Journal of Autonomous Intelligence (2024) Volume 7 Issue 5
doi: 10.32629/jai.v7i5.1581

ORIGINAL RESEARCH ARTICLE

Video analysis and data-driven tactical optimization of sports football matches: Visual recognition and strategy analysis algorithm

Biao Jin

Faculty of Humanities and Social Sciences, Beijing University of Technology, Beijing 100000, China; ab2023000@163.com

ABSTRACT

For the purpose of this research, an original technique to assess football matches is described. The strategy makes use of a set of innovative algorithms for Strategic Analysis (SA) and Visual Recognition (VR). The approach, as mentioned above, has been designed around a virtual reality (VR) platform that is centered on YOLOv5 and successfully monitors the actions of both players and the ball in real-time. With the guidance of Markov Chain Models (MCM), the resulting information is processed and evaluated in order to find correlations in player location and actions. This enables an in-depth comprehension of the tactics and plans the team's management executes. One of the most significant components of the research project is the exploration of multiple approximation techniques with the aim of enhancing frame analysis performance. Furthermore, threshold scaling was executed in order to attain maximum accuracy in detection, and an approach for Steady-State Analysis (SSA) is being created in order to analyze the long-term strategic positions of teammates. This complete method can run on sophisticated knowledge of in-game tactics, and it also serves as a tool for trainers and players who want to increase the effectiveness of the teams they coach and counteract strategies used by the opposing team.

Keywords: visual recognition, strategic analysis, YOLOv5, Markov chain models, football game, machine learning

ARTICLE INFO

Received: 5 February 2023
Accepted: 23 February 2023
Available online: 12 March 2024

1. Introduction

As an outcome of the complicated nature of the sport of football, which has been defined by significant strategic involvement and immediate changes in play, the field of sports data analysis has been presented with a specific set of problems and possibilities[1]. The issues are the main objective of this investigation, resulting in the use of modern technology to go beyond standard review techniques. The current study intends to offer a deep understanding of matches in football by applying the advantages of video analysis and various strategies that are motivated by statistics. The maximum amount of minor technical refinements is often necessary when determining whether or not a team wins or loses when playing contemporary football[2]. A significant impact on the decoding of these ambiguities has been contributed by improved Strategic Analysis (SA) research, which is the primary topic of this investigation. The overall objective of the present study is to further develop our knowledge regarding the sport through the analysis of player and ball motions, team formations, and within-the-game tactics employing sophisticated computer algorithms. This should give an essential benefit to players and trainers during the entire match.

In contrast to having the advantage that the principles are essential, classic football evaluation techniques have drawbacks in terms of their scope, speed, and accuracy. Conventional statistics typically fail to adequately record the constantly changing and reactive aspect of the game[3]; however, hands-on data collection is one of the most excellent time-consuming techniques and is susceptible to error by humans. Because of such limitations, an increasingly refined and computerized technique must be used in order to accurately represent the entire range of the game's level of logistical issues[4]. A substantial part of the approaches that are presently in practice depends on personal observation and basic mathematical modeling, both of which are inappropriate for the fast-paced and complicated environment of the sport of football. Frequently, the techniques mentioned above produce a superficial review of the game's strategical elements because they provide the level of complexity and accuracy essential to an in-depth knowledge of strategies[5].

By addressing the drawbacks of conventional football study, the present investigation has been motivated by the requirement to find solutions to those drawbacks. Applying virtual reality (VR) when combined with a data-driven method of evaluation is an innovative method that is explored in this research article. Automating the technique of SA is an aim of this integration process, which promises findings that are more accurate, objective, and current. The fundamental goal of this project is the creation of an advanced mathematical model for the purpose of performing SA in the sport of football. This model will integrate YOLOv5[6] VR+Markov Chain Models (MCM).

The intended use of this article can be included by:

1) Systematizing participant and ball movement: With the introduction of YOLOv5, a software-driven and real-time tracking framework of participants and the ball, the dependence on traditional techniques will be minimized.
2) Numerical strategic visions: A statistical knowledge of SA and team tactics can be achieved by the application of MCM for the investigation of geographical information.

The recommended technique, "Video Analysis and Data-Driven Tactical Optimization of Sports Football Matches: VR and SA Algorithm," integrates cutting-edge YOLOv5-based VR with MCM-based SA study for football matches. Real-time object detection is implemented in this revolutionary approach with the intent of monitoring participants and the ball as they move, as well as for the aim of collecting and preliminary processing data for SA. The approach has the potential to identify significant tactical knowledge and team tactics through the examination of participant motions and geographic information through the implementation of probabilistic mathematical models. In addition to this goal, it features enhancements with methods for simple visualization and a system for continually acquiring knowledge, which makes the development of the model possible. In the context of game statistical analysis, this analysis method presents a significant development because it presents an in-depth and data-driven approach to analyzing and improving football tactics.

The paper is organized as follows: Section 2 presents the methodology, Section 3 presents the evaluation of the model, and Section 4 concludes the work.

2. Methodology

Football match analysis through video encounters significant challenges, predominantly due to the dynamic nature of the game[7,8]. The key complexity involves players' continuous, rapid movement, often overlapping and interacting in complex patterns[9]. Additionally, varying camera angles, necessary for comprehensive coverage, introduce perspective distortion, complicating consistent player positioning and movement tracking. The demand for real-time analysis adds another layer of complexity[10]. In a sport where strategies and decisions are time-sensitive, relying solely on post-match analysis is inadequate. Real-time data processing and interpretation offer immediate insights, crucial for decision-making during the match[11,12].

Practical football analysis hinges on precise detection, tracking, and identification of players, the ball, and referees. Accurate recognition is essential for understanding player positions and movements and evaluating team formations and strategies[13,14]. Continuous tracking across frames is necessary to construct a detailed picture of game dynamics, encompassing individual performance and team coordination. The proposed methodology involves two phases. The first phase involves YOLOv5 model-based player, ball detection, and MCM-based strategy analysis[15,16].

The following section details the functionalities of these models in detail:

2.1. YOLOv5 for object detection

Object detection is a critical component in the analysis of sports footage, particularly in a fast-paced game like football. It serves as the foundation for further tactical and performance analysis by providing the raw data on which additional assessments and decisions are based. YOLOv5 (You Only Look Once, version 5) is one of the versions in the series of YOLO algorithms that offers enhanced speed and accuracy for real-time object detection tasks[17,18]. A technology designed to identify things, YOLOv5 was significantly increased in terms of detection speed and functionality[19,20]. In a single pass, this technique examines images, concurrently identifying the bounding boxes and class probabilities for these boxes[21,22]. The title of this approach indicates that it analyzes images in sequence. Due to the holistic approach that it requires, YOLOv5 has the capability to perform its functions substantially faster than systems that perform these tasks in chronological order[23–25]. The backbone system, the neck, and the head are the three distinct components that collectively comprise the framework of YOLOv5. The task is assigned to the backbone to extract features from the image being processed, the neck to deal with these features through the use of another layer in order to further enhance the detection of objects across an array of dimensions, and the head to generate the definitive predictions of bounding boxes and class labels[26,27].

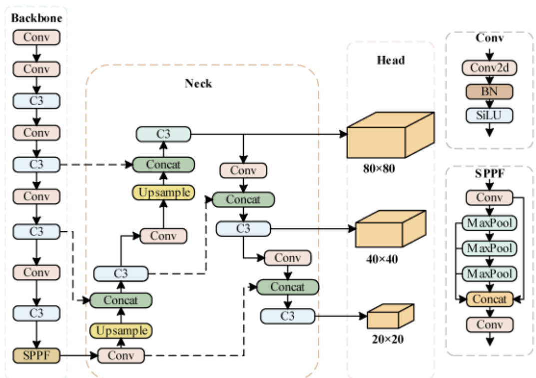

Figure 1. Architecture of YOLOv5.

The YOLOv5 method has been developed on an extensive collection of images, including images of participants, balls, and coaches in several distinct actions (**Figure 1**). In the context of its coaching, it develops the capacity to identify these objects within an intricate framework, including a football pitch, a crowd of people, as well as additional features that are prevalent in a match context[28]. When training an artificial neural network, it is crucial to manipulate the weights of the network in order to decrease the number of detection failures. This can be achieved via the use of enormous data sets which encompass a considerable number of cases. When implemented in the game of football, YOLOv5 performs its functions frame by frame, finding and identifying objects that are relevant, such as participants, balls, and coaches. The method based plays a role in defining bounding boxes for each detected object and labelling them with a class and a trust level. This is an integral component of the steps associated with the object identification method. The physical coordinates that determine the form and position of the object within the context of the frame are commonly referred to as bounding boxes[29,30].

A significant advantage of YOLOv5 is that it finds a balance between speed and accuracy. The investigation of a game of football requires it to take place in real-time or within close proximity to real-time for it to be appropriate in determining changes to the plan[31,32]. This balance is essential. While ensuring a rate of motion that is suitable for real-time study, YOLOv5 possesses the capacity to generate detection results that are comprehensive enough and accurate. YOLOv5 is a powerful tool in football analysis, allowing real-time video footage analysis and immediate response for coaches and analysts. Its accuracy ensures reliable data for SA, minimizing errors from misidentified or undetected objects. The output, bounding boxes, and class labels are inputs for tracking and identification processes, enabling accurate monitoring of player movements and contributing to a comprehensive dataset for post-match analysis[33,34].

2.2. Reference system

First, let's use an examination of the methodology that provides an explanation for the reference framework that is utilized for mapping the relative positions of players from a frame of video to a 2D surface for the goal of performing SA in the game of football:

Pre-Masking the pitch: The process begins by preparing a clean slate of the football pitch, devoid of players, referees, and other transient objects. A digital image of the pitch is pre-masked (**Figure 2**), meaning all elements except for the pitch itself are filtered out. This creates a 'template' that can be used to identify the pitch within the actual video frames captured during a match. The importance of this step lies in its ability to provide a clear view of the pitch, which is essential for accurate homography estimation.

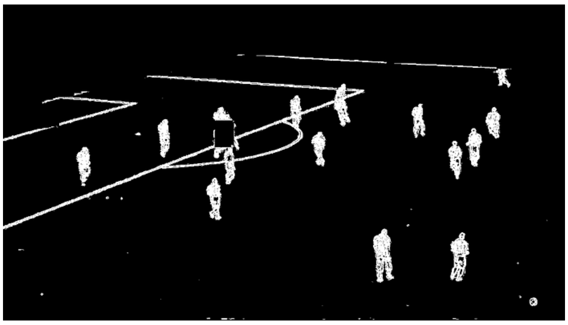

Figure 2. Pre-masked pitch image.

Homography estimation: Homography refers to the transformation that maps the points from one perspective to another. In this context, the aim is to map the positions of players and objects from the camera's perspective to a top-down view of the pitch. To achieve this, the system uses pre-computed pitch images representing the field from various angles and positions. These are a reference for matching the current video frame to the correct angle and position.

Efficient matching via indexing: Once the pitch is isolated, the current frame is treated as a "query" within an indexing system that holds the pre-computed set of pitch images. The system then searches for the closest match to the current frame within this index. By treating the frame as a query, the matching process is made more efficient, enabling the system to quickly find the most similar pitch image that corresponds to the current frame's perspective.

Projection onto a 2D field: After identifying the best match, the homography matrix derived from this matching process is applied to the detection boxes from YOLO, which contain the positions of players and objects on the pitch. This matrix mathematically warps the perspective of these positions so they align with the 2D pitch model, effectively translating the real-world positions onto a flat, 2D plane. Regardless of the camera's motion or direction, this procedure is necessary to ensure that the geographical accuracy of the SA is preserved. It provides an uninterrupted frame reference point to be generated.

Utilization of a flask server: The entire process, from homography estimation to the projection of player positions, is facilitated through a Flask server. Flask is a micro web framework in Python that is well-suited for such tasks due to its ease of use and flexibility. In this setup, the Flask server likely handles the processing of video frames, the application of the homography matrix, and the presentation of the transformed 2D positions in real-time or for subsequent analysis (**Figure 3**). This Flask server could allow for interactive analysis, where coaches and analysts can view and manipulate the data, perhaps overlaying tactical or historical data for comparison. The server architecture also suggests that this system could be accessed remotely, providing a versatile tool for teams to analyze matches from any location.

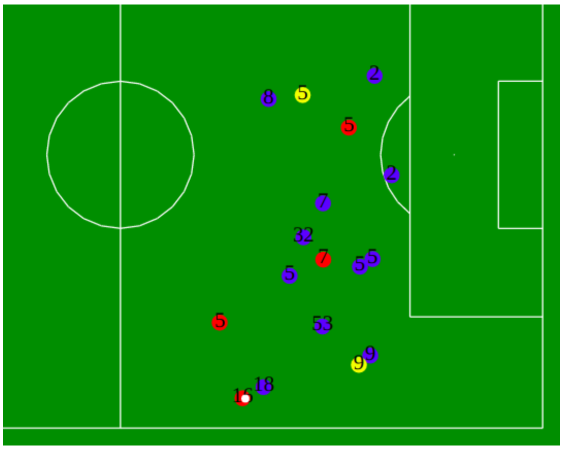

Figure 3. 2D rendered field with player projection.

Final output: The final output, as illustrated by the images provided, is a clean and clear representation of the pitch with players' positions accurately projected onto it. For the objective of determining SA, this result acts as a basis for additional investigation, which may require quantitative analysis of data such as participant actions, geological formations, and distance.

The primary framework that has been presented is a highly advanced system that combines image processing and Machine Learning (ML) methods, which has resulted in an effective tool to perform SA in the sport of football. YOLO is a tool for object detection, and novel techniques have been used for homography prediction. All of this is coordinated within a web-based architecture that is adaptable, allowing real-time and SA features.

2.3. Object tracking using KLT tracker

The design is vital to track objects in the actual practice of sports statistical analysis, especially in the game of football, in order to understand the actions of participants properly, the line-ups they implement, and the techniques that players implement. The Kanade-Lucas-Tomasi (KLT) tracker[35–36] is a technique that is based on features and has been recognized for having the capacity to track objects throughout an array of images or video frames. In the framework of football games, it allows for the accurate tracking of participants and the ball's movement over the duration of the competition. The KLT tracker is based on detecting corner features within an image and tracking these features from frame to frame. The region in the image where there is an essential variation in brightness in every possible direction is commonly referred to as an edge[37,38]. As a result of their susceptibility to noise and blocking, these points of interest tend to be very suitable for tracking purposes.

The KLT tracker tracks footballers and the ball's movement. Given that there is minimum motion and variation in form between frames per second, it constantly predicts the movement of the objects. As the player relocates throughout the playing field, the device that tracks maintains track of player features such as periodic edges and ball positions. At the same time, it presents an ongoing log for each object of significance.

2.4. Player identification with convolutional neural networks (CNN)

For the intended use of monitoring participant actions and tactics, CNN was trained to identify participant numbers that were displayed on the backs of players' jerseys. The plan of action included improving the Street View House Numbers (SVHN) dataset, which had initially been created for something else but appeared to be jersey numbers in style. As a result, the framework was able to adjust to the field successfully, and its resilience and accuracy for determining numbers under other match circumstances were both boosted.

The training model for football games faces challenges in obtaining labelled images due to accessibility and licensing issues. A creative approach is needed to leverage available resources effectively. The identification process starts with a player's jersey number, assigning an identifier (ID) and tracking logic for consistent monitoring across frames. A smoothing mechanism increases the reliability of identification over time.

The model's efficiency can be enhanced by incorporating team information, such as team line-ups and rosters, which allows for a more constrained range of possible numbers and improves the accuracy of player identification (**Figure 4**).

Figure 4. Results of player identification.

2.5. Team recognition via K-means clustering in HSV color-space

Distinguishing between two teams in a football match is approached by employing the K-means clustering algorithm, leveraging the uniformity of team shirt colors. The algorithm operates in the HSV (Hue, Saturation, Value) color-space, which is particularly suited for color-based segmentation tasks due to its alignment with human color perception and its decoupling of color intensity from color information. The process starts by extracting the bounding boxes surrounding each player obtained from the object detection phase. These bounding boxes are then analyzed in the HSV color-space, allowing the K-means algorithm to cluster pixels based on their color characteristics, effectively grouping players by the dominant jersey colors.

To refine the clustering and mitigate the influence of the green playing field, the term frequency-inverse document frequency (TF-IDF) weighting is applied. This statistical measure helps reduce the weight of colors standard across the entire image (like the field's green) and emphasizes the unique colors of each team's jerseys within the bounding boxes. The outcome of this process is the successful segmentation of players into two distinct clusters, each corresponding to a team (**Figure 5**). This color-based team recognition allows for subsequent analysis phases, such as player tracking and tactical assessment, to be performed with an understanding of team affiliations.

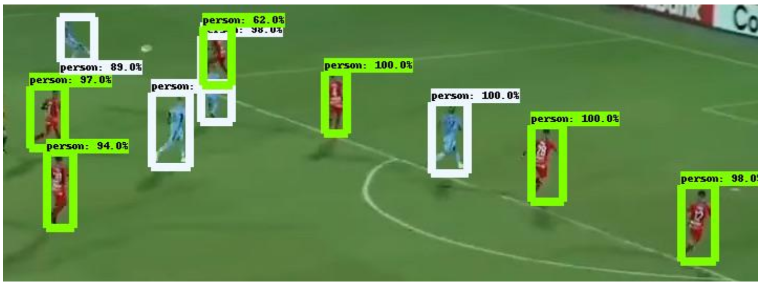

Figure 5. Result of the team recognition algorithm.

2.6. Markov strategy analysis algorithm in football

The MCM Strategy Analysis Algorithm is employed to predict the strategic plays in football. By treating each play or sequence of player movements as a transition between states, the MCM analyzes the probability of shifts from one tactical situation to another, providing a statistical framework for strategy analysis. An MCM is a stochastic model that uses the MCM property, which posits that the future state of a process only depends on the current state and not on the sequence of events that preceded it. In the context of football, this translates to the principle that the next play or player movement is dependent only on the current arrangement of players and ball position.

States and transitions in football analysis: In an MCM applied to football, the pitch is segmented into zones, each representing a state. Let's denote the set of all possible states as $S = \{s_1, s_2, \ldots, s_n\}$. The transition from one state to another occurs due to players' movements and ball dynamics.

Transition probabilities: The transition probability matrix P, where each element P_{ij} represents the probability of transitioning from state s_i to state s_j. This matrix is constructed as follows, Equation (1).

$$P = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1} & P_{n2} & \cdots & P_{nn} \end{bmatrix} \tag{1}$$

The construction of P is based on historical match data, and it is a stochastic matrix, meaning each row sums up to 1, Equation (2)

$$\sum_{j=1}^{n} P_{ij} = 1, \ \forall i \tag{2}$$

Predicting future states: The game's current state game at any time t can be represented as a probability vector $S(t)$, where each element corresponds to the likelihood of being in a particular state. The next state $S(t+1)$ is predicted by, Equation (3)

$$S(t+1) = S(t) \times P \tag{3}$$

Higher-Order Markov models: For more complex analyses, where the future state depends on more than just the current state, higher-order MCMs are used. In a second-order MCM, for instance, the transition probability might rely on the present and the immediately preceding state. This would involve a 3D matrix P_{ijk}, where the transition to state k depends on being in state j and having previously been in state i.

Steady-State Analysis (SSA): The long-term behavior of the system, or the Steady-State Distribution (SSD) π, is particularly insightful. This distribution is found by solving, Equation (4).

$$\pi P = \pi \tag{4}$$

Subject to the normalization condition: $\sum_{i=1}^{n} \pi_i = 1$.

The SSD illustrates the most frequent game zones over a while, displaying strategic structures and participant positioning.

Algorithm 1 Markov SA in Football

1: **Step 1: Data collection and preprocessing**

2: Aggregate game data: Football match data: Player positions, ball motion, and game-related events.

3: Separate pitch into zones: Separate the football pitch into zones. MCM zones represent states.

4: Record state transitions: Trackball movements during the match. Transition probability matrix design requires this data.

5: **Step 2: Constructing the transition probability matrix**

6: Initialize matrix: Set a matrix P with dimensions equal to the sum of zones (states). Each section P_{ij} Signifies the probability of movement from zone i to j.

7: Count transitions: For each pair of zones, count zone transitions.

8: Calculate probabilities: To figure out probabilities, divide movement counts by zone-originating movements. Sum each matrix row to 1.

9: **Step 3: Analyzing game states**

10: Define initial state vector: Generate an original state vector $S(t)$ denoting the probability delivery of the ball in each zone at the start of the analysis period.

11: Predict future states: Increase the present state vector by the evolution matrix to predict the following state: $S(t+1) = S(t) \times P$.

12: Iterate for further predictions: Reiteration the duplication for subsequent time stages to predict further into the game.

13: **Step 4: Higher-Order models (optional)**

14: Extend to higher-order models: Examine earlier states with higher-dimensional move on matrices for a more complex analysis.

15: Adjust state vectors and matrices accordingly: Incorporate numerous past states into state vectors and move on to matrix structures.

16: **Step 5: SSA**

17: Calculate SSD: Solve $\pi P = \pi$ to find the SSD π, which provides the long-term likelihood of the ball being in each zone.

18: Normalize the steady-state vector: Confirm the total of probabilities in π equals 1.

19: **Step 6: Interpretation and tactical insights**

20: Analyze transition patterns: Observe the transition matrix and state predictions to recognize the game's flow.

21: Utilize SSD: Use the SSD to classify zones of high involvement or strategic status.

3. Experiment analysis

3.1. Dataset

The work involved 500 full soccer games from 6 leading European leagues from 2014 to 2017 on Soccernet, accumulating 764 hours. The set features strategy and methods of play from different teams and tournaments. Data quality has been improved by 6637 dynamic temporal tags extracted from online match data. A one-minute resolution is employed to record goals, YELLOW/RED cards, and substitutions. Specific event descriptions help identify match highlights for analysis.

Several crucial steps have been incorporated in the preliminary processing stage of this set of data. For the objective of sustaining an evenly distributed dataset, which is required for future analyses, video frames are first extracted at a comparatively stable frame rate. Because the frames that correlate to the highlighted actions are likely to be of high strategic value, particular focus is taken when identifying frames that match those actions. Several normalization methods are subsequently carried out on these frames. The resolution of the images should be converted to a high-definition layout, and the level of information ought to be verified to ensure that it is sufficient for accurate object detection. As a secondary advantage, the method of processing becomes more accessible by transforming all of the frames to a standard file type, such as JPEG or PNG.

Techniques for color correction are applied in cases where they are considered required due to the fact that the atmospheric circumstances and camera settings for different matches and locations can vary dramatically. Visual errors that could have an effect on the analysis are minimized because of this phase, which is highly significant. Last but not least, the descriptions that have been incorporated into the dataset are carefully reviewed to verify that they are accurate and relevant to the project's core topic of study. The basis for object detection and MCM-SA has been established by comprehensive validation of data, providing innovative and accurate SA in the sport of football.

3.2. Analysis of YOLOv5 for object detection

For the goal of attaining a balance between detection speed and accuracy, YOLOv5 has been programmed with the settings shown in **Table 1**.

Table 1. Hyperparameter for YOLOv5.

Parameter	Description	Configuration value
Input Resolution	Resolution of input images	640 × 640 px
Detection Threshold	Minimum confidence level for detections	0.25
Non-Maximum Suppression (NMS) Threshold	Overlap threshold for suppressing redundant bounding boxes	0.45
Anchor Boxes	Sizes of anchor boxes for detecting objects at different scales	[10 × 13, 16 × 30, 33 × 23, 30 × 61, 62 × 45, 59 × 119, 116 × 90, 156 × 198, 373 × 326]
Model Architecture	A version of YOLOv5 (e.g., small, medium, large, X-Large)	YOLOv5l
Batch Size	Number of images processed simultaneously	16
Learning Rate	The rate at which the model learns during training	0.01
Augmentation Strategies	Techniques for augmenting training data (e.g., flipping, scaling)	Random flipping, brightness/contrast adjustments
Weight Decay	Regularization parameter to prevent overfitting	0.0005
Class Weights	Importance weights for different classes (e.g., player, ball)	Players: 1.0, Ball: 1.5

Objects such as participants, the ball itself, and coaches have categories and are recognized by the model as it analyzes every video frame. It does this by providing bounding boxes with scores of trusts in order to guarantee accurate findings. It is feasible to extract data from the detections, like the coordinates of bounding boxes, the object class, and trust scores. This structured data is vital when performing an SA in the future. The process manages obscured features and fast-changing circumstances. Dynamic thresholding and interpolation of frames improve detection accuracy.

The accuracy of the YOLOv5 model can be impacted by variations in the threshold, as shown in **Table 2** and **Figure 6**, respectively. In contrast, more significant thresholds may result in missing particular actual objects but with fewer false detections. A reduced threshold may capture more objects, but it also runs the risk of obtaining FP. Experimentation and testing using ground truth data are common approaches to identifying the most suitable threshold setting. This setting usually finds a balance between the requirement for accuracy and minimizing the number of FP.

Table 2. Detection accuracy vs threshold value.

Threshold value	Detection accuracy (%)
0.15	82%
0.20	85%
0.25	88%
0.30	86%
0.35	83%

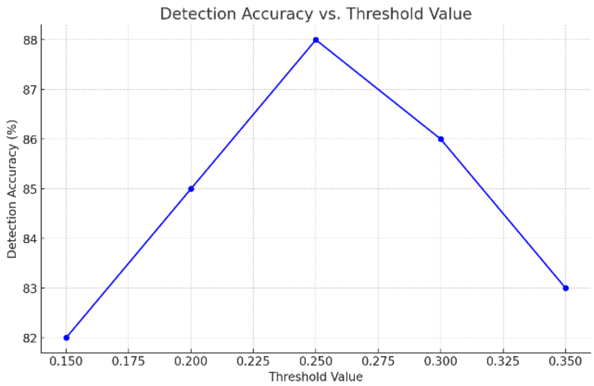

Figure 6. Detection accuracy vs threshold value.

Table 3 and **Figure 7** demonstrate how the accuracy of object detection can be impacted by the use of different approximation methods, especially when dealing with circumstances that are dynamic and fast-paced, such as matches for football. For instance, new techniques such as visual flow or motion-based approximation are able to record the motion of players and the ball's motion with higher precision. Consequently, this eventually results in enhanced detection performance in contrast to basic approximation or iteration that lacks any features in the first place. A contrast of both of these techniques serves as crucial for developing a knowledge of both the advantages and disadvantages related to each method when implemented in the environment of sports video analysis.

Table 3. Detection accuracy vs interpolation models.

Interpolation technique	Description	Detection accuracy (%)
Baseline	No interpolation, direct frame analysis	85%
Linear Interpolation	Simple linear interpolation between frames	87%
Optical Flow	It uses motion vectors to interpolate between frames	89%
Bi-directional	Combines forward and backward frame interpolation	90%
Advanced Motion-based	Advanced algorithm considering motion patterns	92%

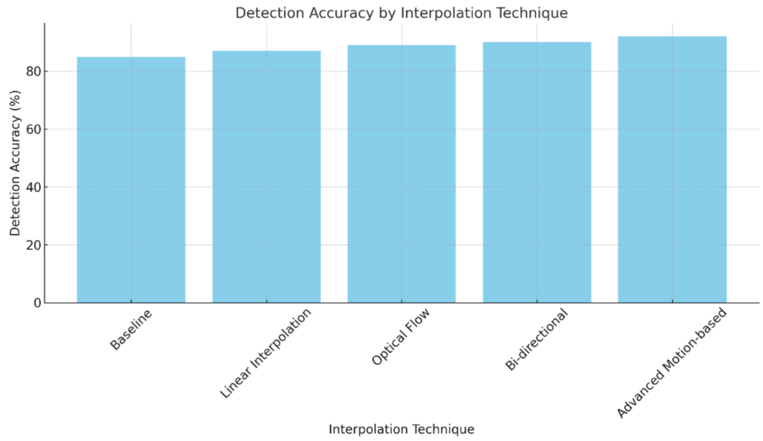

Figure 7. Detection accuracy vs Interpolation models.

Table 4 presents an easy-to-read summary of the results of the model in detecting significant components, such as participants and the ball's trajectory, as well as its accuracy in minimizing FP and FN. It can be concluded from the exceptionally high success percentages for player and ball detections that the framework is highly reliable for use in SA throughout football games. In setting object detection objectives, it is not unusual to encounter a few FP and FN, which help point out areas that need additional refinement of the framework.

Table 4. Performance analysis.

Validation metric	Automated detection (YOLOv5)	Manual annotation (ground truth)	Performance measure
Total frames analyzed	1000	1000	-
Player detections (TP)	4450	4550	-
Ball detections (TP)	980	1000	-
False positives (FP)	50	-	-
False negatives (FN)	70	-	-
Precision	Calculated as TP / (TP + FP)	-	98.7%
Recall	Calculated as TP / (TP + FN)	-	98.4%
Overall accuracy	Calculated as (TP + TN) / (TP + FP + FN + TN)	-	98.6%

3.3. Markov model analysis

Accuracy of state transitions: The precision of the MCM is assessed through the comparison of the predicted transitions between states from the mathematical model to the actual game video. In order to achieve this, a frame-by-frame investigation into the participants and the ball's motion is conducted, with each projected transition being contrasted to the actual motion that is occurring. It is feasible that anomalies are an indication that there is scope for enhancement, while a high level of precision suggests that the model accurately represents the unpredictable nature of the game being played.

The MCM's power to forecast state transitions in football games precisely is shown in **Table 5** and **Figure 8**, respectively. A mathematical model with a high degree of precision throughout sections suggests that it accurately represents the dynamics of the game, while an example with less than perfect accuracy in particular regions reveals that there is scope for development. This examination is essential to obtaining knowledge of the model's merits as well as the points at which it could be enhanced in SA.

Table 5. Accuracy for segment transitions.

Match segment	Total transitions observed	Transitions predicted correctly	Accuracy (%)
Opening play	50	42	84
Mid-Game	70	58	83
Attacking phase	40	35	88
Defensive phase	30	25	83
Set pieces	20	18	90
Closing play	50	43	86

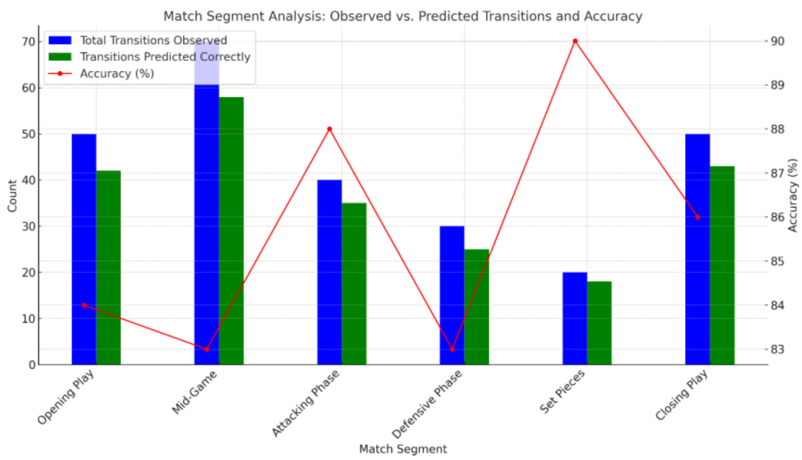

Figure 8. A contrast of state transitions and elements related to accuracy.

Consistency with known tactics: It is feasible for one to contrast the forecasts of the model concerning team behavior with the known tactics and appearances of the teams that were examined. The tactics mentioned above and styles have been obtained from match reports, expert analyses, and past data. Analysts evaluate if the model aligns with these tactics, such as a team's focus on wing play, resulting in higher transition probabilities in pitch zones. Results are presented in **Table 6** and **Figure 9**.

Table 6. Tactic analysis.

Team	Known tactical style	Model's predicted dominant zones	Consistency score (%)	Consistency evaluation
A	Wing play focus	High transitions in wing zones	90%	High
B	Central attacking strategy	High transitions in central zones	85%	High
C	Defensive, counter-attacking	Frequent transitions in defensive zones, quick shifts to attacking zones	70%	Moderate
D	Possession-Based play	Even distribution across all zones	50%	Low
E	High pressing style	High transitions in the opponent's half	88%	High
F	Long ball tactics	Frequent transitions from defensive to attacking zones	65%	Moderate

Figure 9. Team consistency analysis.

SSA: In the sport of football, SSA can perform. The SA tool is a successful tool for comprehending the long-term tactical behaviors of teams. It involves calculating the SSD π of an MCM, where $\pi P = \pi$ and P is the transition probability matrix. The findings of this analysis show the regions of the field where play dominates the most over an extended time. This analysis additionally gives insights into the areas of the game and tactics that a team especially prefers.

For the objective of this analysis, let's assume a playing surface for football that has been divided into six zones for the purposes of clarity and strategically significant data:

1) Zone 1: Left defensive wing
2) Zone 2: Central defensive area
3) Zone 3: Right defensive wing
4) Zone 4: Left attacking wing
5) Zone 5: Central attacking area
6) Zone 6: Right attacking wing

On average, a team will prioritize its play in these zones according to data provided by the SSD across these zones. As an outcome of greater chances of specific zones, it is probable that certain tactical decisions, such as focusing on core attacks or maintaining adequate security, will be decided. The data regarding the number of teams can be found in **Table 7** and **Figure 10**, respectively.

Table 7. Zone-wise state distribution for each team.

Team	SSD by zone (%)	Dominant play area	Tactical insight
A	Z1: 10%, Z2: 15%, Z3: 5%, Z4: 20%, Z5: 30%, Z6: 20%	Z5 (Central Attacking)	Prefers central attacking play
B	Z1: 15%, Z2: 30%, Z3: 15%, Z4: 10%, Z5: 5%, Z6: 25%	Z2 (Central Defensive)	Focuses on deep defense
C	Z1: 30%, Z2: 10%, Z3: 10%, Z4: 25%, Z5: 15%, Z6: 10%	Z1 (Left Defensive)	Strong left-wing play
D	Z1: 16%, Z2: 17%, Z3: 17%, Z4: 17%, Z5: 16%, Z6: 17%	Even Distribution	Balanced play across all areas
E	Z1: 15%, Z2: 10%, Z3: 5%, Z4: 30%, Z5: 20%, Z6: 20%	Z4 (Left Attacking)	Aggressive right-wing play
F	Z1: 5%, Z2: 20%, Z3: 25%, Z4: 10%, Z5: 25%, Z6: 15%	Z3 (Right Defensive)	Central attacking focus

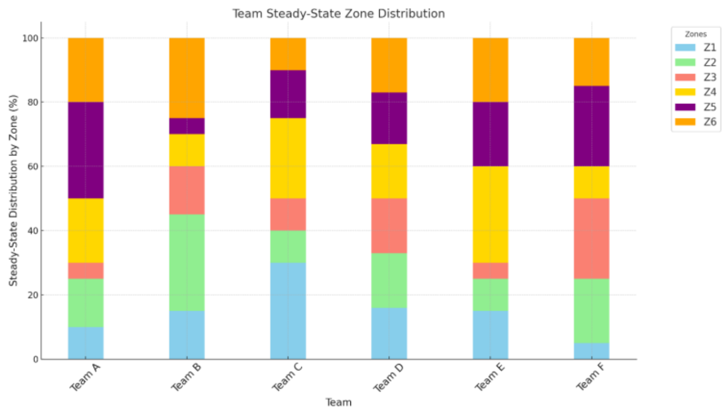

Figure 10. SSD of each team.

It is possible to gain an essential sense of each team's overall tactical strategy and to plan counter-strategies by pointing to **Figure 10**, which successfully highlights each team's strategic objectives and long-term tactical behavior on the field of play.

4. Conclusion and future work

Visual Recognition (VR) and Strategic Analysis (SA) algorithms have been effectively combined in this paper, demonstrating how this pairing can develop tactical analysis in football. It is the support for reliable data collection that the YOLOv5-VR system presents, as it accurately monitors the participants' and the ball's movement. On the basis of this data, the use of Markov Chain Models (MCM) and other SA provided a deep knowledge of the strategies and tactics applied by the team, reaching the volume of data and detail of predictable techniques. Adaptive algorithms and continuous models of learning have been involved in order to deal with challenges such as improving detection accuracy and managing an extensive number of in-game environments. Furthermore, the study verified the importance of using simple-to-use visualization software for the aim of data analysis. The results of this work contribute to the development of the field of sports analytics and provide coaches, analysts, and teams with vital techniques and models that can help them improve their tactical data and success on the sporting field.

On the future horizon, the chance of further integrating with AI and ML predictions even more sophisticated analytical capabilities in maximizing the effectiveness of sports tactics.

Conflict of interest

The author declares no conflict of interest.

References

1. Fernández de la Rosa J. A framework for the analytical and visual interpretation of complex spatiotemporal dynamics in soccer.
2. Mumford S. Football: the philosophy behind the game. John Wiley & Sons; 2019.

3. Li Y. When Moneyball Meets the Beautiful Game: A Predictive Analytics Approach to Exploring Key Drivers for Soccer Player Valuation.
4. Bauer MSP, Ingbert S. Automated Detection of Complex Tactical Patterns in Football.
5. Rønningen MH. The genesis of data-driven decision-making in the world of soccer tactics: deciphering the potential of big data [Master's thesis]. University of Agder; 2021.
6. Jocher G, Stoken A, Borovec J, et al. ultralytics/yolov5: v3. 0. Zenodo; 2020.
7. Suhr JK. Kanade-lucas-tomasi (KLT) feature tracker. Computer Vision (EEE6503); 2009.
8. Cioppa A, Giancola S, Deliege A, et al. SoccerNet-Tracking: Multiple Object Tracking Dataset and Benchmark in Soccer Videos. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Published online June 2022. doi: 10.1109/cvprw56347.2022.00393
9. Shaik AA, Mareedu VDP, Polurie VVK. Learning multiview deep features from skeletal sign language videos for recognition. Turkish Journal of Electrical Engineering & Computer Sciences. 2021; 29(2): 1061-1076. doi: 10.3906/elk-2005-57
10. Ghuge A, Prakash VC, Ruikar SD. Systematic analysis and review of video object retrieval techniques. Control and Cybernetics. 2020; 49(4): 471–498.
11. Ghuge CA, Chandra Prakash V, Ruikar SD. Weighed query-specific distance and hybrid NARX neural network for video object retrieval. The Computer Journal. 2019; 63(11): 1738-1755. doi: 10.1093/comjnl/bxz113
12. Victoria DrAH, Manikanthan SV, H R DrV, et al. Radar Based Activity Recognition using CNN-LSTM Network Architecture. International Journal of Communication Networks and Information Security (IJCNIS). 2023; 14(3): 303-312. doi: 10.17762/ijcnis.v14i3.5630
13. Shaik AA, Mareedu VDP, Polurie VVK. Learning multiview deep features from skeletal sign language videos for recognition. Turkish Journal of Electrical Engineering & Computer Sciences. 2021; 29(2): 1061-1076. doi: 10.3906/elk-2005-57
14. Appathurai A, Sundarasekar R, Raja C, et al. An Efficient Optimal Neural Network-Based Moving Vehicle Detection in Traffic Video Surveillance System. Circuits, Systems, and Signal Processing. 2019; 39(2): 734-756. doi: 10.1007/s00034-019-01224-9
15. Balamurugan D, Aravinth SS, Reddy PCS, et al. Multiview Objects Recognition Using Deep Learning-Based Wrap-CNN with Voting Scheme. Neural Processing Letters. 2022; 54(3): 1495-1521. doi: 10.1007/s11063-021-10679-4
16. Bhavana D, Kishore Kumar K, Bipin Chandra M, et al. Hand Sign Recognition using CNN. International Journal of Performability Engineering. 2021; 17(3): 314. doi: 10.23940/ijpe.21.03.p7.314321
17. Kumar EK, Kishore PVV, Kiran Kumar MT, et al. 3D sign language recognition with joint distance and angular coded color topographical descriptor on a 2 – stream CNN. Neurocomputing. 2020; 372: 40-54. doi: 10.1016/j.neucom.2019.09.059
18. Ghuge C, Prakash V, Ruikar S. An Integrated Approach Using Optimized Naive Bayes Classifier and Optical Flow Orientation for Video Object Retrieval. International Journal of Intelligent Engineering and Systems. 2021; 14(3): 210-221. doi: 10.22266/ijies2021.0630.19
19. Gullapelly A, Banik BG. Exploring the techniques for object detection, classification, and tracking in video surveillance for crowd analysis. Indian Journal of Computer Science and Engineering. 2020; 11(4): 321–326.
20. Saha J, Chowdhury C, Ghosh D, et al. A detailed human activity transition recognition framework for grossly labeled data from smartphone accelerometer. Multimedia Tools and Applications. 2020; 80(7): 9895-9916. doi: 10.1007/s11042-020-10046-w
21. Yadav J, Misra M, Rana NP, et al. Exploring the synergy between nano-influencers and sports community: behavior mapping through machine learning. Information Technology & People. 2021; 35(7): 1829-1854. doi: 10.1108/itp-03-2021-0219
22. Yadav J, Misra M, Rana NP, et al. Netizens' behavior towards a blockchain-based esports framework: a TPB and machine learning integrated approach. International Journal of Sports Marketing and Sponsorship. 2021; 23(4): 665-683. doi: 10.1108/ijsms-06-2021-0130
23. Mohan KK, Prasad CR, Kishore PVV. Yolo V2 with bifold skip: A deep learning model for video-based real-time train bogie part identification and defect detection. Journal of Engineering Science and Technology. 2021; 16(3): 2166–2190.
24. Krishnamohan K, Prasad ChR, Kishore PVV. Train rolling stock video segmentation and classification for bogie part inspection automation: a deep learning approach. Journal of Engineering and Applied Science. 2022; 69(1). doi: 10.1186/s44147-022-00128-x
25. Raju K, et al. A robust and accurate video watermarking system based on SVD hybridation for performance assessment. International Journal of Engineering Trends and Technology. 2020; 68(7): 19–24.
26. Suneetha M, Prasad MVD, Kishore PVV. Sharable and unshareable within class multi view deep metric latent feature learning for video-based sign language recognition. Multimedia Tools and Applications. 2022; 81(19): 27247-27273. doi: 10.1007/s11042-022-12646-0
27. Suneetha M, Prasad MVD, Kishore PVV. Multi-view motion modelled deep attention networks (M2DA-Net) for video-based sign language recognition. Journal of Visual Communication and Image Representation. 2021; 78.
28. Wagdarikar AMU, Senapati RK. A secure communication approach in OFDM using optimized interesting region-

based video watermarking. International Journal of Pervasive Computing and Communications. 2020; 18(2): 171-194. doi: 10.1108/ijpcc-05-2019-0044
29. Janarthanan P, Murugesh V, Sivakumar N, et al. An Efficient Face Detection and Recognition System Using RVJA and SCNN. Hemanth J, ed. Mathematical Problems in Engineering. 2022; 2022: 1-9. doi: 10.1155/2022/7117090
30. Priyadharshini, Gomathi T. Naive Bayes classifier for wireless capsule endoscopy video to detect bleeding frames. International Journal of Scientific and Technology Research. 2020; 9(1): 3286–3291.
31. Ali SA, Prasad MVD, Kishore PVV. Ranked Multi-View Skeletal Video-Based Sign Language Recognition With Triplet Loss Embeddings. Journal of Engineering Science and Technology. 2022; 17(6): 4367–4397.
32. Pande SD, Chetty MSR. Linear Bezier Curve Geometrical Feature Descriptor for Image Recognition. Recent Advances in Computer Science and Communications. 2020; 13(5): 930-941. doi: 10.2174/2213275912666190617155154
33. Depuru S, Nandam A, Ramesh PA, et al. Human Emotion Recognition System Using Deep Learning Technique. Journal of Pharmaceutical Negative Results. 2022; 13(4): 1031–1035.
34. Ali SKA, Prasad MVD, Kumar PP, et al. Deep Multi View Spatio Temporal Spectral Feature Embedding on Skeletal Sign Language Videos for Recognition. International Journal of Advanced Computer Science and Applications. 2022; 13(4). doi: 10.14569/ijacsa.2022.0130494
35. Rani S, Ghai D, Kumar S. Reconstruction of Simple and Complex Three Dimensional Images Using Pattern Recognition Algorithm. Journal of Information Technology Management. 2022; 14: 235–247.
36. Rani S, Ghai D, Kumar S, et al. Efficient 3D AlexNet Architecture for Object Recognition Using Syntactic Patterns from Medical Images. Computational Intelligence and Neuroscience. 2022.
37. Kotkar VA. Scalable Anomaly Detection Framework in Video Surveillance Using Keyframe Extraction and Machine Learning Algorithms. Journal of Advanced Research in Dynamical and Control Systems. 2020; 12(7): 395-408. doi: 10.5373/jardcs/v12i7/20202020
38. Li X, Manivannan P, Anand M. Task Modelling of Sports Event for Personalized Video Streaming Data in Augmentative and Alternative Communication. Journal of Interconnection Networks. 2022; 22(Supp01). doi: 10.1142/s0219265921410279