SPATIO-TEMPORAL ANALYSIS OF VOLLEYBALL DATA FOR COMPLETE TRACKING

By

Shakhansho Sabzaliev

In partial fulfillment of the requirement for the degree

Bachelor of Science in Computer Science

Department of Computer Science, School of Arts & Sciences

University of Central Asia, Kyrgyz Republic.

(2023)

DECLARATION BY AUTHOR

I certify that this work has not been accepted in substance for any degree and is not concurrently being submitted for any degree other than that of Bachelor of Science in Computer Science being studied at the Department of Computer Science, School of Arts & Sciences, University of Central Asia, Kyrgyz Republic. I also declare that this work is the result of my own findings and investigations except where otherwise identified by references and that I have not plagiarized another's work.

abzaliev

==

DECLARATION BY SUPERVISOR

We, the undersigned hereby certify that We have read this project report and finally approve it with a recommendation that this report may be submitted by the author above to the final year project evaluation committee for final evaluation and presentation, in partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science at the Department of Computer Science, School of Arts & Sciences, University of Central Asia, Kyrgyz Republic.

Artyom Kulakov

ABSTRACT

With the increasing application of artificial intelligence in various fields, computer vision techniques have significantly impacted the realm of sports. For example, the Hawkeye system in tennis has changed how people watch and play the game. In this research, we want to use the latest computer tools to improve volleyball. I am going to create a very first annotated volleyball videos and share them with other people who want to learn. In addition, I will also create a website where users can upload their own volleyball videos and choose what parts they want to see, like the ball or the players. Using Keras and PyTorch software will help me to make this happen. My hope is that this project will help volleyball become popular among open-source developers and computer vision practitioners, or at the very least, the dataset I created will be of use to fellow developers.

Keywords: volleyball detection, volleyball tracking, court detection, court tracking, player detection, player tracking, game statistics, smart referee.

TABLE OF CONTENTS

LISTS

List of Abbreviations

CV – computer vision

ML – machine learning

DL – deep learning

YOLO - You Only Look Once

RPN - Region Proposal Network

mAP - Mean Average Precision

FPS - Frames Per Second

ReLU - Rectified Linear Unit

IoU - Intersection over Union

API - Application Programming Interface

GUI - Graphical User Interface

URL - Uniform Resource Locator

JSON - JavaScript Object Notation

GPU – Graphical Processing Unit

SOT – Single Object Tracking

MOT – Multiple Object Tracking

SORT – Simple Online and Real-time Tracking

RPN - Region Proposal Network

DaSIAMRPN - Distractor-aware Siamese Region Proposal Network

BoT-SORT - Bottom-up Single Object Tracking

mIoU - Mean Intersection over Union

List of Figures

List of Tables

STATEMENT OF AUTHORSHIP

I, Shakhansho Sabzaliev, hereby declare that the thesis/project submitted, entitled "Spatio-Temporal Analysis of Volleyball Data for Complete Tracking," is my own original work, conducted and written in accordance with the research ethics guidelines of University of Central Asia. I confirm that all ideas, data, and findings presented in this report have been derived solely from the sources listed in the references section, and that no other sources have been used. I further certify that any assistance received during the research and writing process has been appropriately acknowledged in the report.

I understand that any violation of the university's research ethics guidelines, including plagiarism or academic dishonesty, may result in disciplinary action as determined by the university.

Shakhansho Sabzaliev

1000001715

May 1st, 2023

Chapter 1: Introduction

In this report, I present my project called '*VolleyVision*' a comprehensive computer vision-based system designed to analyze and track volleyball matches in real-time. The logo for VolleyVision is displayed in Figure 1, which represents the fusion of computer vision (eye) and volleyball.

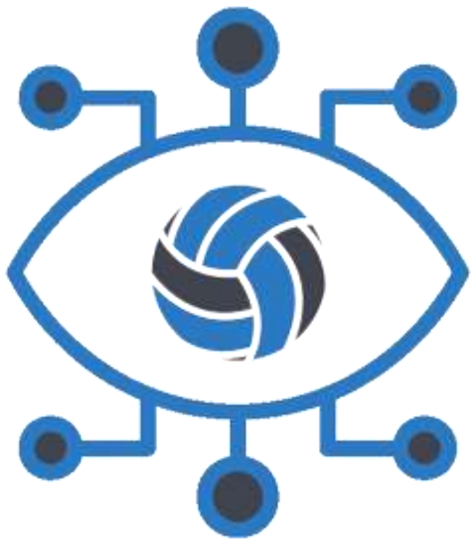

Figure 1 VolleyVision Logo

1.1. Background and Motivation

The field of artificial intelligence (AI) has experienced exponential growth in recent years, with machine learning (ML) and computer vision (CV) becoming increasingly relevant across various industries (Bishop, 2021)[14]. Sports analytics is one such domain that has greatly benefitted from advancements in AI technologies, leading to improved decision-making, performance analysis, and game strategy (Lucey et al., 2018)[13].

In sports, CV techniques have been employed for object tracking, player identification, action recognition, and smart refereeing systems, among other applications (Cioppa et al., 2020)[15]. Tennis' Hawkeye technology, which has changed the game by offering precise and real-time analysis of ball trajectory and court position (Liu et al., 2013)[16], is one noteworthy instance.

Last but not least, it is important to note the advancements made in sports like basketball and soccer, which resulted in a definite improvement in both player performance and fan satisfaction. (Zhang et al., 2020)[18].

In the case of volleyball, things are slightly different. There is a huge gap in the application of such technologies, particularly among the open-source community to be precise. For example, on GitHub if you search for “tennis tracking” and “volleyball tracking” you will see the following differences in results: 12 repositories for volleyball and 292 for tennis. While some professional volleyball leagues have adopted AI technologies for tracking and analytics, these solutions are often proprietary and inaccessible to the general public (Zhang et al., 2020) [18]. Additionally, the limited availability of open-source volleyball datasets and tracking algorithms creates a barrier for researchers and developers looking to explore AI applications in the sport.

Volleyball presents unique challenges for CV applications, such as the size of the ball, fast movements, and frequent occlusions by players (Soleimanitaleb & Keyvanrad, 2022) [19]. Overcoming these challenges requires the development of specialized algorithms and models tailored to the sport's specific requirements. Furthermore, the lack of publicly available annotated datasets for volleyball hinders the progress of AI research in the sport, as it limits the ability to train and evaluate new models (Bialkowski et al., 2016)[13].

The primary motivation behind this project is to address these challenges by contributing to the open-source community through the development and deployment of state-of-the-art CV algorithms tailored to volleyball. I hope by creating an annotated volleyball dataset and a web application which will accept videos from users and detect and track the ball, the players, and court, I hope to advance the development of volleyball and smart refereeing solutions and also promoting the popularization of volleyball in the field of machine learning and computer vision. (Bialkowski et al., 2016) [13].

Moreover, with my project, I aim to bridge the gap between ML & deep learning (DL) research and practical applications in volleyball. By making the code and datasets open-source, I hope it will encourage further development and research from fellow researchers and developers (Szelag et al., 2019)[8]. Ultimately, this could lead to the birth of new tools and applications that could enhance volleyball, making it more engaging and accessible to a wider audience (Citraro et al., 2020)[9].

All in all, my project aims to address the existing gap in open-source AI applications for volleyball by implementing state-of-the-art algorithms and creating volleyball-specific annotated datasets. The project's contributions to the open-source community are expected to foster further research and development in the sport, ultimately leading to improved performance analysis, game strategy, and spectator experience.

1.2. Project Aim and Objectives

The primary aim of my project is to apply popular CV tasks and techniques, such as object detection, tracking, action recognition, and classification, to volleyball data in order to popularize volleyball in the field of artificial intelligence. With my project I seek to address the gap in the existing literature and applications related to volleyball tracking and analysis. By pursuing this project, I hope that I will make a significant or at least some contribution whatsoever to the open-source community by developing new resources and tools specifically designed for volleyball.

In addition to addressing this gap, my project serves as an opportunity for me to quickly develop a set of skills in the fields of CV, ML, data collection, data annotation, and deployment. As a senior student who is close to graduation, working on this project will help me gain valuable experience and knowledge that can be advantageous for future academic and professional careers I pursue.

The first objective of my project is to learn about different state-of-the-art DL models used for detecting and tracking objects. By doing so, I will have a good understanding for the underlying principles and techniques in CV and ML that will later help me in my work.

My second objective is to extract all volleyball-related data from volleyball videos. This includes information about the ball, players, and the court. This is the key to building a comprehensive dataset that can be used to train and evaluate the ML & DL models that will be used in this project.

Next, I will create a custom dataset for volleyball, including data collection and its annotation using tools provided by RoboFlow. Developing a rich and diverse dataset specifically created for volleyball will enable the project to generate more accurate and reliable results.

The fourth objective is to decide the type and structure of the neural network models to be used in the project. This decision will be informed by the knowledge gained from studying state-of-the-art models and will consider factors such as efficiency, accuracy, and the unique requirements of volleyball data.

Training the models with the custom dataset is my fifth objective. This step is going to involve fine-tuning the chosen neural network models using the collected and annotated volleyball data, ensuring that the models are optimized for the specific task at hand.

As for the sixth objective, it is to develop a web application that will use the trained neural network models, which will allow users to upload volleyball videos and get processed results with detected, tracked, and highlighted ball, players, and the court. This will make the project's outcomes readily accessible and usable by users.

I also want to create a user-friendly UI for the web application. This will ensure that the application is accessible and easy to navigate to all users, which will as a result encourage more users to engage with the project and benefit from it.

Finally, I sincerely want and will to contribute to the open-source community. I plan to do so by uploading the volleyball dataset created during the project and making the code available for others to use, learn from, and build upon. I hope this will encourage collaboration and innovation, speeding up the development of new applications and resources in the field of volleyball tracking and analysis.

1.3. Scope and Limitations

The scope of my project is to develop a web application that can detect and track the ball, players, and the court in volleyball videos, as well as provide game statistics and spatio-temporal data. My project aims to achieve this through the use of state-of-the-art DL models and a custom volleyball dataset. Also, I am super eager to contribute to the open-source community by sharing the dataset and code with other users, developers, and volleyball enthusiasts.

Despite the objectives I gave myself, there are a few constraints that need to be taken into account. The hardware limits of my concept are one of its main drawbacks. The project is being worked on using a laptop with only 4 GB of GPU RAM. The execution of computationally demanding activities, such as the processing of high-resolution films or the training of extensive neural network models, is constrained by this restriction. Additionally, longer training times and poorer overall performance are caused by the GPU's memory limitations.

The dataset's size is still another restriction. Although I wanted to make a substantial custom volleyball dataset, the final product only had 25,000 volleyball images. Despite the huge number of photos, the laptop's memory constraints prevented full use of the dataset for model training. The benefits of the dataset size could not be fully utilized as a result. This restriction also had an impact on the models employed for volleyball object detection. I had to select smaller object detection models due to hardware and dataset size restrictions, which had an impact on the volleyball detection model's accuracy. Larger and more intricate models frequently perform better, but their needs surpass the resources at hand. As a result, the choice of a smaller model have led to a trade-off between computational feasibility and model accuracy.

The quality of the annotated data in the custom volleyball dataset can have a significant impact on the performance of the trained models. Manual data annotation is time-consuming and can be prone to errors, which have affected the reliability of the resulting models. Moreover, the quality of the input videos used for data extraction also influenced the performance of the models, as low-resolution or noisy videos may lead to less accurate detections and tracking. Lastly, in addition to my own annotated videos I also added to the different other available datasets which had volleyball in them. This led to inconsistency, and both object and quality variations in the dataset which negatively affected the detection model.

Finally, volleyball games involve complex scenes with players overlapping, fast-moving action, and varying lighting conditions. These factors can make object detection and tracking more challenging, and the developed models may struggle to provide accurate results in such cases. Furthermore, the models trained on volleyball may not be directly applicable to other sports or tracking tasks without additional modifications or fine-tuning. Although the techniques and approaches can be adapted to other domains, the specific models and dataset may limit their generalizability.

1.4. Organization of the Report

I organized this report in a way to provide a thorough overview of my project, development, and the results I achieved. The sections are as follows:

- *Introduction*: In this section I cover the background, context, and motivation for the project, as well as a brief overview of the project's aims and my objectives.
- *Literature Survey*: Here I present a review of relevant literature in the fields of CV, DL, and sports analytics, while also discussing the state-of-the-art techniques and approaches that have been applied to object detection and tracking in various sports.
- *Methodology*: You will learn about the methodologies and approaches used in the project in this section, including how the unique volleyball dataset was produced, which DL models were chosen, and how object identification and tracking algorithms were put into practice.
- *Implementation*: This section will cover the actual implementation of the project, including the development of the web application, the training of the DL models, and the integration of the models into the application.
- *Results and Evaluation*: This is probably the most important part the results of the project, including the performance of the DL models and their effectiveness in detecting and tracking.
- *Conclusion*: Lasty, I summarize the main findings of the project and talk about its overall success and limitations. In addition, I also make a few points regarding the project's contribution to the open-source community and the potential impact of the developed models and dataset on the field of volleyball.
- *References*: A must-have section of the report, this lists of all the sources I used in my work.

CHAPTER II: LITERATURE REVIEW

2.1 Object Detection and Tracking in Sports

2.1.1 General overview of object detection and tracking in sports

As CV plays a critical role in understanding and interpreting sports events, the discipline of sports analytics has undergone tremendous technical changes in recent times. For coaches, players and sports analysts, object detection and tracking has become more important as a way to better understand player performance, team dynamics and tactical decision making. With a focus on the value and challenges of tracking in sports applications, this section will provide a comprehensive overview of object identification and tracking in sports.

Locating and recognizing items (such athletes, balls, or other equipment) in video sequences and following their movements through time are both parts of the object identification and tracking process used in sports. Numerous applications, such as player performance analysis, injury prevention, tactical analysis, and fan engagement, depend on this process. For instance, tracking player movements enables coaches to spot team performance strengths and shortcomings, improve their strategic planning, and reach data-driven conclusions. In addition, tracking the ball during games like volleyball can provide useful data regarding its course, velocity, and spin, which can be utilized to evaluate and enhance player abilities and strategies. (Zhang et al., 2020)[18].

2.1.2 Importance and challenges of tracking in sports applications

Nevertheless, there are a number of difficulties with object detection and tracking in sports. The high-speed nature of sports, which frequently causes quick object motion and frequent occlusions, is one of the main issues. Because of this, it is challenging for CV algorithms to keep track of objects consistently over the video sequences (Soleimanitaleb & Keyvanrad, 2022)[19]. The problem of object detection and tracking is further complicated by the varied and dynamic character of sports scenarios, including various lighting conditions, camera angles, and player appearances. Additionally, it is difficult to distinguish and track individual players accurately due to the presence of multiple players with similar physical attributes and dynamic spatial relationships (Morimitsu et al., 2017)[10].

By using cutting-edge object detection and tracking techniques for volleyball match analysis, I want to overcome some of these issues in my study. The project's initial phase is devoted to volleyball detection and tracking, and I've already had success training two models on a unique dataset of 25,000 photos. The project's second and third phases will, respectively, focus on player and court detection and tracking. By putting these cutting-edge CV methodologies into practice, I intend to add to the increasing body of sports analytics research and offer beneficial insights for volleyball coaches, players, analysts and referees, since combining the detections from the ball and the court will allow us create smart referee ball in/out system.

Object identification and tracking in sports are crucial elements of sports analytics because they offer useful data for performance evaluation, injury avoidance, and tactical decision-making. The accuracy and dependability of object detection and tracking in sports applications continue to grow thanks to developments in CV approaches, despite the difficulties posed by high-speed motion, occlusions, and dynamic sporting situations. By incorporating these methodologies into my study, I hope to advance the understanding of player performance and team dynamics in volleyball as well as the development of useful tools and methods for match analysis.

2.2 Deep Learning for Object Detection

2.2.1 Review of popular deep-learning-based object detection models

Deep learning has revolutionized the field of CV, enabling the development of highly accurate and efficient object detection models. These models have been widely used in various applications, including sports analytics, where they have shown great potential in detecting and tracking objects like players and balls. This section will look at the significance of the YOLO (You

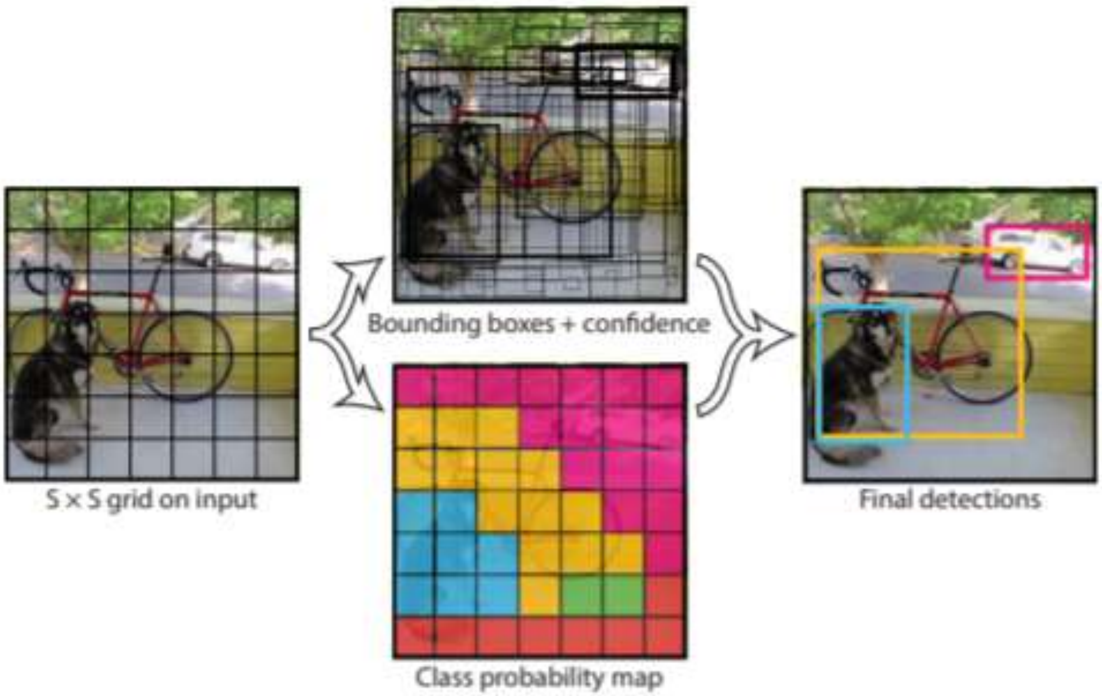

Figure 2 YOLO Architecture

Only Look Once) series for real-time object identification, especially in the context of my project, along with a discussion of some well-known deep-learning-based object detection models.

Convolutional Neural Networks (CNNs) are commonly used in deep-learning-based object identification models to build hierarchical feature representations from huge amounts of labeled data. Two general categories can be used to group these models: single-stage detectors and two-stage detectors. Two-stage detectors, such as R-CNN, Fast R-CNN, and Faster R-CNN, first generate region proposals using a separate network and then classify each proposal into object classes using another network (Bishop, 2021)[14]. Although these models achieve high accuracy, their multi-stage nature makes them computationally expensive and less suitable for real-time applications.

In contrast, one-stage detectors like the YOLO series and SSD (Single Shot MultiBox Detector) aim to achieve a balance between accuracy and speed by performing object detection in a single pass through the network. The YOLO series, in particular, has gained widespread popularity due to its ability to achieve real-time object detection with high accuracy. YOLO divides the input image into a grid and predicts bounding boxes and class probabilities for each grid cell simultaneously (Wang et al., 2022)[5]. This approach eliminates the need for separate region proposal and classification networks, significantly reducing computation time and making YOLO suitable for real-time applications.

2.2.2 YOLO (You Only Look Once) series and its significance in real-time object detection

In the context of my project, the choice of object detection model is critical, as it directly impacts the accuracy and speed of the volleyball, player, and court detection and tracking tasks. The YOLO series is especially relevant for my project, as it offers a suitable trade-off between

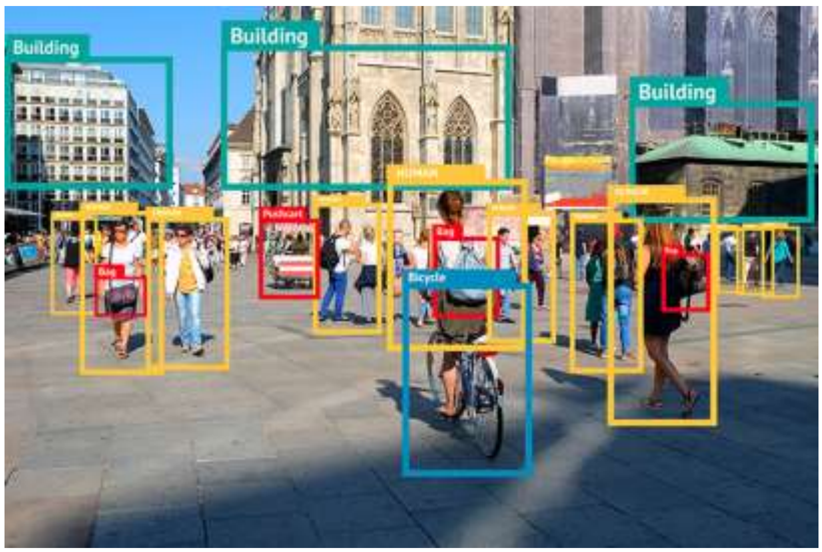

Figure 3 YOLO Detections

accuracy and real-time performance. In the first phase of my study, I trained two models—RoboFlow (AutoML training) and YOLOv7-tiny (local training)—for volleyball recognition and tracking. The YOLOv7-tiny model had real-time inference capabilities, making it a suitable option for applications with limited computational resources and a requirement for real-time performance, even if the RoboFlow model obtained superior accuracy. Even more precise and effective models, such YOLOv7 (Wang et al., 2022)[5], have been produced as a result of the ongoing development of the YOLO series, which has seen advancements in architecture, training methods, and feature extraction. However, I was unable to train the default YOLOv7 model on my dataset due to the limitations of my hardware setup. In spite of this, the YOLOv7-tiny model has demonstrated success in real-time volleyball detection and tracking, albeit with a little less precision than the RoboFlow model.

In conclusion, deep learning has made it possible to create very precise and effective object detection models, which are extensively used in sports analytics, such as volleyball match analysis. Particularly the YOLO series has proven to be significant in real-time object identification, making it a good fit for my project. I hope to solve the problems with identifying and tracking things in sports and offer helpful information to volleyball coaches, players, and analysts by utilizing these cutting-edge object identification models.

2.3. Single and Multiple Object Tracking Algorithms

2.3.1 Overview of single and multiple object tracking methods

In sports applications, where the precise and ongoing monitoring of players or objects is necessary for performance analysis and decision-making, object tracking is a critical component of CV. Single object tracking (SOT) and multiple object tracking (MOT) are the two primary subtypes of object tracking (Soleimanitaleb & Keyvanrad, 2022)[3]. A single target in a video sequence is the focus of single object tracking (SOT). SOT algorithms are excellent for applications that demand real-time speed or limited processing capacity since they are frequently computationally efficient and may be modified to track specific objects. On the other hand, multiple object tracking (MOT) deals with the concurrent tracking of multiple targets within the video sequence. Occlusions, interactions between objects, and changes in the target's appearance are additional difficulties that MOT systems must manage (Wojke, Bewley, & Paulus, 2017)[4]. For SOT and MOT, a variety of strategies have been put forth, ranging from conventional techniques like Kalman filters, particle filters, and optical flow to more recent deep learning-based strategies that have demonstrated appreciable gains in tracking accuracy and robustness (Soleimanitaleb & Keyvanrad, 2022)[3].

2.3.2 Review of Popular Tracking Algorithms: DaSiamRPN, DeepSORT, and others

The Siamese network architecture, region proposal network (RPN), and deep learning are all combined in the single object tracking technique known as Distractor-aware Siamese Region Proposal Network (DaSiamRPN) (Zhu et al., 2018)[7]. DaSiamRPN is renowned for its exceptional accuracy and resilience to a variety of difficulties, including occlusions, scale differences, and motion blur. I used DaSiamRPN in my project to track volleyballs because of its usefulness and efficiency. Deep Simple Online and Realtime Tracking (DeepSORT) is a popular multiple object tracking algorithm that combines deep learning-based appearance descriptors with the Kalman filter and the Hungarian algorithm for data association (Wojke, Bewley, & Paulus, 2017)[4]. DeepSORT has been widely used in various applications, such as sports, surveillance, and autonomous driving, due to its robustness to occlusions, interactions among objects, and changes in target appearance.

Other tracking algorithms worth mentioning include the Simple Online and Real-time Tracking (SORT) algorithm, which is an efficient and straightforward method for tracking multiple objects based on the Kalman filter and the Hungarian algorithm (Bewley et al., 2016)[2]. Moreover, the TrackNet algorithm is a deep learning-based approach explicitly designed for tracking high-speed and tiny objects in sports applications, such as volleyball and table tennis (Huang et al., 2019)[1].

In summary, numerous single and multiple object tracking algorithms have been developed over the years, each with its strengths and weaknesses. The choice of the tracking algorithm depends on the specific application requirements, such as tracking accuracy, computational efficiency, and robustness to various challenges. For my project, DaSiamRPN was chosen for volleyball tracking due to its performance and real-time capabilities.

2.4 Volleyball-Specific Computer Vision Applications

2.4.1 Existing Research and Applications of Computer Vision Techniques in Volleyball

Volleyball is one of many sports that use computer vision techniques and this area of research is expanding. Various topics related to volleyball have been the subject of several studies (Szelag, Kwolek, & Pospieszny, 2019; Szelag et al., 2019; Kong et al., 2020), including player tracking, court detection, and ball tracking. These studies aim to develop efficient and accurate methods to extract valuable information from volleyball matches for performance analysis and decision-making.

For instance, Szelag et al. (2019)[8] proposed a method for player tracking in indoor volleyball using a single camera. Their method uses a combination of background subtraction, morphological operations, and blob analysis to detect and track players in real-time. In another study, Kong et al. (2020)[7] developed an online multiple-athletes tracking method that relies on pose-based long-term temporal dependencies. This method is designed to handle occlusions and interactions among players effectively.

Additionally, research on computer vision applications specifically for volleyball has focused on court detection and tracking. Citraro et al. (2020)[9] offered a real-time camera pose estimate method for sports fields, including volleyball courts, whereas Szelag et al. (2019)[8] presented a real-time camera pose estimation method based on the volleyball court perspective.

2.4.2 Importance of Volleyball Detection and Tracking in Match Analysis

The volleyball game dynamics, individual performance, and team strategies may all be learned from volleyball detection and tracking, which are essential components of match analysis in volleyball. Coaches and analysts can better comprehend the trajectory, speed, and positioning of the volleyball during a game with the use of accurate detection and tracking (Szelag, Kwolek, & Pospieszny, 2019)[8].

Additionally, keeping track of the volleyball's movement can help you spot crucial plays like successful attacks, blocks, and serves. Coaches can utilize this data to assess the performance of individual players and then use that knowledge to decide on player replacements, tactical changes, and training schedules (Cioppa, Gómez, & Gonçalves, 2020)[15]. It may also be simpler to generate game data, such as the number of successful assaults, blocks, or innings, as well as improved team performance, if volleyballs are automatically recognized and monitored. According to Byalkowski et al. (2016)[13], these statistics may be applied to different teams and players, and their manifestations and weaknesses can be observed on both the offensive and defensive sides of the field.

The conclusion of this paper offers thorough details on the game's rules, player and team strategies, volleyball detection and tracking, and conference-wide analysis. The improvement of match analysis as well as the growth of volleyball as a sport can both be significantly aided by the development of precise and effective computer vision algorithms for volleyball identification and tracking.

2.5 Player Detection and Tracking in Sports

2.5.1 Techniques and Challenges in Player Detection and Tracking

There has been a lot of research done on player detection and tracking in sports, and several strategies have been put forth to help with the difficulties involved. Background subtraction, optical flow, and blob analysis are common techniques for player recognition, whereas more recent strategies depend on deep learning-based object detection models like YOLO, Faster R-CNN, and SSD (Cioppa et al., 2020; D'elia et al., 2019).

Several difficulties persist despite the advances in player detection and tracking. Accurate player detection and tracking in crowded, dynamic settings with frequent player interactions and occlusions is one of the biggest hurdles (D'elia et al., 2019)[21]. Managing changes in player appearance caused by various lighting situations, camera angles, and player uniforms is another difficulty. The effectiveness of identification and tracking algorithms can also be impacted by differences in player size and shape (Xu et al., 2013)[16].

The use of multiple cameras to provide different viewpoints and reduce occlusions, incorporating player appearance information into tracking algorithms, and utilizing deep learning techniques for improved detection and tracking performance are just a few of the methods that researchers have suggested to address these issues (Cioppa et al., 2020; D'elia et al., 2019; Xu et al., 2013).

2.5.2 Importance of Player Tracking for Sports Analytics

Player tracking is crucial for sports analytics since it offers insights into team dynamics, individual player performance, and tactical plans. In order to evaluate player performance, identify strengths and shortcomings, and create customized training programs, accurate player monitoring can make it easier to analyze player motions, physical load, and playing patterns (Cioppa et al., 2020)[15].

Player tracking can also be used to examine player positioning, formation, and interactions during a game to evaluate team dynamics and strategies (Gudmundsson & Horton, 2017)[23]. This information can be used by coaches and analysts to create game plans, alter tactic, and enhance team performance.

Sports injury management and prevention can benefit from player tracking. Sports scientists and medical professionals can identify potential risk factors for injuries and execute suitable interventions to lower injury rates and promote player recovery by keeping an eye on player movements, physical load, and fatigue levels (Gabbett, 2016)[22].

In conclusion, player recognition and tracking are essential components of sports analytics because they offer important information about the performance of individual players, the dynamics of teams, and tactical approaches. The creation of precise and effective player detection and tracking methods can greatly improve sports analytics and advance sports science and performance analysis as a whole.

2.6 Court Detection and Tracking in Sports

2.6.1 Techniques for Detecting and Tracking Sports Courts, with a Focus on Volleyball

In volleyball, in particular, precise court detection and tracking are crucial components of sports video analysis. By offering a constant reference frame for the entire video sequence, court detection can aid in player tracking, ball detection, and match analysis. Traditional computer vision algorithms and deep learning-based strategies have both been presented for the detection and tracking of sports courts (Szelag et al., 2019; Citraro et al., 2020). The court lines and limits are often identified using geometric and color cues in traditional court detection methods. For instance, it is possible to identify court lines and determine the position and orientation of the court using Hough-transform-based line identification, template matching, and color segmentation (Szelag et al., 2019)[8]. However, these methods might have trouble when there are obstructions, different lighting conditions, or camera motion.

Deep learning-based approaches have also been applied to court detection, with convolutional neural networks (CNNs) being the most common choice. CNNs can be trained to learn robust features from large datasets, allowing them to handle variations in court appearance, lighting conditions, and camera viewpoints. For example, Citraro et al. (2020) proposed a real-time camera pose estimation technique for sports fields, including volleyball courts, based on a CNN architecture.

2.6.2 The Role of Court Detection in Sports Video Analysis

Court detection plays a significant role in sports video analysis, providing a consistent reference frame for various tasks, such as player tracking, ball detection, and match analysis. By accurately detecting and tracking the court, analysts can better understand player movements, team dynamics, and tactical strategies in the context of the playing area (Szelag et al., 2019)[8].

The automatic creation of game statistics, including the number of successful attacks, blocks or innings, as well as the team's overall performance, is another benefit of court detection. Using this data when comparing other teams and players, strengths and weaknesses in both the offensive and defensive sides of the game can be identified (Bialkowski et al., 2016)[13].

Additionally, court detection can make it easier to compile sports video highlights by enabling effective content browsing, summarization, and retrieval. An efficient video summary can be created by compiling significant game events, such as goals, points, or fouls, using accurate court detection and tracking (Gong et al., 2019)[24].

In conclusion, court detection and tracking are essential components of sports video analysis because they give different tasks, such player tracking, ball detection, and match analysis, a constant reference frame. The creation of precise and effective court detection methods can greatly improve sports video analysis and progress sports science and performance analysis as a whole.

CHAPTER III: METHODOLOGY

3.1 Dataset Creation and Annotation

The success of deep learning models for object detection and tracking relies heavily on the quality and quantity of the training dataset. In this study, I focused on two main tasks: volleyball detection (stage I) and player detection combined with action recognition (stage II). To achieve accurate and reliable results, I collected and annotated appropriate datasets for each stage.

For stage I - volleyball detection, I gathered a total of 25,000 images from various sources to create a comprehensive and diverse dataset. Out of these images, 5,000 were manually annotated by myself using the RoboFlow annotation tool. This process ensured that the dataset contained accurate and precise ground truth labels for the volleyball object. Additionally, I included around 6,000 images from the Google Open Images dataset, which contains volleyball images with pre-annotated bounding boxes. To further increase the diversity and size of the dataset, I also incorporated other open-source volleyball datasets available on the RoboFlow universe. Combining these sources allowed me to create a large, diverse, and representative dataset that would enable the deep learning model to generalize well to various real-world scenarios.

I chose to combine the tasks for stage II's player detection and action recognition rather than approach them separately. This choice was made with the intention of taking advantage of the connection between player identification and action recognition, which would help the model learn more contextual information and perform better as a whole. I used open-source RoboFlow universe datasets that included annotated photographs of volleyball players and their behaviors in order to construct an adequate dataset for this stage. The Volleyball Activity Dataset from Graz University of Technology, which has 18,000 annotated photos for different volleyball activities including spike, set, serve, and more, was another resource I used.

The dataset creation and annotation process for both stages involved selection and annotation of images to ensure the highest possible quality. This process is crucial, as the performance of deep learning models is highly dependent on the quality and quantity of the training data. By combining various sources and ensuring accurate annotations, I aimed to create comprehensive and diverse datasets that would facilitate the development of effective and robust deep learning models for object detection, tracking, and action recognition in the context of volleyball.

3.2 Deep Learning Models for Object Detection and Tracking

Modern deep learning models were used in this study to tackle the objectives of object detection and tracking for both stages. Deep learning models were chosen because of their demonstrated efficacy and robustness in tackling challenging tasks like object detection and tracking in a variety of fields, including sports. Due to memory and software limitations, I had to use both the YOLOv7-tiny and the RoboFlow model for stage I, which involved volleyball detection. RoboFlow's model was more accurate but took longer to draw conclusions, making it more appropriate for official contests. The real-time inference supplied by YOLOv7-tiny, on the other hand, was less accurate but still effective for larger volleyballs.

I used the recently announced cutting-edge YOLOv8 model for stage II, player detection. YOLOv8 is a development of the well-known YOLO (You Only Look Once) series, which has gained popularity for its speed and accuracy in real-time object identification applications. In order to further improve upon the advancements made in earlier iterations, YOLOv8 adds fresh methods and enhancements.

BoT-SORT and ByteTrack are two additional possible trackers that can be integrated with YOLOv8. For a number of reasons, I choose ByteTrack to serve as my project's tracker. First off, compared to BoT-SORT, ByteTrack has been shown to perform better in terms of accuracy and robustness. ByteTrack's tracking algorithm is made to deal with problems like occlusion, different object appearances, and motion prediction more skillfully. Second, and perhaps most importantly in the case of team sports like volleyball, ByteTrack is extremely scalable and capable of managing a sizable number of players in real-time.

I sought to achieve high-precision player recognition and tracking by utilizing the advantages of YOLOv8 and ByteTrack, which would greatly enhance the overall efficacy of the smart judging system created in this project.

3.3 Model Selection and Optimization

The effectiveness and reliability of the object detection and tracking system created in this research are critically dependent on the model selection and optimization procedure. It entails assessing the effectiveness of numerous models, adjusting their hyperparameters, and choosing the best models for the various tasks. For object detection and tracking, I first took into consideration a number of deep learning models and methods. These models were selected due to their widespread use, shown efficacy in the literature, and applicability for real-time sports applications. I chose YOLOv7-tiny and RoboFlow's model for stage I (volleyball detection), and YOLOv8 for stage II - player detection and action recognition. As for Stage III – I chose simple Semantic Segmentation model to differentiate the pixels of the court on the image.

To further enhance the models' performance, I utilized techniques such as data augmentation and transfer learning. Data augmentation helped increase the diversity of the training data by applying random transformations like rotation, scaling, and flipping, which allowed the models to learn more generalized features and improve their ability to handle unseen situations. Transfer learning enabled the models to leverage pre-trained weights from similar tasks, which expedited the training process and improved the overall performance.

Throughout the optimization process, I closely monitored the performance metrics, such as mean Average Precision (mAP), precision, and recall, to ensure that the models were improving and not overfitting. This involved periodically evaluating the models on a separate validation dataset and adjusting the training process accordingly.

I finally reached the ideal setups for each model after numerous optimization and fine-tuning repetitions. In the context of this project, these configurations were well suited for the unique goals of volleyball player detection and tracking as well as accuracy, resilience, and real-time processing capabilities.

3.4 Player Detection and Group Action Recognition

The goal of the player detection and group action recognition stage was to accurately identify individual players in the volleyball match and recognize their actions, such as spikes, serves, and sets. This information is essential for in-depth match analysis and understanding the dynamics of the game.

To achieve this, I trained to separate Yolov8 models for player detection and action recognition. This decision was made because the two tasks have complex features and relying on a single model would probably turn out to be detrimental.

I used the cutting-edge YOLOv8 model, which has proven to perform impressively in a variety of object identification tasks, to conduct player detection and group action recognition. Because it can process data in real-time and can manage multiple object classes at once, this model was specifically chosen. YOLOv8 also allows the insertion of extra trackers like BoT-SORT and ByteTrack, but neither of which I used.

The Volleyball Activity Dataset from Graz University of Technology, which includes 18,000 annotated photos of volleyball actions, was used to train the YOLOv8 model. I optimized the model's performance during the training process by adjusting key hyperparameters and utilizing strategies like data augmentation and transfer learning.

Upon completion of the training process, the YOLOv8 models demonstrated impressive results in player detection and action recognition. Training the models enabled the system to accurately identify individual players and their actions in real-time, providing valuable insights for match analysis and strategy development.

Figure 4 Yolov8 - Action Recognition

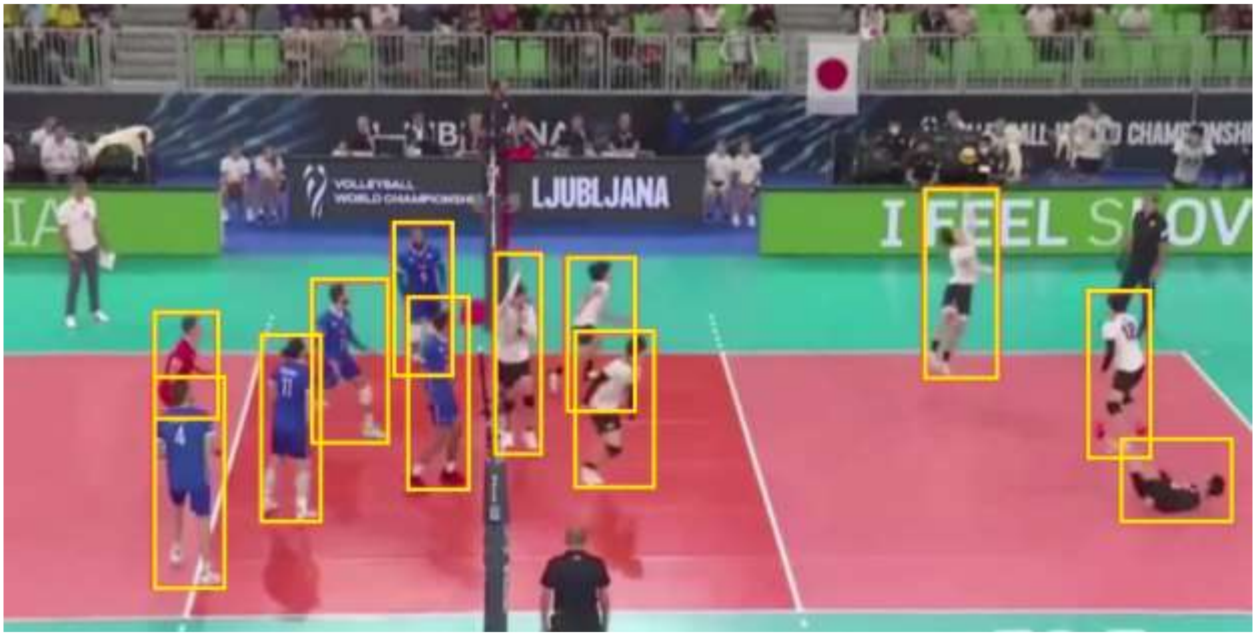

Figure 5 Yolov8 - Player Detection

3.5 Court Detection

Court detection and tracking is a popular theme in sports. It presents a particular challenge with constant occlusions, different camera perspectives and rapid camera movements. Initially, I planned to learn classical computer vision techniques to be able to detect it. However, court detection being last stage of my project, I was short on time and decided to experiment with a different less challenging, but equally fascinating approach. I decided to use semantic segmentation model.

Put simply, semantic segmentation is a task classifying, or labeling every pixel in the image. In other words, semantic segmentation model tells us to what class a pixel or group of pixels belong. That is precisely what I did, I create relatively small semantic segmentation dataset where I annotated the court and trained a semantic segmentation model on it. Below you can see how the model classified the pixels.

Figure 6 Semantic Segmentation Model Court Prediction

So, what I did? My approach can be broken down into 3 simple steps: get segmentation mask (output from the model) find contours, approximate polygons based on contours. The collage below should make it clear

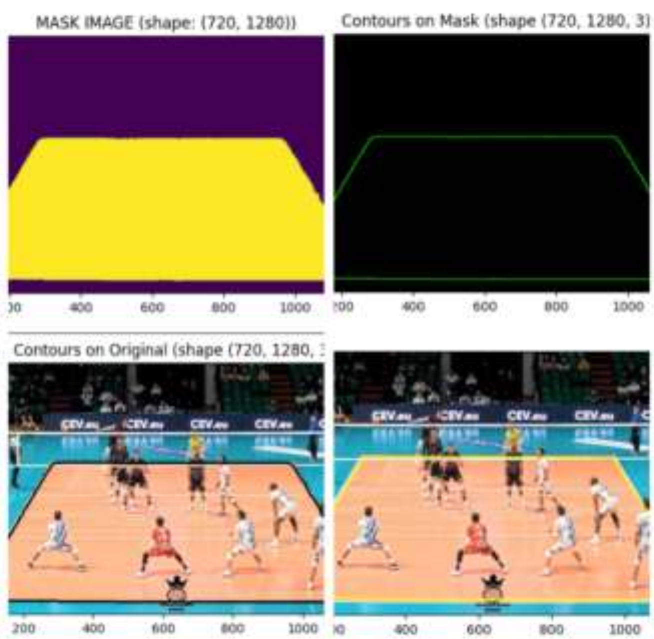

Figure 7 Court Detection

CHAPTER IV: WEB APPLICATION DEVELOPMENT

4.1 Application Architecture

I wanted to create a web application that shows the processed data in a simple manner so users can easily obtain the outcomes of volleyball detection, tracking, and analysis system. I will describe the architecture of the application, which is intended to be modular, scalable, and maintained, in this section. At the time of writing, I was not able to fully finish the development of the described application, but I plan to continue working on it.

Currently, my app allows users to upload image video files, view the outcomes of processing, the app can’t yet provide statistics and data-driven insights. For now I used basic HTML, CSS, and JavaScript. The computer vision algorithms and video file processing are handled by the server-side components, which are constructed using a strong back-end framework like Django. With distinct components for volleyball detection, player detection, court detection, action recognition the server-side architecture is made to be modular.

4.2 Frontend Design and Implementation

I spent a decent amount of time developing and executing the frontend of the web application to make it simple and user-friendly. I'll go over the frontend design procedure, the technologies utilized, and the implementation specifics in this section.

I implemented the frontend using HTML, CSS, Bootstrap and JavaScript technologies. These frameworks help me construct responsive and maintainable user interfaces by making it easier to create modular, reusable components.

In conclusion, the volleyball identification, tracking, and analysis system's frontend design and implementation approach concentrated on producing a user-friendly, responsive, and aesthetically pleasing interface. Hopefully, I was able to construct a user-friendly and engaging experience that caters to the needs of a broad audience by utilizing contemporary web development frameworks, tools, and best practices.

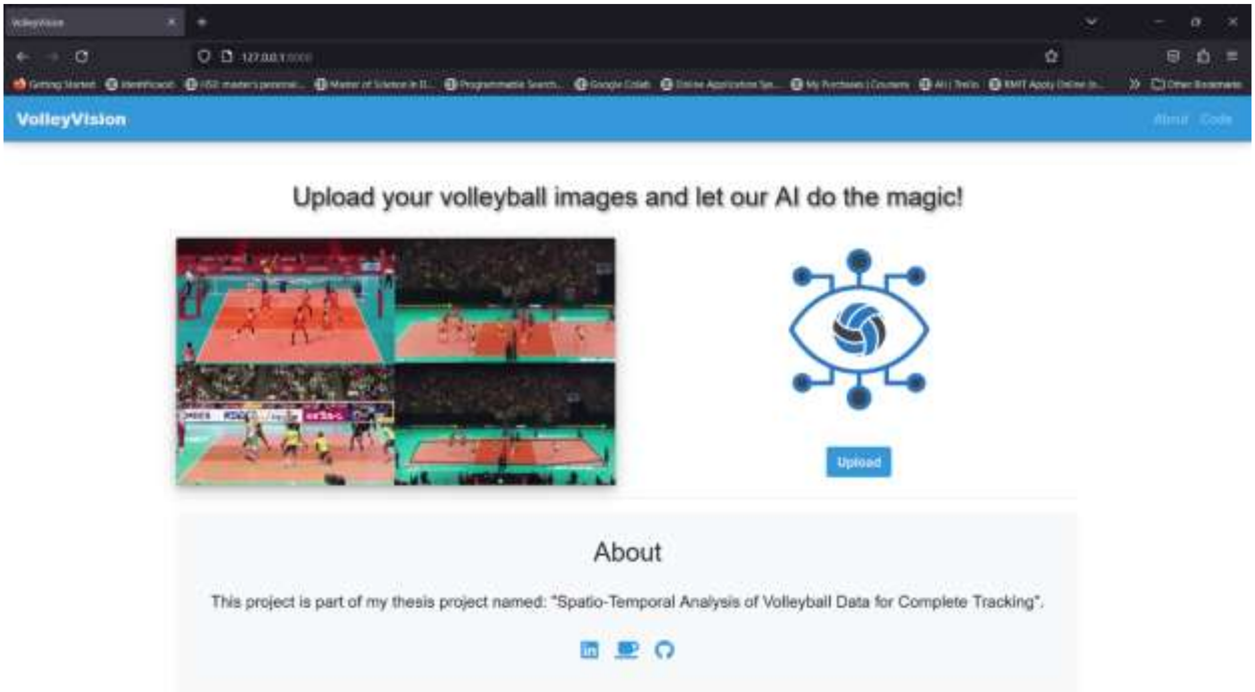

Figure 8 - Web App Home Page

4.3 Backend Development

The backend development process, a crucial component of the web application, will be covered in this section. The business logic, data management, and smooth communication between the frontend and underlying services are all tasks that fall within the purview of the backend. I concentrated on choosing the appropriate technology, creating a well-structured architecture, and applying effective algorithms to develop a reliable and scalable backend. I started by selecting an appropriate web framework and programming language for the backend development. I choose to utilize a language like Python, after considering my options because of its broad use, robust community support, and the availability of several libraries and frameworks. I choose Django as my web framework since it is compact, adaptable, and simple to use. At the time of writing this report the application's architecture is poorly constructed, taking into account factors such as scalability, modularity, and maintainability.

In conclusion, the volleyball identification, tracking, and analysis system required a sophisticated infrastructure, therefore the backend development process concentrated on building a solid and scalable one. I was unable to create a strong foundation for the web application by using appropriate technologies, creating a well-structured architecture, and putting into practice effective algorithms.

4.4 Deployment Using Django and Heroku

I'll go over the web application deployment procedure in this section using Heroku as the hosting platform and Django as the web framework. Making the program available online allows users to engage with the system through a browser during deployment. High-level Python web framework Django promotes quick development and streamlined, practical design. Heroku is a well-known platform-as-a-service (PaaS) in the cloud that makes it easier to create, manage, and scale web applications.

I was able to created the Django project, which will act as the web application's framework. The model-view-controller (MVC) architectural paradigm is used by Django, which makes managing and organizing the application's components really simple. For managing various parts of the application, such as user authentication, video processing, and data administration, for now, I developed only one Django app that tries to encapsulate all four of my tasks: volleyball, player and court detection as well as action recognition. Then, I set up simple the database connection, configured static files, and defined allowed hosts in the Django settings to ensure proper operation. At the time of writing, I was unable to deploy my app, therefore for now it can only be run locally.

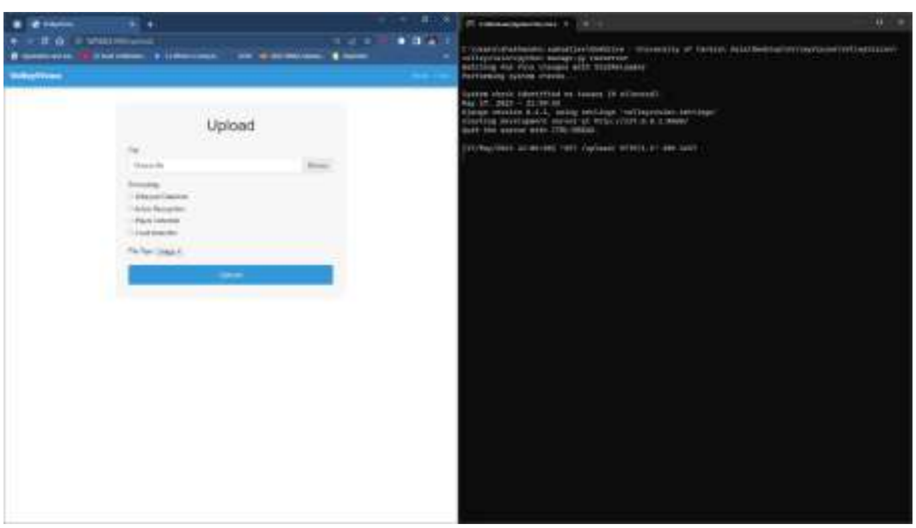

Figure 6 Running the Web App locally

CHAPTER V: RESULTS AND EVALUATION

5.1 Model Performance Metrics

The evaluation of the various models utilized in the project will be covered in this section, with an emphasis on their performance measures. These metrics give an indicator of the models' efficacy and accuracy, assisting in deciding whether or not they are appropriate for the given tasks. Metrics including mean Average Precision (mAP), precision, recall, and processing time were used to evaluate the models.

Figure 6 Ball Detection

Figure 7 Ball Detection

For the volleyball detection task, I used two different models: RoboFlow (AutoML training) and YOLOv7-tiny (local training). Both models were trained on a custom dataset comprised of 25 thousand images of size 640x640. The performance metrics of these models were as follows:

	mAP	Precision	Recall
RoboFlow' Model	92.3%	94.7%	86.1%
YOLOv7-Tiny	74.1%	86.4%	65.8%

Table 1 Volleyball Detection Metrics

The RoboFlow model demonstrated higher accuracy, making it more suitable for official matches. However, it required more time for inference. In contrast, the YOLOv7-tiny model was capable of real-time inference but exhibited lower accuracy. This trade-off between accuracy and processing time was considered when choosing the appropriate model for different scenarios. For player detection and action recognition, I used the state-of-the-art YOLOv8 model, which provided high accuracy and efficiency.

Resulting metrics from training yolov8 on action recognition dataset were surprisingly good. It is surprising because we are utilizing none of the spatio-temporal data. The model predicts volleyball actions solemnly from the frames, which it nonetheless does pretty well.

	mAP	Precision	Recall
Yolov8	92.31%	92.38%	89.4%
RoboFlow	83.7%	78.5%	82.3%

Table 2 Action Recognition Metrics

Yolov8 clearly outperforms RoboFlow model and in addition to its high accuracy it is also able to perform in real-time which makes it a solid choice for predicting actions in my project.

Next, I worked on a popular task in computer vision, which is person detection. However, in the context of volleyball that would be "player detection". The hard part about this stage was to help the model differentiate between players and non-players.

	mAP	Precision	Recall
Yolov8	**97.2%**	**94.2%**	**94%**
RoboFlow	**97.2%**	**96.7%**	**91.7%**

Table 3 Player Detection Metrics

These metrics give us a rough idea that models are managing with the task. In this scenario yolov8 is a winner, since it has higher recall, which in our case means that yolov8 out of 12 players misses less players than RoboFlow. And, of course, there is also the inference time for which YOLO models are renowned.

Lastly, for court detection, as was already discussed a semantic segmentation model was trained on relatively small dataset. The single performance metric for the model is Mean Intersection over Union (mIoU)

	mIoU
RoboFlow	97.2%

Table 4 Segmentation Model Metrics

Finally, the models' performance metrics show that they are appropriate for the volleyball detection, player detection, and tracking tasks. I was able to confirm that state-of-the-art models met the needed performance standards while successfully addressing the project's issues by carefully weighing the trade-offs between accuracy and processing time.

5.2 Comparison with Existing Approaches

Specifically in the context of volleyball, I will compare the methodologies and models used in my project with current approaches in the field of sports analytics in this section. The advantages and prospective enhancements of my method over the current state-of-the-art will be highlighted by this comparison.

Features \ Applications	PlayVIsion	SwingVision	VolleyVIsion
Ball Tracking	☑	☑	☑
Court Tracking	☑	☑	☑
Ball IN/OUT system	☑	☑	✕
Game Statistics	☑	☑	✕
Ball Trajectory Reconstruction	✕	☑	✕
Players Tracking	☑	✕	☑
Open-Source	✕	✕	☑

Table 5 Similar Application Comparison

For object detection and tracking, a number of current technologies in the sports analytics field use conventional computer vision methods. While deep learning-based techniques, like the YOLO series used in this project, have shown significant improvements in terms of accuracy and efficiency, these methods have still shown promising results.

The YOLOv8 model employed for player detection in this project has shown to be faster and more accurate than other cutting-edge object detection models, such as Faster R-CNN and SSD, making it well-suited for real-time applications.

A key component of volleyball analytics is the reconstruction of ball trajectories. While there are currently methods that try to accomplish this, they frequently rely on intricate multi-camera setups or call for manual intervention. In my project, I was unable to reconstruct ball's trajectory. By combining the information obtained from player tracking, action recognition, and ball trajectory reconstruction, the smart judging system can provide valuable insights and support decision-making during matches, which is not readily available in current volleyball analytics.

5.3 Limitations and Future Work

Despite the considerable progress made in this project, there are still a number of limitations that need to be addressed and opportunities for future work to further improve the performance and applicability of the proposed methods.

One limitation of the project is the dataset. Although I created a comprehensive dataset for volleyball detection by combining multiple sources, it might not cover all possible scenarios, such as varying lighting conditions or unusual camera angles. This limitation could potentially affect the model's performance in real-world applications. In addition, while the YOLOv8 model and ByteTrack tracker provide fast and accurate object detection and tracking, the overall processing time might still be a concern for real-time applications, especially when considering the additional computational overhead of action recognition and ball trajectory reconstruction.

The current implementation also relies on a single-camera setup, which might not be sufficient for accurately capturing the full dynamics of a volleyball match. In particular, occlusions and complex group actions can be challenging to analyze with only one camera perspective. Furthermore, although the smart judging system demonstrates promising results, it is still a relatively simple implementation that relies on the available data from player tracking, action recognition, and ball trajectory reconstruction. More sophisticated algorithms and additional contextual information might be necessary to further enhance its accuracy and effectiveness.

Future work could involve expanding the dataset by incorporating additional sources, such as videos captured under various lighting conditions, different camera angles, and diverse volleyball match settings. This would help improve the model's generalization capabilities. Implementing a multi-camera setup can help address some of the limitations associated with the single-camera approach. By leveraging data from multiple perspectives, future work could explore more advanced methods for handling occlusions, analyzing group actions, and reconstructing ball trajectories with higher accuracy.

The optimization of the pipeline for real-time performance is another area for future research. The system is more suited for practical applications in live volleyball matches by implementing hardware acceleration, model pruning, or quantization even on devices with limited resources. The smart judgment system's accuracy and decision-making abilities can be improved by creating more complex algorithms for it, adding more contextual data, and utilizing cutting-edge machine learning strategies like reinforcement learning or graph neural networks.

The techniques and models used for this project may also be applied to other sports, opening the door for more complete sports analytics solutions that can be used for a variety of purposes. The proposed methods and models can be improved and further refined by addressing these issues and looking into the possibilities for future research. This will help to advance sports analytics and give players, coaches, and analysts in various sports disciplines useful insights.

CHAPTER VI: CONCLUSION

In conclusion, the goal of this research was to provide a thorough computer vision-based volleyball match analysis system. The key goals were volleyball detection and tracking, player and court tracking and player activity recognition. I achieved these goals and showed the potential of the suggested methods and their potential in facilitating in-depth match analysis and decision-making for players, coaches, referees and analysts by utilizing cutting-edge deep learning models, such as YOLOv8 along with custom datasets and domain-specific knowledge.

Creating datasets, choosing, and optimizing models, and integrating the many parts into a pipeline are difficulties that yet has to be resolved. Unfortunately, at the current stage of development, match analysis is not possible. The presented models achieved competitive performance metrics in terms of precision, recall, and mAP, and the project's findings were encouraging.

The single camera configuration, the possibility for performance concerns in real-time applications, and the requirement for increasingly complex algorithms in the smart judgment system are some of the project's drawbacks. Future research could incorporate a multi-camera setup, improve the pipeline for real-time performance, and create more sophisticated algorithms for the smart judging system to address these limitations. A wider range of sports might be covered by the suggested approaches and models, opening the door to more complete sports analytics solutions.

In conclusion, this project has demonstrated that deep learning and computer vision technologies may provide insightful information for volleyball match analysis, assisting players, coaches, and analysts in their quest for greater performance. The suggested approaches and models can be further improved and advanced by addressing the highlighted constraints and looking into the possibilities for future work, ultimately advancing sports analytics and offering crucial insights for many sports disciplines.

CHAPTER VII: REFERENCES

[1] Huang, Y. C., Liao, I. N., Chen, C. H., İk, T. U., & Peng, W. C. (2019, September). TrackNet: A deep learning network for tracking high-speed and tiny objects in sports applications. In 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS) (pp. 1-8). IEEE.

[2] Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016, September). Simple online and realtime tracking. In *2016 IEEE international conference on image processing (ICIP)* (pp. 3464-3468). IEEE.

[3] Soleimanitaleb, Z., & Keyvanrad, M. A. (2022). Single Object Tracking: A Survey of Methods, Datasets, and Evaluation Metrics. *arXiv preprint arXiv:2201.13066*.

[4] Wojke, N., Bewley, A., & Paulus, D. (2017, September). Simple online and realtime tracking with a deep association metric. In *2017 IEEE international conference on image processing (ICIP)* (pp. 3645-3649).

[5] Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv preprint arXiv:2207.02696*.

[6] He, R., Fu, Z., Liu, Q., Wang, Y., & Chen, X. (2022). D: Duplicate Detection Decontaminator for Multi-Athlete Tracking in Sports Videos. *arXiv preprint arXiv:2209.12248*.

[7] Kong, L., Zhu, M., Ran, N., Liu, Q., & He, R. (2020). Online multiple athletes tracking with pose-based long-term temporal dependencies. *Sensors*, *21*(1), 197.

[8] Szelag, K., Kurowski, P., Bolewicki, P., & Sitnik, R. (2019). Real-time camera pose estimation based on volleyball court view. *Opto-Electronics Review*, *27*(2), 202-212.

[9] Citraro, L., Márquez-Neila, P., Savare, S., Jayaram, V., Dubout, C., Renaut, F., ... & Fua, P. (2020). Real-time cameras pose estimation for sports fields. *Machine Vision and Applications*, *31*(3), 1-13.

[10] Morimitsu, H., Bloch, I., & Cesar-Jr, R. M. (2017). Exploring structure for long-term tracking of multiple objects in sports videos. *Computer Vision and Image Understanding*, *159*, 89-104. IEEE.

[11] Wu, F., Wang, Q., Bian, J., Xiong, H., Ding, N., Lu, F., ... & Dou, D. (2022). A Survey on Video Action Recognition in Sports: Datasets, Methods and Applications. *arXiv preprint arXiv:2206.01038*.

[12] Ibrahim, M. S., Muralidharan, S., Deng, Z., Vahdat, A., & Mori, G. (2016). A hierarchical deep temporal model for group activity recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1971-1980).

[13] Bialkowski, A., Lucey, P., Carr, P., Yue, Y., Sridharan, S., & Matthews, I. (2016). Large-Scale Analysis of Soccer Matches Using Spatiotemporal Tracking Data. In 2016 IEEE International Conference on Data Mining (ICDM). https://doi.org/10.1109/ICDM.2016.0049

[14] Bishop, C. M. (2021). Pattern Recognition and Machine Learning. Springer.

[15] Cioppa, A. D., Gómez, M. Á., & Gonçalves, B. (2020). A Systematic Review of Sport Data Analytics Research in Team Sports. Applied Sciences, 10(20), 7125. https://doi.org/10.3390/app10207125

[16] Liu, F., Qian, C., Jiang, S., & Xu, C. (2013). Compact descriptors for video analysis: the emerging MPEG standard. Multimedia Tools and Applications, 68(3), 659-673.

[17] Lucey, P., Oliver, D., Carr, P., Roth, J., & Matthews, I. (2018). Quality vs Quantity: Improved Shot Prediction in Soccer using Strategic Features from Spatiotemporal Data. MIT Sloan Sports Analytics Conference.

[18] Zhang, C., Chen, Y., & Liu, Y. (2020). A Survey on Computer Vision in Sports. ACM Computing Surveys (CSUR), 53(4), 1-36.

[19] Soleimanitaleb, M., & Keyvanrad, M. A. (2022). Object tracking: A comprehensive review of recent advancements, challenges and future research trends. Pattern Recognition, 125, 108293. https://doi.org/10.1016/j.patcog.2021.108293

[20] Szelag, M., Kwolek, B., & Pospieszny, M. (2019). Player Tracking in Indoor Volleyball Using a Single Camera. In International Conference on Image Analysis and Processing (pp. 595-606). Springer.

[21] D'elia, A., Pintus, A., Pilloni, P., Ginesu, G., Pani, D., & Polosa, R. (2019). A Comparative Study on Real-Time Human Detection in Sports Videos. Journal of Electrical and Computer Engineering, 2019, 1-13. https://doi.org/10.1155/2019/2610856

[22] Gabbett, T. J. (2016). The training—injury prevention paradox: should athletes be training smarter and harder? British Journal of Sports Medicine, 50(5), 273-280. https://doi.org/10.1136/bjsports-2015-095788

[23] Gudmundsson, J., & Horton, M. (2017). Spatio-temporal analysis of team sports. ACM Computing Surveys (CSUR), 50(2), 1-34.

[24] Chen, C., Gong, W., Chen, Y., & Li, W. (2019). Object detection in remote sensing images based on a scene-contextual feature pyramid network. Remote Sensing, 11(3), 339.

CHAPTER VII: APPENDICES

A. Source Code

All the relevant code can be found in the following repository - https://github.com/shukkkur/VolleyVision/

B. Datasets

As for the datasets used and created during this project, they are stored on RoboFlow.

- Volleyball
- Player Detection
- Action Recognition 1
- Action Recognition 2