

Article

Not peer-reviewed version

Object and Event Detection Pipeline for Rink Hockey Games

J. M. Lopes , L. P. Mota , [S. M. Mota](#) , [José Manuel Torres](#) ^{*} , [Rui S. Moreira](#) , [Christophe Soares](#) , [Ivo Pereira](#) , [Feliz Gouveia](#) , [Pedro Sobral](#)

Posted Date: 4 April 2024

doi: [10.20944/preprints202404.0383.v1](https://doi.org/10.20944/preprints202404.0383.v1)

Keywords: Deep Learning; Rink Hockey; Objects and Events Detection; Artificial Intelligence; Computer Vision; YOLOv7; Team Sports





Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Object and Event Detection Pipeline for Rink Hockey Games

J. M. Lopes ¹, L. P. Mota ¹, S. M. Mota ¹, José Manuel Torres ^{1,2*} , R. S. Moreira ^{1,2},
Christophe Pinto de Almeida Soares ^{1,2} , I. Pereira ¹, F. R. Gouveia ^{1,2} and P. M. Sobral ^{1,2}

¹ ISUS Unit, Faculty of Science and Technology, University Fernando Pessoa, 4249-004 Porto, Portugal; csoares@ufp.edu.pt (C.P.d.A.S.)

² Artificial Intelligence and Computer Science Laboratory, LIACC, University of Porto, 4100-000 Porto, Portugal

* Correspondence: jtorres@ufp.edu.pt

Abstract: All types of sports are potential application scenarios for automatic and real-time visual object and event detection. In rink hockey, the popular roller quad skate variant of hockey team sports, it is of great interest to automatically track player's movements and positions, player's sticks and, also, making other judgments, such as being able to locate the ball. In this work, we introduce a real-time pipeline composed by an object detection model, created specifically for rink hockey games, followed by a knowledge-based event detection module. Even in the presence of occlusions and quick motions, our deep learning object detection model effectively identifies and tracks, in real-time, important visual elements such as: ball; players; sticks; referees; crowd; goalkeeper; and goal. Using a curated dataset composed by a collection of videos of rink hockey, comprising 2525 annotated frames, we trained and evaluated the algorithm performance and compare it to state of the art object detection techniques. Our object detection model, based on YOLOv7, presents a global accuracy of 80%, and presents a good performance in terms of accuracy and speed, according to our results, making it a good choice for rink hockey applications. In our initial tests, the event detection module successfully detected one important event type in rink hockey games, the occurrence of penalties.

Keywords: deep learning; rink hockey; objects and events detection; Artificial Intelligence; computer vision; YOLOv7; team sports

1. Introduction

The rink hockey sports, also known as quad hockey or roller hockey, is a thrilling and fast-paced sport that captivates players and spectators alike. Played on a rink with roller skates, it combines elements of ice hockey and roller skating to create an exhilarating game. Historically, this sport has been very popular in the following four countries: Portugal, Spain, Italy, and Argentina. As the sport evolves and becomes increasingly popular, the use of deep learning and other artificial intelligence techniques has emerged as a valuable tool to enhance various aspects of rink hockey, including game visualization, referee work, and training.

In the context of rink hockey, deep learning techniques can be applied to improve game visualization, allowing for more accurate and detailed analysis of player movements, strategies, and game dynamics. By leveraging computer vision algorithms, cameras placed strategically around the rink can capture high-resolution footage, which is then processed using deep learning models to track the positions and actions of players, the ball, and even the referees in real-time.

This enhanced game visualization has several benefits. First, it provides a valuable tool for coaches, allowing them to analyse and strategize more effectively. They can identify patterns in player movements, study tactical decisions, and gain insights into opponent behaviour. By analysing the data collected through deep learning techniques, coaches can make informed decisions and develop customized training programs to enhance their players' performance.

Furthermore, deep learning-based game visualization can also revolutionize the role of referees in rink hockey. Referees, in this games, can play an important role in ensuring fair play and enforcing the rules of the game. However, due to the fast-paced nature of rink hockey, it can be challenging

for referees to keep up with all the action and make accurate decisions in real-time. By incorporating deep learning techniques, referees can receive real-time assistance through computer vision systems. These systems can track players, identify fouls or infractions, and even provide instant feedback to the referees, thus reducing human error and improving the overall fairness of the game.

In addition to game visualization and referee work, these automatic techniques can also significantly impact player training in rink hockey. Training in this sport requires a combination of technical skills, physical fitness as well as tactical understanding. By leveraging deep learning algorithms, training programs can be personalized and optimized to meet individual player needs. Data collected from game visualization can be used to identify strengths and weaknesses, track progress, and provide targeted feedback to players. Coaches can use this information to design training sessions that address specific areas for improvement and enhance overall performance.

Moreover, deep learning can assist in injury prevention by analysing player movements and identifying potential risks. By monitoring players' performance data, coaches and medical staff can detect signs of fatigue, monitor load management, and make informed decisions to minimize the risk of injuries.

Since there is not much developed content relating to this sport, we decided to create a dataset and test these deep learning and computer vision improvements in the real world. Therefore, object detection, specifically employing the YOLOv7 powerful algorithm, can be used to follow the movement of the hockey ball and players on the rink, providing coaches, players, and fans with important information. In this paper, we will look at the advantages of employing this specific algorithm for object detection in rink hockey, as well as the possible impact on the sport itself.

The rest of the paper is organized as follows. Section 2 discusses the related work. Section 3 presents our approach for automatic analysis of rink hockey games and Section 4 presents the evaluation and results. Lastly, section 5 presents the conclusion and future work.

2. Computer Vision Applied to Hockey Sports

Rink hockey and ice hockey are two team sports that are similar in many ways, but there are also some key differences between them. One of the primary differences is the type of equipment utilized in each sport. Rink hockey is usually played on a dry surface, such as a concrete rink or a sports court, and the playing area is rectangular. Ice hockey, on the other hand, is played on an ice rink, which is typically oval. The field material is a key point that leads to variations in movement patterns and techniques, as players must adapt to different surfaces and levels of friction and stability. The participants use sticks to fire a hard ball/puck into the opposing team's goal while the game is played on a hard, smooth surface, like a basketball court or roller rink. Due to the players' quicker movement on roller skates, rink hockey moves along more swiftly than ice hockey. The regulations of the two sports vary in some ways as well. For instance, each team in rink hockey, in normal conditions, on the playing rink, has one goalkeeper and four field players. Each team in ice hockey typically has six players. Additionally, compared to rink hockey, the offside rule is often applied in ice hockey.

Latest innovations in object detection techniques have resulted in considerable increases in object detection accuracy and speed [1]. Convolutional Neural Networks (CNNs) represent one of the most extensively used techniques for object detection, having obtained state of the art results on diverse object detection benchmarks [2]. The substantial diversity in the appearance of objects due to factors such as lighting conditions, player uniforms, and background noise is one of the primary issues in rink hockey object detection. Researchers have proposed numerous approaches to overcome this difficulty, including employing multiple CNNs with different architectures [3], and using data augmentation techniques to increase the diversity of the training data [4].



(a) Ice Hockey Player



(b) Rink Hockey Players

Figure 1. Ice hockey vs Rink hockey (images: Wikimedia Commons).

In rink hockey, these techniques can be utilized for a variety of purposes, including player tracking, ball tracking, and game analysis. Unfortunately, when it comes to object detection in rink hockey, there is not much work done to date. This regards both the creation of a dataset, and the implementation of deep learning techniques for the associated predictions. However, in relation to ice hockey, there is a lot of content available such as tracking and identification of players, characteristics of the playing field, and the match analysis itself [5–7].

In [5] it is proposed a system to automatically track and identify players in National Hockey League (NHL) streamed videos. Player tracking, team identification, and player identification are the three components of the system. The authors experimented five state-of-the-art tracking algorithms [8–12] on the hockey player tracking dataset. The best tracking performance was achieved using the MOT Neural Solver tracking model [10] re-trained on the hockey dataset, with a Multi-Object Tracking Accuracy (MOTA) score of 94.5%. Additionally, for team identification, the away team jerseys were grouped into a single class, whereas the home team jerseys were grouped by jersey color. For the team identification dataset, a CNN was trained, reaching an accuracy of 97% on the test set. They also presented a novel player identification model that employs a temporal one-dimensional convolutional network to identify players from sequences of player bounding boxes. Using the available NHL game roster data, the player identification model achieved a player identification accuracy of 83%.

In [6] authors developed a cascaded CNN model, comprising two phases, for the detection of ice hockey players. In the same work the jersey color of each of the detected players is extracted in order to determine team affiliations. A filter is considered in the proposed model to exclude distracting information, such as the audience and sideline advertising bars, refining, this way, the detection of the targeted players, and resulting in an accurate detection with a precision of 98.75% and a recall of 94.11% for individual players and an average accuracy of 93.05% for team classification using a dataset of collected images from the 2018 Winter Olympics. The authors argue that, With the custom-built dataset, their player detection model achieves best results when compared to some state-of-the-art approaches such as YOLOv3 [13], Faster R-CNN [14] and CornerNet [15].

Another difficulty in rink hockey object detection is the requirement for real-time performance, as the sport is fast-paced, and object detections must be updated regularly. Researchers have come up with a number of solutions to this problem, including the use of lightweight CNN architectures [16], and the implementation of the object detection pipeline on specialized hardware like graphics processing units (GPUs) [17] or field-programmable gate arrays (FPGAs) [18].

In [16] it is presented a CNN with three branches and a classification network with four cascades. The cascaded networks were initially trained using labeled image patches and applied to an entire image by employing a dilation testing strategy. The authors argue that their approach achieved state-of-the-art accuracy on three types of games (basketball, soccer, and ice hockey) with 1000 fewer

parameters compared to CNNs that are adapted from general object detection networks such as Faster-RCNN.

The work presented in [17] proposed an approach to locate the puck, the key part of an ice hockey game. The small size of the puck is one of the major obstacles. The motion blur caused by the puck's rapid movement, the occlusion between the puck and other objects, and the visual noises (such as the advertisements on the rink) presented additional difficulties for the project. The author designed a two-stage model for detecting minute objects. The first stage is a two-dimensional CNN that summarizes each frame's representation. In the second stage, they stack and feed the fusion of these per-frame representations into a three-dimensional CNN in order to decode the video's temporal information. The proposed approach achieved a Precision of 90.8% and a Recall of 86.7%. The F1 score reached 88.7%, which is higher than the performance of YOLOv3 [13] (F1 score = 0.685) and Mask RCNN [19] (F1 score = 0.749).

In [18] it is presented a hardware accelerator which implements a YOLO CNN for real-time object detection with high throughput and power efficiency. The parameters of the implemented YOLO were retrained and also quantized using binary weight and flexible low-bit activation. The binary weight technique allows storing the whole network model in block RAMs of a field-programmable gate array (FPGA). This way it is reduced the off-chip accesses and is achieved a performance boost. All the convolutional layers have been pipelined for optimize the hardware utilization. Also, the input images are delivered and processed line by line to the hardware accelerator. Similarly, the output of the previous layer is transmitted line-by-line to the subsequent layer. The intermediate data are fully reused across layers, eliminating memory accesses from the outside. With less DRAM accesses it is reduced the power consumption. In addition, since the convolutional layers are fully parameterized, authors argue that it is simple to scale the network. Finally, each convolution layer is mapped to a block of dedicated hardware, making it to outperform other solution designs in energy efficiency and performance.

3. Automatic Pipeline for Object and Event Detection in Rink Hockey Games

One major contribution of this work is the specification, implementation and test of an automatic pipeline capable of detecting and track players, ball, and other in-game objects in real-time during fast-paced rink hockey games as well as some important events, as can be seen in Figure 2.

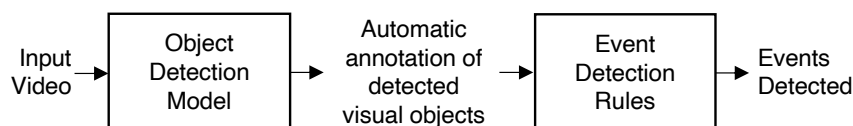


Figure 2. Pipeline for real-time detection of Hockey Objects and Events.

That said, the pipeline was designed to meet the following requirements:

- Object detection capabilities for player tracking, ball tracking, and video analysis.
- Real-time performance for updating object detections regularly during fast-paced games.
- Ability to handle diversity in the appearance of objects due to factors such as lighting conditions, player uniforms, and background noise.
- High accuracy in object detection.
- High speed response in object detection.
- Compatibility with specialised hardware such as GPUs.
- Adaptability to different environments and lighting conditions.
- Capability to detect and represent major events based on the stream of visual objects

3.1. Dataset Organisation

A major contribution of this work is the organisation of the dataset. Our rink hockey dataset that was annotated with the aid of a web-based computer vision annotation tool which helps to label video

and images. The selected solution, Roboflow, is a popular computer vision development platform that facilitates data collection, preprocessing, and model training. This framework allows easy access to public datasets as well as the ability for users to upload their own custom data [20]. Additionally, it supports various annotation formats including JSON, CSV, and XML.

From the total of 2525 annotated frames/images, the image augmentation technique was applied which, by applying horizontal flipping, cropping with a zoom up to 20% and brightness changing between -15% (darker) and +15% (brighter), resulted in a total of 6053 images. Beyond that, through these and other possible modifications, we artificially increased the size of the dataset. This technique is used not only to increase the size of the dataset and hence the robustness and generalization of the model, but also to try to avoid model overfitting, where a trained model performs better on trained data but poorly on unseen data.

The dataset was divided into 3 subsets: train (70%), validation (20%) and test (10%). The dataset included objects labelled with 7 classes: ball; player; stick; referee; crowd; goalkeeper; goal. The main goal of this annotation process was to provide a dataset that can be used to train object detection models, such as YOLOv7 [21], for transfer learning.

As Figure 3 shows, the dataset is slightly unbalanced, with the player and stick classes being the most represented. This is not surprising since these are the most visible objects in all images.

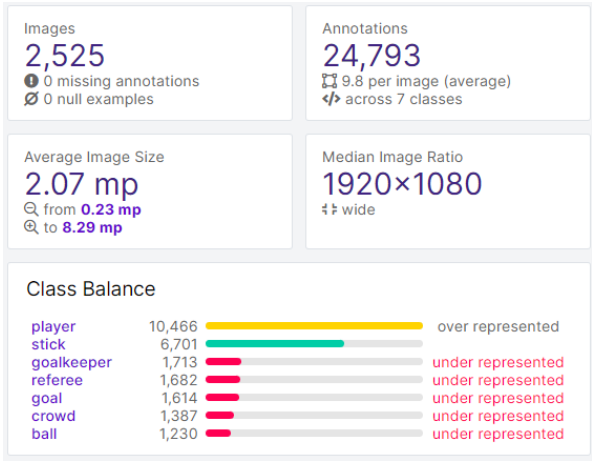


Figure 3. Dataset Health Check.

The resulting dataset is available online¹. Definitely, it can be useful for a wide range of applications, including player tracking and video analysis of rink hockey games. Being publicly available this dataset will serve as a valuable resource for other researchers and practitioners within the field of computer vision and sports analysis.

3.2. Object Detection Model

Yolov7

YOLO (You Only Look Once) [21] is a state-of-the-art object detection system that is highly valuable in the domain of computer vision. In contrast to traditional object detection systems that run a classifier on various parts of the image independently, YOLO innovatively applies a single neural network to the entire image at once. The image is divided into regions, and then the system predicts bounding boxes and probabilities for each region. As such, YOLO significantly enhances the efficiency and speed of real-time object detection, making it a pivotal asset in a wide range of applications. Specifically in the context of hockey, YOLO can be effectively utilized to track the movements of

¹ <https://universe.roboflow.com/visao-computacional/roller-hockey>

players and the puck on the arena. This can furnish extensive insights for the analysis of gameplay, performance review of players, and the development of strategic gameplay. It is noteworthy that YOLO is not a standalone tool, but rather, it refers to a type of neural network architecture and algorithm. It is commonly implemented using comprehensive deep learning libraries, such as TensorFlow or PyTorch.

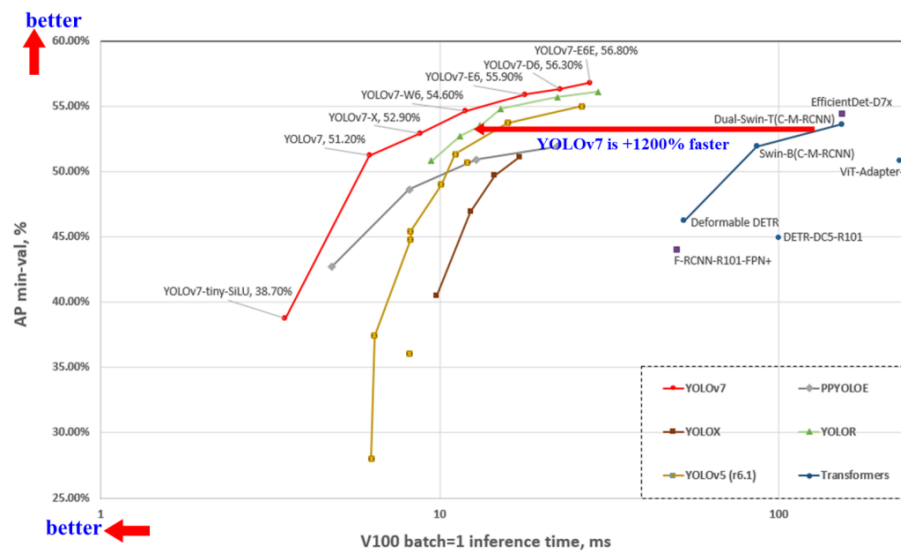


Figure 4. Comparing YOLOv7 with other object detectors [22].

The authors in [23] present another version of YOLO, the YOLOv7. Its base architecture and version are significantly better than prior YOLO versions, such as YOLOv4 and YOLOv5, and it has reached state of the art performance on numerous object detection benchmarks, such as the COCO dataset [24]. The use of a "bag-of-freebies" is one key feature of YOLOv7. This refers to a set of training strategies that do not require the usage of additional computational resources and may thus be utilised to increase the performance of the model without significantly increasing its complexity. Data augmentation, network architecture changes, and multiscale training are examples of these strategies. By using these techniques, this version is able to detect objects at many scales, resulting in better results on object detection tasks while still running in real-time on standard hardware. Overall, apart from its high precision object detection capabilities, YOLOv7 is also known for its efficiency and speed (Figure 4), delivering outstanding performance without losing speed or simplicity, making it a valuable tool for a wide range of applications.

Object Detection for Rink Hockey

In this section, we present the technical details of the developed² object detection system for rink hockey utilising YOLOv7. To facilitate the training of the model, we constructed a dataset composed of various rink hockey scenarios, which were carefully annotated to include different perspectives, lighting conditions, data augmentation and object scales. The YOLOv7 architecture was implemented using the PyTorch framework [25], and the pre-trained weights were fine-tuned on the dataset we were training. Our experiments demonstrated that the model achieved high accuracy in detecting rink hockey objects such as the crowd, goalkeeper, and players. For real-time object detection, we used multiple online videos from professional rink hockey games [26–35] to acquire the feed and processed the frames using the trained YOLOv7 model. Our results demonstrate the capacity of our model to detect visual objects in the video with low latency, thereby making it suitable for real-time applications. Additionally, we leveraged CUDA and cuDNN [36] for GPU acceleration, which

² <https://github.com/LuisMota1999/ProjetoVC>

significantly improved the processing speed of the model. The conjunction with various technologies such as CUDA, OpenCV, and PyTorch has enabled the development of a robust and efficient object detection system for rink hockey.

3.3. Event Detection Rules Module

Event detection follows the object detection in the pipeline showed in fig. 2. The proposed specification of an event detection module, whose input is a timeline stream of detected visual objects, the output of the object detection model, can comprise up to four stages as follows:

1. Filtering: due to the noisy nature of object detection model a filtering operator, such as a windowed spatio-temporal median filter, can be used as a pre-processing stage before the application of a rule based detector
2. Model-based Rule Detection: rule-based system that can detect predefined types of events
3. Representation: the language or taxonomy used to represent the events. Event calculus is a formal way of event representation [37].
4. Revision: can be used to revise or update represented event beliefs due to new information available in the stream of detected visual objects

This module follows the object detection module and, consequently, depends on it according to the directed acyclic graph (DAG) presented in fig. 2. In the preliminary tests performed, a proof-of-concept approach was adopted for this module and the detection of a unique type of event was considered for this article.

4. Experiments and Evaluation

Object Detection

In our experiments we trained a YOLOv7 based model for object detection using footage acquired, essentially, from professional rink hockey games. We performed both qualitative and quantitative evaluations to assess the performance of this model.

The evaluation process consisted of applying YOLOv7 to detect objects of interest in multiple video frames from rink hockey games [38,39] and comparing the results to manually annotated ground truth labels. We used precision (p), recall (r), and F1-score (F_1), the harmonic mean of precision and recall, as the three quantitative metrics to evaluate the performance of the system, according to the following equations:

$$p = \frac{TP}{TP + FP} \quad (1)$$

$$r = \frac{TP}{TP + FN} \quad (2)$$

$$F_1 = 2 * \frac{p * r}{p + r} \quad (3)$$

The model was trained using 100 epochs and a batch size of 32 with all the system being implemented on a Windows desktop with a Zotac Gaming GeForce RTX 3090 Trinity OC 24 GB GDDR6X GPU and 64 GB of RAM, using the Darknet framework in Python.

When testing on frames not seen in the training phase (present in the test subset), the overall performance was considerably good (around 80%).

The results of the evaluation are presented in the following figures. Figure 5 presents a multi-frame inferred plot on the test subset, showing the labels that the trained model detected. Figure 6 represents the trade-off between precision and recall in the classification model. As can be seen from the graph, the precision-recall values are higher for the player, goalkeeper and goal classes, as we had more annotated images for these classes. However, the precision-recall for the ball is lower as we have fewer annotated images for this class.



Figure 5. Multi-frame (4×4=16 frames) inferred plot on the test subset. Upper-left frame is in coordinate (row, col)=(1,1).

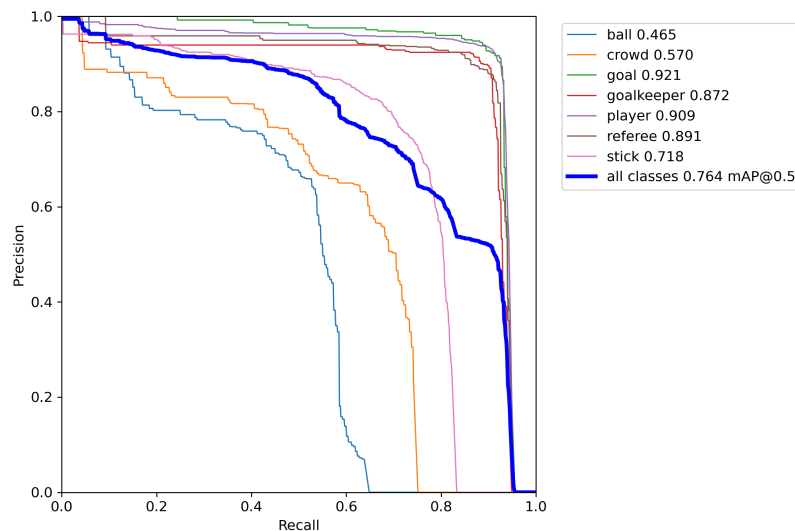


Figure 6. Precision-Recall Curves.

It is important to note that in our evaluation, we only considered detections with a objectness confidence score above 0.3. This threshold was chosen to balance the trade-off between detecting more objects while maintaining a slightly higher precision.

Also, as seen in Figure 6, our trained model got a score of 0.764 mAP@0.5 and, in this context, 0.764 represents the mAP score, i.e., the mean Average Precision (the average of AP over all detected classes, calculated for each class separately) ranging between 0 and 1. The parameter 0.5 represents the IoU (Intersection over Union) selected threshold, in other words, the model is able to detect objects with an overlap of at least 50% with the ground truth. Having that said, in the context of this work, a mAP score of 0.764 is a fairly good score for object detection models.

In follow-up to what was previously said, we can also use the F1-score to evaluate the model performance. The F1-score is a metric that is calculated as the harmonic mean of precision and recall. This metric also has a range of 0 to 1, with 1 representing perfect precision and recall and 0 representing the worst possible value. Beyond that, the precision reflects the number of samples that were predicted

as positive and are actually positive (ground truth) and the recall measures the number of actual positive samples that were predicted correctly as positive. As shown in the Figure 7, the model has a maximum F_1 score of 0.79 at 0.240. Analysing this graph reveals that there is a large plateau with $F_1 > 0.75$ approximately parallel to the confidence axis, which can be considered as a very good indicator of the model behavior.

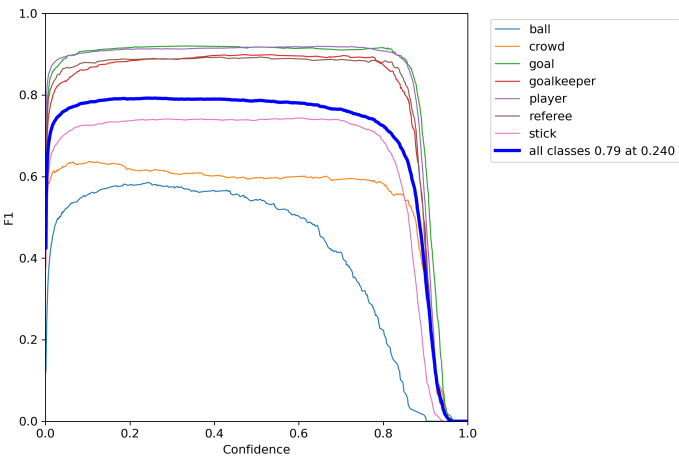


Figure 7. F_1 Curve.

The metrics F1-score, precision and recall are all derived from the values of the confusion matrix which is another useful tool for understanding the performance of a classification model, since it shows where the model is making mistakes and how it performs on different classes. In this case, as can be seen in the normalized confusion matrix obtained Figure 8, when evaluating the test subset, in other words, the images that were not seen by the model while training, we got an accuracy of 80%. This figure also provides the information that the ball and crowd classes were the ones with worst results.

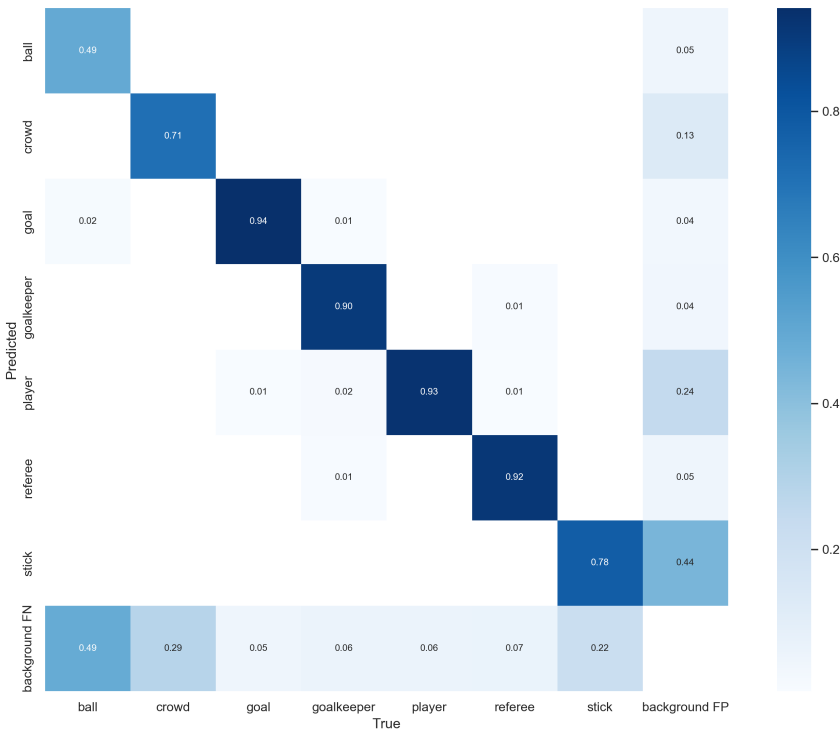


Figure 8. Confusion Matrix.

Event Detection

For the rule production system used for event detection, one rule, illustrated in Algorithm 1, was tested to detect the occurrence of penalties or direct free hits, according to the existing official rink hockey regulations [40].

Algorithm 1 direct_free_hit_or_penalty

```

if  $T(isReferee() \wedge isGoal() \wedge isGoalkeeper() \wedge countPlayer() == 1, t_1, t_2)$  then
  return EVENT_DIRECT_FREE_HIT_OR_PENALTY
end if

```

This rule considers several predicates. The predicate $T(expression, t_1, t_2)$ is true if *expression* is true for the temporal interval between t_1 and t_2 . Remaining predicates are auto explicated. Predicates starting with prefix *is* meaning the presence of, at least, one visual object in that frame for that class. Predicates starting with prefix *count* return the total number of visual objects, in that frame, for that class.

According to the current regulation [40], the player taking the penalty or the direct free hit has a maximum of five seconds to, after the indication of the Main Referee, start the execution with the ball stopped. Thus, in this rule, temporal window duration, $\Delta t = (t_2 - t_1)$ should be appropriately chosen. In our experiments, we set $\Delta t = 4s$.

The penalty and direct free hit sanction executions are, probably, the most important ones in rink hockey. Thus, their detection is rather important as an event. In figure 5, (row, col)=(3,4), is illustrated an example of a penalty event, during the game, that was correctly detected by this rule.

5. Conclusion

Our investigation aspires to evolve the field of research in object and event detection in rink hockey sports since it has the possibility to significantly improve the analysis and understanding of this sport where games take place at a fast-pace.

The use of object detection techniques has been instrumental in achieving the track of the objects in rink hockey videos. However, the diversity in the appearance of objects and the need for real-time performance remain a significant challenge in this field.

This research used a cutting-edge object detection method for rink hockey with the YOLOv7 algorithm followed by knowledge-based event detection module. This approach has shown a good object detection performance in terms of precision, even in the presence of occlusions and quick motions.

For future work, we would point out that it is essential to enlarge the dataset with more annotated data, in order to balance the dataset, and compare it with other algorithms since it's important to keep testing and comparing with other methods and techniques. A final note for the practical application of this work since collaborations are possible with roller hockey clubs and hockey federations to develop and test this approach and enable the continuous evolution of this sport.

References

1. Zou, Z.; Chen, K.; Shi, Z.; Guo, Y.; Ye, J. Object detection in 20 years: A survey. *Proceedings of the IEEE* **2023**.
2. Li, Z.; Liu, F.; Yang, W.; Peng, S.; Zhou, J. A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE transactions on neural networks and learning systems* **2021**.
3. Srivastava, S.; Divekar, A.V.; Anilkumar, C.; Naik, I.; Kulkarni, V.; Pattabiraman, V. Comparative analysis of deep learning image detection algorithms. *Journal of Big Data* **2021**, *8*, 1–27.
4. Shorten, C.; Khoshgoftaar, T.M. A survey on image data augmentation for deep learning. *Journal of big data* **2019**, *6*, 1–48.
5. Vats, K.; Walters, P.; Fani, M.; Clausi, D.A.; Zelek, J.S. Player tracking and identification in ice hockey. *Expert Systems with Applications* **2023**, *213*, 119250.

6. 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.
7. Sousa, T.; Sarmento, H.; Harper, L.D.; Valente-Dos-Santos, J.; Vaz, V. Match analysis in rink hockey: a systematic review. *Human Movement* **2022**, *23*, 33–48.
8. Bergmann, P.; Meinhardt, T.; Leal-Taixe, L. Tracking without bells and whistles. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 941–951.
9. Bewley, A.; Ge, Z.; Ott, L.; Ramos, F.; Upcroft, B. Simple online and realtime tracking. 2016 IEEE international conference on image processing (ICIP). IEEE, 2016, pp. 3464–3468.
10. Brasó, G.; Leal-Taixé, L. Learning a neural solver for multiple object tracking. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 6247–6257.
11. Wojke, N.; Bewley, A.; Paulus, D. Simple online and realtime tracking with a deep association metric. 2017 IEEE international conference on image processing (ICIP). IEEE, 2017, pp. 3645–3649.
12. Zhang, Y.; Wang, C.; Wang, X.; Zeng, W.; Liu, W. Fairmot: On the fairness of detection and re-identification in multiple object tracking. *International Journal of Computer Vision* **2021**, *129*, 3069–3087.
13. Redmon, J.; Farhadi, A. YOLOv3: An incremental improvement. *arXiv preprint arXiv:1804.02767* **2018**.
14. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems* **2015**, *28*.
15. Law, H.; Deng, J. Cornernet: Detecting objects as paired keypoints. Proceedings of the European conference on computer vision (ECCV), 2018, pp. 734–750.
16. Lu, K.; Chen, J.; Little, J.J.; He, H. Lightweight convolutional neural networks for player detection and classification. *Computer Vision and Image Understanding* **2018**, *172*, 77–87.
17. Yang, X. Where is the puck? Tiny and fast-moving object detection in videos. Master's thesis, McGill University, 2021.
18. Nguyen, D.T.; Nguyen, T.N.; Kim, H.; Lee, H.J. A high-throughput and power-efficient FPGA implementation of YOLO CNN for object detection. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* **2019**, *27*, 1861–1873.
19. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask r-cnn. Proceedings of the IEEE international conference on computer vision, 2017, pp. 2961–2969.
20. Roboflow. Roboflow Universe. <https://universe.roboflow.com>, 2023. Accessed: 2023-05-01.
21. Terven, J.; Córdova-Esparza, D.M.; Romero-González, J.A. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *Machine Learning and Knowledge Extraction* **2023**, *5*, 1680–1716. Publisher: Multidisciplinary Digital Publishing Institute, doi:10.3390/make5040083.
22. Kin-Yiu, Wong. YOLOv7 vs YOLOv5 reproducible comparison of accuracy and speed #395. <https://github.com/WongKinYiu/yolov7/issues/395>, 2022. Accessed: 2023-05-01.
23. Wang, C.Y.; Bochkovskiy, A.; Liao, H.Y.M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv preprint arXiv:2207.02696* **2022**.
24. Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft COCO: Common Objects in Context. *Computer Vision – ECCV 2014*; Fleet, D.; Pajdla, T.; Schiele, B.; Tuytelaars, T., Eds.; Springer International Publishing: Cham, 2014; pp. 740–755.
25. Paszke, A.; Gross, S.; Chintala, S.; Chanan, G.; Yang, E.; DeVito, Z.; Lin, Z.; Desmaison, A.; Antiga, L.; Lerer, A. Automatic differentiation in PyTorch. NIPS-W, 2017.
26. Porto, F. Vitória ÉPICA do Hóquei em Patins do FC Porto (Resumo). <https://youtu.be/JRe0iIi55Ig>, 2022. Accessed 19 January 2023.
27. D'hoquei, S. Resum del GSH Trissino vs Credit Agricole Sarzana. https://youtu.be/k3cFN_Y2prc, 2020. Accessed 19 January 2023.
28. D'hoquei, S. Highlights Portugal vs France. <https://youtu.be/Jbjut3NvCGg>, 2022. Accessed 19 January 2023.
29. TV, F. Hóquei em Patins: Física x CRIAR-T - 25 de Abril 18h30. <https://youtu.be/YZMHSuskCg>, 2022. Accessed 19 January 2023.
30. TV, E.R.H. FC Porto x SL Benfica. <https://youtu.be/jnFRrurson0>, 2018. Accessed 19 January 2023.
31. TV, E.R.H. FC Porto (PT) x Sporting CP (PT). <https://youtu.be/uIsaU7ME8h4>, 2018. Accessed 19 January 2023.

32. Barcelona, F. [ESP] LIGA EUROPEA Hockey patines: FC Barcelona Lassa - FC Oporto (3-1). <https://youtu.be/zGXZaarw5BE>, 2017. Accessed 19 January 2023.
33. TV, E.R.H. Euroleague - Barça (SP) x SL Benfica (PT). <https://youtu.be/yTe1Hv9OV-w>, 2020. Accessed 19 January 2023.
34. TV, E.R.H. BENFICA-BARCELONA, 1st Semi-Final of Rink Hockey Euroleague 2015-16, played at Pavilhão. <https://youtu.be/hTNZm1CGww4>, 2016. Accessed 19 January 2023.
35. TV, E.R.H. Match #249 - Spain x Portugal [HD]. <https://youtu.be/fKDWzskg2pE>, 2021. Accessed 19 January 2023.
36. NVIDIA. NVIDIA CUDA Deep Neural Network library (cuDNN). <https://developer.nvidia.com/cudnn>, 2023. Accessed: 2023-09-01.
37. Shanahan, M. The Event Calculus Explained. In *Artificial Intelligence LNAI* **2000**, 1600. doi:10.1007/3-540-48317-9_17.
38. D'hoquei, S. Highlights SL Benfica vs Sporting CP. <https://youtu.be/CS4P8hkM8vk>, 2022. Accessed 19 January 2023.
39. TV, E.R.H. Extended Highlights - Match #243 - France x Portugal [HD]. <https://youtu.be/S0fSwSclAsE>, 2022. Accessed 19 January 2023.
40. Commission, R.H.T. Rink Hockey Official Regulation. Technical report, World Skate, 2020.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.