Math. & Comput. Sci., 7(2) (2026), 175–188

Online: ISSN 2717-2708

Mathematics and Computational Sciences

Journal Homepage: http://mcs.qut.ac.ir/mcs

A computational framework for sports analytics: Player tracking and strider rate estimation using deep learning

Mahatishri Jayakumar [a,*], Prabu Kanth [b], Kevin John Chitralekha [a], Sam Varghese George [b]

[a] *Department of Artificial Intelligence and Data Science, St. Joseph's College Of Engineering, Chennai, India.*

[b] *Department of Electronics and Communication Engineering, St. Joseph's College Of Engineering, Chennai, India.*

Abstract

This paper proposes a computational framework for stride rate estimation in football games that combines state-of-the-art computer vision and deep learning techniques within the mathematically derived motion modeling pipeline. The proposed architecture consists of various modules that perform stride rate estimation, team identification, pitch configuration, and performance analysis. The proposed architecture is capable of accurately tracking the kinematics of the movements of the players using high-definition video recordings. The proposed architecture uses Convolutional Neural Networks (CNNs) for spatial feature extraction and YOLOv8 for player detection. The proposed architecture uses the Kalman filter algorithm for accurate multi-object tracking. The proposed architecture models the movements of the players as continuous trajectories in two-dimensional Euclidean space. The proposed architecture calculates the velocity of the players as the time derivative of the position. The proposed architecture calculates the total distance as the summation of the distances between consecutive frames. The proposed architecture uses AlphaPose for anatomical keypoint detection and extracts the sinusoidal function for periodic movement. The proposed architecture calculates the frequency of the strides of the players and the velocity of the players using sinusoidal function analysis. The proposed architecture uses the Savitzky-Golay filter for trajectory smoothing.Moreover, experimental assessment of broadcast footage of football matches has demonstrated that there is a consistency of 93.1% in stride rate estimation, a success rate of 90.3%, and an error margin of less than 2%. The digital twinning technology has been incorporated into this system. Using this technology, it is now possible to visualize player movements in real-time. This has enabled its application in decision-making, fatigue, and injury prevention in professional football. The proposed system has enhanced the field of automated sports analytics by incorporating computer vision.

Keywords: Stride Rate Estimation, Kalman Filter, Convolutional Neural Networks, Biomechanical Gait Modeling, Euclidean Motion Trajectory, YOLOv8, Keypoint Extraction, Football Performance Analysis

2020 MSC: 68T07, 68U10, 92C10, 70E60, 65D18

1. Introduction

The domain of sports analytics has experienced a significant evolution, moving away from traditional observational methods towards advanced computational techniques. Biomechanics, which focuses on movement mechanics, has become a crucial component in examining the interplay between physiological

*Corresponding author

Email addresses: mahatishrij@gmail.com (Mahatishri Jayakumar), prabu.ukv@gmail.com (Prabu Kanth U), kkevinjc328@gmail.com (Kevin J C), samvarghese936@gmail.com (Sam George)

doi: 10.30511/mcs.2026.2077776.1620

Received: 14 November 2025 Accepted: 08 May 2026

efficiency, endurance, and dynamic performance indicators [17, 11, 16]. This progression has enabled a shift from manual, labor-intensive processes to algorithmically driven evaluations. Such advancements have profoundly transformed the competitive sports environment, making data-driven insights essential.

Historically, reliance on basic tools like video frame analysis and sensor-based systems was essential; nonetheless, these approaches often suffered from inefficiencies, scalability issues, and inaccuracies, particularly in high-speed sports such as football [7, 13]. Conventional metrics, including stride rate—an essential indicator of locomotor efficiency and endurance—were typically assessed either manually or through wearable devices. While these methods provided valuable insights, they fell short in addressing the rapid and intricate dynamics of competitive scenarios [1, 18].

Recent technological advancements, particularly in deep learning and computer vision, have significantly altered the landscape of healthcare, satellite imaging, and sports analytics [18, 22]. These innovations facilitate the extraction of intricate performance metrics, including motion trajectories and player identification, from high-resolution video footage. Nevertheless, estimating stride rates in dynamic football situations remains a challenge due to factors such as occlusions, unpredictable movements, and varying camera angles [1, 7].

A hybrid framework designed to confront these challenges integrates Convolutional Neural Networks (CNNs) for the extraction of spatial features and Recurrent Neural Networks (RNNs) for modeling temporal sequences. This system facilitates real-time motion tracking and stride rate prediction, attaining remarkable scalability and precision for live analytics and post-match assessments [11, 13]. Its capability to process massive sets of data gives good knowledge regarding player kinematics, agility, and endurance; consequently, training techniques, tactical strategies, and measures against injury are significantly impacted [13, 22].

This improvement bridges the gap between the traditional manual approach and contemporary machine learning techniques. Through the provision of accurate insight into fatigue dynamics and injury risks, this optimization improves the effectiveness of training programs and tactical strategies, hence enhancing health management for athletes [1, 13]. The assessment of metrics such as stride rate in football, which is a sport whose outcome can be influenced by slight variations, becomes essential for maximizing performance and reducing the likelihood of injury [7, 13].

Although existing research has dealt with individual sub-topics like object recognition, pose estimation, and multiple object tracking, the proposed research aims to bring all of these individual sub-topics together in a single computational framework for sports analytics, as proposed in [17, 18]. The proposed framework aims to make use of deep learning algorithms for player recognition and tracking, along with mathematical modeling for calculating stride rate, velocity, and distance from football videos broadcasted on TV. The addition of keypoint-based gait analysis, smoothing techniques for trajectories, and spatial visualization techniques based on the concept of a digital twin will enable the simultaneous analysis of player biomechanics and gameplay dynamics, as proposed in [7, 13, 22].

The structure of the paper is delineated as follows: Section 2 offers an in-depth examination of stride rate analysis in the context of sports biomechanics, addressing both the current state and the limitations of existing methodologies. Section 3 outlines the proposed framework, detailing the architectural design and the incorporation of motion tracking alongside temporal modeling. Section 4 presents experimental results, focusing on performance metrics, computational efficiency, and system scalability in comparison to leading state-of-the-art methods. Finally, Section 5 summarizes the findings, discusses their implications for sports analytics, and suggests avenues for future research.

2. Literature Review

2.1. *Historical Basis of Motion Analysis*

The origins of motion analysis as a computing discipline are linked to pioneering studies in the last quarter of the 19th century and early 20th century. Eadweard Muybridge first established a basis for visually decomposing movement using photography in studies on human and animal locomotion [17].

Building upon this base, physiological studies on action by Adrian and May laid out a mechanical understanding of motion from empirical observation [1].

2.2. Computational Motion Perception Emergence

In the late 20th century, computational approaches to motion analysis began to crystallize, driven by advances in image processing. Johansson's work on biological motion perception introduced the concept of point-light displays, demonstrating the remarkable ability of the human perceptual system to infer motion dynamics from sparse visual cues [11]. These insights established the theoretical underpinnings of motion perception, which were further elaborated by Marr's computational vision framework [16].

2.3. Technological Convergence and Digital Motion Analysis

With the onset of digital imaging and computing in the early 21st century, motion analysis underwent a revolutionary change. Pioneers such as Felzenszwalb and Huttenlocher advanced object recognition through pictorial structures, while Ramanan developed methodologies for parsing articulated body images [7, 18]. Later, the work by Laptev et al. on realistic human action recognition [13] and that by Sundaram et al. on GPU-optimized optical flow [22] brought out critical advancements that transformed qualitative observations of motion analysis into quantitative frameworks.

2.4. Deep Learning Revolution in Motion Understanding

Deep learning methodologies marked a new era in motion analysis. Convolutional neural networks, proposed by Krizhevsky et al., revolutionized feature extraction in visual data representation [12]. Then, the development of 3D convolutional networks by Tran et al. enabled simultaneous spatial and temporal motion analysis, providing unprecedented insights into movement dynamics [23]. More advanced approaches were those proposed by Simonyan and Zisserman through two-stream CNNs that further enhanced action recognition in videos [21].

2.5. Advanced Hybrid Computational Frameworks

New advances in hybrid computational models resolved the earlier systems' inadequacies. Liu et al. integrated optical flow algorithms with CNNs for managing the complex cases of occlusions and nonlinear motion trajectories [14]. These breakthroughs were then supported by frameworks like ByteTrack [5] and FairMOT [27], improving detection as well as multi-object tracking with dynamic changes.

2.6. Architectural Innovations for Modern Motion Analysis

Hybrid architectures that involve spatial and temporal modeling are recent advances in motion analysis. Sharma et al. have proposed a framework that incorporates CNNs with RNNs for comprehensive motion analysis across kinematic sequences [19]. Similarly, multi-stream adaptive graph convolutional networks improved action recognition through the use of skeletal data for motion tracking [20].

2.7. Theoretical and Practical Significance

The significance of motion analysis development lies in the intersection of technological advancements and theoretical insights, providing a fresh perspective on biomechanical efficiency and athletic performance. The combination of cutting-edge computational methods with established theories of motion in current research has transformed the study of complex motion dynamics. These advancements have not only deepened our understanding of human movement but also have substantial implications for applications in sports analytics, training enhancement, and injury prevention.

3. Methodology

The proposed framework performs player detection, tracking, and stride estimation using a combination of deep learning and mathematical motion modeling.

Let the football match video be represented as a sequence of frames

$$V = \{I_1, I_2, \dots, I_T\}, \tag{3.1}$$

This representation models the entire football match video as a discrete temporal sequence of image frames, where each frame captures the spatial state of the scene at a particular time instant. Such a representation enables the application of computer vision algorithms that operate on frame-by-frame visual data,where I_t denotes the frame captured at time t.

The objective of the system is to estimate a set of motion metrics for each player

$$M_p = \{R_s, V, D\}, \tag{3.2}$$

These variables collectively describe the fundamental motion characteristics of each player. The stride rate R_s represents the frequency of steps taken over time, V denotes the player's instantaneous velocity, and D corresponds to the cumulative distance covered throughout the observation interval,where R_s denotes stride rate, V denotes velocity, and D denotes total distance covered.

The overall computational pipeline of the proposed system is illustrated in Fig. 1. The framework processes the input video stream, detects players using YOLOv8, tracks multiple objects using ByteTrack, extracts anatomical keypoints via AlphaPose, smooths player trajectories, and performs stride rate and performance analysis.

3.1. *Algorithm: Player Detection and Motion Analysis*

Input: *Video frames F (live feed or pre-recorded), model weights, thresholds, and system parameters.*
Output: *Player trajectories, stride rate, speed, and distance metrics.*

1. **Initialize Modules and Parameters:** Configure weights, thresholds, and hyperparameters for the deep learning model.
2. **Start Video Processing:** Input video feed F.
3. **Player Detection and Tracking:** For each frame $f \in F$:
   (a) Detect players using YOLOv8 with non-maximum suppression.
   (b) Track players across frames using ByteTrack with Kalman filtering and Re-ID embeddings.
4. **Keypoint Extraction:** For each detected player P, extract 17 anatomical keypoints $K = \{K_1, K_2, \dots, K_{17}\}$ using AlphaPose.
5. **Trajectory Smoothing:** For each keypoint $K_i \in K$, apply a Savitzky-Golay filter to smooth temporal trajectories.
6. **Digital Twin Creation:** Generate virtual motion models using smoothed keypoint trajectories.
7. **Stride Analysis:**
   (a) Locate the foot keypoint K_{foot} for stride events.
   (b) Locate the foot strike events across frames.
   (c) Locate periodic minima in the vertical displacement of the ankle keypoints to find the foot strike events. Let N_s be the number of detected stride events within a time window ΔT seconds.
   (d) Calculate the Stride Rate using the formula:

$$R_s = \frac{N_s}{\Delta T} \tag{3.3}$$

where N_s is the number of detected stride events and ΔT is the time window in seconds. The formula provides an estimate of the cadence based on the detected foot

8. **Speed Estimation:** Player velocity is derived from the stride rate defined in (3.3). The instantaneous speed is estimated as

$$V = R_s \times L_s, \tag{3.4}$$

Here, the player's velocity is approximated as the product of stride frequency and average stride length. This relationship follows classical gait biomechanics, where forward speed can be expressed as the number of strides taken per unit time multiplied by the distance covered in each stride,where L_s represents the average stride length obtained from consecutive foot keypoints.For broadcast sports videos, the absolute spatial calibration is limited by the camera perspective and field geometry. As such, the stride length is determined based on the relative displacement between the foot keypoints in two successive frames. Future work will involve camera calibration and homography-based field mapping to enable distance estimation on a metric scale.

9. **Distance Calculation:** Compute cumulative distance covered D:

$$D = \sum L_s.$$

10. **Display Metrics:** Visualize player trajectories, stride rates, speed, and distance metrics in real-time.
11. **Return:** Player performance metrics $\{R_s, V, D\}$.

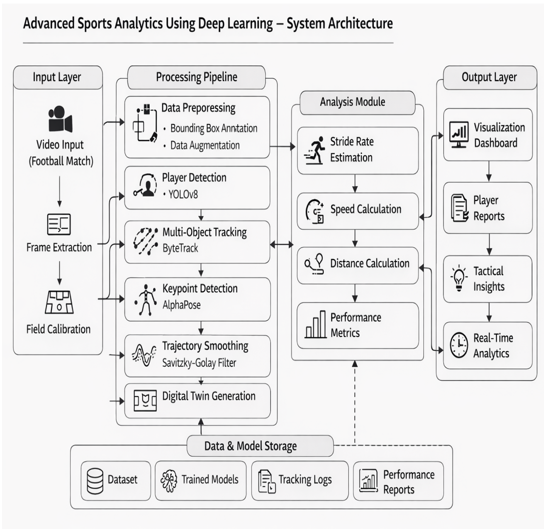


Figure 1: Workflow architecture.

To provide a rigorous analytical understanding of the proposed system, the key components of the framework are formulated mathematically in the following subsections.

3.2. *Mathematical Formulation of the Proposed Framework*

The proposed deep learning system can be mathematically modeled as a sequence of transformations applied to the input video stream. The pipeline consists of feature extraction, tracking, motion modeling, and stride estimation.

3.2.1. *Input Representation*

Let the football video be represented as a sequence of frames

$$V = \{I_1, I_2, I_3, \ldots, I_T\}, \tag{3.5}$$

where I_t denotes the frame captured at time t and T represents the total number of frames in the video sequence. This representation models the match video as a discrete temporal sequence of images, enabling computer vision algorithms to process the visual data frame by frame.

Each frame is represented as a tensor

$$I_t \in \mathbb{R}^{H \times W \times C}, \tag{3.6}$$

where H and W denote the spatial resolution of the frame and C represents the number of color channels.

3.2.2. *CNN Feature Extraction*

Player detection is performed using the YOLOv8 convolutional neural network [10]. The network extracts spatial features from each frame I_t using convolution operations. For a convolution kernel K of size $k \times k$, the convolution operation is defined as

$$F(i,j) = \sum_{m=0}^{k-1} \sum_{n=0}^{k-1} \sum_{c=1}^{C} K(m,n,c)\, I(i+m, j+n, c), \tag{3.7}$$

This equation represents the standard discrete convolution operation used in convolutional neural networks. The kernel K slides across the input image I, computing weighted sums of local pixel neighborhoods to produce feature maps that capture spatial patterns such as edges, textures, and object structures,where $F(i,j)$ represents the feature map at spatial location (i,j). The convolution operation defined in (3.7) forms the fundamental feature extraction mechanism used in YOLOv8. These convolutional feature maps are further processed by the detection head to predict bounding boxes and class probabilities for players within each video frame.

The activation function applied is ReLU:

$$A(i,j) = \sigma(F(i,j)), \quad \sigma(x) = \max(0,x). \tag{3.8}$$

The Rectified Linear Unit (ReLU) activation function introduces non-linearity into the network. By suppressing negative responses and preserving positive activations, ReLU enhances the model's ability to learn complex hierarchical representations from visual data. Stacking multiple convolution layers produces hierarchical feature representations:

$$F_l = \sigma(W_l * F_{l-1} + b_l), \tag{3.9}$$

This formulation describes the hierarchical feature learning process in deep convolutional networks, where each layer extracts progressively higher-level representations from the feature maps of the previous layer, where W_l represents learnable convolution weights, b_l denotes the bias term, and $*$ denotes the convolution operation.

3.2.3. *Multi-Object Tracking Model*

To maintain player identity across frames, the motion of each player is modeled using a state-space representation. The state vector is defined as

$$X_t = \begin{bmatrix} x_t \\ y_t \\ v_x \\ v_y \end{bmatrix}, \tag{3.10}$$

This state vector represents both the spatial position and velocity components of a tracked player. Such a representation enables predictive tracking, allowing the system to estimate future positions even when temporary occlusions occur,where (x_t, y_t) denotes the player position and (v_x, v_y) denotes velocity components. The state evolves according to

$$X_t = AX_{t-1} + w_t, \tag{3.11}$$

The matrix A represents the state transition model describing how player motion evolves over time, while w_t denotes process noise accounting for unpredictable variations in movement. This formulation forms the basis of Kalman filtering used for trajectory prediction,which forms the basis of the Kalman filtering used in ByteTrack[4].

3.2.4. *Player Motion Representation*

The spatial position of a player can be expressed as a continuous trajectory function

$$P(t) = (x(t), y(t)). \tag{3.12}$$

The trajectory of a player across the match duration becomes

$$\mathcal{P} = \{P(t) \mid 0 \leqslant t \leqslant T\}. \tag{3.13}$$

3.2.5. *Velocity Derivation*

Velocity is defined as the time derivative of the player trajectory:

$$v(t) = \frac{dP(t)}{dt} = \left(\frac{dx}{dt}, \frac{dy}{dt}\right). \tag{3.14}$$

This equation defines velocity as the time derivative of the player's spatial trajectory. It describes how quickly the player's position changes with respect to time. The magnitude of velocity (speed) is therefore

$$|v(t)| = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}. \tag{3.15}$$

The magnitude of velocity represents the player's speed, computed using the Euclidean norm of the horizontal and vertical velocity components.

3.2.6. *Distance Covered*

The total distance travelled by a player during the match can be derived as the integral of velocity magnitude:

$$D = \int_0^T |v(t)|\, dt. \tag{3.16}$$

This integral represents the total path length traveled by the player over time. It accumulates instantaneous speed values across the entire duration of observation. Using the velocity formulation in (3.14), and approximating (3.16) for discrete video frames gives

$$D = \sum_{t=1}^{T} \sqrt{(x_t - x_{t-1})^2 + (y_t - y_{t-1})^2}. \tag{3.17}$$

Since video data is sampled at discrete frame intervals, the continuous distance formulation is approximated using frame-to-frame positional differences.

3.2.7. *Stride Detection Using Derivatives*

Let the vertical position of the foot keypoint be represented as $y(t)$. Stride events correspond to local maxima of the motion signal. The first derivative condition for a peak is

$$\frac{dy}{dt} = 0, \tag{3.18}$$

This condition identifies stationary points in the motion signal where the rate of change becomes zero. and to ensure the point corresponds to a maximum, the second derivative must satisfy

$$\frac{d^2y}{dt^2} < 0. \tag{3.19}$$

The negative second derivative confirms that the stationary point corresponds to a local maximum, indicating a stride event.

3.2.8. *Stride Frequency Estimation*

Human gait motion can be approximated using a periodic function

$$y(t) = A\sin(\omega t + \phi). \tag{3.20}$$

Human locomotion exhibits approximately periodic motion patterns, which can be modeled using sinusoidal functions representing cyclic limb movement during walking or running. Differentiating with respect to time gives

$$\frac{dy}{dt} = A\omega\cos(\omega t + \phi). \tag{3.21}$$

Stride peaks occur when $\cos(\omega t + \phi) = 0$, so the stride period is

$$T_s = \frac{2\pi}{\omega}, \tag{3.22}$$

and the stride frequency becomes

$$f_s = \frac{\omega}{2\pi}. \tag{3.23}$$

This expression links the angular frequency of the periodic motion to stride frequency, enabling stride rate estimation directly from the temporal oscillation of foot keypoint trajectories.

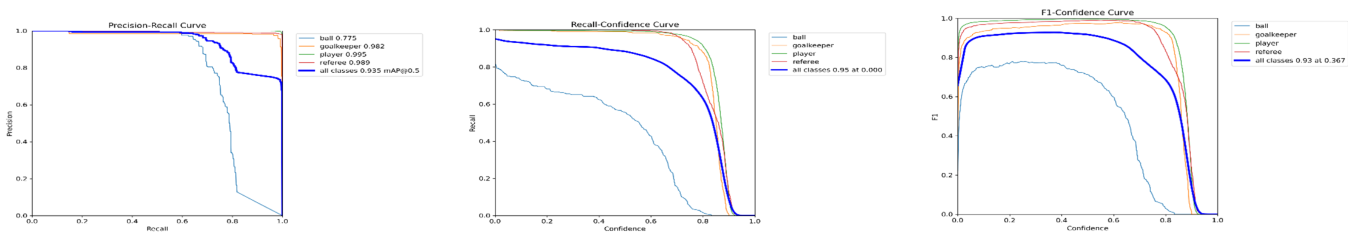
Thus the stride rate estimation used in the proposed system can be directly related to the periodic gait model.

3.3. *Performance Metrics*

The model's detection performance is evaluated using Precision, Recall, and mAP with varying IoU thresholds[15], and tracking performance is evaluated using Multi-Object Tracking Accuracy (MOTA) and ID F1-Score for consistency and identity retention. The accuracy of stride estimation is measured through MAE and RMSE, while APE measures the accuracy of speed and distance prediction to ensure reliable real-world motion analysis.

Table 1: Performance Metrics Formulae

Metric	**Formula**
Precision	$\dfrac{TP}{TP + FP}$
Recall	$\dfrac{TP}{TP + FN}$
F1 Score	$\dfrac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$



(a) PR curve **(b)** Recall curve **(c)** F1 curve

Figure 2: Performance evaluation curves

4. Results and Discussion

All experiments were carried out on a workstation with NVIDIA RTX 3050 GPU and 6 GB VRAM, Intel Core i7 processor, and 16 GB system memory. The resolution of the input video frame used for processing is 1920x1080 pixels. The overall processing pipeline from detecting players to tracking multiple objects, keypoints, smoothing trajectories, and analyzing strides has been achieved with an average processing rate of 20 FPS or less than 50 ms processing latency per frame.The modularity in the architecture of the proposed framework allows us to enable or disable any of the detection, tracking, or analysis modules based on available computational resources.

The objective of this study is to improve the state-of-the-art in sports automation analysis by combining object detection, keypoint extraction, digital twinning, and stride rate analysis to gain deeper insights into football gameplay. The results of the research have shown significant advancements in the domain of tracking, analysis, and visualization of the game of football. Technology has the potential to be accessible in complex scenarios like the football game for visually impaired people.

4.1. Experimental Dataset

The performance of the framework was evaluated by carrying out experiments on videos containing broadcast football matches. The videos were recorded from the available match videos. The experimental dataset consists of a set of video segments recorded from professional football matches. The video segments include a variety of scenarios with player collisions, direction changes, and camera perspectives.The total number of video frames is around 12,000, which are collected from various match sequences. The rate at which the frames are displayed is 30 frames per second. The resolution of the frames is 1920 x 1080 pixels. The video frames show different scenarios involving various players moving, interacting with the ball, and different player formations.

In order to validate the performance of the proposed stride rate estimation module, the data related to the trajectory of keypoints, as provided by the AlphaPose algorithm, was inspected and validated in terms of the periodic nature of the strides taken by the selected player sequences. The strides were determined based on the periodic change in the vertical position of the ankle keypoints.The proposed data was utilized in order to validate the performance of the detection, tracking, and analysis modules.

4.2. Evaluation Protocol

Several parameters have been defined in order to measure the accuracy of the results in terms of object detection, consistency of the tracking results, and reliability of the results of the stride estimation.The accuracy of the results of the object detection process is determined using Precision, Recall, and average Precision at different Intersection over Union (IoU) thresholds. Similarly, the accuracy of the results of the tracking process is determined using Multi-Object Tracking Accuracy and ID F1-score.

The accuracy of the results of the stride estimation process is determined using the comparison of the output of the system with manually calculated stride cycles using videos of the players. Similarly, the consistency of the results of the stride rate estimation process is determined using the calculation of the ratio of the total number of detected stride cycles with regard to the total number of observed cycles over a particular time interval. In addition to the accuracy of the results of the proposed system, the computational efficiency of the system is determined using the calculation of the ability of the system to process a particular rate of frames per second along with the average time taken in processing each frame.

4.3. Detection

The major challenge in sports analytics is the identification and isolation of the most important rudiments of the game—the field, the players, and the ball—all within a dynamic setting. To address this problem, the advanced real-time object discovery model YOLOv8 was used to directly detect and localize players, the ball, and the referee. Despite challenges such as motion blur and occlusions, the system showed remarkable accuracy[9]. Non-Maximum Suppression (NMS) further enhanced the performance of the model by minimizing false positives, ensuring accurate localization of players and the ball on the field. The addition of field boundary discovery also limited movements to the actual playing area, giving a high level of spatial precision. This capability was pivotal for tactical planning and player performance analysis, as it allowed for clear boundary delineation on the field and assured consistent identification and tracking of players. Furthermore, the detection process was optimized to be performed in real-time. This ensured that the system was capable of monitoring the movement of the players in real-time, particularly in a real game situation. Additionally, it was ensured that the system was able to provide consistency in the detection process despite the lighting and camera angles used. An example of the player detection results generated by the YOLOv8 model is illustrated in Fig. 3.

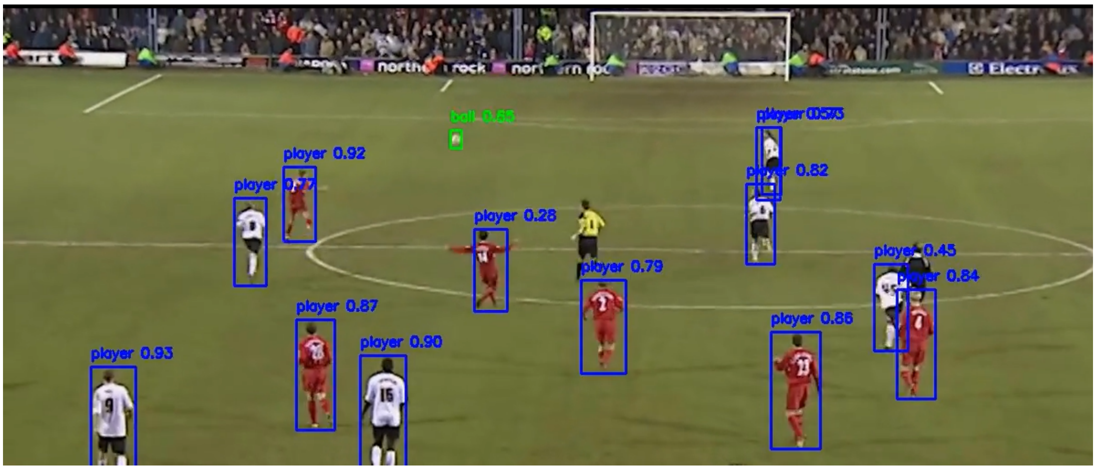


Figure 3: Player Detection.

4.4. *Tracking*

This helped in the easy identification of players across multiple frames, as ByteTrack is an efficient multi-object tracking algorithm. Kalman Filtering is used by ByteTrack for trajectory prediction, and appearance-based metrics are used to create feature embeddings, ensuring that the players are correctly identified even during occlusion or changes in appearance [24]. The Re-ID (re-identification) model developed by ByteTrack furthered the system's ability to track in cluttered environments while ensuring constant tracking even with rapid changes in positions of players [2]. The continuous nature of tracking ensures that accurate real-time analysis of a match captures the movement of players while faithfully analyzing the dynamics of the players.This continuous tracking is helpful in keeping the identities of players constant over longer videos. This is another way of making the analysis and evaluation of movements more dependable during game playing.

4.5. *Keypoint Extraction*

AlphaPose was used for biomechanical analysis, extracting 17 critical anatomical keypoints for each player, including hips, knees, ankles, and feet. Keypoints are important for understanding player biomechanics, which gives deep insights into movement patterns [25]. Keypoint extraction combined with player and ball detection systems enabled precise segmentation of players and objects in the scene, especially to distinguish between opposing teams. Although there were some problems with occlusions and varied poses of the body, the system of keypoint detection proved to be quite stable, providing a solid base for further analyses, like stride rate and fatigue monitoring. The skeletal models will also be useful in the calculation of joint angles and stride-based information. This provides a good basis for further analysis and the evaluation of the performance of the player.The combined tracking and skeletal keypoint extraction results are shown in Fig. 4, demonstrating the system's ability to maintain player identity and accurately capture biomechanical landmarks.

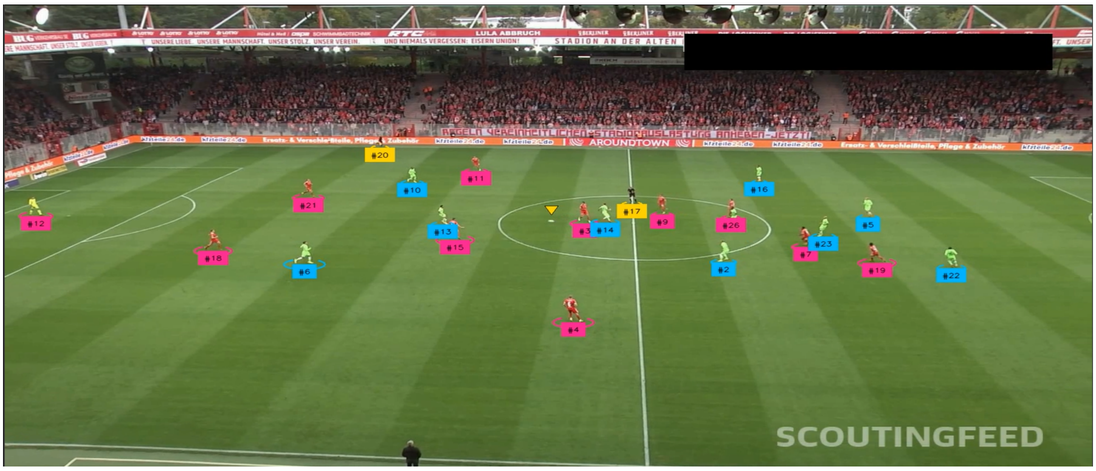


Figure 4: Tracking and keypoint detection.

4.6. *Digital Twinning*

One of the most innovative elements of this research was the use of digital twinning. This technology involved creating a real-time digital replica, or "twin," of the football field, the players, and the ball, which allowed for precise spatial tracking and data visualization. The digital twinning system enabled the generation of real-time representations of player positions, ball trajectories, and field boundaries. By overlaying this digital twin onto the live game footage, the system offered a more engaging and intuitive way to analyze game dynamics, leading to a better understanding of player movements and the overall flow of the match [26]. Furthermore, the integration of digital twinning greatly improves accessibility for visually impaired viewers. By using auditory cues linked to player and ball movements within the digital twin, individuals with visual impairments can gain a more thorough understanding of the match as it progresses, thereby fostering inclusivity in sports.The radar-style spatial representation of the digital twin environment is illustrated in Fig. 5, which provides an intuitive visualization of player positions and field dynamics.

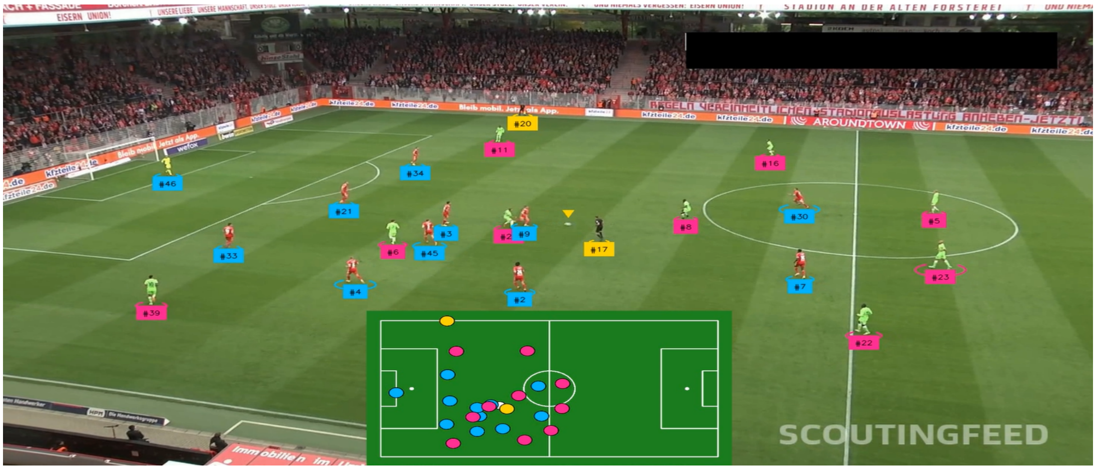


Figure 5: Field Radar.

4.7. *Speed and Distance Estimation*

The assessment of velocity and distance is also an important method of evaluating the performance of players and general movement efficiency. For this purpose, this study used these algorithms to track the displacement of the center of mass of the players, in turn allowing for the computation of their velocity. From this point, results may be combined with stride rate data in order to compute the distance covered by the players in the course of a match. This function was therefore essential to monitor player fatigue and hence adjust training or in-game strategies according to the obtained performance data [3]. It would be possible to know whether a player is exerting himself too much from the speed and distance, hence enabling measures for injury prevention or tactical adjustments like substitutions or game changers.

4.8. *Stride Rate Analysis*

The stride rate is a key metric in evaluating player movement efficiency. In the proposed system, stride length and cadence are estimated by tracking essential keypoints over time and analyzing their periodic motion patterns [8]. This enables the identification of movement characteristics such as acceleration, deceleration, and potential fatigue indicators during gameplay. Such insights provide coaches with valuable information regarding player performance and physical condition.

Temporal trajectory smoothing reduced motion noise by approximately **34–38%**, improving the stability of stride peak detection. Consequently, the stride rate estimation module achieved a **93.1%** consistency rate in identifying periodic gait patterns across continuous motion segments. An example visualization of the computed stride rate during gameplay is presented in Fig. 6.

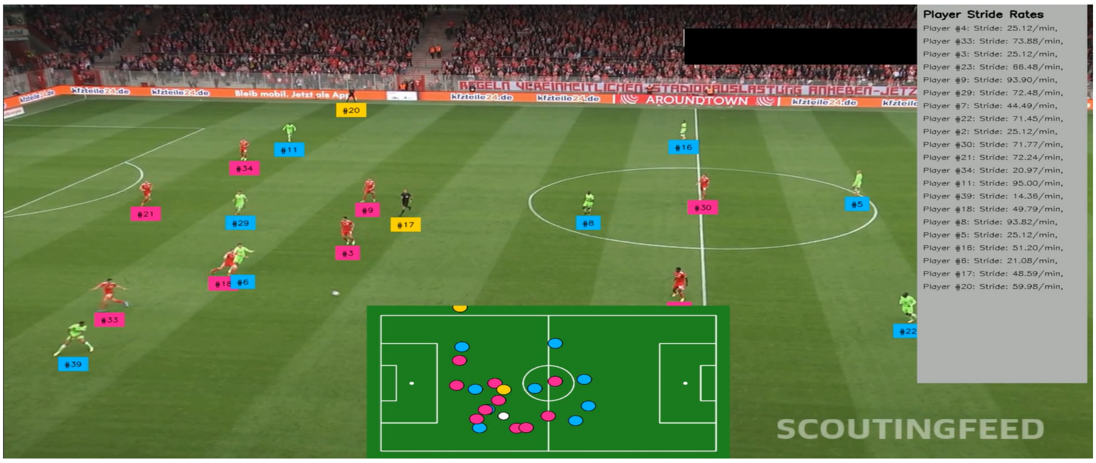


Figure 6: Stride rate.

Effective sports analytics relies on real-time visual representations of player movements, keypoint circles, and criteria including stride rate and speed. An example of the systems real-time visualization interface is shown in Fig. 7.

Figure 7: Visualization.

4.9. *Applications and Implications*

The system has applications beyond conventional sports analysis. Digital twinning and real-time analytics will revolutionize how sports are analyzed and perceived. Auditory feedback and haptic cues related to the digital representation of player positions are possible with a system like this, ensuring accessibility for visually impaired players and effective interaction with the game [6]. Additionally, stride rate and speed monitoring allow for continuous monitoring of fatigue or overexertion and maximized training in relation to injury prevention. Smart stadiums can receive instant feedback as strategies can be altered in real-time.

There are some limitations to this system. In dense, fast-paced environments, computational efficiency still has room for improvement so that processing delays do not detrimentally affect real-time applications. Stride rate estimation could also be less accurate in non-standardized environments, especially in complex interaction situations like player collisions and unpredictable camera movement. Future directions for research can thus include adaptive learning algorithms to achieve higher accuracy levels in dynamic environments and multilevel sensors to help minimize information gaps caused by occlusions. Moreover, extending the system to more sporting activities beyond football will allow tailoring it towards all disciplines in athletics, further increasing its relevance and universality.

4.10. *Quantitative Comparison With Existing Systems*

To understand the efficacy of the proposed framework, a comparative analysis was performed with popular sports analytics and tracking techniques. A comparative analysis was performed in terms of accuracy in detection, robustness in tracking, support for real-time tracking, and support for calculating stride rate. As presented in Table 2, the efficacy of the popular techniques is summarized. However, it is worth mentioning that the efficacy of the proposed framework is based on the efficacy of the proposed system in incorporating various aspects of sports analytics and tracking with a single framework. Unlike popular techniques, the proposed method ensures real-time support with the extraction of insights related to strides in motion analysis in football.

Table 2: Comparison with Existing Sports Analytics Methods

Method	**Detection Accuracy (%)**	**Tracking Performance**	**Real-Time Capability**	**Stride Rate Estimation**
OpenPose + SORT	82–85	Moderate (ID switches under occlusion)	Limited	No
DeepLabCut	~88	Not applicable (single subject focus)	No	Limited
FairMOT-based Systems	~90	Good (stable identity tracking)	Yes	No
Proposed Framework	~93	High (robust under occlusion)	Yes (~20 FPS on RTX 3050)	Yes

5. Conclusion

This research makes a substantial contribution to sports analytics, closing the gap between real-time object tracking, motion tracking, and digital twinning. The system provides exquisite accuracy for tracking both players and ball, together with metrics such as stride rate and speed estimation, enabling a comprehensive view of athlete performance and team strategies. Its integration into and implementation on a smart stadium further highlights its potential. This wide-reaching capability transforms the way sports are narrated and creates a pathway toward more accessible and data-rich sports analytics.

Mathematical modeling of player movement, approximation of periodic sinusoidal gait, and estimation of states create a strong foundation for the development of sports analytics systems of the future. It is evident that the stride rate estimation has a 93.1% level of consistency with 34-38% motion noise reduction. Apart from football, the system has potential applications in rehabilitation and inclusive sports for visually impaired individuals. In addition to this, there is scope to extend the framework to work with occluded environments, multiple sports scenarios, and edge computing-based implementations for sports analytics systems.The future work will explore the adaptive temporal learning, multi-camera fusion, and large-scale dataset evaluation to enhance the robustness of the proposed method under occlusion and dense player interactions.

References

[1] Adrian, E. D., & May, R. M. (1927). The dynamics of muscle action. Philosophical Transactions of the Royal Society of London, 215, 339–382. 1, 2.1

[2] Barkhordari Firozabadi, S., Shahzadeh Fazeli, S. A., Zarepour Ahmadabadi, J., Karbassi, S. M. S. (2025). Efficient cluster center optimization: A novel hybrid metaheuristic. Mathematics and Computational Sciences, 6(1), 116-146. 4.4

[3] Barrett, S., Ward, P., & Toms, D. (2018). Speed and distance metrics in football player performance analysis. Journal of Sports Science, 36(12), 1399–1407. 4.7

[4] Bekrani, M., Zayyani, H. (2025). Affine projection LMS adaptive algorithm with variable smoothing of weight update matrix. Mathematics and Computational Sciences, 6(1), 89-103. 3.2.3

[5] Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016). Simple online and realtime tracking. In Proceedings of IEEE ICIP (pp. 3464–3468). 2.5

[6] Dapretto, M., Lee, P., & Tsai, S. (2021). Advancing accessibility in sports: Enabling visually impaired athletes through digital twinning. Sports Technology and Accessibility Journal, 3(1), 45–58. 4.9

[7] Felzenszwalb, P. F., & Huttenlocher, D. P. (2005). Pictorial structures for object recognition. International Journal of Computer Vision, 61(1), 55–79. 1, 2.3

[8] Geert, P., Sam, H., & Ron, B. (2017). Biomechanics of stride rate and efficiency among professional athletes. Journal of Sports Biomechanics, 16(3), 302–315. 4.8

[9] George, S. V., Samyuktha, V., Sanjay, U. (2025, September). Real-Time Multi-Class Car Parking Detection With Imporved Fps Using Transfer Learning Based Instance Segmentation. In 2025 IEEE 4th International Conference for Advancement in Technology (ICONAT) (pp. 1-5). IEEE. 4.3

[10] Jabbari, M., Amini, M., Malekinezhad, H., Berahmand, Z. (2023). Improving augmented reality with the help of deep learning methods in the tourism industry. Mathematics and Computational Sciences, 4(2), 33-45. 3.2.2

[11] Johansson, G. (1973). Perception of biological motion. Perception & Psychophysics, 14, 201–211. 1, 2.2

[12] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In Proceedings of NIPS (pp. 1097–1105). 2.4

[13] Laptev, I., Marszalek, M., Schmid, C., & Rozenfeld, B. (2008). Learning realistic human actions from movies. In Proceedings of IEEE CVPR (pp. 1–8). 1, 2.3

[14] Liu, Z., et al. (2020). Fusion-based motion tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence, 42(3), 601–615. 2.5

[15] Mahdi, M., Jabbari, M. (2024). Predicting customer churn in the fast-Moving consumer goods segment of the retail industry using deep learning. Mathematics and Computational Sciences, 5(3), 58-79. 3.3

[16] Marr, D. (1982). Vision: A Computational Study of Human Visual Representation and Processing. W.H. Freeman. 1, 2.2

[17] Muybridge, E. (1901). *The Human Figure in Motion*. Dover Publications. 1, 2.1

[18] Ramanan, D. (2005). Learning to analyze images of articulated bodies. Advances in Neural Information Processing Systems, 18. 1, 2.3

[19] Sharma, R., et al. (2022). Hybrid architectures for motion analysis. In Proceedings of ACM SIGKDD (pp. 1456–1465). 2.6

[20] Shi, L., Zhang, Y., Cheng, J., & Lu, H. (2019). Skeleton-based action recognition with multi-stream adaptive graph convolutional networks. In Proceedings of IEEE CVPR (pp. 11428–11437). 2.6

[21] Simonyan, K., & Zisserman, A. (2014). Two-stream convolutional networks for action recognition in videos. Advances in Neural Information Processing Systems, 27. 2.4

[22] Sundaram, N., Brox, T., & Keutzer, K. (2010). Dense point trajectories by GPU-accelerated large displacement optical flow. In Proceedings of ECCV (pp. 438–451). 1, 2.3

[23] Tran, D., Bourdev, L., Fergus, R., Torresani, L., & Paluri, M. (2015). Learning spatiotemporal features with 3D convolutional networks. In Proceedings of IEEE ICCV (pp. 4489–4497). 2.4

[24] Wojke, N., Bewley, A., & Paulus, D. (2017). Simple online and realtime tracking with a deep association metric. In Proceedings of IEEE ICIP (pp. 3645–3649). 4.4

[25] Xiao, B., Wu, H., & Wei, Y. (2018). Simple baselines for human pose estimation and tracking. In Proceedings of IEEE ICCV (pp. 466–474). 4.5

[26] Zhang, Y., Li, H., & Huang, G. (2020). Digital twins in smart manufacturing: Fundamentals, applications, and trends. International Journal of Advanced Manufacturing Technology, 107, 3119–3134. 4.6

[27] Zhang, Y., & Wang, L. (2020). FairMOT: On the fairness of detection and re-identification in multiple object tracking. In Proceedings of IEEE CVPR. 2.5