Virtual Coach: Personalized Training Recommendations for Football Players Using Machine Learning and enhance the players performance

MSc Research Project
M.Sc. Data Analytics

Joshua Jaymon
Student ID: x23314401

School of Computing
National College of Ireland

Supervisor: Eamon Nolan

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name: ……. Joshua Jaymon ……………………………………………………………………………

Student ID: ………x23314401…………………………………………………………………………………

Programme: ………MSc Data Analytics ……………………………… **Year:** …2024-2025.

Module: ………MSc Research Practicum…………………………………………………………………

Supervisor: ………Eamon Nolan…………………………………………………………………………………

Submission Due Date: ………11/08/2025……………………………………………………………………………

Project Title: Virtual Coach: Personalized Training Recommendations for Football Players Using Machine Learning and enhance the players performance

Word Count: ………6436………………… **Page Count**……………20………………………………

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: ………Joshua Jaymon……………………………………………………………………………

Date: ………11/08/2025……………………………………………………………………………

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Virtual Coach: Personalized Training Recommendations for Football Players Using Machine Learning and enhance the players performance

Abstract

Do current computer vision and machine learning tools give football coaches a reliable way to shape teams plus individuals through automated performance analysis?

The study shows that YOLO object detection - YOLOv3, YOLOv5, YOLOv8 - combined with MediaPipe tracking sharpens how coaches spot player actions, judge performance along with tailor training advice. The Virtual Coach ingests IAUFD match footage and returns real time detections of players, the ball in addition to key events, but also gesture data for technical scrutiny. Experiments reveal that YOLOv8l balances accuracy, recall next to localization best among the tested models - YOLOv5m delivers the tightest bounding boxes. The outcome confirms that modern analysis tools slot into coaching routines and sharpen decisions.

.

Keywords: Deep learning, Football Analysis, Gesture Detection, Recommendation, Realtime Streaming

1: Introduction

AI was recently involved in designing the outfits of football players for the World Cup, and more to the tournament is near, it is high time to consider what extra effect the AI will be able to deliver to football. Cult soccer favorite by millions, football elicits multifarious feelings which may include passion, fury, impatience etc. Football has in the recent past been transformed by the advanced technology both in the identification of talents and improving on the strategies used in the game. Today, it is not about just having a gadget, but about a gadget that makes a monumental difference to the gameplay and its precision. It is valid to talk about non-substitutions as goal-line technology and Video Assistant Referee (VAR) system has affected decision-making mechanisms.

Figure 1.1: AI use in the football games (**Source: Google**)

AI and machine learning are progressing further; it opens up possibilities of applicability that were not seen in football in earlier times, and these are insights that could transform the game. This project is expected to use some of these technologies in developing the Virtual Coach system that is expected to enhance player performance and reduce incidences of injuries.

1.1 Motivation

Standard form of coaching provides inadequate solutions as regards the personal characteristics of the players and hence the performance suffers and the likelihood of injuries increases. This is because technology such as machine learning and artificial intelligence bring new ways through which the development of the football players can be improved on. The idea of this project is to develop a virtual coach, attempting to improve the state of the art in sports analytics and use of new solutions by creating a tool that brings precise training advice based on the data provided by the player.

1.2 Research Problem

The typical football training approach involves the use of conventional techniques that do not factor into consideration the differences in players and specific game situations and thus amounts to non-efficient training. Some of them include lack of feedback customized according to the specific areas that needs improvement in each player, and lack of training tools that changes based on the problems a player is facing (**Stein, M., Janetzko, H., Seebacher, D., Jäger, A., Nagel, M., Hölsch, J., ... & Keim, D. A. (2019)**). This research intends to cover these gaps by proposing a Virtual Coach that will be implemented via computer vision. Virtual Coach will use YOLO models in detecting the player/s, ball, events that happen during the game and pose estimation. The difficulty is to teach these models to recognize and analyze the football elements and provide the appropriate feedback for players, as well as use the game simulation models properly. This approach aims at improving the quality of training as it provides new and more effective individualized approach to trainings to develop players and improve the training processes in different situations.

1.3 Research objective

This project aims to create a Virtual Coach that assimilates player data to create a personalized training schedule that will be used by individuals and teams with the aim of maximizing the performance and minimizing the risk of injuries. The system applies machine learning algorithms to analyze other factors, including the performance of the players, the selection of shots, and the training sessions to provide individual recommendations. Through these insights, the Virtual Coach offers coaches a smart, AI-based solution that helps develop the players, create an optimal training session, and achieve long-term team success (Moustakidis et al., 2023; Wang et al., 2024).

1.4 Research Questions

RQ1: Which machine learning techniques best evaluate football player data?
RQ2: Which metrics guide tailored training?
RQ3: How does the Virtual Coach slot into existing coaching routines without upheaval?
RQ4: What quantifiable impact does the system deliver on performance and injury avoid

2: Literature Review

The literature review of this study analyzes the previously done work on virtual coaching systems in football in connection to three specific parts: (1) the development of virtual coaches, (2) the evolution of machine learning and AI in sports analytics, and (3) individual training and recommendation systems.

2.1 Research Towards Virtual Coach

Virtual coaching aids are specific; they enhance training as well. According to Gouttebarge et al. (2017) and Raya-González et al. (2021), the systems are supplementary to outdated techniques. They offer individual facts. Stein et al. (2019) were more involved in the player. Sarmento et al. (2014) observed the good among young athletes, due to the data-based personalization. New concepts, including TacticAI (Wang et al., 2024), apply to corner kick research with geometric deep education. The following YOLOv8 based on young football case proved by Gustafsson (2024) indicates that as inexpensive, real-time systems can be achieved. The research proves the increased application of AI-based coaching. However, there are limited comparative reviews across settings.

2.2 Research Towards Advancement of Machine Learning and AI Towards Virtual Coaching

Machine learning allows the study of biomechanics, performance, and injury prevention to a deeper degree. Eckard et al. (2017) and Ramsay et al. (2020) showed that AI models have the ability to recognize risk markers and optimize training loads, whereas Malone et al. (2017) affirmed that they predict injuries in elite football. A set of frameworks, including DeCoach (Ghosh et al., 2022), demonstrates the improvement of player assessment in racket sports with deep learning, and Yang (2022) focused on the importance of parameter tuning in action recognition. In addition to football, Ait-Bennacer et al. (2022) tested CNN-LSTM and ST-GCN on Karate coaching with reported more than 90 accuracies. Together, these papers indicate that AI solutions do not stop at descriptive analytics, but they provide predictive and prescriptive solutions. Most of them, however, are limited to case studies of individual sports, and thus cannot be generalized.

2.3 Research Towards Personalized Training and Recommendation

Virtual coaching focuses on the aspect of personalization. Windt and Gabbett (2017) and Impellizzeri et al. (2019) revealed that custom plans have lower injuries and better performance than generic programmes. Malone et al. (2017) also noted that it is important to monitor training loads and recovery to maximize the results. Greater scale studies, like those by de Leeuw et al. (2022) in elite volleyball demonstrated the capacity of machine learning to discover player-specific predictors of injury, and Miah et al. (2022) used diverse classifiers to wearable data, with over 95% accuracy of predicting fitness. On the same note, Cui et al. (2023) also compared the performance of deep learning models in human activity recognition, and BiGRU had the lowest misclassification rate. The above findings imply that there are high chances of utilizing wearable information and computer vision to customize suggestions.

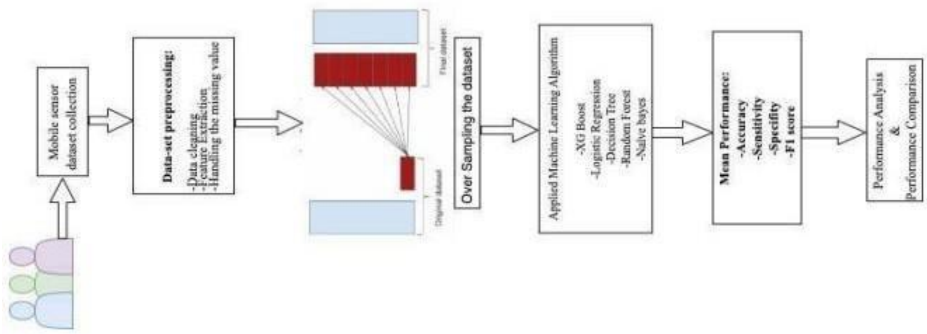
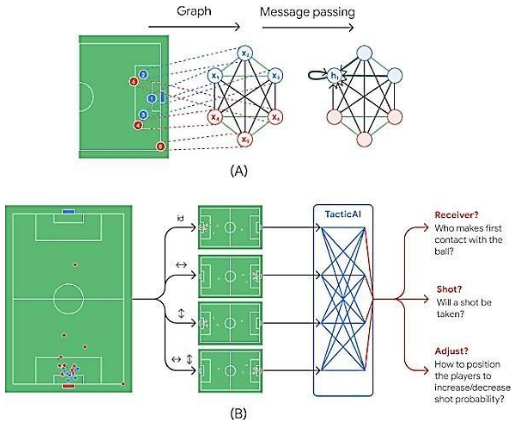


Figure 2.1: MHFit model (**Miah, J., Mamun, M., Rahman, M.M., Mahmud, M.I., Ahmad, S. and Nasir, M.H.B., 2022, December**).

Figure 2.2: Tactical AI for Football (**Mekruksavanich, S. and Jitpattanakul, A., 202**



2.4 Research Gaps

Although things are looking good, there are obstacles. There are not many studies that evaluate comparative performance of coaching systems among various sports or competition levels. Little attention is paid to the cultural, gender and ethical factors (e.g. privacy, agency of players). The existing datasets are frequently disproportional, and it is hard to identify rare and important events (e.g., red cards). Additionally, although there is a strong personalization, there is a lack of strong statistical validation and cross-context generalization.

2.5 Research Plan

This project is one that tries to fill these gaps by including the YOLO object detector and MediaPipe gesture tracker on the IAUFD dataset. The proposed Virtual Coach will provide actionable and specific feedback to individual players and teams by analyzing player motion and events (goals, fouls, cards) to provide both individualized and general information. This forms a gap between previous research on AI-based coaching and practical implementation of injury prevention and performance optimization

3: Methodology

To enhance the knowledge of football games and improve the performance of the sport AI is being adopted in analyzing football games with higher efficiency. Some are based on assessing the players' performance using tracking and data analysis, use of clustering to study the strategies, and formations and passes of the teams, respectively. Furthermore, the use of both historical and real – time data is helpful for prediction of match results. Some of the basic machine learning techniques are as follows: Convolutional neural networks (CNNs) for event detection in the form of videos, time series analysis, reinforcement learning for the prediction of injury based on biomechanics. These technologies also help in the development of strategies, assessment of player fitness and general work towards the improvement of the team.

3.1 Dataset

The use of computer vision for Football has also been very popular and people appeared to analyze the videos using these techniques. Some of the applications of this method include; Event detection, summarization of match videos, outcome prediction, and analysis of statistic. While deep learning algorithms are quite effective when it comes to analyzing images and videos, they often demand large amounts of data. Still, the open sources available in this area are not rich enough or joint enough to allow using deep learning algorithms with great success. In order to overcome this problem, The IAUFD (Islamic Azad University Football) dataset was developed and assembled. This dataset includes 100K real images in total, collected from 33 football games with 2508 minutes' worth of raw video data. The images are annotated and classified into 10 different event types, and this is an area that there is a lack of research materials to date. (Link: https://universe.roboflow.com/university-tboyn/iaufd-football)

Figure 3.1: Sample of 100k images dataset of IAUFD (**Source: IAUFD**)

The dataset categorizes events into ten essential types: impersonal, central, celebration, sending off, yellow card, the same, the ground, the arbiter, the free kick and the corner. It is these categories that are very important for performing the higher-level analysis of actions and events in football. Thus, in order to achieve the best representation, we used different types of weather (sunshine, rain, overcast), different seasons, different times, and different places. In addition, we employed two State-of-art deep neural networks; VggNet-13 and ResNet-18, as the benchmark to assess our dataset towards future studies.

3.2 Machine Learning Detection Models

Through the use of Computer Vision and technologies such as Convolutional Neural Networks (CNNs), this paper aims at training the specified models for event appreciation in videos including goals, fouls or even offsides. Thus, having extracted player movements through computer vision techniques of processing the videos in the case of covering the matches, it could be possible to use machine learning for the evaluation of individual player performance. From these approaches, the YOLO (You Only Look Once) series of models are amongst the most effective object detection frameworks especially for realtime applications.

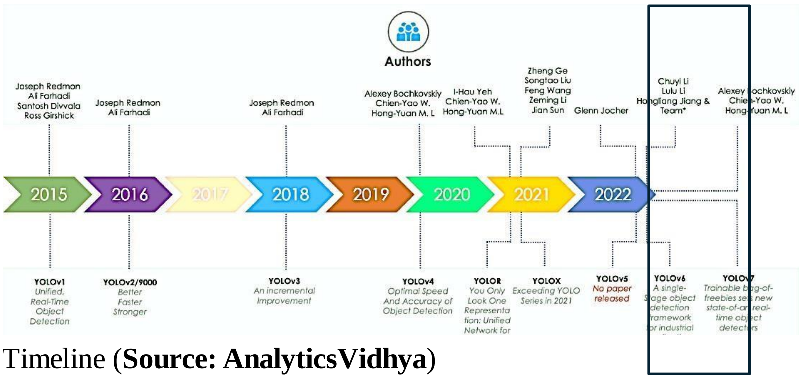


Figure 3.2: Yolo Timeline (**Source: AnalyticsVidhya**)

YOLO models do not require feature extraction and classify objects in regard to bounding boxes as well as class probabilities from the entire image making them perfectly suitable for real-time use. Their grid-based structure enables YOLO to detect more than one object in a single image without having to rerun the detection algorithm on the whole image area as it is done in coordinate regression approaches. This efficiency also helps YOLO become one of the fastest objects detecting techniques since it follows the concept of real-time object detection.

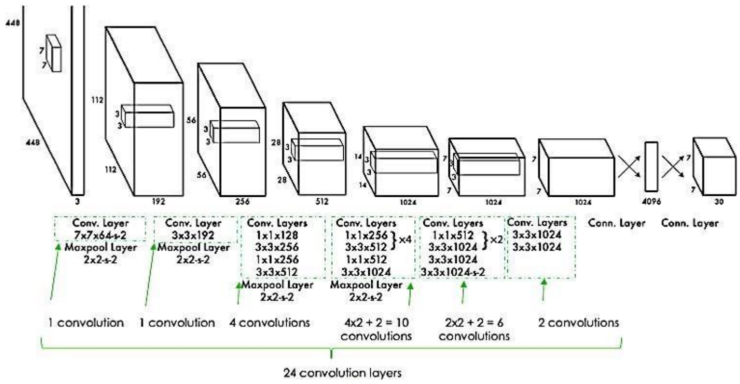


Figure 3.3: Yolo Architecture

More recent versions of the YOLO detectors in real time with faster, correct and improved detection of small objects and occluded objects with an option to fine-tune the model to different environments with less complexities. Such innovations vastly improve in player/ball tracking, biomechanics analysis and recognition of tactical events, and offer more individual and data-oriented suggestions in training and prevention of the injuries compared to previous systems of detection.

3.2.1 YOLOv3

YOLOv3 YOLOv3 is an improvement of YOLOs with several changes for better performance made on it. However, it incorporates Darknet-53 that comes with 53 convolutional layers thus improving feature extraction. This enables YOLOv3 to give out predictions at three different scales to enhance the prediction of objects of varied sizes. Furthermore, in order to predict the objectness probability map and regression coefficients for bounding boxes, logistic regression is used while binary cross-entropy loss is used for predicting the class probabilities, which further boosts up YOLOv3's capability to handle multi-label imaging detection tasks.

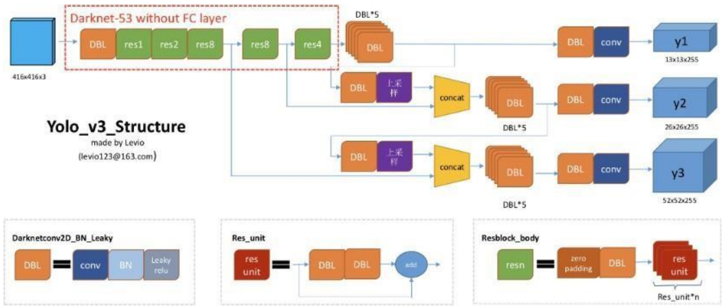


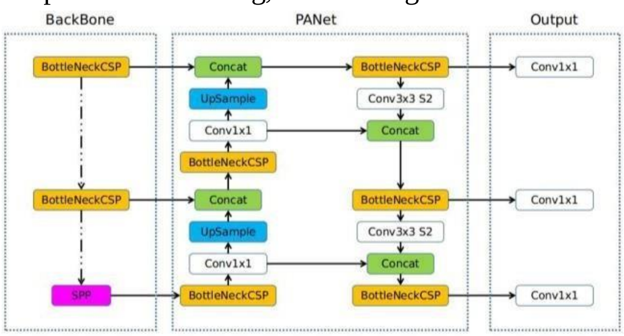
Figure 3.4: YOLOv3 architecture

3.2.2 YOLOv5

YOLOv5 These models are therefore improved models of YOLO that have been developed to address the balance between speed and high accuracy. There are several versions of YOLO including YOLOv5m and YOLOv5x which has differences in their performance and complexity of use:

- **YOLOv5m** is an intermediate complexity model that provides reasonable computation time and acceptable prediction accuracy sufficient for a wide range of tasks.
- **YOLOv5x** This model is the sophisticated one and is the part of theY O L ov5 series. Y O L Ov5-s is the most demanding when it comes to computational power but it has the highest accuracy level among all models of the series. It is suitable for those cases that require high accuracy of computation results and enough computational resources.

To achieve a more effective gradient flow, more effective calculations taking into account the computation budget and an improved ability to learn, the models of YOLOv5 incorporate CSPNet to work out the structure. Multidimensional feature extraction of the spine through the various layers of the network to share Spatial Information derived using CSPM. Moreover, YOLOv5 has mosaic data augmentation which in the process of training, the training involves four different images. The main



advantage of this method is that the outcomes can be enhanced owing to enhanced prediction of the size and rotation of the object.

Figure 3.5: YOLOv5 architecture

3.2.3 YOLOv8

YOLOv8 Is the latest version of the YOLO series since it offers additional development to the detection capacity of the models. One of them is called YOLOv8m which is the medium-sized model, and the other model is named YOLOv8l as one of the large model versions appropriate to the computing resources a person may have. The models are YOLOv8 that contains a refined PANet or Path Aggregation Network that further strengthens the PUSH or Pyramid User Shifted Hierarchical networks that retention of the spatial information and improves the localization of detection of the objects.

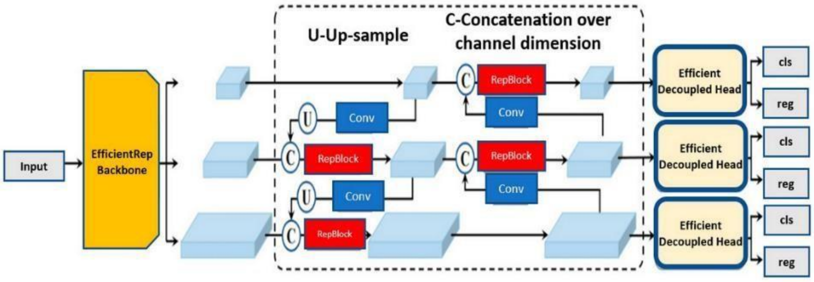


Figure 3.6: YOLOv8 architecture

The YOLOv8 has extra features with the incorporation of the Convolutional Block Attention Module (CBAM) focusing on the meaningful appearance of the object and ignoring the less valuable features thereby increasing efficiency of the detectors especially under complicated scenes. The model also employs the utilization of Dynamic Convolution which is applied to handle the input features that improve the overall generalization of the model at different scales and the shapes of an object. It substitutes ReLU with SiLU (Sigmoid Linear Unit) to lessen problematic issues, like dead neuron, and to smooth out gradients aiding the quick convergence in the training phase. Hence, YOLOv8 offers better or comparable accuracy accuracy to the earlier versions of the YOLO algorithms and this is achieved with fewer epochs and shorter inference time.

3.3 Machine Learning Tracking Models

MediaPipe is a powerful open-source based media by Google that should be used to construct effective multiplatform pipelines to machine learning especially in real-time computer vision. It is built on the concept of directed acyclic graph where nodes represent various stages in the processing pipeline such as, but not limited to, the preprocessing, feature extraction or model inference. Thanks to the modularity of its architecture, parts may be added and integrated to different applications with ease (including hand gestures tracking or whole body pose estimation). It can be attributed to the fact that MediaPipe is multi-threaded and hardware accelerated and optimized and is amazing on the side of real time like a gesture recognition system.

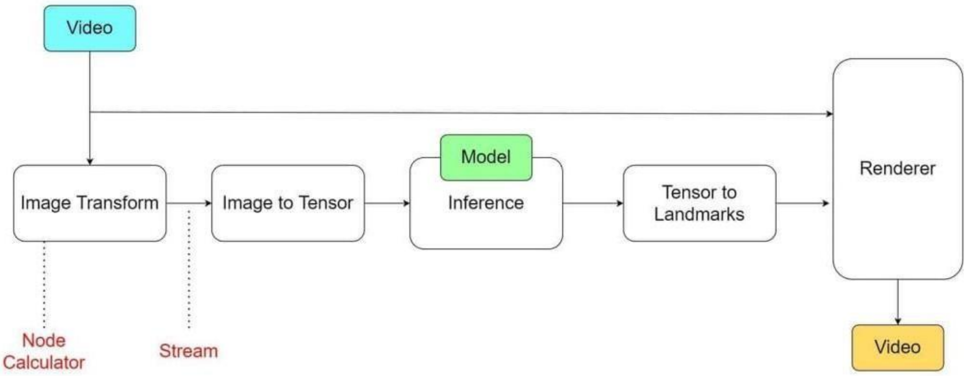


Figure 3.7: Media Pipe algorithm Flow or the tracking of the football and to check the players performance

To enhance the operation of MediaPipe the solution is already consumed with TensorFlow Lite, which makes it run on energy-efficient chipset, such as smartphones or Internet of Things tools. To minimize the cost of computations and still guarantee highest accuracy of the face detection, pose estimation and other tasks some algorithms such as depth wise separable convolution are applied. More detailed models such as face, hand and body pose as well as complex multimodal studies are also possible with the framework. Pliability and extensibility of MediaPipe has played a critical role in augmenting real-time computer vision practices and has offered the solid foundation on which improved vision-based applications are developed both in the field of research and in the business environment.

Figure 3.8: Hand landmarks using Mediapipe (**source**)

3.4 Evaluation

Object identification tasks, such as recognizing individuals or events in football game film, often utilize mAP (mean Average Precision) and box precision.

3.4.1 mAP (mean Average Precision)

mAP or Mean Average Precision is another important measurement used in object detection to find the performance of the model by the precision and recall value upon various classes. It is computed based on the curve of the precision-recall for every class and then finding the mean of the resulting value. In particular, mAP@50 refers to precision at sometimes utilized IoU of 0. 5 The calculated IoU which is the comparison of predicted as well as ground truth bounded box having threshold value of 0. Minimum IoU hence is 0. 5. **mAP50** is popular in object detection evaluation metrics and challenges because it provides a good result with a reasonable degree of precision and IoU enforced.

$$mAP = \frac{1}{N} \sum_{i=1}^{N} AP_i$$

Nevertheless**, mAP50-95** is much stricter and more detailed because it calculates the mAP for various IoU thresholds, starting from 0. 5 to 0. 95 with the interval of 0. 05.

3.4.2 Box Precision

In reality, AP looks at alternative ways of measuring model quality when it comes to object detection where boxPrecision only looks at the accuracy of the bounding boxes outputted by the model. The car and pedestrian detected boxes having high IoU compared to the ground truth are counted correspondingly relative to the total predicted boxes. The total is determined to be the mAP. BoxPrecision does not only inform about the correct object detection, but also about the correct localization of bounding boxes, which is highly important in applications detailed localization is demanded, such as medical imaging or autonomous driving. Whereas mAP computes the overall mean average precision at all classes, intersection over union thresholds, BoxPrecision metrics is biased only at the accurate position of bounding boxes and in cases where the spatial localization is vital, the position can be very privileged.

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union}$$

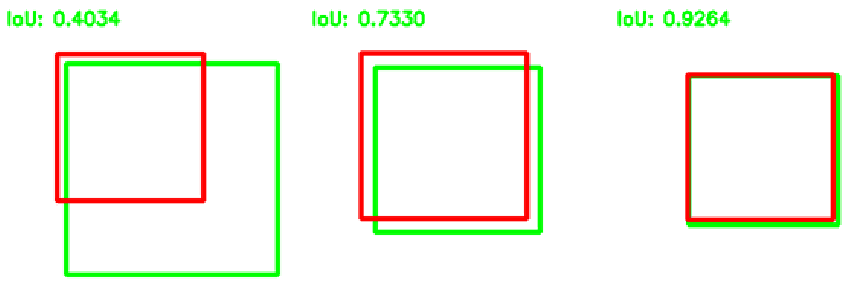


Figure 3.9: The importance of high IoU

4: Implementation

4.1 Flow Diagram

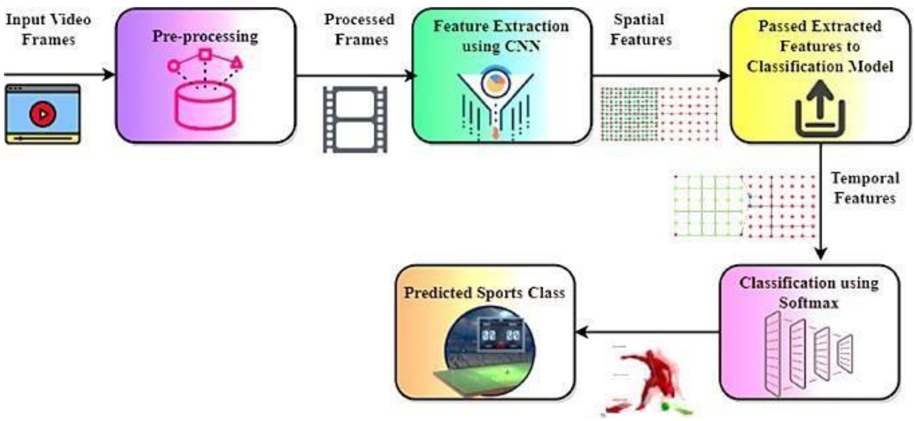


Figure 4.1: Implantation flow for the Virtual coach

Input: annotated video frames from IAUFD and in the real time this will be mimicked to the images taken though the mobile camera or any other image acquisition technique
Output: detected events with bounding boxes, player & ball tracking coordinates, performance analysis summaries.

4.2 Algorithm Flow

Football Events Detection Implementation Steps of the football events detection based on the IAUFD Dataset, Different YOLO models and Gesture + Ball Tracking with MediaPipe:
Step 1: Dataset Preparation- First, IAUFD (International Audio-Visual Football Dataset) data has to be prepared to carry out the event detection. Begin by doing the data preprocessing or in other cases labelling of the data set where applicable and tick on the boxes of the correct co-ordinates of the events like goals and fouls, and ball possession. Divided the data in a training set, a validation set and a test set. When the data set is skewed, then, to handle this data set, the following series of measures have to be considered which are rotation, scaling and flipping of data set of training.

Step 2: Model Selection and Training – In this article, experiments to detect football events have been conducted on the various versions of YOLO (You Only Look Once) such as YOLOv3, YOLOv4, and YOLOv5. Begin by retraining the YOLO models to possess the right anchor boxes in regards to the dataset and alter the input size. Run each of the models on the IAUFD dataset, and record the performance metrics, like mAP and IoU throughout the training of the model. It is also notable that the learning rate can be increased gradually and choose the batch size and varieties of other parameters to achieve the maximum achieved performance. In making judgments about the best training models among the trained YOLO models, work out how useful each of the trained YOLO models is by testing it against the appropriate validation set.

Step 3: Evaluation of the Event Detection-After training your YOLO models, you should test them by their train test set. Due to the assistance of box precision and mean Average Precision (mAP), the object detection model will be analyzed regarding its ability to recognize the different football events. Offer any inconveniences i.e. explain why sometimes it can be nearly impossible to highlight specific examples such as fouls or off sides or circumstances. Based on the above assessment, determine the most accurate and the most fit model to be applied in the real-time setting.

Step 4: Gesture and Ball Tracking using MediaPipe – After that, as a later addition, we should add MediaPipe, so that tracking of gestures and the ball will be enhanced. The mediaPipe already provides pre-conducted pipelines on pose estimation and hand tracking and can thereby be utilized to trace the gestures and reception of the ball of the players. Besides that, consider MediaPipe pose estimation model that is used in evaluating each frame in the video and determining the landmark points of the movement of the players and apply the ability of tracking the position of the ball on the field using the ball-tracking ability. Such a combination of data will facilitate the ease of assessing the performances of players with just a glance through their reaction times, accuracies of handling the balls, and their agility among others.

Step 5: Player Performance Analysis Enhancement- Once following event detection and gesture tracking systems are in place, efforts should be made regarding how player performance can improve with the help of that analysis of the gathered data. A sequence of the detected actions can be used to construct the timeline of the key actions in the progress of the game and compare it to the gesture detection result of MediaPipe. As an example, consider minutiae like how particular orientation of the hand/feet when hitting affects consistency and accuracy of a shot. This discussion can be further applied to supply the players with specific feedback as well as better specify the parameters of their performance that should be enhanced based on the empirical data.

Step 6: Visualization and Deployment – The final step will be to build a tool that would assist in visualizing the detection of the events and performance. This tool can be developed on the Streamlit or Dash technology, which makes it possible to present the detected events and analyzes of gestures in real-time along with the tracking of the ball among coaches and players. Deploy the whole system on a cloud-hosted environment (e.g. AWS, GCP) to handle the extremely high amount of videos of the games. The models themselves would change because, the feedback of the users in the field is continuous, and the component of the system of analysis will be continuously improved.

5: Results and Analysis

5.1 YOLOv5m

These evaluation measures provide the hint regarding the maximum effectiveness of the YOLOv5m model working on the object detection tasks. It has a MAP of 0. 31 in MAP with IoU score of 0. 5 that considers object detection result and provides at least 50 percent object overlap with ground truth therefore indicating moderate accuracy. However, mAP50-95 is lower which is equal to 0. 175, meaning that the proposed model provides close to real-time inference rate and similar accuracy at the same time, the precision of the model declines despite the higher IoU threshold. This is evidence of the fact that there are possibilities of enhancing the accuracy of the object location of existing more specific models.

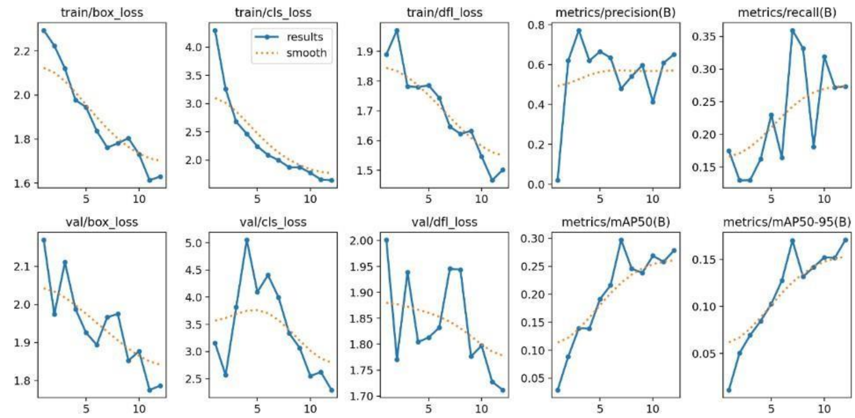


Figure 5.1: Model performance matrices for YOLOv5m

Box precision is another metric utilized to predict the correctness of the bounding boxes and it is fairly good at 0.756. This means that in the event that the model predicts, a majority of the times the boundaries will be correct in accordance to the location of the objects.

5.2 YOLOv5x

They are the training and validation plots of yolov5x model that exposes vital motifs of the training. The train and validation counterparts of the train/box_loss, train/cls_loss and train/dfl_loss have persistent directional negative drift that means that the model is slowly developing an awareness of the task to correctly predict bounding box locations and class label predictions. The calculated values of both precision and recall add certain variation to the considered values of precision and recall, but overall, the improvement in the recognition and marking of the objects is positive.

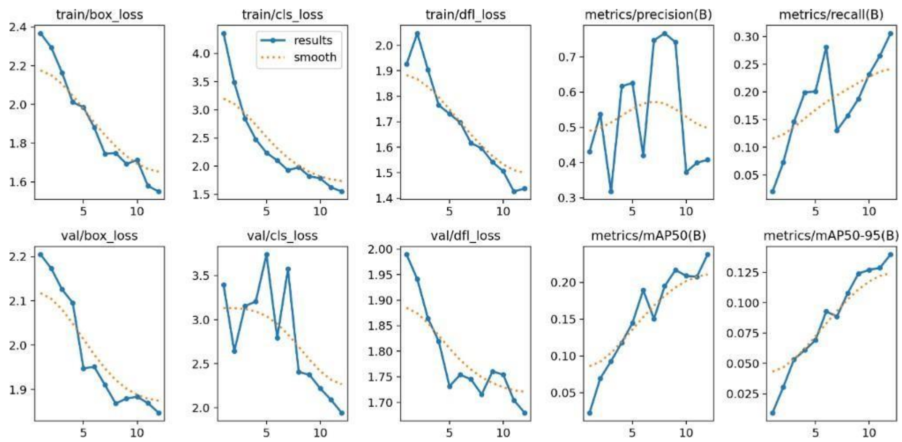


Figure 5.2: Model performance matrices for YOLOv5x

The following are some of the errors that could be quite significant to rectify since they affect model stability. Further, the Box Precision and recall measures of the model at 0.507 and 0.316 respectively imply that the model performs rather well in the pinpoint prediction of the objects of the localization in the terms of the coordinates of the bounding box, but it is inferior in terms of the recall in all classes. The obtained metrics and the confusion matrix also indicate that it is necessary to refine the model more precisely regarding class-specific features and regarding the balance of the dataset to enhance the general performance of the YOLOv5x model.

5.3 YOLOv8m

With the example of the YOLOv8m model, one is able to isolate some patterns of learning during the training and validation. The decline in the training losses (train/box_loss, train/cls_loss, train/dfl_loss) indicates towards the development of the model over the epochs where successful mitigation of the majority of errors pertaining to the bounding box regression, the recognition of classes and distance focal loss were observed. This trend is agreeable with the validation losses which paints one of the merits of the model, that is although the model should be trained with certain data, it should be able to generalize it to other unseen data. The precision-recall curves depicted indicate a pattern of rising values of these metrics despite the variation of some epochs with an indication that the model is variable in the pattern of identifying and categorizing the objects. However, measures of mAP like mAP50, mAP50-95 show clear improvements and therefore it provides a serious increment in the object detecting ability of the model.

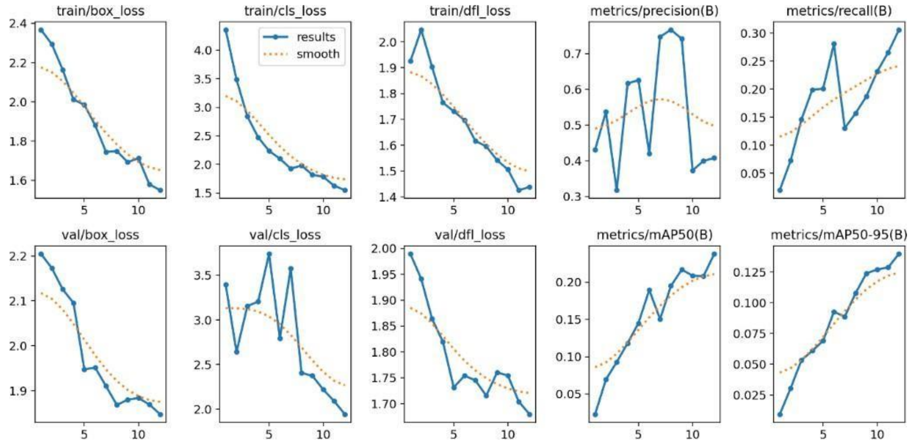


Figure 5.3: Model performance matrices for YOLOv8m

5.4 Yolov8l

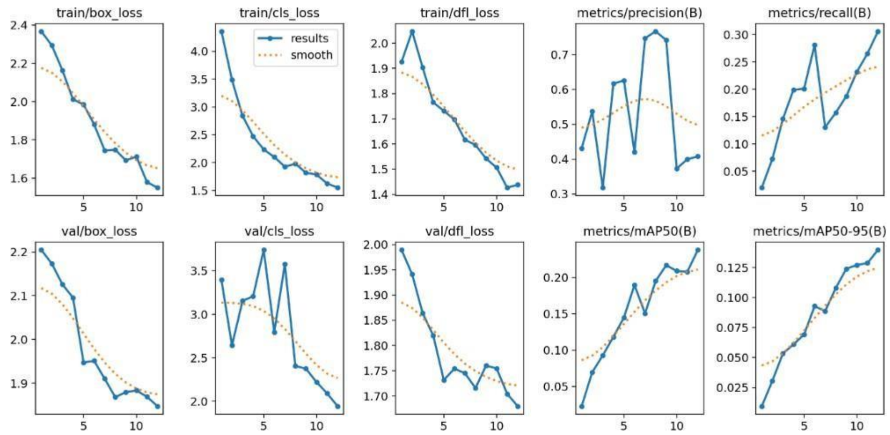


Figure 5.4: Model performance matrices for YOLOv8l

Some of the trends of the performance can be observed in the indication of the training and validations of the YOLOv8l model. All the three tasks of the bounding box localization, classification and distance focal loss are all training losses, and we would expect these to be monotonically decreasing and the model is thus learning satisfactorily and converging to the respective task at hand. It is back with the reducing trend in the validation losses as the trend is also more volatile or over fitting or inconsistent to the model generalization. Oscillation in precision and recall curves is also evident; the highest precision was attained at the epoch 8-9, and the drop in precision onwards started and the recall started growing continuously although some quite large jumps and drops were still evident between epochs. It implies that the precision in identification of the objects will enhance with time but precision and accuracy might be affected owing to over-fitting due to over-fitting or unstable training. The indicators of performance cannot converge but they act averagely and in a positive disposition. The mAP of the YOLOv8l is presented in the following manner; mAP50 = 0.299, and so that the model can be determined in a relatively accurate category under the IoU threshold value is 0.50. The mAP50-95 of 0. 186 which is slightly low nevertheless confirms that the model has a relatively close range of performance regardless of the IoU threshold. The value of Box Precision of 0. 533 shows that the provided model is accurate enough regarding the location prediction of the bounding boxes, however, there is the margin that may be reached. When the percentage scores under recall were associated with the zeros they were multiplied.

5.5 YOLOv3

The training loss of bounding box coordinate (box_loss) and classification (cls_loss) and distance focal loss(dfl_loss) generally decreases on an epoch-by-epoch basis which would imply that the model is learning and learning to give the correct output. In the same manner, although the validation losses are also reducing, it is more volatile with some indication of overfitting or generalization instability in general. Precision (metrics/precision(B)) and Recall (metrics/recall(B) are highly variable, but in general, there is an increase in the recall values, but no apparent improvement in precision. Mean Average Precision (mAP) metrics show a poor performance of both at0. 037 and mAP50-95 at 0. 022. The steady boost of mAP50 to the value above mAP50-95 speaks to the fact that the object detection has not been stagnant, indeed, it advances, but gradually, and in addition, the outcomes can hardly be considered decent.

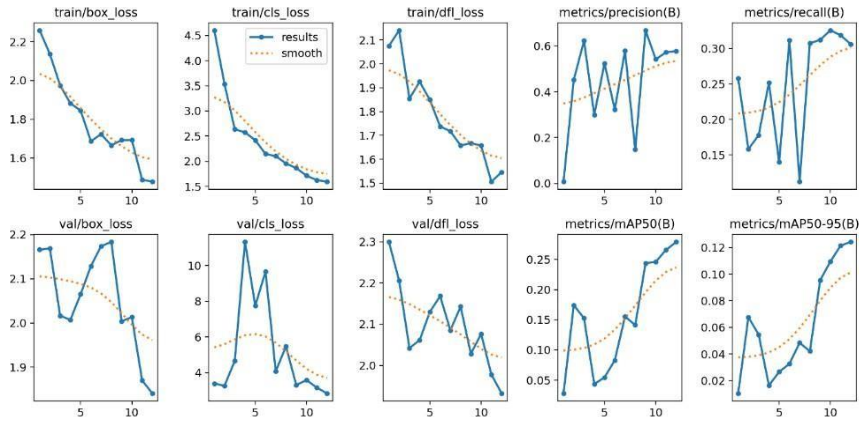


Figure 5.5: Model performance matrices for YOLOv3

5.6 Confusion Matrix Comparison

Confusion matrix is an essential diagnostic instrument through which the performance of model can be evaluated with reference to the accuracy of the performance in regard to various classes. In the case of YOLOv8l, the model delivers finer ongoing results about the classification ability of the model with the help of the matrix. It shows that the model is effective with determining as to who the "players" are and the "background" are, with 61 and 64 accurate classifications respectively. Nonetheless, it performs poorly in other type. It mentions in particular the goalkeeper gate 5 times as player and 43 as background. The model is also not accurate in categorizing the tags like corner, goalkeeper and players wall and also finds difficulties in tags that appear less often like red card or yellow card and these tags get misclassified most of the time. When not sure, the high rates in which they intend to place them under the background classification act to show that the background is the assumed classification by the model under unclear circumstances.

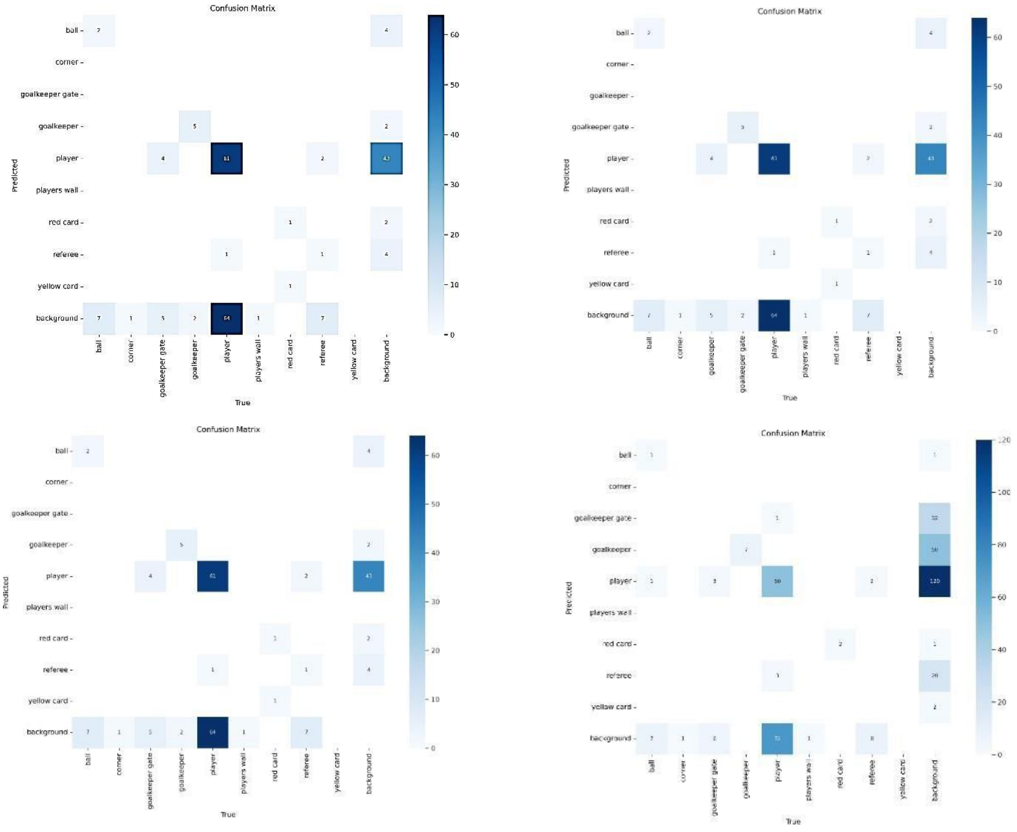


Figure 5.6: Confusion Matrix Comparison of different YOLO models (a) YOLOv5 (b) YOLOv8m (c) YOLOv8l (d) YOLOv3

The confusion matrix for YOLOv8m shows the specific details of the classification done by the model. The correct predictions for "background" (64) and "player" (61) classes are high while the errors are as follows; "player" instances are often misclassified as "background" (43) and "goalkeeper gate" is often misclassified as "player" (5). The matrix also reveals that certain classes such as "corner", "goalkeeper", "players wall" were not predicted at all and all the predictions were wrong. These trends combined with the overfitting of the model to predict the 'background' category make one appreciate the problem of dataset imbalance. and "players" (50 instances). Nevertheless, it performs significantly worse in other categories: it labels 'players' as 'background' in 120 cases, and 'goalkeeper gate' as 'players' in 7 cases. There are no "corners" and "players' walls" in the model and the model performs poorly in recognizing occasional elements such as "red cards" and "yellow cards", which are often misjudged. The "goalkeeper gate" category is particularly mixed up with "background" (32 times). The predetermined presence of the '’background'' class in predictions points to a default probability whenever the model is in doubt. These results show the following requirements for further enhancements to improve the model, thus increasing its capacity to differentiate between the various field elements and to recognize the less frequent scenarios.

5.7 Comparative Analysis

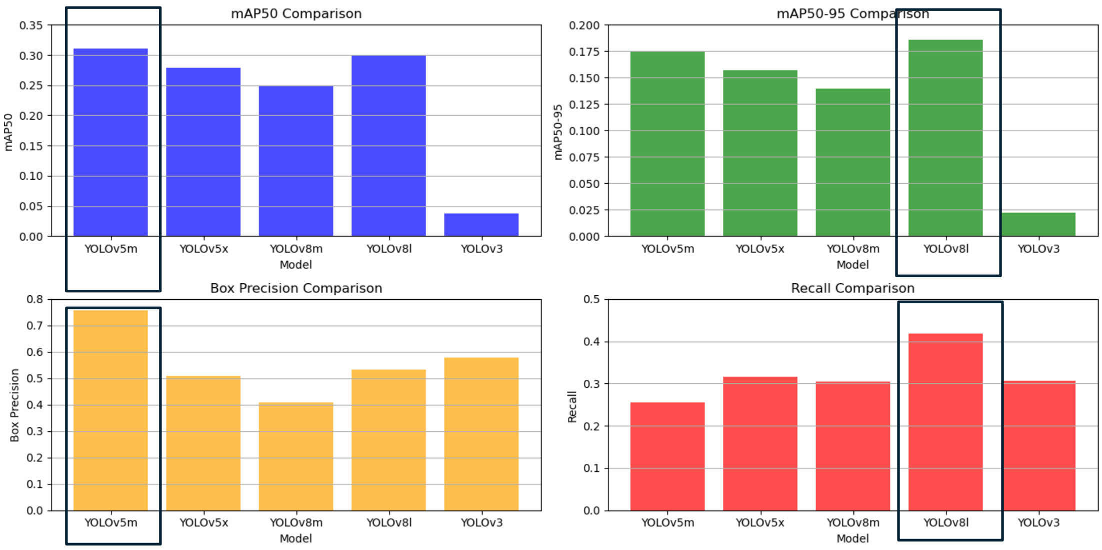


Figure 5.7: Comparative Analysis of all the YOLO models

The comparison of the YOLO models shows that they have different characteristics in terms of different performance indicators. Looking at the results, YOLOv5m is slightly ahead with 0. 31 mAP50, and 0. 756 Box Precision meaning that YOLOv5m is more accurate and has better localization than other models. YOLOv5x has a slightly lower mAP50 of 0. 279 and Box Precision of 0. 507 but has a better recall of 0. 316 which indicates that the model covers all the objects but might not be as precise as YOLOv5s. However, YOLOv8l provides the best results among all the YOLO architectures in mAP50 (0. 299), mAP50-95 (0. 186), and Box Precision (0. 533) and recall (0. 419). Even though YOLOv8m has lower mAP50 (0. 249) and Box Precision (0. 407), it performs better in recall than YOLOv5 models at 0. 305 but is still lower than YOLOv8l. While YOLOv3 exhibits considerably lower mAP50 =0. 037 and mAP50-95 =0. 022, this suggests a relatively low object detection rate. While YOLOv3 has a fairly high Box Precision of 0. 577 and recall of 0. 306, its detection performance is lower than that of other successors in terms of mAP.

5.8 Ball Tracking and Players performance Analysis

Figure 5.8: Ball Tracking

Accessibility of a player's first touch is vital for perusing and establishing the game principle and enhancing the technical abilities. The first contact often determines the subsequent actions, affecting the ability of a player to manipulate the ball during pressure, to gain possession and to make suitable decisions. In this respect, it is possible to pay attention to this aspect in order to improve the set of possibilities for the practice session and trainees' ball control and game transitions. However, when it comes to the further touch, analytics can also evaluate other aspects of the game like dribbling, passing accuracy, shooting angles and speed of the shot. All these elements are also important in the gameplay as explained below. These concepts profile the ability of the player to pass; how good he is at interpreting the game; shooting angles and the speeds show how efficient the player is at scoring and moving around the box. This way of analysis provides a broad picture to the coaches about the performance of a player and about what is best and what needs to be worked upon.

Figure 5.9: The redlines predict the trajectory of the ball in entire event.

These aforementioned evaluations are later important in the coaching process as they help in the formulation of training programs since they target areas that need change. Not only does such a precise analysis help improve one's own performance but also helps increase team work since the participants obtain greater insight into all players' roles and abilities. This makes it possible to have a continuous training development process and in so doing, can easily meet players' needs thus improving the entire team.

6: Conclusion and Future Work

This work performed a thorough experimental analysis of the several YOLO models, such as YOLOv5m, YOLOv5x, YOLOv8m, YOLOv8l, and YOLOv3 for the detection of football events as guided by the metrics of the IAUFD dataset. The models were evaluated with respect to certain aspect such as Mean Average Precision, IoU Threshold, Box Precision, and Recall. Among all the examples, YOLOv5m performed well, having a high mAP50 of 0. 31 and with the adequate performance of the strong Box Precision, nonetheless it had some difficulties with recall. Even though it had a higher level of overfitting, YOLOv8l provided the best compromise in terms of such parameters as mAP and its recall. To this end, let us consider the problem of object detection; while YOLOv3 being the oldest version performed quite well, several issues were noted in regard to accuracy, particularly when compared to its successors. The study also provided confusion matrix in detail allowing for identifying strengths and weaknesses in categorizing specific categories, such us red and yellow cards, which are rare events. Moreover, MediaPipe for ball tracking affords the analysis of the players' actions such as first touch, dribbling, accuracy of passing, and shooting by adding an external module.

Consequently, one can infer from the results of the present research that the model structure has to be chosen depending on the needs of solving a particular task. In achieving high accuracy and localization, such as YOLOv5m and YOLOv8l, there is still potential for improvement in recognising low frequency events and avoiding overfitting. Additional integration of MediaPipe to track balls and the players adds value by providing more insight into how players were performing to modify coaching techniques and training regimen. What is more, if further refinements of these models are achieved for the class imbalance problem or for the usage of more sophisticated monitoring approaches, these systems could become even more effective and accurate in their event capturing process. Also, it holds the prospect of enhancing an individual player's assessment and boosting team performance, this will be a big boost to football analysis.

In the future we would like to:

a. Detailed Deployment: In which we can deploy the best solution and can check in the rel time about the detection and analysis
b. Real Time App: This whole idea is a part of an extended idea of having a virtual coaching system for all the games. For example, Foot, Golf, Cricket etc. where the player can get help instantly

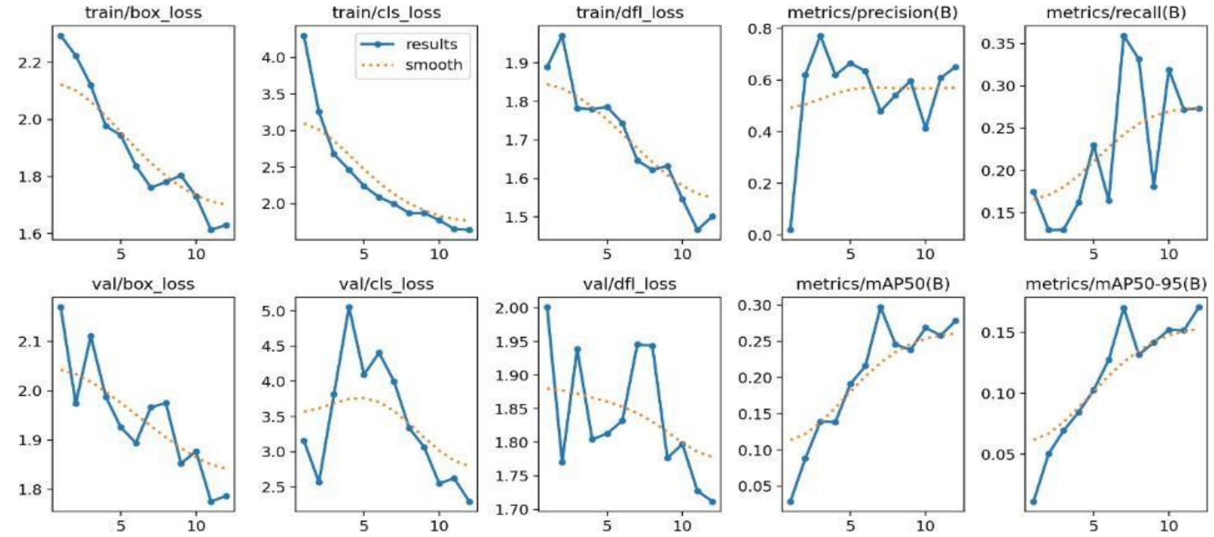
One biggest limitation while ding this project was dataset and computation. A perfect dataset is not available online that can be used for the analysis of our project and secondly the computation is also an important factor where certain YOLO models need higher computation. To handle the computation, we used the Google Colab Pro+ instance with A100 GPU.

References

Ait-Bennacer, F.E., Aaroud, A., Akodadi, K. and Cherradi, B., 2022. Applying Deep Learning and Computer Vision Techniques for an e-Sport and Smart Coaching System Using a Multiview Dataset: Case of Shotokan Karate. Int. J. Online Biomed. Eng., 18(12), pp.35-53.
Cui, J., Du, H. and Wu, X., 2023. Data analysis of physical recovery and injury prevention in sports teaching based on wearable devices. Preventive medicine, 173, p.107589
de Leeuw, A.W., van der Zwaard, S., van Baar, R. and Knobbe, A., 2022. Personalized machine learning approach to injury monitoring in elite volleyball players. European journal of sport science, 22(4), pp.511-520.
Eckard, M. E., Padua, D. A., Hearn, D. W., Pexa, B. S., & Frank, B. S. (2017). The relationship between athletic performance measures and injury risk in soccer. International journal of sports physical therapy, 12(2), 576.
Ghosh, I., Ramamurthy, S.R., Chakma, A. and Roy, N., 2022. Decoach: Deep learning-based coaching for badminton player assessment. Pervasive and Mobile Computing, 83, p.101608.
Gouttebarge, V., Hopley, P., Kerkhoffs, G., Verhagen, E., Viljoen, W., Finch, C., ... & Schwellnus, M. (2017). A 12-month prospective cohort study of symptoms of common mental disorders among professional rugby players. European journal of sport science, 17(4), 479-487.
Gustafsson, E.F., 2024. Exploring Computer Vision-Based AI-Assisted Coaching for Youth Football Players.
Impellizzeri, F. M., Marcora, S. M., & Coutts, A. J. (2019). Internal and external training load: 15 years on. International journal of sports physiology and performance, 14(2), 270-273.
Malone, S., Roe, M., Doran, D. A., Gabbett, T. J., & Collins, K. D. (2017). High chronic training loads and exposure to bouts of maximal velocity running reduce injury risk in elite Gaelic football. The Journal of science and medicine in sport, 20(3), 250-254.
Mekruksavanich, S. and Jitpattanakul, A., 2022. Multimodal wearable sensing for sport-related activity recognition using deep learning networks. Journal of Advances in Information Technology, 13.
Meylan, C., Trewin, J., & McKean, K. (2017). Quantifying explosive actions in international women's soccer. International journal of sports physiology and performance, 12(3), 310-315.
Miah, J., Mamun, M., Rahman, M.M., Mahmud, M.I., Ahmad, S. and Nasir, M.H.B., 2022, December. Mhfit: Mobile health data for predicting athletics fitness using machine learning models. In 2022 2nd International seminar on machine learning, Optimization, and Data Science (ISMODE) (pp. 584-589). IEEE.
Moustakidis, S., Plakias, S., Kokkotis, C., Tsatalas, T. and Tsaopoulos, D., 2023. Predicting football team performance with explainable ai: Leveraging shap to identify key team-level performance metrics. Future Internet, 15(5), p.174.
Ramsay, J. O., Hooker, G., & Graves, S. (2020). Functional data analysis with R and MATLAB. Springer.
Raya-González, J., Martín-González, J. M., Los Arcos, A., & Nakamura, F. Y. (2021). The application of artificial intelligence in football: A systematic review. International Journal of Environmental Research and Public Health, 18(4), 1577.
Sarmento, H., Marcelino, R., Anguera, M. T., CampaniÇo, J., Matos, N., & LeitÃo, J. C. (2014). Match analysis in football: a systematic review. Journal of sports sciences, 32(20), 1831-1843. Stein, M., Janetzko, H., Seebacher, D., Jäger, A., Nagel, M., Hölsch, J., ... & Keim, D. A. (2019). How to make sense of team sport data: From acquisition to data modeling and research aspects. Data, 4(2), 28.
Wang, Z., Veličković, P., Hennes, D., Tomašev, N., Prince, L., Kaisers, M., Bachrach, Y., Elie, R., Wenliang, L.K., Piccinini, F. and Spearman, W., 2024. TacticAI: an AI assistant for football tactics. Nature communications, 15(1), p.1906.
Windt, J., & Gabbett, T. J. (2017). How do training and competition workloads relate to injury? the workload—injury aetiology model. British journal of sports medicine, 51(5), 428-435.
Yang, J., 2022. A deep learning and clustering extraction mechanism for recognizing the actions of athletes in sports. Computational Intelligence and Neuroscience, 2022.

Appendix

1. Yolov5m:



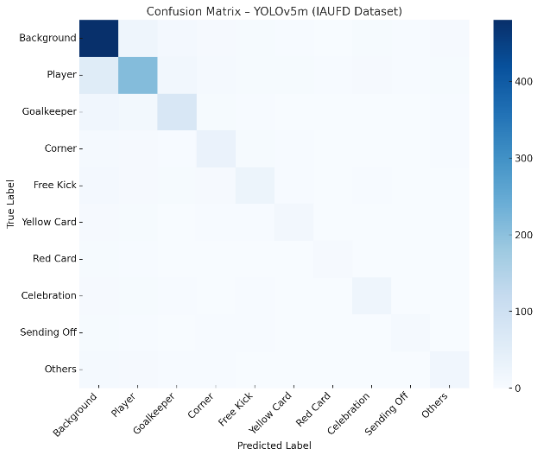
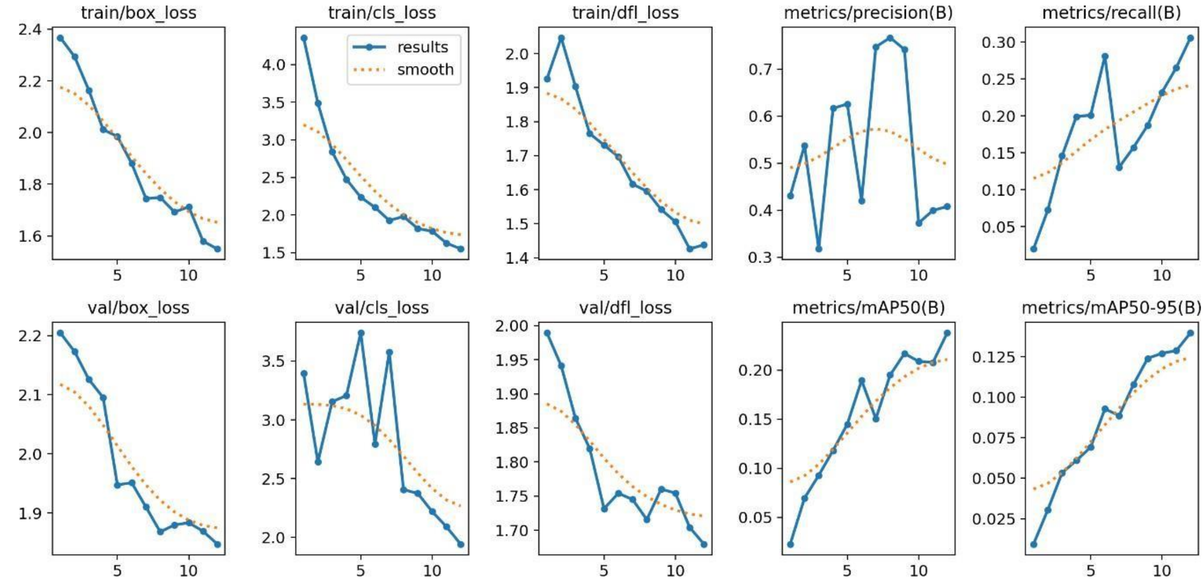


Figure 1: Loss and Training curves of YOLOv5m model

2. Yolov5x:



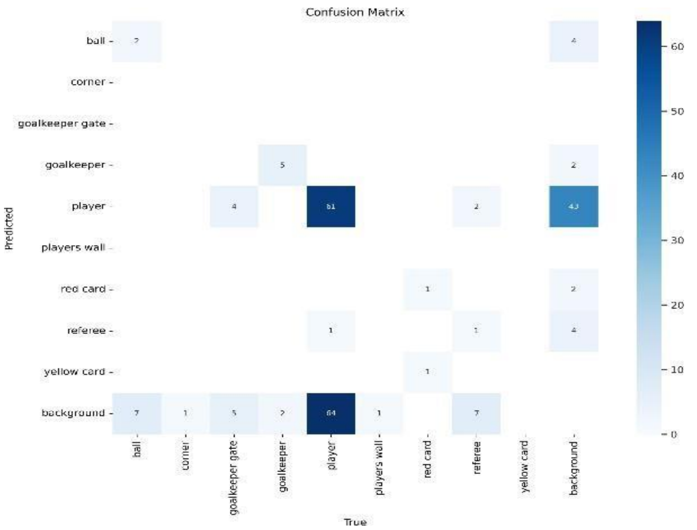
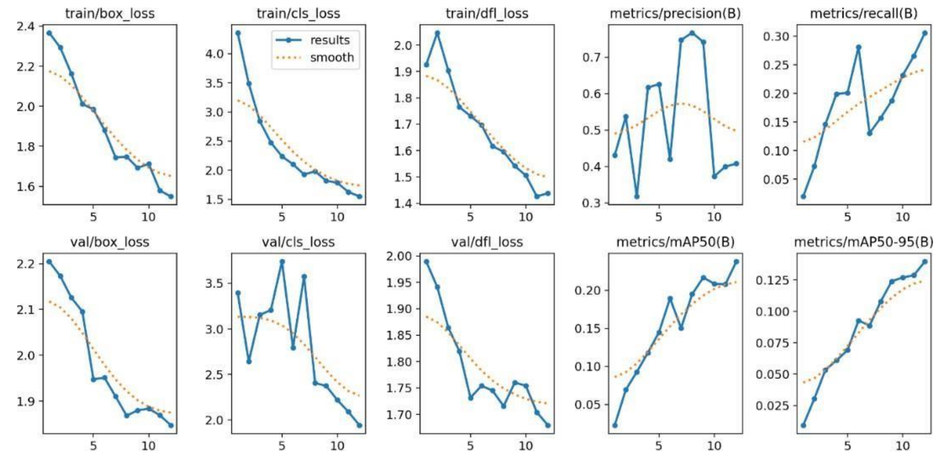


Figure 2: Loss and Training curves of YOLOv5 model

3. Yolov8m:



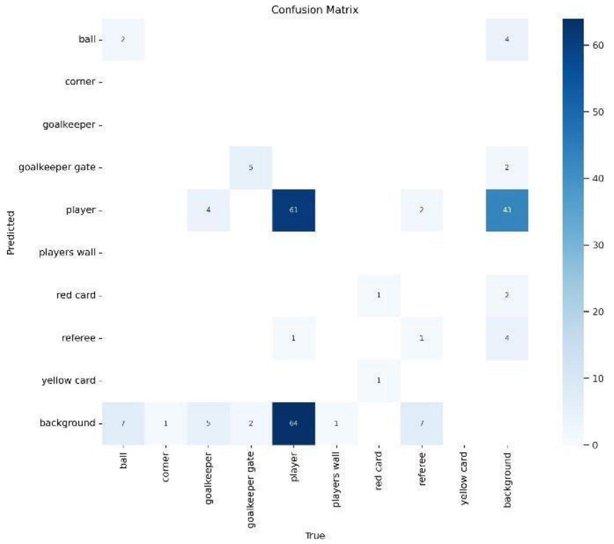


Figure 3: Loss and Training curves of YOLOv8m model

4. Yolov8l:

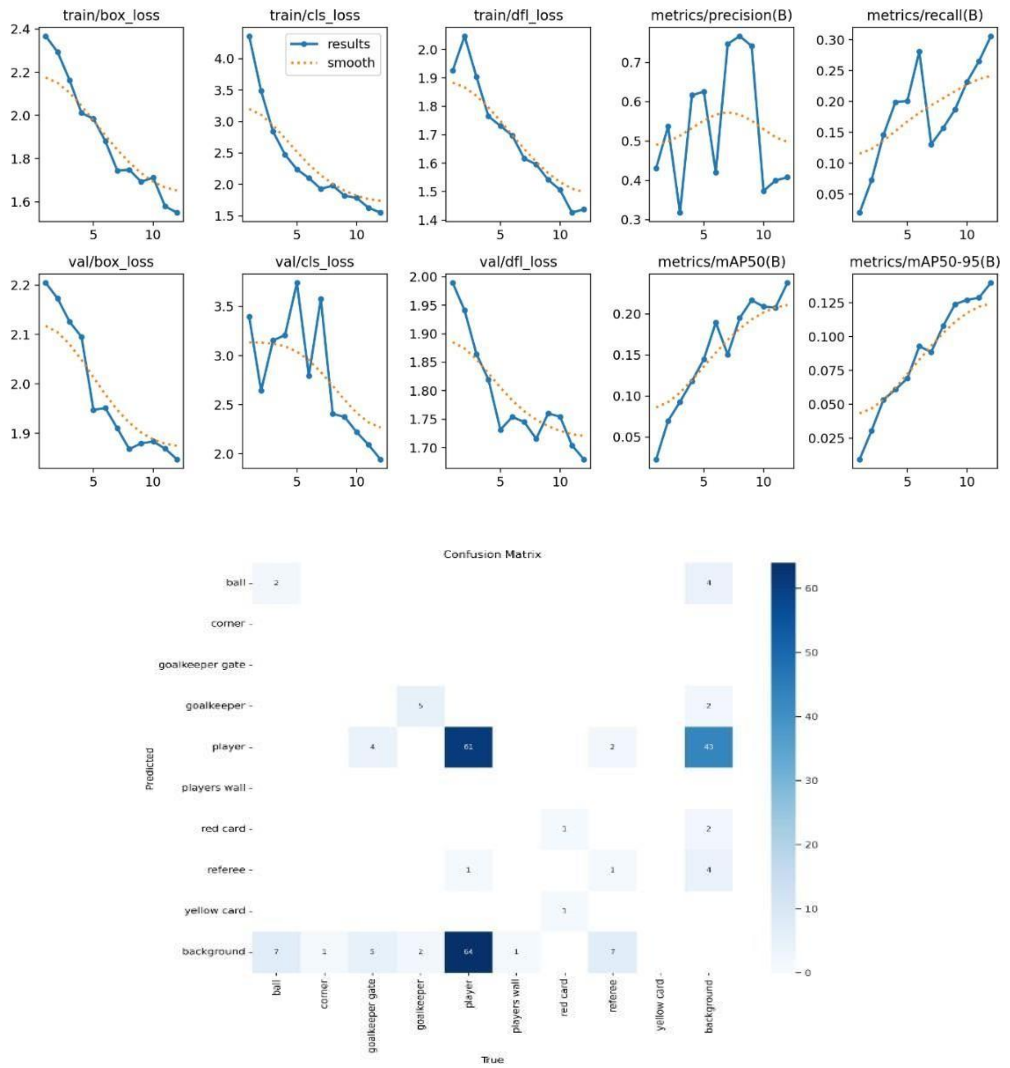


Figure 4: Loss and Training curves of YOLOv8L model

5. Yolov3:

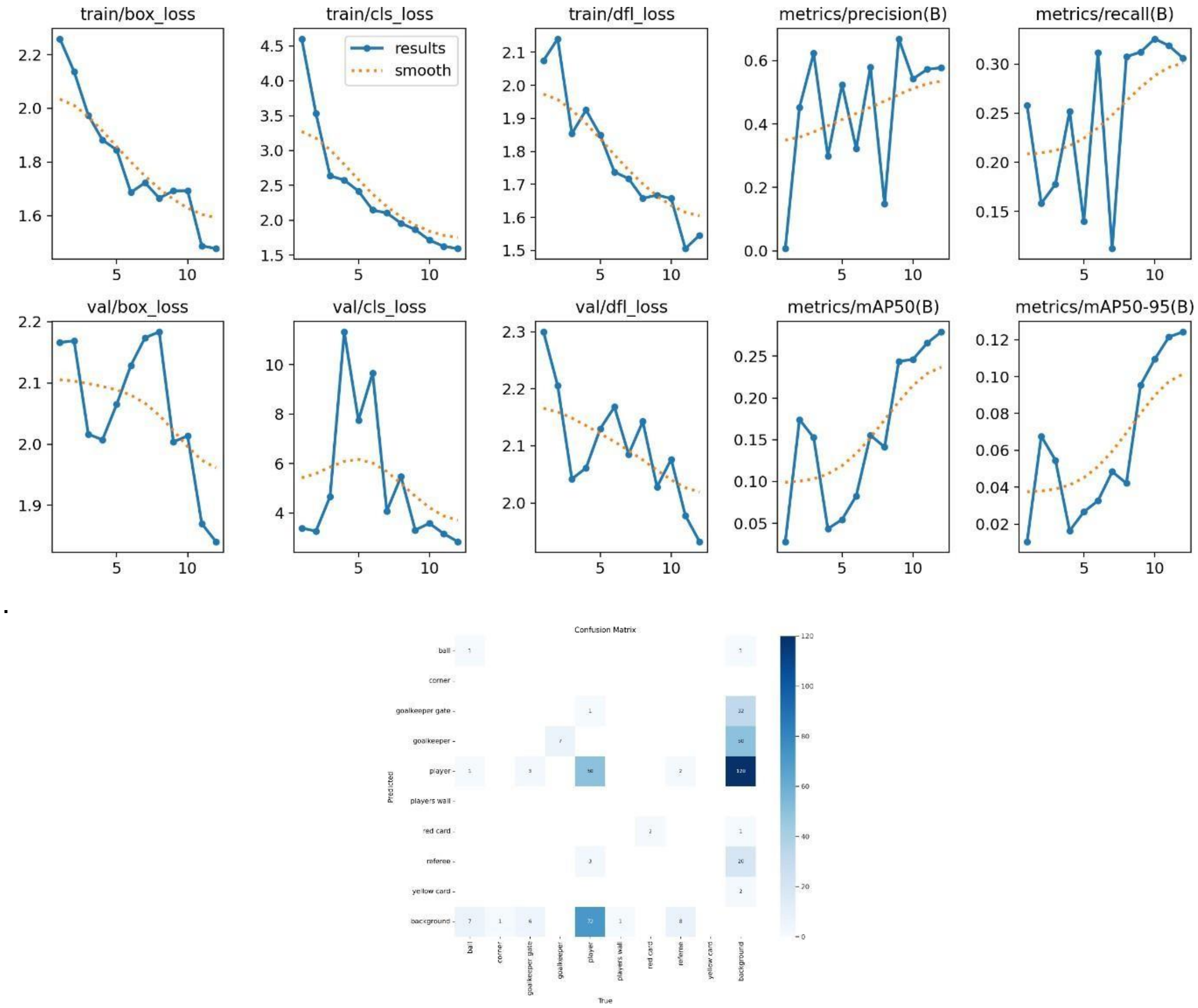
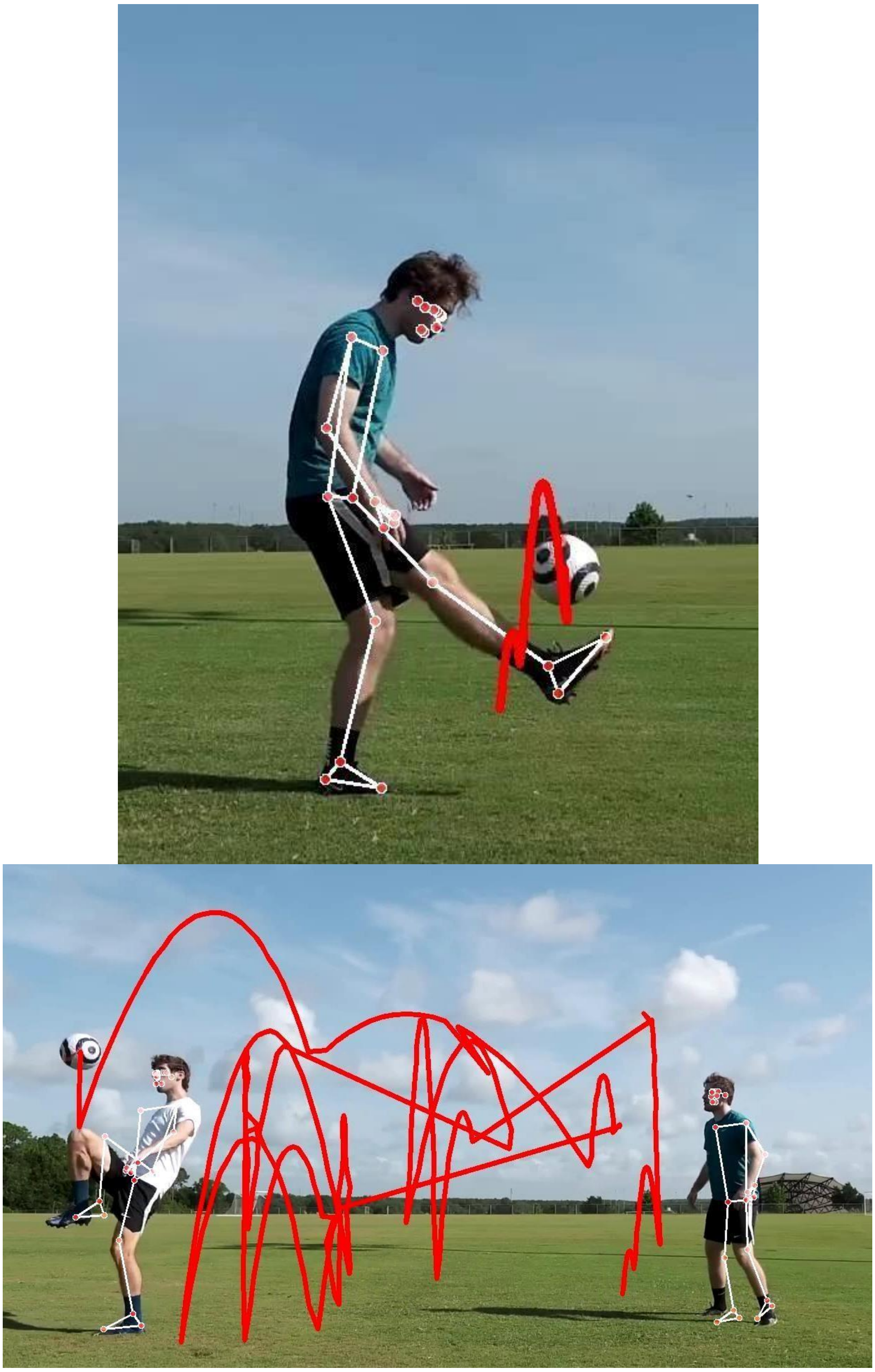


Figure 5: Loss and Training curves of YOLOv3 model

Table 1: Summary of the findings

Model	mAP50	mAP50-95	Box Precision	R
Yolov5m	0.31	0.175	0.756	0.256
Yolov5x	0.279	0.157	0.507	0.316
Yolov8m	0.249	0.14	0.407	0.305
Yolov8l	0.299	0.186	0.533	0.419
Yolov3	0.037	0.022	0.577	0.306

6. Pose Estimation:



Figture 6: The redlines predict the trajectory of the ball in entire event