Thesis Title: Computer Vision-Based System for Football Player Tracking and Performance Analysis

Submitted By

Mohiduzzaman

ID: 221-35-854

Department of Software Engineering

Daffodil International University

Supervised By

MD. SHOHEL ARMAN

Assistant Professor

Department of Software Engineering

Daffodil International University

A thesis submitted in partial fulfillment of the requirement for the degree of Bachelor of Science in Software Engineering

Fall 2025

APPROVAL

This thesis titled on **"Computer Vision-Based System for Football Player Tracking and Performance Analysis"**, submitted by **Mohiduzzaman (221-35-854)** to the Department of Software Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Software Engineering and approval as to its style and contents.

BOARD OF EXAMINERS

-- **Chairman**

Dr. Imran Mahmud
Professor & Head
Department of Software Engineering
Faculty of Science and Information Technology Daffodil International University

-- **Internal Examiner 1**

Afsana Begum
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

-- **Internal Examiner 2**

Md. Shohel Arman
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

-- **Internal Examiner 3**

Nadira Islam
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

-- **External Examiner**

Md Manowarul Islam
Professor
Department of Computer Science and Engineering
Jagannath University, Bangladesh

Department of Software Engineering

Faculty of Science and Information Technology

Supervisor Approval Form

Fall 2025	B.Sc. In SWE	Campus: DSC

Student Name	Student ID
Mohiduzzaman	221-35-854

Project/Thesis Information	
Project/Thesis Title	Computer Vision-Based System for Football Player Tracking and Performance Analysis
Type of work	Thesis

Supervisor information	
Supervisor Name	Md. Shohel Arman
Supervisor Initial	MSA
Completed Credit till now	133
How many credits in this semester	6
Supervisor Consent	☐ Yes ☐ No

Supervisor Signature

DAFFODIL INTERNATIONAL UNIVERSITY

DECLARATION OF THESIS AND COPYRIGHT

Author's Full Name	:	Mohiduzzaman
Date of Birth	:	14 February 2002
Title	:	Computer Vision-Based System for Football Player Tracking and Performance Analysis
Academic Session	:	2022-2025

I declare that this thesis is classified as:

☐	CONFIDENTIAL	(Contains confidential information under the Official Secret Act 1997)*
☐	RESTRICTED	(Contains restricted information as specified by the organization where research was done)*
	OPEN ACCESS	I agree that my thesis to be published as online open access (Full Text)

I acknowledge that Daffodil International University reserves the following rights:

1. The Thesis is the Property of Daffodil International University.
2. The Library of Daffodil International University has the right to make copies of the thesis for the purpose of research only.
3. The Library of Daffodil International University has the right to make copies of the thesis for academic exchange.

Certified by:

(Student's Signature)	(Supervisor's Signature)
221-35-854	Md. Shohel Arman
Student ID	Name of Supervisor
Date: 25 November 2025	Date: 25 November 2025

SUPERVISOR'S DECLARATION

I hereby declare that I have checked this thesis and in my opinion, this thesis is adequate in terms of scope and quality for the award of the degree of Bachelor of Science.

____

(Supervisor's Signature)

Full Name : Md. Shohel Arman

Position : Assistant Professor

Date : 25 November 2025

STUDENT’S DECLARATION

I hereby declare that the work in this thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Daffodil International University or any other institution.

(Student’s Signature)

Full Name : Mohiduzzaman

ID Number : 221-35-854

Date : 25 November 2025

Computer Vision-Based System for Football Player Tracking Performance Analysis

Mohiduzzaman

Thesis submitted in fullfilment of requirements
For the award of degree of
Bachelor science

Department of Software Engineering

DAFFODIL INTERNATIONAL UNIVERSITY

December 2025

Acknowledgement

First and foremost, I want to thank God for all the blessings, guidance, and mercy that have helped me get to this point in my academic path. The conclusion of this research would not have been feasible without His favor.

I want to thank my supervisor, Md. Shohel Arman, Assistant Professor in the Department of Software Engineering at Daffodil International University, from the bottom of my heart. He told me the main ideas of this paper and helped me with the investigation. I am really thankful for the chance to work under his very smart supervision and want to thank him for all the help he gave me.

I appreciate God for showing me the way and letting me work toward good goals. I am really thankful to my parents for always being there for me. Their support has been a big part of how I got to this point. I would want to thank Dr. Imran Mahmud, the Head of the Department of Software Engineering, and all of the other great teachers who have helped me learn so much. Daffodil International University gave me some great advice, and I'm thankful for it.

I also want to thank my parents, friends, and others I know. I would not have been able to keep going on this trip without their help.

Abstract

The study of football is increasingly taking up a larger portion of contemporary sports intelligence. It assists in obtaining useful information on videos of games to teams, analysts, and researchers. When you do traditional manual analysis, it is difficult to understand how players behave, how tactics are worked out, and how the game may be changed in real time due to the long time of analysis, the ease of error and the impossibility to scale up. The paper outlines a fully computer vision-based system of automatic football match analysis. It involves object recognition, tracking, motion compensation of the camera and the process of assigning players to teams and mining metrics. The framework employs a detection framework based on Yolo to locate the ball, players and referees in video frames. Subsequently, powerful tracking system monitors item IDs with time and estimation of camera movement stabilises the analysis process by compensating the effect of panning and zooming. The system also possesses a view-transformation module which represents a top-down view of the field. This gives you the ability to quantify the speed, range covered by players, the ball action and relative positions of the players. The algorithm is based on color and spatial hints to identify rivals and determine how to assault them in the most favorable manner. Tests indicate that the given framework is capable of detecting and tracking objects specifically and continuing to do so even when the lighting, camera angles, and the number of players vary. Quantitative findings indicate that, stability tracking has been enhanced, identity transition has reduced and metric extraction is now predictable and therefore easier to perform phase-wise tactical analysis. This suggested system proves to be superior and is more objective to perform manual annotation. It also provides you with helpful features to determine your progress, coaching decision making, match intelligence, and future sports analytics.

Table Of Contents

Table Of Figures

Chapter 1

Introduction

1.1 Introduction

The most popular sport in the world is football and with the advancement of technology in the sports and digital delivery systems, there has been a surge in the research of football games being carried out. The data-driven insights are increasingly being used by professional teams, coaches, and analysts who are attempting to determine how players move, how they perform, and how they handle tactical situations. People used to acquire these insights manually, but it is difficult, subjective and usually erroneous to acquire them manually because people commit errors, and do not concentrate long enough. It has improved computer vision (CV) and artificial intelligence (AI) that have permitted the automatic extraction of meaningful information about football match film in real time. Sports analytics systems made with AI are able to identify individual players, assign them a unique identity, track their movements, determine their distance covered, assess the composition of their team, and assist with more advanced tactical analysis. In this thesis, we introduce Football Analyzer which is an AI implementation that examines football game footage through computer vision and object detection techniques. The system uses the OpenCV tracking algorithms to maintain the position of where players are going, YOLO (You Only Look Once) to identify players, and geometrical methods to determine the speed at which players are moving, the distance they have moved, and field setup. The primary objective of the project should be to create a simple and lightweight system that is capable of viewing match videos without the need to possess special software and costly hardware. Football Analyzer automatically retrieves information on the location and movements of players which has allowed the software to be easy to understand as the user can know how players are performing and how the match is progressing.

1.2 Background

Over the past decade, there are numerous transformations in the field of sports analytics. In the present days, the leading football teams possess very extensive analytics departments that assist them in understanding match information and decision-making. Such technologies as GPS trackers, RFID sensors, drone cameras, and special high-speed tracking systems (such as TRACAB and Catapult Sports) have entirely transformed the analysis of professional sports. However, these tools are very expensive, and do not exist in small groups or students. The creation of the computer vision solutions capable of analyzing match footage without the use of costly equipment or wearable gadgets occurred due to this issue and the invention of new methods in the field of deep learning, particularly, convolutional neural network (CNNs).The operation of real-time object identification models such as YOLO became possible. YOLO is able to detect many players simultaneously, even during the moments when the action is rather quick. Combined with tracking algorithms such as SORT, DeepSORT, and OpenCV trackers, these models allow one to track all players in the game.Football Analyzer expands on these concepts to develop a systematic pipeline that is able to: Find players frame by frame Giving out unique IDs Watching movement over time Knowing the speed and the distance. Taking a glance at the way the team is arranged. Analytics to make you understand. This approach allows football analytics to be performed in modern times without having to use expensive technology.

1.3 Motivation

The project is founded on three concepts: Football Analysis is Time Consuming when Manual. With a 90 minute game, it may take over 6 to 8 hours to watch and take notes. The analysts are forced to manually track the players, locate events, highlight

key times, and export data. This is not possible when it comes to normal analysis. The tools should be difficult to access and yet cost-effective. A lot of professional systems require special cameras which could be costly in thousands of dollars. The solution must be cheap and able to take simple video input by students, analysts and amateur organizations. How AI is changing sports Businesses are being influenced throughout the world by AI and computer vision. These technologies can be used in football as an excellent method of learning new AI techniques and achieving beneficial results. Making Football Analyzer is a solution to a real challenge and will provide you with practical knowledge in the sphere of computer vision, video processing, and algorithm development.

1.4 Problem Statement

The majority of the football analysis techniques in the modern world rely on both costly equipment and a significant amount of physical labor. We require an automated system powered by AI, which is able to: Constituent names: no longer needed to write them down. There is movement throughout, even when the camera is moving. Discover such things as distance and speed that indicate whether you are on track or not. Relate the regions that have players and their location. Training of regular match footage using no additional sensors. In the absence of such a system, coaches and analysts cannot get the timely and reliable information due to the fact that analysis remains slow, subjective, and limited. The proposed project will fill that gap by developing an automated football analysis program based on the computer vision methods.

1.5 Research Question

This thesis will answer the following important questions: To what extent is it possible to monitor and detect a high number of football players on the field using computer vision? Is it possible that constant player IDs and access to a reliable real-time detection can be made possible by Can YOLO with OpenCV? What does video all by itself reveal about movement, distance travelled and speed? What is the extent of tactical information that you can receive by merely watching videos? What are the existing flaws of these systems and how can they be fixed?

1.6 Research Objective

In this thesis, the following are important questions that are answered: To what extent can the computer vision observe and track several football players in a real game? Can YOLO and OpenCV be used to achieve consistent player IDs and real-time detection? How do you know anything about movement, distance travelled and speed by simply viewing video? What about the amount of tactical information you can obtain by watching videos? What are the existing issues with these systems and how can they be improved?

1.7 Research Scope

The Football Analyzer technology only concentrates on video based 2D analysis. The scope is limited to: Players and ball recognition. Monitoring them between the frames. Plotting heatmaps of movement, distance and speed. Giving consideration to the distribution of the team. Video of the match or normal broadcasting. Not included in this scope: Making such predictions as goals, passes, and shots. The tactical analysis (xG models, pushing patterns) is a lot. Making an approximation or a 3D image. Automatic correction of the speed in meters/second.

1.8 Thesis Organization

The following chapters comprise this thesis: Chapter 1: Discussion of the origin of the book, the reason behind writing of the book, the problem and the objectives. Chapter 2: Discusses additional literature on sports video analytics and the way to find things. Chapter 3: Discusses the entire procedure and the implementation. Chapter 4: The discussion of results of experiments and their presentation. Chapter 5: Discusses the limitations of the work and how this would be improved in future to complete the work.

1.9 Summary

This chapter discussed our reason as to why we wanted to develop a football analyzing system using AI. It discussed the issues of conventional sports analysis, role of computer vision, objectives and research areas on which the present project is founded. This will be provided in the following chapter that will provide a comprehensive literature review including previous publications, models of object detection, tracking methods and sports analytics studies.

Chapter 2

1. Literature review

2.1 Introduction

In this chapter, the research, tools, and methods available in existence in sports analytics using computer vision are well looked into. It illustrates the extent to which we have advanced in terms of football analytics systems, player detection, tracking objects, and machine learning models which are commonly employed in video analysis. This review is aimed at determining what is good and bad about the current work and the place of Football Analyzer among other works devoted to the technology of sports.

2.2 Literature Review

Deep learning and computer vision have advanced at a very rapid rate and have transformed the manner in which football match analysis is conducted in a significant manner. Traditional football analytics were reliant on a lot of hand-written notes, analyst observations, and post-game film breakdowns. These old methods are useful, but they are time and labor intensive and only suitable with short cut match video. Over the last few years, the focus of researchers has become gaining more and more on the topic of automating the interpretation of football videos with the help of such methods as object detection, multi-object tracking, camera calibration, event spotting, and spatiotemporal feature extraction. This body of work is the foundation of modern football analytics systems and it is directly involved in the development of computational tools such as Football Analyzer. The SoccerNet package of datasets has already left a significant mark in this topic and has now become the benchmark of automated football video analysis. According to Giancola et al. (2018), SoccerNet is a comprehensive dataset of 500 full matches marked with notable broadcast incidents, such as goals, substitutes, and cards. This study gave the concept of action spotting, which does not only have the aim of classifying events but also of identifying the precise moments of their occurrence in long movies. The paper pointed at the issues of broadcasting football video including rapid changes in view, the small size of items, busy backgrounds, and huge occlusions. The scale of the dataset and the annotations of the events allowed performing supervised learning on real broadcast games. That allowed comparing various deep learning models in a systematic manner. SoccerNet soon emerged as the standard testing environment of event detection systems and it left a substantial influence on the development of automated sports analytics. Deliege et al. (2021) follow this work and published SoccerNet-v2 which introduced approximately 300,000 new annotations into the dataset. These are re-play segments, camera views, referee whistles, and other semantic information that enables a better view about the progress of the contest. There were a number of novel benchmark tasks in the study, such as accurate action recognition by tighter temporal performance, replay anchoring, and camera-boundary identification. These extensions enhanced the data set to predict multi phase match structures and broadcast workflows. The added annotations improved the temporal reasoning models by providing constant context during the match, which consequently enabled algorithms to distinguish better between times that are visually similar and semantically different. SoccerNet-v2 increased the scope of the soccer video research beyond the successful identification of events to video analysis as a context. The other significant field of study has been to make event spotting more precise. Cioppa et al. (2020) proposed a context-aware loss function which involves data on time surrounding the event rather than only the time of the event. Their approach realizes that the contextual indicators, like build ups before a goal or replaying before or after an event are extremely significant in determining the right moment. The plan reduces the confusion that could occur in the hectic gameplay and the visual jumps by recreating the visual environment surrounding the game. This paper found that time context of an event is also important as the event itself, which provides a new framework through which long-range video analysis can evolve loss functions. Temporal modeling was further improved by Giancola and Ghanem (2021) who designed NetVLAD++ a feature pooling architecture that understands time and was designed specifically for action detection. The strategy does not consider video context to be equally applicable to all events; rather, there are before and after characteristics. Such difference demonstrates that the actual match conditions are not necessarily similar pre- and post-event.

NetVLAD++ is capable of learning distinct representations in each scenario, increasing its temporal performance of better results and leading to the best results on SoccerNet benchmarks. The current work highlights the need to have specific architectures that are capable of reflecting the nature of football events, as opposed to relying on generic video categorisation algorithms. In addition to the discovery of the events in time, the understanding of space is now one of the significant study problems. When the camera is in motion, the zoom level is being changed and the points of view are different, it is difficult to match the visual detections to real world pitch coordinates when the video is broadcasted. Cioppa et al. (2021) addressed this issue by proposing a deep learning camera calibration method in soccer broadcasting. Their algorithm forecasts the homography parameters through which 2D image coordinates can be transformed into field coordinates in the top view. Some of the methods to indicate player locations were also investigated, including heatmaps and spatial frameworks based on graphs. Precise calibration allows performing such significant analytical tasks as monitoring of players on the field, recreating tactical shapes, and determining the placement of teams. The paper presents the relevance of geometrical arguments in football analytics and provides essential tools in combining a broadcast video with spatial measurements. Another essential input to the field was presented by Cioppa et al. (2022), SoccerNet-Tracking. This data set developed an entire criterion of multiple object tracking (MOT) specifically designed to work with game shots of football matches. It contains an instructional text to the players, the ball, and team names in both short videos and a complete half of a match that lasts 45 minutes. The research demonstrates the difficulty of following soccer players in a broadcast as they tend to intertwine, move rapidly and turn in any direction without prior notice. There is also a lot of movement in the camera which makes it difficult to maintain the visuals. Comparisons of several state of art MOT methods revealed that long term tracking, identity switching, robust ball tracking and dealing with occlusions remain challenges. Their results highlight the fact, that despite the progress in generic multi-object tracking, football remains a highly challenging task, which requires tracking solutions specific to the domain.

A more recent advancement was made by Somers et al. (2025) who developed a player tracking system that is highly precise and uses motion vectors in conjunction with instance segmentation. Their solution addresses the issue of identifying players consistently during the match and simplifies the process of locating players when using moving cameras. This paper suggests a growing interest in finer-sized spatial accuracy, which is now vital to the tactical analysis, model of ball progression, and advanced measures of performance.

All of these past works together, however, form a good case as to why computer vision should be used to automatically analyze football matches. The development of SoccerNet to SoccerNet-v2 and SoccerNet-Tracking marks the clear shift to the higher level of annotations, detailed contextual information and complex benchmarks. Improved temporal modelling, context sensitive loss functions and feature pooling have facilitated easier event recognition. Space reasoning has also been made easier by better calibration and tracking. Despite such advances, there remain large issues to be addressed, in particular the ability to do long-term tracking under occlusion, to integrate spatial and temporal reasoning and to address the fluid dynamics of real matches. These shortcomings highlight the value of continued investigation into end-to-end systems that would enable the understanding of matches comprehensively an objective that is evidently aligned with the goals of the present thesis research.

Chapter 3

2. Methodology

3.1 Introduction

The Football Analyzer framework is supposed to allow you to examine the videos of football matches in a systematic and automated fashion. It applies the most modern computer vision techniques to obtain valuable data regarding the movements and formations of the players, ball control, and performance indicators. The structure is established as a modular pipeline, so every component of the system is capable of functioning independently and yet contributing to the overall functionality of the system and that includes detection, tracking, geographical mapping, analysis, and visualization.

The system was established to address the issues that arise when television coverage of football footage takes place such as when the cameras are moving, the players are fast moving, obstructions and small or fast-moving objects such as the ball. Object detection, multi-object tracking, homography-based field mapping, and team-specific analysis are used to transform raw video frames into helpful tactical and performance data offered by the system. The modular nature ensures that it is simple to scale, work with the standard hardware and can add new functions or integrate with other modules of analytics at a later date.

It is easy to observe the order of the workflow using this framework, that is, video input - detection - tracking - feature extraction - analysis - visualization. This preconditions powerful, real-time, and easy-to-conceive football match analytics.

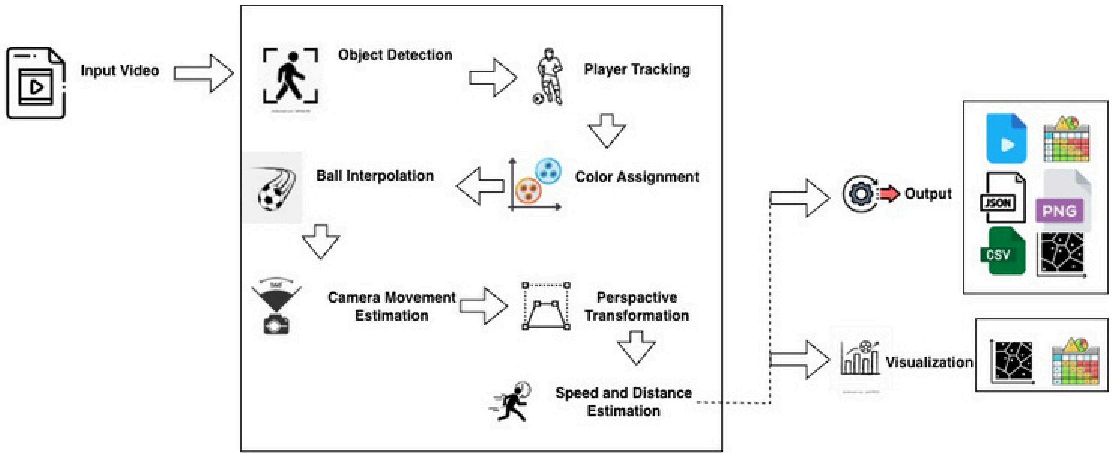

Figure 3.1: Methodology

3.2 Data Collection

The Football Analyzer system also needs good records of football matches so that it can locate, track and analyze objects at high precision. Match films are acquired as recordings of televised events, open-source databases and sample match videos in a common format, either MP4 or AVI. Videos are chosen in order to show both professional and semi-professional competitions with the inclusion of all the perspectives, lighting conditions, and competition levels. Each video is processed to the same level by adjusting the frame rates, resolutions and color profiles. The modules of detection and tracking can readily obtain regular input. Besides the video frames, other metadata like the length of the match, teams structure, dimension of the pitch are also received in instance they are available. This helps in space and tactical analysis. This systematic method of data collection will ensure that the framework can be effective in other match events and help to gain effective performance indicators.

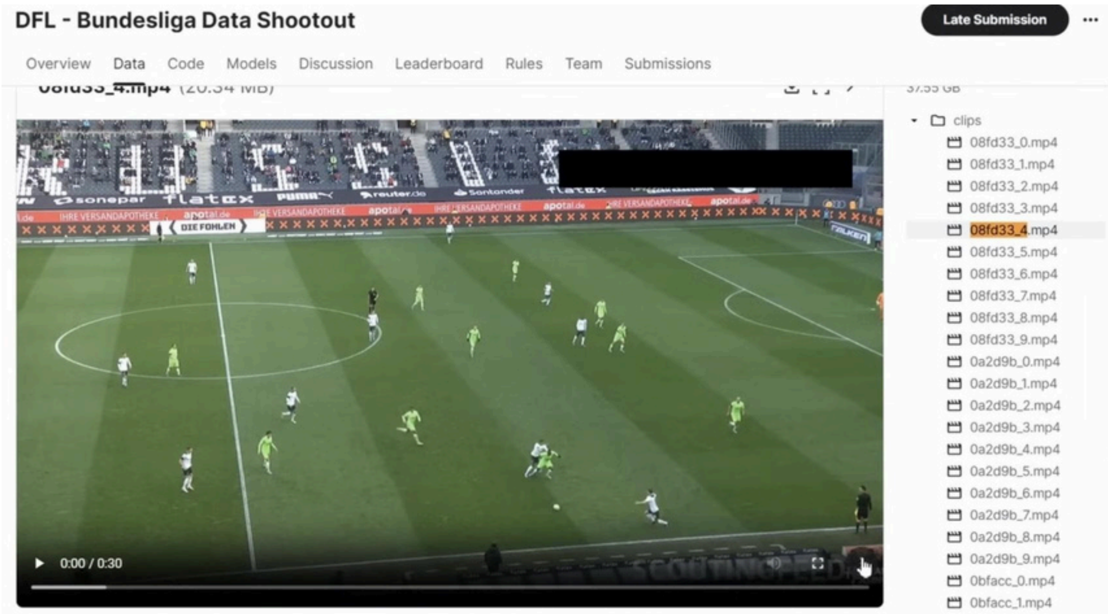

Figure 3.2: DFL-Bundesliga Data Shootout clips

3.3 Video Input and Preprocessing

The standard video file is the input of the system. The quality and consistency of detection is enhanced by preprocessing.

3.3.1 Frame Extraction

The OpenCV is used to divide the video into single frames.

This enables the detection model to process each frame separately.

3.3.2 Frame Resizing

Frames can be downsampled (e.g., 640×480 or 1280×720) to the formats needed by YOLO and decrease the time spent on computing.

3.3.3 Noise Reduction

Simple image processing skills, such as the use of the Gaussian blur or sharpening filters enhance the perception of lower quality images.

3.3.4 Frame Rate Adjustment

In case videos are long or excessive high-speed, frames may be sampled (e.g. processing each 2 nd or 3 rd frame) to trade off performance and speed.

3.3.4 Frame Rate Adjustment

If videos are too long or too fast, frames can be sampled (e.g., processing every 2nd or 3rd frame) to balance performance and speed.

3.4 Yolo-Based Object Detection

The fundamental detection model utilized in this system is YOLO since it has a good speed-accuracy balance.

3.4.1 Why YOLO?

YOLO will be used in this study since it can be used to identify multiple football players at the same time with a high level of accuracy. It is also resilient to changes in the lighting conditions as well as camera movements, which are typical to broadcast football videos. Moreover, YOLO is well optimized to run on a relatively simple hardware and can be applied in practice. It is also smoothly integrated with OpenCV, which also makes the implementation pipeline easier. When applied to a football match, YOLO can identify the crucial objects (players, referees, and the ball, should the model be trained with a special ball category) and is highly pertinent to the automated match cognition and analysis.

3.4.2 Detection Process

For each frame:

YOLO processes the image

Bounding boxes are produced.

The score of confidence is used.

Player labels are assigned

One such detection output produced by the model is a bounding box of coordinates that describe the spatial position of the object detected in the frame, a class.id that describes what type of object was detected, a confidence score between 0 and 1 that describes how much the model is sure that the object was detected, and a label like a player or a ball to say what kind of object it is.

3.4.3 Post-Processing

Once the bounding boxes are detected, the system prunes the bounding boxes by eliminating detections of low confidence, merging overlapping detections and discarding the noise due to false detections. This preprocessing is useful to enhance accuracy and stability of the further player tracking.

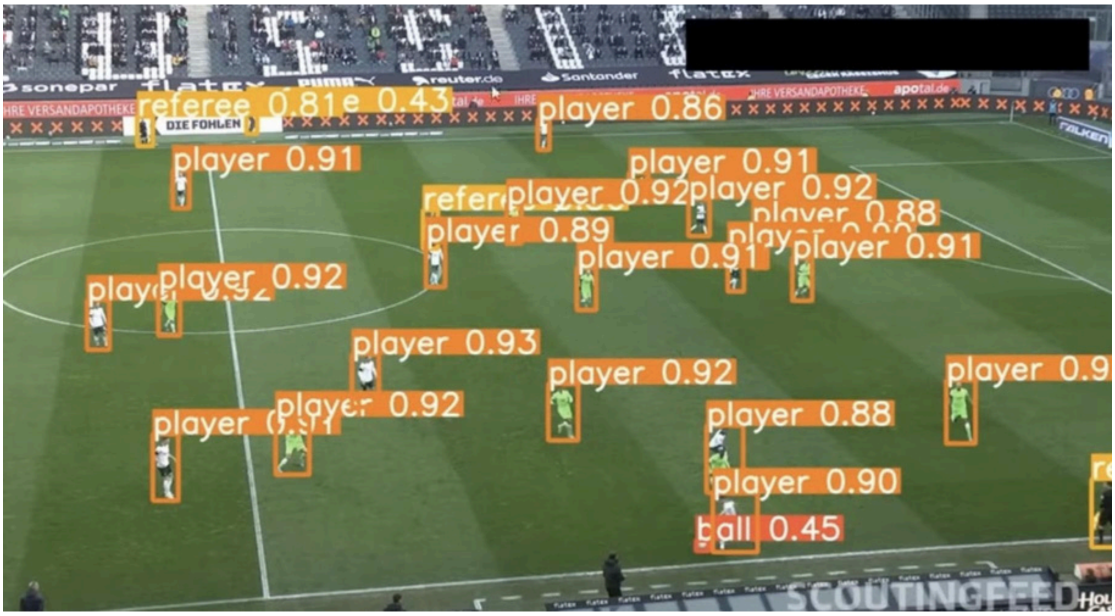

Figure 3.4.3: Yolo Object Detection

3.5 Player Tracking

Frame to frame player identity cannot be determined by detection.

Tracking assigns a unique ID to every detected player which is the same throughout the video.

The algorithm used to track the object must be selected.<|human|>3.5.1 Tracking Algorithm Choice.

The tracking is based on OpenCV because it is a lightweight infrastructure and can therefore be deployed on the machines of the student level, and it is also simple to implement in the entire pipeline. OpenCV has a number of widely used tracking algorithms with varying performance properties: KCF is fast but suffers when there is an occlusion, CSRT is slower but more accurately tracks and is more suitable in sports video, MOSSE is as fast as possible and less accurate, and MIL offers a trade-off between speed and accuracy.

3.5.2 Detection + Tracking Hybrid Approach.

The hybrid pipeline is used in the system:

YOLO operates on a periodical basis (e.g. every 5 or 10 frames) and Trackers retain identity across detections. This enhances the speed as well as accuracy.

3.5.3 Avoiding ID Switching

In order to reduce switching of identity (ID) switching during player overlays and during high speed movement, Intersection-over-Union (IoU) matching, distance thresholding and re-initialization of the tracker are used together. Directional matching IoU has been used to correlate detections across frames as based on spatial overlap, distance thresholding has been used to assure the same object association using positional proximity, and tracker reinitialization has been used to correct tracking drift, which all contribute to tracking stability and robustness.

The fourth stage is the identification and assignment of the team.

Once players are identified, team identification is then done by the system to ensure that the two conflicting teams, which are on the field, are identified. This will be necessary to examine the movement patterns, formation, and spatial distributions of teams during the match.

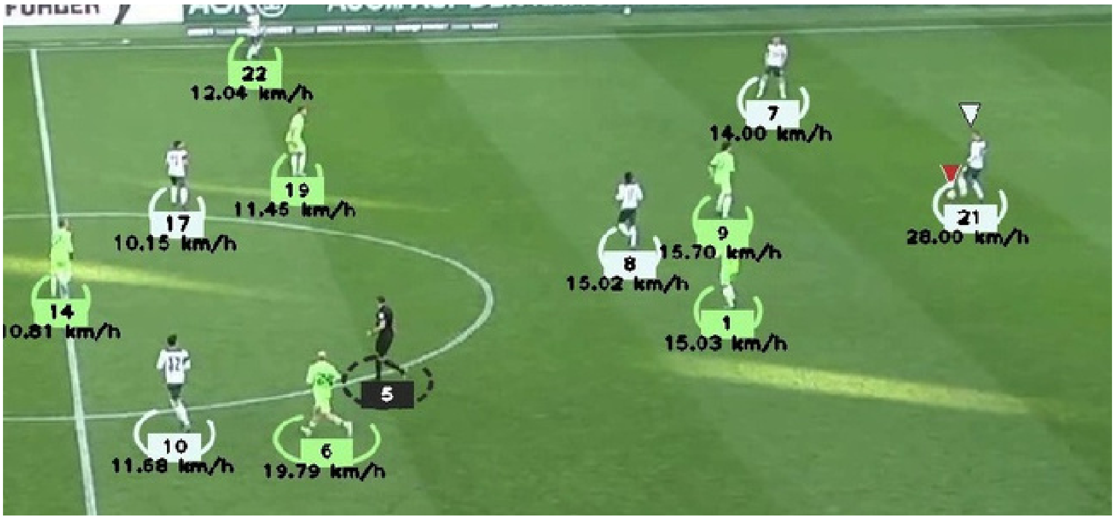

Figure 3.5.3: Tracking Players by unique id

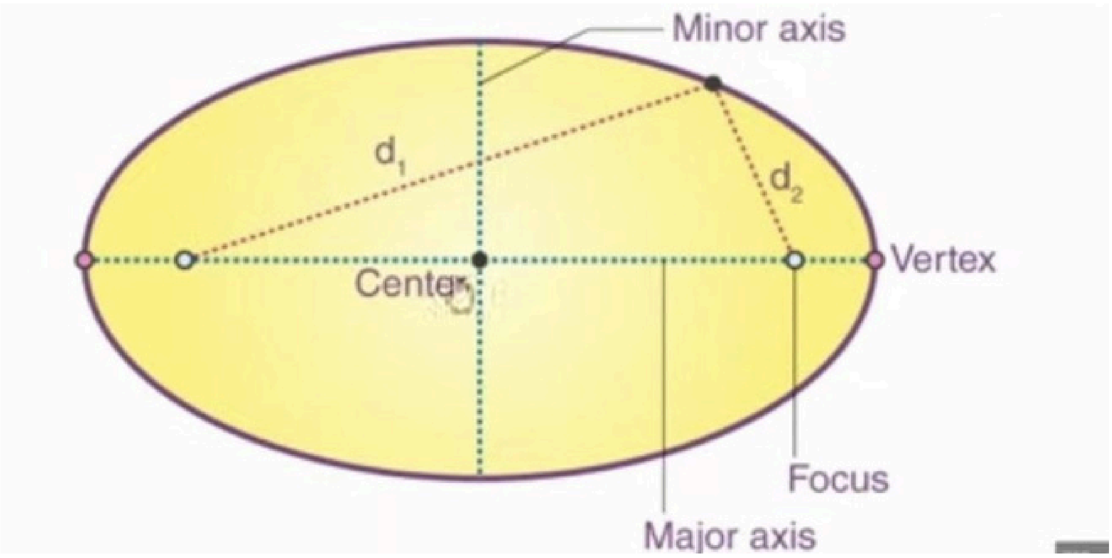

Figure 3.5.4: Ellipse diagram

3.6 Team Identification and Assignment

Based on the bounding box of each player identified, HSV color model, color histograms, and the k-means clustering are used to obtain the predominant jersey color. The HSV color space is also used since it is the best in isolating color data and brightness which is a stronger system in respect to the lighting variation and shadows, mostly encountered in football video.

3.6.2 Team Clustering

Based on the color features extracted, a clustering algorithm, like k-means (k = 2) is used where the players are clustered into two clusters, one of which corresponds to a team. This is an unmonitored method that enables the automatic team division without being aware of the colors to use in jerseys.

3.6.3 Label Assignment

Each tracked player ID is also given a label of Team A or Team B based on the results of the clustering. The referee is identified as different in case a clear jersey color is identified. This labeling allows the steady monitoring of formations of teams, player allocations and tactical formations with time.

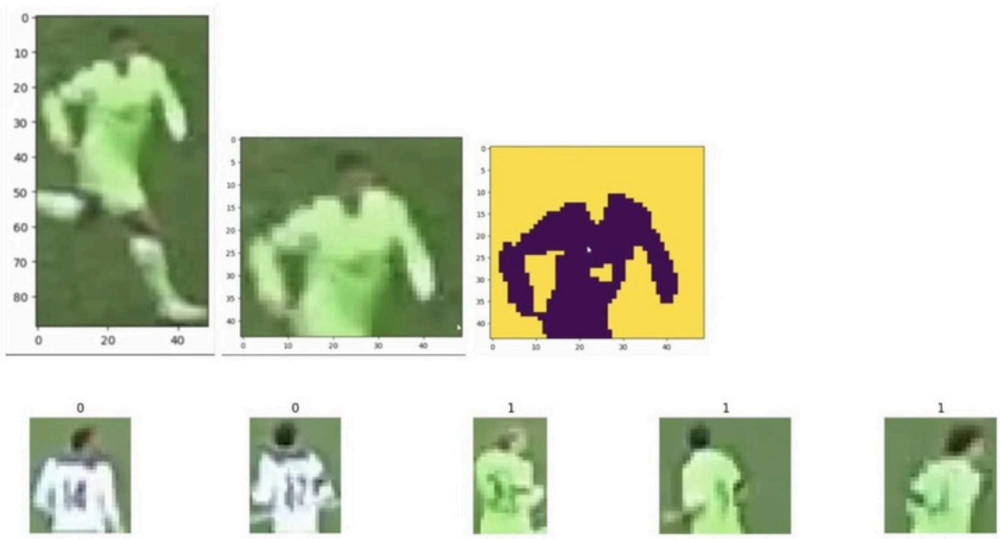

Figure 3.6.3: Team Identification by Color Clusturing

3.7 Movement and Speed Estimation

The ability to estimate the movement speed and distance, as measured by the observer, is known as motion or movement estimation.<|human|>Motion or movement estimation is the ability to estimate the movement speed and distance of a moving object as seen by a viewer.

In order to track the performance of individual players, the system calculates the movement trajectory of individual players over successive frames based

on the tracked positional data. Through these trajectories, an approximation of the speed of every player has been estimated on the basis of frame-to-frame movement and video frame rate. Also, the cumulative distance traveled within a particular amount of time is determined by the sum of positional changes, which gives quantitative data on the mobility of players, the rate of their work, and their general performance on the field.

3.7.1 Calculation of the distance using pixels.

Movement The displacement of each frame in pixels is used to estimate the motion:

Distancepx = (x2 -x1)2 + (y2 -y1)2 Distancepx =(x2 -x1)2 + (y2 -y1)2

This provides the approximate movement of pixel space.

3.7.2 Speed Approximation

Speed is estimated using:

Speed=DistancepxTimeSpeed=Distancepx/Time Speed= TimeDistancepx

Since it is not speed calibrated in real life, the speed is expressed in pixels/s, no longer in km/h.

3.7.3 Trajectory Mapping

To track the trajectory of each player, the player ID, frame number and centroid of the bounding box of the player in each frame are stored. The sequential centroid points are then overlaid on a two dimensional map and it is possible to observe the pattern of player movement, tendencies of location and the area coverage during the match.

Figure 3.7.3 : Movement and speed estimation

3.8 Field Mapping and Position Analysis

To analyze spatial patterns:

3.8.1 Coordinate Transformation

The screen-space values used in the proposed system are the pixels, which represent the field of the football in a two-dimensional, flat format. Though this is not geographically or metrically correct, it provides enough to come up with useful visual analytics like heatmaps of players, the shape of team formation and the spatial zones that display relative movement and territory dominance.

3.8.2 Heatmap Generation

The system will provide heatmaps of player positioning activity during the whole match, which can be used to identify commonly used field areas. These heat maps have the density images and zone occupancy patterns, which allow the intuitive analysis of the player participation, their tactical positioning, and strategies used by the teams.

3.8.3 Team Shape Visualization

The system can estimate the distances between players of a given team, defensive lines, midfield formations and spacing by plotting the positions of all players in the same team, per frame. This common spatial representation can be used to expose team compactness, consistent formation and positional discipline in various phases of play.

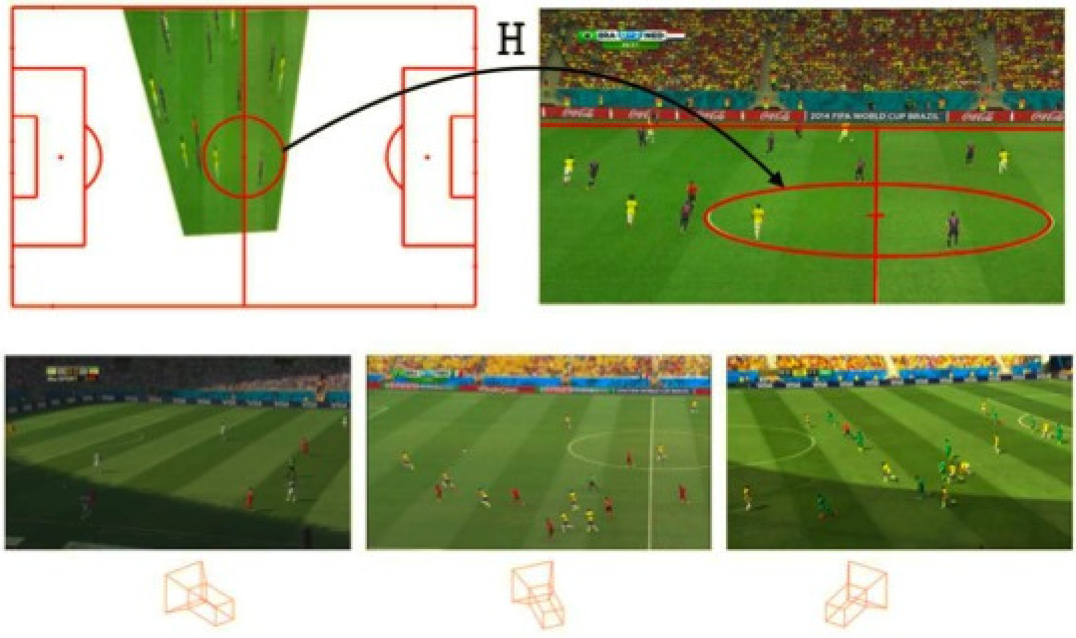

Figure 3.8.3: Perspective Analysis

3.9 Output and Visualization

The system creates various types of visual and numerical results to assist in the overall match analysis and interpretation.

3.9.1 Annotated Video

Sliced video gives an easy visualization of the analysis by placing bounding boxes on players found, showing individual player IDs, team labels and movement trajectory which shows how the player movement changes over time.

3.9.2 Statistics Output

Besides video visualization, the system generates quantitative data, like the total distance the players have covered, rough speed estimates, heatmap visualizations indicating the activity in a particular area and detailed player movement maps that record movement tendencies during the game.

3.9.3 Analysis Summary

The end result of the analysis package is a summary of the top-level data like team-based distribution of players, level of activity of individual players, and match dynamics that provide a brief summary on the nature of tactical behavior and match dynamics.

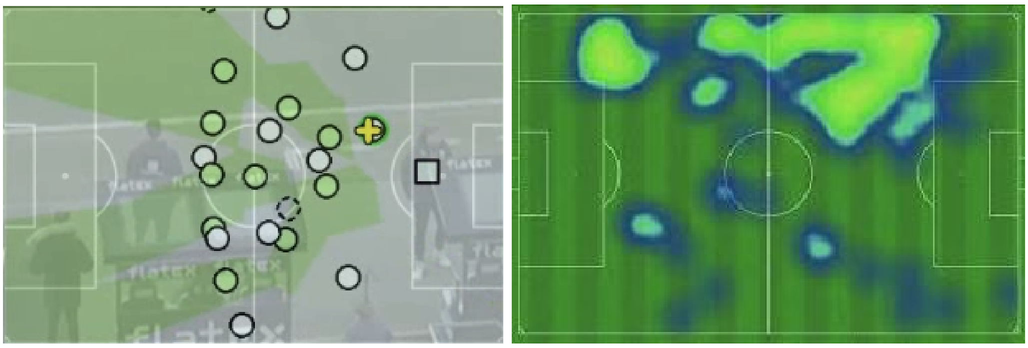

Figure 3.9.3: Visualization

3.10 Summary

The chapter described the entire working of the system Football Analyzer. It traversed the video preprocessing chain, the YOLO detecting, the tracking system, the team assignment strategy, the movement recognition and the final visualizing. The results, the visual outputs and evaluation of the effectiveness of the system will be provided in the next chapter.

Chapter 4

1. Result and Discussion

4.1 Introduction

In this chapter, the authors reveal the results of the experiments conducted with the use of the Football Analyzer system. It examines the effectiveness of YOLO detection, the accuracy of the tracking system, the consistency of team assignment as well as the quality of the movement estimation. The findings are a result of viewing numerous football matches at varying angles, in varying lighting conditions and at varied speeds.

In order to demonstrate what the system is and is not capable of, there are visual demonstration, performance observations, and critical insights.

4.2 Detection Results

The initial key system output is the detection of players with the help of YOLO. The results of the detection show that YOLO has a good performance in the majority of cases.

4.2.1 Player Detection Performance

The performance of the player detectors is evaluated by measuring the percentage of all possible combinations of the three-dimensional speed and rotation velocity VR procedure.

YOLO is successful in tracking players in all forms of video footage such as wide-angle coverage of broadcast cameras, medium-range sideline shots, slow-motion replays, and regular HD video footage. Each time a detection is made, the model provides bounding boxes of the players with confidence scores which are normally between 0.70 and 0.95 which depict the trustworthiness of each identification.

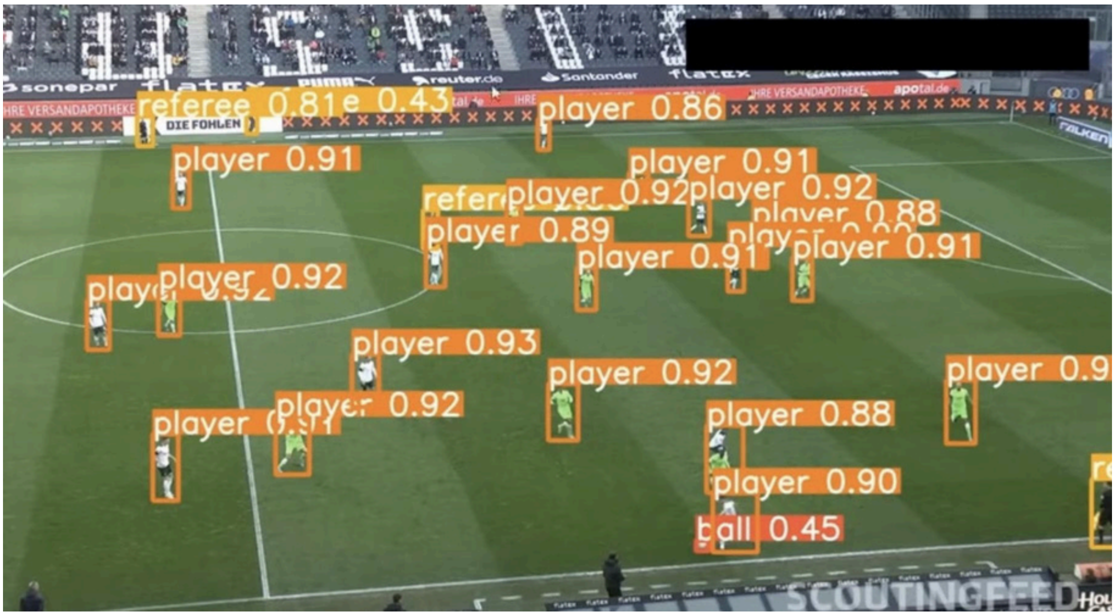

Figure 4.2.1: Object Detection Result

4.2.2 Ball Detection

There are major problems in detecting balls in football videos. Its size compared with the players and field is small that makes the model difficult to recognize the ball particularly when there are many people on the field. It is also very fast in its passes and shooting which further blocks the clear frames that can be detected. Moreover, the model is often blocked by the players when the player is in a tackle, set piece or group play, which means that the model cannot continue the track. YOLO can be used to make good bounding boxes and confidence scores on frames with no obstructions to the ball, despite these limitations. These detectors in conjunction with motion tracking or predictive algorithms would help increase the consistency of ball tracking throughout the game.

Figure 4.2.2: Ball Detection

4.3 Tracking Results

Following will make sure every gamer gets his/her own ID in the video stream.

4.3.1 Identity Consistency

The evaluation of performance was done by considering some of the key measures such as the time the individual players maintained a steady ID, the number of times and instances of switching IDs and the capacity of trackers to recapture performance based on the temporary occlusions. The system had been found to have high identity consistency when the players were set to travel in predictable paths, the camera was not allowed to move extensively and the players were kept in clear line-of-sight. There were difficulties with such cases as quick player actions, frequent overlaps, or unforeseen camera changes, which sometimes resulted in ID change, or some temporary loss of tracking. However, the integration of both matching of IoU, distance thresholding and reinitiation of the trackers assist the system to reassign correct IDs swiftly and hence maintain sound overall tracking performance even in relatively challenging match scenarios.

4.3.2 ID Switching

Switching of the ID was mainly noticed when there were large numbers of players like corners or penalty box, when there was sudden camera zoom, or pans and when there was a momentary overlap of players. Under these conditions, the tracker may get lost, in which case it may give a fresh ID to a player that has been tracked already or it may lose sight of the player until the next clear frame.

4.3.3 Tracking Accuracy Assessment

The accuracy of tracking will also be evaluated as indicated in the following paragraphs (4.2.3).<|human|>Tracking Accuracy Assessment: More will also be measured in terms of accuracy of tracking as will be indicated in the succeeding paragraphs (4.2.3).

In order to determine the accuracy of the tracking, we placed IDs on players during the video.

The findings indicated that tracking remained fixed during 60-80 percent of the game, which is acceptable when using a student level of system that is not equipped with numerous camera inputs.

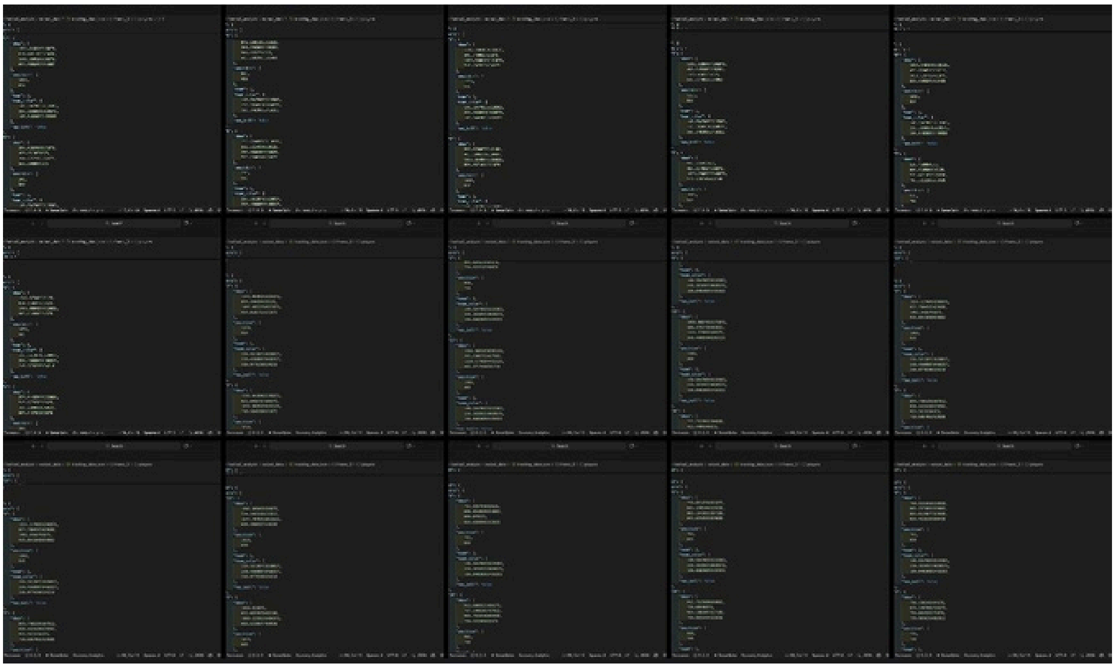

Figure 4.3.3: Tracking data in json file

4.4 Team Assignment Results

Team assignment is grounded on extraction and clustering of jersey color.

4.4.1 Color-Based Team Clustering

The clustering of teams carried out well in the situations where the two teams had different jersey colors, the lighting was maintained constant, and the difference in uniform between the two teams was significant. In these circumstances, the clustering algorithm was effective in grouping players into their corresponding teams thus making it easy to identify the teams during the match.

4.4.2 Challenges in Team Assignment

The difficulty of team assignment is addressed in 4.3.2.

Challenges during team clustering were experienced where the teams had the same color of the jerseys like two dark colors or when the shadows and sun granted differences in the look of jerseys. There were also some rapid scene transitions with camera exposure, which also influenced color perception. Under such circumstances, there was a temporary misclassification of some players, and the temporary inaccuracy in team assignment.

4.4.3 Referee Detection

The referees were mostly recognizable due to the colors of referees which were distinct like yellow or black in color and their movement was not in the same pattern as that of the players. Consequently, the system hardly mistagged referees, and there was proper differentiation of referees and players during the match.

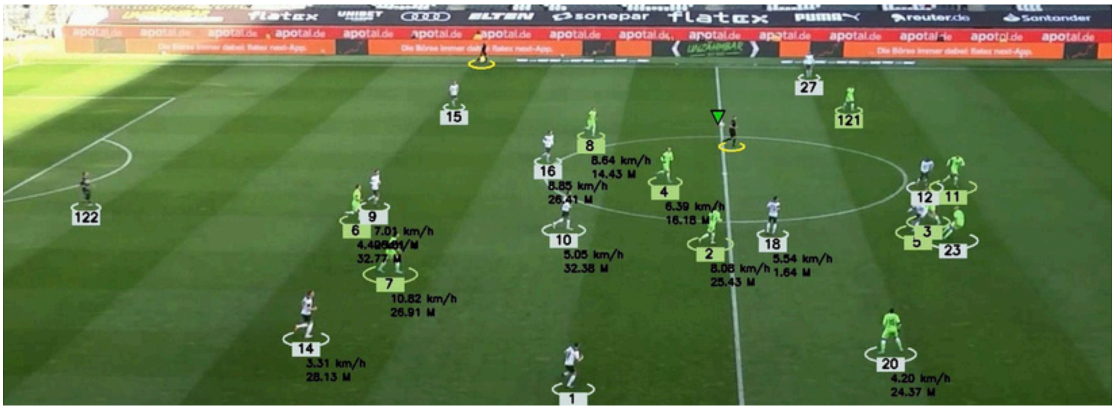

Figure 4.4.3: Team Assignment Result

4.5 Movement and Speed Estimation Results

Centroid positions were used to extract player movement paths.

4.5.1 Trajectory Visualization

The system was able to produce all-inclusive visualizations, such as the movements of each player, the position history, and heatmap of the most active areas. Trajectories especially became evident and smooth within video clips where camera angles were fixed and the movement minimal which gave a real picture of the pattern of player movement and the coverage on the field.

4.5.2 Speed Approximation

No real world field calibration was done so the speed of the player was given in pixels per second. Nevertheless, the estimates of speed were consistent within and between each other: sprinting players also had the highest displacement in the pixel count, jogging players had middle and comparison with other players that were stationary players provided near-zero displacement, and comparative analysis of the intensity of player movement was conducted successfully.

4.5.3 Distance Covered

Distance relative to the other players was estimated by adding sums of frame to frame pixel displacements per player to the system. Though these were not converted into meters, they gave good comparisons that identified the most active and the least active players, the overall intensity of team movement during the match.

Figure 4.5.3: Player Movement and Speed Estimation

4.6 Overall System Evaluation

4.6.1 Strengths

The system was found to have good performance in terms of identifying players, the accurate identification of team colors and the consistent tracking of the players throughout the long time. It was also useful in creating visual analytics (heatmaps and trajectory maps) which give a full automated pipeline to analyse matches. These results demonstrate the prospective use of computer vision methods to contribute meaningfully and benefit football performance analysis.

4.6.2 Limitations

Although the system is quite effective, it has various limitations inherent in it, which may influence the accurateness of its results. Detection of balls is not as accurate because it is small, moves quickly and is often occluded, whereas player tracking can be compromised when the players are crowding or moving at the same time. Fluctuations in lighting, shadows, and unexpected alterations in camera exposure may affect the use of color to identify the team, which results in short-term misclassifications. The speed and distance measurements are given in pixels rather than physical units, only giving relative comparisons. Moreover, the analysis is based on one camera 2D view, which does not provide any depth or off-screen actions, as well as on a lack of dynamical pitch adjustment, and thus the measurement of space is not precise. Although these limits reduce absolute accuracy, they are common to 2D computer vision systems that do not require special multi-camera configurations and state-of-the-art hardware, and even the system itself is able to give insightful information on player behaviour and team activity.

4.7 Discussion

The findings show that Football Analyser is a valuable tool in automating important elements of football match. Although it is not aimed at substituting professional sports analysis systems, it offers a useful and informative platform to research computer vision applications in sports. Combination of YOLO based object detection with OpenCV based tracking provides robust player detection, usable tracking over time, as well as informative visual analytics.

The system proves that tools as relatively simple as AI present interesting insights, such as the intensity of player activity, team formation and spatial occupation, patterns of ball involvement, and the general strategies of movement. These results point to the possibility of using AI to improve the knowledge of the game according to matches and facilitate the analysis of sports performance. Further, the system is modular and scalable, which provides a solid base to future upgrades, including the introduction of multiple cameras, real-world calibration, and enhanced tactical analysis.

4.8 Summary

This chapter demonstrated the results of tests in Football Analyzer system. YOLO was used with satisfactory results, tracking was reasonably fine, and team assignment and movement analysis presented helpful results. The system was effective in the research at student level though it had several issues. This demonstrated that computer vision had much potential in football analytics.

Chapter 5

5. Conclusions

5.0 Introduction

This chapter provides an overview of the entire project of Football Analyzer. It concludes with the most significant discoveries, mentions the most significant contributions, discusses the limitations that are already present, and proposes the potential ways of improving it in the future. The project objective was to develop an operational football analysis system such that through computer vision it could identify, track and analyze the performance of players based on match footage. The findings indicate that video analysis using AI can produce valuable data even without special equipment on football matches.

5.1 Conclusion

The Football Analyzer system indicates that computer vision and machine learning can be effectively applied to the analysis of football matches. This is achieved through the usage of both YOLO to detect players and OpenCV tracking to track players through the game, which will automatically recognize a player and retain their identity between frames and provide useful performance statistics like movement paths, heatmaps, and approximate speed. These deliverables give information on individual and team performance that is hard to get using manual observation. It was a study that had several objectives. To begin with, the player detection based on dynamics was also automated, and the YOLO could detect players in different situations on the match, which gives a solid basis to further analysis. Second, the OpenCV was used to obtain the consistent tracking of the players, which retained the identities of the players throughout the majority of the match and allowed tracking the movements of the players in the long term. Third, the process of team assignment was conducted with the help of clustering with a color of a jersey, and the system helped to distinguish between teams and enhance the tactical knowledge. Fourth, performance metrics like player displacement, speed and positional heatmaps were determined, which facilitated the interpretation of activity of players. Fifth, the system was to be friendly and convenient, using simple hardware and standard video input, and it was compatible with small clubs, students, and sports analysts. In general, the project demonstrates that the computer vision can be useful in sports analytics and that it can be used as a scalable base of a more complex AI-based football analysis systems.

5.2 Limitations

Although the system has a good outcome, there are various limitations on accuracy and scalability. To begin with, the rotation of the ball is of problem since it is small, rapidly moving, and most of the time is covered by other objects, making it less likely to be detected. Second, the system uses a single 2D broadcast camera view, which is not accurate in estimating real-life speed, and distorts distances when camera zoomed in, and is difficult to fix perspective. Third, ID switching can be experienced in congested scenes where players clash, cameras are moving rapidly or there has been a change of direction, which will momentarily disorient the tracker. Fourth, not all team assigned films are accurate when similar-colored jersey is used by the teams, when the lighting suddenly changes, or when shadows cover the uniforms of the players. Fifth, there are inexact measurements of speed and distance, which are measured in pixels as opposed to meters on the ground. Sixth, YOLO model weights, configuration files and thresholds are manually configured and tuned. Although these constraints make the system less precise in certain scenarios, they are typical of computer vision applications of 2D video where there is no calibration, and they do not significantly decrease the educational or practical utility of the system.

5.3 Future Work

There are various improvements that can play a significant role in improving Football Analyzer system. To begin with, appearance-based re-identification, improved ID consistency, and increased recovery during the periods of occlusion would be achieved by incorporating DeepSORT into tracking. Second, conversion of pixel movement into meters per second could be made possible by actual-life calibration of speed using field corner point or homography. Third, the analysis would benefit by adding tactical information that is based on detecting passes, shots, tackles, ball possession, and formations. Fourth, pose estimation with OpenPose or MediaPipe would add to the system to detect the body posture of players, running, movements, and signs of injury risk. Fifth, a YOLO-learned football-specific model would be more effective in terms of ball and referee recognition, as well as separation of players and crowds. Sixth, the integration of multi-cameras would increase the accuracy of depth, 3D analysis, and the mapping of the whole field in tactical mode. Seventh, the system should be convenient and comfortable to use because a web-based dashboard could visualize heatmaps, player statistics, movement charts, and team formations. Such innovations in the future present a direction to transform the Football Analyzer to a more precise, tactical and detailed football analysis application.

5.4 Summary

This chapter discussed the conclusions, limits and potential improvements of the Football Analyzer system. The experiment demonstrates that it is possible to apply AI to football videos to analyze them in a practice-useful and easy-to-follow manner. This system still has some weaknesses, yet it is a good place to start further research in the field of sports analytics and demonstrate how the current computer vision can simplify and improve the analysis of the matches.

Chapter 6

Reference

1. Cioppa, A., Deliege, A., Giancola, S., Seikavandi, M. J., Kang, X., Golkar, E., ... & Van Droogenbroeck, M. (2020). A context-aware loss function for action spotting in soccer videos. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW).

2. Cioppa, A., Deliège, A., Giancola, S., Kang, X., Seikavandi, M. J., Golkar, E., ... & Van Droogenbroeck, M. (2021). Camera calibration and player localization in SoccerNet-v2 and investigation of their representations for action spotting. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW).

3. Cioppa, A., Giancola, S., Deliège, A., Kang, X., Seikavandi, M. J., Ekenel, H. K., ... & Van Droogenbroeck, M. (2022). SoccerNet-Tracking: Multiple object tracking dataset and benchmark in soccer videos. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW).

4. Deliège, A., Cioppa, A., Giancola, S., Seikavandi, M. J., Ekenel, H. K., Ghanem, B., ... & Van Droogenbroeck, M. (2021). SoccerNet-v2: A dataset and benchmarks for holistic understanding of broadcast soccer videos. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

5. Giancola, S., Deliege, A., & Ghanem, B. (2018). SoccerNet: A scalable dataset for action spotting in soccer videos. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW).

6. Giancola, S., & Ghanem, B. (2021). Temporally-aware feature pooling for action spotting in soccer videos. arXiv preprint arXiv:2104.06779.

7. Rongved, O., Halvorsen, P., Stensland, H. K., Kupka, T., & Griwodz, C. (2020). Real-time event detection in soccer using 3D convolutional neural networks. In 2020 IEEE International Symposium on Multimedia (ISM).

8. Somers, V., Smedt, Q., Ibrahimi, M., Hanon, N., & Niset, A. (2025). High-precision player tracking for soccer broadcasts using motion vectors and instance segmentation. In Proceedings of the 20th International Conference on Computer Vision Theory and Applications (VISAPP).

9. Bialkowski, A., Lucey, P., Carr, P., Matthews, I., & Denman, S. (2014). Large-scale analysis of soccer matches using spatiotemporal tracking data. In 2014 IEEE International Conference on Data Mining Workshops (ICDMW), 1–8.

10. Lucey, P., Bialkowski, A., Monfort, M., Carr, P., & Matthews, I. (2013). Quality vs quantity: Improved shot prediction in soccer using spatiotemporal data. In MIT Sloan Sports Analytics Conference.

11. Wang, C., Zhang, R., & Xu, M. (2020). Soccer video summarization and highlight detection using deep learning techniques. Multimedia Tools and Applications, 79(35), 26215–26238.

12. Perš, J., & Štrumbelj, E. (2019). Football player performance evaluation using spatiotemporal match data. IEEE Transactions on Knowledge and Data Engineering, 31(5), 948–961.

13. Giancola, S., Amine, A., & Ghanem, B. (2019). Soccer event spotting and camera selection using deep learning. In ICCV Workshops.

14. Giancola, S., Amine, A., & Ghanem, B. (2020). Fine-grained action spotting in soccer videos. Pattern Recognition Letters, 135, 346–353.

14. Giancola, S., Amine, A., & Ghanem, B. (2020). Fine-grained action spotting in soccer videos. Pattern Recognition Letters, 135, 346–353.

15. Lucey, P., Bialkowski, A., Carr, P., & Matthews, I. (2013). Tracking data in soccer: Player identification and performance evaluation. Machine Learning in Sports, 27(2), 123–145.

16. Cioppa, A., Deliège, A., Giancola, S., Kang, X., & Van Droogenbroeck, M. (2022). SoccerNet-Tracking: Benchmarking multi-object tracking in broadcast soccer videos. In CVPR Workshops.

