Article

Detection of Ice Hockey Players and Teams via a Two-Phase Cascaded CNN Model

Tianxiao Guo [1], Kuan Tao [2,*], Qingrui Hu [1] and Yanfei Shen [2,*]

[1] School of Sport Science, Beijing Sport University, 100084, Beijing, PRC; 1004320180262@bsu.edu.cn (T.X.G); 1004320180260@bsu.edu.cn (Q.R.H)
[2] School of Sports Engineering, Beijing Sport University, 100084, Beijing, PRC; taokuan@bsu.edu.cn (K.T); syf@bsu.edu.cn (Y.F.S)
* Correspondence: syf@bsu.edu.cn (Y.F.S); taokuan@bsu.edu.cn (K.T)

Featured Application: This study accomplishes automated detection of ice hockey players labeled with team information, which is particularly designed for analyses of trajectory tracking data in ice hockey, and is important for performance evaluation of individual player and team tactical decisions.

Abstract: Accurate detection of players and teams in ice hockey games is crucial to the tracking of individual players on court and team tactical decisions, which is therefore becoming an important task for coaches and other analysts. However, hockey is a fluid sport due to complex dynamics and frequent substitutions by both teams, resulting in various body positions of players. Few traditional player detection models from other team sports take these characteristics into account, especially for the detection of teams without prior annotations. Here, we design a two-phase cascaded Convolutional Neural Network (CNN) model coupling between individual players position information and team uniform colors to hierarchically detect players in ice hockey games. Our model filters most of disturbing information, such as audience and sideline advertising bars, in Phase I, and refines the detection of targeted players in Phase II, which is efficient and accurate with a precision rate of 91.30% and a recall rate of 95.60% for individual players detection, and an average accuracy of 93.05% in team classification from a self-built dataset of collected images in the 2018 Winter Olympics. Meanwhile, we also present results based on the images and real-time detection from broadcasting videos of 2019-20 NHL regular games covering all 31 teams to show the robustness of our model.

Keywords: Player detection; Team detection; Player tracking data; Ice hockey

1. Introduction

Ice hockey is a popular team sport in North America and Northern Europe that is described as a fluid sport [1, 2] with players frequently substituting on and off the court without timeouts. Although hockey games are fascinating to watch, using analytical approaches to assess player performance is still at an early age due to low score [2] and complex dynamics [3, 4]. Evaluating performance of individual players and their contribution to the overall performance of the team [5, 6] is a major challenge in the field of sports analysis. Several metrics were proposed for performance analysis in different team sports, i.e. "Expected-Point-Value" in basketball [7, 8], "Expected-Goal-Value" in soccer [9] and American football [10], etc. In professional ice hockey leagues, such as the National Hockey League (NHL) in North America, winning the final championship is the greatest honor and desire of all players and teams. As a result, a number of natural concerns arise, such as how to assemble a winning team with players maximizing their capacity, how to schedule the most effective tactics after comparing different offensive and defensive formats. The key factor to answer these questions is to take advantage of enormous data.

As ice hockey is complicated from a spatial-temporal perspective, the most valuable data is trajectory tracking data, which encodes vital information on the actions and intentions of players [3]. Many models used deep learning diagrams to analyze player dynamics and team dynamics based on trajectory data in different team sports. Le et al. [7] utilized deep imitation learning to generate alternative strategies for defensive teams on soccer. Similar work was also carried out in soccer [11]. Miller et al. [12] analyzed NBA team strategies through probabilistic theme modeling that captured the structure of player trajectories. Wang et al. and Mehrasa et al. [3, 13] used a convolutional neural network to classify offensive plays in basketball games, while Tian et al. [14] distinguished defensive patterns through a number of machine learning models based on team trajectory data.

One of the most essential steps to collect trajectory data is the detection of the targeted players. Although deep learning approaches [15-21] have been widely applied for the detection of objects, detection of players is more difficult because of complex game dynamics and sparsely distributed players appearing from broadcasting videos. A variety of player detection studies are introduced via non-intrusive methods. Lara et.al [22] used two calibrated cameras to capture the location of individual tennis player to assist as an auxiliary training medium. Lu et.al [23] and Parisot et.al [24] achieved player detection from broadcasting images via a single calibrated camera.

However, despite numerous literature on trajectory tracking models, few focus on ice hockey. The main reason is that traditional models fail to recover several common characteristic features in ice hockey games, such as severe occlusions and mass of physical confrontations between players [25-27]. In addition, due to high movement speed and abrupt direction change, ice hockey players always present body positions with stretched aspect ratios. These features challenge the efficiency and accuracy of detection of individual players and teams, which is regarded as the most vital component of analyzing trajectory data.

In this paper, we design a two-phase cascaded Convolutional Neural Network (CNN) model coupling between individual players position information and team uniform colors to detect players in ice hockey games. Phase I of the architecture roughly detects the targeted players by filtering most of disturbing information, such as audience and sideline advertising bars, while Phase II incorporates detailed information as overlapped area (occlusion) of body position caused by individual players and uniform color from different teams to further refine results following the outputs of Phase I. We constructed a image collection from the 2018 Winter Olympics and divided it into a training (4048 samples) and test (1327 samples) dataset and calculated the distribution of aspect ratios of all players from the training data to train for a suitable bounding box from a deep learning framework to resolve the challenging conditions with players exhibiting various postures.

The results on test dataset show the detection of both individual players and teams with high accuracy and efficiency. Meanwhile, we also present results based on the images and real-time detection from broadcasting videos of 2019-20 NHL regular games covering all 31 teams to show the robustness of our model. Therefore, our proposed two-phase cascaded CNN model is particularly designed to detect individual players and teams in ice hockey games respectively, which aims at tracking personal trajectory data to elevate player performance, and recognizing team offensive and defensive patterns to make decisive tactical decisions.

2. Methods

2.1. Two-phase cascaded CNN model

The schematic diagram of our proposed two-phase cascaded neural network topology is shown in Figure 1, which is categorized as two phases that are branched from the backbone network with Phase I colored in yellow and Phase II colored in blue, respectively. The main procedures comprise three steps: (1) First, image patches are inputted to the backbone network and processed by a convolution layer Conv-B1 consisting of 16 filters of size 3×3, which is then followed by an activation layer with Rectified Linear Unit (ReLU) applied to speed up the convergence of learning rate and avoid the optimized function being trapped into saddle points or local minima. Furthermore, a pooling layer Pool-B1 of size 3×3 is used for subsampling the output image patches from Conv-B1/ReLU to ensure the extraction of the most informative local features based on max-pooling criteria. (2) Second, the feature maps from the backbone network are compressed by Pool-S1 through average-pooling in Phase I, which could further extract characteristic features and smooth the feature maps. In order to avoid overfitting, Pool-S1 is subsequently followed by another layer to drop units from the neural network [28] randomly, which is called Dropout. Furthermore, we apply a convolution layer Conv-S1 equipped with an activation function ReLU to produce feature maps with size of 1×1, which is then processed by Softmax layer to show the likelihood of input images to be recognized as a player. (3) Third, the architecture of Phase II follows almost the same procedures as Phase I. The only difference lies in the inputs, with raw images used in Phase I and selected outputs from Phase I used in Phase II. Hence, Phase II refines the results by reducing the numbers of false positive samples significantly.

In both classification branches, we define a cross-entropy loss function, which is common in other cases under deep learning framework [3, 23, 29, 30], to attune the weights of the proposed model. The training set is denoted as $S = \{(x_{i,j}, y_{i,j})\}$, $(1 \le i \le N, 1 \le j \le K)$ with $x_{i,j} \in \mathbb{R}^d$ representing the feature map of the $i-th$ sample at the $j-th$ cascaded phases and $y_{i,j} \in \{0,1\}$ standing for the binary label accordingly. The probability for the positive sample prediction is:

$$p_i(y_i = 1|x_i, w) = \prod_{j=1}^{K} p_{i,j}(y_{i,j} = 1|x_{i,j}, w), \tag{1}$$

where w are the weights of model. Likewise, The probability of a negative sample belongs to:

$$p_i(y_i = 0|x_i, w) = 1 - \prod_{j=1}^{K} p_{i,j}(y_{i,j} = 1|x_{i,j}, w). \tag{2}$$

Therefore the loss function is defined as:

$$L_p(w) = -\sum_{i=1}^{N} [y_i \log(p_i(y_i = 1|x_i, w)) + (1 - y_i) \log(p_i(y_i = 0|x_i, w))]. \tag{3}$$

The whole proposed cascaded CNN model is trained to minimize $L_p(w)$, meaning the lowest possibility of false positive/negative attains and targeted players or non-player elements are accurately classified.

One of the distinct features of our two-phase cascaded CNN model is that two separate classification phases are optimized as a unified block. Phase I aims at recognizing the ground-truth (usually annotated in the training set) as positive samples and randomly selecting other elements as negative ones, thus the rough outputs contain players and non-players elements, while the false positive samples would be labeled as negative samples during the training of Phase II. The model parameters weights (w) are treated globally and not updated until the finish of all procedures for both phases. This layered design of network topology exerts great impacts on reducing computational cost and improving performance, as the first phase eliminates most of non-player elements and confirms the detection of players in fidelity, and the second phase focuses on complicated false positive samples, such as the background of audience and sideline cameras. At the meantime, multiple convolution filters and pooling layers in both classification branches allow the maximum extraction of all features of true targeted players from raw images with high efficiency, including players with stretched aspect ratios and occlusions caused by different players.

2.2. Parameter settings

Since ice hockey players exhibit different postures depending on real-time team decisions, the bounding box of targeted detection should be specifically adjusted due to the overall statistics of players on court from the perspective of spatial scale and aspect ratio. Therefore, besides model parameters, we fixed two crucial physical parameters which were related to the postures for hockey players from various games by analyzing training set.

Parameter 1 The size of input image patches.

This parameter is the priority to be settled to obtain a suitable size of input image patches, since the number of layers in the neural network will increase exponentially for a large size of input image while small-size patches are not sufficient to extract discriminative features. Notice that the aspect ratios of ice hockey players are more widely distributed, compared to other popular team sports like basketball and soccer. In the broadcasting of ice hockey games, the photographic distance of cameras are much more close to the ground than in the soccer games, which results in the diverse size differences of different players. Meanwhile, due to the high moving speed of players and a mass of physical confrontations in ice hockey games, a large range of aspect ratios are presented than those in broadcasting basketball games. Therefore the size of input image patches was determined as 42×25 after calculating all players in the training set with results showing that the aspect ratios of players mainly equaled to 1.65 (height/width) and 1.40 (height/width) depending on different postures of players.

Parameter 2 Zooming scales of the original image.

Due to the filming angle and the location of broadcasting cameras, the sizes of players vary from image to image, depending on the player locations on court. For instance, players in half-court area are in larger size than in defending or attacking zone. As mentioned in *Parameter 1*, the size of input image patches has been determined by 42×25 through the statistics of aspect ratio. In order to detect players with diverse pixel sizes, input image patches should be cropped from original images by appropriate zooming scales to ensure that the sizes of cropped players are around 42×25. According to the statistics about player pixel sizes of our training set, we determined the zooming scales of original images by 30%, 40% and 50% respectively, covering most of player sizes.

3. Dataset construction from the 2018 Winter Olympics

Our dataset consists of six ice hockey broadcast games from the 2018 Winter Olympics in PyeongChang including five men's and one women's games, namely Olympic Athletes from Russia (OAR) vs Germany (GER), Canada (CAN) vs GER, Czech Republic (CZE) vs OAR, CAN vs The United States (USA), Finland (FIN) vs OAR and Switzerland (SUI) vs Japan (JPN). These videos are recorded by the official pan-tilt-zoom broadcast cameras. Most of the highlights and playback scenes have been manually removed from our dataset and some of the close shots from cameras are also excluded in order to ensure that players of interest are well positioned in the video. Images with resolution of 1280×720 were edited from these game videos where locations of players were manually annotated in the form of $\{x_{bbox}, y_{bbox}, h_{bbox}, w_{bbox}\}$, with x_{bbox} and y_{bbox} meaning x-coordinate and y-coordinate of the left-top point of the bounding box and h_{bbox} and w_{bbox} representing the pixel values of height and width per bounding box. In addition, we calculated several statistics to visualize our data which are shown in Figure 2.

Figure 2a and 2b show that both the scales and heights are distributed diversely with nearly half of scales are close to 1 and the height per player ranges from 80 pixels to 150 pixels. Meanwhile, the total number of players in each image is approximately subject to a Gaussian distribution with mean value of six players (Figure 2c). Notice that the main purpose of our proposed model is to detect individual players and teams, which indicates that team uniform color should be considered when segmenting the whole dataset. Hence the images from four games in which team uniform colors are in sharp contrast are divided into train set (705 images with 4048 of annotated players in total) and test set (212 images with 1327 of annotated players in total), respectively.

4. Experimental results

In this section, the numerical experiments on self-built dataset from the 2018 Winter Olympics in PyeongChang are conducted to clarify the validation of our proposed two-phase cascaded CNN model on the detection of individual players and teams. Furthermore, we apply the model to the broadcasting videos of 2019-20 NHL regular games covering all 31 teams (source from https://v.qq.com/) to verify the robustness. We achieve the detection of player and teams based on images (segmented frame by frame from videos) with efficiency and accuracy and accomplish the real-time tracking of players from video inputs.

4.1. Detection of individual players from self-built dataset in the 2018 Winter Olympics

The experiments in silico show the results about detection of individual players from four games, namely GRE vs OAR, CZE vs OAR, FIN vs OAR, CAN vs USA, which are obtained by our model. The first row in Figure 3 represents the Phase I of detection with green bounding boxes standing for the outputs from the first branch of classification. As we can see, misclassified situations arise when players are under complex context, such as mixed with audience (see the game FIN vs OAR and CAN vs USA) or the backgrounds are complicated, i. e. stripes appearing on the sideline advertising bars could easily be recognized as player number on jersey (see the game GER vs OAR and CZE vs OAR). However, these misclassified situations disappear in Phase II because as the true positive samples, targeted players have larger confidence score than those in false positive samples. Hence the second branch of classification singles out all false positive examples accurately.

Table 1 Performances about detection of individual players

Game	Correct	False	Missing	Precision	Recall	F-score
GER vs. OAR	270	35	8	0.8852	0.9712	0.9262
CZE vs. OAR	354	26	21	0.9316	0.9440	0.9377
CAN vs. USA	448	44	20	0.9106	0.9573	0.9333
FIN vs. OAR	196	16	10	0.9245	0.9515	0.9378

We also evaluate the performances of our two-phase cascaded CNN model according to precision rate, recall rate and F-score, which are shown in Table 1. Above all the selected games, the average precision rate and recall rate reach to 0.913±0.021 and 0.956±0.012, respectively. Since our model is capable of detecting players in blurred images and recognizing players with various body positions, high recall rates in the test set are thus achieved. The F-scores in all four games exceed 92%, demonstrating the robustness and the effectiveness of our detection of individual players.

4.2 Team classification from ice hockey games in the 2018 Winter Olympics

We further verify the validation of our model on the detection of teams by exploring team classifications based on color information of jerseys. We extract the features of uniform colors by designing a customized color space, which is presented in SM Figure 1, and detailed information is presented in the text of Supplementary Materials. This customized color space, indeed, distinguishes most of colors such as yellow, blue, red, etc. However, according to the statistics about uniform color from 64 countries of the International Ice Hockey Federation (IIHF), 87.5% of home team uniforms are white, which is a great challenge for our customized color space since white is a combination of other primary or secondary colors. In order to resolve the issue, we introduced a weighted mechanism that described the proportion of color components delicately (see Supplementary Materials).

Table 2 Performances of team classification

Games	GT	INF		
		TA (%)	TB (%)	OTHER (%)
GER (TA) vs. OAR (TB)	TA	95.83	4.17	0
	TB	0	100	0
CZE (TA) vs. OAR (TB)	TA	74.18	25.82	0
	TB	0	100	0
CAN (TA) vs. USA (TB)	TA	97.85	1.29	0.86
	TB	4.76	95.24	0
FIN (TA) vs. OAR (TB)	TA	80.46	0	19.54
	TB	0	100	0

Table 2 summarizes the evaluation of team classification where GT stands for the ground truth, INF stands for the inferred teams predicted by our model that comprise TA and TB representing team A and team B, respectively. OTHER means that the predicted results belong to a third category, for instance, a referee or misclassified teams. As is seen from Table 2, our approach attains an average accuracy of 93.05% for the detection of teams. Meanwhile, Figure 4 presents several examples about team classification for four different components of uniform. Although the colored bounding boxes matched perfectly with team uniform colors when the dominant color component belongs to one channel in our customized color space, such as GER, OAR, FIN and USA, the other two teams CZE and CAN are not lined with the color of their jerseys yet our model still distinguishes these different national squads. The reason is that their main color component of jersey is white.

The novelty of applying the weighted mechanism on customized color space for our model is that no extra preliminary annotations of team information in training set are needed, meaning the detection of teams is adaptive after the targeted players are predicted and correspondingly their team uniform colors are extracted. The detections of individual players and teams inspire us to confirm the robustness of our model in other ice hockey game broadcasting video data.

4.3 *Detection of player and teams simultaneously on 2019-20 season NHL games*

NHL is the most popular ice hockey league in North America that include 31 teams with each team scheduling for 82 games per regular season. Due to the high impact and high-level competence of NHL, enormous data from games is worthy exploring in order to make winning tactics and improve player performance. Therefore we build another test set containing the newest 39 games from 2019-20 season that cover all 31 teams in the league to verify the robustness of the two-phase cascaded CNN model. Since individual players and teams are well detected according to Figure 3 and Figure 4, the main purpose here is to detect individual players labeled with a predicted team.

We first calculate the color components of home and away uniforms of all NHL teams (source from nhluniforms.com), results are shown in Figure 5. We extract the mean hue (the attribute of color that enables an observer to classify it as red, green, blue, purple, etc., and excludes white, black, and shades of grey.) values of different uniforms and normalize them within a range of 0 to 1 in order to preliminarily describe the major color component, or essential color. An evident trend is obtained in Figure 5 that for the majority of teams, the mean hue value of home team uniforms is higher than that of away team uniforms, which is consistent with the intuition that uniforms of away teams are always white-based in NHL (this conclusion is opposite to the color-of-team-uniform principle for national squads, as mentioned in Section 4.2). Therefore, the significant differences of home and away team uniforms ensure the plausibility of our weighted mechanism on customized color space. Nonetheless, note that the uniforms of Los Angeles Kings and Detroit Red Wings are exceptions since the hue values of essential color equal to zero (major color components are black for Kings and red for Red Wings), so that the corresponding bars are colored in black and red respectively in Figure 5.

Furthermore, we select 24 sample images for all 31 teams as the results of detection in Figure 6. Most of players are accurately classified according to their teams, however, some misclassified samples exist because the illumination conditions in stadium and the positions of cameras are personalized team by team. Also, the game tempo is quicker for NHL games that allows the appearance of severe overlapping of players and mass of physical collisions, which challenge the detection accuracy of both individual players and team classification. In general, we accomplish the detection of players and label the team information without prior annotations in the training set, which motivate us to realize the real-time tracking of player of interest and to analyze the optimized moving trajectory and make the best team decisions.

4.4 *Real-time detection from videos of 2019-20 season NHL games*

Although ice hockey enjoys world-wide popularity, the existing data analyses of ice hockey games is of relatively small quantity compared to other popular team sports. Probably because of some features as stretched aspect ratios and occlusions are difficult to tackle with, the lack of sufficient data is a priority among all reasons with few sports data analyzing platforms provide profound explorations. However, enormous event data is produced per ice hockey game due to the quicker game tempo and more attack-defense-transition rounds. Hence, to elaborate a real-time detection of player and teams is of significant importance that is prerequisite to track moving trajectories of each player, and analyze moving patterns of teams, which could be applied into a variety of fields, i.e., visualization of player trajectory [31], heat map analysis [32], event recognition [4, 32, 33] and performance assessment [34, 35]. We aim to realize automated detection of players with recognized team information through the proposed model coupling with player locations and uniform color components.

Here, a broadcasting video of Pittsburgh Penguins vs Anaheim Ducks in the 2019-20 season is illustrated as an example (see SM Movie 1). The results about detection are sequentially presented in Figure 7 which consist of 30 frames edited from the game video about 10 seconds. Red and yellow bounding boxes mean the automated detections of players labeled with team affiliations. In some frames, referee is also picked out with a blue bounding box (see Figure 7, take 2nd 18:54 and 2nd 18:53 for examples), though this is not mandatory for the output of the model. This video exhibits a power play from Pittsburgh Penguins by terminating with a goal, which is a characteristic offensive pattern occurring in most ice hockey games. It is therefore practicable for coaches to further devise tactical decisions weighing the pros and cons on the positions of players in offensive team and defensive patterns from the opponent team.

5. Discussion

Accurate and automated detection of players in ice hockey games is crucial for improving player performance and team tactical decisions, which is therefore urgent for coaches and other analysts. In this paper, we proposed a two-phase cascaded CNN model to achieve the detection of individual players labeled with team affiliations on the self-built dataset from 2018 Winter Olympics in PyeongChang and the 2019-20 season NHL regular games. The precision and recall rate for the individual players detection in ice hockey games in Winter Olympics attain 91.30% and 95.60% correspondingly, while the average accuracy of team classification reaches 93.05%. Meanwhile, our model is applied on the dataset containing 2019-20 season NHL regular games covering all teams in the league, in which we accomplish the concurrent detections of players and their team affiliations from image and video inputs.

Our model is designed in particular for the player tracking data collection in ice hockey, which is capable of dealing with some characteristic features such as attaining high detection accuracy in fuzzy images, etc. In addition, the main architecture of proposed deep learning network is efficient and straightforward, as two phases in the network are subtle about the different confidence scores by targeted players and other non-informative contexts which avoids updating parameters redundantly and repetitively. Nevertheless, some further improvements have to be implemented in future study such as the recognition of players with extremely stretched aspect ratios and precise team classification when players are in physical contacts. Also, it is inspiring to note that the trajectories of players can be reconstructed so that the locations and movements of players are visualized temporally (Figure 8). We believe that advanced game statistics are available by developing our current models, which will shed light on insightful analyses and predictions in ice hockey.

Supplementary Materials: Supplementary materials include one supplementary Text, one supplementary Figure and one supplementary Movie.

Author Contributions: Conceptualization, T.X.G, K.T and Y.F.S; Methodology, T.X.G, K.T and Q.R.H; Writing—original draft preparation, T.X.G, K.T and Q.R.H; Writing—review and editing, T.X.G, K.T and Y.F.S. All authors have read and agreed to the published version of the manuscript.

Funding: Y.F.S received funding from the National Key R&D Program of China 2018YFC2000600 and the National Natural Science Foundation of China 11901037; K.T received funding from the Fundamental Research Funds for the Central Universities in China (Beijing Sport University), grant number 2020042.

Conflicts of Interest: The authors declare no conflict of interest.

Figure Legends:

Figure 1 Architecture of the two-phase cascaded CNN model. The backbone network along with Phase I and II classification branches are encircled by a yellow, blue and green dotted line, respectively. The suffixes "B1", "B2" and "S1", "S2" mean the convolution or pooling layer of Phase I (B1 and S1) or Phase II (B2 and S2) in both backbone network and classification branches, respectively.

Figure 2 Statistics of dataset. (a) The distribution about the scales of players. Scales are defined as the ratio of height and width. (b) The distribution about the heights of players with each number calculated as pixel values per player. (c) The total number of players per image.

Figure 3 Performance of player detection in two classification phases. The green bounding boxes mean the predictions of targeted player.

Figure 4 Detection of teams. The colored bounding boxes indicate the inferred teams suggested by the model.

Figure 5 Hue value statistics of uniforms of 31 NHL teams.

Figure 6 Detection of individual players and teams simultaneously from all teams in NHL.

Figure 7 Real-time detection of individual players and teams from a 10-second video. "2nd" means second quarter of the game and counting time above each frame represents the remaining time of the quarter.

Figure 8 The envisioned framework for tracking data analysis in ice hockey. Steps on solid line are complete while Step 4 to 6 on the dashed line constitute future work.

References

1. Weissbock, J., *Forecasting Success in the National Hockey League using In-Game Statistics and Textual Data*. 2014, University of Ottawa Canada.
2. Schuckers, M. and J. Curro. *Total Hockey Rating (THoR): A comprehensive statistical rating of National Hockey League forwards and defensemen based upon all on-ice events*. in *7th annual MIT sloan sports analytics conference*. 2013.
3. Mehrasa, N., et al., *Deep learning of player trajectory representations for team activity analysis*, in *11th MIT Sloan Sports Analytics Conference*. 2018.
4. Tora, M.R., J. Chen, and J.J. Little, *Classification of Puck Possession Events in Ice Hockey*. 2017 IEEE conference on computer vision and pattern recognition workshops (CVPRW), 2017: p. 147--154.
5. Chan, T.C.Y., J.A. Cho, and D.C. Novati, *Quantifying the Contribution of NHL Player Types to Team Performance*. Interfaces, 2012. **42**(2): p. 131--145.
6. Rajšp, A. and I. Fister, *A Systematic Literature Review of Intelligent Data Analysis Methods for Smart Sport Training*. Applied Sciences, 2020. **10**(9): p. 3013.

7. LeHoangM, et al., *Data-Driven Ghosting using Deep Imitation Learning*. 2017.
8. Cervone, D., et al. *POINTWISE: Predicting Points and Valuing Decisions in Real Time with NBA Optical Tracking Data*. in *Proceedings of the 8th MIT Sloan Sports Analytics Conference*. 2014. Boston, MA, USA.
9. Lucey, P., et al., *quality vs quantity: Improved shot prediction in soccer using strategic features from spatiotemporal data*, in *Proc. 8th annual MIT sloan sports analytics conference*. 2014.
10. Macdonald, B., *An Expected Goals Model for Evaluating NHL Teams and Players.* Proceedings of the 2012 MIT Sloan Sports Analytics Conference, 2012.
11. Bialkowski, A., et al., *Identifying team style in soccer using formations learned from spatiotemporal tracking data*, in *2014 IEEE international conference on data mining workshop*. 2014.
12. Miller, A.C. and L. Bornn, *Possession sketches: Mapping nba strategies*, in *11th MIT Sloan Sports Analytics Conference* 2017.
13. Wang, K.-C. and R. Zemel, *Classifying NBA offensive plays using neural networks*, in *Proceedings of MIT Sloan Sports Analytics Conference*. 2016.
14. Tian, C., et al., *Use of Machine Learning to Automate the Identification of Basketball Strategies Using Whole Team Player Tracking Data.* Applied Sciences, 2019. **10**(1).
15. Tang, P., et al., *Pcl: Proposal cluster learning for weakly supervised object detection.* IEEE transactions on pattern analysis and machine intelligence, 2018. **42**(1): p. 176-191.
16. He, K., et al. *Mask r-cnn*. in *Proceedings of the IEEE international conference on computer vision*. 2017.
17. Yang, F., W. Choi, and Y. Lin. *Exploit All the Layers: Fast and Accurate CNN Object Detector with Scale Dependent Pooling and Cascaded Rejection Classifiers*. in *computer vision and pattern recognition*. 2016.
18. Redmon, J., et al. *You Only Look Once: Unified, Real-Time Object Detection*. in *computer vision and pattern recognition*. 2016.
19. Ren, S., et al. *Faster R-CNN: towards real-time object detection with region proposal networks*. in *neural information processing systems*. 2015.
20. Gidaris, S. and N. Komodakis. *Object Detection via a Multi-region and Semantic Segmentation-Aware CNN Model*. in *international conference on computer vision*. 2015.
21. Girshick, R., et al. *Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation*. in *computer vision and pattern recognition*. 2014.
22. Lara, J.P.R., et al., *Validation of a video-based system for automatic tracking of tennis players.* International Journal of Performance Analysis in Sport, 2018. **18**(1): p. 137--150.
23. Lu, K., et al., *Light Cascaded Convolutional Neural Networks for Accurate Player Detection*, in *British Machine Vision Conference (BMVC)*. 2017.
24. Parisot, P. and C. De Vleeschouwer, *Scene-specific classifier for effective and efficient team sport players detection from a single calibrated camera.* Computer Vision and Image Understanding, 2017. **159**(159): p. 74--88.
25. Huda, N.U., et al., *Estimating the Number of Soccer Players Using Simulation-Based Occlusion Handling.* computer vision and pattern recognition, 2018: p. 1824-1833.
26. Nandashri, D. and P. Smitha, *An Efficient Tracking of Multi Object Visual Motion using Hungarian Method.* International journal of engineering research and technology, 2015. **4**(4).
27. Liu, J., et al., *Automatic player detection, labeling and tracking in broadcast soccer video.* Pattern Recognition Letters, 2009. **30**(2): p. 103 - 113.
28. Srivastava, N., et al., *Dropout: a simple way to prevent neural networks from overfitting.* Journal of Machine Learning Research, 2014. **15**(1): p. 1929--1958.

29. Rahman, T.C., M.E.H.; Khandakar, A.; Islam, K.R.; Islam, K.F.; Mahbub, Z.B.; Kadir, M.A.; Kashem, S., *Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection Using Chest X-Ray*. Appl. Sci., 2020.
30. Kandel, I.C., M., *A Novel Architecture to Classify Histopathology Images Using Convolutional Neural Networks.* Appl. Sci., 2020.
31. Pileggi, H., et al., *SnapShot: Visualization to Propel Ice Hockey Analytics*, in *IEEE Transactions on Visualization and Computer Graphics*. 2012. p. 2819-2828.
32. Imai, T., et al., *Play Recognition Using Soccer Tracking Data Based on Machine Learning.* network-based information systems, 2018: p. 875--884.
33. Kaarthick, P.K.S.B., *An Automated Player Detection and Tracking in Basketball Game.* Cmc-computers Materials & Continua, 2019. **58**(3): p. 625--639.
34. Schulte, O., Z. Zhao, and M. Javan. *Apples-to-Apples : Clustering and Ranking NHL Players Using Location Information and Scoring Impact*. in *Proceedings of the MIT Sloan Sports Analytics Conference.* 2017.
35. Schulte, O., et al., *A Markov Game model for valuing actions, locations, and team performance in ice hockey.* Data Mining and Knowledge Discovery, 2017. **31**(6): p. 1735-1757.