Unlocking Ice Hockey Prowess: Pose-Centric Analysis With MMaction2, Yolov10, and BoT-Sort for Sports Education

Boyang Zhang, Independent Scholar, Finland

The European Conference on Education 2024
Official Conference Proceedings

Abstract
In contemporary fixed-field team sports, the integration of computer vision techniques has become a popular practice for scrutinizing team strategies and tactics. This study embraces a comprehensive approach that amalgamates action recognition using MMaction2, object detection by Yolov10, and multi-object tracking using Bot-Sort to unravel the intricacies of ice hockey strategies. With a keen focus on pivotal actions such as passing and shooting movements, this research utilizes the McGill Hockey Player Tracking Dataset (MHPTD) to unearth valuable insights into ice hockey strategies. In the methodology, we provide a concise understanding of ice hockey dynamics, highlighting the significance of key actions in team sports. Social Network Analysis (SNA) is performed to understand ice hockey strategies, emphasizing the role of key actions, and fostering innovative approaches to sports education and analysis. We emphasize that key actions, such as precise passing and strategic offensive moves, could offer a profound impact on SNA. Moreover, we use Louvain method for community detection which allows us to uncover latent structures within the player network that extend beyond connections. By adding action patterns to the community, we reveal clusters of players who collaborate closely on the field and exhibit similar playing styles and strategic preferences. This study offers coaches and educators actionable insights for designing training programs and understanding team tactics. By integrating cutting-edge technologies with educational principles, our research contributes to performance analysis and modern coaching methods, enriching the landscape of fixed-field team sports education.

Keywords: Object Detection, Pose Detection, Social Network Analysis

Introduction

In the analysis of modern fix-field sports, computer vision technologies have been commonly used to interpret team strategies. The integration of the advanced computer vision techniques allows for a detailed look at the movements and interactions in sports. This study sets up a comprehensive approach combining action recognition using MMaction2 which are designed by MMaction2 contributors (2020), object detection with Yolov10 (Wang et al., 2024), and multi-object tracking through Bot-Sort Aharon et al. (2022) to unveil the game strategies in ice hockey. By focusing on critical actions like passing and shooting, this study aims to analysis the interactions between players by pose recognition and Social Network Analysis (SNA) to uncover ice hockey team strategies.

As to action recognition, MMActions2 is a powerful open-source toolbox designed for video understanding, particularly excelling in action recognition tasks, which can be used in analyzing group fixed-field sports, such as soccer, basketball, American football, volleyball and ice hockey. These fixed-field sports involve complex player interactions, rapid movements, and diverse actions occurring within a defined space (Zhang & Löwendahl, 2023), for example, like shooting, and passing action. MMActions2 provides a robust solution for identifying and classifying actions which can be used for detailed sports analysis, allowing coaches and analysts to dissect player behaviors and team strategies efficiently.

Figure 1: Shooting position detected by MMaction2

Yolov10 is so far the latest advancement in the You Only Look Once (YOLO) series for real-time object detection in 2024, Yolov10 combines the tradition of balancing computational efficiency with detection performance (Wang et al., 2024). Yolov10 introduces important enhancements in post-processing and model architecture. A key innovation in Yolov10 is the adoption of consistent dual assignments for Non-Maximum Suppression (NMS) free training, which improves efficiency by eliminating redundant predictions and reducing inference latency. What's more, Yolov10 employs a holistic efficiency-accuracy driven design strategy, optimizing various model components to minimize computational overhead while enhancing performance (Wang et al., 2024). Ice hockey is a fixed-field sport characterized by its fast pace and dynamic player interactions. Yolov10 can be utilized for object detection in ice hockey.

As to BoT-SORT which is an advanced tracker designed for multi-object tracking (MOT) tasks, Bot-SORT combines motion and appearance information to track objects across frames, ensuring robust and accurate tracking in challenging scenarios (Aharon et al., 2022).

Moreover, the MOT tasks contain the tracking of multiple objects across a series of video frames, specifically for sport analysis, each players with unique object_ID can be tracked and Re-ID for re-identification, which is to re-track the object ID which is the specific player, like Temmu Selanne with object ID "01". which ensures the same player is consistently tracked across frames, even when it undergoes occlusions or exits and re-enters the scene. Another key advantage of Bot-SORT is the Camera Motion Compensation (CMC), which addresses the challenges posed by camera movement (Aharon et al., 2022). CMC adjusts the predicted positions of objects to account for camera motion, thereby maintaining accurate tracking despite the dynamic background.

Theoretical Background

Skeleton-based action recognition models use human skeletal movements to identify and classify actions in video sequences (Duan et al., 2023). For example, in the Adaptive Graph Convolutional Network (AGCN), it uses a unique approach of the two-stream adaptive graph convolutional network (2s-AGCN), adapting to diverse action patterns and enhancing flexibility and recognition accuracy (Shi et al., 2019). PoseC3D employs a 3D heatmap stack representation of skeletons, which improves robustness against pose estimation noise and supports multi-person scenarios (Duan et al., 2021). The Spatial-Temporal Graph Convolutional Network (STGCN) (Yan et al. 2018) and its enhanced version, STGCN++ (Duan et al., 2022), further refines the approach by automatically learning spatial and temporal patterns. The usage of MMAction2 which can integrate above models to facilitate model implementation, training, and evaluation. MMAction2 enables coaches to leverage AGCN, PoseC3D, STGCN, and STGCN++ for advanced action recognition tasks.

Nowadays Social Network Analysis (SNA) has been commonly used in sport research (Wäsche et al. 2017), by researching the structure and dynamics of the relationship in a network. The SNA in sports analysis can provide a deep understand of player interactions and team dynamics (Zhang & Löwendahl, 2023), by identifying key actions like passing and shooting in SNA, it can help coaches to reveal patterns and connections among players, understand team strategies. Further, utilizing SNA can lead to effective training and strategy planning in team sport (Bruner et al., 2021).

Methodology

The MHPTD (McGill Hockey Player Tracking) dataset is selected to perform the analysis, it includes high-definition NHL gameplay video clips with continuous frames and divided into segments of 30 and 60 frames per second (Zhao, Li, & Chen, 2020), adheres to the Multiple Object Tracking (MOT) challenge format to facilitate the evaluation of computer vision algorithms for tracking the movement of ice hockey players' movements over time. The dataset was annotated in yolo classes format to facilitate comprehensive analysis, and there are six classes: player_color1, goalie_color1, player_color2, goalie_color2, referee, and ball.

Yolov10 is selected to perform the detection, there are 200 epochs in the training. In below Figure2, it provides a comprehensive analysis of training and validation losses, performance metrics, and learning rates over epochs. The training losses for object models consistently decrease, indicating effective learning and reduced errors over time. For instance, the train/box om loss decreases from 5.57 at epoch 1 to 2.67 at epoch 5, and the train/cls om loss decreases from 5.08 to 2.46 over the same period. Similarly, the training losses for object-object models follow a similar trend, with the train/box loss decreasing from 5.47 to 2.44, and the train/cls loss from 9.20 to 3.17. Validation losses provide further insights into the model's performance. The validation losses decrease, indicating the model's improving ability to generalize to unseen data. Performance metrics include precision, recall, mAP50, and mAP50-95 reflect the model's detection accuracy. Precision is reaching a high of 0.9147 by epoch 199. Recall showed a similar trend, the peaking of recall is at 0.8154 by epoch 200. The Mean Average Precision (MAP) at IoU threshold 0.50 (mAP50) achieves 0.79 by epoch 200. The mAP50-95, averaging over multiple IoU thresholds, and achieving 0.62 at the end of the training.

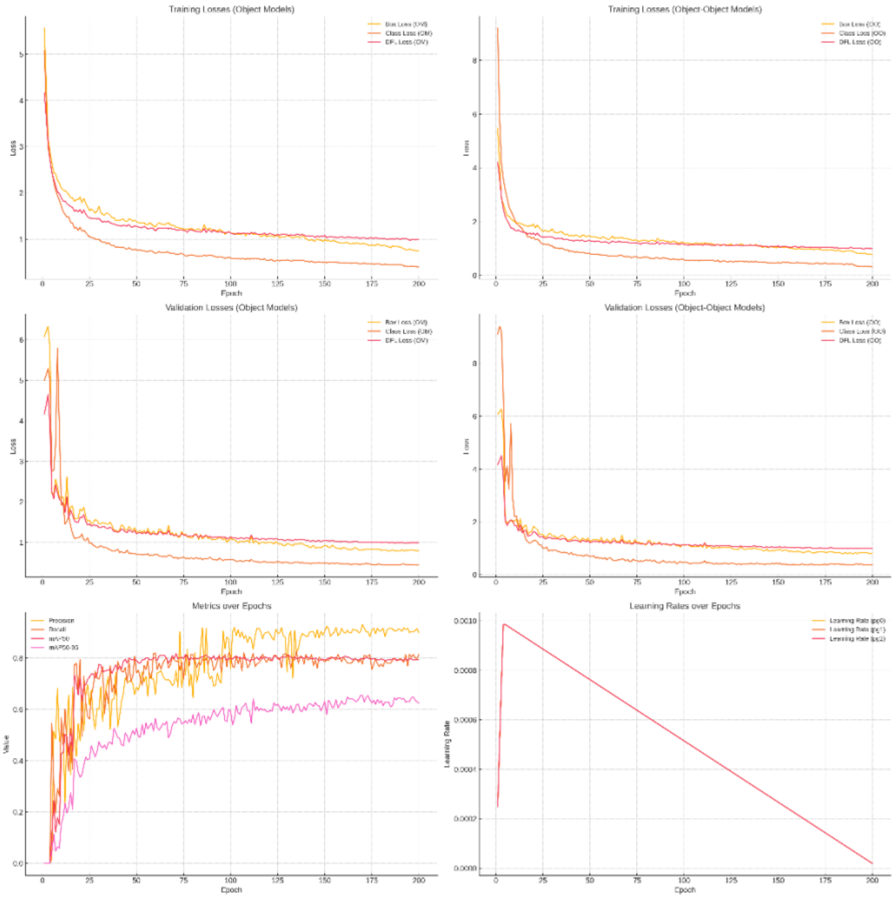

Figure 2: Yolov10 Training and Validation Performance Metrics

To achieve skeleton-based action recognition for identifying shooting and passing actions of ice hockey players, we employed the MMAction2 framework and followed a systematic methodology. Initially, key points were extracted from MHPTD dataset, and actions are annotated in classes: passing and shooting. These key points are converted into the COCO format to ensure compatibility with MMAction2. The dataset was organized into structured annotations indicating the actions. The model was trained using MMAction2's training protocols. After that, the trained model was deployed for inference on new video clips to detect shooting and passing actions. This method integrates advanced computer vision techniques and the MMAction2 framework to facilitate detailed analysis of player actions in passing and shooting actions.

Findings and Discussions

In this chapter, we will discuss the usage of object detection, pose detection and SNA in hockey games. Yolov10 provides the precise identification and classification of ice hockey players. Each detected object was assigned a class_ID, representing different categories such as players from different teams, referees, and the ball. Within each class_ID, unique object_IDs were assigned to individual instances, enabling detailed tracking using Bot-SORT and analysis of each hockey player over time. For example, in the player_black class, each player was assigned a specific object_ID, such as player_black ID1, ID3, ID5, facilitating the examination of individual player movements and actions. This granularity in object detection and classification of object_ID by Yolov10 allows for comprehensive analysis of player behaviors and interactions. With the usage of Bot-SORT tracker, we can track the player's positions over time accurately, providing valuable insights into understanding players' movements and locations. By ensuring consistent and reliable tracking, even in complex situations when camera moves or players occlusion, Bot-SORT enhances the ability of re-identification of players, it can track players' movements continuously over game time.

Figure 3: Different object_ID for the same class (player_black) with ID1, ID3, and ID5

In the next step, each ice hockey players' movements are analyzed by MMAction2, which provides, the skeleton based action recognition of each individual player' movement, and it enables the extraction of key points of the skeletal structure of each individual player, the player with specific object_ID, allowing for precise tracking of body movements and poses. By combining the object detection data from Yolov10 with the detailed movement analysis like passing or shooting from MMAction2, we can achieve a comprehensive understanding of player behaviors and interactions inside the fixed-field sports.

In Social Network Analysis, there are nodes and edges, by building from hockey player's movements (passing and shooting), a directed network can be created with all the active players and the ball using Gephi (Figure 4). In the network, the nodes are the players, edges are the passing and shooting connection which have the direction. Social Network Analysis has been proven as a valuable methodology in sport analytics (Zhang & Löwendahl, 2023). By integrating SNA with advanced Yolov10 for object detection frameworks, coaches can track and classify player movements accurately. Each detected player is assigned a unique object_ID within their respective class_ID, allowing for detailed temporal analysis of individual and group behaviors. In hockey team analysis, player_color1_ID1 can be tracked throughout a game to analyze his/her movement patterns, passing frequency, and interactions with other team players. This approach enables a comprehensive understanding of team strategies, player roles, and overall game dynamics.

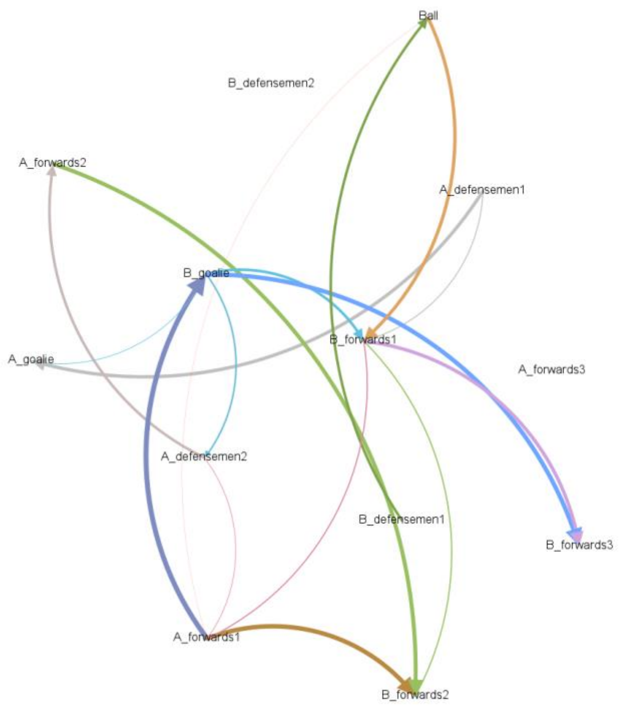

Figure 4: Social Network Analysis of Ice Hockey Players

After building the directed network, Louvain method (Blondel et al., 2008) is used for community detection which enables us to uncover latent structures within the player network. Louvain method identifies clusters who frequently interact with each other, and provides hierarchical structures within the game, for example, the macro level strategies of offensive and defensive groupings.

Conclusion

The object of this research is to analysis the interactions between different team players by object detection, pose recognition and SNA during an ice hockey game to uncover strategic patterns for coaches to interpret the game dynamics. By focusing on key poses such as the passing and shooting, information of interaction and timestamp can be analyzed of each individual player. The BoT-SORT tracker provides the players' position tracking with accuracy. The tracking activities enable coaches to understand the movement of players.

To better understand the game dynamics, Table 1 shows the interaction and timestamp involving a key player, for example TeamA1. The table outlines each interaction, as passing or shooting, along with the corresponding timestamp of when each action occurred during the game. By examining these interactions, we can gain insights into the strategic roles of different players and the overall team strategy. In Table1, TeamA1 frequently passed the ball to TeamA2 at multiple timestamps (0:01, 0:14, 0:27, 1:24). This suggests a strong connection or reliance between these two players, which could indicate a planned strategy where TeamA2 is positioned to support or advance the ball after receiving it from TeamA1. Moreover, TeamA1 attempted to shoot at the opposing team's goal multiple times (at 0:03, 0:07, 0:22, 0:36, and 2:00). These attempts are critical moments where the team tried to convert the ball possession into scoring opportunities. The fact that most of these shots were aimed directly at the opposing team's goalie or defense players highlights the offensive role of TeamA1. What's more, there are several instances where TeamA1 passed the ball to their own goalie (at 1:06, 1:19, 1:29, 1:36, 1:40). This indicates a strategy where the team resets or repositions itself during the game, to relieve pressure from the opposing team or to reorganize the play. From the strategic insights' perspective, TeamA1 is a central figure in gameplay, acting as a distributor which uses the passing action, and as an offensive threat which uses the shooting action. The role of TeamA1 is significant in creating scoring opportunities and maintaining ball control.

Source	**Target**	**Interaction**	**Timestamp**
TeamA1	TeamA2	Passing	0:01
TeamA1	TeamB3	Shooting	0:03
TeamA1	TeamB_goalie	Shooting	0:07
TeamA1	TeamA2	Passing	0:14
TeamA1	TeamB2	Shooting	0:22
TeamA1	TeamA2	Passing	0:27
TeamA1	TeamB2	Shooting	0:36
TeamA1	TeamA_goalie	Passing	1:06
TeamA1	TeamA3	Passing	1:14
TeamA1	TeamA_goalie	Passing	1:19
TeamA1	TeamA2	Passing	1:24
TeamA1	TeamA_goalie	Passing	1:29
TeamA1	TeamA_goalie	Passing	1:36
TeamA1	TeamA_goalie	Passing	1:40
TeamA1	TeamA3	Passing	1:46
TeamA1	TeamA3	Passing	1:55
TeamA1	TeamB_goalie	Shooting	2:00

Table 1: Interaction Data Between Different Team Players

In conclusion, this research contributes valuable insights into the ice hockey strategic dynamics. The method of object detection, pose detection, SNA can be utilized in multiple contexts within sport analytics, and extend to a variety of sports, particularly in fix-field sports, such as socker, basketball, American football, and volleyball. The ability to track and analyze player positions, detect poses, and uncover underlying social networks within the game can provide coaches with the understanding on how players move, collaborate, offense, and defense. Furthermore, understanding the patterns of player interactions can help coaches to design offensive and defensive strategies, optimize key players roles, refine player positioning and movements.

References

Aharon, N., Orfaig, R., & Bobrovsky, B.-Z. (2022). Bot-sort: Robust associations multi-pedestrian tracking. *arXiv preprint* arXiv:2206.14651.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment, 2008*(10), P1000.

Bruner, M., Mclaren, C., Mertens, N., Steffens, N. K., Boen, F., McKenzie, L., Haslam, S., & Fransen, K. (2021). Identity leadership and social identification within sport teams over a season: A social network analysis. *Psychology of Sport and Exercise, 59,* 102106.

Contributors, M. (2020). OpenMMLab's next generation video understanding toolbox and benchmark. Retrieved from https://github.com/open-mmlab/mmaction2

Duan, H., Wang, J., Chen, K., & Lin, D. (2022). Pyskl: Towards good practices for skeleton action recognition.

Duan, H., Xu, M., Shuai, B., Modolo, D., Tu, Z., Tighe, J., & Bergamo, A. (2023). Skeletr: Towards skeleton-based action recognition in the wild. In *Proceedings of the International Conference on Computer Vision (ICCV) 2023*.

Duan, H., Zhao, Y., Chen, K., Shao, D., Lin, D., & Dai, B. (2022). Revisiting skeleton-based action recognition. *arXiv preprint* arXiv:2104.13586.

Shi, L., Zhang, Y., Cheng, J., & Lu, H. (2019). Two-stream adaptive graph convolutional networks for skeleton-based action recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 12026–12035.

Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024). Yolov10: Real-time end-to-end object detection. *arXiv preprint* arXiv:2405.14458.

Wäsche, H., Dickson, G., Woll, A., & Brandes, U. (2017). Social network analysis in sport research: An emerging paradigm. *European Journal for Sport and Society, 14*(1), 1–28.

Yan, S., Xiong, Y., & Lin, D. (2018). Spatial temporal graph convolutional networks for skeleton-based action recognition. In *Thirty-second AAAI conference on artificial intelligence*.

Zhang, B., & Löwendahl, T. (2023). Exploring the dynamics of ice hockey strategies using yolov8 and gephi in sports education. In *The European Conference on Education 2023: Official Conference Proceedings*. IOP Publishing.

Zhao, Y., Li, Z., & Chen, K. (2020). A method for tracking hockey players by exploiting multiple detections and omni-scale appearance features. *Project Report.*