# A Deep Learning Network for Detecting Head Impacts in Ice Hockey from 2D Game Video

Amirhossein Azadi, Parisa Dehghan, Ryan Graham, T. Blaine Hoshizaki

***Abstract*** This paper presents a deep learning approach for detecting head impacts in ice hockey from 2D game videos. Traumatic brain injuries, including concussions and repetitive head impacts, pose significant health risks in sports. Understanding the relationship between head impact features (magnitude and frequency) and outcomes such as mental health decline, cognitive deficits, and Chronic Traumatic Encephalopathy (CTE) requires extensive datasets. Tracking head impacts during games is challenging, and available tools are impractical for most leagues due to cost and equipment constraints. Utilising game videos for head impact detection offers a viable solution. A methodology combining player detection and tracking with a Long-term Recurrent Convolutional Network (LRCN) for head impact detection is proposed. Our player detection model achieved high precision and recall scores, facilitating accurate tracking. The YOLOv8x object detection model yielded precision, recall, mAP50, and mAP50-95 scores of 0.97, 0.97, 0.99, and 0.95, respectively. The StrongSORT tracking algorithm used for player tracking minimised ID switches, important for precise tracking in dynamic sports environments. The LRCN-based head impact detection model showed promising results, with an accuracy of 87% and a loss of 0.04. Future work involves refining dataset creation to address data imbalance and exploring alternative deep learning models like CONVLSTM and 3DCNN for improved performance.

***Keywords*** Deep Learning, Head Impact Detection, Ice Hockey, Object Tracking, Traumatic Brain Injuries

## I. INTRODUCTION

Scientists continue to undertake research to better understand the complex relationship between head impact severity and brain injuries, particularly Traumatic Brain Injuries (TBIs) such as concussions in sport [1]. The brain's complexity and individual factors like age, genetics, and health conditions pose challenges to understanding brain injuries. Traumatic brain injury (TBI) is a heterogeneous injury resulting from various physical forces, leading to a spectrum of neurological and psychiatric symptoms. Concussion, despite its *mild* (mTBI) classification, can have severe consequences if not treated properly or if individuals experience repeated mTBIs. Unlike more severe forms of TBI, mTBIs often do not result in any visible brain abnormalities like bleeding. Instead, they manifest through functional disruptions, leading to symptoms such as dizziness, nausea, and cognitive disturbances, which can lead to a wide range of neurological and psychiatric symptoms, both acute and chronic [2-4]. Annually, over 500,000 adolescents in the United States alone experience mTBIs, with many of these injuries subclinical and may contribute to accumulated brain damage from repetitive head impacts (RHI), and risk for neurodegenerative disorders [5]. RHI refers to the repeated exposure to head impacts (clinical and subclinical) over time, often experienced by athletes participating in contact sports like American football and ice hockey. Head impact magnitude and frequency play important roles in understanding and assessing the potential risks associated with brain trauma. Factors that affect the magnitude of impact include the velocity, head location, and the type of impact event (e.g., shoulder collision, fall to ice). Impact frequency can lead to cumulative effects on the brain, emphasising the importance of understanding both single and repeated head impacts on athletes' health [2-3].

The need to understand the link between brain trauma, especially repetitive traumatic brain injury, and neurological health has led to the emphasis on large, validated datasets [13] [17]. Objectively capturing both impact magnitude and frequency are challenging and the current available methods using head impact sensors or laboratory reconstructions both present with characteristics that limit their ability to create datasets large

Amirhossein Azadi is a PhD candidate (Amirhossein.azadi@uottawa.ca), Parisa Dehghan is a PhD candidate, Ryan Graham is an Associate Professor, and T. Blaine Hoshizaki is a Professor of Human Kinetics in the Department of Health Sciences at the University of Ottawa in Canada.

enough to disentangle intricate relationships [8-10] [11-15]. With the advent of big data and the integration of artificial intelligence (AI) in medical research, new possibilities have emerged for investigating these complex relationships in extensive, heterogeneous datasets [18-19]. Computer vision, is a branch of AI, that enables machines to process and interpret visual data. Its utility in sports includes several applications including game analytics. [18-23]. The application of computer vision in sports has led to most sports leagues and organizations filming their games for all age groups and competition levels leading to the wide accessibility of game footage. Automated methods using computer vision for measuring brain trauma and track head impacts is necessary for creating large data sets. In this regard, deep learning algorithms are key to automating the detection of head impacts from video footage.

Human Action Recognition (HAR) in sports is an important technological application that employs computer vision techniques to identify and understand player actions. This process is invaluable across various sporting events, including warm-ups, specific training, or competitions, as it assists coaches, medical staff, and journalists by enhancing performance analysis, preventing injuries, and accumulating match statistics [25]. The potential of HAR extends over a range of sports, from individual activities like skiing and swimming to team sports such as ice hockey, basketball, soccer, and volleyball [25]. One of the applications of HAR is in detecting a head impact event [24]. Recognising an event or action from 2D video is challenging due to the variations in pixel intensity, camera views, and complex patterns. Players might change their position, block each other from the camera, enter and exit the camera's field of vision, and operate at different angles [25]. These impacts in sports are brief and associated with distinct pre-impact reactions, player positions during the impact, and immediate post-impact reactions. Consequently, traditional methods of HAR often fail to capture the features due to differing illumination, scale, posture, perspectives, and complex human body movements during an activity like head impact [24] [26-28]. To address some of these challenges it would be advantageous to incorporate computer vision models that provide both spatial and temporal information to a HAR application of identifying head impacts from game footage. Spatial features are related to the visual attributes and pixel arrangements within an image, while temporal features pertain to time-based characteristics and changes over a sequence of frames.

Various deep learning architectures have been proposed to recognise human action from video. Long-term Recurrent Convolutional Networks (LRCN) combine Convolutional Neural Networks (CNNs) with Long Short-Term Memory (LSTM) networks for sequence learning tasks [38]. However, they can be computationally intensive and prone to overfitting [34]. Convolutional Long Short-Term Memory (ConvLSTM) networks use convolutional operations within LSTM units, making them suitable for spatial-temporal data but leading to increased computational complexity and larger model sizes [39]. Temporal Segment Networks (TSNs) segment videos and sample brief clips, however they have been shown to miss critical actions and struggle to integrate temporal information effectively [40]. The Temporal Shift Module (TSM) optimises temporal modelling by shifting channels in a convolutional network but can have limitations in capturing longer temporal dependencies [24] [36]. Additionally, pose estimation enhances human activity recognition by providing skeletal representations. However, given the nature of ice hockey, where players frequently obscure each other from the camera's view and head impacts occur within the brief span of 1 or 2 frames, the method of pose estimation is often ineffective in many parts of the game [41-43].

The purpose of this study was to develop, train and validate an activity detection network to detect head impacts from 2D videos in ice hockey games. This was accomplished by employing spatial and temporal techniques in computer vision. These sophisticated algorithms can be used to create big datasets to facilitate research to investigate the relationship between impact forces and brain injuries [13] [17]. Additionally, this method could be used to improve the management of brain injuries in sports by enabling real-time monitoring and assessment [24].

## II. METHODS

The process of creating and validating an activity detection network is comprised of multiple steps including player detection, player tracking, video segmentation, data sets preparation and head impact detection.

### *Player Detection*

To detect players in the visual field, methods in object detection were used to identify players on the ice. Specifically, the YOLOv8x model, the latest version of the YOLO (You Only Look Once) object detection algorithm, was compiled using the Adam optimiser, and categorical cross-entropy loss function [47-48]. This model was chosen as it treats object detection as a single regression problem, allowing it to make predictions with a single network evaluation, which significantly increases its speed compared to methods that require multiple evaluations for different image regions [49]. The dataset employed in this task consisted of 80,000 annotated images extracted from 25 National Hockey League (NHL) game videos [31]. The dataset was split into a training set comprising 65,000 images and a validation set of 15,000 images. The validation set was solely utilised to evaluate the performance of the player detection model, and the output of the network on this dataset determined the results. Due to the frequent occlusion events in ice hockey, the object detection model was trained with multiple labels, including player, goal keeper, and referee, along with their corresponding jersey colour and occlusion statuses (either occluded or not-occluded), providing essential information for subsequent tracking analysis (details of label selection will be discussed in the tracking section). Annotating the images was accomplished by drawing bounding boxes around each object and assigning its appropriate label. The YOLOv8x object detection framework underwent training for 250 epochs, and the image resolution was 1280x720 pixels. The performance of player detection was evaluated using presicion, recall and Mean Average Precision (mAP) of 50 and 50-90. The mAP was used to calculate the average precision (AP) for each class and then take the mean over all classes. 50-90 mAP refers to the mAP calculated at different Intersection over Union (IoU) thresholds, specifically at 50% and 90%. An IoU, a commonly used metric in computer vision, calculated the overlap between a predicted bounding box (the one output by an object detection model) and a ground truth bounding box (representing the actual location of the object in the image). If there is no overlap between two bounding boxes, the IOU cost is 1, and if the two bounding boxes have 100% overlap, the IOU cost is 0.

***Player Tracking***

Our tracking algorithms employed the Non-Maximum Suppression (NMS) method to eliminate redundant or overlapping bounding boxes of detections [32]. The player detection output consists of bounding boxes in each frame, some of which may overlap. This overlap can cause issues in player tracking, as the tracking algorithm may mistakenly treat overlapping boxes as the same player and delete duplicates. However, each box represents a distinct player. To address this, we propose changing the labels of the bounding boxes from a single label (Player) to different labels (P: Player, KP: Goalkeeper, Ref: Referee, OC: Occlusion, NOC: Non-Occlusion, and the jersey colour). This change allows the tracking method to treat each bounding box as a separate class, preventing the deletion of any box. The NMS process was employed to ensure that no bounding box is missed during occlusion events, thereby improving our player tracking accuracy. To continuously track the players throughout the video, the StrongSORT object tracking algorithm, that combines deep neural networks with traditional computer vision techniques, was employed in this study [32]. This was chosen as this algorithm improves tracking accuracy and robustness in challenging scenarios by using deep neural networks to extract features robust to appearance changes, such as illumination, viewpoint, and occlusion. Additionally, StrongSORT utilises a Kalman filter to model target object motion and estimate its position in subsequent frames [32]. At each frame, StrongSORT integrated both the detected players from the current frame and predictions of those players which was derived from the 2-3 previous frames. Utilising the feature embeddings of the detections and the predictions, along with considering motion costs, a cost matrix is constructed by StrongSORT [32], meaning linear assignment techniques are employed to link detections to their corresponding predictions correctly. Feature embeddings are high-dimensional vectors that represent the detected objects' attributes, such as appearance, shape, and other visual characteristics. This function measured the similarity between the visual features of track bounding boxes and the current frame detection bounding boxes. However, in ice hockey, where players from each team often share similar appearances due to identical jerseys and helmets, the default cost function in StrongSORT is not robust enough to differentiate between multiple bounding boxes that are similar in appearance. For this reason, an adapted cost function using IoU was incorporated into this phase. The IoU cost was incorporated to provide additional information about the location of track and detection bounding boxes that helps distinguish between multiple similar bounding boxes, and decreasing the likelihood of ID switches between frames, especially in occlusion events. To evaluate player tracking, a series of 15 videos, each lasting 10 seconds were used, and the algorithm was run to test ID switches, referring to the number of times the tracking algorithm assigns a different identity (ID) to the same object or player throughout the frames. The StrongSORT algorithm, employing its default cost function, was compared with those obtained using the adapted IoU cost function.

Fig. 1. Sample of whole frame analysis (left) using the entire image frame and player analysis approach (right) focusing on a narrowed visual of an individual player.

***Video Segmentation***

To prepare an input dataset for the activity detection network, a preprocessing phase of video segmentation was performed. A player analysis technique was chosen as it narrows its focus specifically to individual players as opposed to a whole frame analysis that uses all the pixels in the frames as input to the network (Figure 1). The advantage is that irrelevant pixels from other parts of the frame are excluded, increasing the accuracy of our network at identifying when the activity has occurred and specify the individual players involved [29]. This computer vision technique was used to divide long video into smaller sets of frames of 50 frames, and analyse each set independently, allowing for a finer temporal resolution. This means that the network can more precisely determine the timing of head impacts. This approach was used to detect and track each player individually within every frame and segment. The result is that short video clips for each player during every video segment was created. Segmenting the video and using a player analysis approach provides information about which player experienced a head impact and at which frame it occurred.

***Head Impact Dataset***

A dataset was developed to train the head impact activity detection networks. Head impacts documented through video analysis from the National Hockey League (NHL) game videos from the 2016-17 season were utilised [46]. The identified head impact events were cropped around the player area for 50 frames, ensuring that the head impact frame was the 25$^{th}$ frame. A total of 150 head impact events categorized by event type such as head-to-board, head-to-elbow, head-to-foot, head-to-glass, head-to-head, head-to-shoulder, head-to-puck/stick, head-to-ice, head-to-gloves, were included. To account for a small sample size in head impact events, we increased our data set using augmentation, where multiple videos can be created for each head impact event by extracting sequences of frames from various starting points. Two-25 frame videos were extracted from each video, from frames 5 to 30 and 20 to 45, with a frame size of 300x300. Additionally, each of these videos were flipped to augment the number of head impact events. For this study, 750 non-head impact videos were used. Non-head impact events were also classified into several categories, such as normal movement, impact with boards or glass, falls to the ice, body collisions, and occlusion scenes. Although some of these categories may resemble head impacts, they are not head impacts, and they served to prevent the network from overfitting to detect head impact events. To enable the network to differentiate between head and non-head impact events, our diverse data set was developed to ensure that the network would be exposed to a broad range of scenarios and conditions that occur during a game. In total, the dataset for the head impact detection was comprised of 1350 video clips, which were split into a training set and a validation set, with 75% and 25% of the clips, respectively. The validation dataset is used to assess the performance of the head impact detection dataset.

***Head Impact Activity Detection Network***

An activity detection network was utilised for the detection of head impacts within each video segment. Considering its cons as computationally intensive and prone to overfitting, for this study, an LRCN architecture was chosen as this approach utilises convolutional layers for spatial feature extraction from the frames, with the extracted spatial features then fed to LSTM layer(s) at each time step for temporal sequence modelling [38]. As mentioned before, LRCN has some drawbacks, but they are not significant enough to prevent its use as a method. `The code provided in Figure 2 implements a sequential model architecture, incorporating Time-Distributed layers for applying convolutional operations and pooling across the temporal dimension of the input data. Dropout layers are included to prevent overfitting during training [50]. Multiple convolutional layers with varying filter sizes and activation functions capture spatial features effectively, while LSTM layers enable the model to capture long-term dependencies in the data. The input data has a shape of (None, 25, 300, 300, 3). None represents the batch size (which can vary during training or inference), 25 is the sequence length (number of time steps), 300 and 300 are the height and width of the input images, and the last number is the number of channels (e.g., RGB channels). The final output shape is (None, 2), indicating two output classes (**Error! Reference source not found.**). The model included the Adam optimiser and categorical cross-entropy loss function. Additionally, class weights were computed to address class imbalance during training, as there are more non-head impact than head impact datasets. The training loop iterated over 50 epochs, with early stopping set to 10 to prevent overfitting. Early stopping is a technique used to stop the training process if the model's performance on a separate validation dataset stops improving or starts deteriorating. The performance of the head impact activity detection neural network was evaluated using metrics of validation accuracy and validation loss. Validation accuracy was used to assess the percentage of correctly predicted labels on a separate validation dataset, which was not used for training. Validation loss was used to measure how well the model performed on the validation dataset, calculated as the difference between the predicted output and the actual target labels. A good model is indicated by both high validation accuracy and low validation loss, showing accurate and consistent predictions. For example, a model with a validation accuracy above 90% and low, stable validation loss is typically considered good. A fair model may have moderate accuracy (75-90%) and higher loss, indicating reasonable predictions but with room for improvement [24].

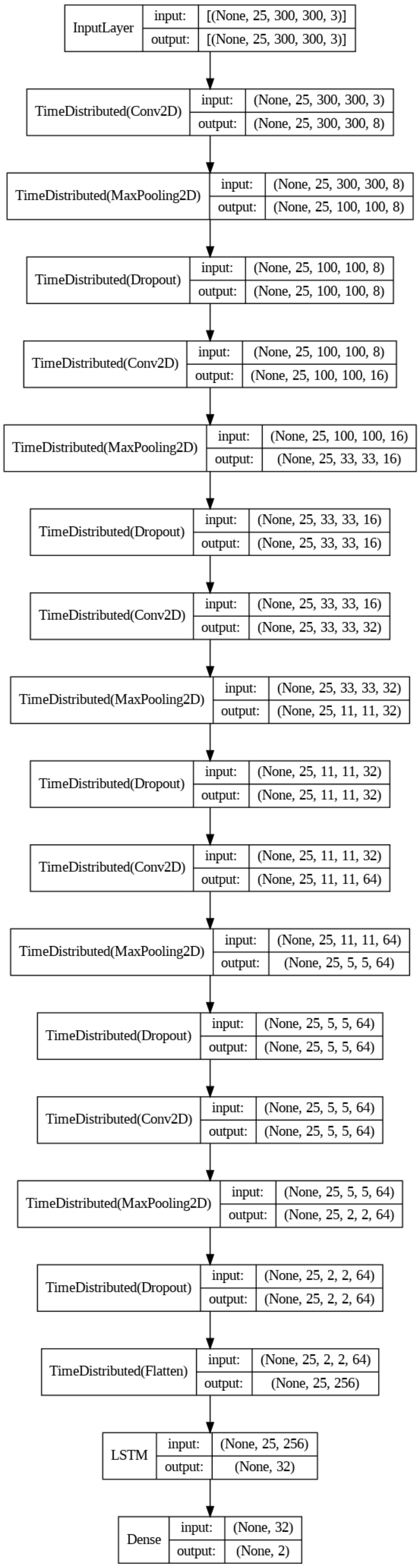


Fig. 2. LRCN Architecture for the Head Impact Activity Detection Network.

## III. RESULTS

### *Player Detection*

After 250 epochs of training, the YOLOv8x model yielded results, with precision, recall, mAP50, and mAP50-95 scores of 0.97, 0.97, 0.99, and 0.95, respectively. These metrics reflect the model's ability to accurately identify and locate players, goal keepers and referees in the input images. A precision score of 0.97 indicates high accuracy in object predictions, while a recall score of 0.97 demonstrates the model's proficiency in capturing a substantial portion of the actual positive instances. The mAP50 and mAP50-95 scores of 0.99 and 0.95 signify the model's strong performance in detecting objects at varying confidence thresholds. The output of an unseen video is shown in the **Error! Reference source not found.**.

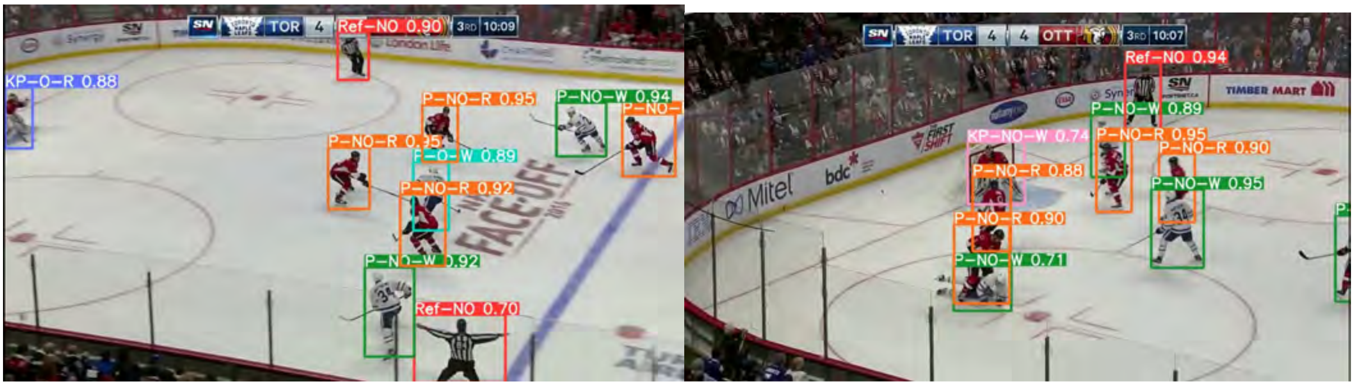


Fig. 3. Images of the output of the player detection network using a test video. The labels consist of three parts: the first letter corresponds to the category (P: Player, KP: Goal Keeper, Ref: Referee), the second letter indicates occlusion status (OC: Occlusion, NOC: Non-Occlusion), and the third letter denotes the jersey colour (R: Red, W: White).

***Player Tracking***

Player tracking was evaluated from a series of 15, 10 seconds videos, including scenarios with occlusions and head impact events. The STRONGSORT algorithm, using the IoU incorporated cost function resulted in no ID switches observed in any of the videos, unlike the default version, which revealed an average of 9 ID switches during the tracking process. This difference emphasises the significance of our adapted cost function in mitigating ID switches, thereby contributing to the improved accuracy and consistency of the object (player) tracking system (**Error! Reference source not found.**).

***Head Impact Activity Detection Network***

After 50 epochs of training, the head impact detection network achieved fair results. The validation accuracy reached 87%, and the validation loss reached 0.04. This outcome highlights the model's ability to discern and classify head impact events effectively. The integration of early stopping contributes to model efficiency and prevents unnecessary computational expenditure beyond the point of optimal performance.

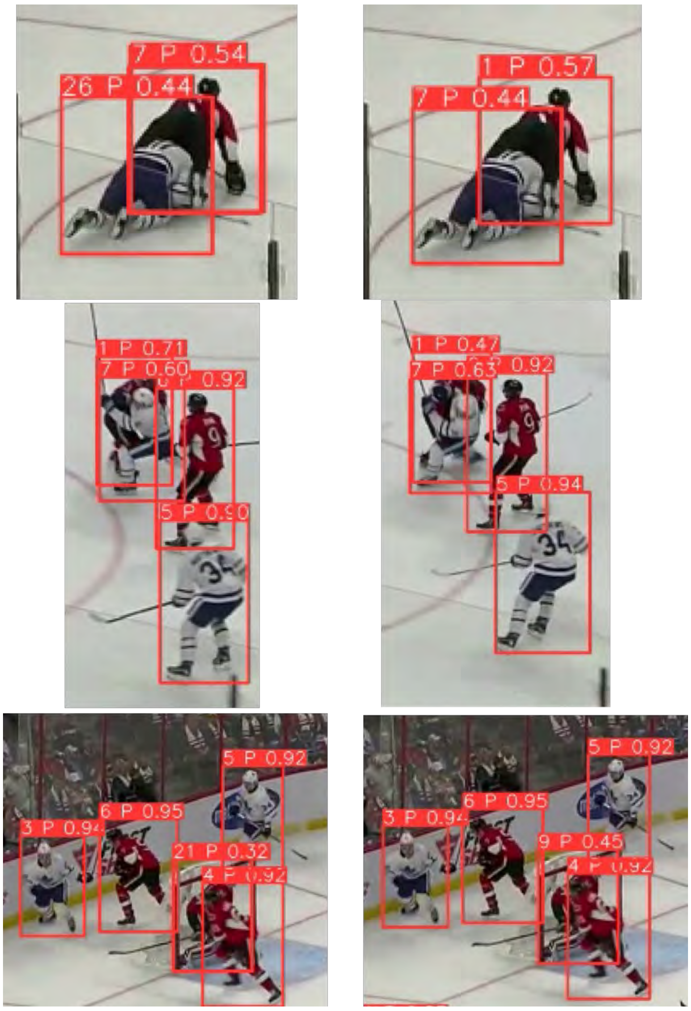


Fig. 4. Image of video output of the players' ID using StrongSORT default cost function (left) and the adapted cost function (right).

## IV. Discussion

Understanding the relationship between brain trauma and the outcome of injury is extremely complex, particularly concerning the compounding effect of RHI in contact sports. An approach to gaining better insight is to leverage widespread game video and available techniques in AI to create large datasets for complex analysis. Identifying and tracking head impacts throughout the game is challenging, often relying on observation, but is the first step in tracking head trauma experienced over time. The aim of this study was to establish an approach for automating the detection of head impacts from 2D videos using computer vision techniques.

The outcome of our player detection algorithm exhibits very good results, including high precision, recall, mAP50, and mAP50-95 scores. However, exploring alternative object detection algorithms, such as Faster R-CNN, could provide valuable insights into potential improvements in tracking accuracy. Faster R-CNN offers precise and high-quality object detection, leveraging CNN features effectively [33] [35]. Hence, a comparative analysis of multiple object detection algorithms can offer a comprehensive understanding of their strengths and weaknesses, aiding informed decisions regarding their implementation for player detection in ice hockey games.

While results obtained in this analysis of unseen 10-second videos, showed the absence of player ID switches, certain conditions pose challenges. These conditions include player substitutions, extended periods of players being out of view for over 40 frames, or specific occlusion events, which may lead to ID switches or incorrect ID assignments. To further enhance tracking performance, the authors agree that post-tracking ID refinement is deemed necessary in the future. Previous studies have highlighted the effectiveness of using jersey numbers for ID confirmation after tracking [45]. Additionally, incorporating player trajectories and contextual information, such as team formations and player positions, have potential in providing cues for accurate ID assignment.

Implementing these approaches could enhance tracking performance, minimising ID switches, and improving our overall tracking accuracy.

Our head impact activity detection network demonstrates good performance on the limited datasets used for its development. There are several factors that should be addressed to enhance its efficacy for future studies. One key consideration is dataset creation. In this study, the area around the players was manually cropped using video analysis software, resulting in a dataset with varying zoom levels. This manual selection of pixels around each player created a diverse dataset in terms of zooming view, with smaller crops leading to zoomed-in videos and larger crops to zoomed-out views of events. This manual dataset creation resulted in inconsistency between videos, and consequently our network could only be trained on a subset of the cropped videos. Our current study has created a network that can detect and track the players, making it possible to create more datasets automatically. Moving forward, our future work will create a larger, and more diverse dataset that can generalise our network into all visual features, such as zooming in and out in all head impact event categories.

One challenge in this study involved the limited number of head impact events in our dataset, which was notably smaller compared to the dataset for our non-head impact events. To address this issue and increase the head impact dataset, augmenting methods were performed including techniques like flipping video frames and generating two 25-frame head impact videos from a single 50-frame video. However, our dataset size remained insufficient for comprehensive generalization across all possible scenarios and events during ice hockey games. Increasing our head impact event dataset to include number and diversity of impact scenarios will increase the robustness of our model and output results and solve any issues with data imbalance.

A final important factor for future integration into our model is to account for the variability in visual features found among different head impact event categories in ice hockey. For example, we found that categories such as shoulder and elbow collisions are more temporal, while others, such as stick and puck impacts, are more spatial. Therefore, we aim to determine the optimal number of frames and image resolution for each category. Our results showed that using 25 frames, which allows for observing the players' reactions before and after the head impact, and an image size of 300 x 300, based on GPU capabilities, yielded good results. However, for future studies, exploring a smaller range of time duration, such as eight frames, and larger image sizes to assess their impact on detection performance will be beneficial.

## V. CONCLUSIONS

Detecting head impacts in ice hockey from 2D game videos presents significant challenges including pixel intensity variations, camera views, and complex player movements. To address these challenges, a methodology incorporating spatial and temporal features in player detection, tracking, and head impact detection was proposed. The YOLOv8x model demonstrated high precision, recall, mAP50, and mAP50-95 scores for player detection. Object tracking using the StrongSORT algorithm with a modified cost function significantly reduced ID switches, improving player tracking accuracy. For head impact detection, LRCN architectures were used to, achieve a validation accuracy of 87%. However, challenges remain, including dataset, model generalisation, and variability in head impact event categories, suggesting areas for future improvement. Future work will focus on exploring alternative object detection algorithms, refining tracking algorithms, and expanding the head impact detection dataset. Additionally, investigating different neural network architectures, such as CONVLSTM and 3DCNN, may further enhance head impact detection performance. Our findings describe a novel methodology for a head impact activity detection network, establishing a valid approach for automating the detection and tracking of head impacts from 2D videos. This initial step in a network development can be applied to a larger scaled automated system for collecting large brain trauma exposure datasets.

## VI. REFERENCES

[1] Casper, S. T. (2018). Concussion: A History of Science and Medicine, 1870-2005. Headache: The Journal of Head and Face Pain, 58(6), 795-810.

[2] Karton, C., Blaine Hoshizaki, T., & Gilchrist, M. D. (2020). A novel repetitive head impact exposure measurement tool differentiates player position in National Football League. Scientific reports, 10(1), 1200.

[3] Karton, C., & Hoshizaki, T. B. (2018). Concussive and subconcussive brain trauma: the complexity of impact biomechanics and injury risk in contact sport. Handbook of clinical neurology, 158, 39-49.

[4] Post, A., & Hoshizaki, T. B. (2012). Mechanisms of brain impact injuries and their prediction: a review. Trauma, 14(4), 327-349.
[5] Nowinski, C. J., Rhim, H. C., McKee, A. C., Zafonte, R. D., Dodick, D. W., Cantu, R. C., & Daneshvar, D. H. (2024). 'Subconcussive'is a dangerous misnomer: hits of greater magnitude than concussive impacts may not cause symptoms. British Journal of Sports Medicine.
[6] Hughes, M. D., & Bartlett, R. M. (2002). The use of performance indicators in performance analysis. Journal of sports sciences, 20(10), 739-754.
[7] O'donoghue, P. (2009). Research methods for sports performance analysis. Routledge.
[8] Liu, Y., Domel, A. G., Yousefsani, et al. (2020). Validation and comparison of instrumented mouthguards for measuring head kinematics and assessing brain deformation in football impacts. Annals of Biomedical Engineering, 48, 2580-2598.
[9] Daneshvar, D. H., Baugh, C. M., Nowinski, C. J., McKee, A. C., Stern, R. A., & Cantu, R. C. (2011). Helmets and mouth guards: the role of personal equipment in preventing sport-related concussions. Clinics in sports medicine, 30(1), 145-163.
[10] Kuo, C., Wu, L., Loza, J., Senif, D., Anderson, S. C., & Camarillo, D. B. (2018). Comparison of video-based and sensor-based head impact exposure. PloS one, 13(6), e0199238.
[11] Larsen, K. (2022). Comparison of Brain Strain Magnitudes Calculated Using Head Tracking Impact Parameters and Body Tracking Impact Parameters Obtained from 2D Video (Doctoral dissertation, Université d'Ottawa/University of Ottawa).
[12] Chen, W., Post, A., Karton, C., Gilchrist, M. D., Robidoux, M., & Hoshizaki, T. B. (2023). A comparison of frequency and magnitude of head impacts between Pee Wee and Bantam youth ice hockey. Sports biomechanics, 22(6), 728-751.
[13] Hoshizaki, B., Post, A., Kendall, M., Karton, C., & Brien, S. (2013). The relationship between head impact characteristics and brain trauma. J Neurol Neurophysiol, 5(1), 1-8.
[14] Karton, C., & Hoshizaki, T. B. (2021). Biomechanics of sport-related neurological injury. Clinics in Sports Medicine, 40(1), 19-38.
[15] Karton, C., Post, A., et al. (2021). Exposure to brain trauma in six age divisions of minor ice hockey. Journal of biomechanics, 116, 110203.
[16] Clara Karton, Thomas Hoshizaki, and Michael Gilchrist. The influence of impactor mass on the dynamic response of the hybrid iii headform and brain tissue deformation. In Mechanism of concussion in sports. ASTM International, 2014.
[17] Agoston, D. V., & Langford, D. (2017). Big Data in traumatic brain injury; promise and challenges. Concussion, 2(4), CNC44.
[18] Rodger, J. A. (2015). Discovery of medical Big Data analytics: Improving the prediction of traumatic brain injury survival rates by data mining Patient Informatics Processing Software Hybrid Hadoop Hive. Informatics in Medicine Unlocked, 1, 17-26.
[19] Yurgil, K. A., Barkauskas, D. A., et al. (2014). Association between traumatic brain injury and risk of posttraumatic stress disorder in active-duty Marines. JAMA psychiatry, 71(2), 149-157.
[20] Duong, T. V., Bui, H. H., Phung, D. Q., & Venkatesh, S. (2005, June). Activity recognition and abnormality detection with the switching hidden semi-markov model. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) (Vol. 1, pp. 838-845). IEEE.
[21] Blank, M., Gorelick, L., Shechtman, E., Irani, M., & Basri, R. (2005, October). Actions as space-time shapes. In Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1 (Vol. 2, pp. 1395-1402). IEEE.
[22] Ke, Y., Sukthankar, R., & Hebert, M. (2007, June). Spatio-temporal shape and flow correlation for action recognition. In 2007 IEEE conference on computer vision and pattern recognition (pp. 1-8). IEEE.
[23] Du, Y., Chen, F., & Xu, W. (2007). Human interaction representation and recognition through motion decomposition. IEEE Signal Processing Letters, 14(12), 952-955.
[24] Rezaei, A., & Wu, L. C. (2022). Automated soccer head impact exposure tracking using video and deep learning. Scientific reports, 12(1), 9282.
[25] Host, K., & Ivašić-Kos, M. (2022). An overview of Human Action Recognition in sports based on Computer Vision. Heliyon, 8(6).
[26] Ivasic-Kos, M., Host, K., & Pobar, M. (2021). Application of deep learning methods for detection and tracking of players. Deep Learning Applications.
[27] Lu, N., Wu, Y., Feng, L., & Song, J. (2018). Deep learning for fall detection: Three-dimensional CNN combined with LSTM on video kinematic data. IEEE journal of biomedical and health informatics, 23(1), 314-323.

[28] Ngo, C. W., Ma, Y. F., & Zhang, H. J. (2005). Video summarization and scene detection by graph modeling. IEEE Transactions on circuits and systems for video technology, 15(2), 296-305.
[29] Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., & Fei-Fei, L. (2014). Large-scale video classification with convolutional neural networks. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (pp. 1725-1732).
[30] Wang, L., Qiao, Y., & Tang, X. (2015). Action recognition with trajectory-pooled deep-convolutional descriptors. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4305-4314).
[31]Prakash, H., Chen, Y., Rambhatla, S., Clausi, D., & Zelek, J. VIP-HTD: A Public Benchmark for Multi-Player Tracking in Ice Hockey.
[32] Du, Y., Zhao, Z., Song, Y., Zhao, Y., Su, F., Gong, T., & Meng, H. (2023). Strongsort: Make deepsort great again. IEEE Transactions on Multimedia.
[33] Girshick, R. (2015). Fast r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448).
[34]Pascanu, R., Mikolov, T., & Bengio, Y. (2013, May). On the difficulty of training recurrent neural networks. In International conference on machine learning (pp. 1310-1318). Pmlr.
[35] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28.
[36] Lin, J., Gan, C., & Han, S. (2019). Tsm: Temporal shift module for efficient video understanding. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 7083-7093).
[37] Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., & Woo, W. C. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. Advances in neural information processing systems, 28.
[38]Donahue, J., Anne Hendricks, L., Guadarrama, S., Rohrbach, M., Venugopalan, S., Saenko, K., & Darrell, T. (2015). Long-term recurrent convolutional networks for visual recognition and description. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2625-2634).
[39]Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., & Woo, W. C. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. Advances in neural information processing systems, 28.
[40]Wang, L., Xiong, Y., Wang, Z., Qiao, Y., Lin, D., Tang, X., & Van Gool, L. (2018). Temporal segment networks for action recognition in videos. IEEE transactions on pattern analysis and machine intelligence, 41(11), 2740-2755.
[41] Fang, H. S., Xie, S., Tai, Y. W., & Lu, C. (2017). Rmpe: Regional multi-person pose estimation. In Proceedings of the IEEE international conference on computer vision (pp. 2334-2343).
[42] Fani, M., Neher, H., Clausi, D. A., Wong, A., & Zelek, J. (2017). Hockey action recognition via integrated stacked hourglass network. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 29-37).
[43] Neher, H. (2018). Hockey pose estimation and action recognition using convolutional neural networks to ice hockey (Master's thesis, University of Waterloo).
[44] Lu, N., Wu, Y., Feng, L., & Song, J. (2018). Deep learning for fall detection: Three-dimensional CNN combined with LSTM on video kinematic data. IEEE journal of biomedical and health informatics, 23(1), 314-323.
[45] Vats, K., Walters, P., Fani, M., Clausi, D. A., & Zelek, J. S. (2023). Player tracking and identification in ice hockey. Expert Systems with Applications, 213, 119250.
[46] Butterfield, J., Post, A., Karton, C., Robidoux, M. A., Gilchrist, M., & Hoshizaki, T. B. (2023). A video analysis examination of the frequency and type of head impacts for player positions in youth ice hockey and FE estimation of their impact severity. Sports Biomechanics, 1-17. Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
[47] Mao, A., Mohri, M., & Zhong, Y. (2023, July). Cross-entropy loss functions: Theoretical analysis and applications. In International Conference on Machine Learning (pp. 23803-23828). PMLR.
[48] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).
[49]Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research, 15(1), 1929-1958.