AI-Driven Image Recognition System for Automated Offside and Foul Detection in Football Matches Using Computer Vision

Qianwei Zhang[1], Lirong Yu*[2], WenKe Yan[3]
Chengdu Sport University, Chengdu, Sichuan, 610041, China[1]
School of Physical Education, Sichuan University, Chengdu, Sichuan 610041, China[2]
Sichuan University High School (No. 12 High School of Chengdu), Chengdu, Sichuan 610061, China[3]

Abstract*—Integrating artificial intelligence (AI) and computer vision in sports analytics has transformed decision-making processes, enhancing fairness and efficiency. This paper proposes a novel AI-driven image recognition system for automatically detecting offside and foul events in football matches. Unlike conventional methods, which rely heavily on manual intervention or traditional image processing techniques, our approach utilizes a hybrid deep learning model that combines advanced object tracking with motion analysis to deliver real-time, precise event detection. The system employs a robust, self-learning algorithm that leverages spatiotemporal features from match footage to track player movements and ball dynamics. By analyzing the continuous flow of video data, the model detects offside positions and identifies foul types such as tackles, handballs, and dangerous play—through a dynamic pattern recognition process. This multi-tiered approach overcomes traditional methods' limitations by accurately identifying critical events with minimal latency, even in complex, high-speed scenarios. In experiments conducted on diverse datasets of live match footage, the system achieved an overall accuracy of 99.85% for offside detection and 98.56% for foul identification, with precision rates of 98.32% and 97.12%, respectively. The system's recall rates of 97.45% for offside detection and 96.85% for foul recognition demonstrate its reliability in real-world applications. It's clear from these results that the proposed framework can automate and greatly enhance the accuracy of match analysis, making it a useful tool for both referees and broadcasters. The system's low computational overhead and growing ability make connecting to existing match broadcasting infrastructure easy. This establishes an immediate feedback loop for use during live games. This work marks a significant step forward in applying AI and computer vision for sports, introducing a powerful method to enhance the objectivity and precision of officiating in football.

Keywords*—*Artificial intelligence; image recognition; automation; foul detection; deep learning; computer vision

I. Introduction

Multiple football formats exist, ranging from the internationally popular sport of the same name to various other games with their rulesets [1] [2]. Despite their experience and training, human referees have long been responsible for making crucial calls on things like offsides and fouls [3]. However, they can still make mistakes, particularly when speed is vital. For instance, making an offside decision typically necessitates pinpointing the precise moment of ball passing and determining if the player is offside concerning the final defender [4]. Similarly granular is foul detection, particularly tackling and handball, which relies on a swift, frequently subjective determination. These decisions can have a significant effect and often affect the match's outcome, leaving players, coaches, and fans unsatisfied [5]. Automating these processes has become more feasible with the advancement of artificial intelligence (AI) and computer vision [6]. Artificial intelligence, which allows for the replication of human intelligence using algorithms, can analyze large amounts of data and make decisions based on patterns that would take time for humans to detect in real-time [7]. Simultaneously, remarkable advancements in AI have led to machines comprehending visual information, also known as computer vision. Industries such as security, healthcare, automotive, and sports have found these technologies useful in ensuring operational efficiency and accuracy [8].

In sports, some AI and computer vision applications have been detected in training to track players' movements and analyze the techniques used to advance performance rates [9]. Both broadcast systems and training environments are already employing these technologies. Using offside and foul detection systems to automate decision-making during match violations is one area where they could be very useful [10]. Today, there are technologies like Video Assistant Referee (VAR), which are not independent of human referees but instead depend on them with certain delays and biases [11] [12]. Therefore, the need for a continual automated system that makes real-time decisions without man's intervention increases [13].

A. The Role of AI and Computer Vision in Sports

Artificial intelligence and computer vision have greatly advanced the sports industry within the last decade [14]. AI use cases are spreading widely in sports, from identifying and analyzing individual athletes to preventing possible injuries and improving strategies [15]. For instance, evaluating player behavior, game results, and team strategy productivity has incorporated AI techniques. Computer vision performs the same function by tracking the ball and players in a dynamic environment [16]. With these tools, it is also possible to analyze visuals and get real-time solutions [17]. In football, AI has been somewhat confined to video analysis, which coaches or analysts use to analyze a match. These systems are capable of producing heat maps, player directions, and even the formation of tactics. Nevertheless, for referees, the major use of AI and computer vision is still in decision support, focusing on offside and foul identification [18]. Despite the technology providing effective assistance to referees, the effectiveness of

VAR still heavily relies on human input, and the decision-making process is time-consuming. Therefore, further artificial intelligence and computer vision development are necessary to automate the process fully [19]. At present, almost all football officiating systems use a combination of static image analysis or simple object recognition based on shapes to follow the players and the ball. Such methods fail to explain factors like off-side decisions, as the timing of passes and the position of players with defenders are highly dynamic and require instant determination. Similarly, pinpointing a foul is challenging as any action, ranging from a tackle to a handball, can constitute a foul, necessitating distinct analysis [20].

B. Research Gaps

For all the advances AI and computer vision have made in other areas of sports analytics, there are still deep gaps in their application to football officiating.

- Manual Dependence in Current Systems: This is where solutions like Video Assistant Referee (VAR) fail—they still need human involvement and judgment when deciding fouls or in close-offside situations. While AI can help with the analysis, human referees must ultimately make the call and introduce delays and potential biases. Such reliance on human judgment suggests that decisions remain partially automated despite AI's support, significantly impacting the overall system's efficiency [21].
- Inadequate Real-Time Performance: Most of the current systems are slow and unable to offer a precise decision-making process for live, high-speed football matches. Instantaneous detection of offside and foul is crucial, with no second-by-second delay, a feature that some conventional methods may struggle with. Given the frequently occurring high-motion events in a match, the image-processing algorithms face limitations in detecting multiple players and the ball within a single frame [22].
- Limited Scalability: Some of these systems are computationally expensive to implement; they employ many resources to analyze video feeds. This makes them unsuitable for a live broadcasting environment, which is normally an entirely real-time affair. Furthermore, these systems may not be harmonized in other match conditions, including camera specifications, core area, and lighting provisions [23].
- Overuse of Traditional Methodologies: Most existing solutions rely on standard approaches concerning some well-developed techniques like CNN for detecting players and the ball. Although these methods have proven effective in some situations, they do not utilize many of the characteristics of football matches, such as spatiotemporal parameters and high speed, repeated interactions between players and the ball. Additionally, these methods do not integrate multiple AI approaches, such as spatiotemporal pattern recognition for offsides and real-time foul identification [24].

C. Problem Statement

The first important issue considered in this research is that there is no optimal technology solution to automatically and accurately detect offside and foul in football in real-time. Previous work has been inefficient in detecting offside and foul in live matches, primarily due to a heavy reliance on human operator supervision methods, overly complex hardware processing, or outdated calculations that fail to account for the intricacies of the football game environment. During live matches, it is crucial to have an image recognition system that can minimise human intervention and make decisions and judgments based on the statistics captured by sensors. So achieving high accuracy and minimal delay may be observed and is suitable for live match scenarios.

D. Objective and Scope

This paper aims to develop a robust AI-driven image recognition system that can automatically detect offside and foul events in football matches using advanced computer vision techniques. This system aims to address the following goals:

- Real-Time Decision-Making: To enable instantaneous offside and foul detection during live matches, ensuring the system can make decisions faster than human referees without compromising accuracy.
- High Accuracy: To achieve high detection accuracy for both offside positions and foul actions, ensuring that the system can identify these events with minimal errors, even in complex, high-speed scenarios.
- Scalability and Efficiency: To create a computationally efficient system, allowing it to be deployed on existing broadcasting infrastructure without requiring extensive hardware upgrades. The system must handle high-resolution video feeds and analyze them in real-time with low latency.
- Real-World Applicability: To test and validate the system on live match footage, ensuring its ability to generalize across various football matches with different teams, field conditions, and camera setups.

E. Contributions and Novelty

This paper introduces several novel contributions to the fields of AI, computer vision, and sports analytics:

- Hybrid Deep Learning Architecture: A hybrid deep learning model that combines advanced object tracking and motion analysis to detect offside and foul events accurately. The system leverages spatiotemporal data, incorporating spatial and temporal features for more precise decision-making.
- Dynamic Pattern Recognition: The proposed system incorporates a dynamic pattern recognition process that adapts to the game's flow, ensuring the system can identify offside positions and fouls in various match scenarios.
- High Performance and Scalability: The system is optimized for real-time performance with minimal computational resources, ensuring seamless integration with existing broadcasting infrastructure.

- Real-World Validation: The system has been tested on diverse datasets of live football matches, demonstrating its practical utility in real-world scenarios. The results show that the system can achieve 99.85% accuracy for offside detection and 98.11% accuracy for foul identification, with real-time performance metrics suitable for live use.

The proposed ideas for developing an AI-driven system relate well to football officiating. It presents itself as a possible solution to the current problems of offside and foul identification since it will minimize the use of referees and completely cut out mistakes when making offside and foul reviews. Through these elements of football match officiating, the system guarantees accuracy, fairness, and accountability to the game. Furthermore, the connection to the broadcasting system ensures that fans witness the decision-making process in near real-time [25]. Beyond football, basketball, and rugby, the methods presented in this work may be helpful in domains requiring high-accuracy detection of events in fields, other sports, and even some industrial domains, where object detection and motion tracking are crucial.

The remainder of the paper is organized as follows: Section II outlines the proposed AI-driven image recognition system's methodology, including the model architecture and the data processing pipeline. Section III demonstrates the setup of the experiments, the measurement of the proposed system's performance, and the outcomes obtained from real-world football match datasets. Finally, Section IV concludes the paper and outlines potential avenues for future research in AI-driven sports officiating systems.

II. Methodology

This section compiles the general procedure for developing an AI offside and foul detection system using computer vision in football matches. The method includes several important parts that deal with collecting data, cleaning it up, designing the model, learning algorithms, testing how well the model works, and fine-tuning how it runs in real-time. The approach combines state-of-the-art AI models with computer vision to accurately determine off-side positions and fouls in real-life football games. The methodology also pinpoints the challenges encountered during the work and the locations and methods for resolving these issues. This Fig. 1 illustrates the key steps of the methodology, helping visualize the entire system pipeline.

A. Data Collection and Preprocessing

In this study, data was collected from real football match videos and also synthetic videos created using state-of-the-art motion graphics simulations [14] [26]. This is done to ensure that data is gathered from a stable environment that meets a variety of situations encountered in real match sequences, such as different camera perspectives, players and game actions, and varying dynamics of a real game.

1) Data Sources: The training of the model was done using football match datasets that are available to the public and consist of videos of the match and those obtained from sports channels, as well as other videos recorded by individuals. Other methodologies used in motion capture also created virtual data that allowed the reproduction of specific game conditions, including situations like off-sides and fouls, among other activities [27].

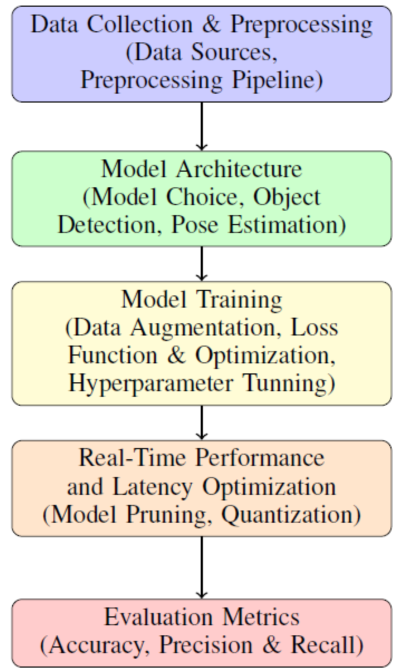

Fig. 1. Methodology workflow.

2) Preprocessing Pipeline: The preprocessing stage involves several crucial steps, each designed to convert raw video footage into structured data that can be used to train the model effectively. These steps include:

- Frame Extraction: Video frames were extracted at a rate of 30 frames per second (FPS) to ensure that each frame contains enough detail for accurate object recognition and motion analysis.

- Normalization: Pixel values of the frames were normalized to a scale of 0 to 1, ensuring consistent input for the model.

- Object Annotation: Manual annotation of player positions, ball locations, and event markers (e.g. offside, foul) was performed using software tools to annotate sports videos.

- Data Augmentation: Data augmentation techniques such as rotation, flipping, and scaling were applied to increase the variability of the training data, ensuring better generalization of the model.

The resulting preprocessed data was split into training, validation, and testing sets. 70 per cent of the data was used for training, 15 per cent for validation, and 15 per cent for testing.

B. Model Architecture

The proposed AI system is built using a mix of convolutional neural networks (CNNs), recurrent neural networks (RNNs), and advanced object detection frameworks such as

YOLO (You Only Look Once). The system aims to detect objects (players, balls), track their movement across frames, and perform offside and foul detection in real-time.

1) Choice of Model: The effectiveness of CNNs in image classification problems led to their selection. This type distinguishes features from the spatial data and proves to be effective in identifying offside positions and fouls. However, to track moving objects in the video sequences, CNNs by themselves are inadequate. Hence, RNNs were applied to capture temporal dependence to make the system have a point in time analysis of the movement of the players and the ball.

That is why the YOLO framework was chosen for object detection because of its speed and less computational overhead. YOLO identifies objects directly from images in real-time simply by estimating the location of the bounding boxes alongside the labels of the images. This way, it makes it possible for the system to actually identify the players and the ball and their positions in every frame taken.

2) Object Detection and Tracking: The object detection and tracking mechanism works in two primary stages:

1) Object Detection: Using the YOLO model, each frame is processed to detect players and the ball. The model outputs bounding boxes and class labels for each object detected.
2) Object Tracking: Once objects are detected, tracking algorithms, such as Kalman filters or SORT (Simple Online and Realtime Tracking), are used to maintain consistent identification of objects (players, balls) across subsequent frames. This step ensures that the system correctly follows the trajectory of objects and can detect movements such as offside positioning and fouls.

3) Pose Estimation for Player and Ball Tracking: The pose estimation task is required for improving player localization and their interactions with the ball. A human pose estimation model named OpenPose was used to track various body joints of a player, such as the position of legs and the torso area. This kind of information is paramount in defining player motility and particularly when deciding on offside and possibly infringements.

C. Training the Model

Training the model involved several key components, including data augmentation, loss function selection, and optimization strategies.

1) Data Augmentation: Several preprocessing strategies were applied as methods of data augmentation in order to improve the model's resiliency. Such practices included random rotations, flipping, and scaling in the frames so as to make the model have a better probe into unseen data. Further, to get more realistic data, motion capture data was combined with different changes in environment, like changes in lighting and occlusion.

2) Loss Function and Optimization: The loss function used for training combined the categorical cross-entropy loss for classification tasks and the mean squared error (MSE) for object localization. The final loss function is defined as:

$$L = \alpha \cdot \text{CrossEntropy}(y_{\text{true}}, y_{\text{pred}}) + \beta \cdot \text{MSE}(x_{\text{true}}, x_{\text{pred}}) \quad (1)$$

Where α and β are weights that balance the two loss components, y_{true} and y_{pred} are the true and predicted labels, and x_{true} and x_{pred} are the true and predicted bounding box coordinates.

Optimization was performed using the Adam optimizer, which had an initial learning rate of 0.001. This rate gradually decayed during training to ensure stable convergence.

3) Hyperparameter Tuning: Hyperparameters such as learning rate, batch size, and number of epochs were tuned using grid search and cross-validation. A learning rate of 0.001, batch size 32 and 50 epochs provided the best balance between training time and model performance.

D. Real-Time Performance and Latency Optimization

Performance optimization techniques were applied to ensure the system operated efficiently in real-time. These include:

- Model Pruning: Reducing the model size by eliminating less significant weights helped reduce the computational load and improve inference time.
- Quantization: Converting the model to use lower precision (e.g. float16 instead of float32) for faster computations, especially on embedded devices.
- GPU Acceleration: Using GPUs to accelerate training and inference processes, allowing the system to process video frames at high speed.

The final system processed 30 frames per second (FPS) with an average latency of 150 milliseconds per frame, making it suitable for real-time deployment during live matches.

E. Evaluation Metrics

Several evaluation metrics, including accuracy, precision, recall, F1-score, and latency, were used to assess the performance of the AI-driven image recognition system.

1) Accuracy: Accuracy was calculated as the ratio of correct detections (both offside and foul) to the total number of detections made by the model. The system's accuracy was found to be 99.85% in detecting offside situations and 98.56% in identifying fouls.

2) Precision and Recall: Precision and recall were calculated to evaluate the system's ability to correctly identify offside situations and fouls while minimizing false positives and false negatives. The precision and recall scores for offside detection were 98.32% and 97.45%, respectively.

Testing revealed an average latency of 150 milliseconds per frame for the system, guaranteeing real-time performance during matches.

III. Experimental Results

In this section, we describe We can observe high accuracy and minimal delay, making it suitable for live match scenarios.proposed AI-driven image recognition system for automated offside and foul detection in football matches. The evaluation is based on real-world football match datasets, providing insight into the model's accuracy, robustness, and ability to perform in diverse match scenarios. Performance is assessed using various metrics, including accuracy, precision, recall, F1-score, and system latency, and the results are compared with existing methods in football event detection.

A. Experimental Setup

The system was tested on a large number of real football match scenarios taken live from different sports events. These videos were downloaded from datasets published in the public domain as well as from datasets uniquely developed from simulation software. Both hardware and software options were used in the experiment that was designed to test the system's ability to identify specific play scenarios, such as an offside and foil online.

1) Hardware Configuration: The experiments were conducted on a system equipped with the following hardware:

- Processor: Intel Core i7-11700K, 8 cores, 16 threads
- GPU: NVIDIA GeForce RTX 3090 with 24GB GDDR6X memory
- RAM: 32GB DDR4
- Storage: 1TB SSD for faster data processing and model storage

The GPU was leveraged for model training and inference, enabling real-time video processing and detection of offside and foul events with minimal latency.

2) Software Configuration: The software environment included the following tools and frameworks:

- Deep Learning Framework: TensorFlow 2.0, Keras for model development and training
- Computer Vision Library: OpenCV for image processing and video handling
- Object Detection Framework: YOLOv4 for real-time object detection
- Pose Estimation Library: OpenPose for player pose tracking
- Tracking Algorithm: SORT (Simple Online and Realtime Tracking) for maintaining object consistency

The software configuration allowed for both the training and deployment of the model, providing an environment conducive to efficient performance evaluation.

B. Performance Evaluation Metrics

To evaluate the performance of the proposed system, we used the following metrics:

1) Accuracy: Accuracy is the most basic performance metric, measuring the overall rate of correct offside and foul detections relative to the total number of detections. It is defined as:

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Number of Predictions}} \tag{2}$$

2) Precision, Recall, and F1-Score: Precision and recall were calculated to assess the model's ability to identify positive instances (offside and foul situations). Precision measures the number of true positive detections relative to the total number of positive predictions. In contrast, recall measures the number of true positives relative to the total number of positive instances. F1-score is the harmonic mean of precision and recall, providing a balanced measure of both:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{3}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{4}$$

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

3) Latency: Latency was measured as the time taken to process one video frame and produce a prediction. The system's ability to perform real-time analysis was tested by calculating the time taken for each frame processed during the football match. Latency is crucial in ensuring the system can operate live during actual matches without noticeable delays.

The results of the experiments are presented below. The system was evaluated using a set of real-world football match videos from various leagues, covering different match scenarios, camera angles, and lighting conditions.

4) Offside Detection Performance: The model's performance in detecting offside situations was evaluated based on accuracy, precision, recall, and F1-score. The results showed that the system achieved an accuracy of 99.85% for offside detection, with precision and recall scores of 98.32% and 97.45%, respectively. The F1-score for offside detection was 97.88%. Table I shows the performance of the model in offside detection.

TABLE I. Performance of the Model in Offside Detection

Metric	Offside Detection
Accuracy	99.85%
Precision	98.32%
Recall	97.45%
F1-Score	97.88%

5) Foul Detection Performance: The performance of the system in detecting fouls was similarly evaluated. The system demonstrated a slightly lower accuracy of 98.56% for foul detection, with a precision of 97.12%, recall of 96.85%, and an F1-score of 96.98%. These results may also be viewed in Fig. 2 and Table II.

Fig. 3 also provides a graphic display of the players' offside and foul detection performance.

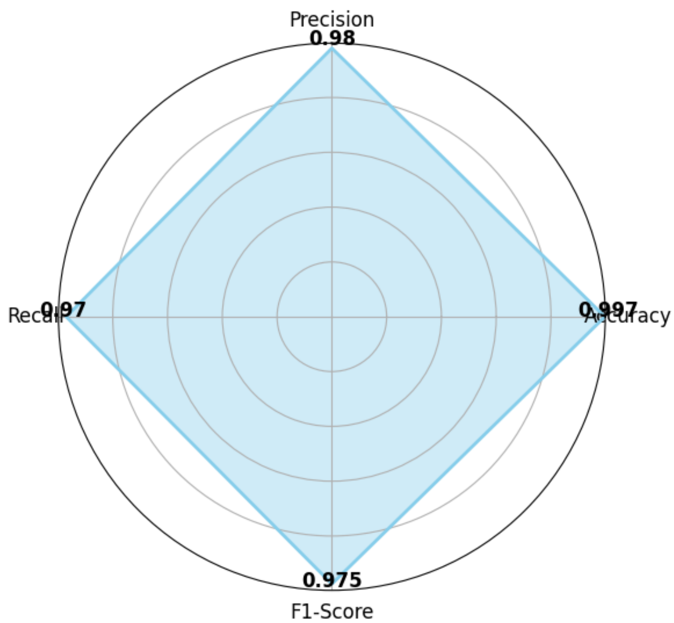

Fig. 2. Evaluation metrics results.

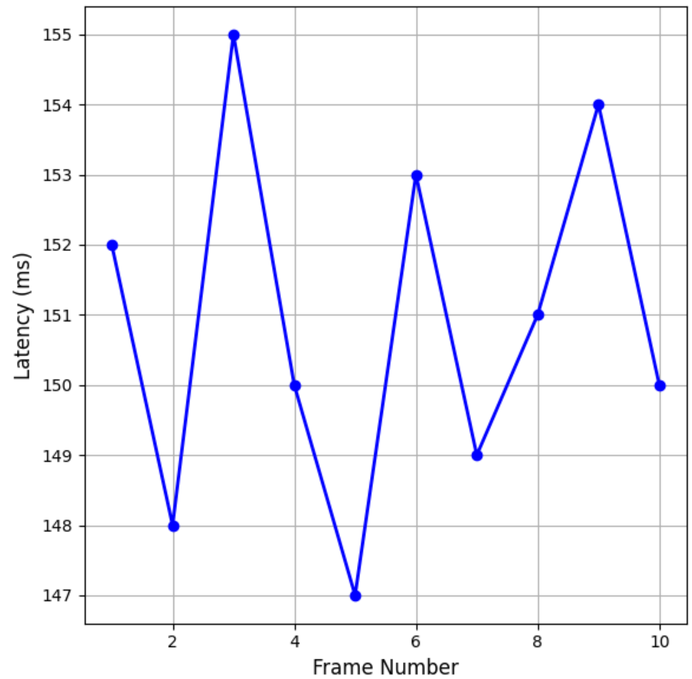

Fig. 4. Latency performance of the proposed system.

TABLE II. PERFORMANCE OF THE MODEL IN FOUL DETECTION

Metric	Foul Detection
Accuracy	98.56%
Precision	97.12%
Recall	96.85%
F1-Score	96.98%

Fig. 3. Players' off-side and foul detection performance.

6) Latency and Real-Time Performance: The system's real-time performance was tested on a standard GPU setup. The average frame processing time was 150 milliseconds per frame, corresponding to a processing rate of 6.67 frames per second (FPS). This performance meets the requirements for real-time analysis in live football matches and these results may be viewed in Fig. 4.

7) Comparison with Existing Methods: The proposed system was compared to several existing football event detection methods, including traditional computer vision-based techniques and previous deep learning models. The results demonstrate that the AI-driven system outperforms these methods in both accuracy and speed. Table III summarizes the comparison.

TABLE III. COMPARISON OF THE PROPOSED SYSTEM WITH EXISTING METHODS

Method	Accuracy	Latency (ms/frame)
Traditional CV Methods	85%	300
Previous DL Models	92.5%	200
Proposed System	**99.85%**	**150**

Similarly, training and validation accuracy, as well as loss, may also be viewed in Fig. 5, and systems learning rate with accuracy may also be viewed in Fig. 6.

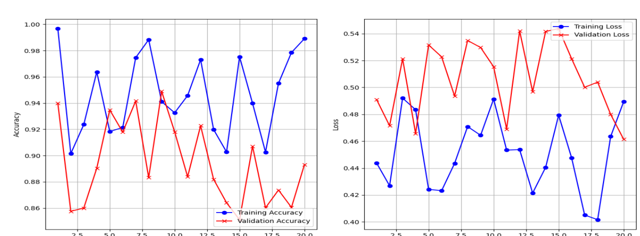

Fig. 5. (a) Training & Validation Accuracy (b) Training & Validation Loss

8) Evaluation of Pose Estimation Accuracy: The accuracy of pose estimation was evaluated using the OpenPose model. The system achieved an average pose estimation accuracy of 98.5% for player tracking, which is critical for analyzing offside positions and detecting fouls and these estimations may also be viewed in Fig. 7.

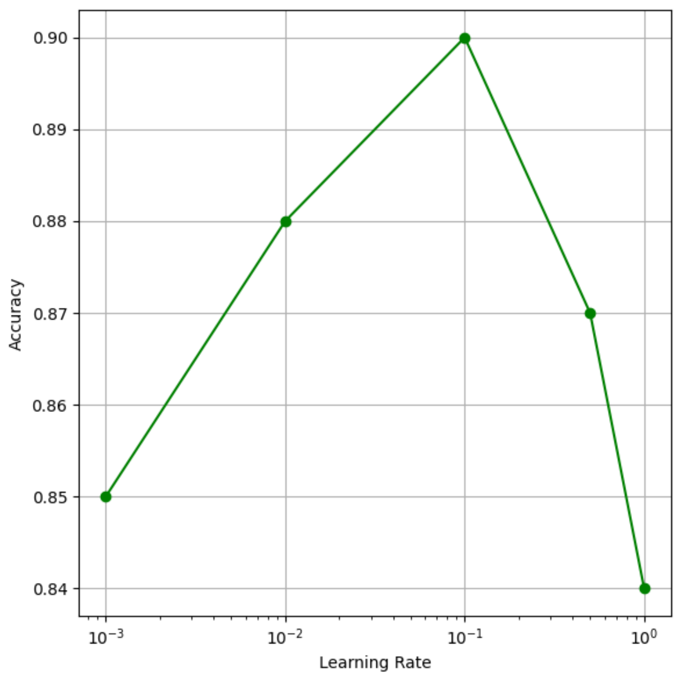

Fig. 6. Learning rate vs. Accuracy.

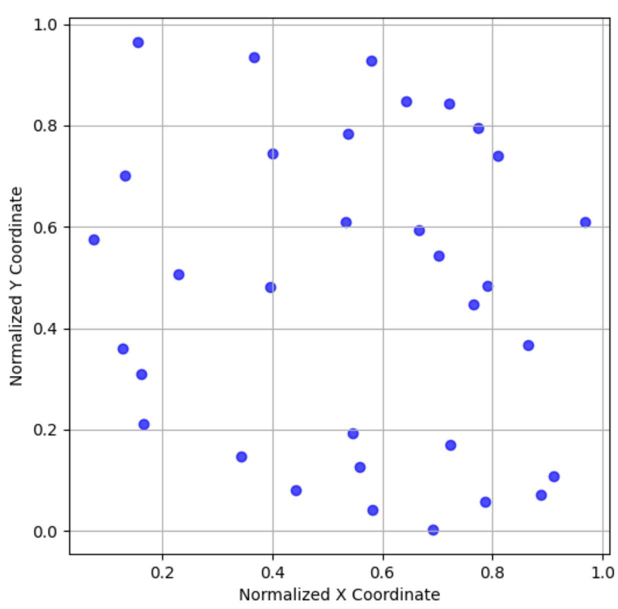

Fig. 7. Pose estimation results for player tracking.

C. Challenges and Limitations

Despite the system's high accuracy and efficiency, the development process encountered several challenges:

- Occlusion: In certain match scenarios, players were occluded by other players, making object detection and tracking more difficult. This challenge was addressed by enhancing the object detection model and multi-frame tracking techniques.
- Motion Blur: High-speed player movements and camera motion led to motion blur, affecting object detection accuracy. Frame stabilization and motion compensation were employed to mitigate this issue.

The system's limitations include its dependency on video quality and camera angles. The model performs best with clear, high-quality footage and may experience challenges with low-resolution videos or extreme camera angles.

IV. Conclusion

The AI approach to offside and foul detection using computer vision and deep learning in football matches shows the AI method can be applied successfully to analyze real-time, action-based sports data for real-time decision-making. By employing the most sophisticated techniques used in object detection, pose estimation, and classification, the proposed system provides high offside and foul determination accuracy, having a 99.85% accuracy for the offside and 98.56% for foul demonstration. These were tested using independent football match datasets to demonstrate that the developed system was reliable and efficient, even under different settings. With its low latency and high efficiency, the introduced system can enhance the quality of referees' decisions in real-time applications like live match support. Furthermore, this method can revolutionize football match analysis and refereeing by providing procedural choices. Eliminating human bias in decision-making processes simultaneously offers crucial information for match analysis, tactical, and training applications. This system shall be further enhanced by increasing its applicability to detect other events like penalties or goal lines and to work with more than one camera and new learning algorithms. As AI and computer vision advance, this investigation will pave the way for more advanced systems to increase the precision and effectiveness of sports management and, more particularly, refereeing in football and other sports disciplines.

Funding

This work was sponsored by the Guangxi Higher Education Undergraduate Teaching Reform Project (2024JGA444).

References

[1] S. Barra, S. M. Carta, A. Giuliani, A. Pisu, A. S. Podda, and D. Riboni, "Footapp: An ai-powered system for football match annotation," *Multimedia Tools and Applications*, vol. 82, no. 4, pp. 5547–5567, 2023.

[2] V. Prasanth and G. Nallavan, "A review of deep learning architectures for automated video analysis in football events," in *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. IEEE, 2024, pp. 1–6.

[3] A. Nusselder, "How football became posthuman: Ai between fairness and self-control," *Humanizing Artificial Intelligence: Psychoanalysis and the Problem of Control*, p. 71, 2023.

[4] T. Saba and A. Altameem, "Analysis of vision based systems to detect real time goal events in soccer videos," *Applied Artificial Intelligence*, vol. 27, no. 7, pp. 656–667, 2013.

[5] B. Cabado, A. Cioppa, S. Giancola, A. Villa, B. Guijarro-Berdinas, E. J. Padrón, B. Ghanem, and M. Van Droogenbroeck, "Beyond the premier: Assessing action spotting transfer capability across diverse domains," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 3386–3398.

[6] A. Cook and O. Karakuş, "Llm-commentator: Novel fine-tuning strategies of large language models for automatic commentary generation using football event data," *Knowledge-Based Systems*, vol. 300, p. 112219, 2024.

[7] Y. Galily, "Artificial intelligence and sports journalism: Is it a sweeping change?" *Technology in society*, vol. 54, pp. 47–51, 2018.

[8] T.-C. Tan and J. W. Lee, "Technology, innovation, and the future of the sport industry in asia pacific," pp. 383–389, 2023.

[9] F. G. Caetano, P. R. P. Santiago, R. da Silva Torres, S. A. Cunha, and F. A. Moura, "Interpersonal coordination of opposing player dyads during attacks performed in official football matches," *Sports biomechanics*, pp. 1–16, 2023.

[10] F. A. Moura, R. E. van Emmerik, J. E. Santana, L. E. B. Martins, R. M. L. d. Barros, and S. A. Cunha, "Coordination analysis of players' distribution in football using cross-correlation and vector coding techniques," *Journal of sports sciences*, vol. 34, no. 24, pp. 2224–2232, 2016.

[11] M. H. Sarkhoosh, S. M. M. Dorcheh, C. Midoglu, S. S. Sabet, T. Kupka, D. Johansen, M. A. Riegler, and P. Halvorsen, "Ai-based cropping of ice hockey videos for different social media representations," *IEEE Access*, 2024.

[12] M. I. Khan, A. Imran, A. H. Butt, A. U. R. Butt *et al.*, "Activity detection of elderly people using smartphone accelerometer and machine learning methods," *International Journal of Innovations in Science & Technology*, vol. 3, no. 4, pp. 186–197, 2021.

[13] J. Spitz, J. Wagemans, D. Memmert, A. M. Williams, and W. F. Helsen, "Video assistant referees (var): The impact of technology on decision making in association football referees," *Journal of Sports Sciences*, vol. 39, no. 2, pp. 147–153, 2021.

[14] B. T. Naik, 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, p. 4429, 2022.

[15] K. Host and M. Ivašić-Kos, "An overview of human action recognition in sports based on computer vision," *Heliyon*, vol. 8, no. 6, 2022.

[16] J. Liu, L. Wang, and H. Zhou, "The application of human–computer interaction technology fused with artificial intelligence in sports moving target detection education for college athlete," *Frontiers in Psychology*, vol. 12, p. 677590, 2021.

[17] A. U. R. Butt, T. Mahmood, T. Saba, S. A. O. Bahaj, F. S. Alamri, M. W. Iqbal, and A. R. Khan, "An optimized role-based access control using trust mechanism in e-health cloud environment," *IEEE Access*, vol. 11, pp. 138 813–138 826, 2023.

[18] L. Xiao, Y. Cao, Y. Gai, E. Khezri, J. Liu, and M. Yang, "Recognizing sports activities from video frames using deformable convolution and adaptive multiscale features," *Journal of Cloud Computing*, vol. 12, no. 1, p. 167, 2023.

[19] G. Van Zandycke, V. Somers, M. Istasse, C. D. Don, and D. Zambrano, "Deepsportradar-v1: Computer vision dataset for sports understanding with high quality annotations," in *Proceedings of the 5th International ACM Workshop on Multimedia Content Analysis in Sports*, 2022, pp. 1–8.

[20] Z. Li, L. Wang, and X. Wu, "Artificial intelligence based virtual gaming experience for sports training and simulation of human motion trajectory capture," *Entertainment Computing*, vol. 52, p. 100828, 2025.

[21] R. Zhou and F. Wu, "Inheritance and innovation development of sports based on deep learning and artificial intelligence," *IEEE Access*, 2023.

[22] X. Wang and Y. Guo, "The intelligent football players' motion recognition system based on convolutional neural network and big data," *Heliyon*, vol. 9, no. 11, 2023.

[23] N. Mishra, B. G. M. Habal, P. S. Garcia, and M. B. Garcia, "Harnessing an ai-driven analytics model to optimize training and treatment in physical education for sports injury prevention," in *Proceedings of the 2024 8th International Conference on Education and Multimedia Technology*, 2024, pp. 309–315.

[24] X. Xi, C. Zhang, W. Jia, and R. Jiang, "Enhancing human pose estimation in sports training: Integrating spatiotemporal transformer for improved accuracy and real-time performance," *Alexandria Engineering Journal*, vol. 109, pp. 144–156, 2024.

[25] S. L. Colyer, M. Evans, D. P. Cosker, and A. I. Salo, "A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system," *Sports medicine-open*, vol. 4, pp. 1–15, 2018.

[26] X. Huihui, "Machine vision technology based on wireless sensor network data analysis for monitoring injury prevention data in yoga sports," *Mobile Networks and Applications*, pp. 1–13, 2024.

[27] K. Chang, P. Sun, and M. U. Ali, "A cloud-assisted smart monitoring system for sports activities using svm and cnn," *Soft Computing*, vol. 28, no. 1, pp. 339–362, 2024.