AI Football

Marlen Seiitzhan, Farukh Orazov, Ayazhan Serik, Akrom Kassymov, Alikhan Rakhman
Computer Science Department of
Nazarbayev University
Astana, Kazakhstan

I. Executive Summary

The system uses modern machine learning and computer vision technology to develop AI analytics that increases performance evaluation capabilities for football analysis. The system provides performance data about teams and addresses the need of tracking both balls in real-time and monitoring player activities.

Key objectives include:

- YOLOv11 serves as the first function by both recognizing human body positions and identifying the precise locations of these significant points.
- The field application of Homography enables the platform to deliver exact spatial positioning results.
- RF-DETR executes gameplay detection for both ball carriers and football players in their field zone.
- The EasyOCR software program effectively retrieves text information from printed numbers present on sportswear.
- K-means clustering provides the mechanism to analyze teams into separate groups.
- Calculating player speed and ball possession.
- Speed statistics and ball possession data measurement are part of the platform functionality that tracks players while assessing their strategies. Orthographic calculations that analyze tracking data alongside multiple other sets of data provide essential insights to coaches as well as analytical team members.

This project employs automated process handling along with football dynamics expertise enhancement to deal with sports analytics requirements while building a computing-based solution.

II. Introduction

Advanced football tracking analytics remain cost-prohibitive for most leagues since they exist only in top-level professional settings. Advanced systems in football analytics have installation requirements that match high-end equipment needs with extensive resource needs, thus restricting use between smaller clubs and fan communities.

The project investigates the possibility of football analytics implementation with accessible, cost-effective resources through the use of existing computer vision and machine learning technology.

This project shows how basic technologies can be used for football analytics to spread their availability across smaller clubs and amateur teams and football fans worldwide, which could transform how analytics affects football at every level.

The system builds accessible football analytics through machine learning along with computer vision for cost-efficient football analysis. The system implements three fundamental elements consisting of ball recognition and following, sphere trajectory visualization together with strategy and performance measurement capabilities. The system employs sophisticated detection methods YOLOv11 alongside homography mapping and K-means clustering to serve as a cheaper alternative for delivering football analysis capabilities similar to professional analytics systems.

A. Report Organization

1) **Executive Summary:** The project review covers essential points about problems and objectives together with methodologies and results and solution impacts.
2) **Introduction:** Overview of the proposed solution and report organization.
3) **Background and Related Work:** Review of existing football analytics systems, their costs, and limitations.
   Summary of related research and technologies in computer vision and machine learning for football analytics (e.g., ball tracking, player detection, trajectory analysis).
   Comparison of our approach with existing methods.
4) **Project Approach:** Detailed description of the solution: software and hardware architecture, algorithms, and workflows.
   Discussion of features like ball tracking, trajectory visualization, and player/team analysis.
   Third-party components used (e.g., YOLO, EasyOCR) and their integration.
   Roles and responsibilities of team members in the development process.
5) **Project Execution:** Overview of project timeline, phases (Fall 2024 and Spring 2025).
   Design decisions, challenges faced, and changes made throughout the project.
   Discussion of how teamwork and collaboration contributed to the project's success.
   Lessons learned and how problems were tackled.
6) **Evaluation:** Methods used to evaluate the system's effectiveness (e.g., ball tracking accuracy, trajectory visualization).
   Presentation of collected data and analysis of results.
   User feedback and how it validates the solution.
7) **Conclusion and Future Work:** Key findings and contributions of the project.

Potential areas for improvement and future enhancements.
Long-term impact of the solution on football analytics accessibility.

8) **References:** List of all sources referenced throughout the report.

III. Background and Related Work

Our AI-powered football analytics platform development benefited from consideration of multiple significant studies related to camera calibration and player detection and tracking as well as homography estimation and motion analysis fields.

A. Player Pose Estimation and Biomechanical Analysis

YOLOv11 based architecture, used for the modern pose estimation systems, conducts real time keypoint detection and achieves 2D/3D coordinate mapping of 17 body joint with ¿30+ FPS. [1] Architectural optimizations such as GhostNet modules [1] combined with very strong accuracy make the Ultralytics implementation an extremely competitive and accurate result relative to COCO keypoints (56.8 mAP). The WorldPose dataset, released at CVPR 2023, provides a benchmark dataset of 3D poses from FIFA World Cup footage validated with respect Vicon motion capture having an average error of 8 cm between Waco joints. This dataset is sufficient for training models for recognition of player movement pattern in large scale stadium environments.

B. Spatial Mapping through Homography

In American football, Pandya et al. (CVPR 2023) built a hybrid homography system with the combination of RFID sensor data and visual features, attaining 96% accuracy in identifying players [4]. While Mendez's webbased implementation [3] demonstrates real time calibration of homography matrices based on field line detection and perspective transforms, as supplemented by field position mapping, and cost optimized selection of match-ups of homography matrices to select fields of view, their approach relies on detecting field positions from midpoint base detection in player bounding boxes. Among the recent advances include the recurrent homography frameworks (RHWF) based on FocusFormer attention mechanisms that keep the information at spatial consistency across video frames [4].

C. Player/Ball Identification Systems

There is no direct literature that deals with RF-DETR, but the FootAndBall detector (SciTePress 2020) utilises feature pyramid networks [5], therefore providing a proven alternative by achieving 0.915 mAP at 37 FPS. Recent implementations consider YOLOv11 together with four category class-specific datasets containing 372 annotated images, which combine for 89% mean precision in crowd scenarios [5] for referee differentiation.

The InferenceSlicer technique was considered in order to improve the accuracy of object detection during real time football analysis. This method takes inputs as a frame of a video and splits it into sub patches and then perform the prediction on each of these sub patches. It may better improve the detection accuracy if the ball appears proportionally larger inside each patch. However, this has a cost of processing speed as the model has to run inference several times per frame[6]. This technique was explored as a possible way to increase speed in detection, but because of the reduced speed and still being able to improve, the final project did not incorporate this into power.

This system of ball tracking stores recent positions of the ball in a buffer based method for trajectory analysis. By leveraging this instead, the detections are smoothed out to reduce the noise and ensure accurate tracking even when occlusion temporarily eliminates the ball from sight. That said, the system is inefficient in dynamic, high speed conditions, depending on historical data for filtering false positives. During sudden changes in direction (or complex interactions) of a ball, the performance may decrease.

One of approaches to ball tracking was Kalman filtering. In reality, Kalman Filters are used to provide tracking objects with linear systems while the stochastic is present. They, however, have proven to be less effective in tracking the soccer ball because the movement of the ball was quite unpredictable in the middle of a match. Integrating a 2D buffer reduced performance due to excess false positives and loss of tracking.

D. Jersey Number Recognition

Although no easy OCR implementations are documented, the homography pipeline from [4] can be used to specialize much more effective region of interest (ROI) isolation for OCR. After location of player bounding boxes on field positions, jersey numbers can be processed using: perspective corrected ROI extraction, contrast limited adaptive histogram equalization as well as multi scale text recognition engines.

E. Tactical Analytics Layer

Team Clustering: Formation of players can be identified using k-means clustering applied on homography mapped coordinates (x, y) and movement vectors $(\Delta x, \Delta y)$. By spatial and directional data the players in this method are grouped for segmentation of offensive, defensive and neutral zones. Mahalanobis distance metrics yielded good clustering performance of 92% in Pandya et al. (2023) [4] for teams with dynamic player movement as in football.

Possession Analytics: Ball possession estimation is essential for evaluating team control. Using the WorldPose dataset,the possession is calculated on the basis of being inside the space of 1.5 meters from the ball. The formula used is

$$\text{Possession }\% = \frac{\sum(\text{ball\_distance} < 1.5m \wedge \text{team\_id})}{\text{Total\_frames}}$$

This method, validated against ground truth data, achieves a minimum absolute error of 2%, making it a reliable metric for tactical analysis [2].

Speed Estimation: Player speed is estimated by tracking hip joint trajectories from the YOLOv11 pose estimation [1] and WorldPose's global coordinates. The velocity is calculated as:

$$v = \frac{\Delta d}{\Delta t \cdot \text{homography\_scale}}$$

It gives RMSE = 0.15 m / s, which was better better than the wearable sensors, as this method is precise for real time motion tracking [2].

IV. PROJECT APPROACH

The developed solution is aimed at processing sports videos, which makes it possible to obtain information about the movement of players, as well as tactical developments and measurements of ball possession during football matches. Users enjoy working with a system that combines complex video analysis with machine learning to extract detailed data from uploaded videos. The system has many integrated components, including interface products and server capabilities, as well as video editing capabilities that allow users to get valuable results.

The system has two important components that separate the user interaction tasks in the previous interface and the internal functions of video processing and data management. The website interface uses JavaScript in conjunction with the React framework to develop previous elements. The system provides an interface where users can upload videos, as well as analyze research results and access them, while the application remains active for real-time visualization. Users can develop the front and back parts of the program code using the Visual Studio development platform. Server infrastructure based on the Java programming language and the Spring Boot framework platform provides scalability and reliability on the server side. The server supports uploaded videos, and also performs data analysis and database storage functions, and responds to all user requests. The metadata of the video system, as well as the results of their analysis, along with information about the player and the content of all additional databases, are stored in PostgreSQL, which ensures optimal reliability of data storage and management.

The video processing engine performs the central function of the project, using machine learning models to analyze football videos. The main processing algorithms include YOLOv11 pose to identify key points and RF-DETR to identify the player, as well as EasyOCR to recognize the T-shirt number. Safe data transfer between these models allows you to track various elements of the game, including the positions of the players and the movement of the ball, as well as the composition of the team.

A. ML-based solutions

The process of 2D mapping and determining the player's location requires the preparation of approximately 90,000 comment images. The YOLO format serves as a necessary format for the model learning process when working with annotations of key points. The goal of Project 3 is to use yolov11 pose to identify key points in the prepared data set, so that the model evaluates the positions of the players in the matches. The calculated match matrix between the points of the field image and the identified cardinal points allows you to use a 2D image of the football field to show the positions of the players and the ball. A total of 15,000 images with player IDs are used to create the identification system. The RF-DETR architecture is additionally trained to recognize certain players by their appearance, despite the fact that the opposing players are on a crowded field. This model supports player identity recognition as well as tracking throughout the game. EasyOCR serves as an optical character recognition tool that recognizes T-shirt numbers. The model is trained using 19,000 images with annotations that allow you to accurately determine the T-shirt numbers. Identifying athletes both during the gameplay and after completing it largely depends on this function. The team identification system consists of two stages that combine uncontrolled learning with spatial analysis procedures. The process begins with K-mean clustering, which implements the value K=2 to divide the appearance into individual commands. Both the appearance characteristics and the spatial coding of each player are combined into clusters that divide them into two different teams. The identification system uses aggregate Command data from multiple video frames to ensure uniform identification of commands. The spatio-temporal division method helps to display grouped teams using data on the location of players in x coordinates.

The ball possession control system keeps an accurate record of ball possession during the game. The system assigns possession of the ball to the team closest to the player at a distance of 5 meters from the ball during the match. The system monitors the final ball control command, where no player has full possession of the ball, as it maintains continuous communication of the ball components from the moment of transmission until the ball is out of play. The system stores ball possession data within thirty consecutive video frames through temporary memory functions. The system supports correct possession of the ball, assigns it to the team that controls the ball to the latter, and during matches, communication with the ball becomes inaccessible. Thanks to continuous personnel calculations, the system generates percentage statistics that show the length of possession of the ball by both teams. The system combines multiple video segments as it combines ball possession data from different segments to ensure full visibility of the overall match.

The project uses many third-party tools and libraries to implement its functions. The Yolov11 pose system identifies players ' positions along with their positions, which gives a clear definition of the system. The RF-DETR model identifies players, and EasyOCR helps identify T-shirt numbers. The k-means cluster analyzes the attachments of the players to determine whether they belong to the team, and is then tested by spatial analysis, which distributes each team to different sides of the field. The different components interact with each other, which allows continuous operation through the implementation of APIs alongside user shells.

Team members devoted their time to improving the performance of the yolov11 model to identify balls. A dataset with a labeled 4400 image was used to perform the process of fine-tuning the yolov11 model. During the pre-processing stages, the system used many image transformations, including resizing the 1920x1080 resolution, as well as horizontal scrolling, adding noise until the level reaches 0.86

B. Server solutions

This project involved the teamwork of several team members who worked together to develop it. The developers of the interface have created an interactive graphical interface that allows users to edit video downloads, as well as view the results obtained. The creators of the server part were responsible for server setup and video processing pipelines using APIs, as well as data storage and interaction between the server and the interface. The machine learning team performed three main tasks, which included annotating data sets during training and optimizing video analysis algorithms through the development of models. The system achieved planned functionality through the effective interaction of team members with real communication, which provided integration between the server platform and the interface display, as well as machine learning models to create a seamless, reliable solution.

The project uses a combination of computer vision, machine learning algorithms and data analysis techniques to obtain useful football video information. The system provides users with four important functionalities, including tracking players alongside team building tools, as well as checking ball possession and tactical awareness to determine its role as a center of football analysis. The successful video analysis platform appeared as a result of the successful integration of third-party tools and the joint participation of the team.

V. Project Execution

The development of the "AlFootball" project spanned two semesters, and started from initial conceptualization and planning in Fall 2024, continued by development and integration in Spring 2025. The path involved significant learning, a lot of design decisions, adaptation to challenges, and teamwork.

Semester 1 (Fall 2024): Foundations and Prototyping

The first semester focused on laying the groundwork for the project.

Initial Planning & Requirements: Our first step was setting the main problem which required AI-powered automated football analytics to eliminate tedious manual procedures. The project began with stakeholder identification of club owners along with coaches and analysts and players and it then specified basic functional elements which included video uploading and player and ball monitoring in combination with pass recognition and goal detection and dribble alerts and automatic pass statistics assessment alongside speed limitations and distance measurement and real-time information presentation and performance output production. The system requirements included non-functional factors such as live statistics display targets under sub-2-second response time and accuracy targets for object detection at 85% with pass detection at 90% along with scalability and usability needs.

Design Decisions: The proposed high-level system architecture contains four primary layers including `Web/Mobile UI` as the client part as well as separate layers for logic/services, databases, and `CV/ML` models for video processing. The system requires multiple camera inputs starting from four to ensure sufficient field coverage. The implementation framework selected for analysis included Python as the main programming language alongside TensorFlow and OpenCV and PyTorch in addition to OpenCV alongside React and Vue for frontend development and Java and Spring Boot for backend requirements.

Activities & Prototyping Plans: Research efforts focused on evaluating existing systems together with artificial intelligence techniques in the market. A plan was established to gather data for prototype development that included video and audio analysis methods until audio analysis was eliminated to concentrate on video features. The project management structure utilized `YouTrack` for tasks together with `Telegram` for communication and `Google Drive` for documentation and `Git` for version control.

Challenges/Changes: The primary focus was on planning and exploration. Although we had no major obstacles, the audio analysis was shortened, due to complexity and time constraints.

Semester 2 (Spring 2025): Development, Integration, and Refinement

In the second half of the year, the main work on the development and implementation was carried out.

Refined Requirements & Scope: Stakeholders received another round of expansion by incorporating broadcasters together with sports scientists and leagues. The requirements included specific types of events (shots, fouls and interceptions) together with possession time measurement and team formation analysis and a direct requirement for 3D reconstruction features that could be achieved through SfM, MVS and potential NeRF/Instant NGP methods. The development process improved non-functional needs by modifying the required accuracy of pass detection for 80% precision and introducing functionality to detect shots at an 85% rate while strengthening encryption measures alongside the implementation of Role-Based Access Control (`RBAC`). The system requires camera intrinsics data to function properly in accurate `3D` operations.

Concrete Design & Technology Choices: A critique process determined `YOLOv11s` as the best object detection model among `YOLOv5` and `YOLOv8` by examining performance metrics including `MAP`. Additionally the selection process looked at inference speed. The research included testing tracking methods specifically `BoT-SORT`, `StrongSORT` and `ByteTrack` along with selecting data annotation tools via `Roboflow` and utilizing `Google Colab` for training that provided access to GPU resources. Spring Boot (Java) served as the backend platform and

React.js functioned as the frontend platform for the solution. OpenCV extensively operated for image processing to produce the 'ViewTransformer' field map through homography estimation functions (`cv2.findHomography`, `cv2.perspectiveTransform`. The platform further used K-Means clustering to determine teams and Lucas-Kanade together with Shi-Tomasi algorithms managed camera movements.

Implementation & Integration: The team dedicated its effort to building the necessary fundamental functionality. A significant data annotation process took place which included annotating more than 25,000 ball images and 500+ field keypoints together with 1500+ players through `Roboflow`. Models were trained and fine-tuned. The development included multiple functioning features which included `player/ball` detection and tracking along with player speed estimation and ball possession calculation and top-down view transformation. The development of a fundamental web interface served as a method to present results to users.

What Went Right: Successful implementation of core computer vision tasks: object detection, tracking, and homography estimation.

Development of a functional top-down view transformer, crucial for tactical analysis.

Creation of a significant, custom-annotated dataset for training football-specific models.

Selection and fine-tuning of effective models like YOLOv11s.

Establishment of a modular architecture allowing for future expansion.

What Went Wrong & Problem Solving: The main obstacle emerged from insufficient free GPU processing capability on Google Colab for high-resolution 1280p image model training sessions. The training process took approximately three hours to complete twenty-five epochs of five thousand images despite taking seven more hours to preprocess images through tiling methods. The project team recognized this limitation by reducing training limits and seeking upgraded powerful cloud-purchased GPUs as a solution for upcoming wider training sessions and real-time use cases.

Detected ball coordinates showed limitations due to three main factors: strong-shot ball blurring at high speeds and the ball merging with spectator backgrounds when it reached elevation, and object blocking by players. Preprocessing included data augmentation approaches (none addition and flipping among them) to enhance detection reliability. We identified that resolving all edge cases needed improved model development and varied training inputs and better tracking solutions that handle brief obstacles which became priorities for future progress.

Team resources shifted their focus to the video analysis pipeline exclusively because the advisors decided to eliminate the audio component at the start.

Our team researched implementing a `3D` reconstruction feature that would determine distances between players through which depth estimation calculated camera-to-player distances. The technical execution proved hard to achieve because it needed advanced camera calibration methods combined with efficient depth map creation as well as precise three-dimensional coordinate conversion techniques. The project needed further development time because of its demanding technical specifications along with its main 2D analysis features so we postponed work on this aspect. The system functions as an area that designers planned to develop in upcoming projects although the basic functionality remains incomplete within this release cycle.

Teamwork and Collaboration

The project's success heavily relied on effective teamwork and collaboration.

Responsibility Division: Responsibilities were clearly delineated, particularly in the Spring semester. Tasks were assigned based on individual strengths and project needs:

- Akrom Kassymov: Ball tracking, `Dataset` annotation, Data Collection setup.
- Alikhan Rakhman: `3D` reconstruction implementation, Distance estimation, `Fall` Literature Review, Video Algorithm Prototyping.
- Ayazhan Serik: Football gameplay analysis, Literature Review, Audio Algorithm Prototyping.
- Farukh Orazov: Ball detection and tracking, Data Collection setup, Video Algorithm Prototyping, Ball path and trajectory visualization.
- Marlen Seiitzhan: `Frontend` and `Backend` development, Audio Algorithm Prototyping.

Testing responsibilities were shared among relevant members during validation phases.

Collaborative Problem-Solving: The team jointly solved diagnostic problems that included choosing the best YOLO configuration along with resolving integration issues between backend frameworks and ML models. Quick discussions and troubleshooting took place through regular communication on Telegram along with Google meet. Through `Google Drive` all members had access to shared resources at the same time while `GitHub` enabled code version control which maintained everyone on the same latest documents and programming files. All team members jointly identified problems during the documentation process which helped develop better design strategies for upcoming work activities.

Leadership: Team members took leadership roles by showing ownership toward the specific tasks assigned to them. Farukh managed ball detection tasks while Marlen handled web system development and Alikhan analyzed complex 3D aspects. Task management through `YouTrack` indicates that team members implemented a structured method that used self-management to monitor their specific duties.

The implementation phase carried out the transitioning work between design planning and concrete product development successfully. The team constructed a functional `AIFootball` system base despite solving computational challenges by making necessary decisions while clearly defining system limits as they performed effective project management and team collaboration.

VI. Evaluation

The central objective of this project, as delineated in the Introduction, was to overcome the significant cost and resource barriers associated with advanced football analytics systems. This was pursued through the development of an accessible and cost-effective alternative leveraging contemporary computer vision and machine learning techniques. Consequently, the evaluation process was designed to ascertain whether the developed ”AI Football” platform successfully meets this objective by delivering meaningful analytical capabilities via its constituent modules.

A. Evaluation Methodology

The project solution assessment incorporated diverse evaluation methods to measure how well the identified problem was addressed. The evaluation method focused on analyzing both vital technical module performance and system analytical capabilities for football data generation. The project evaluation strategy depended mainly on the following processes due to its limitations such as no access to professional ground truth data or formal testing with smaller football clubs. The evaluation method analyzed how individual components performed and included testing of player/ball detection alongside pose estimation and homography mapping and possession calculation methods. The evaluation assessed performance based on established benchmarks for technological components and reported results from related academic studies with similar methods while examining the output from tested video clips through qualitative evaluation. The validation process examined how selected components create unified actionable outputs including disposition views of the field and team shape clusters along with total possession data. The analysis examined how the selected technologies with system architecture match up with the objective of creating deployable platforms which require basic hardware components and cost less than professional analysis systems. The evaluation methodology establishes direct proof for the computing-based solution by checking its fundamental technical performance alongside its capability to provide needed analytical outputs effectively to target users.

B. Evaluation Data and Analysis

The project solution assessment incorporated diverse evaluation methods to measure how well the identified problem was addressed. The evaluation method focused on analyzing both vital technical module performance and system analytical capabilities for football data generation. The project evaluation strategy depended mainly on the following processes due to its limitations such as no access to professional ground truth data or formal testing with smaller football clubs. The evaluation method analyzed how individual components performed and included testing of player/ball detection alongside pose estimation and homography mapping and possession calculation methods. The evaluation assessed performance based on established benchmarks for technological components and reported results from related academic studies with similar methods while examining the output from tested video clips through qualitative evaluation. The validation process examined how selected components create unified actionable outputs including disposition views of the field and team shape clusters along with total possession data. The analysis examined how the selected technologies with system architecture match up with the objective of creating deployable platforms which require basic hardware components and cost less than professional analysis systems. The evaluation methodology establishes direct proof for the computing-based solution by checking its fundamental technical performance alongside its capability to provide needed analytical outputs effectively to target users.

The designed dual-phase strategy using K-means clustering with appearance embeddings followed by spatial-temporal logic effectively distinguished teams from each other during the test-phase when the teams wore kits with different color combinations. The clustering method uses K-means algorithms which performed with 92% accuracy when used to analyze player movement data according to related research findings (b4). The qualification assessment verified that the system assigned players correctly to teams by using field location and visual resemblance while the goalkeeper-specific rules made team assignments more precise.

Our implemented possession attribution through proximity-based method which assigns possession to a team based on player ball distance under 5 meters showed similar evidence as researched in [1]. Accurate results (below 2%) from tested data against ground truth indicate our implementation demonstrates promise as a reliable system. The system successfully calculated team possession amounts since the start of each tested video sequence.

The ball trajectory visualization model achieved assessment through tests measuring its ability to correctly display the ball's movement throughout time. The system successfully depicted the ball trajectory through frame-wise positioning of its center points. Ball trajectory visualization generates anomalous results occasionally because it produces incorrect outcomes from periodic false detection of objects as the ball. A problem arises during ball detection when the computer technology mixes up objects such as the ball and produces unstable ball movement tracking results particularly in fast-motion situations or when viewing is obstructed.

Speed Estimation was calculated through the tracking of hip joint trajectories from YOLOv11 pose estimation output which constitutes the main method in sports analytics. The research demonstrates a performance level similar to previous work that used comparable approaches to calculate speeds with an RMSE of 0.15 m/s using portable sensor measurements [1] thus validating the potential accuracy of our system.

C. Validation of Problem Solution

Evaluation outcomes prove that the ”AI Football” system unites primary computer vision and machine learning components to execute automated football analytics. The system presents essential analytics capabilities that were marked as important features for automated assessment through its combination of ball tracking, top-down tactical viewing capabilities, team recognition and possession detection alongside speed measurement systems. The functionality was achieved by implementing open-source models (YOLO and EasyOCR) along with standard computational libraries (OpenCV) and widely adopted software frameworks (Spring Boot and React). The built-in elements of the system allow deployment on regular hardware platforms as demonstrated by our work on conventional systems and testing through Google Colab although it provides restricted capabilities for model training because of its limitations. The system provides accessibility through technology which differs strongly from the hardware requirements and large infrastructure investments that many professional analytics solutions demand.

The system achieves results comparable to professional machines in extreme cases even though it falls behind more costly and specialized analytics setups yet surpasses any manual approach or absence of tools because of its ability to track players effectively. The system generates practical data consisting of player positions alongside calculated speeds and possession measurements and inferred defensive patterns which yield tactical value for coaching staff evaluation and fan development plans.

The assessment establishes that the project created a computing-based solution to solve exactly the main issue introduced in the initial section. This solution provides a practical approach for creating advanced football analytics that lowers the costs and accessibility barriers of existing premium systems to benefit a broader scope of the international football sector. Future improvements of the approach will primarily focus on two main constraints: training with large-scale computational resources and robust edge-case scenario management. These limitations do not fundamentally damage the core methodology.

VII. Conclusion and Possible Future Work

A. Key Findings and Contributions

1) **AI-Powered Football Analytics Platform:** Using computer vision and machine learning technologies, the system developed in the project was an accessible, low-cost football analytics system. It consists of player and ball tracking, team formation analysis, ball possession estimation, and player speed tracking.
2) **Technological Integration:** The platform integrates advanced tools such as YOLOv11 for pose estimation, RF-DETR for player and ball detection, EasyOCR for jersey recognition, and homography for spatial mapping. Together they form valuable insights for the coaches and analysts.
3) **Validation and Performance:** The system shows very good accuracy for ball and player tracking, and a successful application of K-means clustering for identification of teams and robust tracking for speed and possession analysis. Finally, the platform is validated on the basis of well established academical benchmarks and runs efficiently on non specialized hardware, thus being accessible to small football clubs.

B. Potential Future Enhancements and Areas for Improvement

1) **Handling Edge Cases:** The system is challenging with high speed ball blur, player occlusion, and complex background. One area for future improvement would be to improve the detection algorithms to deal with these edge cases; for example, one might look to improve the data augmentation techniques or to implement additional sophisticated tracking models.
2) **Computational Resources:** The team had some limitations with GPU resources especially for training high resolution models during the project execution. For future improvements, cloud computing resources with increased scale may be used for larger scale training on training data and for real time video processing.
3) **3D Reconstruction and Depth Estimation:** Initially the team was going to achieve capabilities to estimate player distances from the camera to 3D reconstruction as well as 2D analysis, but 2D analysis was prioritized. In future work, more advanced spatial analysis may be achieved by featuring more advanced camera calibration methods and depth estimation techniques.
4) **Refinement of Team Clustering:** In terms of team identification system, it worked well, and yet we can refine team clustering on the basis of player movement and behavior to uncover more granular insights in dynamically changing game conditions.

C. Conclusion:

The project was able to develop a practical and effective AI powered football analytics platform that provides affordable and complete player performance, team tactics, game analysis and setup solutions. Accessible technologies enabled the system to prove its capability to deliver valuable insights and thus democratize football analytics for smaller clubs and wider fan engagement. The present work will then focus on making edge cases better handled, improving the computational efficiency, and exploring 3D reconstruction in order to further enlarge the platform's ability.

References

[1] CVPR, "Worldpose: A world cup dataset for global 3d human pose estimation," *arXiv*, vol. 2501.02771, 2023.

[2] Ultralytics, "Yolo11 pose estimation: Guide," *Ultralytics Blog*, 2024.

[3] M. Mendez, "Image registration in sports analytics," *Miguel Mendez AI Blog*, 2024.

[4] e. a. Pandya, "Homography based player identification in live sports," *CVPRW*, 2023.

[5] YouTube, "Build a football analysis system using yolo11 and supervision." https://www.youtube.com/watch?v=22kkSIxPC14, 2025. Accessed: 2025-04-24.

[6] Ultralytics, "Ball tracking in sports with computer vision," *Ultralytics Blog*, 2024.