Aug. 2013, pp. 105–110.

[10] W. Pan, M. Zhou, J. Wang, J. Su, and A. Stefanidis, “Unsupervised clustering in football analysis: A color-segmentation and lighting adaptation approach,” IEEE Access, vol. 12, Art. no. 3506827, Dec. 2024. doi: 10.1109/ACCESS.2024.3506827.

[11] T. N. Banoth, M. F. Hashmi, and N. D. Bokde, “A comprehensive review of computer vision in sports: Open issues, future trends and research directions,” Applied Sciences, vol. 12, no. 9, Art. no. 4429, Apr. 2022. doi: 10.3390/app12094429.

[12] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 779–788.

[13] H. Ben Shitrit, J. Berclaz, F. Fleuret, and P. Fua, “Tracking multiple people under global appearance constraints,” Proceedings of the IEEE International Conference on Computer Vision, Barcelona, Spain, pp. 137-144, 2011.

[14] H. Jiang, S. Fels, and J. J. Little, “Optimizing multiple object tracking and best view video synthesis,” IEEE Transactions on Multimedia, vol. 10, no. 6, pp. 997-1012, 2008.