Football Match Analysis Using AI

MennaTullah Mohammed Abdelbadea [Computer Science], Faculty of Computer & Information Sciences, Ain Shams University

Nourelhuda Ashraf Abdelhaleem [Computer Science] Faculty of Computer & Information Sciences, Ain Shams University

Ismaeel shreif Ahmed [Computer Science] Faculty of Computer & Information Sciences, Ain Shams University

Ahmed Esmat Ahmed [Computer Science] Faculty of Computer & Information Sciences, Ain Shams University

Ahmed Hany Fathy [Computer Science] Faculty of Computer & Information Sciences, Ain Shams University

Abdelrahman Fayez Mohammed [Computer Science] Faculty of Computer & Information Sciences, Ain Shams University

Dr. Mohammad Essam [Bioinformatics Department]. Faculty of Computer & Information Sciences, Ain Shams University .

Dr. Ahmed Salah, Lecturer, Computer Science Department, Faculty of Computer and Information Sciences, Ain Shams University.

Abstract— Football is a globally beloved sport, and the ability to analyze matches with precision can provide valuable insights for teams, coaches, and fans. The "Football Matches Analysis using AI" project aims to enhance the understanding of football games by leveraging advanced artificial intelligence techniques. This project is designed to address the need for automated, accurate, and detailed analysis of football matches, which traditionally relies on manual methods that can be time-consuming and subjective.

The system developed in this project incorporates several key features: detection and tracking of players and the ball, team mapping, and event detection. By using sophisticated computer vision algorithms and machine learning models, the system can accurately detect and track players and the ball throughout the match. Team mapping algorithms identify the players' team affiliations, while event detection mechanisms highlight significant events such as goals, fouls, and substitutions.

Furthermore, the system includes defined functions to calculate critical metrics such as ball possession and the number of passes for each team. These metrics provide a comprehensive overview of the match dynamics and can be used to generate detailed performance reports.
The final results demonstrate the effectiveness and accuracy of the developed system in analyzing football matches. The AI-based approach not only automates the analysis process but also ensures consistent and objective results, making it a valuable tool for football analysts, coaches, and enthusiasts. The project shows great potential for further enhancements and applications in sports analytic

Introduction

Despite the sport's popularity, there has been limited attention and investment in sports analytics and Computer Vision technology in the league. This project aims to fill this gap by contributing to the advancement of these technologies and achieving acceptable accuracies with limited resources. By doing so, we can enhance the league's performance and make a significant impact on the sport's development. This project is essential because it can help improve the overall quality of the league, attract more fans and investors, and ultimately contribute to the growth of football.

Problem Definition

The problem at hand is the lack of attention and investment in sports analytics and Computer Vision technology football. Despite football being the most popular sport in the world, the research in this area is not given much concern in the open source community. This has resulted in limited resources being allocated to enhance the performance of the leagues. Our aim is to contribute to the advancement of these technologies and achieve acceptable accuracies with the limited resources available. With additional resources.

Objective

Our goal is to use Computer Vision to achieve the following:

1. player-Team identification by Mapping each player to his own team.
2. Continuously tracking the ball with player throughout the entire video clip.
3. Detect happened events like (fouls, corners, penalties, Red/Yellow cards & goals)
4. Generate statistics for each team as passes completed, (successful/missed), shots, yellow/red cards, ball possession, scored goals.

Field and scientific background of the project :

The field of this project related to computer vision, machine learning and deep learning. Specifically, the focus is on YOLO, a type of deep learning algorithm used for to detect objects within images,
CNNs, a type of deep learning architecture particularly well-suited for image recognition tasks.

CNNs used for recognizing and classify objects within images. For example, they can distinguish between different types of ball, player, or any other objects, locate and identify objects within an image or video by drawing bounding boxes around them. YOLO used for real-time object detection in videos and live streams, such as tracking players movements across the field, detect the ball's position.

To understand YOLO and CNNs, it's important to have some knowledge of neural networks and deep learning which are subsets of AI.Neural networks are a class of machine learning algorithms inspired by the structure and function of the human brain and they are mathematical models designed to recognize patterns within data, while deep learning refers to the use of multiple layers of these networks to improve accuracy and efficiency. In recent years, deep learning has led to significant advancements in computer vision tasks such as object detection, image classification.

A survey of the work done in the field

The application of AI in sports analytics has garnered significant interest, leading to numerous research efforts and commercial solutions. Football, being a globally popular sport, has been a focal point for many of these advancements. Key areas of research and development include player and ball tracking, event detection, and the generation of actionable insights from video data.

Various studies have employed convolutional neural networks (CNNs) to detect and track players and the ball. For instance, the YOLO (You Only Look Once) algorithm has been widely used for real-time object detection due to its balance of speed and accuracy. Further advancements include integrating tracking algorithms like SORT (Simple Online and Realtime Tracking) to maintain the identity of players and the ball across video frames.

Identifying events such as goals, fouls, and corners is critical for comprehensive match analysis. Techniques involving Temporal Convolutional Networks (TCNs) and Long Short-Term Memory (LSTM) networks have been effective in analyzing sequences of frames to detect and classify these events. These models capture temporal dependencies, making them suitable for understanding the progression of a match.

Combining statistical analysis with intuitive visual representations enhances the accessibility of insights. Dashboards and interactive tools have been developed to present data such as player heat maps, ball possession statistics, and event timelines. These tools leverage web technologies to offer real-time updates and user-friendly interfaces.

Related Work

Authors	Dataset	Method(s)	Accuracy
Ragd Alhejaily, Rahaf Alhejaily, Mai Almdahrsh, Shareefah Alessai, and Saleh Albelwi.	Common Objects in Context (COCO)	Team Assignment	
		CAE (Type of CNN) +K-means	92%
		Ball Detection	
		Pretrained Yolov5 +TL	98%
James Hong, Haotian Zhang, Michael Gharbi, Matthew Fisher	Competition split (FS-Comp) Soccer Net	E2E-Spot	74.84%
Ruslan Baikulov	Soccer Net	CNN	86.47%
Anthony Cioppa, Silvio Giancola, Adrien Deliege, Le Kang, Xin Zhou	Soccer Net	ByteTrack	50.257%

Existing similar systems

After Searching, we found somecommercial and open-source platforms provide functionalities similar to those of our project:

- Opta Sports: A prominent sports analytics platform that manually collects data to offer detailed statistics and insights into player and team performance. Opta's data is widely used by professional teams, broadcasters, and analysts. Opta Analytics suite of advanced metrics, which measure the quality of passes, shots and individual playing styles, is enabling you to take your analysis of football to the next level by providing underlying insights into the actions of all the key performers on the pitch

- Playmaker AI: This platform utilizes AI to provide detailed football data and insights. Playmaker. AI’s features include scouting and player development tools, player avatars that describe roles more precisely than traditional positions, and comprehensive individual and team reports. The platform supports football clubs, agents, media companies, and researchers by offering extensive data on player performance and facilitating in-depth analysis to enhance strategic decision-making.

- Second Spectrum: Known for its use of AI and machine learning to deliver real-time tracking and analysis, Second Spectrum offers advanced visualizations and insights. It serves both basketball and football, providing detailed analytics for teams and broadcasters.

System Architecture

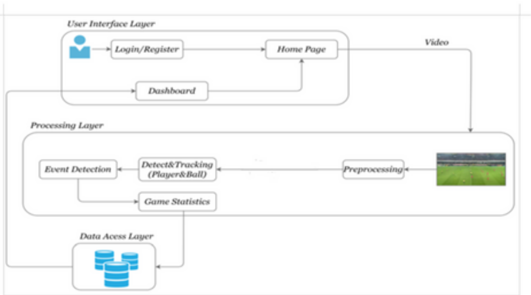


Figure 3.1 - System Architecture

User Interface Layer :

Upon opening the website, users are presented with a home screen where they can choose to log in with existing credentials or register for a new account. After successful authentication, users are directed to the home screen where they have the ability to view and edit their personal information. Additionally, the website offers football video Analysis that allows users to upload a video of a Match and give him dashboard of important events.

Processing Layer :

The input video enter the processing stage in in which video frames are read one by one and then start detecting and tracking object in each frame and detect if there is an event in the frame, then we store analytics we extract from the frame to the memory, for the event detection part, we crop the most important event on the match as a match highlights, after repeating the process to the end of the frames then we start saving the analytics of the whole video and present it in the game statistics dashboard

Data Access Layer :

It is the layer in which the website's database is located. This database holds the information of the registeredusers, includingthe dashboards and saved highlights of the match.

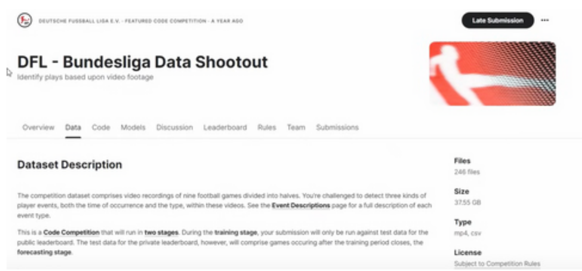
Datasets:

1- Soccernet :

The Soccernet dataset is a comprehensive collection of data related to soccer (football) matches. It is often used for research in computer vision, machine learning, and sports analytics, The dataset is composed of 500 complete soccer games from six main European leagues, covering three seasons from 2014 to 2017 and a total duration of 764 hours

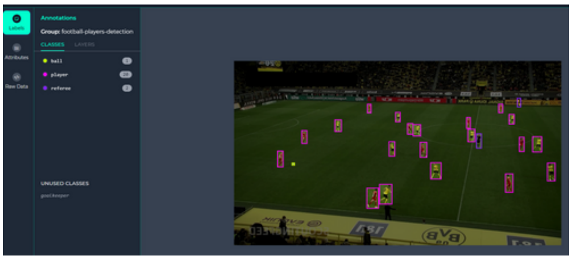
2- DFL - Bundesliga Data Shootout

- dataset in a Kaggle competition for football
- it consists of 200 high quality clips of 30 seconds
- The dataset covers various matches filmed with different cameras



3- DFL - Bundesliga Data Shootout

- consist of 663 total images, 612 for training, 38 for validation and 13 images for testing.
- We use this dataset to finetune our YOLO model to better identify players, referees and ball.



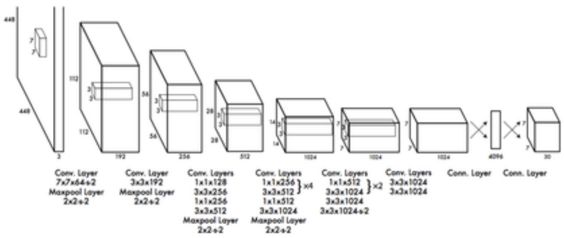
Preprocessing:

- Frame Sampling to reduce computation time for the video.
- Resizing video frames into 448x448 for YOLO.
- Converting frames from BGR format to RGB format.
- Stacking frames to create a single, more informative input for a model. This approach helps in capturing temporal information and motion, which can be crucial for tasks that depend on understanding changes over time.

Techniques, Models and Algorithms used:

- Object Detection :
  - Yolov5 algorithm (is very fast—it runs at 45 frames per second). Yolov5 model was trained on the Common Objects in Context (COCO)-> detect 80 different obj.
  - used weights trained on the COCO dataset(123,000 images in the sports ball class) ,no need for training step.
  - The model was trained and fine-tuned using a combination of footballs and cricket balls of different sizes to increase its accuracy and differentiate between referees, players, goalkeepers, and normal viewers.

YOLO Model Architecture



- Resizes the input image into 448x448 before going through the convolutional network.
- A 1x1 convolution is first applied to reduce the number of channels, which is then followed by a 3x3 convolution to generate a cuboidal output.
- The activation function under the hood is ReLU, except for the final layer, which uses a linear activation function.
- Some additional techniques, such as batch normalization and dropout, respectively regularize the model and prevent it from overfitting.

- **Tracking Using ByteTrack :**
  - ByteTrack is a multi-object tracking algorithm.
  - It takes object detections from YOLO (or other detectors) and links them across video frames to create consistent tracks for each object.
  - It handles occlusions, disappearances, and new object appearances.

Model Architecture

- Input: A list of detections from YOLO (same format as output above).
- Output: A list of tracks, each containing:
  - Object ID (unique identifier for each tracked object).
  - Bounding box information for each frame the object appears in.
  - Additional information (optional): velocity, trajectory, etc
- Two-Step Matching: ByteTrack uses a two-step approach to associate detections with tracks.
- Step 1: High-Confidence Matching:
  - Uses Kalman filter to predict object positions in the next frame.
  - Matches predicted positions with high-confidence detections using IoU (Intersection over Union) for motion similarity.
  - IoU reflects the overlap between bounding boxes.
- Step 2: Low-Confidence Matching:
- Objects not matched in step 1 (potentially occluded) are considered.
- These are matched with remaining detections, even those with lower confidence scores.
- This helps track objects during occlusions where confidence might drop.

Team Mapping :

Our implementation of a team mapping mechanism for football matches using AI represents a significant advancement in sports analytics. By refining the algorithm to process video frame by frame rather than in its entirety, we achieved notable efficiency gains. This approach enabled precise calculations of possession statistics for each team, including the number of passes executed, as well as individual player possession metrics. These developments mark a substantial enhancement in our ability to extract detailed insights from match footage, demonstrating the transformative potential of AI in sports analysis.

Event Detection :

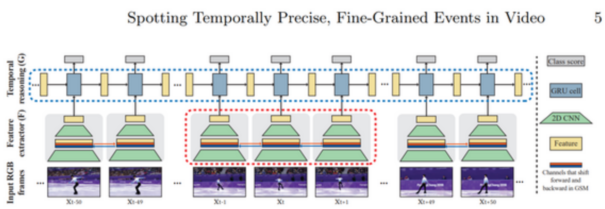
JHONG Action Spotting :

- Introducing the task of spotting temporally precise, fine-grained events in video, which involves detecting the exact moment events occur in a video

- Precise spotting refers to the task of accurately localizing specific events or actions within video sequences with high temporal precision. Unlike traditional video action recognition, which focuses on identifying the presence of actions in videos, precise spotting aims to pinpoint the exact moments when these actions occur.

- Use 2 Data sets (Soccer Net , Sports Dataset Compilation).

Model Architecture



1. Local Spatial-Temporal Feature Extractor (F):

- Extracts spatial-temporal features from each frame in the video sequence.
- Utilizes a lightweight 2D Convolutional Neural Network (CNN) with Gate Shift Modules (GSM) to capture subtle motion and visual differences between neighboring frames.
- Ensures that each frame's features capture both spatial and temporal information crucial for event recognition.

2. Long-term Temporal Reasoning Module (G):

- Utilizes a 1-layer bidirectional Gated Recurrent Unit (GRU) network to gather long-term temporal information.
- Processes dense per-frame features obtained from the feature extractor to capture temporal dependencies and context across the entire video sequence.
- Outputs class predictions for each frame, including a 'background' class for frames where no event is detected.

3. Per-frame Cross-Entropy Loss:

- Optimizes the model's classification performance using per-frame cross-entropy loss.
- Compares the model's predictions with ground-truth labels for each frame and computes the loss based on the classification error.
- Helps in training the model to accurately spot events in the video sequence.

Experimental Results:

Evaluation:

Algorithm	Setup	HOTA	DetA	AssA	MOTA
DeepSORT	w/ GT	69.552	82.628	58.668	**94.844**
FairMOT	w/ GT	-	-	-	-
ByteTrack	w/ GT	**71.500**	**84.342**	**60.718**	94.572
DeepSORT	w/o GT	36.663	40.022	33.759	33.913
FairMOT	w/o GT	43.911	46.317	41.778	50.698
ByteTrack	w/o GT	47.225	44.489	50.257	31.741
FairMOT-ft	w/o GT	**57.882**	**66.565**	**50.492**	**83.565**

Figure 4.6 comparison between bytetrack and other open-source models

Result:
After applying tracking and detection the output of each video frame be like:

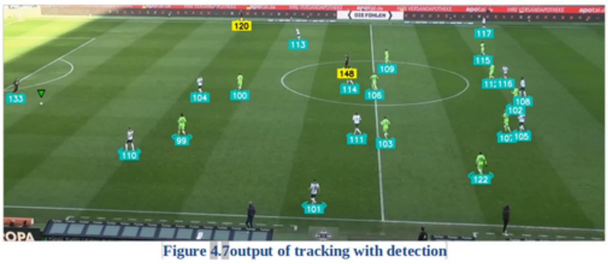


Figure 4.7output of tracking with detection

Evaluation:

Table 4: **Average-mAP @ t for tolerances in seconds.** SOTA in **bold**. We show the top results from the CVPR 2021 and 2022 SoccerNet Action Spotting challenges. ‡ indicates challenge results — trained on the train, validation, and test splits. Shown and unshown refer to whether actions are visible; E2E-Spot is better at detecting the former, but Soares et al. [51] is superior at the latter.

	Test split		Challenge split		
Average-mAP @ tolerances	Tight (1–5 s)	Loose (5–60 s)	Tight (1–5 s)	Shown	Unshown
RMS-Net [57]	28.83	63.49	27.69	-	-
NetVLAD++ [22]	-	-	43.99	-	-
Zhou et al. [71] (2021 challenge; 1st)	47.05	73.77	49.56	54.42	45.42
‡Soares et al. [51] (2022 challenge; 1st)	-	-	‡**67.81**	‡72.84	‡**60.17**
E2E-Spot 200MF	61.19	73.25	63.28	70.41	45.98
E2E-Spot 800MF	61.82	74.05	66.01	72.76	51.65
‡E2E-Spot 800MF (2022 challenge; 2nd)	-	-	‡66.73	‡**74.84**	‡53.21

Figure 4.9 action spotting model compared with other open-source models

Penalty	0.6117246389973663	Ball out of play	0.7597325074474494
Kick-off	0.1914447608875885	Throw-in	0.739045113850471 4
Goal	0.747110202356527	Foul	0.3319909517860364 4
Substitution	0.3160992350081842 7	Indirect free-kick	0.2229626790874183 8
Offside	0.2952431938061168 5	Direct free-kick	0.5463329480875319
Shots on target	0.3955649021521312 5	Corner	0.830785950189399 7
Shots off target	0.4949210089405469	Yellow card	0.6416569714054904
Clearance	0.3591693015295976	Red card	0.0
		Yellow->red card	0.0

Figure 4.10 accuracy per event class

References :

- Hong, J., Zhang, H., Gharbi, M., Fisher, M., and Fatahalian, K., "Spotting Temporally Precise, Fine-Grained Events in Video," in Proceedings of the European Conference on Computer Vision (ECCV), 2022.
- Iromul, "Ball Action Spotting," in Proceedings of the European Conference on Computer Vision (ECCV), 2023.
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A., "You Only Look Once: Unified, Real-Time Object Detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788.
- Zhang, Y., Wang, C., Wang, X., Zeng, W., and Liu, W., "ByteTrack: Multi-Object Tracking by Associating Every Detection Box," in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2021, pp. 1516-1525.
- Wojke, N., Bewley, A., and Paulus, D., "Simple Online and Realtime Tracking with a Deep Association Metric," in Proceedings of the IEEE International Conference on Image Processing (ICIP), 2017, pp. 3645-3649.