Review

A Comprehensive Review of Computer Vision in Sports: Open Issues, Future Trends and Research Directions

Banoth Thulasya Naik [1] , Mohammad Farukh Hashmi [1] and Neeraj Dhanraj Bokde [2,*]

1 Department of Electronics and Communication Engineering, National Institute of Technology, Warangal 506004, India; thulasyramsingh@student.nitw.ac.in (B.T.N.); mdfarukh@nitw.ac.in (M.F.H.)
2 Department of Civil and Architectural Engineering, Aarhus University, 8000 Aarhus, Denmark
* Correspondence: neerajdhanraj@cae.au.dk

Abstract: Recent developments in video analysis of sports and computer vision techniques have achieved significant improvements to enable a variety of critical operations. To provide enhanced information, such as detailed complex analysis in sports such as soccer, basketball, cricket, and badminton, studies have focused mainly on computer vision techniques employed to carry out different tasks. This paper presents a comprehensive review of sports video analysis for various applications: high-level analysis such as detection and classification of players, tracking players or balls in sports and predicting the trajectories of players or balls, recognizing the team's strategies, and classifying various events in sports. The paper further discusses published works in a variety of application-specific tasks related to sports and the present researcher's views regarding them. Since there is a wide research scope in sports for deploying computer vision techniques in various sports, some of the publicly available datasets related to a particular sport have been discussed. This paper reviews detailed discussion on some of the artificial intelligence (AI) applications, GPU-based work-stations and embedded platforms in sports vision. Finally, this review identifies the research directions, probable challenges, and future trends in the area of visual recognition in sports.

Keywords: sports; ball detection; player tracking; artificial intelligence; computer vision; embedded platforms

1. Introduction

Automatic analysis of video in sports is a possible solution to the demands of fans and professionals for various kinds of information. Analyzing videos in sports has provided a wide range of applications, which include player positions, extraction of the ball's trajectory, content extraction, and indexing, summarization, detection of highlights, on-demand 3D reconstruction, animations, generation of virtual view, editorial content creation, virtual content insertion, visualization and enhancement of content, gameplay analysis and evaluations, identifying player's actions, referee decisions and other fundamental elements required for the analysis of a game.

The task of player detection (identification) and tracking is very difficult because of many challenges, which include the similar appearance of subjects, complex occlusions, an unconstrained field environment, background, unpredictable movements, unstable camera motion, issues with calibration of low textured fields, and the editing performed for broadcasting video, lower pixel resolution of players who are distant and smaller in the frame, and motion blur, among others. The simultaneous detection of players and ball and tracking them at once is quite challenging, because of the zigzag movements of the ball and player, change of the ball from player to player, severe occlusion between players and the ball. Hence, this paper presents a survey of detection, classification, tracking, trajectory prediction and recognizing the team's strategies in various sports. Detection and tracking of players is the only major requirement in some sports such as cycling and swimming. Hence,

this study presents a survey of detection, classification, tracking, trajectory prediction and recognizing the team's strategies in various sports. Detection and tracking of the player is the only major requirement in some sports such as cycling, swimming, among others. As a result, as illustrated in Figure 1, this research classifies all sports into two categories: player-centered and ball-centered sports, with extensive analysis in Section 4.
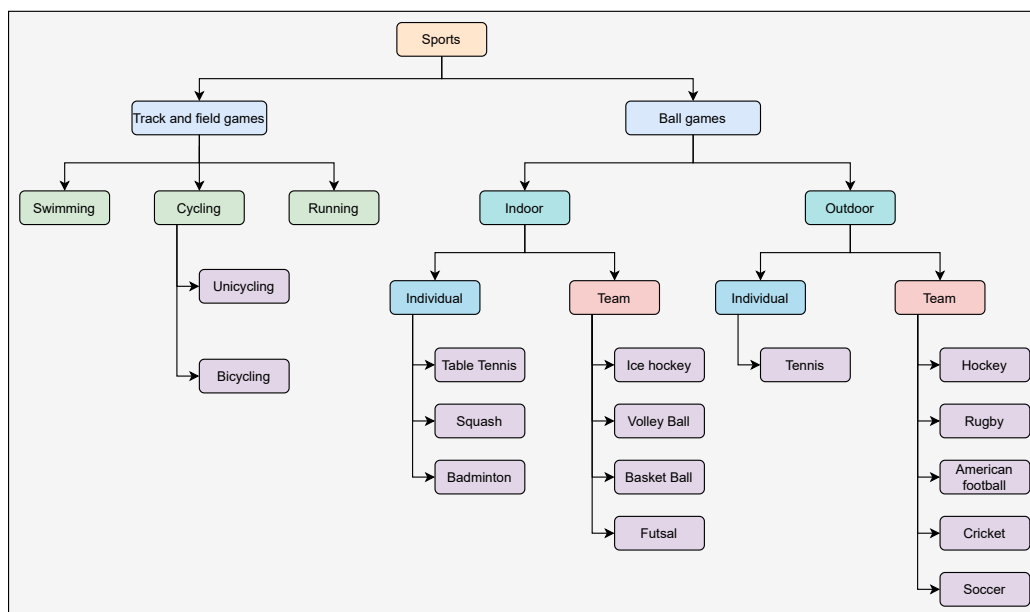

Figure 1. Classification of different types of sports.

Recent developments in video analysis of sports have a focus on the features of computer vision techniques, which are used to perform certain operations for which these are assigned, such as detailed complex analysis such as detection and classification of each player based on their team in every frame or by recognizing the jersey number to classify players based on their team will help to classify various events where the player is involved. In higher-level analysis, such as tracking the player or ball, many more such evaluations are to be considered for the evaluation of a player's skills, detecting the team's strategies, events and the formation of tactical positions such as midfield analysis in various sports such as soccer, basketball, and also various sports vision applications such as smart assistants, virtual umpires, assistance coaches, have been discussed in Section 7. A higher-level semantic interpretation is an effective substitute, especially in situations when reduced human intervention and real-time analysis are desired for the exploitation of the delivered system outputs.

The main task of video summarization or highlight extraction is extracting key events of the game which provides users with an ability to view highlights as per their interests. For this purpose, it is necessary to detect and classify gestures, recognize the actions of the referee/umpire, track players and the ball in key events like the time of goal scoring to analyze and classify different types of shots performed by players. The framework for processing and analyzing task-specific events in sports applications, such as playfield extraction, detection, and tracking of the player/ball is shown in Figure 2, and detailed analyses of playfield extraction are discussed in Section 3.

A detailed review of research in the above-mentioned domains is presented in this article and the data were compiled from papers that focus on computer vision-based approaches that were used for each application, followed by inspecting key points and weaknesses, thereby investigating whether these methodologies in their current state of implementation can be utilized in real-time sports video analysis systems.
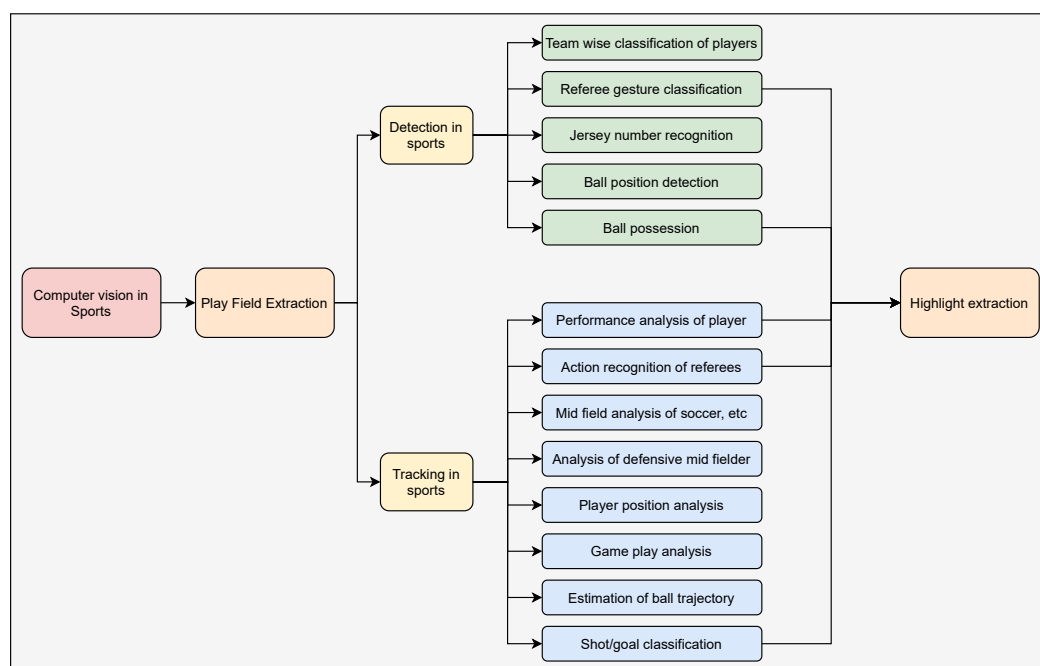
Figure 2. Framework of processing and analysis of different applications in sports video.

Features of the Proposed Review

Some of the surveys and reviews published in different sports video processing and their main contributions, i.e., whether the paper discussed (marked as '✔' if discussed in Table 1) hand-crafted and machine learning algorithms, the type of sport, different tasks in sports, whether it provided datasets or not and finally the aim of the review are discussed and summarized in Table 1 and listed below.

- Tan et al. [1] researched badminton movement analysis such as Badminton smashing, badminton service recognition, badminton swing, and shuttle trajectory analysis.
- Bonidia et al. [2] presented a systematic review of sports data mining, which discusses the current research body, themes, the dataset used, algorithms, and research opportunities.
- Rahmad et al. [3] presented a survey on video-based sports intelligence systems used to recognize sports actions. They provided video-based action recognition frameworks used in the sports field and also discussed deep learning implementation in video-based sports action recognition. They proposed a flexible method that classifies actions in different sports with different contexts and features as part of future research.
- Eline and Marco [4] presented a summary of 17 human motion capture systems which reports calibration specs as well as the specs provided by the manufacturer. This review helps researchers to select a suitable motion capture system for experimental setups in various sports.
- Ebadi et al. [5] presented a survey on state-of-the-art (SOTA) algorithms for player tracking in soccer videos. They analyzed the strengths and weaknesses of different approaches and presented the evaluation criteria for future research.
- Thomas [6] proposed an analysis of computer vision-based applications and research topics in the sports field. The study summarized some of the commercially available systems such as camera tracking and player tracking systems. They also incorporated some of the available datasets of different sports.
- Cust et al. [7] presented a systematic review on machine learning and deep learning for sports-specific movement recognition using inertial measurement units and computer vision data.

- Kamble et al. [8] presented an exhaustive survey on ball tracking categorically and reviewed several used techniques, their performance, advantages, limitations, and their suitability for a different sports.
- Shih [9] focused on the content analysis fundamentals (e.g., sports genre classification, the overall status of sports video analytics). Additionally reviewed are SOTA studies with prominent challenges observed in the literature.
- Beal et al. [10] explored AI techniques that have been applied to challenges within team sports such as match outcome prediction, tactical decision making, player investments, and injury prediction.
- Apostolidis et al. [11] suggested a taxonomy of the existing algorithms and presented a systematic review of the relevant literature that shows the evolution of deep learning-based video summarization technologies.
- Yewande et al. [12] explored a review to better understand the usage of wearable technology in sports to improve performance and avoid injury.
- Rana et al. [13] offered a thorough overview of the literature on the use of wearable inertial sensors for performance measurement in various sports.

Table 1. Summary of previous surveys and reviews in different sports.

Articles	Hand Crafted Algorithms	Machine Learning Algorithms	Sport and Application				Discussion about Dataset	Aim of Review
			Sport	Detection	Tracking	Classification and Movement Recognition		
[1]	✗		Badminton	✗	✗	✔	✗	Motion analysis
[2]	✗	✔	-	✔	✗	✔	✗	Sport data mining
[3]	✗	✔	-	✔	✗	✔	✔	-
[7]	✗	✔	-	✗	✗	✔	✗	-
[4]	✗	✗	-	✗	✗	✔	✗	Motion Capture
[8]	✔	✔	Soccer	✔	✔	✗	✗	Ball Tracking
[5]	✗	✔	Soccer	✔	✔		✗	Player detection/tracking
[6]	✗	✔	-	✔	✔	✗	✔	Availability of datasets for sports
[9]	✗	✗	-		✗		✗	Content-Aware Analysis
[10]	✗	✔	-	✗	✗	✔	✗	-
[11]	✗	✔	-	✗	✗	✗	✔	Video Summarization
[12]	✔	✗	-	-	-	-	✗	Wearable technology in sports
[13]	✔	✗	-	-	✔	✔	✗	Wearable technology in sports

The proposed survey mainly focuses on providing a proper and comprehensive survey of research carried out in computer vision-based sports video analysis for various applications such as detection and classification of players, tracking players or balls and predicting the trajectories of players or balls, recognizing the team's strategies, classifying various events on the sports field, etc. and in particular, establishing a pathway for next-generation research in the sports domain. The features of this review are:

- In contrast to recently published review papers in the sports field, this article comprehensively reviews statistics of studies in various sports and various AI algorithms that have been used to cover various aspects observed and verified in sports.
- It provides a roadmap of various AI algorithms' selection and evaluation criteria and also provide some of the publicly available datasets of different sports.
- It discusses various GPU-based embedded platforms for real-time object detection and tracking framework to improve the performance and accuracy of edge devices.
- Moreover, it demonstrates various applications in sports vision and possible research directions.

The rest of this paper is organized as follows. Section 2 provides statistical details of research in sports. Section 3 presents extraction data vis-a-vis various sports playfields, followed by a broader dimension that covers a wide range of sports and is reviewed in Section 4. Some of the available datasets for various sports along and embedded platforms have been reviewed in Sections 5 and 6. Section 7 provides various application-specific tasks in the field of sports vision. Section 8 covers potential research directions, as well as different challenges to be overcome in sports studies. Last but not least, Section 9 concludes by describing the final considerations.

2. Statistics of Studies in Sports

Detection of the positions of the players at any given point of time is the basic step for tracking a player, which is also needed for graphics systems in sports for analysis and obtaining pictures of key moments in a game. Equipment and methods used in commercial systems for broadcast analysis vary from those depending on a manual operator clicking on the players' feet with a calibrated camera image to an automated technique involving segmentation and identification of areas which likely correspond to players. For the performance improvement of sport teams in soccer, volleyball, hockey, badminton among others, analyzing the movements of players individually, and the real-time team formation, may provide a crucial real-time insight for the team coach.

The research articles discussed in this review were obtained from various reputed publishers such as IEEE, Elsevier, MDPI, Springer among many others, and top-tier computer vision conferences such as CVPR, ICCV, and ECCV, ranging from high impact factor online sources in the domain of player/ball/referee detection and tracking in sports, classification of objects in sports, behavior, and performance analysis of players, gesture recognition of referees/umpires, automatic highlight detection, score updating among others. Figure 3 provides overall information of sports research publications in the past five years considered in this comprehensive survey article.

Figure 4 provides the statistics of studies of various sports in various applications such as detecting/tracking the player and ball, trajectory prediction, classification, and video summarization. which are published in various standard journals as presented in Figure 3.
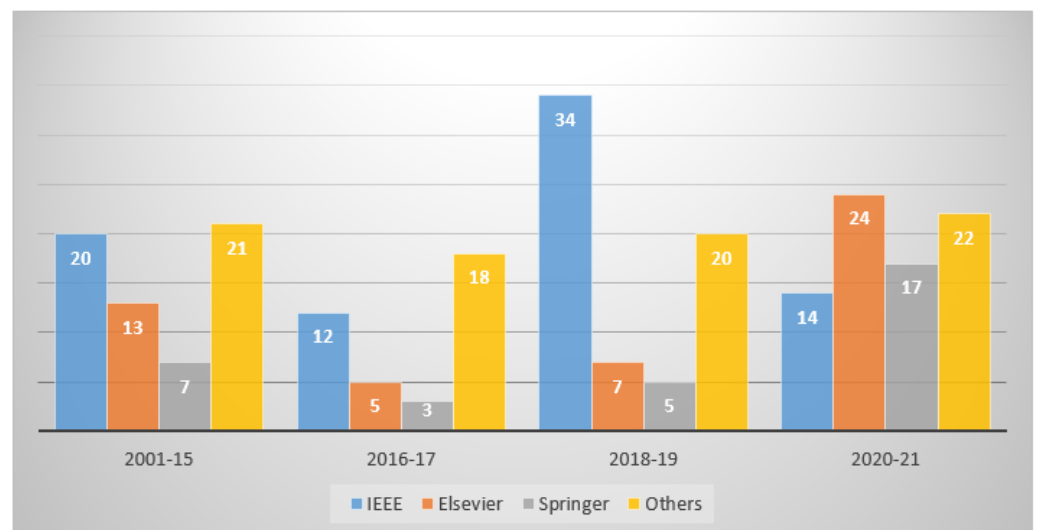

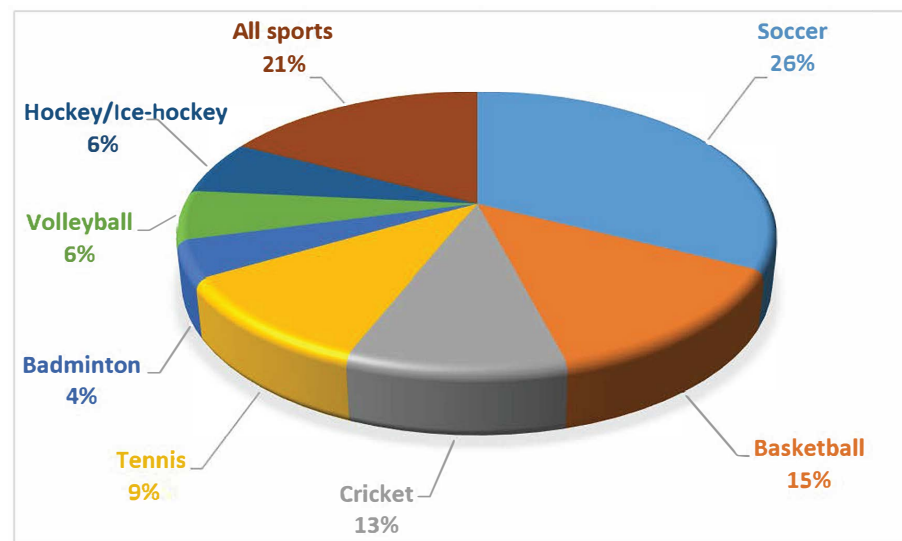**Figure 3.** Sports research progress in past five years.

Figure 4. Sports wise research progress.

3. Play Field Extraction in Various Sports

Detection of the sports field plays an important role in sports video analysis. Detection of the playfield region has two objectives. One is to detect the playfield region from non-playfield areas, while the other is to identify primary objects from the background by filtering out redundant pixels such as grass and court lines. This provides a reduced pixel which requires processing and reduction in errors to simplify player or ball detection and tracking phases, event extraction, pose detection, etc. The challenges here include distinguishing the color of the playfield from that of the stadium, lighting conditions and sometimes weather, viewing angles, and the shadows. Therefore, accurate segmentation of the playfield cannot be achieved just by processing the color of the playfield under certain situations and making it constant without updating the statistics throughout the game. There is also an added noise when the player's clothes matching that of the ground, and there appear shadows at the base of a player from different sources of light. A Gaussian-based background subtraction technique [14], which is implemented using computer vision methods, generates the foreground mask as shown in Figure 5.
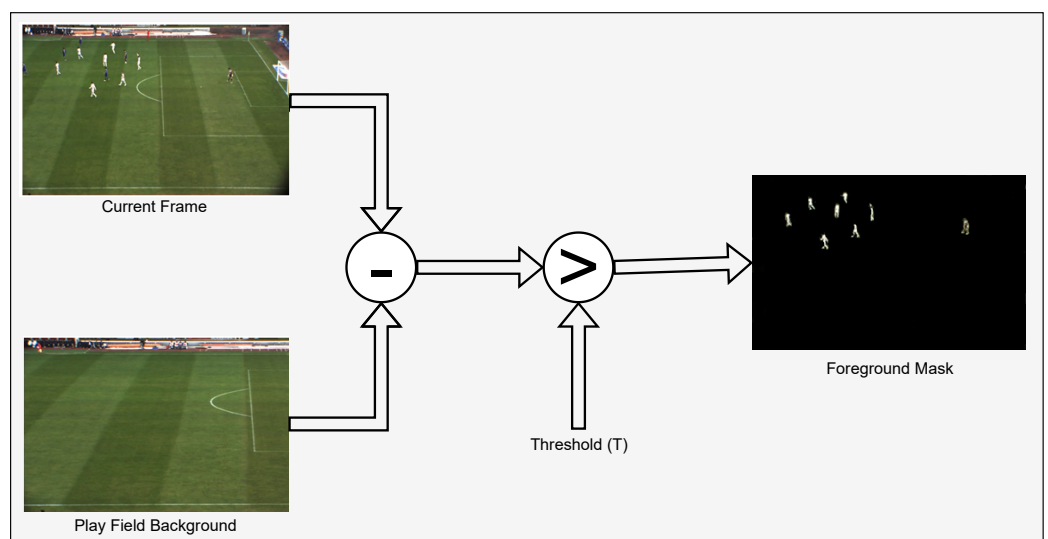

Figure 5. Background subtraction model.

Researchers have used a single dominant color for detecting the playfield. Accordingly, some studies have utilized the features of images in which illumination is not affected by transforming the images from RGB space to HIS [15–17], YCbCr [18], normalized RGB [19–21].

For a precise capture of the movements of the players, tracking the ball and actions of referees on the field or court, it is necessary to calibrate the camera [5,8] and also to use an appropriate number of cameras to cover the field. Though some algorithms are capable of tracking the players, some other objects also need to be tracked in dynamically complex situations of interest for detailed analysis of the events and extraction of the data of the subject of interest. Reference [22] presented an approach to extract the playing field and track the players and ball using multiple cameras in soccer video. In [23,24], an architecture was presented, which uses single (Figure 6a) and multi-cameras (Figure 6c) to capture a clear view of players and ball in various challenging and tricky situations such as severe occlusions and the ball being missing from the frames. To estimate the players' trajectory and team classification in [25,26] a bird's eye view of the field is presented to capture players, precisely as shown in Figure 6b. Various positions of the camera for capturing the entire field are presented in [27,28] to detect and track the players/ball and estimate the position of the players.
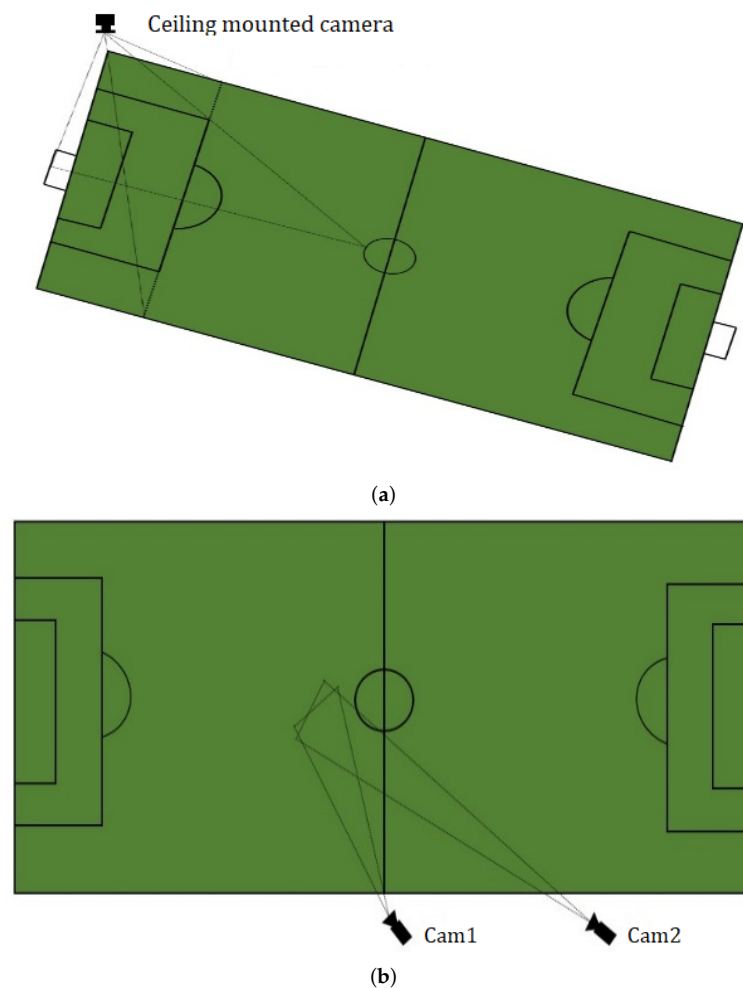

(a)

(b)

Figure 6. *Cont.*

(**c**)

Figure 6. Camera placements in the playfield. (**a**) Ceiling-mounted camera [23]. (**b**) Birds eye view of the field [25,26]. (**c**) Multiple cameras placed to cover the complete playfield [27].

Morphological operations-based techniques can separate the playfield and non-playfield regions, but they cannot detect the lines in the playfield. The background subtraction-based techniques generate foreground regions by subtracting the background frame from the current frame (i.e., by detecting moving objects in the frame); however, they fail to detect the playfield lines as shown in Figure 5. So, the best way to detect the playfield lines is by labeling the data as playfield lines (as shown in Figure 7a), advertisements (as shown in Figure 7b), and the non-playfield regions as shown in Figure 7c.

Training the model using a dataset that is labeled as playfield lines, advertisements, and the non-playfield region as shown in Figure 7 can detect and classify the playfield lines, advertisements, and the non-playfield region, which reduces the detection of false positives and false negatives.

Figure 7. Background-labeled samples from the dataset (**a**) playfield lines, (**b**) advertisements, (**c**) non-playfield region.

4. Literature Review

In this section, the overview of traditional computer-vision methods implemented for major application specifics in sports (such as detection, event classification/recognition, tracking and trajectory prediction) investigated by the researchers and their significant limitations is discussed.

4.1. Basketball

Basketball is a sport played between two teams consisting of five players each. The task of this sport is to score more points than the opponent. This sport has several activities with the ball such as passing, throwing, bouncing, batting, or rolling the ball from one player to another. Physical contact with an opponent player may be a foul if the contact impedes the players' desired movement. The advancements in computer vision techniques have effectively employed fully automated systems to replace the manual analysis of basketball sports. Recognizing the player's action and classifying the events [29–31] in basketball videos helps to analyze the player's performance. Player/ball detection and tracking in basketball videos are carried out in [32–37] but fail in assigning specific identification to avoid identity switching among the players when they cross. By estimating the pose of the player, the trajectory of the ball [38,39] is estimated from various distances to the basket. By recognizing and classifying the referee's signals [40], player behavior can be assessed and highlights of the game can be extracted [41]. The behavior of a basketball team [42] can be characterized by the dynamics of space creation presented in [43–48] that works to counteract space creation dynamics with a defensive play presented in [49]. By detecting the specific location of the player and ball in the basketball court, the player movement can be predicted [50] and the ball trajectory [51–53] can be generated in three dimensions which is a complicated task. It is also necessary to study the extraction of basketball players' shooting motion trajectory, combined with the image feature analysis method of basketball shooting, to reconstruct and quantitatively track the basketball players' shooting motion trajectory [54–57]. However, it is difficult to analyze the game data for each play such as the ball tracking or motion of the players in the game, because the situation of the game changes rapidly, and the structure of the data is complicated. Therefore, it is necessary to analyze the real-time gameplay [58]. Table 2 summarizes various proposed methodologies used to complete various challenging tasks in basketball sport including their limitations.

Table 2. Studies in basketball.

Studies in Basketball				
Ref.	Problem Statement	Proposed Methodology	Precision and Performance Characteristics	Limitations and Remarks
[31]	Recognizing actions of basketball players by using image recognition techniques	Bi-LSTM Sequence2Sequence	The metrics used to evaluate the method are the Spearman rank-order correlation coefficient, Kendall rank-order correlation coefficient, Pearson linear correlation coefficient, and Root Mean Squared Error and achieved 0.921, 0.803, 0.932, and 1.03, respectively.	The methodology failed to recognize difficult actions due to which accuracy is reduced. The accuracy of action recognition can be improved with a deep convolutional neural network.
[54]	Multi-future trajectory prediction in basketball.	Conditional Variational Recurrent Neural Networks (RNN)—TrajNet++	The proposed methodology was tested on Average Displacement Error and Final Displacement Error metrics. The methodology is robust if the number it achieves is smaller than 7.01 and 10.61.	The proposed methodology fails to predict the trajectories in the case of uncertain and complex scenarios. As the behavior of the basketball or players is dynamic, belief maps cannot steer future positions. Training the model with a dataset of different events can rectify the failures of predictions.

Table 2. *Cont.*

		Studies in Basketball		
Ref.	Problem Statement	Proposed Methodology	Precision and Performance Characteristics	Limitations and Remarks
[58]	Predicting line-up performance of basketball players by analyzing the situation of the field.	RNN + NN	At the point guard (pg) position 4 candidates were detected and at the center (c) position 3 candidates were detected. The total score of pg candidates is 13.67, 12.96, 13.42, 10.39, and where the total score of c candidates is 10.21, 14.08, and 13.48, respectively.	-
[32]	Multiplayer tracking in basketball videos	YOLOv3 + Deep-SORT, Faster-RCNN + Deep-SORT, YOLOv3 + DeepMOT, Faster-RCNN + DeepMOT, JDE	Faster-RCNN provides better accuracy than YOLOv3 among baseline detectors. The joint Detection and Embedding method performs better in the accuracy of tracking and computing speed among multi-object tracking methods.	Tracking in specific areas such as severe occlusions and improving detection precision improves the accuracy and computation speed. By adopting frame extraction methods, in terms of speed and accuracy, it can achieve comprehensive performance, which may be an alternative solution.
[40]	Recognizing the referee signals from real-time videos in a basketball game.	HOG + SVM, LBP + SVM	Achieved an accuracy of 95.6% for referee signal recognition using local binary pattern features and SVM classification.	In the case of a noisy environment, a significant chance of occlusion, an unusual viewing angle, and/or variability of gestures, the performance of the proposed method is not consistent. Detecting jersey color and eliminating all other detected elements in the frame can be the other solution to improve the accuracy of referee signal recognition.
[30]	Event recognition in basketball videos	CNN	mAP for group activity recognition is 72.1%	The proposed model can recognize the global movement in the video. By recognizing the local movements, the accuracy can be improved.
[59]	Analyzing the behavior of the player.	CNN + RNN	Achieved an accuracy of 76.5% for four types of actions in basketball videos.	The proposed model gives less accuracy for actions such as passing and fouling. This also gives less accuracy of recognition and prediction on the test dataset compared to the validation dataset.
[33]	Tracking ball movements and classification of players in a basketball game	YOLO + Joy2019	Jersey number recognition in terms of Precision achieved is 74.3%. Player recognition in terms of Recall achieved 89.8%.	YOLO confuses the overlapped image for a single player. In the subsequent frame, the tracking ID of the overlapped player is exchanged, which causes wrong player information to be associated with the identified box.
[29]	Event classifications in basketball videos	CNN + LSTM	The average accuracy using a two-stage event classification scheme achieved 60.96%.	Performance can be improved by introducing information such as individual player pose detection and player location detection
[49]	Classification of different defensive strategies of basketball payers, particularly when they deviate from their initial defensive action.	KNN, Decision Trees, and SVM	Achieved 69% classification accuracy for automatic defensive strategy identification.	Considered only two defensive strategies 'switch' and 'trap' involved in Basketball. In addition, the alternative method of labeling large Spatio-temporal datasets will also lead to better results. Future research may also consider other defensive strategies such as pick-and-roll and pick-and-pop.
[38]	Basketball trajectory prediction based on real data and generating new trajectory samples.	BLSTM + MDN	The proposed method performed well in terms of convergence rate and final AUC (91%) and proved deep learning models perform better than conventional models (e.g., GLM, GBM).	To improve the accuracy time series the prediction has to consider. By considering factors such as player cooperation and defense when predicting NBA player positions, the performance of the model can be improved.

Table 2. *Cont.*

	Studies in Basketball			
Ref.	Problem Statement	Proposed Methodology	Precision and Performance Characteristics	Limitations and Remarks
[51]	Generating basketball trajectories.	GRU-CNN	Validated on a hierarchical policy network (HPN) with ground truth and 3 baselines.	The proposed model failed in the trajectory of a three-dimensional basketball match.
[41]	Score detection, highlights video generation in basketball videos.	BEI+CNN	Automatically analyses the basketball match, detects scoring, and generates highlights. Achieved an accuracy, precision, recall, and F1-score of 94.59%, 96.55%, 92.31%, and 94.38%.	The proposed method is lacking in computation speed which achieved 5 frames per second. Therefore, it cannot be implemented in a real-time basketball match.
[34]	Multi-person event recognition in basketball videos.	BLSTM	Event classification and event detection were achieved in terms of mean average precision, i.e., 51.6% and 43.5%.	A high-resolution dataset can improve the performance of the model.
[44]	Player behavior analysis.	RNN	Achieved an accuracy of 80% over offensive strategies.	The methodology fails in many factors such as complexity of interaction, distinctiveness, and diversity of the target classes and other extrinsic factors such as reactions to defense, unexpected events such as fouls, and consistency of executions.
[39]	Prediction of the 3-point shot in the basketball game	RNN	Evaluated in terms of AUC and achieved 84.30%.	The proposed method fails in the case of high ball velocity and the noisy nature of motion data.

4.2. Soccer

Soccer is played using football, and eleven players in two teams compete to deliver the ball into the other team's goal, thereby scoring a goal. The players confuse each other by changing their speed or direction unexpectedly. Due to them having the same jersey color, players look almost identical and are frequently possess the ball, which leads to severe occlusions and tracking ambiguities. In such a case, a jersey number must be detected to recognize the player [60]. Accurate tracking [61–72] by detection [73–76] of multiple soccer players as well as the ball in real-time is a major challenge to evaluate the performance of the players, to find their relative positions at regular intervals, and to link spatiotemporal data to extract trajectories. The systems which evaluate the player [77] or team performance [78] have the potential to understand the game's aspects, which are not obvious to the human eye. These systems are able to evaluate the activities of players successfully [79] such as the distance covered by players, shot detection [80,81], the number of sprints, player's position, and their movements [82,83], the player's relative position concerning other players, possession [84] of the soccer ball and motion/gesture recognition of the referee [85], predicting player trajectories for shot situations [86]. The generated data can be used to evaluate individual player performance, occlusion handling [21] by the detecting position of the player [87], action recognition [88], predicting and classifying the passes [89–91], key event extraction [92–101], tactical performance of the team [102–106], and analyzing the team's tactics based on the team formation [107–109], along with generating highlights [110–113]. Table 3 summarizes various proposed methodologies to resolve various challenging tasks in soccer with their limitations.

Table 3. Studies in Soccer.

		Studies in Basketball		
Ref.	Problem Statement	Proposed Methodology	Precision and Performance Characteristics	Limitations and Remarks
[71]	Player and ball detection and tracking in soccer.	YOLOv3 and SORT	Methodology achieved tracking accuracy of 93.7% on multiple object tracking accuracy metrics with a detection speed of 23.7 FPS and a tracking speed of 11.3 FPS.	This methodology effectively handles challenging situations, such as partial occlusions, players and the ball reappearing after a few frames, but fails when the players are severely occluded.
[72]	Player, referee and ball detection and tracking by jersey color recognition in soccer.	DeepPlayerTrack	The model achieved a tracking accuracy of 96% and 60% on MOTA and GMOTA metrics, respectively, with a detection speed of 23 FPS.	The limitation of this method is that, when a player with the same jersey color is occluded, the ID of the player is switched.
[77]	Tracking soccer players to evaluate the number of goals scored by a player.	Machine Learning and Deep Reinforcement Learning.	Performance of the player tracking model measured in terms of mAP achieved 74.6%.	The method failed to track the ball at critical moments such as passing at the beginning and shooting. It also failed to overcome the identity switching problem.
[94]	Extracting ball events to classify the player's passing style.	Convolutional Auto-Encoder	The methodology was evaluated in terms of accuracy and achieved 76.5% for 20 players.	Concatenation of the auto-encoder and extreme learning machine techniques will improve classification of the event performance.
[101]	Detecting events in soccer.	Variational Auto- encoder and EfficientNet	Achieved an F1-score of 95.2% event images and recall of 51.2% on images not related to soccer at a threshold value of 0.50.	The deep extreme learning machine technique which employs the auto-encoder technique may enhance the event detection accuracy.
[82]	Action spotting soccer video.	YOLO-like encoder	The algorithm achieved an mAP of 62.5%.	-
[78]	Team performance analysis in soccer	SVM	Prediction models achieved an overall accuracy of 75.2% in predicting the correct segmental and the outcome of the likelihood of the team making a successful attempt to score a goal on the used dataset.	The proposed model failed in identifying the players that are more frequently involved in match events that end with an attempt at scoring i.e., a 'SHOT' at goal, which may assist sports analysts and team staff to develop strategies suited to an opponent's playing style.
[85]	Motion Recognition of assistant referees in soccer	AlexNet, VGGNet-16, ResNet-18, and DenseNet-121	The proposed algorithm achieved 97.56% accuracy with real-time operations.	Though the proposed algorithm is immune to variations of illuminance caused by weather conditions, it failed in the case of occlusions between referees and players.
[106]	Predicting the attributes (Loss or Win) in soccer.	ANN	The proposed model predicts 83.3% for the winning case and 72.7% for loss.	-
[109]	Team tactics estimation in soccer videos.	Deep Extreme Learning Machine (DELM).	The performance of the model is measured on precision, recall, and F1-score and achieved 87.6%, 88%, and 87.8%, respectively.	Team tactics are estimated based on the relationship between tactics of the two teams and ball possession. The method fails to estimate the team formation at the beginning of the game.
[88]	Action recognition in soccer	CNN-based Gaussian Weighted event-based Action Classifier architecture	Accuracy in terms of F1-score achieved was 52.8% for 6 classes.	By classifying the actions into subtypes, the accuracy of action recognition can be enhanced.
[62]	Detection and tracking of the ball in soccer videos.	VGG – MCNN	Achieved an accuracy of 87.45%.	It could not detect when the ball moved out of play in the field, in the stands region, or from partial occlusion by players, or when ball color matched the player's jersey.
[95]	Automatic event extraction for soccer videos based on multiple cameras.	YOLO	The U-encoder is designed for feature extraction and has better performance in terms of accuracy compared with fixed feature extractors.	To carry out a tactical analysis of the team, player trajectory needs to be analyzed.

Table 3. *Cont.*

			Studies in Basketball	
Ref.	**Problem Statement**	**Proposed Methodology**	**Precision and Performance Characteristics**	**Limitations and Remarks**
[80]	Shot detection in a football game	MobileNetV2	The MobileNetV2 method performed better than other feature extractor methods.	Extracting the features with the MobileNetV2 and then using 3D convolution on the extracted features for each frame can improve detection performance.
[86]	Predicting player trajectories for shot situations	LSTM	Performance is measured in terms of F1-score and achieved 53%.	The model failed to predict the player trajectory in the case of players confusing each other by changing their speed or direction unexpectedly.
[108]	Analyzing the team formation in soccer and formulating several design goals.	OpenCV is used for back-end visualization.	The formation detection model achieved a max accuracy of 96.8%.	The model is limited to scalability as it cannot be used on high-resolution soccer videos. The results are bounded to a particular match, and it cannot evaluate the tactical schemes across different games. Visualization of real-time team formation is another drawback as it limits the visualization of non-trivial spatial information. By applying state-of-the-art tracking algorithms, one can predominantly improve the performance of tactics analysis.
[60]	Player recognition with jersey number recognition.	Spatial Constellation + CNN	Achieved an accuracy of 82% by combining Spatial Constellation + CNN models.	The proposed model failed to handle the players that are not visible for certain periods. Predicting the position of invisible players could improve the quality of spatial constellation features.
[89]	Evaluating and classifying the passes in a football game.	SVM	The proposed model achieves an accuracy of 90.2% during a football match.	To determine the quality of each pass, some factors such as pass execution of player in a particularly difficult situation, the strategic value of the pass, and the riskiness of the pass need to be included. To rate the passes in sequence, it is necessary to consider the sequence of passes during which the player possesses the ball.
[84]	Detecting dribbling actions and estimating positional data of players in soccer.	Random forest	Achieved an accuracy of 93.3%.	The proposed methodology fails to evaluate the tactical strategies.
[103]	Team tactics estimation in soccer videos.	SVM	The performance of the methodology is measured in terms of precision, recall, and F1-score and achieved 98%, 97%, and 98%.	The model fails when audiovisual features could not recognize quick changes in the team's tactics.
[93]	Analyzing past events in the case of non-obvious insights in soccer.	k-NN, SVM	To extract the features of pass location, they used heatmap generation and achieved an accuracy of 87% in the classification task.	By incorporating temporal information, the classification accuracy can be improved and also offers specific insights into situations.
[61]	Tracking the players in soccer videos.	HOG + SVM	Player detection is evaluated in terms of accuracy and achieved 97.7%. Classification accuracy using k-NN achieved 93% for 15 classes.	-
[79]	Action classification in soccer videos	LSTM + RNN	The model achieves a classification rate of 92% on four types of activities.	By extracting the features of various activities, the accuracy of the classification rate can be improved.

4.3. Cricket

In many aspects of cricket as well, computer vision techniques can effectively replace manual analysis. A cricket match has many observable elements including bat-

ting shots [114–121], bowling performance [122–127], number of runs or score depending on ball movement, detecting and estimating the trajectory of the ball [128], decision making on placement of players' feet [129], outcome classification to generate commentary [130,131], detecting umpire decision [132,133]. Predicting an individual cricketer's performance [134,135] based upon his past record can be critical in team member selection at international competitions. Such process are highly subjective and usually require much expertise and negotiation decision-making. By predicting the results of cricket matches [136–140] such as the toss decision, home ground, player fitness, player performance criteria [141], and other dynamic strategies the winner can be estimated. The video summarization process gives a compact version of the original video for ease in managing the interesting video contents. Moreover, the video summarization methods capture the interest of the viewer by capturing exciting events from the original video [142,143]. Table 4 summarizes various proposed methodologies with their limitations to resolve various application issues in cricket.

Table 4. Studies in Cricket.

		Studies in Cricket		
Refs.	**Problem Statement**	**Proposed Methodology**	**Precision and Performance Characteristics**	**Limitations and Remarks**
[117]	Shot classification in cricket.	CNN—Gated Recurrent Unit	It is evaluated in terms of precision, recall, and F1-score and achieved 93.40%, 93.10%, and 93% for 10 types of shots.	By incorporating unorthodox shots which are played in t20 in the dataset may improve the testing accuracy.
[125]	Detecting the action of the bowler in cricket.	VGG16-CNN	It was evaluated in terms of precision, recall, and F1-score and the maximum average accuracy achieved is 98.6% for 13 classes (13 types of bowling actions).	Training the model with the dataset of wrong actions can improve detection accuracy.
[129]	Movement detection of the batsman in cricket.	Deep-LSTM	The model was evaluated in terms of mean square error and achieved a minimum error of 1.107.	-
[142,143]	Cricket video summarization.	Gated Recurrent Neural Network + Hybrid Rotation Forest-Deep Belief Networks YOLO	The methodology was evaluated in terms of precision, F1-score, accuracy and achieved 96.82%, 94.83%, and 96.32% for four classes. YOLO is evaluated on precision, recall, and F1-score and achieved 97.1%, 94.4%, and 95.7% for 8 classes.	Decision tree classifier performance is low due to the existence of a huge number of trees. Therefore, a small change in the decision tree may improve the prediction accuracy. Extreme Learning Machines have faced the problem of overfitting, which can be overcome by removing duplicate data in the dataset.
[134]	Prediction of individual player performance in cricket	Efficient Machine Learning Techniques	The proposed algorithm achieves a classification accuracy of 93.73% which is good compared with traditional classification algorithms.	Replacing machine learning techniques with deep learning techniques may improve the performance in prediction even in the case of different environmental conditions.
[114]	Classification of different batting shots in cricket.	CNN	The average classification in terms of precession is 0.80, Recall is 0.79 and F1-score is 0.79.	To improve the accuracy of classification, a deep learning algorithm has to be replaced with a better neural network.
[130]	Outcome classification task to create automatic commentary generation.	CNN + LSTM	Maximum of 85% of training accuracy and 74% validation accuracy	Due to the unavailability of the standard dataset for the ball by ball outcome classification in cricket, the accuracy is not up to mark. In addition, better accuracy leads to automatic commentary generation in sports.

Table 4. *Cont.*

		Studies in Cricket		
Refs.	Problem Statement	Proposed Methodology	Precision and Performance Characteristics	Limitations and Remarks
[132]	Detecting the third umpire decision and an automated scoring system in a cricket game.	CNN + Inception V3	It holds 94% accuracy in the Deep Conventional Neural Network (DCNN) and 100% in Inception V3 for the classification of umpire signals to automate the scoring system of cricket.	To build an automated umpiring system based on computer vision application and artificial intelligence, the results obtained in this paper are more than enough.
[133]	Classification of cricket bowlers based on their bowling actions.	CNN	The test set accuracy of the model is 93.3% which demonstrates its classification ability.	The model lacks data for detecting spin bowlers. As the dataset is confined to left-arm bowlers, the model misclassifies the right-arm bowlers.
[115]	Recognition of various batting shots in a cricket game	Deep-CNN	The proposed models can recognize a shot being played with 90% accuracy.	As the model is dependent on the frame per second of the video, it fails to recognize when the frames per second increases.
[131]	Automatic highlight generation in the game of cricket.	CNN + SVM	Mean Average Precision of 72.31%	The proposed method cannot clear metrics to evaluate the false positives in highlights.
[133]	Umpire pose detection and classification in cricket.	SVM	VGG19-Fc2 Player testing accuracy of 78.21%	Classification and summarization techniques can minimize false positives and false negatives.
[116]	Activity recognition for quality assessment of batting shots.	Decision Trees, k-Nearest Neighbours, and SVM.	The proposed method identifies 20 classes of batting shots with an average F1-score of 88% based on the recorded movement of data.	To assess the player's batting caliber, certain aspects of batting also need to be considered, i.e., the position of the batsman before playing a shot and the method of batting shots for a particular bowling type can be modeled.
[136,137]	Predicting the outcome of the cricket match.	k-NN, Naïve Bayesian, SVM, and Random Forest	Achieved an accuracy of 71% upon the statistics of 366 matches.	Imbalance in the dataset is one of the causes which produces lower accuracy. Deep learning methodologies may give promising results by training with a dataset that included added features.
[124]	Performance analysis of the bowler.	Multiple regression	Variation in ball speed has a feeble significance in influencing the bowling performance (the p-value being 0.069). The variance ratio of the regression equation to that of the residuals (F-value) is given as 3.394 with a corresponding p-value of 0.015.	-
[135]	Predicting the performance of the player.	Multilayer perceptron Neural Network	The model achieves an accuracy of 77% on batting performance and 63% on bowling performance.	-

4.4. Tennis

Worldwide, Tennis has experienced gain a huge popularity. This game need a meticulous analysis to reducing human errors and extracting several statistics from the game's visual feed. Automated ball and player tracking belongs to such class of systems that requires sophisticated algorithms for analysis. The primary data for tennis are obtained from ball and player tracking systems, such as HawkEye [144,145] and TennisSense [28,146]. The data from these systems can be used to detect and track the ball/player [147–150], visualizing the overall tennis match [151,152] and predicting trajectories of ball landing positions [153–155], player activity recognition [156–158], analyzing the movements of the player and ball [159], analyzing the player behavior [160] and predicting the next shot movement [161] and real-time tennis swing classification [162]. Table 5 summarizes various proposed methodologies to resolve various challenging tasks in tennis with their limitations.

Table 5. Studies in Tennis.

		Studies in Tennis		
Ref.	Problem Statement	Proposed Methodology	Precision and Performance Characteristics	Limitations and Remarks
[145]	Monitoring and Analyzing tactics of tennis players.	YOLOv3	The model achieved an mAP of 90% with 13 FPS on high-resolution images.	Using a lightweight backbone for detection, modules can improve the processing speed.
[158]	Player action recognition in tennis.	Temporal Deep Belief Network (Unsupervised Learning Model)	The accuracy of the recognition rate is 94.72%	If two different movements are similar, then the model fails to recognize the current action.
[162]	Tennis swing classification.	SVM, Neural Network, K-NN, Random Forest, Decision Tree	Maximum classification accuracy of 99.72% achieved using NN with a Recall of 1. The second-highest classification accuracy of 99.44% was achieved using K-NN with a recall of 0.98.	If the play styles of the players are different but the patterns are the same, in that case, models failed to classify the current swing direction.
[156]	Player activity recognition in a tennis game.	Long Short Term Memory (LSTM)	The average accuracy of player activity recognition based on the historical LSTM model was 0.95, and that of the typical LSTM model was 0.70.	The model lacks real-time learning ability and requires a large computing time at the training stage. The model also lacks online learning ability.
[147]	Automatic detection and classification of change of direction from player tracking data in a tennis game.	Random Forest Algorithm	Among all the proposed methods, model 1 had the highest F1-score of 0.801, as well as the smallest rate of false-negative classification (3.4%) and average accuracy of 80.2%	In the case of non-linear regression analysis, the classification performance of the proposed model is not up to the mark.
[153]	Prediction of shot location and type of shot in a tennis game.	Generative Adversarial Network (GAN) (Semi-Supervised Model)	The performance factor is measured based on the minimum distance recorded between predicted and ground truth shot location.	The performance of the model deviates from the different play styles as it is trained on the limited player dataset.
[159]	Analyzing individual tennis matches by capturing spatio-temporal data for player and ball movements.	For data extraction, a player and ball tracking system such as HawkEye is used.	Generation of 1-D space charts for patterns and point outcomes to analyze the player activity.	The performance of the model deviates from different matches, as it was trained only on limited tennis matches.
[157]	Action recognition in tennis	3-Layered LSTM	The classification accuracies are as follows: Improves from 84.10 to 88.16% for players of mixed abilities. Improves from 81.23 to 84.33% for amateurs and from 87.82 to 89.42% for professionals, when trained using the entire dataset.	The detection accuracy can be increased by incorporating spatio-temporal data and combining the action recognition data with statistical data.
[161]	Shot prediction and player behavior analysis in tennis	For data extraction, player and ball tracking systems such as HawkEye are used and a Dynamic Bayesian Network for shot prediction is used.	By combining factors (Outside, Left Top, Right Top, Right Bottom) together, speed, start location, the player movement assessment achieved better results of 74% AUC.	As the model is trained on limited data (only elite players), it cannot be performed on ordinary players across multiple tournaments.
[148]	Ball tracking in tennis	Two-Layered Data Association	Evaluation results in terms of precision, recall, F1-score are 84.39%, 75.81%, 79.87% for Australian open tennis matchwa and 82.34%, 67.01%, 73.89% for U.S open tennis matches.	The proposed method cannot handle multi-object tracking and it is possible to integrate audio information to facilitate high-level analysis of the game.
[160]	Highlight extraction from rocket sports videos based on human behavior analysis.	SVM	The proposed algorithm achieved an accuracy of 90.7% for tennis videos and 87.6% for badminton videos.	The proposed algorithm fails to recognize the player, as the player is a deformable object of which the limbs perform free movement during action recognition.

4.5. Volleyball

In volleyball, two teams of six players each are placed on either side of a net. Each team attempts to ground a ball on the opposite team's court and to score points under the defined rules. So, detecting and analyzing the player activities [163–165], detecting play patterns and classifying tactical behaviors [166–169], predicting league standings [170], detecting and classifying spiking skills [171,172], estimating the pose of the player [173], tracking the player [174], tracking the ball [175], etc., are the major aspects of volleyball analysis. Predicting the ball trajectory [59] in a volleyball game by observing the motion of the setter player has been conducted. Table 6 summarizes various proposed methodologies to resolve various challenging tasks in volleyball sport with their limitations.

Table 6. Studies in volleyball.

	Studies in Volleyball			
Ref.	Problem Statement	Proposed Methodology	Precision and Performance Characteristics	Limitations and Remarks
[173]	Group activity recognition by tracking players.	CNN + Bi-LSTM	The model achieved an accuracy of 93.9%.	The model fails to track the players if the video is taken from a dynamic camera. Temporal action localization can improve the accuracy of tracking the players in severe occlusion conditions.
[174]	Recognizing and classifying player's behavior.	SVM	The achieved recognition rate was 98% for 349 correct samples.	-
[167]	Classification of tactical behaviors in beach volleyball.	RNN + GRU	The model achieves better classification results as prediction accuracies range from 37% for forecasting the attack and direction to 60% for the prediction of success.	By employing a state-of-the-art method and training on a proper dataset that has continuous positional data, it is possible to predict tactics behavior and set/match outcomes.
[175]	Motion estimation for volleyball	Machine Vision and Classical particle filter.	Tracking accuracy is 89%	Replacing methods with deep learning algorithms gives better results.
[168]	Assessing the use of Inertial Measurement Units in the recognition of different volleyball actions.	KNN, Naïve Bayes, SVM	Unweighted Average Recall of 86.87%	By incorporating different frequency domain features, the performance factor can be improved.
[59]	Predicting the ball trajectory in a volleyball game by observing the the motion of the setter player.	Neural Network	The proposed method predicts 0.3 s in advance of the trajectory of the volleyball based on the motion of the setter player.	In the case of predicting the 3D body position data, the method records a large error. This can be overcome by training properly annotated large data on state-of-art-methods.
[164]	Activity recognition in beach volleyball	Deep Convolutional LSTM	The approach achieved a classification accuracy of 83.2%, which is superior compared with other classification algorithms.	Instead of using wearable devices, computer vision architectures can be used to classify the activities of the players in volleyball.
[170]	Volleyball skills and tactics analysis	ANN	Evaluated in terms of Average Relative Error for 10 samples and achieved 0.69%.	-
[165]	Group activity recognition in a volleyball game	LSTM	Group activity recognition of accuracy of the the proposed model in volleyball is 51.1%.	The performance of architecture is poor because of the lack of hierarchical considerations of the individual and group activity dataset.

4.6. Hockey/Ice Hockey

Hockey, also known as Field hockey, is an outdoor game played between two teams of 11 players each. These players use sticks that are curved at the striking end to hit a small and hard ball into their opponent's goal post. So, detecting [176] and tracking the

player/hockey ball, recognizing the actions of the player [177–179], estimating the pose of the player [180], classifying and tracking the players of the same team or different teams [181], referee gesture analysis [182,183] and hockey ball trajectory estimation are the major aspects of hockey sport.

Ice hockey is another similar game to field hockey, with two teams with six players each, wearing skates and competing on an ice rink. All players aim to propel a vulcanized rubber disk, the puck, past a goal line and into a net guarded by a goaltender. Ice hockey is gaining huge popularity on international platforms due to its speed and frequent physical contact. So, detecting/tracking the player [184–186], estimating the pose of the player [187], classifying and tracking with different identification the players of the same team or different teams, tracking the ice hockey puck [188], and classification of puck possession events [189] are the major aspects of the ice hockey sport. Table 7 summarizes various proposed methodologies to resolve various challenging tasks in hockey/ice hockey with their limitations.

Table 7. Studies in hockey.

Studies in Hockey				
Ref.	Problem Statement	Proposed Methodology	Precision and Performance Characteristics	Limitations and Remarks
[176]	Detecting the player in hockey.	SVM, Faster RCNN, SSD, YOLO	HD+SVM achieved the best results in terms of accuracy, recall, and F1-score with values of 77.24%, 69.23%, and 73.02%.	The model failed to detect the players in occlusion conditions.
[188]	Localizing puck Position and Event recognition.	Faster RCNN	Evaluated in terms of AUC and achieved 73.1%.	Replacing the detection method with the YOLO series can improve the performance.
[181]	Identification of players in hockey.	ResNet + LSTM	Achieves player identification accuracy of over 87% on the split dataset.	Some of the jersey number classes such as 1 to 4 are incorrectly predicted. The diagonal numbers from 1 to 100 are falsely classified due to the small number of training examples.
[177]	Activity recognition in a hockey game.	LSTM	The proposed model recognizes the activities such as free hits, goals, penalties corners, and long corners with an accuracy of 98%.	As the proposed model is focused on spatial features, it does not recognize activities such as free hits and long corners as they appear as similar patterns. By including temporal features and incorporating LSTM into the model, the model is robust to performance accuracy.
[180]	Pose estimation and temporal-based action recognition in hockey.	VGG19 + LiteFlowNet + CNN	A novel approach was designed and achieved an accuracy of 85% for action recognition.	The architecture is not robust to abrupt changes in the video, e.g., it fails to predict hockey sticks. Activities such as a goal being scored, or puck location, are not recognized.
[187]	Action recognition in ice hockey using a player pose sequence.	CNN+LSTM	The performance of the model is better in similar classes such as passing and shooting. It achieved 90% parameter reduction and 80% floating-point reduction on the HARPET dataset.	As the number of hidden units to LSTM increases, the number of parameters also increases, which leads to overfitting and low test accuracy.
[178]	Human activity recognition in hockey.	CNN+LSTM	An F1-score of 67% was calculated for action recognition on the multi-labeled imbalanced dataset.	The performance of the model is poor because of the improper imbalanced dataset.
[179]	Player action recognition in an ice hockey game	CNN	The accuracy of the actions recognized in a hockey game is 65% and when similar actions are merged accuracy rises to 78%.	Pose estimation problems due to severe occlusions when motions blur due to the speed of the game and also due to lack of a proper dataset to train models, all causing low accuracy.

4.7. Badminton

Badminton is one of the most popular racket sports, which includes tactics, techniques, and precise execution movements. To improve the performance of the player, technology plays a key role in optimizing the training of players; technology determines the movements of the player [190] during training and game situations such as with action recognition [191–193], analyzing the performance of player [194], detecting and tracking the shuttlecock [195–197]. Table 8 summarizes various proposed methodologies to resolve various challenging tasks in badminton with their limitations.

Table 8. Studies in badminton.

		Studies in Badminton		
Ref.	Problem Statement	Proposed Methodology	Precision and Performance Characteristics	Limitations and Remarks
[195]	Shuttlecock detection problem of a badminton robot.	Tiny YOLOv2 and YOLOv3	Results show that, compared with state-of-art methods, the proposed networks achieved good accuracy with efficient computation.	The proposed method fails to detect different environmental conditions. As it uses the binocular camera to detect a 2D shuttlecock, it cannot detect the 3D shuttlecock trajectory.
[191]	Automated badminton player action recognition in badminton games.	AlexNet+CNN, GoogleNet+CNN and SVM	Recognition of badminton actions by the linear SVM classifier for both AlexNet and GoogleNet using local and global extractor methods is 82 and 85.7%.	The architecture can be improved by fine-tuning in an end-to-end manner with a larger dataset on features extracted at different fully connected layers.
[192]	Badminton activity recognition	CNN	Nine different activities were distinguished: seven badminton strokes, displacement, and moments of rest. With accelerometer data, accurate estimation was conducted using CNN with 86% precision. Accuracy is raised to 99% when gyroscope data are combined with accelerometer data.	Computer vision techniques can be employed instead of sensors.
[193]	Classification of badminton match images to recognize the different actions were conducted by the athletes.	AlexNet, GoogleNet, VGG-19 + CNN	Significantly, the GoogleNet model has the highest accuracy compared to other models in which only two-hit actions were falsely classified as non-hit actions.	The proposed method classifies the hit and non-hit actions and it can be improved by classifying more actions in various sports.
[196]	Tracking shuttlecocks in badminton	An AdaBoost algorithm which can be trained using the OpenCV Library.	The performance of the proposed algorithm was evaluated based on precision and it achieved an average precision accuracy of 94.52% with 10.65 fps.	The accuracy of tracking shuttlecocks is enhanced by replacing state-of-the-art AI algorithms.
[190]	Tactical movement classification in badminton	k-Nearest Neighbor	The average accuracy of player position detection is 96.03 and 97.09% on two halves of a badminton court.	The unique properties of application such as the length of frequent trajectories or the dimensions of the vector space may improve classification performance.

4.8. Miscellaneous

Player detection and tracking is the major requirement in athletic sports such as running, swimming [198,199], and cycling. In sports such as table tennis [200], squash [201,202], and golf [203], ball detection and tracking and player pose detection [204] are challenging tasks. In ball-centric sports such as rugby, American football, handball, baseball, ball/player detection [205–211] and tracking [212–221], analyzing the action of the player [23,222–227], event detection and classification [228–232], performance analysis of player [233–235], referee identification and gesture recognition are the major challenging tasks. Video highlight generation is a subclass of video summarization [236–239] which may be viewed as a subclass of sports video analysis. Table 9 summarizes various proposed methodologies to resolve various challenging tasks in various sports with their limitations.

Table 9. Studies in various sports.

		Studies in Various Sports		
Ref.	Problem Statement	Proposed Methodology	Precision and Performance Characteristics	Limitations and Remarks
[211]	Beach sports image recognition and classification.	CNN	The model achieved a recognition accuracy of 91%.	Lightweight networks of deep learning algorithms can improve the recognition accuracy and can also be implemented in real-time scenarios.
[199]	Motion image segmentation in the sport of swimming	GDA + SVM	The performance of the Symmetric Difference Algorithm was measured in terms of recall and achieved 76.2%.	Using advanced optimization techniques such as Cosine Annealing Schedulers with deep learning algorithms may improve the performance.
[200]	Identifying and recognizing wrong strokes in table tennis.	k-NN, SVM, Naïve Bayes	Performs various ML algorithms and achieves an accuracy of 69.93% using the Naïve Bayes algorithm.	A standard dataset can improve the accuracy of recognizing the wrong strokes in table tennis.
[212]	Multi-player tracking in sports	Cascade Mask R-CNN	The proposed Deep Player Identification method studies the patterns of jersey number, team class, and pose-guided partial feature. To handle player identity switching, the method correlates the coefficients of player ID in the K-shortest path with ID. The proposed framework achieves state-of-art performance.	When compared with existing methods, the computation cost is higher and can be considered a major drawback of the proposed framework. To refine 2D detection, temporal information needs to be considered and can be transferred to tracking against a real-time performance such as soccer, basketball, etc.
[215]	Individual player tracking in sports events.	Deep Neural Network	Achieved an Area Under Curve (AUC) of 66%	Tracking by jersey number recognition may increase the performance of the model.
[204]	Skelton-based key pose recognition and classification in sports	Boltzmann machine+CNN Deep Boltzmann machine + RNN	The proposed architecture successfully analyses feature extraction, motion attitude model, motion detection, and behavior recognition of sports postures.	The architecture is bound to individual-oriented sports and can be further implemented on group-based sports, in case of challenges such as severe occlusion, misdetection due to failure in blob detection in object tracking.
[224]	Human action recognition and classification in sports	VGG 16 + RNN	The proposed method achieved an accuracy of 92% for ten types of sports classification.	The model fails in the case of scaling up the dataset for larger classification which shows ambiguity between players and similar environmental conditions. Football, Hockey; Tennis, Badminton; Skiing, Snowboarding; these pairs of classes have similar environmental features; thus, it is only possible to separate them based on relevant actions which can be achieved by state-of-the-art methods.
[237]	Replay and key event detection for sports video summarization	Extreme Learning Machine (ELM)	The framework is evaluated on a dataset that consists of 20 videos of four different sports categories. It achieves an average accuracy of 95.8%, which illustrates the significance of the method in terms of key-event and replay detection for video summarization.	The performance of the proposed method drops in the case of the absence of a gradual transition of a replay segment. It can be extended by incorporating artificial intelligence techniques.
[228]	Event detection in sports videos for unstructured environments with arbitrary camera angles.	Mask RCNN + LSTM	The proposed method is accurate in unsupervised player extraction, which is used for precise temporal segmentation of rally scenes. It is also robust to noise in the form of camera shaking and occlusions.	It can be extended to doubles games with fine-grained action recognition for detecting various kinds of shots in an unstructured video and it can be extended to analyze videos of games such as cricket, soccer, etc.
[229]	Human motion quality assessment in complex motion scenarios.	3-Dimensional CNN	Achieved an accuracy of 81% on the MS-COCO dataset.	Instead of the Stochastic Gradient Descent technique for learning rate, using the Cosine annealing scheduler technique may improve the performance.

Table 9. *Cont.*

		Studies in Various Sports		
Ref.	Problem Statement	Proposed Methodology	Precision and Performance Characteristics	Limitations and Remarks
[213]	Court detection using markers, player detection, and tracking using a drone.	Template Matching + Particle Filter	The proposed method achieves better accuracy (94%) in the case of two overlapping players	As the overlapping of players, increases the accuracy of detection and tracking decreases due to similar features of players on the same team. The method uses a template matching algorithm, which can be replaced with a deep learning-based state-of-art algorithm to acquire better results.
[214]	Target tracking theory and analyses its advantages in video tracking.	Mean Shift + Particle Filter	Achieves better tracking accuracy compared to existing algorithms such as TMS and CMS algorithms.	If the target scales change then the tracking of players fails due to the unchanged window of the mean-shift algorithm. Furthermore, it cannot track objects which are similar to the background color. The accuracy of tracking players can be improved by replacing them with artificial intelligence algorithms.
[236]	Automatically generating a summary of sports video.	2D CNN + LSTM	Describes a novel method for automatic summarization of user-generated sports videos and demonstrated the results for Japanese fencing videos.	The architecture can be improved by fine-tuning in an end-to-end manner with a larger dataset for illustrating potential performance and also to evaluate in the context of a wider variety of sports.
[227]	Action Recognition and classification	SVM	Achieved an accuracy of 59.47% on the HMDB 51 dataset.	In cases where the object takes up most of the frame, the human detector cannot completely cover the body of the object. This leads to the system missing movements of body parts such as hands and arms. In addition, recognition of similar movements is a challenge for this architecture.

4.9. Overview of Machine Learning/Deep Learning Techniques

There are multiple ways to classify, detect, and track objects to analyze the semantic levels involved in various sports. They pave the way for player localization, jersey number recognition, event classification and trajectory forecasting of the ball in a sports video with a much better interpretation of an image as a whole.

The selected AI algorithm is better if it is tested and benchmarked on different data. To evaluate the robustness of AI algorithms, some metrics are required which measure the performance of particular AI algorithms to enable better selection. Figure 8 depicts the road map of the machine learning algorithms' general information, methods, and evaluation criteria for a particular task and required libraries/tools for training the model. Figure 9 depicts the roadmap of the deep learning algorithm selection, training, and evaluation criteria for a particular task and required libraries/tools for training the model. Figure 10 shows taxonomy of various deep learning techniques of classification [240–245], detection [246] and prediction [247–249] algorithms, unsupervised learning [250,251], tracking [252–261], and trajectory prediction [262–269]. Since various tasks in sports such as classification/detection, tracking, and trajectory prediction show great advantages in various sports. A bi-layered parallel training architecture in distributed computing environments was introduced in [270], which discusses the time-consuming training process of large-scale deep learning algorithms.
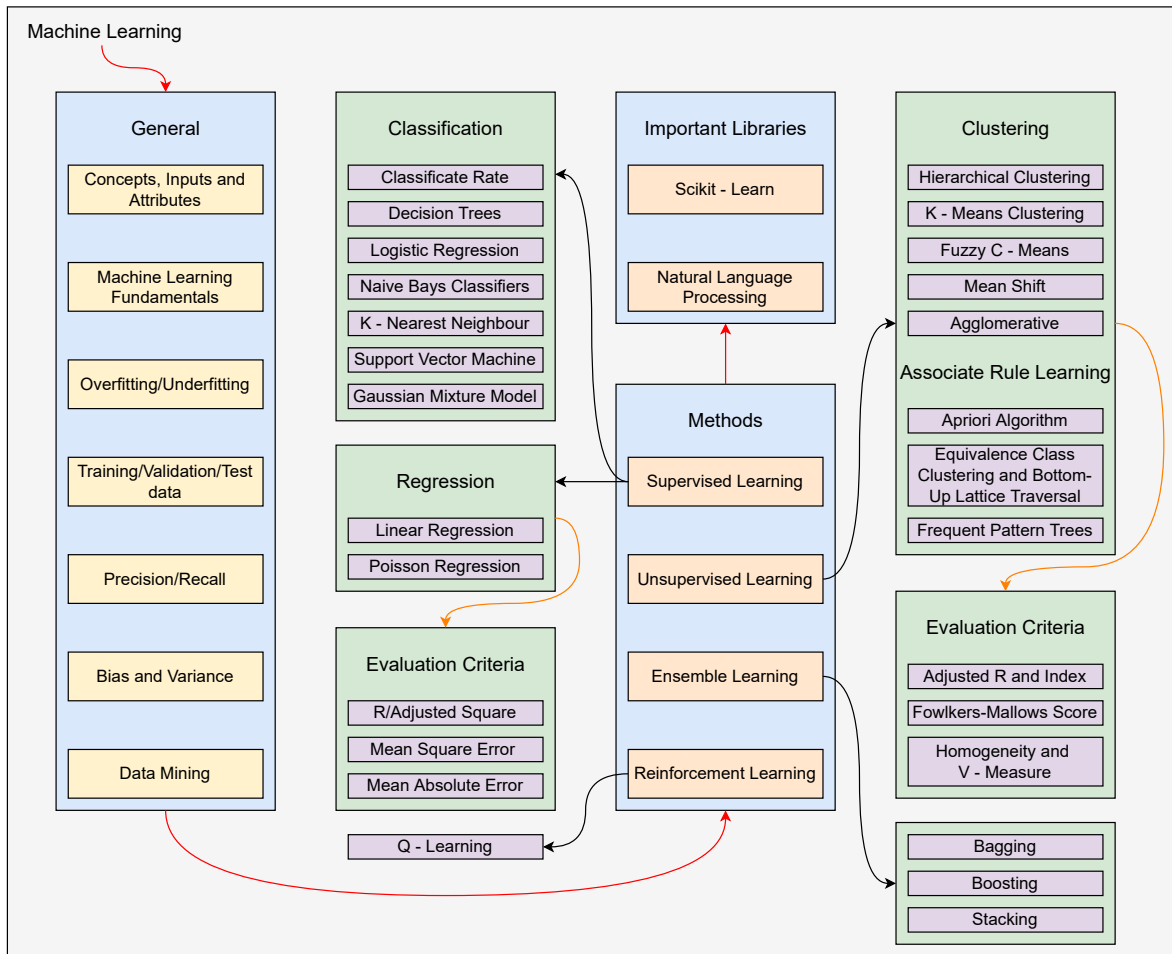
Figure 8. Block diagram of the road map to machine learning architecture selection and training.
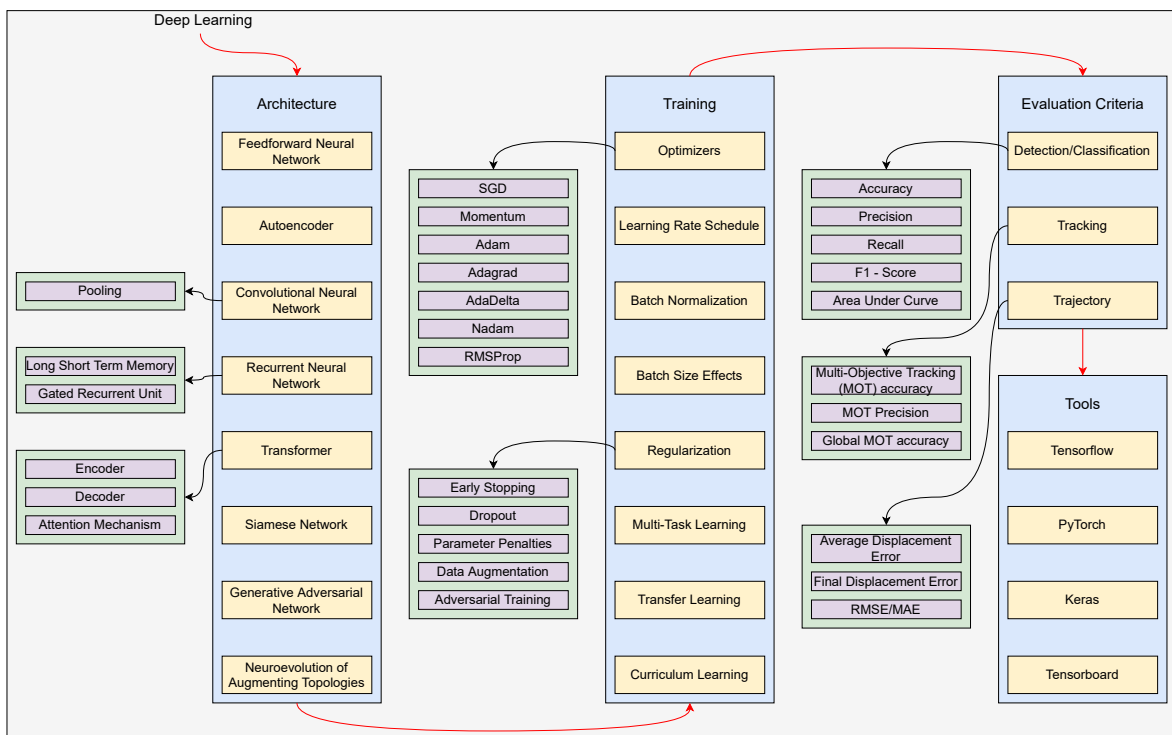

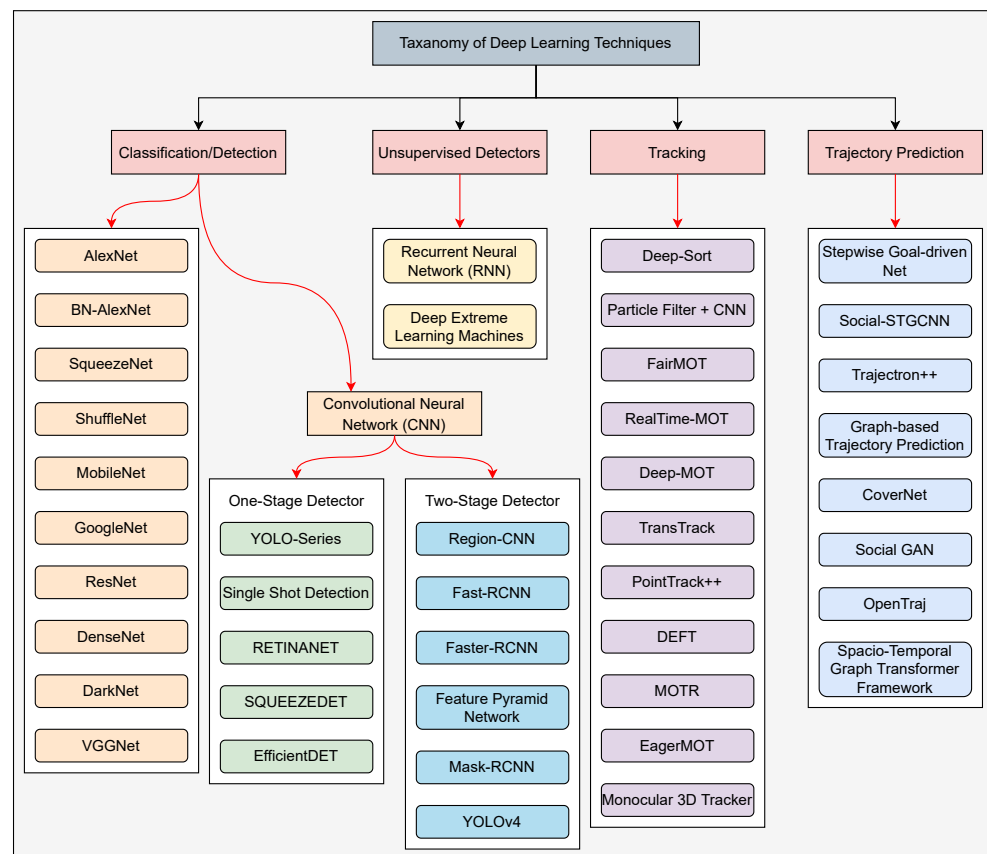**Figure 9.** Block diagram of the road map to deep learning architecture selection and training.

Figure 10. Overview of deep learning algorithms of classification/detection, tracking and trajectory prediction.

5. Available Datasets of Sports

In this section, a brief description of some sports video datasets which are available publicly with annotations is provided. Utilizing these shared datasets provides a platform for comparison of the performance of algorithms with common data for improving the transparency in research in this domain. Additionally, sharing the data among the users (researchers) reduces time-consuming efforts in capturing and annotating large quantities of videos in diversified areas. This allows users to obtain benchmark scores for the algorithms developed.

These shared datasets can be categorized into two types: videos or still images, which are typically taken with moving cameras, particularly of individual athletes or of team sports, for recognition of player actions [96,165,225,226,271–274], event detection and classification [34,98,275], which are often captured using several setups of static cameras, for detection and tracking of players/balls [276–278], pose estimation [279], and sports event summarization [280] of team plays. One dataset is focused on the spectators' actions in sports rather than those of the players. These datasets which are available for analysis largely perform a great variety of actions. Table 10 describes the available datasets of various sports, the mode of the dataset, annotated parameters, number of frames, and length of the video.

Table 10. Details of the available datasets.

		Details of the Available Dataset			
Refs.	Sport	Dataset	Mode of the Dataset	Annotated Parameters	Length of the Video and Number of Images
[101]	Soccer	Image type Football Keyword Dataset	Event detection and Classification.	Events such as free kicks, penalty kicks, tackles, red cards, yellow cards.	Dataset was categorized as training, testing, and validation with 5000, 500, and 500 images.
[271]	Basketball	Basketball dataset	Action Recognition	Dribbling, Passing, Shooting	Dataset consists of video of 8 h duration, 3399 annotations and 130 samples of each class.
[121]	Cricket	Video type Cricket Strokes Dataset	Cricket Stroke Localization.	Annotated with strokes played in untrimmed videos.	Highlights dataset and the Generic dataset comprised of Cricket telecast videos at 25 fps.
[272]	Table Tennis	Video type TTNet dataset	Ball detection and Event Spotting	Ball bounces, Net hit, Empty Events	5 Videos of 10–25 min duration for training and 7 short videos for testing
[96,272]	Soccer	Image type SoccerNet and SoccerNetv2 dataset	Action spotting in soccer videos	Goal, Yellow/Red Card, Substitution	Handles a length of video of about 764 h and 6637 moments which are split into three major classes (Substitution, Goal, and Yellow/Red Card).
[165]	Volleyball	Volleyball dataset	Group activity recognition	Person-level actions, Temporal dynamics of a person's action and Temporal evolution of group activity	1525 frames were annotated.
[273]	Hockey	The Spectators of Hockey (S-HOCK)	Analyzing Crowds at the Stadium	spectator categorization such as position, head pose, posture and action	Video type dataset. 31 s 30 fps
[281]	487 classes of sports	Sports-1M dataset	Sports classification and activity recognition.	Activity labels	5 m 36 s
[276]	Soccer	Player position in soccer video dataset	Player tracking system	Trajectories of players	45 min
[225]	Volleyball	Indoor volleyball dataset	Activity Detection and Recognition	Seven activities such as the serve, reception, setting, attack, block, stand, and defense/move are annotated to each player in this dataset.	23 min 25 fps
[274]	Various jump games, various throw games, bowling, tennis serving, diving, and weightlifting.	Olympic Sports Dataset	Recognition of complex human activities in sports	Different poses in different sports	Video type dataset. It contains 16 sports classes, with 50 sequences per class.
[34]	Basketball	NCAA Dataset	Event Recognition	Event classification, Event detection and Evaluation of attention.	The video type dataset. Length 1.5 h long Annotated with 11 types of events.
[277]	Soccer	ISSIA Soccer Dataset	Objective method of Ground Truth Generation	Player and Ball trajectories	2 min 25 fps
[280]	Basketball	APIDIS Basketball Dataset	sport-event summarization	Basketball events such as the position of players, referees, and the ball.	16 min, 22fps

Table 10. *Cont.*

			Details of the available dataset		
Refs.	Sport	Dataset	Mode of Dataset	Annotated Parameters	Length of Video and Number of Images
[279]	Badminton, Basketball, Football, Rugby, Tennis, Volleyball	Martial Arts, Dancing and Sports dataset	3D human pose estimation.	It is annotated with 5 types of actions.	Size of the dataset is 53,000 frames.
[226]	Kicking, golf swing, lifting, diving, riding horses, skateboarding, running, walking, swing-side.	UCF Sports Action Dataset	Action Recognition	Action localization and the class label for activity recognition.	13k clips and 27 h of video data
[278]	Soccer	SSET	Shot segmentation, Event detection, Player Tracking	Far-view shot, Medium-view shot, Close-view shot, Out-of-field shot, and Playback shot.	350 soccer videos of total length 282 h 25 fps

The parameters which are annotated in the ISSIA dataset relate to the positions of the ball, player, and referee in each video from each camera. The images shown in Figure 11 are a few sample frames from the ISSIA dataset.

Figure 11. Instances from the ISSIA dataset [277].

The parameters which are annotated in the TTNet dataset are the ball bouncing moments, the ball hitting the net, and empty events. The images shown in Figure 12 are a few sample frames from the TTNet dataset.

Figure 12. Instances from the TTNet dataset [275].

For the creation of the APIDIS dataset, videos were captured from seven cameras from above and around the court. The events which are annotated in this dataset are player positions, movements of referees, baskets, and the position of the ball. The images shown in Figure 13 are a few samples from the APIDIS dataset.

Figure 13. Instances from the APIDIS dataset [280].

6. GPU-Based Work Stations and Embedded Platforms

To find the target, GPU-constrained devices such as Raspberry Pi, Latte Panda, Odroid Xu4 and Computer Vision were used. The disadvantages of machine learning techniques are that they provide poor or inaccurate results and have issues in predicting an unknown future data, whereas deep learning algorithms provide accurate results and also make predictions from unknown future data. Segmentation, localization and image classification are visual recognition systems which have prominent research contributions.

Among embedded AI computing platforms, NVidia Jetson devices provide low-power computing and high-performance support for artificial intelligence-based visual recognition systems. Jetson modules are configured with OpenCV, cuDNN, CUDA Toolkit, L4T with LTS Linux kernel and TensorRT. The Intel Movidius Neural Compute Stick uses the Intel Movidius Neural Compute SDK in GPU-Constrained devices to deploy AI algorithms.

Wang et al. [203] presented a high-speed stereo vision system that can track the motion of a golf ball at a speed of 360 km/h under indoor lighting conditions. They implemented the algorithm on a field-programmable gate array board with an advanced RISC machine CPU [62] which implemented a deep learning approach to track the soccer ball on NVIDIA GTX1050Ti GPU [43] and a deep learning algorithm on GTX 1080 ti GPU, based on CUDA 9.0 and Caffe to analyze the technical features in basketball video. Table 11 shows the basic comparison between GPU-based devices and GPU-Constraint Devices and possible deep learning algorithms to implement on various devices.

Table 11. Comparison between Jetson Modules and GPU-Constrained Devices.

	Jetson TX1	Jetson TX2	Jetson AGX Xavier	Raspberry Pi Series	Latte Panda	Odroid Xu4
CPU	Quad-core ARM Processor	Dual-core Denver CPU and quad-core ARM 57	64 bit CPU and 8-core ARM	63-bit quad-core ARM	Intel Cherry Train quad-core CPU	Cortex A7 octa-core CPU
GPU	NVIDIA Maxwell with CUDA cores	NVIDIA Pascall with CUDA cores	Tensor cores + 512-core Volta GPU	-	-	-
Memory	4GB Memory	8 GB Memory	16 GB Memory	1 GB Memory	4 GB Memory	Stacked memory of 2 GB
Storage	16 GB Flash Storage	32 GB storage	32 GB storage	Support Micro SD card	64 GB storage	Support Micro SD Card
Possible DL Algorithms to Implement	YOLO v2 and v3, tiny YOLO v3, SSD, Faster–RCNN and Tracking algorithm like YOLO v3 + Deep SORT, YOLO v4, YOLOR			YOLO, YOLO v2 and SSD-MobileNet etc.		

A Field Programmable Gate Array (FPGA) has also been used in sports involving 3D motion capturing, object movement analysis and image recognition, etc. Table 12 describes how different researchers performed various studies of sports on hardware platforms such as FPGA and GPU-based devices and their results in terms of performance measures are listed.

Table 12. Performance of various studies on hardware platforms.

Ref.	GPU-Based Work Station	Embedded Platform	Problem Statement	Performance Measures	Result
[29]	NVidia Titan X GPU.	-	Event classifications in basketball videos.	Average Accuracy	58.10%
[38]	NVIDIA GTX 960	-	Basketball trajectory prediction based on real data and generating new trajectory samples.	Measured in terms of AUC	91.00%
[282]	-	FPGA	Recognizing swimming styles of a swimmer.	The result shows the three-level identification system in Average, Minimum and Maximum offset.	4.14%, 2.16%, 5.77%.
[198]	-	-	-	Recall and Specificity	85% and 96.6%
[33]	NVidia GeForce GTX 1080Ti	-	Tracking ball movements and classification of players in a basketball game	Precision and Recall	74.3% and 89.8%
[282]		FPGA	Ball detection and tracking to reconstruct trajectories in basketball.	Accuracy	>90%
[43]	NVidia GTX 1080 ti GPU	-	Analyzing the behavior of the player.	Accuracy	83%
[283]	-	FPGA	Detecting the movement of the ball in basketball.	Average rate vs Frame range	Varies from 12 to 100% for different frame ranges.
[215]	NVidia GTX 1080Ti	-	Individual player tracking in sports events.	Achieved an Area Under Curve (AUC)	66%
[88]	NVidia GeForce GTX 1080Ti	-	Action recognition in soccer	Accuracy in terms of F1-score	52.80%
[284]	-	FPGA	Movement classification in basketball based on Virtual Reality Technology to improve basketball coaching.	Accuracy	93.50%
[62]	NVidia GTX 1050Ti	-	Detection and tracking of the ball in soccer videos.	Accuracy	87.45%
[285]	-	FPGA	Action recognition based on arm movement in basketball.	Accuracy	92.30%

7. Applications in Sports Vision

A fan who is digitally connected becomes the biggest online influencer of sports venues. Teams and stadium owners provide plenty of personalized experiences through their custom apps, mobile phone support for content with offers and live updates of game information using digital boards to increase the engagement of fans and in turn generate opportunities for new revenues [286]. Figure 14 depicts where AI technology can be used within the sporting landscape.
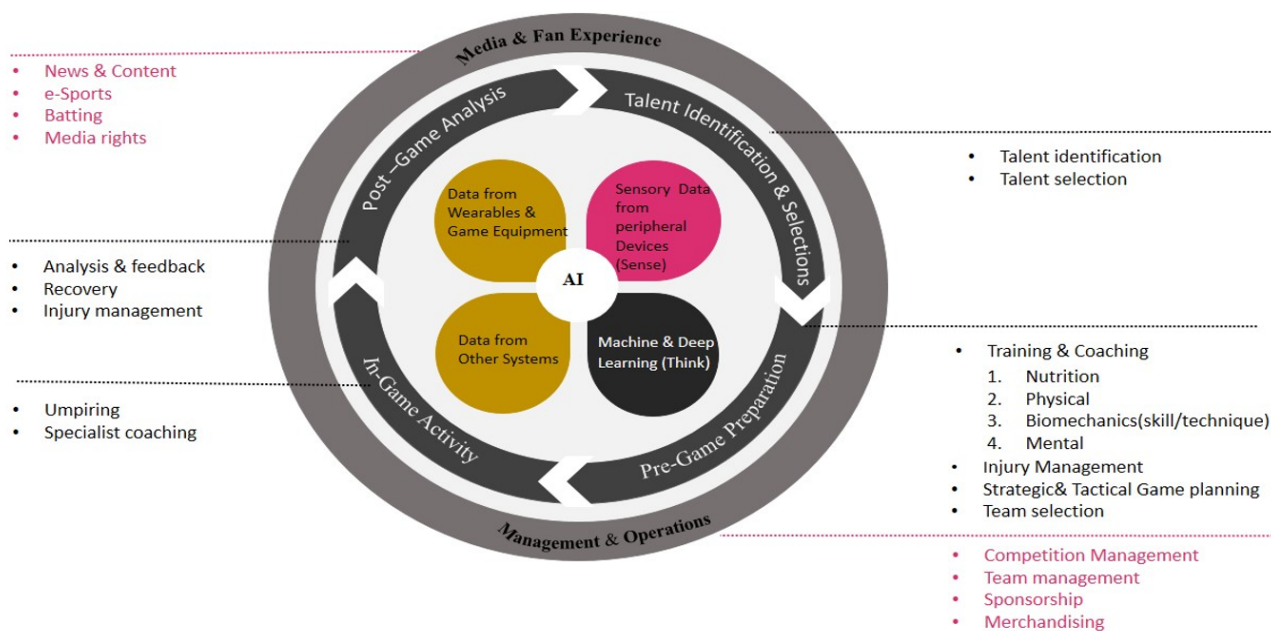

Figure 14. AI technology framework for the sport industry.

Modern artificial intelligence fields are not good sparring partners, but they can be valuable as research tools. One of the most effective methods to improve this is to learn from one's failures. The suggested method to improve playing abilities is to review games, but how can one detect the mistakes? How can one come up with better alternatives? This challenge is solved by the field of artificial intelligence analysis tools such as AlphaGO [287–289], which provide probability distribution of smart moves and their assessment.

7.1. Chabots and Smart Assistants

Recently, sports bodies such as the NHL and NBA have started using virtual assistants to respond to inquiries made by fans in a wide range of topics such as ticketing, arena logistics, parking, and other game-related information. If the bots are not capable, such scenarios are handled by human intervention and they maintain customer services for that.

7.2. Video Highlights

The challenges facing the industry include not just the creation of content but also delivering it to customers through multiple devices and screens for viewing different content at different times. There is a serious demand from fans for in-depth analysis and also for commentary. Many others like action-packed highlights and some behind-the-scenes content as well. Introducing AI enables the solving of challenging tasks in various sports and it provides an exciting viewing experience to the audience, attracting more viewers.

7.3. Training and Coaching

Effective methods for improving the analysis of the performance of athletes and also assisting coaches with team guidance to gauge the tactics of the opponents are gaining popularity.

An application that uses AI contains huge a dataset of game performances and training-related information, which is backed up with the knowledge of several coaches and sports scientists. They act as an accumulated source of current knowledge on the dissemination of the latest techniques, tactics, or knowledge for professional coaches.

With the evolution of knowledge on any tactic or technique, the knowledge base of AI is updated. The accumulated data can be used for training and educating sports coaches, scientists, and also athletes, which in turn leads to improved performance.

7.4. Virtual Umpires

In cricket and tennis, Video Assistant Referees (VAR) and Decision Review Systems (DRS) have used Hawk-eye, slow-motion replays, and some other technologies. However, the catch is that these involve a request from players or team for review when an umpire's or referee's decision has some uncertainty and then involves other parties to assist the main umpire. The whole process is time-consuming and takes away the momentum and excitement of the game.

The latest camera technology supporting AI software creates a situation where an umpire's role is limited to on-field behavior management of players rather than making critical decisions. For example, in the case of tennis, computer vision is used for detecting placement and speed of the ball; therefore, the need for a line umpire is eliminated. The future scope can be an umpire's earpiece and glasses assisting the decision instantly and eliminating the necessity of reviews.

7.5. AI Assistant Coaches

AI can be way more capable in situations involving dynamic planning and analysis of the scenes where a coach would rely on previous data and experience, and it is not effective enough to frame dynamically changing strategies in comparison to a machine. A future can be imagined in which a machine with AI running alongside the gameplay dynamically predicts and creates strategies, helping the teams to gain an edge over others.

One example where the evolution of AI can be seen is Chess. The evolution of chess technology is shown in Figure 15. The Russian Garry Kasparov who was considered the world's number 1 for about 19 years with an Elo rating (skill level measurement) 2851 was surpassed by Magnus Carlsen with Elo rating 2882 in 2014.
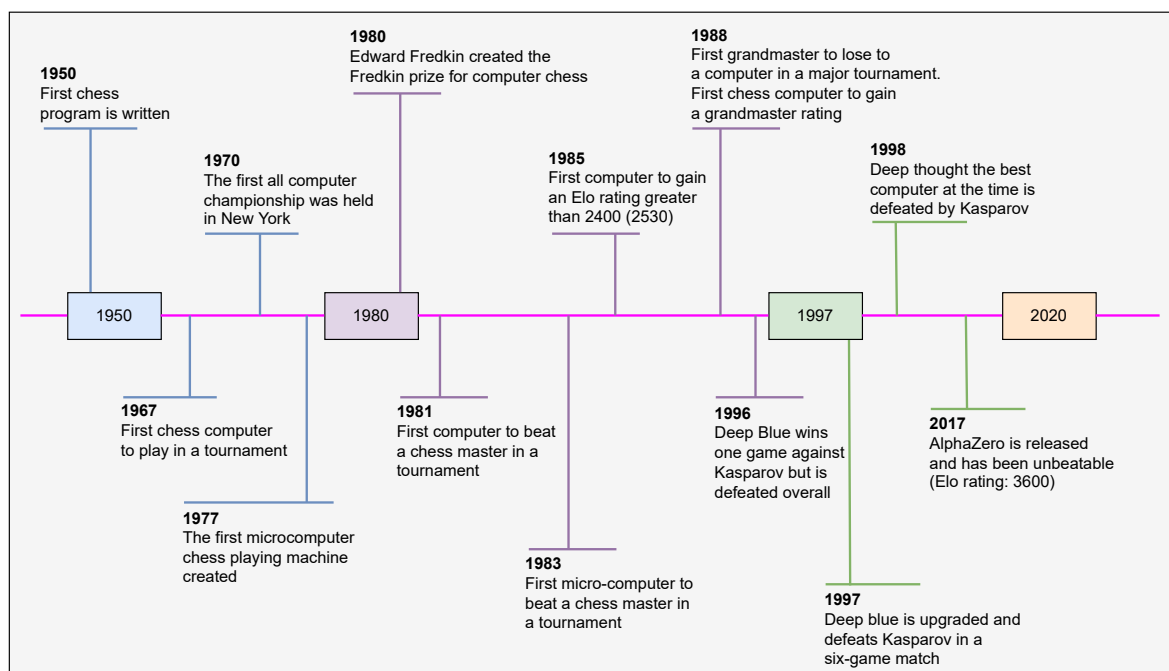

Figure 15. Evolution of chess technology demonstrates the speed of AI adoption.

In computers, Deep Blue's rating which was 2700+ was surpassed by Deep Mind's Alpha Zero with an estimated Elo 3600, which was developed by Google's sibling DeepMind. It was developed by a reinforcement learning technique called self-play. It took just 24 h to achieve it and proves the capabilities of the machine.

7.6. Available Commercial Systems for Player and Ball Tracking

Hawkeye [290–292] is the technology for ball tracking in cricket, tennis, and soccer. The area of primary application is officiating in these sports to enhance broadcast videos. Figure 16 shows the visualization performance of the commercial systems.

Figure 16. Hawkeye technology in cricket, tennis and soccer [290–292].

STATS SportVU [293] and ChyronHego TRACAB [294] are the technologies for player tracking in sports. The area of primary application is to track players in various sports, analyze their performance and assist coaches for training. Figure 17 shows the player position and pose estimation using commercial systems. SportsVu is a computer vision technology that provides real-time optical tracking in various sports. It provides in-depth performance of any team, such as tracking every player from both teams to provide comprehensive match coverage, collecting data to provide tactical analysis of the match, and highlighting the performance deviations to reduce injuries in the game.

Figure 17. TRACAB Gen5 Technology for player tracking [294].

8. Research Directions in Sports Vision

Based on the investigation of the available articles in sports, we were able to come out with various research topics and identifies research directions to be taken for further research in sports. They are categorized based on the task specifics in sports applications (such as major sports in which player/ball/referee detection and tracking, pose estimation, trajectory prediction are required) as shown in Figure 18 to provide promising and potential research directions for future computer vision/video processing in sports.

As sporting activities are dynamic, the accuracy and reliability of single- or multiplayer tracking [63,215] in real-time sports video can be enhanced by proposing a framework that learns object identities with deep representations which resolve the problem of identity switching among players [212]. By considering the temporal information, the performance of the tracking algorithm can become robust to overcome problems such as severe occlusions and miss-detection.

The accuracy of classifying different defensive strategies of soccer [49] can be improved by labeling large spatio-temporal datasets and by classifying the actions into subtypes [88]. The performance measures of team tactics analysis [93,95] of soccer videos can be enhanced by analyzing player trajectories. By incorporating the temporal information, the classification accuracy can be improved while it also offers more specific insights into situations such as pass events in the case of non-obvious insights in soccer videos [93]. Accurate pose detection as shown in Figure 19a is still a major challenge to identify whether the player is running, jumping, or walking as shown in Figure 19b, and also to handle the severe occlusions or identity switches among players.
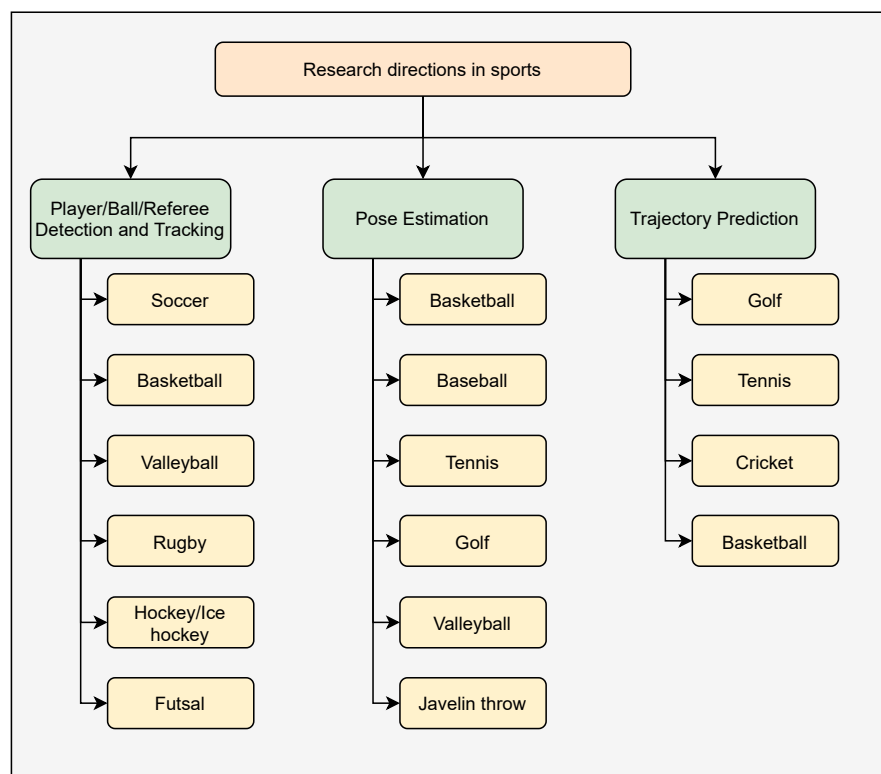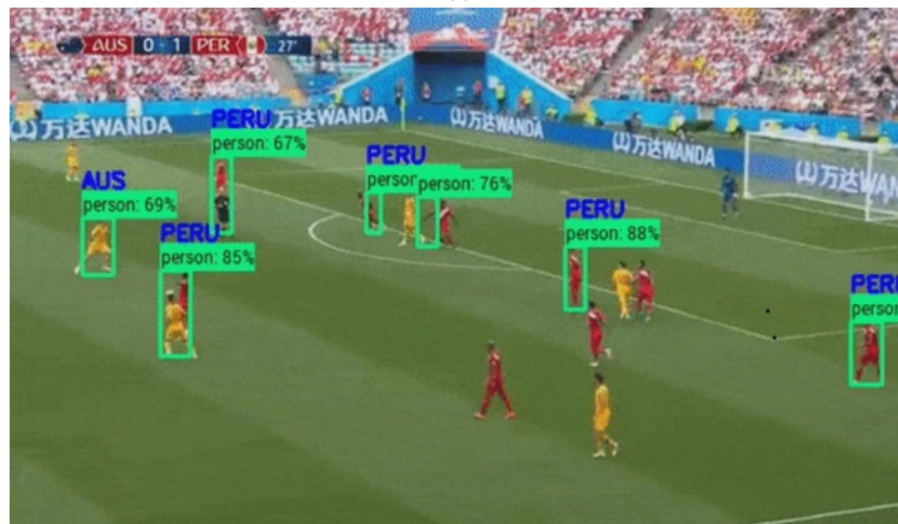

Figure 18. Major task specifics in sports applications.

To assess the batter's caliber, certain aspects of batting need to be considered, i.e., position of batsman before playing a shot, and the method of batting shots for a particular bowling type needs to be modeled [116]. Classification and summarization techniques can minimize false positives and false negatives to detect and classify umpire poses [133]. Detecting various moments such as whether the ball hit the bat and precise detection of the player and wicket keeper at the moment of run-outs, as shown in Figure 20a, is still a

major issue in cricket. Predicting the trajectory of balls bowled by spin bowlers as shown in Figure 20b can be resolved accurately by labeling large datasets and modeling using SOTA algorithms.
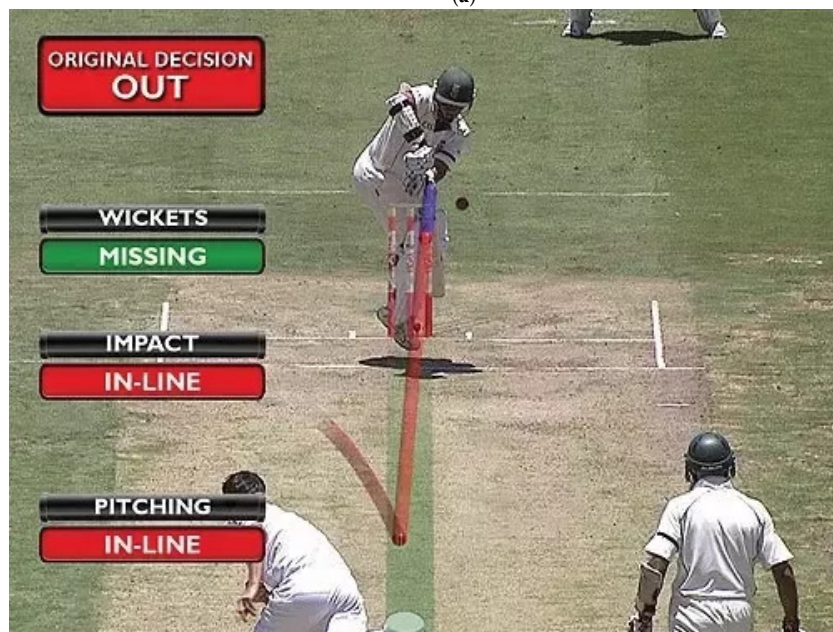

(**a**)

(**b**)

Figure 19. Instances from soccer matches. (**a**) Detecting body pose and limbs. (**b**) Handling severe occlusions among players.

The recognition accuracy of player actions in badminton games [191] can be improved by SOTA computer vision algorithms and fine-tuning in an end-to-end manner with a larger dataset on features extracted at different fully connected layers. In the implementation of an automatic linesman system in badminton games [295], the algorithm is not robust to the far views of the camera, where illumination conditions heavily impact the system while the speed of the shuttlecock is also a major factor for poor accuracy. So, it is necessary to track the path, which becomes simpler for the referee to decide if a shuttle lands out or in as shown in Figure 21.

(**a**)

(**b**)

Figure 20. Instances from cricket matches. (**a**) Precise detection at the moment of run outs. (**b**) Predicting the trajectory of the ball in-line or out-line, etc.
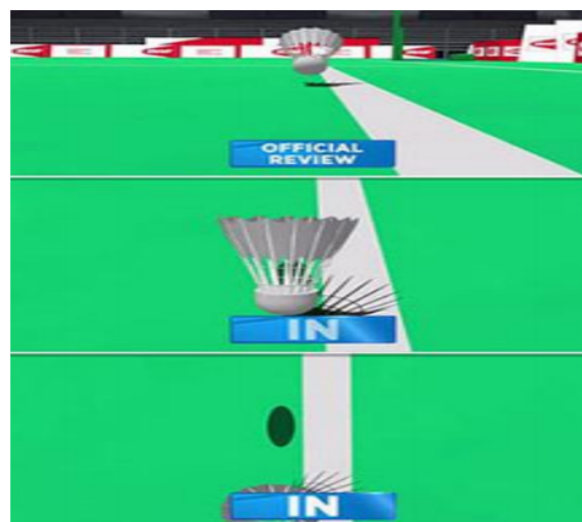

Figure 21. Exact spot on which the shuttlecock lands.

8.1. Open Issues and Future Research Areas

Computer vision plays a vital role in the area of sports video processing. To analyze sports events, there are many issues open for research. Calibration and viewpoints of the camera to capture the sports events such as close-up views, far views, and wide views in a degree of occlusions are still issues that have not been satisfactorily addressed.

Detecting the ball in various sports helps to detect and classify various ball-based events such as goals, possession of the ball, and many other events. Due to the size, speed, velocity, and unstructured motion of the ball compared to the players and playfield in various sports, it is still an open issue to detect and track the ball. Various AI algorithms have been developed to achieve better performance in various sports such as soccer [62,84], basketball [33,38], tennis [148], and badminton [195,196] in terms of detecting and tracking concerning various aspects of the ball.

Tracking players and the ball is one of the most open areas for research which includes various issues such as fast and frequent movements of the players, the similar appearance of players due to jersey color in team sports, often partial and full occlusions of players. Various algorithms use linear motions for multi-player tracking, resulting in poor performance but solves data association problems with appearance models. However, this algorithm fails in various conditions such as severe occlusions, the ambiguity of appearance between players, etc. [296,297] applied context information to track the players in soccer and volleyball sports.

8.2. Future Research Trends according to Methodologies in Sports Vision

In this section, we aim to set forth the methodological approach to various components of detection, classification, and tracking in sports. By considering the deep analysis of sports studies, it is clear that the performance of the algorithm depends on the type (annotation parameters) of dataset used, which is carried out based on loss functions and evaluation metrics. The major difficulties in the real-time use of AI algorithms in various research areas of sports are accuracy, computation speed, and size of the model. Considering all these aspects, the development of future trends based on contemporary ideas are presented below.

- Due to the continuous movements of the player, jersey numbers encounter serious deformation, and various image sizes and low resolution make it difficult to read the jersey number [205]. Players' similar appearances and severe occlusions make it difficult to track and identify players, referees, and goalkeepers reliably, which causes the critical problem of identity switching among players [212]. To solve these challenges, a framework is needed to propose that learning objects' identities with deep representations and improve tracking using identity information is necessary.

- The algorithms employed to detect and track the state of the ball such as whether it is controlled by a player (dribbling), moving on the ground (passed from one player to another player), or flying in the air to categorize the movement as a rolling pass or lobbing pass are not robust concerning the size, shape, and velocity of the ball and other parameters under different environmental conditions. A few researchers have come forward with novel ideas to deal with the above-mentioned aspects [62,68,69,107,203,210,285]; however, the state-of-the-art research is still at a nascent stage.

- Conventional architectures of detecting, classifying, and tracking are replaced with more promising and potential modern learning paradigms such as Online Learners and Extreme Learning Machines.

- Various activities in sports such as players' fatigue information can be acquired from wearable sensors and monitoring health conditions of each player in the play-field, tracking players and the ball, predicting the trajectory of the ball, game play analysis and evaluations; identifying players' actions and other fundamental elements can be accomplished with the help of big data and information technologies. This improves the sports industry's operational efficiency and leverages its immense potential. Due to advancements in big data analysis [298,299], and the Internet of Things (IoT) [300],

personalized care monitoring will become a new direction and breakthrough in the sports industry.

8.3. Different Challenges to Overcome in Sports Studies

- Classification of jersey numbers in sports such as soccer and basketball is quite simple [205] as they have plain jerseys, but in the case of the sports such as hockey and American football, the jerseys are massive and have sharp contours. Therefore, jersey number recognition is quite hard. By implementing proper bounding box techniques and digit recognition methods, better performance of jersey number recognition in every sport can be achieved.
- Action recognition in sports videos [222,223] is explicitly a non-linearity problem, which can be obtained by aligning feature vectors, by providing a massive amount of discriminative video representations. It can provide a method to capture the temporal structure of a video that is not present in the dynamic image space and analyzing salient regions of frames for action recognition.
- Provisional tactical analysis related to player formation in sports such as soccer [102], basketball, rugby, American football, and hockey, and pass prediction [89–91,98], shot prediction [81,301], the expectation for goals given a game's state [82] or possession of the ball, or more general game strategies, can be achieved through AI algorithms.
- Recognition of the fine-grained activity of typical badminton strokes can be performed by using off-the-shelf sensors [192], and it can be replaced with automatic detection and tagging of aspects/events in the game and use of CCTV-grade digital cameras without additional sensors.
- The identity of the player is lost when the player moves out of the frame and, to retain the identity when the players reappear in subsequent frames, the player must be recognized. The key challenges for player recognition are posing variations, i.e., change or rotation of the image in different poses on 2D or 3D perspectives [56,302], which is the most difficult recognition challenge, especially in case of resolution effects, variable illuminations, or lighting effects and severe occlusions.

9. Conclusions

Sports video analysis is an emerging and very dynamic field of research. This study comprehensively reviewed sports video analysis for various applications such as tracking players or balls in sports and predicting the trajectories of players or balls, players' skill, and team strategies analysis, detecting and classifying objects in sports. As per the requirements of deploying computer vision techniques in various sports, we provided some of the publicly available datasets related to a particular sport. Detailed discussion on GPU-based work stations, embedded platforms and AI applications in sports are presented. We have present various classical techniques and AI techniques employed in sports, their performance, pros, cons, and suitability to particular sports. We list probable research directions, existing challenges, and current research trends with a brief discussion and also widely used computer vision techniques in various sports.

Individual player tracking in sports is very helpful for coaches and personal trainers. Though the sports include particularly challenging tasks such as similarities between players, generation of blurry video segments in some cases, partial or full occlusions between players, the invisibility of jersey numbers in some cases, computer vision is the best possible solution to achieve player tracking.

Classification of jersey numbers in sports such as soccer and basketball is quite simple as they have plain jerseys, but in the case of sports such as hockey and American football, the jerseys are massive and come with sharp contours, due to which jersey number recognition is quite hard. By implementing proper bounding box techniques and digit recognition methods, better performance vis-a-vis jersey number recognition in every sport can be achieved. As the appearance of players varies from sport to sport, an algorithm trained on

one sport may not work when it is tested on another sport. The problem may be solved by considering a dataset that contains a small set of samples from every sport for fine-tuning.

In the case of multi-player tracking in real-time sports videos, severe occlusions cause a critical problem of identity switching among the players. The continuous movement of players makes it difficult to read jersey numbers. A player's similar appearance to another and severe occlusions make it difficult to track and identify players, referees, and goalkeepers reliably. Multiple object tracking in sports is a key prerequisite for the realization of advanced operations in sports, such as player movement and their position in sports, which will give good objective criteria to the team manager for developing a new plan to improve team performance as well as evaluating each player accurately.

Commercially used multi-camera tracking systems of players rely on some mixtures of manual and automated tracking and player labeling. Optical tracking systems are a good approach for tracking players occluding each other or players with a similar appearance. The algorithm may detect false positives from out of the court such as fans wearing team uniforms, as the appearance of fans is similar to that of players. This can be eliminated by estimating the play area or broadcast camera parameters with extra spatiotemporal locations of player positions.

Action recognition in sports videos is explicitly a non-linearity problem, which can be obtained by aligning feature vectors, by providing a massive amount discriminative video representations to capture the temporal structure of the video that is not present in the dynamic image space and analyzing the salient regions of the frames for action recognition.

The algorithms employed so far for detecting and tracking ball movements began with estimating the 3D ball position in trajectory. Employing these methods is very critical, as they include a lot of mathematical relations and require reliable reference objects to construct the path of the trajectory. Kalman filter- and particle filter-based methods are robust concerning the size, shape, and velocity of the ball. However, the methods fail to establish the track when the ball reappears after occlusion. Trajectory-based methods solve the problem of occlusion and are robust in obtaining data regarding missing and merging balls but fail to obtain data regarding the case of the size and shape of the ball. Data association methods are best suited for detecting and tracking small size balls in small courts such as tennis courts but are not suited for challenges in sports such as basketball, soccer, and volleyball. AI algorithms predict the precise trajectories of the ball from a knowledge of previous frames and are immune to challenges such as air friction, ball spin, and other complex ball movements. A precise database that includes different sizes and shapes of the ball has to be introduced to detect the ball position and enable tracking algorithms to perform efficiently.

Detection and tracking of players, balls and assistant referee as well as semantic scene understanding in computer vision applications of sports is still an open research area due to various challenges such as sudden and rapid changes in movements of the players and ball, similar appearance, players with extreme aspect ratios (players have extremely small aspect ratios in terms of height and width when they fall down on the field) and frequent occlusions. The future scope of computer vision research in sports, therefore, is handling limitations more accurately on different AI algorithms.

As the betting process involves financial assets, it is important to decide which team is likely to win; therefore, bookmakers, fans, and potential bidders are all interested in estimating the odds of the game in advance. So, provisional tactical analysis of field sports related to player formation in sports such as soccer, basketball, rugby, American football, and hockey, as well as pass prediction, shot prediction, and expectations of goals in a given game state or a possession, or more general game strategies, need to be analyzed in advance.

Tracking algorithms that are used in various sports cannot be compared on a common scale as experiments, requirements, situations, and infrastructure in every scenario differ. Determining the performance benchmark of algorithms quantitatively is quite difficult due to the unavailability of a comparable database with ground truths of different sports differing in many aspects. In addition to these, there are additional parameters such as

different video capturing devices and their parameter variations which lead to difficulty in building an object tracking system in the sports field.

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial Neural Network
AI	Artificial intelligence
AUC	Area under Curve
BEI-CNN	Basketball Energy Image - Convolutional Neural Network
Bi-LSTM	Bi-directional Long Short Term Memory
CNN	Convolutional Neural Network
CPU	Central Processing Unit
CUDA	Compute Unified Device Architecture
DELM	Deep Extreme Learning Machine
DeepMOT	Deep Multi Object Tracking
Deep-SORT	Simple Online Real Time Tracking with Deep Association
DRS	Decision Review System
ELM	Extreme Learning Machine
Faster-RCNN	Faster-Regional with Convolutional Neural Network
FPGA	Field Programmable Gate Array
GAN	Generative Adversarial Network
GDA	Gaussian Discriminant Analysis
GPU	Graphical Processing Unit
GRU-CNN	Gated Recurrent Unit - Convolutional Neural Network
GTX	Giga Texel Shader eXtreme
HOG	Histogram of Oriented Gradients
HPN	Hierarchical Policy Network
KNN	K-Nearest Neighbor
LSTM	Long Short Term Memory
Mask R-CNN	Mask Region-based Convolutional Neural Network
NBA	National Basketball Association
NHL	National Hockey League
R-CNN	Region-based Convolutional Neural Network
ResNet	Residual neural Network
RISC	Reduced Instruction Set Computer
RNN	Recurrent Neural Networks
SOTA	State-of-the-art
SSD	Single-Shot Detector

SVM	Support Vector Machine
VAR	Video Assistant Referee
VGG	Visual Geometry Group
YOLO	You Only Look Once

References

1. Tan, D.; Ting, H.; Lau, S. A review on badminton motion analysis. In Proceedings of the International Conference on Robotics, Automation and Sciences (ICORAS), Melaka, Malaysia, 5–6 November 2016; pp. 1–4.
2. Bonidia, R.P.; Rodrigues, L.A.; Avila-Santos, A.P.; Sanches, D.S.; Brancher, J.D. Computational intelligence in sports: A systematic literature review. *Adv. Hum.-Comput. Interact.* **2018**, *2018*, 3426178. [CrossRef]
3. Rahmad, N.A.; As'ari, M.A.; Ghazali, N.F.; Shahar, N.; Sufri, N.A.J. A survey of video based action recognition in sports. *Indones. J. Electr. Eng. Comput. Sci.* **2018**, *11*, 987–993. [CrossRef]
4. Van der Kruk, E.; Reijne, M.M. Accuracy of human motion capture systems for sport applications; state-of-the-art review. *Eur. J. Sport Sci.* **2018**, *18*, 806–819. [CrossRef] [PubMed]
5. Manafifard, M.; Ebadi, H.; Moghaddam, H.A. A survey on player tracking in soccer videos. *Comput. Vis. Image Underst.* **2017**, *159*, 19–46. [CrossRef]
6. Thomas, G.; Gade, R.; Moeslund, T.B.; Carr, P.; Hilton, A. Computer vision for sports: Current applications and research topics. *Comput. Vis. Image Underst.* **2017**, *159*, 3–18. [CrossRef]
7. Cust, E.E.; Sweeting, A.J.; Ball, K.; Robertson, S. Machine and deep learning for sport-specific movement recognition: A systematic review of model development and performance. *J. Sports Sci.* **2019**, *37*, 568–600. [CrossRef]
8. Kamble, P.R.; Keskar, A.G.; Bhurchandi, K.M. Ball tracking in sports: A survey. *Artif. Intell. Rev.* **2019**, *52*, 1655–1705. [CrossRef]
9. Shih, H.C. A survey of content-aware video analysis for sports. *IEEE Trans. Circuits Syst. Video Technol.* **2017**, *28*, 1212–1231. [CrossRef]
10. Beal, R.; Norman, T.J.; Ramchurn, S.D. Artificial intelligence for team sports: A survey. *Knowl. Eng. Rev.* **2019**, *34*, e28. doi: 10.1017/S0269888919000225. [CrossRef]
11. Apostolidis, E.; Adamantidou, E.; Metsai, A.I.; Mezaris, V.; Patras, I. Video Summarization Using Deep Neural Networks: A Survey. *Proc. IEEE* **2021**, *109*, 1838–1863. [CrossRef]
12. Adesida, Y.; Papi, E.; McGregor, A.H. Exploring the role of wearable technology in sport kinematics and kinetics: A systematic review. *Sensors* **2019**, *19*, 1597. [CrossRef] [PubMed]
13. Rana, M.; Mittal, V. Wearable sensors for real-time kinematics analysis in sports: A review. *IEEE Sens. J.* **2020**, *21*, 1187–1207. [CrossRef]
14. Kini, S. Real Time Moving Vehicle Congestion Detection and Tracking using OpenCV. *Turk. J. Comput. Math. Educ.* **2021**, *12*, 273–279.
15. Davis, M. Investigation into Tracking Football Players from Single Viewpoint Video Sequences. Bachelor's Thesis, The University of Bath, Bath, UK, 2008; p. 147.
16. Spagnolo, P.; Mosca, N.; Nitti, M.; Distante, A. An unsupervised approach for segmentation and clustering of soccer players. In Proceedings of the International Machine Vision and Image Processing Conference (IMVIP 2007), Washington, DC, USA, 5–7 September 2007; pp. 133–142.
17. Le Troter, A.; Mavromatis, S.; Sequeira, J. Soccer field detection in video images using color and spatial coherence. In Proceedings of the International Conference Image Analysis and Recognition, Porto, Portugal, 29 September–1 October 2004; pp. 265–272.
18. Heydari, M.; Moghadam, A.M.E. An MLP-based player detection and tracking in broadcast soccer video. In Proceedings of the International Conference of Robotics and Artificial Intelligence, Rawalpindi, Pakistan, 22–23 October 2012; pp. 195–199.
19. Barnard, M.; Odobez, J.M. Robust playfield segmentation using MAP adaptation. In Proceedings of the 17th International Conference on Pattern Recognition, Cambridge, UK, 26 August 2004; Volume 3, pp. 610–613.
20. Pallavi, V.; Mukherjee, J.; Majumdar, A.K.; Sural, S. Graph-based multiplayer detection and tracking in broadcast soccer videos. *IEEE Trans. Multimed.* **2008**, *10*, 794–805. [CrossRef]
21. Ul Huda, N.; Jensen, K.H.; Gade, R.; Moeslund, T.B. Estimating the number of soccer players using simulation-based occlusion handling. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 1824–1833.
22. Ohno, Y.; Miura, J.; Shirai, Y. Tracking players and estimation of the 3D position of a ball in soccer games. In Proceedings of the 15th International Conference on Pattern Recognition, Barcelona, Spain, 3–7 September 2000; pp. 145–148.
23. Santiago, C.B.; Sousa, A.; Reis, L.P.; Estriga, M.L. Real time colour based player tracking in indoor sports. In *Computational Vision and Medical Image Processing*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 17–35.
24. Ren, J.; Orwell, J.; Jones, G.A.; Xu, M. Tracking the soccer ball using multiple fixed cameras. *Comput. Vis. Image Underst.* **2009**, *113*, 633–642. [CrossRef]
25. Kasuya, N.; Kitahara, I.; Kameda, Y.; Ohta, Y. Real-time soccer player tracking method by utilizing shadow regions. In Proceedings of the 18th ACM international conference on Multimedia, Firenze Italy, 25–29 October 2010; pp. 1319–1322.
26. Homayounfar, N.; Fidler, S.; Urtasun, R. Sports field localization via deep structured models. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 June 2017; pp. 5212–5220.

27. Leo, M.; Mosca, N.; Spagnolo, P.; Mazzeo, P.L.; D'Orazio, T.; Distante, A. Real-time multiview analysis of soccer matches for understanding interactions between ball and players. In Proceedings of the 2008 International Conference on Content-Based Image and Video Retrieval, Niagara Falls, ON, Canada, 7–9 July 2008; pp. 525–534.

28. Conaire, C.O.; Kelly, P.; Connaghan, D.; O'Connor, N.E. Tennissense: A platform for extracting semantic information from multi-camera tennis data. In Proceedings of the 16th International Conference on Digital Signal Processing, Santorini, Greece, 5–7 July 2009; pp. 1–6.

29. Wu, L.; Yang, Z.; He, J.; Jian, M.; Xu, Y.; Xu, D.; Chen, C.W. Ontology-based global and collective motion patterns for event classification in basketball videos. *IEEE Trans. Circuits Syst. Video Technol.* **2019**, *30*, 2178–2190. [CrossRef]

30. Wu, L.; Yang, Z.; Wang, Q.; Jian, M.; Zhao, B.; Yan, J.; Chen, C.W. Fusing motion patterns and key visual information for semantic event recognition in basketball videos. *Neurocomputing* **2020**, *413*, 217–229. [CrossRef]

31. Liu, L. Objects detection toward complicated high remote basketball sports by leveraging deep CNN architecture. *Future Gener. Comput. Syst.* **2021**, *119*, 31–36. [CrossRef]

32. Fu, X.; Zhang, K.; Wang, C.; Fan, C. Multiple player tracking in basketball court videos. *J. Real-Time Image Process.* **2020**, *17*, 1811–1828. [CrossRef]

33. Yoon, Y.; Hwang, H.; Choi, Y.; Joo, M.; Oh, H.; Park, I.; Lee, K.H.; Hwang, J.H. Analyzing basketball movements and pass relationships using realtime object tracking techniques based on deep learning. *IEEE Access* **2019**, *7*, 56564–56576. [CrossRef]

34. Ramanathan, V.; Huang, J.; Abu-El-Haija, S.; Gorban, A.; Murphy, K.; Fei-Fei, L. Detecting events and key actors in multi-person videos. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 3043–3053.

35. Chakraborty, B.; Meher, S. A real-time trajectory-based ball detection-and-tracking framework for basketball video. *J. Opt.* **2013**, *42*, 156–170. [CrossRef]

36. Santhosh, P.; Kaarthick, B. An automated player detection and tracking in basketball game. *Comput. Mater. Contin.* **2019**, *58*, 625–639.

37. Acuna, D. Towards real-time detection and tracking of basketball players using deep neural networks. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017; pp. 4–9.

38. Zhao, Y.; Yang, R.; Chevalier, G.; Shah, R.C.; Romijnders, R. Applying deep bidirectional LSTM and mixture density network for basketball trajectory prediction. *Optik* **2018**, *158*, 266–272. [CrossRef]

39. Shah, R.; Romijnders, R. Applying Deep Learning to Basketball Trajectories. *arXiv* **2016**, arXiv:1608.03793.

40. Žemgulys, J.; Raudonis, V.; Maskeliūnas, R.; Damaševičius, R. Recognition of basketball referee signals from real-time videos. *J. Ambient Intell. Humaniz. Comput.* **2020**, *11*, 979–991. [CrossRef]

41. Liu, W.; Yan, C.C.; Liu, J.; Ma, H. Deep learning based basketball video analysis for intelligent arena application. *Multimed. Tools Appl.* **2017**, *76*, 24983–25001. [CrossRef]

42. Yao, P. Real-Time Analysis of Basketball Sports Data Based on Deep Learning. *Complexity* **2021**, *2021*, 9142697. doi: 10.1155/2021/9142697. [CrossRef]

43. Chen, L.; Wang, W. Analysis of technical features in basketball video based on deep learning algorithm. *Signal Process. Image Commun.* **2020**, *83*, 115786. [CrossRef]

44. Wang, K.C.; Zemel, R. Classifying NBA offensive plays using neural networks. In Proceedings of the Proceedings of MIT Sloan Sports Analytics Conference, Boston, MA, USA, 11–12 March 2016; Volume 4, pp. 1–9.

45. Tsai, T.Y.; Lin, Y.Y.; Jeng, S.K.; Liao, H.Y.M. End-to-End Key-Player-Based Group Activity Recognition Network Applied to Basketball Offensive Tactic Identification in Limited Data Scenarios. *IEEE Access* **2021**, *9*, 104395–104404. [CrossRef]

46. Lamas, L.; Junior, D.D.R.; Santana, F.; Rostaiser, E.; Negretti, L.; Ugrinowitsch, C. Space creation dynamics in basketball offence: Validation and evaluation of elite teams. *Int. J. Perform. Anal. Sport* **2011**, *11*, 71–84. [CrossRef]

47. Bourbousson, J.; Sève, C.; McGarry, T. Space–time coordination dynamics in basketball: Part 1. Intra-and inter-couplings among player dyads. *J. Sports Sci.* **2010**, *28*, 339–347. [CrossRef] [PubMed]

48. Bourbousson, J.; Seve, C.; McGarry, T. Space–time coordination dynamics in basketball: Part 2. The interaction between the two teams. *J. Sports Sci.* **2010**, *28*, 349–358. [CrossRef] [PubMed]

49. Tian, C.; De Silva, V.; Caine, M.; Swanson, S. Use of machine learning to automate the identification of basketball strategies using whole team player tracking data. *Appl. Sci.* **2020**, *10*, 24. [CrossRef]

50. Hauri, S.; Djuric, N.; Radosavljevic, V.; Vucetic, S. Multi-Modal Trajectory Prediction of NBA Players. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 3–8 January 2021; pp. 1640–1649.

51. Zheng, S.; Yue, Y.; Lucey, P. Generating Long-Term Trajectories Using Deep Hierarchical Networks. In Proceedings of the 30th International Conference on Neural Information Processing Systems, Barcelona, Spain, 5–10 December 2016; pp. 1551–1559.

52. Bertugli, A.; Calderara, S.; Coscia, P.; Ballan, L.; Cucchiara, R. AC-VRNN: Attentive Conditional-VRNN for multi-future trajectory prediction. *Comput. Vis. Image Underst.* **2021**, *210*, 103245. [CrossRef]

53. Victor, B.; Nibali, A.; He, Z.; Carey, D.L. Enhancing trajectory prediction using sparse outputs: Application to team sports. *Neural Comput. Appl.* **2021**, *33*, 11951–11962. [CrossRef]

54. Li, H.; Zhang, M. Artificial Intelligence and Neural Network-Based Shooting Accuracy Prediction Analysis in Basketball. *Mob. Inf. Syst.* **2021**, *2021*, 4485589. [CrossRef]

55. Chen, H.T.; Chou, C.L.; Fu, T.S.; Lee, S.Y.; Lin, B.S.P. Recognizing tactic patterns in broadcast basketball video using player trajectory. *J. Vis. Commun. Image Represent.* **2012**, *23*, 932–947. [CrossRef]

56. Chen, H.T.; Tien, M.C.; Chen, Y.W.; Tsai, W.J.; Lee, S.Y. Physics-based ball tracking and 3D trajectory reconstruction with applications to shooting location estimation in basketball video. *J. Vis. Commun. Image Represent.* **2009**, *20*, 204–216. [CrossRef]

57. Hu, M.; Hu, Q. Design of basketball game image acquisition and processing system based on machine vision and image processor. *Microprocess. Microsyst.* **2021**, *82*, 103904. [CrossRef]

58. Yichen, W.; Yamashita, H. Lineup Optimization Model of Basketball Players Based on the Prediction of Recursive Neural Networks. *Int. J. Econ. Manag. Eng.* **2021**, *15*, 283–289.

59. Suda, S.; Makino, Y.; Shinoda, H. Prediction of volleyball trajectory using skeletal motions of setter player. In Proceedings of the 10th Augmented Human International Conference, Reims, France, 11–12 March 2019; pp. 1–8.

60. Gerke, S.; Linnemann, A.; Müller, K. Soccer player recognition using spatial constellation features and jersey number recognition. *Comput. Vis. Image Underst.* **2017**, *159*, 105–115. [CrossRef]

61. Baysal, S.; Duygulu, P. Sentioscope: A soccer player tracking system using model field particles. *IEEE Trans. Circuits Syst. Video Technol.* **2015**, *26*, 1350–1362. [CrossRef]

62. Kamble, P.; Keskar, A.; Bhurchandi, K. A deep learning ball tracking system in soccer videos. *Opto-Electron. Rev.* **2019**, *27*, 58–69. [CrossRef]

63. Choi, K.; Seo, Y. Automatic initialization for 3D soccer player tracking. *Pattern Recognit. Lett.* **2011**, *32*, 1274–1282. [CrossRef]

64. Kim, W. Multiple object tracking in soccer videos using topographic surface analysis. *J. Vis. Commun. Image Represent.* **2019**, *65*, 102683. [CrossRef]

65. Liu, J.; Tong, X.; Li, W.; Wang, T.; Zhang, Y.; Wang, H. Automatic player detection, labeling and tracking in broadcast soccer video. *Pattern Recognit. Lett.* **2009**, *30*, 103–113. [CrossRef]

66. Komorowski, J.; Kurzejamski, G.; Sarwas, G. BallTrack: Football ball tracking for real-time CCTV systems. In Proceedings of the 16th International Conference on Machine Vision Applications (MVA), Tokyo, Japan, 27–31 May 2019; pp. 1–5.

67. Hurault, S.; Ballester, C.; Haro, G. Self-Supervised Small Soccer Player Detection and Tracking. In Proceedings of the 3rd International Workshop on Multimedia Content Analysis in Sports, Seattle, WA, USA, 12–16 October 2020; pp. 9–18.

68. Kamble, P.R.; Keskar, A.G.; Bhurchandi, K.M. A convolutional neural network based 3D ball tracking by detection in soccer videos. In Proceedings of the Eleventh International Conference on machine vision (ICMV 2018), Munich, Germany, 1–3 November 2018; Volume 11041, p. 110412O.

69. Naidoo, W.C.; Tapamo, J.R. Soccer video analysis by ball, player and referee tracking. In Proceedings of the 2006 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on IT Research in Developing Countries, Somerset West, South Africa, 9–11 October 2006; pp. 51–60.

70. Liang, D.; Liu, Y.; Huang, Q.; Gao, W. A scheme for ball detection and tracking in broadcast soccer video. In Proceedings of the Pacific-Rim Conference on Multimedia, Jeju Island, Korea, 13–16 November 2005; pp. 864–875.

71. Naik, B.; Hashmi, M.F. YOLOv3-SORT detection and tracking player-ball in soccer sport. *J. Electron. Imaging* **2023**, *32*, 011003. [CrossRef]

72. Naik, B.; Hashmi, M.F.; Geem, Z.W.; Bokde, N.D. DeepPlayer-Track: Player and Referee Tracking with Jersey Color Recognition in Soccer. *IEEE Access* **2022**, *10*, 32494–32509. [CrossRef]

73. Komorowski, J.; Kurzejamski, G.; Sarwas, G. FootAndBall: Integrated Player and Ball Detector. In Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, Valletta, Malta, 27–29 February 2020; Volume 5, pp. 47–56. [CrossRef]

74. Pallavi, V.; Mukherjee, J.; Majumdar, A.K.; Sural, S. Ball detection from broadcast soccer videos using static and dynamic features. *J. Vis. Commun. Image Represent.* **2008**, *19*, 426–436. [CrossRef]

75. Leo, M.; Mazzeo, P.L.; Nitti, M.; Spagnolo, P. Accurate ball detection in soccer images using probabilistic analysis of salient regions. *Mach. Vis. Appl.* **2013**, *24*, 1561–1574. [CrossRef]

76. Mazzeo, P.L.; Leo, M.; Spagnolo, P.; Nitti, M. Soccer ball detection by comparing different feature extraction methodologies. *Adv. Artif. Intell.* **2012**, *2012*, 512159. [CrossRef]

77. Garnier, P.; Gregoir, T. Evaluating Soccer Player: From Live Camera to Deep Reinforcement Learning. *arXiv* **2021**, arXiv:2101.05388.

78. Kusmakar, S.; Shelyag, S.; Zhu, Y.; Dwyer, D.; Gastin, P.; Angelova, M. Machine Learning Enabled Team Performance Analysis in the Dynamical Environment of Soccer. *IEEE Access* **2020**, *8*, 90266–90279. [CrossRef]

79. Baccouche, M.; Mamalet, F.; Wolf, C.; Garcia, C.; Baskurt, A. Action classification in soccer videos with long short-term memory recurrent neural networks. In Proceedings of the International Conference on Artificial Neural Networks, Thessaloniki, Greece, 15–18 September 2010; pp. 154–159.

80. Jackman, S. Football Shot Detection Using Convolutional Neural Networks. Master's Thesis, Department of Biomedical Engineering, Linköping University, Linköping, Sweden, 2019.

81. Lucey, P.; Bialkowski, A.; Monfort, M.; Carr, P.; Matthews, I. quality vs quantity: Improved shot prediction in soccer using strategic features from spatiotemporal data. In Proceedings of the 8th Annual MIT Sloan Sports Analytics Conference, Boston, MA, USA, 28 February–1 March 2014; pp. 1–9.

82. Cioppa, A.; Deliege, A.; Giancola, S.; Ghanem, B.; Droogenbroeck, M.V.; Gade, R.; Moeslund, T.B. A context-aware loss function for action spotting in soccer videos. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 13126–13136.

83. Beernaerts, J.; De Baets, B.; Lenoir, M.; Van de Weghe, N. Spatial movement pattern recognition in soccer based on relative player movements. *PLoS ONE* **2020**, *15*, e0227746. [CrossRef] [PubMed]

84. Barbon Junior, S.; Pinto, A.; Barroso, J.V.; Caetano, F.G.; Moura, F.A.; Cunha, S.A.; Torres, R.d.S. Sport action mining: Dribbling recognition in soccer. *Multimed. Tools Appl.* **2022**, *81*, 4341–4364. [CrossRef]

85. Kim, Y.; Jung, C.; Kim, C. Motion Recognition of Assistant Referees in Soccer Games via Selective Color Contrast Revelation. EasyChair Preprint no. 2604, EasyChair, 2020. Available online: https://easychair.org/publications/preprint/z975 (accessed on 2 November 2021).

86. Lindström, P.; Jacobsson, L.; Carlsson, N.; Lambrix, P. Predicting player trajectories in shot situations in soccer. In Proceedings of the International Workshop on Machine Learning and Data Mining for Sports Analytics, Ghent, Belgium, 14–18 September 2020; pp. 62–75.

87. Machado, V.; Leite, R.; Moura, F.; Cunha, S.; Sadlo, F.; Comba, J.L. Visual soccer match analysis using spatiotemporal positions of players. *Comput. Graph.* **2017**, *68*, 84–95. [CrossRef]

88. Ganesh, Y.; Teja, A.S.; Munnangi, S.K.; Murthy, G.R. A Novel Framework for Fine Grained Action Recognition in Soccer. In Proceedings of the International Work-Conference on Artificial Neural Networks, Munich, Germany, 17–19 September 2019; pp. 137–150.

89. Chawla, S.; Estephan, J.; Gudmundsson, J.; Horton, M. Classification of passes in football matches using spatiotemporal data. *ACM Trans. Spat. Algorithms Syst.* **2017**, *3*, 1–30. [CrossRef]

90. Gyarmati, L.; Stanojevic, R. QPass: A Merit-based Evaluation of Soccer Passes. *arXiv* **2016**, arXiv:abs/1608.03532.

91. Vercruyssen, V.; De Raedt, L.; Davis, J. Qualitative spatial reasoning for soccer pass prediction. In *CEUR Workshop Proceedings*; Springer: Berlin/Heidelberg, Germany, 2016; Volume 1842.

92. Yu, J.; Lei, A.; Hu, Y. Soccer video event detection based on deep learning. In Proceedings of the International Conference on Multimedia Modeling, Thessaloniki, Greece, 8–11 January 2019; pp. 377–389.

93. Brooks, J.; Kerr, M.; Guttag, J. Using machine learning to draw inferences from pass location data in soccer. *Stat. Anal. Data Min. ASA Data Sci. J.* **2016**, *9*, 338–349. [CrossRef]

94. Cho, H.; Ryu, H.; Song, M. Pass2vec: Analyzing soccer players' passing style using deep learning. *Int. J. Sports Sci. Coach.* **2021**, *17*, 355–365. [CrossRef]

95. Zhang, K.; Wu, J.; Tong, X.; Wang, Y. An automatic multi-camera-based event extraction system for real soccer videos. *Pattern Anal. Appl.* **2020**, *23*, 953–965. [CrossRef]

96. Deliège, A.; Cioppa, A.; Giancola, S.; Seikavandi, M.J.; Dueholm, J.V.; Nasrollahi, K.; Ghanem, B.; Moeslund, T.B.; Droogenbroeck, M.V. SoccerNet-v2: A Dataset and Benchmarks for Holistic Understanding of Broadcast Soccer Videos. *arXiv* **2020**, arXiv:abs/2011.13367.

97. Penumala, R.; Sivagami, M.; Srinivasan, S. Automated Goal Score Detection in Football Match Using Key Moments. *Procedia Comput. Sci.* **2019**, *165*, 492–501. [CrossRef]

98. Khan, A.; Lazzerini, B.; Calabrese, G.; Serafini, L. Soccer event detection. In Proceedings of the 4th International Conference on Image Processing and Pattern Recognition (IPPR 2018), Copenhagen, Denmark, 28–29 April 2018; pp. 119–129.

99. Khaustov, V.; Mozgovoy, M. Recognizing Events in Spatiotemporal Soccer Data. *Appl. Sci.* **2020**, *10*, 8046. [CrossRef]

100. Saraogi, H.; Sharma, R.A.; Kumar, V. Event recognition in broadcast soccer videos. In Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing, Hyderabad, India, 18–22 December 2016; pp. 1–7.

101. Karimi, A.; Toosi, R.; Akhaee, M.A. Soccer Event Detection Using Deep Learning. *arXiv* **2021**, arXiv:2102.04331.

102. Suzuki, G.; Takahashi, S.; Ogawa, T.; Haseyama, M. Team tactics estimation in soccer videos based on a deep extreme learning machine and characteristics of the tactics. *IEEE Access* **2019**, *7*, 153238–153248. [CrossRef]

103. Suzuki, G.; Takahashi, S.; Ogawa, T.; Haseyama, M. Decision level fusion-based team tactics estimation in soccer videos. In Proceedings of the IEEE 5th Global Conference on Consumer Electronics, Kyoto, Japan, 11–14 October 2016; pp. 1–2.

104. Ohnuki, S.; Takahashi, S.; Ogawa, T.; Haseyama, M. Soccer video segmentation based on team tactics estimation method. In Proceedings of the International Workshop on Advanced Image Technology, Nagoya, Japan, 7–8 January 2013; pp. 692–695.

105. Clemente, F.M.; Couceiro, M.S.; Martins, F.M.L.; Mendes, R.S.; Figueiredo, A.J. Soccer team's tactical behaviour: Measuring territorial domain. *J. Sports Eng. Technol.* **2015**, *229*, 58–66. [CrossRef]

106. Hassan, A.; Akl, A.R.; Hassan, I.; Sunderland, C. Predicting Wins, Losses and Attributes' Sensitivities in the Soccer World Cup 2018 Using Neural Network Analysis. *Sensors* **2020**, *20*, 3213. [CrossRef]

107. Niu, Z.; Gao, X.; Tian, Q. Tactic analysis based on real-world ball trajectory in soccer video. *Pattern Recognit.* **2012**, *45*, 1937–1947. [CrossRef]

108. Wu, Y.; Xie, X.; Wang, J.; Deng, D.; Liang, H.; Zhang, H.; Cheng, S.; Chen, W. Forvizor: Visualizing spatio-temporal team formations in soccer. *IEEE Trans. Vis. Comput. Graph.* **2018**, *25*, 65–75. [CrossRef]

109. Suzuki, G.; Takahashi, S.; Ogawa, T.; Haseyama, M. Team tactics estimation in soccer videos via deep extreme learning machine based on players formation. In Proceedings of the IEEE 7th Global Conference on Consumer Electronics, Nara, Japan, 9–12 October 2018; pp. 116–117.

110. Wang, B.; Shen, W.; Chen, F.; Zeng, D. Football match intelligent editing system based on deep learning. *KSII Trans. Internet Inf. Syst.* **2019**, *13*, 5130–5143.

111. Zawbaa, H.M.; El-Bendary, N.; Hassanien, A.E.; Kim, T.h. Event detection based approach for soccer video summarization using machine learning. *Int. J. Multimed. Ubiquitous Eng.* **2012**, *7*, 63–80.

112. Kolekar, M.H.; Sengupta, S. Bayesian network-based customized highlight generation for broadcast soccer videos. *IEEE Trans. Broadcast.* **2015**, *61*, 195–209. [CrossRef]

113. Li, J.; Wang, T.; Hu, W.; Sun, M.; Zhang, Y. Soccer highlight detection using two-dependence bayesian network. In Proceedings of the IEEE International Conference on Multimedia and Expo, Toronto, ON, Canada, 9–12 July 2006; pp. 1625–1628.

114. Foysal, M.F.A.; Islam, M.S.; Karim, A.; Neehal, N. Shot-Net: A convolutional neural network for classifying different cricket shots. In Proceedings of the International Conference on Recent Trends in Image Processing and Pattern Recognition, Solapur, India, 21–22 December 2018; pp. 111–120.

115. Khan, M.Z.; Hassan, M.A.; Farooq, A.; Khan, M.U.G. Deep CNN based data-driven recognition of cricket batting shots. In Proceedings of the International Conference on Applied and Engineering Mathematics (ICAEM), Taxila, Pakistan, 4–5 September 2018; pp. 67–71.

116. Khan, A.; Nicholson, J.; Plötz, T. Activity recognition for quality assessment of batting shots in cricket using a hierarchical representation. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*; ACM Digital Library: New York, NY, USA, 2017; Volume 1, p. 62. [CrossRef]

117. Sen, A.; Deb, K.; Dhar, P.K.; Koshiba, T. CricShotClassify: An Approach to Classifying Batting Shots from Cricket Videos Using a Convolutional Neural Network and Gated Recurrent Unit. *Sensors* **2021**, *21*, 2846. [CrossRef] [PubMed]

118. Gürpınar-Morgan, W.; Dinsdale, D.; Gallagher, J.; Cherukumudi, A.; Lucey, P. You Cannot Do That Ben Stokes: Dynamically Predicting Shot Type in Cricket Using a Personalized Deep Neural Network. *arXiv* **2021**, arXiv:2102.01952.

119. Bandara, I.; Bačić, B. Strokes Classification in Cricket Batting Videos. In Proceedings of the 2020 5th International Conference on Innovative Technologies in Intelligent Systems and Industrial Applications (CITISIA), Sydney, Australia, 25–27 November 2020; pp. 1–6.

120. Moodley, T.; van der Haar, D. Scene Recognition Using AlexNet to Recognize Significant Events Within Cricket Game Footage. In Proceedings of the International Conference on Computer Vision and Graphics, Valletta, Malta, 27–29 February 2020; pp. 98–109.

121. Gupta, A.; Muthiah, S.B. Viewpoint constrained and unconstrained Cricket stroke localization from untrimmed videos. *Image Vis. Comput.* **2020**, *100*, 103944. [CrossRef]

122. Al Islam, M.N.; Hassan, T.B.; Khan, S.K. A CNN-based approach to classify cricket bowlers based on their bowling actions. In Proceedings of the IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON), Dhaka, Bangladesh, 28–30 November 2019, pp. 130–134.

123. Muthuswamy, S.; Lam, S.S. Bowler performance prediction for one-day international cricket using neural networks. In Proceedings of the IIE Annual Conference Proceedings. Institute of Industrial and Systems Engineers (IISE), New Orleans, LA, USA, 30 May–2 June 2008, p. 1391.

124. Bhattacharjee, D.; Pahinkar, D.G. Analysis of performance of bowlers using combined bowling rate. *Int. J. Sports Sci. Eng.* **2012**, *6*, 1750–9823.

125. Rahman, R.; Rahman, M.A.; Islam, M.S.; Hasan, M. DeepGrip: Cricket Bowling Delivery Detection with Superior CNN Architectures. In Proceedings of the 6th International Conference on Inventive Computation Technologies (ICICT), Lalitpur, Nepal, 20–22 July 2021; pp. 630–636.

126. Lemmer, H.H. The combined bowling rate as a measure of bowling performance in cricket. *S. Afr. J. Res. Sport Phys. Educ. Recreat.* **2002**, *24*, 37–44. [CrossRef]

127. Mukherjee, S. Quantifying individual performance in Cricket—A network analysis of Batsmen and Bowlers. *Phys. A Stat. Mech. Its Appl.* **2014**, *393*, 624–637. [CrossRef]

128. Velammal, B.; Kumar, P.A. An Efficient Ball Detection Framework for Cricket. *Int. J. Comput. Sci. Issues* **2010**, *7*, 30.

129. Nelikanti, A.; Reddy, G.V.R.; Karuna, G. An Optimization Based deep LSTM Predictive Analysis for Decision Making in Cricket. In *Innovative Data Communication Technologies and Application*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 721–737.

130. Kumar, R.; Santhadevi, D.; Barnabas, J. Outcome Classification in Cricket Using Deep Learning. In Proceedings of the IEEE International Conference on Cloud Computing in Emerging Markets (CCEM), Bengaluru, India, 19–20 September 2019; pp. 55–58.

131. Shukla, P.; Sadana, H.; Bansal, A.; Verma, D.; Elmadjian, C.; Raman, B.; Turk, M. Automatic cricket highlight generation using event-driven and excitement-based features. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 1800–1808.

132. Kowsher, M.; Alam, M.A.; Uddin, M.J.; Ahmed, F.; Ullah, M.W.; Islam, M.R. Detecting Third Umpire Decisions & Automated Scoring System of Cricket. In Proceedings of the 2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2), Rajshahi, Bangladesh, 11–12 July 2019; pp. 1–8.

133. Ravi, A.; Venugopal, H.; Paul, S.; Tizhoosh, H.R. A dataset and preliminary results for umpire pose detection using SVM classification of deep features. In Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence (SSCI), Bangalore, India, 18–21 November 2018; pp. 1396–1402.

134. Kapadiya, C.; Shah, A.; Adhvaryu, K.; Barot, P. Intelligent Cricket Team Selection by Predicting Individual Players' Performance using Efficient Machine Learning Technique. *Int. J. Eng. Adv. Technol.* **2020**, *9*, 3406–3409. [CrossRef]

135. Iyer, S.R.; Sharda, R. Prediction of athletes performance using neural networks: An application in cricket team selection. *Expert Syst. Appl.* **2009**, *36*, 5510–5522. [CrossRef]

136. Jhanwar, M.G.; Pudi, V. Predicting the Outcome of ODI Cricket Matches: A Team Composition Based Approach. In Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECMLPKDD 2016), Bilbao, Spain, 19–23 September 2016.

137. Pathak, N.; Wadhwa, H. Applications of modern classification techniques to predict the outcome of ODI cricket. *Procedia Comput. Sci.* **2016**, *87*, 55–60. [CrossRef]

138. Alaka, S.; Sreekumar, R.; Shalu, H. Efficient Feature Representations for Cricket Data Analysis using Deep Learning based Multi-Modal Fusion Model. *arXiv* **2021**, arXiv:2108.07139.

139. Goel, R.; Davis, J.; Bhatia, A.; Malhotra, P.; Bhardwaj, H.; Hooda, V.; Goel, A. Dynamic cricket match outcome prediction. *J. Sports Anal.* **2021**, *7*, 185–196. [CrossRef]

140. Karthik, K.; Krishnan, G.S.; Shetty, S.; Bankapur, S.S.; Kolkar, R.P.; Ashwin, T.; Vanahalli, M.K. Analysis and Prediction of Fantasy Cricket Contest Winners Using Machine Learning Techniques. In *Evolution in Computational Intelligence*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 443–453.

141. Shah, P. New performance measure in Cricket. *ISOR J. Sports Phys. Educ.* **2017**, *4*, 28–30. [CrossRef]

142. Shingrakhia, H.; Patel, H. SGRNN-AM and HRF-DBN: A hybrid machine learning model for cricket video summarization. *Vis. Comput.* **2021**. [CrossRef]

143. Guntuboina, C.; Porwal, A.; Jain, P.; Shingrakhia, H. Deep Learning Based Automated Sports Video Summarization using YOLO. *Electron. Lett. Comput. Vis. Image Anal.* **2021**, *20*, 99–116.

144. Owens, N.; Harris, C.; Stennett, C. Hawk-eye tennis system. In Proceedings of the International Conference on Visual Information Engineering, Guildford, UK, 7–9 July 2003; pp. 182–185.

145. Wu, G. Monitoring System of Key Technical Features of Male Tennis Players Based on Internet of Things Security Technology. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 4076863. [CrossRef]

146. Connaghan, D.; Kelly, P.; O'Connor, N.E. Game, shot and match: Event-based indexing of tennis. In Proceedings of the 9th International Workshop on Content-Based Multimedia Indexing (CBMI), Lille, France, 28–30 June 2011; pp. 97–102.

147. Giles, B.; Kovalchik, S.; Reid, M. A machine learning approach for automatic detection and classification of changes of direction from player tracking data in professional tennis. *J. Sports Sci.* **2020**, *38*, 106–113. [CrossRef]

148. Zhou, X.; Xie, L.; Huang, Q.; Cox, S.J.; Zhang, Y. Tennis ball tracking using a two-layered data association approach. *IEEE Trans. Multimed.* **2014**, *17*, 145–156. [CrossRef]

149. Reno, V.; Mosca, N.; Marani, R.; Nitti, M.; D'Orazio, T.; Stella, E. Convolutional neural networks based ball detection in tennis games. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 1758–1764.

150. Archana, M.; Geetha, M.K. Object detection and tracking based on trajectory in broadcast tennis video. *Procedia Comput. Sci.* **2015**, *58*, 225–232. [CrossRef]

151. Polk, T.; Yang, J.; Hu, Y.; Zhao, Y. Tennivis: Visualization for tennis match analysis. *IEEE Trans. Vis. Comput. Graph.* **2014**, *20*, 2339–2348. [CrossRef]

152. Kelly, P.; Diego, J.; Agapito, P.; Conaire, C.; Connaghan, D.; Kuklyte, J.; Connor, N. Performance analysis and visualisation in tennis using a low-cost camera network. In Proceedings of the 18th ACM Multimedia Conference on Multimedia Grand Challenge, Beijing, China, 25–29 October 2010; pp. 1–4.

153. Fernando, T.; Denman, S.; Sridharan, S.; Fookes, C. Memory augmented deep generative models for forecasting the next shot location in tennis. *IEEE Trans. Knowl. Data Eng.* **2019**, *32*, 1785–1797. [CrossRef]

154. Pingali, G.; Opalach, A.; Jean, Y.; Carlbom, I. Visualization of sports using motion trajectories: Providing insights into performance, style, and strategy. In Proceedings of the IEEE Visualization 2001, San Diego, CA, USA, 24–26 October 2001; pp. 75–544.

155. Pingali, G.S.; Opalach, A.; Jean, Y.D.; Carlbom, I.B. Instantly indexed multimedia databases of real world events. *IEEE Trans. Multimed.* **2002**, *4*, 269–282. [CrossRef]

156. Cai, J.; Hu, J.; Tang, X.; Hung, T.Y.; Tan, Y.P. Deep historical long short-term memory network for action recognition. *Neurocomputing* **2020**, *407*, 428–438. [CrossRef]

157. Vinyes Mora, S.; Knottenbelt, W.J. Deep learning for domain-specific action recognition in tennis. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 114–122.

158. Ning, B.; Na, L. Deep Spatial/temporal-level feature engineering for Tennis-based action recognition. *Future Gener. Comput. Syst.* **2021**, *125*, 188–193. [CrossRef]

159. Polk, T.; Jäckle, D.; Häußler, J.; Yang, J. CourtTime: Generating actionable insights into tennis matches using visual analytics. *IEEE Trans. Vis. Comput. Graph.* **2019**, *26*, 397–406. [CrossRef]

160. Zhu, G.; Huang, Q.; Xu, C.; Xing, L.; Gao, W.; Yao, H. Human behavior analysis for highlight ranking in broadcast racket sports video. *IEEE Trans. Multimed.* **2007**, *9*, 1167–1182.

161. Wei, X.; Lucey, P.; Morgan, S.; Sridharan, S. Forecasting the next shot location in tennis using fine-grained spatiotemporal tracking data. *IEEE Trans. Knowl. Data Eng.* **2016**, *28*, 2988–2997. [CrossRef]

162. Ma, K. A Real Time Artificial Intelligent System for Tennis Swing Classification. In Proceedings of the IEEE 19th World Symposium on Applied Machine Intelligence and Informatics (SAMI), Herl'any, Slovakia, 21–23 January 2021; pp. 21–26.

163. Vales-Alonso, J.; Chaves-Diéguez, D.; López-Matencio, P.; Alcaraz, J.J.; Parrado-García, F.J.; González-Castaño, F.J. SAETA: A smart coaching assistant for professional volleyball training. *IEEE Trans. Syst. Man Cybern. Syst.* **2015**, *45*, 1138–1150. [CrossRef]

164. Kautz, T.; Groh, B.H.; Hannink, J.; Jensen, U.; Strubberg, H.; Eskofier, B.M. Activity recognition in beach volleyball using a Deep Convolutional Neural Network. *Data Min. Knowl. Discov.* **2017**, *31*, 1678–1705. [CrossRef]

165. Ibrahim, M.S.; Muralidharan, S.; Deng, Z.; Vahdat, A.; Mori, G. A hierarchical deep temporal model for group activity recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1971–1980.

166. Van Haaren, J.; Ben Shitrit, H.; Davis, J.; Fua, P. Analyzing volleyball match data from the 2014 World Championships using machine learning techniques. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 627–634.

167. Wenninger, S.; Link, D.; Lames, M. Performance of machine learning models in application to beach volleyball data. *Int. J. Comput. Sci. Sport* **2020**, *19*, 24–36. [CrossRef]

168. Haider, F.; Salim, F.; Naghashi, V.; Tasdemir, S.B.Y.; Tengiz, I.; Cengiz, K.; Postma, D.; Delden, R.v.; Reidsma, D.; van Beijnum, B.J.; et al. Evaluation of dominant and non-dominant hand movements for volleyball action modelling. In Proceedings of the Adjunct of the 2019 International Conference on Multimodal Interaction, Suzhou, China, 14–18 October 2019; pp. 1–6.

169. Salim, F.A.; Haider, F.; Tasdemir, S.B.Y.; Naghashi, V.; Tengiz, I.; Cengiz, K.; Postma, D.; Van Delden, R. Volleyball action modelling for behavior analysis and interactive multi-modal feedback. In Proceedings of the 15th International Summer Workshop on Multimodal Interfaces, Ankara, Turkey, 8 July 2019; p. 50.

170. Jiang, W.; Zhao, K.; Jin, X. Diagnosis Model of Volleyball Skills and Tactics Based on Artificial Neural Network. *Mob. Inf. Syst.* **2021**, *2021*, 7908897. [CrossRef]

171. Wang, Y.; Zhao, Y.; Chan, R.H.; Li, W.J. Volleyball skill assessment using a single wearable micro inertial measurement unit at wrist. *IEEE Access* **2018**, *6*, 13758–13765. [CrossRef]

172. Zhang, C.; Tang, H.; Duan, Z. WITHDRAWN: Time Series Analysis of Volleyball Spiking Posture Based on Quality-Guided Cyclic Neural Network. *J. Vis. Commun. Image Represent.* **2019**, *82*, 102681. [CrossRef]

173. Thilakarathne, H.; Nibali, A.; He, Z.; Morgan, S. Pose is all you need: The pose only group activity recognition system (POGARS). *arXiv* **2021**, arXiv:2108.04186.

174. Zhao, K.; Jiang, W.; Jin, X.; Xiao, X. Artificial intelligence system based on the layout effect of both sides in volleyball matches. *J. Intell. Fuzzy Syst.* **2021**, *40*, 3075–3084. [CrossRef]

175. Tian, Y. Optimization of Volleyball Motion Estimation Algorithm Based on Machine Vision and Wearable Devices. *Microprocess. Microsyst.* **2020**, *81*, 103750. [CrossRef]

176. Şah, M.; Direkoğlu, C. Review and evaluation of player detection methods in field sports. *Multimed. Tools Appl.* **2021**. [CrossRef]

177. Rangasamy, K.; As'ari, M.A.; Rahmad, N.A.; Ghazali, N.F. Hockey activity recognition using pre-trained deep learning model. *ICT Express* **2020**, *6*, 170–174. [CrossRef]

178. Sozykin, K.; Protasov, S.; Khan, A.; Hussain, R.; Lee, J. Multi-label class-imbalanced action recognition in hockey videos via 3D convolutional neural networks. In Proceedings of the 19th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), Busan, Korea, 27–29 June 2018; pp. 146–151.

179. Fani, M.; Neher, H.; Clausi, D.A.; Wong, A.; Zelek, J. Hockey action recognition via integrated stacked hourglass network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 29–37.

180. Cai, Z.; Neher, H.; Vats, K.; Clausi, D.A.; Zelek, J. Temporal hockey action recognition via pose and optical flows. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA, 16–17 June 2019.

181. Chan, A.; Levine, M.D.; Javan, M. Player Identification in Hockey Broadcast Videos. *Expert Syst. Appl.* **2021**, *165*, 113891. [CrossRef]

182. Carbonneau, M.A.; Raymond, A.J.; Granger, E.; Gagnon, G. Real-time visual play-break detection in sport events using a context descriptor. In Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Lisbon, Portugal, 24–27 May 2015; pp. 2808–2811.

183. Wang, H.; Ullah, M.M.; Klaser, A.; Laptev, I.; Schmid, C. Evaluation of local spatio-temporal features for action recognition. In Proceedings of the British Machine Vision Conference, London, UK, 7–10 September 2009.

184. Um, G.M.; Lee, C.; Park, S.; Seo, J. Ice Hockey Player Tracking and Identification System Using Multi-camera video. In Proceedings of the IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Jeju, Korea, 5–7 June 2019; pp. 1–4.

185. Guo, T.; Tao, K.; Hu, Q.; Shen, Y. Detection of Ice Hockey Players and Teams via a Two-Phase Cascaded CNN Model. *IEEE Access* **2020**, *8*, 195062–195073. [CrossRef]

186. Liu, G.; Schulte, O. Deep reinforcement learning in ice hockey for context-aware player evaluation. *arXiv* **2021**, arXiv:1805.11088.

187. Vats, K.; Neher, H.; Clausi, D.A.; Zelek, J. Two-stream action recognition in ice hockey using player pose sequences and optical flows. In Proceedings of the 16th Conference on Computer and Robot Vision (CRV), Kingston, QC, Canada, 29–31 May 2019; pp. 181–188.

188. Vats, K.; Fani, M.; Clausi, D.A.; Zelek, J. Puck localization and multi-task event recognition in broadcast hockey videos. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 4567–4575.

189. Tora, M.R.; Chen, J.; Little, J.J. Classification of puck possession events in ice hockey. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 22–25 July 2017; pp. 147–154.

190. Weeratunga, K.; Dharmaratne, A.; Boon How, K. Application of computer vision and vector space model for tactical movement classification in badminton. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 June 2017; pp. 76–82.

191. Rahmad, N.; As'ari, M. The new Convolutional Neural Network (CNN) local feature extractor for automated badminton action recognition on vision based data. *J. Phys. Conf. Ser.* **2020**, *1529*, 022021. [CrossRef]

192. Steels, T.; Van Herbruggen, B.; Fontaine, J.; De Pessemier, T.; Plets, D.; De Poorter, E. Badminton Activity Recognition Using Accelerometer Data. *Sensors* **2020**, *20*, 4685. [CrossRef]

193. binti Rahmad, N.A.; binti Sufri, N.A.J.; bin As'ari, M.A.; binti Azaman, A. Recognition of Badminton Action Using Convolutional Neural Network. *Indones. J. Electr. Eng. Inform.* **2019**, *7*, 750–756.

194. Ghosh, I.; Ramamurthy, S.R.; Roy, N. StanceScorer: A Data Driven Approach to Score Badminton Player. In Proceedings of the IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Austin, TX, USA, 13–20 September 2020; pp. 1–6.

195. Cao, Z.; Liao, T.; Song, W.; Chen, Z.; Li, C. Detecting the shuttlecock for a badminton robot: A YOLO based approach. *Expert Syst. Appl.* **2021**, *164*, 113833. [CrossRef]

196. Chen, W.; Liao, T.; Li, Z.; Lin, H.; Xue, H.; Zhang, L.; Guo, J.; Cao, Z. Using FTOC to track shuttlecock for the badminton robot. *Neurocomputing* **2019**, *334*, 182–196. [CrossRef]

197. Rahmad, N.A.; Sufri, N.A.J.; Muzamil, N.H.; As'ari, M.A. Badminton player detection using faster region convolutional neural network. *Indones. J. Electr. Eng. Comput. Sci.* **2019**, *14*, 1330–1335. [CrossRef]

198. Hou, J.; Li, B. Swimming target detection and tracking technology in video image processing. *Microprocess. Microsyst.* **2021**, *80*, 103535. [CrossRef]

199. Cao, Y. Fast swimming motion image segmentation method based on symmetric difference algorithm. *Microprocess. Microsyst.* **2021**, *80*, 103541. [CrossRef]

200. Hegazy, H.; Abdelsalam, M.; Hussien, M.; Elmosalamy, S.; Hassan, Y.M.; Nabil, A.M.; Atia, A. IPingPong: A Real-time Performance Analyzer System for Table Tennis Stroke's Movements. *Procedia Comput. Sci.* **2020**, *175*, 80–87. [CrossRef]

201. Baclig, M.M.; Ergezinger, N.; Mei, Q.; Gül, M.; Adeeb, S.; Westover, L. A Deep Learning and Computer Vision Based Multi-Player Tracker for Squash. *Appl. Sci.* **2020**, *10*, 8793. [CrossRef]

202. Brumann, C.; Kukuk, M.; Reinsberger, C. Evaluation of Open-Source and Pre-Trained Deep Convolutional Neural Networks Suitable for Player Detection and Motion Analysis in Squash. *Sensors* **2021**, *21*, 4550. [CrossRef]

203. Wang, S.; Xu, Y.; Zheng, Y.; Zhu, M.; Yao, H.; Xiao, Z. Tracking a golf ball with high-speed stereo vision system. *IEEE Trans. Instrum. Meas.* **2018**, *68*, 2742–2754. [CrossRef]

204. Zhi-chao, C.; Zhang, L. Key pose recognition toward sports scene using deeply-learned model. *J. Vis. Commun. Image Represent.* **2019**, *63*, 102571. [CrossRef]

205. Liu, H.; Bhanu, B. Pose-Guided R-CNN for Jersey Number Recognition in Sports. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Long Beach, CA, USA, 16–17 June 2019; pp. 2457–2466. [CrossRef]

206. Pobar, M.; Ivašić-Kos, M. Detection of the leading player in handball scenes using Mask R-CNN and STIPS. In Proceedings of the Eleventh International Conference on Machine Vision (ICMV 2018), Munich, Germany, 1–3 November 2018; Volume 11041, pp. 501–508.

207. Van Zandycke, G.; De Vleeschouwer, C. Real-time CNN-based Segmentation Architecture for Ball Detection in a Single View Setup. In Proceedings of the 2nd International Workshop on Multimedia Content Analysis in Sports, Nice, France, 25 October 2019; pp. 51–58.

208. Burić, M.; Pobar, M.; Ivašić-Kos, M. Adapting YOLO network for ball and player detection. In Proceedings of the 8th International Conference on Pattern Recognition Applications and Methods, Prague, Czech Republic, 19–21 February 2019; Volume 1, pp. 845–851.

209. Pobar, M.; Ivasic-Kos, M. Active Player Detection in Handball Scenes Based on Activity Measures. *Sensors* **2020**, *20*, 1475. [CrossRef]

210. Komorowski., J.; Kurzejamski., G.; Sarwas., G. DeepBall: Deep Neural-Network Ball Detector. In Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, Valletta, Malta, 27–29 February 2019; Volume 5, pp. 297–304. [CrossRef]

211. Liu, W. Beach sports image detection based on heterogeneous multi-processor and convolutional neural network. *Microprocess. Microsyst.* **2021**, *82*, 103910. [CrossRef]

212. Zhang, R.; Wu, L.; Yang, Y.; Wu, W.; Chen, Y.; Xu, M. Multi-camera multi-player tracking with deep player identification in sports video. *Pattern Recognit.* **2020**, *102*, 107260. [CrossRef]

213. Karungaru, S.; Matsuura, K.; Tanioka, H.; Wada, T.; Gotoda, N. Ground Sports Strategy Formulation and Assistance Technology Develpoment: Player Data Acquisition from Drone Videos. In Proceedings of the 8th International Conference on Industrial Technology and Management (ICITM), Cambridge, UK, 2–4 March 2019; pp. 322–325.

214. Hui, Q. Motion video tracking technology in sports training based on Mean-Shift algorithm. *J. Supercomput.* **2019**, *75*, 6021–6037. [CrossRef]

215. Castro, R.L.; Canosa, D.A. Using Artificial Vision Techniques for Individual Player Tracking in Sport Events. *Proceedings* **2019**, *21*, 21.

216. Buric, M.; Ivasic-Kos, M.; Pobar, M. Player tracking in sports videos. In Proceedings of the IEEE International Conference on Cloud Computing Technology and Science (CloudCom), Sydney, Australia, 11–13 December 2019; pp. 334–340.

217. Moon, S.; Lee, J.; Nam, D.; Yoo, W.; Kim, W. A comparative study on preprocessing methods for object tracking in sports events. In Proceedings of the 20th International Conference on Advanced Communication Technology (ICACT), Chuncheon, Korea, 11–14 February 2018; pp. 460–462.

218. Xing, J.; Ai, H.; Liu, L.; Lao, S. Multiple player tracking in sports video: A dual-mode two-way bayesian inference approach with progressive observation modeling. *IEEE Trans. Image Process.* **2010**, *20*, 1652–1667. [CrossRef]

219. Liang, Q.; Wu, W.; Yang, Y.; Zhang, R.; Peng, Y.; Xu, M. Multi-Player Tracking for Multi-View Sports Videos with Improved K-Shortest Path Algorithm. *Appl. Sci.* **2020**, *10*, 864. [CrossRef]

220. Lu, W.L.; Ting, J.A.; Little, J.J.; Murphy, K.P. Learning to track and identify players from broadcast sports videos. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1704–1716.

221. Huang, Y.C.; Liao, I.N.; Chen, C.H.; İk, T.U.; Peng, W.C. Tracknet: A deep learning network for tracking high-speed and tiny objects in sports applications. In Proceedings of the 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Taipei, Taiwan, 18–21 September 2019; pp. 1–8.

222. Tan, S.; Yang, R. Learning similarity: Feature-aligning network for few-shot action recognition. In Proceedings of the International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 14–19 July 2019; pp. 1–7.

223. Ullah, A.; Ahmad, J.; Muhammad, K.; Sajjad, M.; Baik, S.W. Action recognition in video sequences using deep bi-directional LSTM with CNN features. *IEEE Access* **2017**, *6*, 1155–1166. [CrossRef]

224. Russo, M.A.; Kurnianggoro, L.; Jo, K.H. Classification of sports videos with combination of deep learning models and transfer learning. In Proceedings of the International Conference on Electrical, Computer and Communication Engineering (ECCE), Chittagong, Bangladesh, 7–9 February 2019; pp. 1–5.

225. Waltner, G.; Mauthner, T.; Bischof, H. Indoor Activity Detection and Recognition for Sport Games Analysis. *arXiv* **2021**, arXiv:abs/1404.6413.

226. Soomro, K.; Zamir, A.R. Action recognition in realistic sports videos. In *Computer Vision in Sports*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 181–208.

227. Xu, K.; Jiang, X.; Sun, T. Two-stream dictionary learning architecture for action recognition. *IEEE Trans. Circuits Syst. Video Technol.* **2017**, *27*, 567–576. [CrossRef]

228. Chaudhury, S.; Kimura, D.; Vinayavekhin, P.; Munawar, A.; Tachibana, R.; Ito, K.; Inaba, Y.; Matsumoto, M.; Kidokoro, S.; Ozaki, H. Unsupervised Temporal Feature Aggregation for Event Detection in Unstructured Sports Videos. In Proceedings of the IEEE International Symposium on Multimedia (ISM), San Diego, CA, USA, 9–11 December 2019; pp. 9–97.

229. Li, Y.; He, H.; Zhang, Z. Human motion quality assessment toward sophisticated sports scenes based on deeply-learned 3D CNN model. *J. Vis. Commun. Image Represent.* **2020**, *71*, 102702. [CrossRef]

230. Chen, H.T.; Chou, C.L.; Tsai, W.C.; Lee, S.Y.; Lin, B.S.P. HMM-based ball hitting event exploration system for broadcast baseball video. *J. Vis. Commun. Image Represent.* **2012**, *23*, 767–781. [CrossRef]

231. Punchihewa, N.G.; Yamako, G.; Fukao, Y.; Chosa, E. Identification of key events in baseball hitting using inertial measurement units. *J. Biomech.* **2019**, *87*, 157–160. [CrossRef] [PubMed]

232. Kapela, R.; Świetlicka, A.; Rybarczyk, A.; Kolanowski, K. Real-time event classification in field sport videos. *Signal Process. Image Commun.* **2015**, *35*, 35–45. [CrossRef]

233. Maksai, A.; Wang, X.; Fua, P. What players do with the ball: A physically constrained interaction modeling. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 972–981.

234. Goud, P.S.H.V.; Roopa, Y.M.; Padmaja, B. Player Performance Analysis in Sports: With Fusion of Machine Learning and Wearable Technology. In Proceedings of the 3rd International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 27–29 March 2019; pp. 600–603.

235. Park, Y.J.; Kim, H.S.; Kim, D.; Lee, H.; Kim, S.B.; Kang, P. A deep learning-based sports player evaluation model based on game statistics and news articles. *Knowl.-Based Syst.* **2017**, *138*, 15–26. [CrossRef]

236. Tejero-de Pablos, A.; Nakashima, Y.; Sato, T.; Yokoya, N.; Linna, M.; Rahtu, E. Summarization of user-generated sports video by using deep action recognition features. *IEEE Trans. Multimed.* **2018**, *20*, 2000–2011. [CrossRef]

237. Javed, A.; Irtaza, A.; Khaliq, Y.; Malik, H.; Mahmood, M.T. Replay and key-events detection for sports video summarization using confined elliptical local ternary patterns and extreme learning machine. *Appl. Intell.* **2019**, *49*, 2899–2917. [CrossRef]

238. Rafiq, M.; Rafiq, G.; Agyeman, R.; Choi, G.S.; Jin, S.I. Scene classification for sports video summarization using transfer learning. *Sensors* **2020**, *20*, 1702. [CrossRef]

239. Khan, A.A.; Shao, J.; Ali, W.; Tumrani, S. Content-Aware summarization of broadcast sports Videos: An Audio–Visual feature extraction approach. *Neural Process. Lett.* **2020**, *52*, 1945–1968. [CrossRef]

240. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]

241. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In Proceedings of the International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015.

242. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.

243. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

244. Iandola, F.N.; Moskewicz, M.W.; Ashraf, K.; Han, S.; Dally, W.J.; Keutzer, K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <1 MB model size. *arXiv* **2016**, arXiv:abs/1602.07360.

245. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2021**, arXiv:1704.04861.

246. Murthy, C.B.; Hashmi, M.F.; Bokde, N.D.; Geem, Z.W. Investigations of object detection in images/videos using various deep learning techniques and embedded platforms—A comprehensive review. *Appl. Sci.* **2020**, *10*, 3280. [CrossRef]

247. Cao, D.; Zeng, K.; Wang, J.; Sharma, P.K.; Ma, X.; Liu, Y.; Zhou, S. BERT-Based Deep Spatial-Temporal Network for Taxi Demand Prediction. *IEEE Trans. Intell. Transp. Syst.* **2021**, Early Access. [CrossRef]

248. Wang, J.; Zou, Y.; Lei, P.; Sherratt, R.S.; Wang, L. Research on recurrent neural network based crack opening prediction of concrete dam. *J. Internet Technol.* **2020**, *21*, 1161–1169.

249. Chen, C.; Li, K.; Teo, S.G.; Zou, X.; Li, K.; Zeng, Z. Citywide traffic flow prediction based on multiple gated spatio-temporal convolutional neural networks. *ACM Trans. Knowl. Discov. Data* **2020**, *14*, 1–23. [CrossRef]

250. Zaremba, W.; Sutskever, I.; Vinyals, O. Recurrent Neural Network Regularization. *arXiv* **2014**, arXiv:abs/1409.2329.

251. Jiang, X.; Yan, T.; Zhu, J.; He, B.; Li, W.; Du, H.; Sun, S. Densely connected deep extreme learning machine algorithm. *Cogn. Comput.* **2020**, *12*, 979–990. [CrossRef]

252. Zhang, Y.; Wang, C.; Wang, X.; Zeng, W.; Liu, W. Fairmot: On the fairness of detection and re-identification in multiple object tracking. *Int. J. Comput. Vis.* **2021**, *129*, 3069–3087 [CrossRef]

253. Wojke, N.; Bewley, A.; Paulus, D. Simple online and realtime tracking with a deep association metric. In Proceedings of the IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 3645–3649.

254. Hu, H.N.; Yang, Y.H.; Fischer, T.; Darrell, T.; Yu, F.; Sun, M. Monocular Quasi-Dense 3D Object Tracking. *arXiv* **2021**, arXiv:2103.07351.

255. Kim, A.; Osep, A.; Leal-Taixé, L. EagerMOT: 3D Multi-Object Tracking via Sensor Fusion. In Proceedings of the 2021 IEEE International Conference on Robotics and Automation (ICRA), Xi'an, China, 30 May–5 June 2021; pp. 11315–11321.

256. Chaabane, M.; Zhang, P.; Beveridge, J.R.; O'Hara, S. Deft: Detection embeddings for tracking. *arXiv* **2021**, arXiv:2102.02267.

257. Zeng, F.; Dong, B.; Wang, T.; Chen, C.; Zhang, X.; Wei, Y. MOTR: End-to-End Multiple-Object Tracking with TRansformer. *arXiv* **2021**, arXiv:2105.03247.

258. Wang, Z.; Zheng, L.; Liu, Y.; Li, Y.; Wang, S. Towards real-time multi-object tracking. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; pp. 107–122.

259. Xu, Y.; Osep, A.; Ban, Y.; Horaud, R.; Leal-Taixé, L.; Alameda-Pineda, X. How to train your deep multi-object tracker. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 6787–6796.

260. Sun, P.; Jiang, Y.; Zhang, R.; Xie, E.; Cao, J.; Hu, X.; Kong, T.; Yuan, Z.; Wang, C.; Luo, P. Transtrack: Multiple-object tracking with transformer. *arXiv* **2021**, arXiv:2012.15460.

261. Xu, Z.; Zhang, W.; Tan, X.; Yang, W.; Su, X.; Yuan, Y.; Zhang, H.; Wen, S.; Ding, E.; Huang, L. PointTrack++ for Effective Online Multi-Object Tracking and Segmentation. *arXiv* **2021**, arXiv:2007.01549.

262. Gupta, A.; Johnson, J.; Fei-Fei, L.; Savarese, S.; Alahi, A. Social gan: Socially acceptable trajectories with generative adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 2255–2264.

263. Phan-Minh, T.; Grigore, E.C.; Boulton, F.A.; Beijbom, O.; Wolff, E.M. Covernet: Multimodal behavior prediction using trajectory sets. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 14074–14083.

264. Li, X.; Ying, X.; Chuah, M.C. Grip: Graph-based interaction-aware trajectory prediction. In Proceedings of the IEEE Intelligent Transportation Systems Conference (ITSC), Auckland, NZ, USA, 27–30 October 2019; pp. 3960–3966.

265. Salzmann, T.; Ivanovic, B.; Chakravarty, P.; Pavone, M. Trajectron++: Dynamically-feasible trajectory forecasting with heterogeneous data. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; pp. 683–700.

266. Mohamed, A.; Qian, K.; Elhoseiny, M.; Claudel, C. Social-stgcnn: A social spatio-temporal graph convolutional neural network for human trajectory prediction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 14424–14432.

267. Amirian, J.; Zhang, B.; Castro, F.V.; Baldelomar, J.J.; Hayet, J.B.; Pettré, J. Opentraj: Assessing prediction complexity in human trajectories datasets. In Proceedings of the Asian Conference on Computer Vision, Kyoto, Japan, 30 November–4 December 2020; pp. 1–17.

268. Yu, C.; Ma, X.; Ren, J.; Zhao, H.; Yi, S. Spatio-temporal graph transformer networks for pedestrian trajectory prediction. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; pp. 507–523.

269. Wang, C.; Wang, Y.; Xu, M.; Crandall, D.J. Stepwise Goal-Driven Networks for Trajectory Prediction. *arXiv* **2021**, arXiv:abs/2103.14107.

270. Chen, J.; Li, K.; Bilal, K.; Li, K.; Philip, S.Y. A bi-layered parallel training architecture for large-scale convolutional neural networks. *IEEE Trans. Parallel Distrib. Syst.* **2018**, *30*, 965–976. [CrossRef]

271. Gu, X.; Xue, X.; Wang, F. Fine-Grained Action Recognition on a Novel Basketball Dataset. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; pp. 2563–2567.

272. Giancola, S.; Amine, M.; Dghaily, T.; Ghanem, B. Soccernet: A scalable dataset for action spotting in soccer videos. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 1711–1721.

273. Conigliaro, D.; Rota, P.; Setti, F.; Bassetti, C.; Conci, N.; Sebe, N.; Cristani, M. The s-hock dataset: Analyzing crowds at the stadium. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 2039–2047.

274. Niebles, J.C.; Chen, C.W.; Li, F.-F. Modeling temporal structure of decomposable motion segments for activity classification. In Proceedings of the European Conference on Computer Vision, Crete, Greece, 5–11 September 2010; pp. 392–405.

275. Voeikov, R.; Falaleev, N.; Baikulov, R. TTNet: Real-time temporal and spatial video analysis of table tennis. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 13–19 June 2020; pp. 884–885.

276. Pettersen, S.A.; Johansen, D.; Johansen, H.; Berg-Johansen, V.; Gaddam, V.R.; Mortensen, A.; Langseth, R.; Griwodz, C.; Stensland, H.K.; Halvorsen, P. Soccer video and player position dataset. In Proceedings of the 5th ACM Multimedia Systems Conference, Singapore, 19 March 2014; pp. 18–23.

277. D'Orazio, T.; Leo, M.; Mosca, N.; Spagnolo, P.; Mazzeo, P.L. A semi-automatic system for ground truth generation of soccer video sequences. In Proceedings of the Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance, Genova, Italy, 2–4 September 2009; pp. 559–564.

278. Feng, N.; Song, Z.; Yu, J.; Chen, Y.P.P.; Zhao, Y.; He, Y.; Guan, T. SSET: A dataset for shot segmentation, event detection, player tracking in soccer videos. *Multimed. Tools Appl.* **2020**, *79*, 28971–28992. [CrossRef]

279. Zhang, W.; Liu, Z.; Zhou, L.; Leung, H.; Chan, A.B. Martial arts, dancing and sports dataset: A challenging stereo and multi-view dataset for 3D human pose estimation. *Image Vis. Comput.* **2017**, *61*, 22–39. [CrossRef]

280. De Vleeschouwer, C.; Chen, F.; Delannay, D.; Parisot, C.; Chaudy, C.; Martrou, E.; Cavallaro, A. Distributed video acquisition and annotation for sport-event summarization. *NEM Summit* **2008**, *8*. Available onoline: http://hdl.handle.net/2078.1/90154 (accessed on 12 February 2020).

281. Karpathy, A.; Toderici, G.; Shetty, S.; Leung, T.; Sukthankar, R.; Fei-Fei, L. Large-scale video classification with convolutional neural networks. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 1725–1732.

282. Dou, Z. Research on virtual simulation of basketball technology 3D animation based on FPGA and motion capture system. *Microprocess. Microsyst.* **2021**, *81*, 103679. [CrossRef]

283. Yin, L.; He, R. Target state recognition of basketball players based on video image detection and FPGA. *Microprocess. Microsyst.* **2021**, *80*, 103340. [CrossRef]

284. Bao, H.; Yao, X. Dynamic 3D image simulation of basketball movement based on embedded system and computer vision. *Microprocess. Microsyst.* **2021**, *81*, 103655. [CrossRef]

285. Junjun, G. Basketball action recognition based on FPGA and particle image. *Microprocess. Microsyst.* **2021**, *80*, 103334. [CrossRef]

286. Avaya. Avaya: Connected Sports Fans 2016—Trends on the Evolution of Sports Fans Digital Experience with Live Events. Available online: https://www.panoramaaudiovisual.com/wp-content/uploads/2016/07/connected-sports-fan-2016-report-avaya.pdf (accessed on 12 February 2020).

287. Duarte, F.F.; Lau, N.; Pereira, A.; Reis, L.P. A survey of planning and learning in games. *Appl. Sci.* **2020**, *10*, 4529. [CrossRef]

288. Lee, H.S.; Lee, J. Applying artificial intelligence in physical education and future perspectives. *Sustainability* **2021**, *13*, 351. [CrossRef]

289. Egri-Nagy, A.; Törmänen, A. The game is not over yet—go in the post-alphago era. *Philosophies* **2020**, *5*, 37. [CrossRef]

290. Innovations, H.E. Hawk-Eye in Cricket. 2017. Available online: https://www.hawkeyeinnovations.com/sports/cricket (accessed on 12 February 2020).

291. Innovations, H.E. Hawk-Eye Tennis System. 2017. Available online: https://www.hawkeyeinnovations.com/sports/tennis (accessed on 12 February 2020).

292. Innovations, H.E. Hawk-Eye Goal Line Technology. 2017. Available online: https://www.hawkeyeinnovations.com/products/ball-tracking/goal-line-technology (accessed on 12 February 2020).

293. SportVU, S. Player Tracking and Predictive Analytics. 2017. Available online: https://www.statsperform.com/team-performance/football/optical-tracking/ (accessed on 12 February 2020).

294. ChyronHego. Product Information Sheet TRACAB Optical Tracking. 2017. Available online: https://chyronhego.com/wp-content/uploads/2019/01/TRACAB-PI-sheet.pdf (accessed on 12 February 2020).

295. Leong, L.H.; Zulkifley, M.A.; Hussain, A.B. Computer vision approach to automatic linesman. In Proceedings of the IEEE 10th International Colloquium on Signal Processing and its Applications, Kuala Lumpur, Malaysia, 9–10 March 2014; pp. 212–215.

296. Zhang, T.; Ghanem, B.; Ahuja, N. Robust multi-object tracking via cross-domain contextual information for sports video analysis. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan, 25–30 March 2012; pp. 985–988.

297. Xiao, J.; Stolkin, R.; Leonardis, A. Multi-target tracking in team-sports videos via multi-level context-conditioned latent behaviour models. In Proceedings of the British Machine Vision Conference, Nottingham, UK, 1–5 September 2014.

298. Wang, J.; Yang, Y.; Wang, T.; Sherratt, R.S.; Zhang, J. Big data service architecture: A survey. *J. Internet Technol.* **2020**, *21*, 393–405.

299. Zhang, J.; Zhong, S.; Wang, T.; Chao, H.C.; Wang, J. Blockchain-based systems and applications: A survey. *J. Internet Technol.* **2020**, *21*, 1–14.

300. Pu, B.; Li, K.; Li, S.; Zhu, N. Automatic fetal ultrasound standard plane recognition based on deep learning and IIoT. *IEEE Trans. Ind. Inform.* **2021**, *17*, 7771–7780. [CrossRef]

301. Messelodi, S.; Modena, C.M.; Ropele, V.; Marcon, S.; Sgrò, M. A Low-Cost Computer Vision System for Real-Time Tennis Analysis. In *Proceedings of the International Conference on Image Analysis and Processing*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 106–116.

302. Liu, Y.; Liang, D.; Huang, Q.; Gao, W. Extracting 3D information from broadcast soccer video. *Image Vis. Comput.* **2006**, *24*, 1146–1162. [CrossRef]