Developing an automated system to measure player velocity for head impacts evaluation using 2D video in youth ice hockey games

Parisa Dehghan

Supervisors:

Dr. Blaine Hoshizaki, Dr. Allison Clouthier

Committee members:

Dr. Robert Laganière, Dr. Ryan Graham

Thesis submitted to the University of Ottawa
in partial fulfillment of the requirements for
the Doctor of Philosophy Degree in Human Kinetics

School of Human Kinetics
Faculty of Health Science
University of Ottawa

© Parisa Dehghan, Ottawa, Canada, 2025

Table of Contents

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Professor Blaine Hoshizaki, for his unwavering support, invaluable guidance, and profound insights throughout my doctoral journey. His mentorship, encouragement, and dedication to excellence have shaped both my research and my growth as a scholar.

I am also sincerely grateful to my co-supervisor, Dr. Allison Clouthier, for her thoughtful feedback, constructive advice, and steadfast support, which have been instrumental in refining my work.

My heartfelt thanks go to my committee members, Dr. Ryan Graham and Dr. Robert Laganière, for their guidance, constructive feedback, and valuable contributions that strengthened the quality of this research.

I am thankful to our lab coordinator, Dr. Clara Karton, for her assistance and for keeping our research environment running smoothly. I also appreciate my lab teammates for their collaboration and camaraderie, which made this academic journey a memorable one.

A very special thank you goes to my best colleague, lab member, friend, and my rock, my husband, Amirhossein Azadi. His unwavering help, encouragement, and love have made every step of this journey not only possible but also much sweeter.

Finally, I owe my deepest appreciation to my mother, Tayebeh, whose endless love, support, and belief in me have been a constant source of strength.

Co-Authorship

This dissertation includes one published and two unpublished manuscripts. The authorship details are as follows:

1. **Dehghan, P., Azadi, A., Clouthier, A., & Hoshizaki, T. B.** *Object Detection for Ice Surface Localisation in Youth Hockey.* Published in the peer-reviewed journal **IRCOBI** and presented at the IRCOBI Conference, Sweden, 2024.

2. **Dehghan, P., Azadi, A., Karton, C., Clouthier, A., & Hoshizaki, T. B.** *Development of an Automated System to Obtain the Planar Velocity of Individual Players from 2D Video of Youth Ice Hockey Games.* Submitted to the peer-reviewed journal **Sports Engineering**.

3. **Dehghan, P., Azadi, A., Karton, C., Laganière, R., Graham, R., Clouthier, A., & Hoshizaki, T. B.** *Validation of Deep Learning-Based Velocity Estimation from Single, Side-View Ice Hockey Game Footage Using Top-Down Drone Measurements.* Submitted to the peer-reviewed journal **Sport Biomechanics.**

For each manuscript, **Parisa Dehghan** completed the tasks of conceptualization and research design, data acquisition, data analysis and interpretation, manuscript drafting, and critical revision prior to publication or submission.

Abstract

Traumatic brain injury (TBI) is a major concern in contact and collision sports such as football, rugby, ice hockey, and boxing, where athletes are frequently subjected to repetitive head impacts. These impacts range from mild concussions to severe neurological trauma and are often associated with long-term cognitive, behavioural, psychological, and physical impairments. In ice hockey, the combination of high-speed skating and physical contact significantly increases the risk of concussion. Concussion rates as high as 1.81 per 1000 athletic exposures have been reported, emphasizing the need for improved injury monitoring and prevention, particularly at the youth level, which constitutes the majority of participants in the sport. Young athletes are especially vulnerable due to their developmental stage and lack of access to monitoring technologies commonly available in professional settings. Many head impacts in youth hockey go undetected or unreported, impeding early intervention and increasing the risk of long-term effects. Effective injury mitigation requires a deeper understanding of the biomechanical factors contributing to brain trauma, including impact magnitude, frequency, inter-impact intervals, and cumulative exposure duration. Among these, velocity at the moment of impact plays an important role in determining the kinematics and energy transferred to the brain. However, the relationship between specific impact characteristics and clinical outcomes remains poorly understood. Developing predictive models requires large, structured datasets that capture biomechanical and contextual variables in real-world game settings. To address this need, and as part of a larger project aimed at developing a fully automated pipeline to extract head-impact characteristics from 2D video, this study focused on a primary step required for estimating head-impact velocity. Specifically, we developed and validated a fully automated pipeline for estimating planar player velocity using single-camera, side-view 2D video footage, an accessible and scalable solution tailored to youth hockey environments. The pipeline eliminates the need for specialized tracking systems and consists of three core components, structured into three studies:

Study 1: Object Detection for Ice Surface Localization in Youth Hockey

A YOLOv5 object detector was trained on 9,900 annotated side-view frames from youth games at the Peewee, Midget, and Atom levels to identify specific rink landmarks. When at least four landmarks were detected by the network, a custom script computed a homography matrix to project

the side-view frame into a standardized top-down view. This approach demonstrated strong performance, achieving an F1 score of 0.99 and mAP@0.5 of 98.5% for object detection, along with an average Intersection over Union (IOU) of 0.96 for localization accuracy.

Study 2: Development of an Automated System to Obtain the Planar Velocity of Individual Players From 2D Video of Youth Ice Hockey Games

A YOLOv8 model, pretrained on 80,000 NHL images and fine-tuned on 4,700 youth hockey frames, detects and classifies players by jersey color (light/dark). Player tracking is performed using a modified StrongSORT algorithm, which reduced identity switch errors from 172 to 53 on 100 clips and improved Multi-Object Tracking Accuracy (MOTA) from 89.0% to 94.5%. Trajectories are transformed into top-down coordinates using the output of Study 1, and velocity is estimated based on displacement over 5-frame intervals (200 ms at 25 FPS), enabling real-time analysis of motion patterns relevant to head impact risk. A YOLOv8 model, pretrained on 80,000 annotated NHL player images and fine-tuned on 4,700 annotated youth hockey frames, was used to detect and classify players by jersey color (light/dark). The model achieved high detection accuracy, with a precision of 96% and a recall of 97%. Player tracking was performed using the StrongSORT algorithm enhanced with our custom cost function, which reduced identity switch errors from 172 to 53 across 100 clips (90 frames each at 30 FPS) and improved Multi-Object Tracking Accuracy (MOTA) from 89.0% to 94.5%. Player trajectories from side-view footage were transformed to a top-down view using frame-by-frame homography calculated in Study 1, and velocities were estimated based on displacement over 5-frame intervals (200 ms at 25 FPS).

Study 3: Validation of Deep Learning-Based Velocity Estimation from Single, Side-View Ice Hockey Game Footage Using Top-Down Drone Measurements

The objective of Study 3 was to validate the player velocities estimated by the proposed pipeline in Study 2 under real-world conditions, using drone-mounted top-down video as the ground truth. Velocities in the top-down footage were manually extracted in Kinovea by tracking a helmet-mounted marker across synchronized 10-frame intervals. Side-view estimates from both fixed and panning camera configurations were compared against the ground truth across four rink regions: two defensive faceoff circles and two central zones (near and far from the camera). Movement types included lateral, forward, and diagonal skating paths. A zoomed-in side camera configuration was also tested in the far zone to assess the impact of reduced landmark visibility. Results showed that both fixed and panning camera setups achieved mean absolute errors under 10%, while the zoomed-in configuration remained within 20%, a range still acceptable for categorizing impact severity. Bland–Altman analysis indicated strong agreement between the fixed/panning configurations and the top-down reference, providing preliminary validation of the pipeline. Sensitivity analyses were conducted by perturbing player coordinates in both axes by ±5, ±10, and ±15 pixels to evaluate susceptibility to localization error. The results revealed greater sensitivity in regions farther from the camera. Additionally, velocity was recalculated using temporal windows of 5, 10, 20, and 30 frames, showing that increasing the frame window consistently reduced error, particularly in the vertical direction where perspective distortion is more pronounced.

This study proposed a preliminarily validated pipeline offering a cost-efficient, scalable, and accessible solution for estimating player velocity from standard broadcast footage of youth ice hockey games. Its fully automated structure and reliance on commonly available video make it particularly suited for large-scale injury surveillance and biomechanics research in environments without advanced tracking infrastructure or specialized sensor-based systems. As part of a larger project to develop an automated data capture system for documenting head impact characteristics, this study focused on creating and validating an AI-driven method for estimating player velocity in the context of head impacts. Velocity is a key factor influencing the energy transferred to the head during collisions. By generating large datasets and applying advanced AI algorithms, the system enables the identification of patterns linking impact characteristics to injury outcomes.

List of Tables

List of Figures

List of Equation

Definition of Terms

TBI - Traumatic Brain Injury

SDH - Subdural Hematoma

EDH - Epidural Hematoma

SAH - Subarachnoid Hemorrhage

DAI - Diffuse Axonal Injury

mTBI - Mild Traumatic Brain Injury

LMHI - Low-Magnitude Head Impact

CTE - Chronic Traumatic Encephalopathy

NfL - Neurofilament Light Chain

AI - Artificial Intelligence

MPS - Maximum Principal Strain

FE - Finite Element

GPS - Global Positioning System

IOU - Intersection Over Union

2D - Two-Dimensional

YOLO - You Only Look Once (object detection algorithm)

YOLOv8 - You Only Look Once, version 8

DLT - Direct Linear Transform

VGG - Visual Geometry Group (from VGG16 architecture)

CPU - Central Processing Unit

CNN - Convolutional Neural Network

R-CNN - Region-based Convolutional Neural Network

RPN - Region Proposal Network

SSD - Single Shot Detector

COCO - Common Objects in Context (dataset)

NHL - National Hockey League

SORT - Simple Online and Realtime Tracking

RHI - Repetitive Head Impacts

CTE - Chronic Traumatic Encephalopathy

MBIM - Model-Based Image Matching

mAP - Mean Average Precision

RANSAC - Random Sample Consensus

NMS - Non-Maximum Suppression

MOT - Multi-Object Tracking

MOTA - Multiple Object Tracking Accuracy

ID - Identity (in tracking context, as in ID switches)

AP - Average Precision

FPS - Frames Per Second

MAE - Mean Absolute Error

U11, U13, U18 - Under-11, Under-13, Under-18 (age categories)

TP - True Positive: correctly predicted positive.

FP - False Positive: incorrectly predicted positive

TN - True Negative: correctly predicted negative.

FN - False Negative: incorrectly predicted negative

GT- Ground Truth

IMU- inertial measurement unit

NISL-Neurotrauma Impact Science Laboratory

Chapter 1. Introduction

1.1. Introduction and Background

Traumatic brain injury (TBI) is a major public health concern and one of the leading causes of morbidity and mortality among young individuals [1]. In the United States alone, approximately two million people sustain a TBI each year, contributing to nearly one-third of all injury-related deaths [2]. TBI affects diverse populations, including athletes, military personnel, and civilians involved in unintentional incidents [3], [4] . According to the most recent consensus, as proposed by the Demographics and Clinical Assessment Working Group of the International and Interagency Initiative toward Common Data Elements for Research on Traumatic Brain Injury and Psychological Health, TBI is defined as “an alteration in brain function, or other evidence of brain pathology, caused by an external force.” [5],[6]. These forces may include direct impact to the head, rapid acceleration or deceleration, penetrating trauma, or blast exposure [7]. Such mechanisms can trigger both structural and functional changes in the brain, initiating a complex neurometabolic cascade characterized by ionic imbalance, glutamate excitotoxicity, cytoskeletal disruption, and mitochondrial dysfunction [6], [8], [9]. These processes result in transient neurological impairments such as confusion, disorientation, and loss of consciousness [8]. TBI encompasses a broad range of clinical presentations, from focal injuries such as subdural and epidural hematomas (SDH, EDH), contusions, and subarachnoid hemorrhage (SAH), to diffuse injuries such as diffuse axonal injury (DAI) [9]. Among these, mild traumatic brain injury (mTBI), commonly referred to as concussion, is the most frequently reported form, particularly in contact sports [7]. Although classified as “mild,” concussions can result in serious acute and long-term cognitive, emotional, and neurological consequences [10], [11]. This is especially concerning in sports like ice hockey, where the incidence of head impacts is notably high [12]. Ice hockey poses unique risks due to its high-speed, high-contact nature. Players are regularly exposed to head impacts through collisions, falls, and contact with the boards or other players. Concussion incidence rates in professional ice hockey have been reported as high as 1.81 per 1000 player-games [13], highlighting the sport’s potential for brain trauma [14]. However, not only severe, but asymptomatic injuries also pose a risk. Research has shown that even low-magnitude head impacts

(LMHIs), which may not produce immediate symptoms, can alter brain function, compromise white matter integrity, and increase susceptibility to future injury [9], [15], [16]. These impacts have been linked to chronic traumatic encephalopathy (CTE), depression, cognitive decline, and early-onset dementia [9], [17]. This evidence is supported by biomarkers such as tau protein and neurofilament light chain (NfL), which have been proposed as indicators of subclinical injury [15]. Elevated levels of these biomarkers have been associated with long-term changes in brain structure and increased risk of CTE [18], [19]. A critical concern is that these asymptomatic impacts often go unreported, as they do not cause immediate, recognizable symptoms and are therefore easily overlooked [20]. These risks are magnified in youth sports, where over 600,000 young athletes are registered in ice hockey programs across North America [21]. Pediatric athletes are especially vulnerable due to the brain's ongoing development, including essential processes such as myelination and synaptic refinement [22], [23]. Repetitive trauma during this developmental window has been associated with persistent impairments in cognition, motor control, and emotional regulation [23], [24], [25].

Despite the known risks, objective monitoring systems are rarely implemented at the youth level. Many concussions go undetected due to limited symptom awareness, peer pressure, or fear of removal from play [26], [27]. Furthermore, most youth programs lack the infrastructure and tools necessary for systematic injury surveillance. Therefore, there is a need for scalable, objective tools capable of measuring brain trauma in youth sports across all age categories in order to improve athlete safety.

To effectively measure brain trauma, it is essential to advance our understanding of the biomechanical mechanisms through which head impacts lead to brain injury. Sport-related brain trauma is influenced by multiple biomechanical variables, including impact velocity, anatomical location (e.g., rear, rear boss, side, front boss, front), direction (e.g., below, center, above), and the compliance of the impact surface (e.g., boards, glass, ice, or a player's body) [26], [27]. While all of these factors are known to affect injury outcomes, the precise relationships between specific impact characteristics and the severity of brain trauma remain poorly understood. Enhancing our understanding of these relationships is important for developing more effective injury prevention strategies and, ultimately, for improving athlete safety. The development and analysis of large datasets, combined with advanced artificial intelligence (AI) algorithms, represent a promising approach for uncovering the complex, nonlinear patterns among these variables [28]. To support such analysis, an automated and scalable method is needed to accurately capture head impact characteristics and systematically record them in large-scale datasets. This system must be accessible across all age groups and capable of functioning as a monitoring tool that logs head impact events and their associated biomechanical parameters.

The current study is part of a broader research initiative to develop an automated system for extracting head-impact characteristics from 2D ice-hockey video. Specifically, it focuses on developing an automated system to estimate player velocities, which are essential for calculating impact velocity (i.e., the closing velocity) between players involved in a collision. Velocity plays a central role in the dynamics of a collision, as it directly determines the amount of energy transferred to the head and is a key determinant of injury severity [29]. Recent research by Azadi et al. (2025) [30] demonstrated that inbound player velocity was the strongest predictor of Maximum Principal Strain (MPS), a widely used biomechanical metric derived from finite element (FE) models that quantifies brain tissue deformation. In that study, velocity outperformed other variables, including impact location, surface compliance, and athlete age, in predicting injury severity. Accurate velocity measurement is therefore essential for deepening our understanding of injury mechanisms and for developing robust, predictive models of brain injury risk. Despite its significance, velocity is rarely measured in youth hockey due to the cost and complexity of available technologies. Systems such as radar guns [31], GPS trackers [32], and optical motion capture [33] are expensive, intrusive, and require specialized setups, rendering them impractical for widespread use, particularly at the younger age categories. While manual video analysis is a potential alternative, it is labor-intensive and not scalable for processing the large volumes of data needed in research and injury surveillance. Moreover, existing 2D video analysis tools such as Kinovea [34] have several key limitations. They rely heavily on manual calibration, which introduces user bias and inconsistency, particularly when using side-view footage prone to parallax errors and spatial distortion. Additionally, they do not support real-time analysis or automation, and the overall process is time-consuming and inefficient for large-scale applications [34]. These limitations are especially problematic in youth hockey, where video quality is often suboptimal and infrastructure for structured data collection is limited.

As such, there is a need for a scalable, automated, accessible, and accurate solution for estimating head-impact velocity to support effective head-trauma monitoring and the creation of large datasets essential for advancing brain-injury prevention in sport. To achieve this, the first critical step is to develop a pipeline to estimate players' planar velocities. Once the head-impact detection network [35] detects an impact event and identifies the players involved, their planar velocities can be used to compute the closing velocity between them, which serves as an estimate of head-impact velocity. This approach is particularly suitable for scenarios in which planar velocity is the dominant component, such as body-to-body collisions.

1.2. Research Objectives and Methodological Framework

This research aimed to develop a fully automated, scalable, accurate, and cost-effective system for estimating player velocity from single-camera, side-view 2D video, specifically tailored to youth ice hockey games, to support improved assessment of head impacts. The system provides 2D planar velocity estimates of each player, accounting for both horizontal components of movement. To achieve this objective, the following three interrelated studies were conducted:

1.2.1. Study I – Ice Rink Localization Using 2D Video:

An automated homography-based [36], [37] transformation framework was developed to map 2D side-view coordinates to a top-down perspective using YOLO-based [37] rink landmark detection. This method was specifically tailored for youth hockey rinks and applied frame by frame to accommodate varying video conditions. By transforming each frame into a consistent top-down view, this study established a uniform coordinate system that treats all players, regardless of their distance from the camera, equally. The transformation also compensates for perspective distortion inherent in side-view footage, thereby enabling more accurate velocity calculations. The corresponding paper, including detailed methodology and results, is presented in Chapter 3, which is based on a paper published in *International Research Council on the Biomechanics of Injury (IRCOBI*) [38] and presented at the corresponding conference in Sweden in 2024.

1.2.2. Study II – 2D Planar Velocity Estimation:

A robust and accurate multi-player detection and tracking system was implemented by combining YOLOv8 [37]for detection with a modified StrongSORT [38], [39] for tracking. An IOU-assisted function, based on the intersection over union of closely positioned bounding boxes, was incorporated into StrongSORT to enhance identity consistency across frames. This system enabled continuous tracking of individual players, even in crowded and dynamic youth hockey settings, using side-view broadcast footage. Leveraging the homography transformation obtained in Study I, each frame was projected into a top-down view, and player velocity was computed over a multi-frame sliding window, yielding precise frame-by-frame planar velocity estimates across both horizontal axes. The corresponding paper, containing detailed methodology and results, is presented in Chapter 4 and has been submitted to *Sports Engineering*, where it is currently under peer review.

1.2.3. Study III – Velocity Validation with Ground Truth Data from a Top-View Camera (Drone):

The velocity estimation pipeline was validated against ground truth velocity data collected using a top-down drone camera that recorded a skater performing under real-world conditions. Error analyses were conducted across different rink zones, considering variations in player distance relative to the camera, movement types in various rink regions, and different camera configurations (e.g., panning, fixed, and zoomed-in views). The detailed methodology and results of this study are presented in Chapter 5 and have been submitted to *Sports Biomechanics*, where the paper is currently under peer review.

1.3. Significance and Broader Implications

This research addresses key limitations in accessibility, scalability, and automation, offering a practical and low-cost alternative to sensor-based tracking technologies that are often unsuitable for youth-level sports. To the best of our knowledge, no existing method has been specifically tailored to the unique characteristics of youth hockey games. Compared to professional settings, youth games differ in several critical ways, including smaller player sizes, darker lighting conditions, and lower-resolution video quality. This study was designed with these challenges in mind and is purpose-built to operate under the constraints commonly found in youth-level hockey footage. By enabling accurate measurement of player velocity, an essential biomechanical factor in brain injury risk, the proposed system facilitates large-scale documentation of head impact characteristics. This capability substantially enhances the potential for injury surveillance, retrospective analysis, and targeted prevention strategies in community and amateur sports environments.

Chapter 2. Basis for the Selected Research Approach

This chapter presents selected works we conducted and reviewed, which served as exploratory trials to guide decision-making. Some tested approaches proved ineffective, and their limitations informed the methodological strategies ultimately adopted in each study.

2.1. Sport field localization

Traditional methods for sports field localization generally involve a three-step process: (i) extracting geometric features such as lines and circles from video frames, (ii) matching these features to a predefined field template, and (iii) estimating a homography transformation using algorithms like the Direct Linear Transform (DLT). For example, Gupta et al. [40] introduced a method that combines points, lines, and ellipses to compute homographies as shown in Figure 2-1. Their approach relied on manually annotated correspondences between keyframes and the field template, which were propagated across subsequent frames through spatial matching. While effective for tracking over long sequences (up to 1,000 frames), this method is limited by its dependence on manual initialization, sensitivity to keyframe selection, and high computational cost, making it unsuitable for real-time applications. To overcome these limitations, recent research has shifted toward deep learning-based approaches. Homayounfar et al. [41], [42] applied a semantic segmentation network based on the VGG16 architecture [43] to detect field lines and estimated camera pose by minimizing vanishing point error. Sharma et al. [44]employed a pix2pix-based model to generate edge maps, which were then aligned with a database of synthetic edge images for field registration.

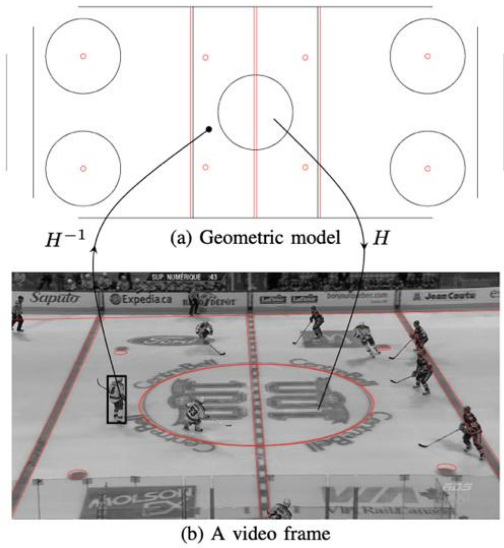

Figure 2-1. The key-frames used to estimate homography between two images[41]

Tarashima [45] introduced a multi-task learning approach to jointly segment and register zones in basketball court images, and Nie et al. [46] developed a dense grid-based representation for field localization in sports such as football and ice hockey. Despite these advancements, there is not universally adopted standard for field localization. Most existing methods are designed for sports with well-defined markings, such as soccer or basketball, and are not easily transferable to ice hockey. Ice hockey poses unique challenges for computer vision due to the inconsistent quality and clarity of rink markings, particularly at the youth level. Markings on the ice surface can vary in brightness, contrast, and visibility, with many rinks having faded or poorly maintained lines. The reflective nature of the ice further introduces glare and lighting variability, making consistent feature detection difficult. These conditions complicate the application of standard geometric feature extraction techniques and require custom solutions tailored to the sport's visual and structural characteristics.

Jiang et al. [47] offered one of the few publicly available deep learning-based pipelines for field localization in soccer. However, their ice hockey dataset was not made publicly accessible. The proposed system consisted of two main stages: an initial homography estimation network, followed by a registration error prediction module that iteratively refined the alignment between video frames and a predefined field template. While the approach was conceptually well-structured, our practical evaluation revealed several limitations. When tested on a standard side-view soccer image, the system failed to produce an accurate alignment, resulting in a visibly misaligned overlay, as illustrated in Figure 2-2 (left). Additionally, the refinement process required approximately two minutes per frame on a 12-core CPU, making it unsuitable for real-time or large-scale deployment. Even after optimization, the predicted field layout remained misaligned with key landmarks, as shown in Figure 2-2 (right), highlighting shortcomings in both accuracy and computational efficiency.

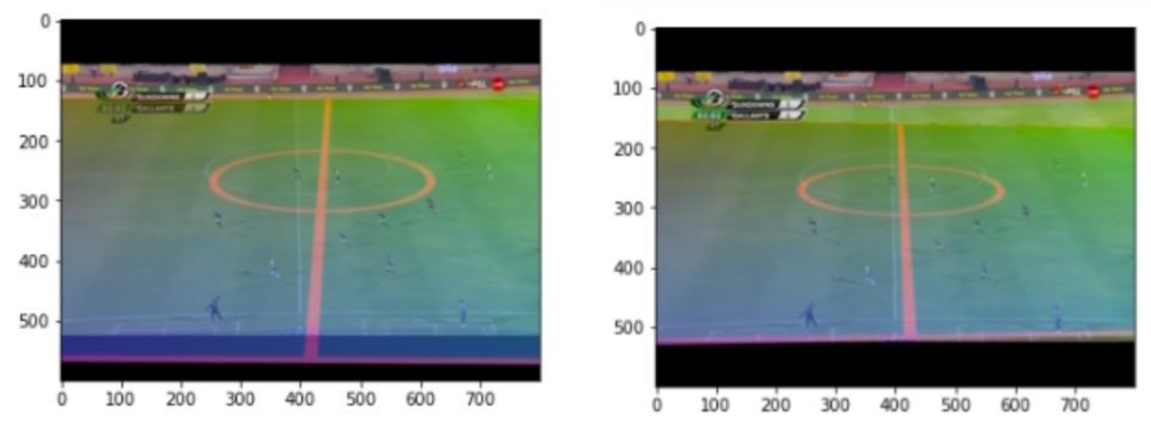

Figure 2-2. Field localization results using Jiang et al.'s pipeline. The left image shows the overlay result after the initial homography estimation (first step of the pipeline). The right image shows the result after optimization using the registration error prediction module (second step of the pipeline).

Another method explored for rink localization was inspired by findings from the literature, which suggested learning-based homography estimation using deep neural networks [48], [49], [50]. To evaluate this approach, we constructed a dataset comprising 694 frames from youth hockey games, each captured from varying camera angles and manually labeled with ground-truth homography matrices. A ResNet-18 [51] model was trained on this dataset, with 80% used for training, 10% for validation, and 10% for testing. However, due to significant variation in camera viewpoints, elevation, and focus typical of youth game recordings, the model exhibited poor generalization. The dataset size proved insufficient to capture the full variability present in real-world conditions, and the resulting performance was unsatisfactory. Consequently, it became evident that this method would require a substantially larger and more diverse dataset to be effective. Given these limitations, this approach was also disregarded in favour of a more practical, real-time method.

Due to the limitations of existing field localization methods, particularly their inefficiency, lack of adaptability to ice hockey, reliance on private datasets and code, and inability to operate in real time, we developed a custom approach tailored to the rink environment. Ice hockey poses specific challenges, including curved rink boundaries, reflective ice surfaces, obstructed views caused by boards and glass, and rapid camera movements required to follow fast-paced gameplay. These factors hinder the effectiveness of traditional line-based or optimization-based localization methods and make frame-by-frame overlay both difficult and time-consuming. To achieve accurate and real-time localization under these conditions, we designed a lightweight pipeline based on object detection, focusing on consistent and semantically meaningful rink landmarks such as faceoff dots, goalposts, and lines. This automated, frame-by-frame approach eliminates the need for manual initialization and enables both precision and computational efficiency suitable for youth hockey analysis. The detailed methodology and results are presented in our first paper, *"Object Detection for Ice Surface Localization in Youth Hockey,"* which was published in the peer-reviewed journal of *IRCOBI* [38] and presented at the *IRCOBI* Conference in Sweden, 2024.

2.2. Player Detection and Tracking in Sports from 2D Videos

Player detection and tracking in sports videos, particularly in fast-paced games like ice hockey, presents substantial challenges due to rapid player motion, frequent occlusions, the uniform visual appearance of players, and variable lighting or camera movement. Accurate detection and consistent tracking are essential for downstream applications such as velocity estimation. The following section reviews relevant literature on player detection and tracking, which informed the design decisions in our system.

Early approaches to player detection and tracking relied on classical computer vision techniques such as background subtraction, optical flow [51], Kalman filters [52], and Multi-Hypothesis Tracking (MHT) [53]. Background subtraction aimed to isolate moving players from static backgrounds but struggled in dynamic environments like ice hockey rinks, where the reflective, uniform ice surface often produced false positives. Optical flow estimated motion vectors across frames but performed poorly during rapid or non-linear movements and in the presence of occlusions. Kalman filters were effective for predicting smooth trajectories but failed when abrupt direction changes occurred. Particle filters improved robustness by maintaining multiple hypotheses of player locations. MHT further extended this by generating and resolving multiple tracking paths, but it was computationally intensive and prone to identity switches, particularly in crowded scenarios with frequent player interactions. These limitations paved the way for deep learning-based methods, which offer greater accuracy, adaptability, and efficiency in complex scenes.

The advent of Convolutional Neural Networks (CNNs) significantly advanced object detection in sports analytics. Faster R-CNN [54], which uses a Region Proposal Network (RPN) to identify potential object regions before classifying them, became a benchmark for high-accuracy detection, even under conditions of partial occlusion. However, due to its two-stage architecture, Faster R-CNN often suffers from slower inference speeds. The Single Shot Detector (SSD) [42] improved computational efficiency by directly predicting bounding boxes across the entire image, making it more suitable for real-time applications, though it tends to underperform when detecting small or distant players. YOLO introduced a highly efficient single-stage detection pipeline capable of real-time processing. Although earlier versions of YOLO struggled with small object detection, YOLOv3 addressed this limitation through multi-scale detection heads, improving its performance in sports scenarios involving small, distant players, such as youth hockey footage captured from wide angles [55].

To evaluate detection methods in our context, we applied Faster R-CNN with a ResNet-101 backbone, pretrained on the COCO dataset, to a dataset of approximately 8,000 annotated frames from NHL hockey games, each containing bounding box annotations for individual players. The dataset was split into 80% for training, 10% for validation, and 10% for testing. The model achieved 92% training accuracy and 80.6% validation accuracy, performing reasonably well in terms of player detection However, its two-stage architecture resulted in slower inference, making it less suitable for real-time or large-scale deployment. In comparison, YOLO, a widely adopted single-stage detector, offered a better balance of accuracy and speed, particularly excelling in detecting small player instances in crowded or low-resolution footage. Given its real-time capabilities and efficient performance, YOLO was selected as the core detection framework for our system. The corresponding loss and accuracy curves over 20 epochs are presented in Figure 2.3.

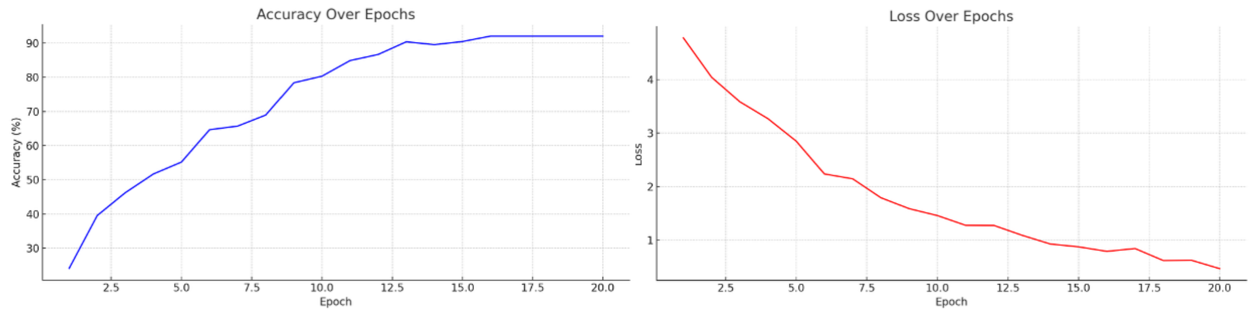

Figure 2-3. Accuracy and loss curves during training of Faster R-CNN with a ResNet-101 backbone over 20 epochs.

2.3. Deep Learning-Based Tracking Approaches

Detection alone is insufficient in multi-player sports, where maintaining consistent player identities across frames is essential for tracking trajectories and calculating velocities. SORT [56] (Simple Online and Realtime Tracking) introduced an efficient method for associating detections across frames using Kalman filters for motion prediction and the Hungarian algorithm for data association. While effective in many real-time scenarios, SORT depends heavily on accurate detections and is sensitive to occlusions and close-proximity interactions, both of which are common in ice hockey. To improve robustness, DeepSORT [56]extends SORT by incorporating appearance-based features extracted using CNNs, enabling more reliable identity tracking through occlusions and abrupt motion changes. This approach has demonstrated success in sports such as basketball and soccer [39], [57], [58]. However, ice hockey presents unique challenges: players often wear visually similar helmets and uniforms, and jersey numbers are frequently obscured, particularly in youth leagues. This visual uniformity reduces the effectiveness of appearance-based re-identification.

To address these challenges, especially occlusion and uniform appearance, we developed a restructured dataset aimed at improving detection under occlusion and proposed enhancements to StrongSORT [39], an advanced tracker that integrates motion, appearance, and temporal continuity. These contributions are detailed in our second study, titled *"Development of an Automated System to Obtain the Planar Velocity of Individual Players from 2D Video of Youth Ice Hockey Games,"* which has been submitted to the peer-reviewed journal *Sports Engineering* and is currently under review. Additionally, a third study, titled *"Validation of Deep Learning-Based*

Velocity Estimation from Single, Side-View Ice Hockey Game Footage Using Top-Down Drone Measurements," presents a validation of the developed velocity estimation pipeline against real-world measurements obtained from drone footage. This study is complete and ready for submission.

Chapter 3. Study I

This chapter is based on the peer-reviewed journal paper published in IRCOBI and presented at the IRCOBI Conference 2024 in Sweden. It addresses the objectives of Study I, which focused on developing an automated method for rink localization, specifically, performing frame-by-frame homography transformation for youth ice hockey game footage. Although evaluated on a limited dataset of approximately 30 youth-level games, the method demonstrated reliable ice-surface localization and supported estimation of a homography matrix for each frame, providing an initial proof of concept for this localization approach. These results offered a practical foundation for subsequent components of the broader pipeline, including player tracking and planar velocity estimation.

Object Detection for Ice Surface Localisation in Youth Hockey

Parisa Dehghan, Amirhossein Azadi, Allison Clouthier, T. Blaine Hoshizaki

Abstract

Sports-related brain injury is a pressing issue, particularly in high-impact sports like ice hockey where impact velocity plays an important role in determining the magnitude of head impacts and subsequent risk of injury. However, existing methods for measuring impact velocity, such as GPS tracking and manual video analysis, are costly making them inaccessible, especially for youth leagues. This study introduces an automated, cost-effective method using computer vision to determine player velocity from 2D video. The initial step involves localising the field, achieved through a novel approach employing YOLOv5 to detect specific landmarks on the ice surface. With a dataset of over 9,900 annotated images, YOLOv5 demonstrates exceptional performance, achieving an F1 score and precision-recall of 0.99 at an 80% confidence level, and mAP scores of 98.5% and 64.5% at IOU thresholds of 0.5 and 0.5:0.95, respectively. By detecting at least four landmarks per frame, homography matrices were calculated to obtain a top-down view, completing the localisation process. This approach achieved an average IOU of 0.96, validating its accuracy in field localisation and demonstrating its potential for improving accessibility and cost-efficiency in measuring impact velocity in ice hockey.

Keywords Head impacts, homography matrix, ice hockey, localisation, object detection.

P. D. is a PhD student in Human Kinetics, A. A. is a PhD student in Human Kinetics, A.C. is an Assistant Professor of Biomedical Engineering, and T. B. H. is a Full Professor of Biomechanics in the Department of Human Kinetics, at the University of Ottawa in Ottawa, Canada.

3.1. Introduction

For the most part, the majority of sports-related traumatic brain injuries (TBI) that contribute to mental health deficits and increased risk for neurodegenerative disease do not result in loss of consciousness or serious symptoms, especially in youth sport and remain under reported. Effective diagnosis, management and prevention require accurate and reliable injury data. Automated data capture technology using video provides an opportunity to obtain biomechanical measures of head impact characteristics to measure brain tissue damage in sport. In ice hockey, the incidence of concussion is notably high, reaching up to 1.81 concussions per 1000 athletic exposures[59]Despite this, the actual incidence among youth hockey players remains uncertain due to ineffective monitoring of this demographic [60]. Hockey is recognised as one of the sports with the highest rates of head injury among youth. Athletes aged 5 to 18 years account for 65% of all sports and recreation-related head injuries treated in US emergency departments [61]. Repetitive head impacts (RHI), both high and low magnitude, continue to be associated with long-term brain health issues including depression, early cognitive decline, and chronic brain disease and chronic traumatic encephalopathy (CTE) [62], [63].

Biomechanical measurements including impact velocity, location, mass, compliance, and direction can be used to obtain the dynamic response of the head and subsequently calculate potential brain tissue damage [64], [65], [66]. The magnitude and interaction of these parameters collectively influence brain motion during impact. Impact velocity is an important variable used to calculate the amount of energy transferred to the brain during impact and resulting injury severity [67]. Various methods, including GPS tracking devices [68], [69], video analysis [70], [71], time gates [72], model-based image matching (MBIM) [73], and Doppler effect devices [74], have been employed to measure velocity in sports. However, these methods are expensive and necessitate specialised technology, and few sports groups, especially youth, have access to such technology [75]. Manual video analysis is time-consuming and labor-intensive, making them impractical for use by sport organizations to monitor head trauma or creating the large databases needed for research. An objective, low-cost method for accurately documenting head impact characteristics during games, especially at the youth level is necessary. Automating data collection systems, facilitated by the application of computer vision techniques using 2D video, presents an opportunity to document head impact characteristics for a large number of sports and athletes. A challenging but important contributor to brain trauma magnitude involves the calculation of player velocity at the time of impact. Accurate estimation of player velocities requires detection, identification, and tracking of players to monitor their trajectories across the field [76]. Players' absolute velocities can be obtained using a top-down view of the ice surface to establish location. The vast majority of game video used in youth sport do not have a top-down view of the playing surface. However, a top-down view can be achieved using homography, a transformation matrix that maps each frame onto a reference plane. Localising the ice rink in each frame is essential for this calculation.

To calculate head impact trauma in sport the accuracy of a novel method for ice rink localisation using various angles from 2-d video to obtain player head impact velocities was investigated. Calculating player velocities is critical to automating head impact data capture in sports like ice hockey. Automating head trauma data collection from video provides for economical head impact documentation and the creation of a comprehensive dataset. Big data supports research investigating the associations between impact characteristics and clinical outcomes contributing to our understanding of brain trauma and injury.

This paper utilised game footage from the primary broadcast camera of youth games (peewee, midget, and Atom), positioned in the arena stands above the centre ice line. A novel method has been proposed in this paper for localising the ice surface, which uses an object detection algorithm to detect landmarks on the ice surface. Specifically, we employed YOLOv5 [77] for landmark detection on the ice rink, aiming to identify at least four points to localise the rink in the top view frame [78]. This paper includes a) Proposing a novel localisation method utilising YOLOv5 [77] for field landmark detection, b) introducing an ice rink dataset comprising over 9,000 images containing landmarks of the playing surface and c) leveraging computer vision techniques in developing an objective tool for measuring brain trauma in ice hockey.

The prevention of sports-related TBI is an important public health concern. Technologies such as deep neural networks and computer vision can facilitate the creation of extensive datasets to investigate the complex interplay between head trauma and mental health. Furthermore, it only requires a 2D video of the game and does not necessitate specialised equipment, making it accessible to youth participants who make up the majority of ice hockey players [61].

Using broadcast footage for computer vision tasks eliminates the need for additional specialised hardware to produce analytics. However, utilising broadcast footage poses its own unique challenges. Cameras used in broadcasts frequently pan, tilt, and zoom to follow the action, necessitating compensation for this movement in the analytics process. This becomes simpler when videos are captured from stationary cameras. Since hockey broadcast videos often lack camera parameters, determining location information must be automated from the video feed itself. Consequently, rink registration is conducted to determine how pixels in video frames correspond to the top-down view of the rink. The sports field localisation task involves determining the planar transform between the view of the hockey rink captured in the broadcast video frame and an overhead view of a hockey rink template [79], [80]. This transform is defined by a homography matrix [78]between the two views of the plane of the ice surface. The homography matrix, H, is a 3 × 3 matrix with eight degrees of freedom, representing the transformation between two planes, typically scaled by a factor s. Equation 1 provides an example of a standard homography matrix, which establishes the correspondence between a point [x′ y′ 1]T on the rink model and a point [x y 1]T on the ice surface in the broadcast frame. The elements of H incorporate translation, rotation, and scale factors [81]. Our method dynamically derives the homography matrix for each video frame based on visible landmarks, allowing for accurate transformation to a top-down view. This frame-specific computation ensures precise localisation and alignment, which are essential for tasks like velocity calculation and player tracking.

$$
s\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = H\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad \text{Eq.1}
$$

A standard method for computing the transformation between the plane of the ice surface in the broadcast video and the rink model involves detecting field markings in the frame, including points, lines, and line intersections, and matching them with their counterparts in the model. After the detection of multiple key points or landmarks in the frame, the RANSAC algorithm is employed. This involves the selection of a random subset of four points from these landmarks and the estimation of the homography matrix using them, typically through the Direct Linear Transform (DLT) algorithm [26]. The fit of all detected landmarks to this estimated homography is then evaluated, with well-fitting points being classified as inliers. This iterative process ensures robustness, with repetitions made multiple times using different sets of points. Ultimately, the homography with the highest number of inliers is chosen. Through the utilisation of RANSAC, outliers are effectively managed, resulting in the identification of the most accurate set of points for the homography and thereby rectifying perspective distortion [81]. An illustration of this process is depicted in Figure 3.1, demonstrating how an image appears when transformed to the top-down view, as well as how the top-down template appears when transformed onto the image.

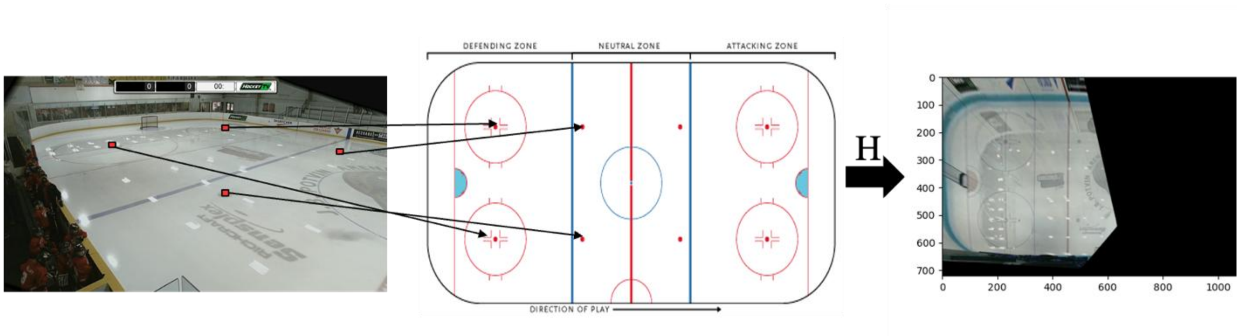

Figure 3-1. Example of warping a video frame onto the overhead rink template (and vice versa) using homography.

Several computer vision techniques have proposed various methods for the localisation of ice surface. A method that combining point, line, and ellipse matches was used to estimate homography in ice hockey rinks, building upon the DLT algorithm [27]. A VGG16 semantic segmentation network to detect lines and estimate the camera position using branch and bound was also used [24]. Researchers have partitioned ice hockey footage into video shots and utilised a ResNet18-based regressor to determine homography between each frame and the ice-rink model [28]. Researchers have also employed the pix2pix network to extract lines from soccer broadcast video and compared them with a database of synthetic edge images for field localisation [29-30]. Lines were identified and categorised on a soccer field [31], while a line segment detector to identify intersections and match them to a template was utilised [32]. Zones from soccer and basketball datasets were detected initialising camera pose estimation through a dictionary lookup and refining it with a spatial transformer network [33]. Reference [34-36] utilised Semantic segmentation or keypoint segmentation was used for sports field localisation, while a two-step method for refining homography estimates by combining the warped template and frame and minimising estimation error has been proposed [37].

The diversity of approaches published in sports field localisation underscores the absence of consensus on an optimal solution. This is compounded by the limited accessibility to algorithms and datasets from other research groups, resulting in each new contribution introducing a novel approach instead of building upon existing methods. Many projects achieve high accuracy using proprietary datasets, hindering meaningful comparisons. Despite introducing various methods, recent advancements suggest that deep network architectures offer superior performance with faster computation [82]. This study proposed a novel localisation method utilising YOLOv5 [77] as an object detection algorithm to identify landmarks on the ice surface, as detailed in the methodology section. This presents an important step for accurately determining player velocity.

3.2. Methods

3.2.1. Data Collection

A dataset was created to support ice rink localisation techniques. It was derived from a larger video repository maintained by the Neurotrauma Impact Science Laboratory (NISL) at the University of Ottawa, which contains youth ice hockey games that had previously been systematically reviewed to document head-impact events. For this thesis, the focus was restricted to rink landmark localisation. Accordingly, multiple games across several youth age categories (e.g., Novice, Atom, and Peewee) were selected from the NISL repository. The games were recorded using a stationary camera positioned near centre ice that panned to capture the full play area, providing comprehensive views of the ice surface across a range of viewpoints. From the selected games, frames were randomly extracted and retained when the designated rink landmarks were visibly discernible. The resulting dataset comprised 9,963 images, each annotated with the predefined landmark labels. Manual annotations were performed in Roboflow [83] , with eight distinct rink landmarks labelled as separate objects of interest in each image. These landmarks encompass critical points such as the intersection of blue lines, goal lines, and centre lines with the boards, the intersection of goal posts with goal lines, the intersection of centre face-off circles with centre lines, and all face-off spots, as depicted in Figure3.2. Well-defined guidelines were implemented to ensure accuracy, and training was provided for the labelers, significantly minimizing potential errors. Once the initial annotation was completed, another researcher reviewed a random sample of the annotations to check for inconsistencies. During the annotation process, markings with similar shapes were grouped together. In the programming section, it was described how each member of the same group was specified and categorised into individual classes based on their relative positions on the ice surface. Bounding boxes were drawn around each object during annotation to ensure accurate representation. The original images maintained an average size of 720 x 1280 pixels. The training process validation for this dataset is observed in Figure3.3. This figure presents annotated image samples, describing the classes defined in Figure 3.2 as they appear in each frame.

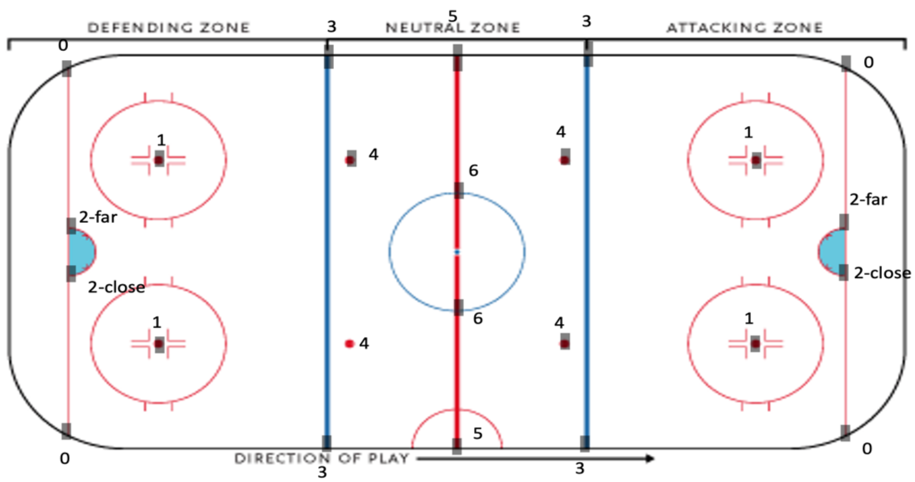

Figure 3-2. Top view of an ice rink with annotated landmarks labelled with their corresponding class names.

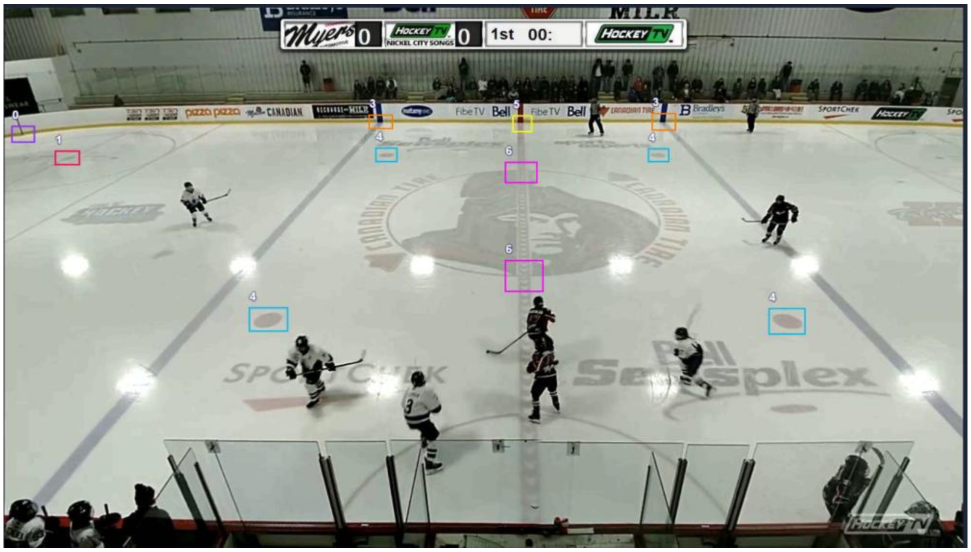

Figure 3-3. Ice surface's landmark dataset.

3.2.2. Data Preprocessing

The hockey images underwent preprocessing to prepare them for use in the YOLOv5 object detection algorithm. This process included resizing the images to 650 x 650 and applying filters to enhance contrast and remove noise. Subsequently, the dataset was augmented by rotating and flipping images, increasing the number of images to 14650. The images were randomly divided into two sets: training data (80%), and validation data (20%). The training data were used to train the machine learning model. The validation data were withheld during training and employed to provide an unbiased assessment of the model's performance throughout the training process. This information was used to fine-tune the model's hyperparameters.

3.2.3. Training the YOLOv5 model

The subsequent stage involved using a dataset to train the YOLOv5 model for ice hockey object detection. YOLO (You Only Look Once) is a prevalent deep learning-based algorithm for object detection known for its wide global receptive field, grid division, anchor frame matching, and multi-semantic fusion detection mechanism. Unlike traditional methods, which typically involve a multi-stage process of feature extraction followed by object identification and classification using separate algorithms like sliding window approaches or region-based methods such as R-CNN [84], YOLO directly predicts bounding boxes and object probabilities using Convolutional

Neural Networks (CNNs) [85], thereby significantly improving detection accuracy. YOLOv5, built on the PyTorch [86] framework, boasts rapid detection speeds, reaching up to 140 frames per second. It incorporates the Mosaic data augmentation technique from YOLOv4, enhancing the detection of small targets by combining input images through random scaling, cropping, and arrangement. YOLOv5 has four variants, YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x, each differing in network depth. Deeper networks yield more feature maps but also involve more complex computations. In this study, YOLOv5s pretraining weights were employed for training. The training process involved feeding images into the algorithm and adjusting neural network weights to improve its capability in detecting and classifying landmarks on the ice rink.

3.2.4. Experimental Environment

The system used Windows 10 Pro with PyTorch (1.8.0) with an AMD Ryzen 9 3900X 12-core processor@3.8 GHz with 32G memory CPU and a NVIDIA GeForce RTX 3090 GPU with CUDA (11.7).

3.2.5. Evaluation of the Object Detection Model

Once the YOLOv5 model had completed its training phase, it underwent evaluation using various metrics to assess its performance, including precision, recall, and F1-score. Additionally, the loss of validation and loss of training, representing the error between predicted and actual values during model training, were monitored, with lower values indicating better performance. These metrics served as indicators of the accuracy and efficiency of the model in detecting and classifying landmarks within the ice hockey images.

Precision, one of the essential components of the F1 score, measures the accuracy of positive predictions made by the model, indicating how many of the predicted landmarks are correct. Mean Average Precision (mAP), calculated at Intersection over Union (IOU) thresholds of 0.5 and 0.5-0.95, provided additional insight into the accuracy of predicted bounding boxes. Recall evaluates the model's ability to find all the relevant landmarks, indicating how many were successfully detected. The F1-score is a combined measure of precision and recall, providing a balanced assessment of the model's overall performance in detecting landmarks on the ice surface. The precision-recall curve (PR) is another metric that evaluates the trade-off between precision and recall at different confidence levels. These metrics collectively provided a comprehensive evaluation of the model's effectiveness in detecting landmarks on the ice surface.

In this paper the following terms were defined as: True positive (TP) refers to instances where the model correctly predicts the presence and location of a landmark within an image, with an IOU score exceeding 0.5. False positive (FP) occurs when the model incorrectly predicts the presence of a landmark instance within an image. This may happen if the model predicts a landmark where none exists or assigns a different label to the predicted landmark compared to the ground truth. False negative (FN) arises when the model erroneously predicts the absence of a landmark instance within an image with an IOU score below 0.5. This could occur if the model overlooks a landmark instance present in the ground truth or assigns a smaller bounding box to the predicted landmark compared to the ground truth.

$$\text{Recall} = \frac{TP}{TP+FN}, \qquad \text{Eq. 2}$$

$$\text{Precision} = \frac{TP}{TP+FP}, \qquad \text{Eq. 3}$$

$$\text{F1-score} = \frac{2*\text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}}, \qquad \text{Eq. 4}$$

By meticulously examining the true positive, false positive, and false negative parameters and their associated scores, we can gain valuable insights into the model's performance. This process not only helped us identify areas for improvement but also allowed us to refine the model's capabilities for more accurate landmark detection and classification.

3.2.6. Programming

Once objects were detected across the frames, those containing at least four detected objects were selected. The centre point of each bounding box enclosing these objects was identified to represent the coordinates of the respective landmarks. In alignment with Figure 3.2, landmarks with similar shapes were grouped into classes. A Python script was developed to distinguish between different objects within the same class. This differentiation was based on various factors, including the objects' relative coordinates to each other, their positions along the centre line, and their relative placements within the image boundaries, such as mid-top or mid-bottom. For each landmark on the ice surface, a subclass name was assigned, as illustrated in Figure 3.4. After initial detection of each object by YOLOv5 in a frame, our programme specified the corresponding label for each subclass. Therefore, if we have class four detected in a frame four times, our programme determined which one belongs to 4-right-top, which one is 4-left-top, which one is 4-right-bottom, and which one is 4-left-bottom. All landmarks shown in Figure 3.2 are also present in Figure 3.4; however, in Figure 3.4, the position of each landmark is explicitly identified and assigned its corresponding subclass name**.**

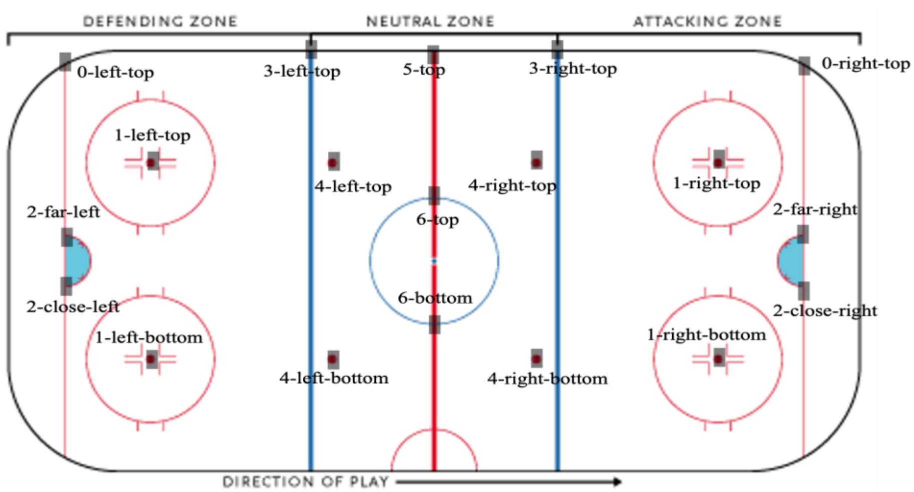

Figure 3-4. Reference frame including the landmarks with their true label according to their position.

A homography matrix was derived by identifying correspondences between at least four landmark points on a top-view reference image and their matching points in each game-video frame. This matrix defined the planar projective transformation used to map the broadcast view to the reference view. The 612 × 258 reference image was selected as a proportional, full-rink top-view plane. Because the homography operates in pixel coordinates and metric values are obtained by scaling using known rink dimensions, this resolution primarily defines the discretization of the top-view coordinate system. Applying the resulting transformation matrix to each video frame maps the broadcast view to a top-view perspective, thereby enabling rink localisation.

3.2.7. Evaluation of the localisation

Two common metrics for evaluating localisation performance are IOUpart and IOUwhole. IOUpart measures the overlap between the ground-truth visible portion of the field (red) and the predicted transformed region (yellow), whereas IOUwhole assesses overlap over the entire field area after transformation (Figure 3.5.). To evaluate localisation performance, 50 frames were randomly selected from nine youth games. Ground-truth localisation was defined using manually identified rink landmarks in each evaluated frame, which served as the reference correspondences. Using these manual landmark coordinates, a ground-truth top-view rink region was generated by projecting the reference rink geometry into the frame. Similarly, the predicted top-view rink region was obtained using the automatically detected landmarks and the corresponding estimated homography. IOU was computed as the ratio of the intersection area to the union area between the ground-truth and predicted regions, i.e., IOU = $|\frac{A \cap B}{A \cup B}|$, where A and B denote the two regions being compared. In this study, $\mathrm{IOU}_{\mathrm{part}}$ was used as the primary localisation metric because it specifically quantifies alignment over the portion of the rink that is visible in each frame.

a) $\mathrm{IOU}_{\mathrm{part}}$ b) $\mathrm{IOU}_{\mathrm{whole}}$

Figure 3-5. Illustration of IOU Part versus Whole, highlighting the distinction in bounding box overlap assessment [87].

3.3. Results

The number of epochs was set to 300. In Figure 3.6, the performance of YOLOv5 on landmark's detection is presented through various metrics during training and validation. The curves depict the loss associated with bounding box predictions (Train/Box Loss, Val/Box Loss), object presence predictions (Train/Obj Loss, Val/Obj Loss), and class predictions (Train/Class Loss, Val/Class Loss), reflecting the model's ability to accurately predict these parameters. Notably, in both training and validation, as epochs increased, the loss consistently decreased, highlighting the network's robust performance and effective learning process. Upon evaluation, the YOLOv5 model achieved a validation precision score of 0.98 and a recall score of 0.97. The mAP_0.5 (mean Average Precision at IOU 50%) and mAP_0.5:0.95 (Mean Average Precision at IOU 50% to 95% in increments of 0.05) of the landmark classification model were 98.5% and 64.5%, respectively. As illustrated in Figure 3.6, as the number of epochs increased, the loss curve gradually stabilised, indicating improved model training effectiveness. These results confirmed the effectiveness of our approach in accurately predicting landmarks from frames taken at various camera angles. F1 and precision-recall curves, evaluating integrated Precision and Recall (PR) indices, are depicted in Figure 3.7. It is evident that almost all eight classes performed very well, with F1 and PR evaluations achieving 0.99 at a confidence level of 80%. The spike observed in the Precision-Confidence Curve for the "2-close" class can occur due to fluctuations in data distribution or the model's adjustments, and it usually smooths out as training progresses. Overall accuracy and performance across all classes are the main concerns, and as long as the trend shows improvement and stability, occasional spikes were considered acceptable.

Homography matrices were calculated and applied to each frame by correlating the detected landmarks with their corresponding coordinates on the reference frame (see Figure 3.2), enabling localisation. Subsequently, the IOU score was computed, yielding an average of 96.1. This high IOU score indicates the accuracy of the localisation process, affirming the method's effectiveness in precisely mapping objects to their respective landmarks on the ice surface. Figure 3.8 and 3.9 depict test images from two video frames of youth games and their corresponding warped frames in the top view, achieved through our network. Figure 3.8 and 3.9 demonstrate the accuracy of our localisation and transformation process by comparing the detected landmarks in the original image (left) with the transformed top-down view (right). Focusing on specific landmarks in the left image, they align precisely between the original and transformed images. For instance, objects in class 6 in the left image correspond with their locations in the top-down view on the right, overlaying correctly on the points at the top and bottom of the centre circle, indicating high precision. It is noteworthy that this transformation provides a top-down view of the frame, allowing us to calculate horizontal plane velocity, and does not include the vertical component.

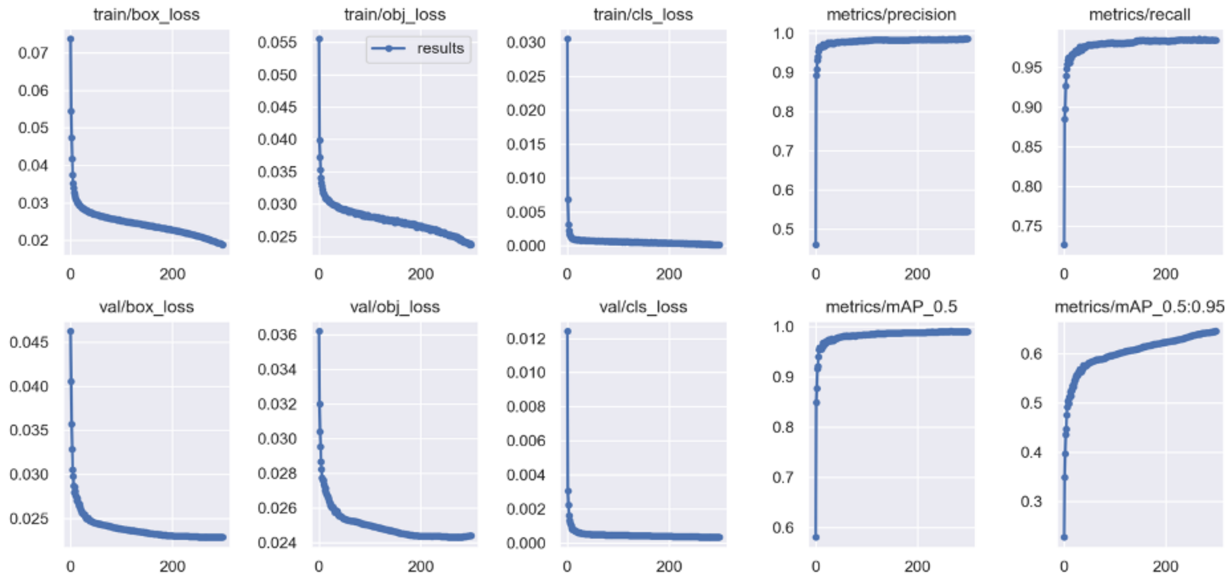

Figure 3-6. Train and validation loss curves, mAP_0.5 and mAP_0.5:0.95 curves, precision curve, recall curve.

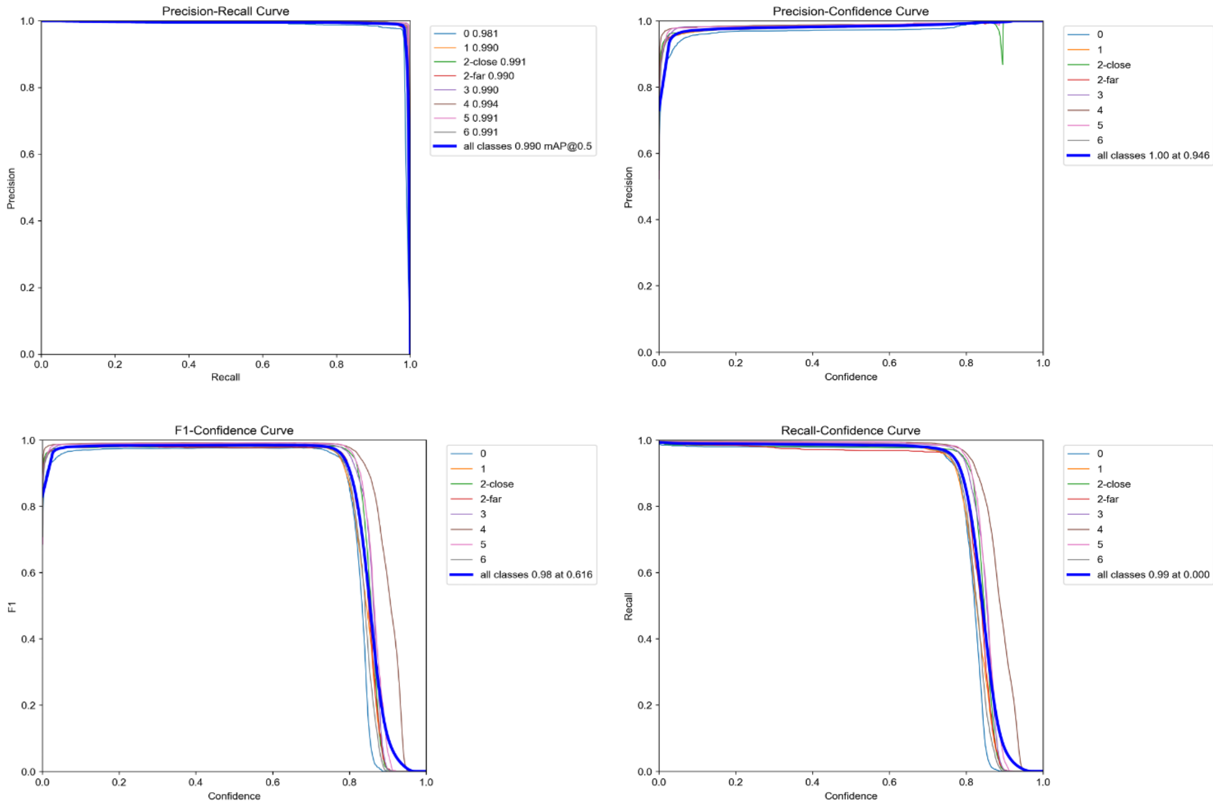

Figure 3-7. Precision-Recall Curve, Precision-Confidence Curve, F1-Confidence Curve, Recall-Confidence Curve.

Figure 3-8. Test image of a sample of a video frame including centre line.

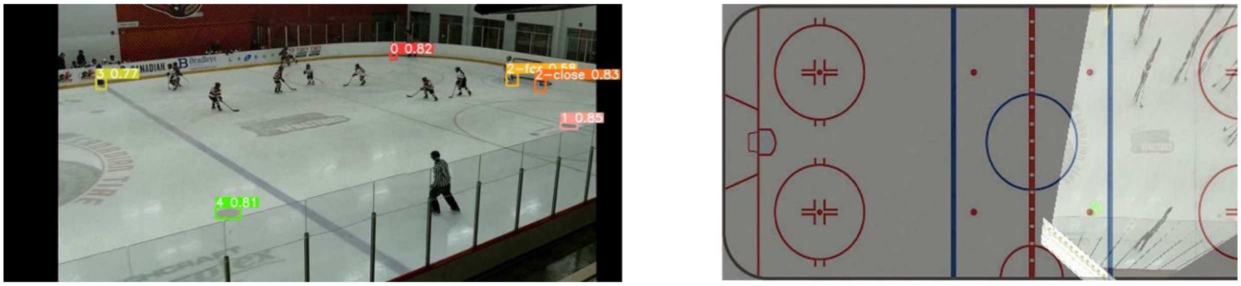

Figure 3-9. Test image of a sample of a video frame recording on right side of the rink.

3.4. Discussion

Medical science professionals [44-46] have identified head trauma especially in youth contact sports as a contributor to increases in mental health problems and risk of neurodegenerative disease. Youth, who represent the vast majority of athletes playing sporting activities have few financial or clinical resources to effectively monitor and manage head trauma. Impact velocity is an important characteristic of head impact severity with accurate impact velocities necessary to obtain accurate head impact severity measures. A top-view frame of players is essential for calculating accurate player plane velocity. This research investigated a novel method employing object detection techniques from various angles using 2d video for localising the ice surface in youth ice hockey games. This localisation enabled an accurate measurement of player velocity during head impacts, a critical factor in determining the magnitude of impact and assessing brain trauma. The entire rink was mapped by matching the visible landmarks from the current frame with those on the top-down view. Applying the homography matrix ensured the whole rink was accurately aligned, as shown in Figures 8 and 9. This study yielded promising results, achieving an IOU score of 0.96 and localising youth ice hockey players without needing multiple cameras or specialised setups, which is described as excellent when compared to research in the field [24,29,37]. The high PR scores observed for landmark detection are likely influenced by the constrained nature of the task and the distinctive visual characteristics of the rink markings in the dataset. In more variable deployments, such as games recorded from different camera positions, with increased zooming, poorer lighting conditions, or differing arena geometries, detection may be more challenging, and somewhat lower performance would be expected. Expanding the training dataset to include examples from these conditions would likely improve robustness and generalisation, as the model would be exposed to a wider range of viewing angles, scales, illumination patterns, and rink appearances.

It is important to note that the homography transformation is generally more accurate near the reference points and might exhibit some error as we move further from the points of reference as a result the extrapolation needed in those areas, The high IOU score indicates that our overall transformation was accurate across the entire rink. The accuracy of the transformation process will be accessed using the player's coordinates from the top-down view to calculate their horizontal velocities and compared direct measures of player horizontal velocities.

In the proposed pipeline, landmark coordinates were defined as the centres of the predicted bounding boxes rather than as regressed single points. This choice was motivated by the appearance of rink landmarks in side-view footage, where markings often occupy small painted regions rather than sharply defined point features. Under these conditions, predicting a bounding region and using its centre can be more stable under changes in viewpoint, motion blur, and partial occlusion. Moreover, homography estimation is primarily governed by the consistency and spatial coverage of multiple correspondences across the rink, rather than by single-pixel accuracy at any one landmark. Consequently, modest localisation errors in individual bounding-box centres can be attenuated by estimating the transformation from multiple detected landmarks and by applying a robust fitting strategy (DLT with RANSAC) to reduce the influence of outliers and noisy detections. As the number of detected landmarks increases, RANSAC is more likely to recover a larger inlier set and refit an overdetermined homography with broader spatial coverage, which generally reduces sensitivity to per-landmark noise and improves the stability of the estimated transformation.

The proposed method involves the application of biomechanical measures from 2d video to create large data sets supporting the investigation of the relationship between head trauma and neurodegenerative disease. The method employed to calculate impact velocities using 2d video support further development of an economic and efficient method to identify and document head trauma, creating large data sets to contribute to research designed to increase safety in sports medicine.

While our research has yielded promising results, it is essential to acknowledge its limitations. One such limitation is that homography estimation requires a minimum of four non-collinear landmark correspondences; therefore, frames with fewer than four detected landmarks cannot be directly processed. In the NISL youth-game recordings used in this study, the camera viewpoint was typically a wide, long-shot broadcast angle with limited zooming, which meant that the four-landmark criterion was satisfied in most frames. However, in datasets that include frequent zoom-in views or tighter crops, the number of visible landmarks can drop below four more often. In such cases, expanding the landmark set (i.e., detecting additional rink markings or structural cues) is recommended to increase the likelihood of having at least four detectable correspondences even during zoomed-in segments. In addition, temporal "gap-filling" strategies could be used for short sequences of frames with insufficient detections, for example, by reusing the most recent valid homography from a neighbouring frame, interpolating homographies across adjacent valid frames, or applying temporal smoothing, under the assumption that camera motion and rink geometry change gradually over short time intervals. These extensions may improve practical coverage while preserving the core homography-based localisation framework.

3.5. Conclusion

Accurate measures of impact velocity in sport head injuries is fundamental to establishing associated risk of brain injury. This research presents a novel method for localising the ice surface in youth ice hockey games employing object detection techniques from various angles using 2 d video. By leveraging computer vision technology, specifically YOLOv5[77], we successfully detected landmarks on the ice surface, enabling accurate localisation without requiring specialised setups or multiple camera arrays. Upon evaluation, the YOLOv5 model achieved a validation precision score of 0.98 and a recall score of 0.97. The mAP (IOU [0.5]) and mAP (IOU [0.5: 0.95]) of the landmark classification model were 98.5% and 64.5%, respectively. As the number of epochs increased, the loss curve gradually stabilised, indicating improved model training effectiveness. These results support the effectiveness of our approach in accurately predicting landmarks from frames taken at various camera angles. F1 and precision-recall curves, evaluating integrated Precision and Recall (PR) indices, confirmed that almost all eight classes perform exceptionally well, with F1 and PR evaluations achieving 0.99 at a confidence level of 80%. These results validate the effectiveness of our approach in accurately predicting landmarks from frames taken at various camera angles.

Subsequently, the IOU score was computed, yielding an average of 96.1. This high IOU score indicates the accuracy of the localisation process, affirming the method's effectiveness in precisely mapping objects to their respective landmarks on the ice surface. The study achieved promising results, with a high IOU score of 0.96, indicating the accuracy of the localisation process.

The proposed method for obtaining impact velocities from various angles using 2 d video provides a promising approach to documenting, managing and studying head impact velocity and the risk of brain injury in youth sport. This study contributes to sports medicine by facilitating the measurement of player velocity during head impacts, crucial for assessing brain trauma opening avenues for further analysis of player performance, injury risk assessment, and trauma management in ice hockey.

Chapter 4. Study II

This chapter is based on the journal paper submitted to *Sports Engineering* and accepted as an oral presentation at the CMBBE Symposium, to be held in September 2025 in Spain. It addresses the objectives of Study II, which focused on developing a robust and accurate detection and tracking system for estimating the planar velocity of players in youth ice hockey games from 2D side-view broadcast footage. The detection and tracking pipeline provided player coordinates from side-view footage, which were then transformed to a top-down perspective using the homography matrix calculated in Study I. By analyzing changes in these transformed coordinates over a sliding multi-frame window, the planar velocity of each player was estimated.

Development of an Automated System to Obtain the Planar Velocity of Individual Players From 2D Video of Youth Ice Hockey Games

Parisa Dehghan, Amirhossein Azadi, Clara Karton, Allison Clouthier, Blaine Hoshizaki

P. Dehghan is a Ph.D. student in Human Kinetics at the Faculty of Health Sciences, University of Ottawa, Canada

A. Azadi is a Ph.D. student in Human Kinetics at the Faculty of Health Sciences, University of Ottawa, Canada

C.Karton is a Research Associate in the Neurotrauma Impact Science Laboratory in the Department of Human Kinetics, University of Ottawa, Canada

A. Clouthier is an Assistant Professor of Biomedical Engineering in the Department of Human Kinetics at the University of Ottawa, Canada

T. B. Hoshizaki is a Full Professor of Biomechanics in the Department of Human Kinetics, Faculty of Health Sciences, University of Ottawa, Canada

Abstract

Head impacts are common in youth sports like ice hockey, often causing brain injuries. However, most concussions go unreported due to the lack of objective monitoring in youth games. Velocity, a key metric for coaching and performance, also affects energy transfer to the head and injury risk. A cost-effective, user-friendly tool for velocity measurement would help with injury prevention and performance analysis. We developed a real-time system to estimate planar velocity of players in youth ice hockey using computer vision. A YOLOv8 model, pre-trained on 80,000 NHL images, was fine-tuned on a 4,700-image youth dataset, achieving 98.5% mAP at IOU 0.5 by labeling players based on jersey color (Light/Dark). Players were tracked with a modified StrongSORT algorithm, reducing ID switch errors from 172 to 53 across 100 clips (90 frames, 30 FPS). MOTA improved from 89.0% to 94.5%. Using a validated rink localization method (IOU = 96%), player positions were transformed to a top-down view. Velocity was estimated by tracking displacement over five-frame intervals in 30 FPS videos. Our system offers a scalable, real-time solution for monitoring motion and assessing impact risks in youth hockey, aiding injury prevention and analysis.

Keywords Homography, Tracking, Velocity Estimation, Object detection

4.1. Introduction

Traumatic brain injury (TBI) is a significant public health concern, particularly in contact sports such as ice hockey, football, rugby, and boxing where high-impact collisions are an inherent part of gameplay [10]. Among these sports, ice hockey stands out due to its fast-paced nature, frequent physical contact, and the potential for sudden decelerations, collisions with other players, and impacts with the boards, glass, or ice [12]. These factors contribute to a high incidence of head injuries, with research indicating that ice hockey has one of the highest concussion rates among contact sports, reaching up to 1.81 concussions per 1,000 athletic exposures [59].

The consequences of TBIs in ice hockey extend far beyond immediate symptoms. Repeated head impacts, both concussive and sub-clinical, have been linked to long-term cognitive, behavioral, and neurological impairments. Chronic traumatic encephalopathy (CTE), depression, early-onset dementia, and motor impairments are among the most severe outcomes associated with repeated brain trauma [24], [25]. TBI severity is influenced by several factors, including the magnitude, frequency, interval between impacts, cumulative trauma, and exposure over an athlete's career [88], [89], [90]. Among these, velocity is an important determinant of impact magnitude, as it directly influences the energy transfer during impacts [91]. However, despite its importance, the relationships between these factors and brain injury outcomes remain poorly understood, largely due to the lack of comprehensive datasets documenting impact characteristics.

The lack of a monitoring system for documenting head impact characteristics is particularly concerning in youth sports, which constitutes the majority of players globally [92]. Concussions in youth hockey are often underreported, as young athletes may fail to recognize symptoms or feel pressured to continue playing despite injuries [93].This issue is compounded by the fact that youth athletes are more vulnerable to long-term damage from repeated head impacts compared to adults, as their brains are still undergoing critical developmental stages [11], [23]. Unlike professional leagues such as the NHL, youth hockey programs rarely have access to advanced diagnostic tools or sideline medical evaluations, making it challenging to accurately assess injury incidence and implement effective prevention strategies.

At the professional level, player velocity is typically measured using specialized setups such as GPS [94], [95], wearable sensors [96], radar systems [97], or multiple high-speed cameras [98], [99] installed around the arena. These technologies provide highly accurate velocity measurements, including depth velocity (movement toward or away from the camera). However, such systems are prohibitively expensive and logistically challenging for youth hockey programs, which often rely on limited resources. Alternative methods, such as using video analysis software like Kinovea [100], estimate velocity by analyzing side-view footage. However, these methods have limitations, including the need for manual tracking of objects or points of interest by the user, which can introduce human error [100]. Additionally, accuracy is compromised if the camera angle is not perfectly perpendicular to the motion being analyzed, as perspective distortion can occur, leading to skewed results. As a result, there is a pressing need for an objective, scalable, and cost-effective tool to monitor and document brain trauma and impact characteristics, such as velocity, in youth athletes. In the current research, we developed an automated system as part of a larger project aimed at monitoring and documenting head impact characteristics in youth ice hockey.

The primary goal of this system is to automatically measure the planar velocity of hockey players from 2D side-view video footage. Our system leverages computer vision techniques to estimate planar velocity from 2D video footage, providing a practical solution for tracking player motion in resource-constrained environments.

4.1.1. Background

Player Tracking and Detection

Early player tracking in ice hockey relied on classical computer vision techniques such as optical flow [101] and Kalman filtering [102]. Optical flow estimates motion by analyzing pixel displacements between consecutive frames but struggles with occlusions, rapid movements, and lighting variations [103]. Kalman filtering improved motion prediction by incorporating temporal dynamics but had limitations in handling abrupt directional changes, which are common in hockey [104]. To address these limitations, particle filters [105] were introduced, which consider multiple hypotheses for player positions, enhancing robustness against noise and occlusions. However, these methods were computationally expensive and struggled with real-time processing, making them less suitable for high-speed sports like ice hockey.

The rise of deep learning revolutionized player detection through convolutional neural networks (CNNs) [106]. Models like YOLO (You Only Look Once) [37] and Faster R-CNN [54]have been widely adopted for detecting players in hockey footage. A key limitation of these models is that they are often trained on NHL footage [107], [108], [109], which features high-resolution broadcasts and consistent lighting. As a result, these models tend to generalize poorly to youth hockey, where lighting conditions, camera quality, and rink markings can differ significantly.

Multi-Object Tracking (MOT) in Ice Hockey

Tracking multiple players in ice hockey presents significant challenges due to fast movements, frequent occlusions, and uniform similarities among teammates. Early tracking methods, such as SORT (Simple Online and Real-time Tracker), relied on bounding box IOU (Intersection over union) matching and Kalman filtering but struggled with ID switches when players moved close to each other or when occlusions occurred [110]. DeepSORT [111]improved upon this by incorporating a CNN-based appearance descriptor, which helped with player re-identification and reduced ID switches[112]. However, because it relies on appearance features, DeepSORT is less effective in ice hockey, where players on the same team wear nearly identical uniforms, making color- and texture-based identification unreliable. Vats et al.[113] compared five tracking algorithms, Tracktor [114], FairMOT [115], DeepSORT [112], SORT [110], and MOT Neural Solver [116], on NHL footage. They found that MOT Neural Solver achieved the highest Multiple Object Tracking Accuracy (MOTA) but was computationally expensive due to the overhead of following long-term tracklets. StrongSORT [39] is an enhanced version of DeepSORT, leveraging advanced techniques such as improved appearance feature extraction and association methods to achieve significantly higher tracking accuracy. It employs motion-based matching and appearance embedding filtering to minimize ID switches [112]. particularly in fast-paced hockey scenarios. However, challenges persist in long-term tracking and jersey number-based re-identification, especially in youth hockey, where lower video quality and varying lighting conditions add further complexity. While commercial solutions exist for player tracking, no publicly available datasets, network architectures, or performance benchmarks exist for youth ice hockey.

Rink Localization and Homography Transformation

Once youth players are detected and tracked, there is a need to transform their side-view coordinates into a top-down view. For this purpose, we built on our previous work [38], which introduced a novel ice rink localization method using a stationary side-positioned camera. To achieve precise rink localization, key landmarks, such as blue line intersections, face-off circles, and goal lines, were manually annotated across 9,000 images. We then employed YOLOv5 for landmark detection, achieving a mean Average Precision (mAP) of 98.5% at IOU@0.5. Using at least four detected landmarks per frame, we calculated a homography [36] matrix to transform side-view footage into a top-down perspective. This transformation allowed us to track player positions accurately in the 2D plane. To validate the method's accuracy, the transformed frames were compared to a top-down rink template, achieving an IOU score of 0.96, confirming the accuracy of this approach for tracking player movements and computing velocities.

This paper advances player tracking and detection in youth ice hockey through the following contributions:

1) **Enhanced Player Detection**

A dataset of 4700 annotated youth hockey images (U13, U18, and U11 levels), labeled by jersey color (Light/Dark), was developed to improve player detection. A YOLOv8 model, pre-trained on 80,000 NHL player images[117], was fine-tuned for the unique challenges of youth hockey, such as varying jersey colors, lighting conditions, and lower video quality.

2) **Improved Multi-Object Tracking (MOT)**

The StrongSORT algorithm was modified to better handle occlusions in youth hockey, leveraging state-of-the-art techniques to enhance tracking robustness.

3) **Velocity Calculation**

Players' horizontal velocity was calculated every 0.16 seconds using side-view footage. A homography transformation was applied to track precise player positions and project them onto a top-view perspective.

The proposed system offers an accurate and cost-effective solution for player detection, tracking, and velocity calculation. It generates valuable datasets for injury research, concussion risk assessment, and performance monitoring in youth ice hockey.

4.2. Methods

4.2.1. Player Detection

Building on our previous work [35], this study refines player detection and tracking by adapting deep learning models to youth hockey. The original YOLOv8-based player detection model was trained on an 80,000-image NHL dataset [117], annotated for players, goalkeepers, referees, jersey colors, and occlusion status (occluded/non-occluded). While the model achieved high precision (97%) on professional-level hockey footage, it struggled to generalize to youth hockey due to differences in lighting conditions, filming angles, and player appearances. To address these limitations, a new dataset specifically tailored for youth hockey games was developed. This dataset comprises 4,700 annotated images collected from 20 U11, U13, and U18 games. Of these, 4,200 images (89%) were randomly allocated for training, while the remaining 500 images (11%) were reserved for validation. A simplified labeling strategy was adopted to exploit the consistent team colour schemes typically observed in youth ice hockey footage. In most of the recorded games, one team wore predominantly light jerseys (most often white) and the opposing team wore predominantly dark jerseys (e.g., red, black, or orange). Rather than labeling individual players or introducing multiple team-specific classes, all player bounding boxes in the training dataset were annotated using only two categories based on jersey appearance: a light-jersey player class and a dark-jersey player class. This choice not only reduced annotation complexity but also provided a practical advantage during detector post-processing. In standard object-detection pipelines, Non-Maximum Suppression (NMS) is applied after the network produces multiple candidate boxes (with confidence scores) to remove highly overlapping, lower-confidence boxes and retain a single best box per object. Because NMS is typically performed independently within each class, separating players into light and dark jersey classes helps preserve detections when players overlap or skate in close proximity. For example, if two players come together (e.g., during a battle along the boards), their detections are less likely to be merged into a single box or suppressed as "duplicates" when they belong to different jersey classes; similarly, overlapping detections from opposing teams are less likely to suppress one another. This, in turn, improves detection continuity during brief collisions or partial occlusions and supports more stable downstream tracking.

The dataset was annotated using Roboflow[118], with five distinct classes: Goalie_Dark, Goalie_Light, Player_Dark, Player_Light, and Referee. This class-based labeling approach enhances tracking stability, particularly in occlusion-heavy scenarios where individual player identification is often challenging.

To optimize detection accuracy, the YOLOv8 model, pre-trained on professional NHL data[117], was fine-tuned on this youth-specific dataset. Training parameters included an image size of 1280x1280 pixels, 250 epochs, and a batch size of 1. While the model retained detailed class assignments internally, detected players were mapped to a generic "Player" category during tracking for simplified visualization and analysis.

4.2.2. Detection Metrics

The performance of the player detection component was evaluated using standard object detection metrics: precision, recall, Average Precision (AP), and Mean Average Precision (mAP). Precision measures the proportion of correctly identified players relative to all detections made, reflecting the system's ability to avoid false positives. Recall quantifies the proportion of actual players that were correctly detected, indicating the system's capacity to minimize missed detections. Average Precision (AP) represents the area under the precision–recall curve for a given

Intersection over Union (IOU) threshold, summarizing the trade-off between precision and recall. In this study, AP was computed at an IOU threshold of 0.7 to evaluate spatial localization accuracy. Mean Average Precision (mAP) generalizes AP across multiple IOU thresholds, with mAP@50 calculated at IOU = 0.50 (a lenient criterion) and mAP@50–95 averaged over thresholds from 0.50 to 0.95 in increments of 0.05 (a stricter, more comprehensive measure). These metrics together provide a balanced assessment of both detection accuracy and robustness across varying spatial tolerances.

4.2.3. Tracking: Enhanced Multi-Object Tracking for Ice Hockey (IOU-Assisted StrongSORT)

In this study, a multi-object tracking (MOT) system was developed to ensure consistent player identification across consecutive video frames. The system builds upon the StrongSORT algorithm[39], an advanced tracking method that integrates deep learning with traditional computer vision techniques. StrongSORT tracks players by combining appearance features with motion predictions using a Kalman filter. However, ice hockey presents unique challenges for MOT, such as frequent occlusions, rapid player movements, and overlapping bounding boxes, which often result in ID switches (loss of player identity).

A major challenge in ice hockey player tracking is the overlapping of bounding boxes, particularly during close interactions and occlusions. Standard Non-Maximum Suppression (NMS) removes overlapping detections by retaining only the bounding box with the highest confidence score while discarding others that exceed a pre-set overlap threshold (typically 50% IOU). However, this approach can cause tracking failures in hockey, where players frequently cluster together. To address this limitation, we introduced a modified NMS approach tailored for ice hockey.

To improve the tracking algorithm, we leveraged the class-based labeling strategy used during detection. By classifying players based on their jersey colors (e.g., Player_Dark, Player_Light), we were able to distinguish between overlapping bounding boxes, ensuring that occluded players were still detected. This step preserved multiple bounding boxes in cases of occlusion without immediately removing them. However, some of these bounding boxes belonged to the same player and needed to be filtered.

To refine the detections while maintaining bounding boxes for other players, we removed redundant bounding boxes corresponding to the same player. Specifically, bounding boxes with an IOU greater than 0.85 were eliminated, as they were determined to belong to the same player. This threshold was established through trial and error to effectively remove duplicate detections while ensuring valid detections for other players remained intact. This approach significantly improved player detection accuracy without introducing excessive false positives.

4.2.4. Minimizing ID Switches for Stable Tracking

Tracking players consistently across frames requires matching new detections with previously established tracks. StrongSORT's default matching process prioritizes appearance-based features when assigning player identities. However, in ice hockey, uniform and helmet similarities create high visual ambiguity, leading to frequent ID switches, where a tracked player's identity is mistakenly reassigned to another player. These errors disrupt player movement continuity and negatively affect velocity calculations, as stable tracking over at least five consecutive frames are necessary for reliable speed estimation.

IOU-assisted association cost to reduce ID switches

To mitigate identity switches (ID switches) during multi-player tracking, the data-association stage of the tracker was augmented with an additional spatial term based on the Intersection-over-Union (IOU) between predicted track boxes and current-frame detections. In each frame, StrongSORT forms a cost matrix C ∈ R^{N×M} where each entry Cij reflects the dissimilarity between an existing track i (propagated to the current frame using its motion model) and a candidate detection j. Although StrongSORT incorporates motion cues, the primary association score is typically dominated by appearance embeddings, which can become ambiguous in youth hockey footage where multiple players share highly similar uniforms, helmets, and body shape, and where broadcast-like viewpoints compress fine-grained visual differences. Under these conditions, the appearance distance between two different players can be comparable to the distance between the correct player across frames, increasing the likelihood of incorrect matches and subsequent ID switches. To address this, an IOU-based cost term was computed for every track–detection pair:

$$C\_IOU_{ij} = 1 - IOU(\hat{b}_i, b_j) \quad \text{Eq. 6}$$

In this expression, $\hat{b}i$ is the predicted bounding box of track i in the current frame (e.g., from the Kalman-filter prediction) and bj is the observed detection box. This spatial cost provides a direct geometric consistency check: the correct detection for a track is expected to overlap strongly with the track's predicted location, whereas mismatched detections, especially those belonging to nearby players, tend to produce lower IOU. The final association cost was then obtained by combining spatial and appearance terms (after normalization to comparable ranges), for example via a weighted sum:

$$Cij = \alpha \cdot C\_appij + (1 - \alpha) \cdot C\_IOUij \qquad \text{Eq. 7}$$

Here, C_appij is the appearance-based distance and $\alpha \in [0,1]$ controls the trade-off. This combined cost matrix is subsequently used by the assignment solver (e.g., Hungarian matching) to select the globally optimal set of tracks–detection matches under one-to-one constraints. Integrating IOU into the association step is particularly beneficial during fast-paced play and brief interactions (e.g., board battles, crossings, and partial occlusions) where appearance features can be unstable due to motion blur, pose changes, and truncation. In these cases, spatial proximity and overlap often remain more reliable than appearance similarity, allowing the tracker to preserve correct identities through short-term ambiguity. In practice, this modification reduced erroneous reassignments between adjacent, visually similar players and improved track continuity, thereby lowering the frequency of ID switches and reducing downstream fragmentation in player trajectories. While the method still depends on detection quality and can be challenged by prolonged occlusions, the added IOU constraint provides a simple and effective spatial regularizer that complements appearance embeddings in visually repetitive sports environments.

4.2.5. Tracking Metrics

The performance of the tracking system was evaluated using standard multi-object tracking metrics, including Multiple Object Tracking Accuracy (MOTA) and ID switches (IDs). MOTA is a comprehensive measure that accounts for false positives, missed detections (false negatives), and identity switches, and is calculated as:

$$\text{MOTA} = 1 - \frac{FN + FP + ID_{sw}}{GT} \qquad \text{Eq. 8}$$

where FN is the number of false negatives, FP is the number of false positives, ID is the total number of identity switches, and GT is the number of ground truth instances. A higher MOTA score indicates better overall tracking performance. ID switches occur when the tracker assigns a different identity to the same player across consecutive frames, and minimizing this metric is critical for maintaining consistent player trajectories, particularly in high-occlusion and fast-paced environments like ice hockey. Together, these metrics quantify the system's ability to maintain accurate and stable player identities over time.

To assess the effectiveness of the adapted tracking system, 100 video clips (3 seconds of a 30 FPS video) were extracted from 20 youth-level ice hockey games. Two configurations were tested:

1) Original StrongSORT
2) IOU-Assisted StrongSORT (Proposed)

Each configuration was applied to the 100 video clips, and ID switches were manually reviewed by two reviewers. An ID switch was logged whenever a player's identity incorrectly changed across consecutive frames, with a focus on maintaining ID stability over 90-frame clips.

4.2.6. Velocity Calculation

Once players are detected and tracked, their positions are estimated by computing the midpoint of the bottom line of each bounding box in the side-view footage. This midpoint serves as a reference point for tracking player movement. Using a homography matrix, derived from landmark detection in each frame, these midpoints are transformed to a top-down perspective, enabling accurate 2D plane localization of each player [38].For each unique player ID, the transformed coordinates are tracked over five consecutive frames. Since the video frame rate is 30 FPS, the time interval for each 5-frame segment is 0.16 seconds. To convert the displacement in the transformed coordinates to real-world distances, the measurements for a standard North American ice rink (as depicted in Figure 4.1) were used for calibration. The average velocity for each player is then calculated by measuring the displacement of their transformed coordinates over this 5-frame period, providing a frame-to-frame velocity profile for each player.

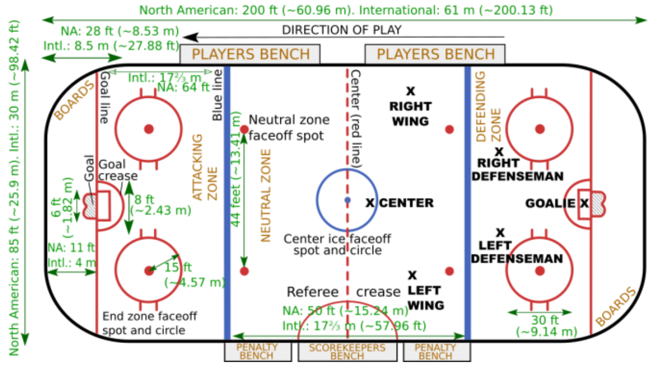

Figure 4-1. Ice hockey rink dimensions [45].

4.3. Results

4.3.1. Player Detection

The performance of the player detection model was assessed using precision, recall, average precision (AP) at an IOU threshold of 0.7, and Mean Average Precision (mAP) at IOU thresholds of 50 (mAP@50) and 50-95 (mAP@50-95), as summarized in Table 4.1. The fine-tuned youth model demonstrated significant improvements in detection accuracy compared to the NHL-trained baseline model when tested on youth hockey footage.

Table 4.1. PLAYER DETECTION PERFORMANCE

Model	Precision	Recall	AP @ 0.7	mAP @ 50	mAP @50-95
Youth Model	0.96	0.97	0.94	0.97	0.82
NHL Model (Baseline)	0.45	0.41	0.18	0.40	0.17

The youth-specific model outperformed the NHL-trained model across all key metrics, highlighting the importance of domain-specific data for accurate player detection in youth hockey. The high precision and recall values (0.96 and 0.97, respectively) indicate that the model effectively identifies players with minimal false positives and false negatives. The mAP@50 of 0.97 further confirms the model's robustness in detecting players under varying conditions.

4.3.2. Tracking Performance

The adapted StrongSORT configuration, incorporating IOU-based cost functions, significantly reduced ID switches, lowering the total number from 172 to 53, as depicted in Table 4.2. This represents a 69.2% reduction in ID switches, demonstrating the effectiveness of the proposed modifications in maintaining player identities during fast-paced gameplay and occlusions. Additionally, MOTA in occlusion-heavy clips improved from 89.0% to 94.5%, further validating the robustness of the tracking system.

Table 1.2. TRACKING PERFORMANCE

Tracking Configuration	Average ID switches (Total across 100 clips)	MOTA
Original StrongSORT	172	89.0%
IOU-Assisted StrongSORT	53	94.5%

4.3.3. Velocity Calculation and Motion Tracking

As illustrated in Figure 4.2, the system successfully detects players in the side-view footage, with the midpoints of their bounding boxes representing the players' coordinates. Through the automated homography transformation process, these positions are converted to a top-down view, enabling precise motion tracking. The velocity of each player is calculated for every consecutive set of frames, generating a continuous velocity profile, as depicted in Figure 4.3.

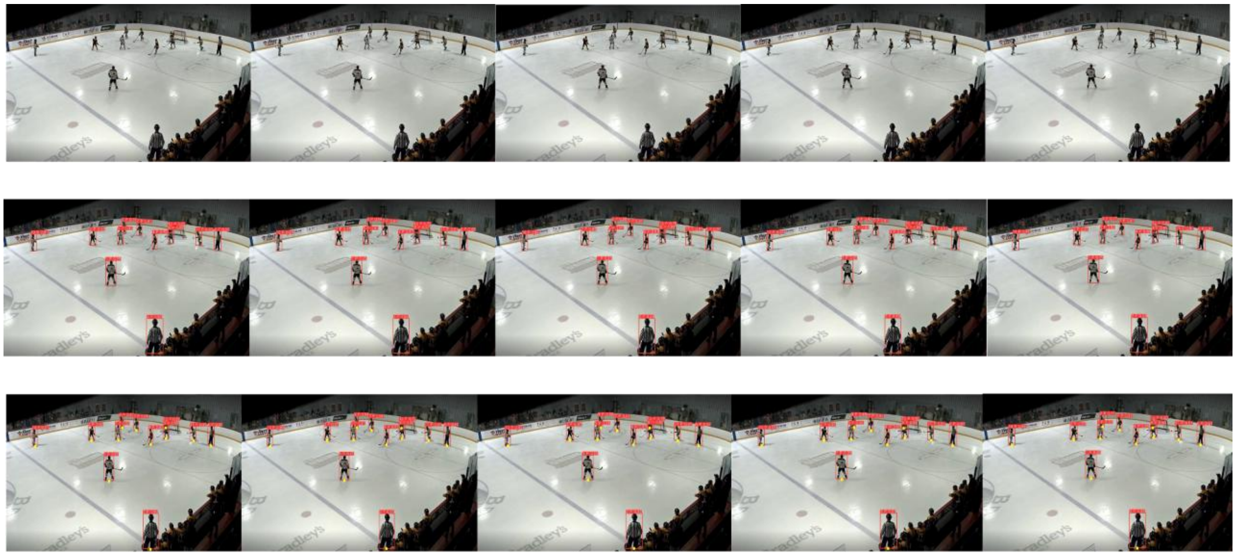

Figure 4-2. The first row displays five consecutive frames from a youth ice hockey game. The second row shows the corresponding frames with players detected using the fine-tuned YOLOv8 model, where bounding boxes are drawn around each player. The third row illustrates the midpoint of each bounding box, representing the player's estimated point of contact with the ice.

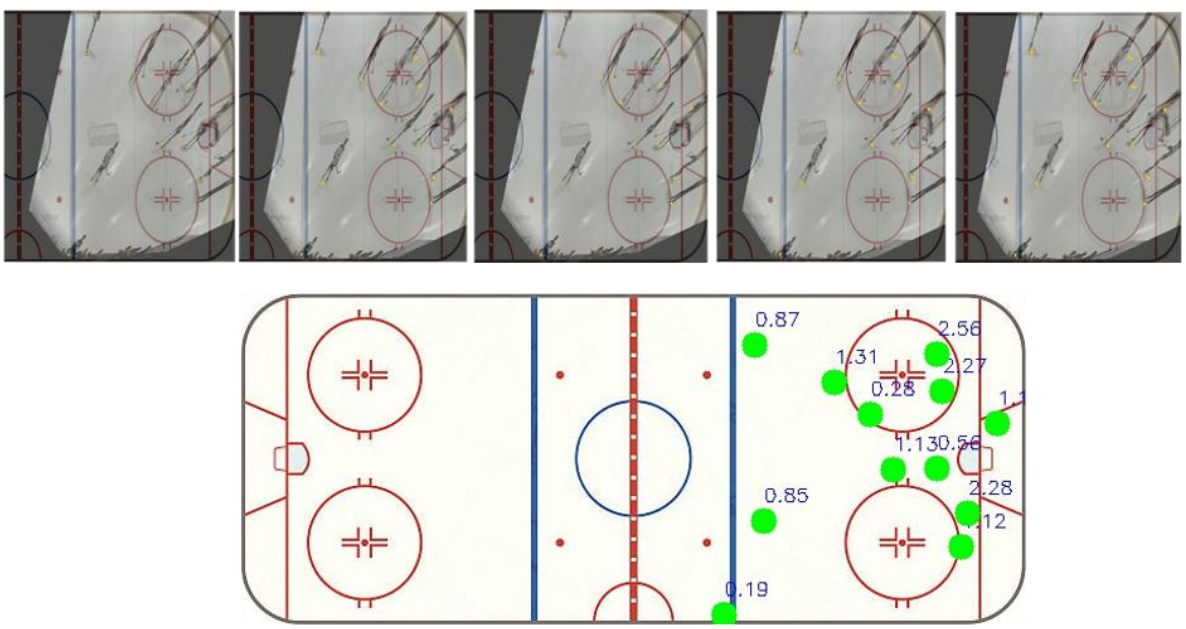

Figure 4-3. This figure displays the top-down view of five consecutive frames from Figure 4.2, with points representing the midpoints of the bounding boxes' bottom lines, indicating player positions. The bottom image shows the average velocity of players in meters per second (m/s) over the 5-frame interval.

4.4. Discussion

The development of an automated system to obtain the planar velocity of individual players using 2D videos in youth ice hockey represents a significant advancement in sports analytics and injury prevention. This study addresses a critical gap in the monitoring and documentation of head impact characteristics, particularly in youth hockey, where access to advanced diagnostic tools is limited. The proposed system leverages computer vision techniques to provide a cost-effective, scalable, solution for tracking player motion, offering valuable insights into player dynamics and velocity for head impact assessment. To our knowledge, this is the first system capable of calculating planar velocity in youth ice hockey, providing a unique advantage over existing methods.

The fine-tuned YOLOv8 model demonstrated significant improvements in detection accuracy when applied to youth hockey footage, achieving precision and recall values of 0.96 and 0.97, respectively. These results highlight the importance of domain-specific training data, as the NHL-trained baseline model struggled to generalize to youth hockey due to differences in lighting, camera angles, and player appearances. The high mAP@50 score of 0.97 further confirms the model's robustness in detecting players under varying conditions in youth games, including smaller player sizes, darker lighting, and different camera viewing angles, making it a reliable tool for player tracking in youth hockey.

The modified StrongSORT algorithm, incorporating IOU-based cost functions and momentum-based matching, significantly reduced ID switches by 69.2% compared to the original StrongSORT implementation. This improvement is particularly important in ice hockey, where frequent occlusions and rapid player movements often lead to tracking failures. The reduction in ID switches enhances the continuity of player tracking, which is essential for accurate velocity calculations and long-term motion analysis. Additionally, the improvement in MOTA from 89.0% to 96.5% in occlusion-heavy clips further validates the robustness of the proposed tracking system. The system's ability to calculate player velocity every 0.16 seconds using homography transformation provides a detailed and continuous velocity profile for each player. While this study used a 5-frame interval for velocity calculation, the system is flexible and can be adapted to other frame intervals as needed. The ability to accurately track player movements and calculate velocities has significant implications for injury prevention in youth ice hockey. By integrating this system into a broader framework for head impact monitoring, velocity data can be combined with other impact characteristics, such as location, event type, and direction, to provide a comprehensive understanding of brain trauma. This data can help researchers and coaches identify patterns associated with high-risk impacts, assess the severity of collisions, and implement targeted interventions to reduce the incidence of traumatic brain injuries (TBIs). For example, velocity profiles can inform rule changes, training programs, and equipment modifications aimed at minimizing injury risks.

4.4.1. Limitations and Future Work

While the proposed system represents a significant step forward, it is important to acknowledge its limitations and identify areas for future improvement:

- Validation of Velocity Accuracy: The system's velocity calculations need to be validated against real-world velocity data obtained from top-down cameras or other ground truth methods. This will ensure the accuracy of the midpoint assumption (using the bottom line of bounding boxes) in transformation process.
- Generalization to Other Sports: While the system is tailored for ice hockey, its underlying principles could be adapted for other sports with similar tracking challenges, such as soccer, basketball, or rugby. Future studies could explore the applicability of the system in these contexts.

4.5. Conclusion

This study presents a novel and effective approach to player tracking and velocity calculation in youth ice hockey using 2D video footage. By leveraging deep learning and computer vision techniques, the system provides a scalable, and cost-effective solution for monitoring player dynamics and assessing injury risks. The improvements in detection accuracy, tracking stability, and velocity calculation demonstrate the system's potential to enhance injury prevention strategies and contribute to the long-term health and safety of youth athletes.

Future work will focus on addressing the current limitations, including validating velocity accuracy, generalizing the system to other sports, and improving long-term tracking capabilities.

By expanding the system's functionalities and applications, this research lays the groundwork for advancing sports analytics and injury prevention efforts, ultimately benefiting youth athletes worldwide.

Chapter 5. Study III

This chapter is based on a journal paper submitted to *Sports Biomechanics*. It addresses the objectives of Study III, which focused on validating the velocity estimates from our proposed pipeline using side-view footage, as commonly found in youth hockey broadcasts. These estimates were compared against ground-truth velocities from a top-down, rink-mounted camera under real-world conditions, covering various movement types and player distances. The results provide preliminary validation, showing reasonable agreement between the two measurement methods.

Validation of Deep Learning-Based Velocity Estimation from Single, Side-View Ice Hockey Game Footage Using Top-Down Drone Measurements

Parisa Dehghan, Amirhossein Azadi, Clara Karton, Robert Laganière, Ryan Graham, Allison Clouthier, Blaine Hoshizaki

Abstract

This study presents a validation of a previously proposed deep learning-based pipeline for estimating player planar velocity from side-view broadcast footage in ice hockey. The automated system consists of three modules: (1) a rink localization module that computes a frame-specific homography to transform side-view coordinates to a top-down perspective; (2) a player detection and tracking module that extracts player trajectories; and (3) a velocity calculation module that computes average velocity over 10-frame intervals based on transformed coordinates. To validate this method, drone-mounted top-down recordings were used as ground truth. Drone-based player velocity was manually computed in Kinovea by tracking a helmet-mounted marker over synchronized 10-frame intervals. Estimated velocities from both fixed and panning side-view cameras were compared against these reference values across four rink regions: two defensive faceoff spots and two central zones (near and far from the camera). The evaluation included lateral, forward (toward the camera), and diagonal movements. A zoomed-in camera configuration was also tested in the far central zone to examine the impact of reduced landmark visibility. Results showed that both panning and fixed camera setups achieved mean absolute errors below 10%, while the zoom-in configuration produced errors under 20%. A sensitivity analysis assessed the influence of player localization error by perturbing coordinates by ±5, ±10, and ±15 pixels in both horizontal and vertical directions. Velocity estimation was repeated across frame intervals of 5, 10, 20, and 30. Findings revealed that longer frame windows reduced the impact of localization noise and that regions farther from the camera were more sensitive to perturbations. This validation demonstrates the pipeline's accuracy and practicality as a scalable, low-cost solution for velocity measurement in hockey, enabling broader access to head impact data collection in youth sports.

Keywords Validation, Velocity, Head impact, Homography

5.1. Introduction

Among contact sports, ice hockey is characterized by high-speed skating, frequent body collisions, and abrupt directional changes, which contribute to a high incidence of head impacts[59]. Such impacts can cause a diverse range of traumatic brain injuries (TBIs), including concussions, as well as more complex and progressive neurological disorders like chronic traumatic encephalopathy (CTE)[12]. Concussions, in particular, are prevalent in ice hockey. At the youth and collegiate levels, the incidence of concussion in men's ice hockey has been reported up to 1.63 per 1,000 athlete exposures, while women's ice hockey reports higher rates, of 2.27 per 1,000 exposures [119]. The prevalence of these injuries has resulted in research efforts to understand their biomechanical mechanisms.

Multiple factors have been identified to influence the severity and outcome of head impacts. Key variables include the magnitude of the impact, direction (e.g., frontal, lateral, occipital), compliance of the impacted surface (such as boards, ice, or another player), duration of exposure, and the cumulative number of impacts sustained over a player's career[88], [89], [90]. Among these factors, player velocity at the time of head impact is particularly important, as it directly affects the amount of kinetic energy transferred to the head, contributing to the potential for injury[91]. In a recent study, Azadi et al. (2025) [30] identified inbound velocity as the most influential predictor of Maximum Principal Strain (MPS), a widely accepted biomechanical metric for estimating brain tissue deformation and injury risk. Inbound velocity surpassed other predictor variables including impact location, surface compliance, and player age group. Although velocity is an influential factor in brain injuries, the precise relationship between impact characteristics, such as velocity, and injury outcomes remains poorly understood. To advance our understanding, it is important to develop automated methods that can accurately measure player velocity and systematically collect large datasets. With the support of deep learning techniques and large-scale data, we can begin to uncover the complex patterns underlying brain injury mechanisms. These findings highlight the need for an objective, automated velocity measurement system capable of efficiently collecting large datasets on head impact characteristics, including velocity, and their corresponding injury outcomes. Such a tool would enhance our understanding of the mechanisms behind sports-related brain trauma, ultimately supporting more effective injury prevention strategies and informed rule changes.

Traditional velocity measurement techniques such as radar guns [97], GPS tracking systems [94], [95], and multi-camera motion capture [98], [99] provide expensive and complex methods of measuring impact velocity. These technologies are primarily employed in professional sports and laboratory settings. These limitations emphasize the need for accessible, scalable alternatives that can be readily applied in community sports settings across all age groups, including youth and children. Video-based motion analysis has emerged as a promising alternative. Tools like Kinovea [100], which support two-dimensional (2D) motion analysis using conventional video footage, are commonly employed in sports research. However, these tools rely heavily on manual calibration and user input, making them impractical for large-scale or fully automated studies. Additionally, Kinovea's dependence on manual frame-based axis calibration can introduce geometric inaccuracies, especially when the camera is moving or when subjects are located at varying distances from the camera [34]. Although such tools can be effective for biomechanical modeling and simulations, where player velocities are typically categorized into broad velocity bins (e.g., 0–2 m/s, 2–4 m/s, 4–6 m/s) [120], [121], manual velocity estimation across large datasets is labour-intensive and limits scalability.

Recent advances in computer vision and deep learning have created opportunities for automating the extraction of kinematic data from video footage. In our previous work [38], we developed a novel deep learning-based pipeline to estimate planar player velocity from side-view game recordings, providing average velocity measurements over specific consecutive frame segments. Although this method was initially validated through computational accuracy metrics at each step of the pipeline, further validation is required to assess its accuracy in real-world conditions. In this validation study, we aim to evaluate the accuracy of the proposed velocity estimation method across different rink regions and movement types. Specifically, we compare the velocity estimates generated by the side-view pipeline against drone-based top-down measurements, which serve as the ground truth. This validation is essential to understand how the system performs when players move at varying distances and angles relative to the side-view camera.

5.1.1. Background of the Proposed Pipeline

The player velocity estimation pipeline consists of two primary modules: (1) rink localization and (2) player detection and tracking. The rink localization module is responsible for computing a homography matrix for each frame, which enables the transformation of the side-view footage into a reference top-down view. The player detection and tracking module is designed to localize and follow players in each frame. By applying the homography transformation to the player coordinates, their positions are mapped onto the top-down view, from which the average velocity is calculated over a certain number of frames.

5.1.2. Rink Localization Network

The rink localization module is built upon a YOLOv5-based [37] object detection network trained to identify specific, high-contrast landmarks within the ice hockey rink. These include the intersections of blue lines and center lines with the boards, all faceoff spots, goalposts, and the intersection of the center line with the center circle. These landmarks serve as anchor points to establish geometric correspondence with a standardized top-down rink template. A custom alignment algorithm matches the detected landmarks in each side-view frame to their known coordinates in the top-down view, enabling the computation of a frame-specific homography matrix. This transformation matrix is then used to warp the side-view frame into a virtual top-down perspective, ensuring spatial consistency across frames and improving the accuracy of kinematic calculations.

5.1.3. Player Detection and Tracking

Player positions are identified in each frame using a YOLOv8 [37] object detector. To maintain temporal coherence and accurately track player trajectories, the pipeline integrates a proposed IOU-assisted StrongSORT [39] tracking algorithm, which combines appearance-based features with spatial similarity measures to preserve player identity across frames. For each detected player, the midpoint of the bounding box's bottom edge is used to approximate the skate contact point on the ice surface. These 2D coordinates, initially extracted in the side-view frame, are then projected onto the top-down coordinate system using the homography [36] matrix calculated in the previous step. Player velocity is estimated by calculating the Euclidean distance between player positions over a certain frame interval, such as 10 frames, divided by the corresponding time duration. This process operates as a sliding window, where the average velocity is calculated over each 10-frame segment (from the current frame to the 10th frame ahead). This approach provides an estimate of the player's planar (horizontal) velocity in meters per second (m/s). An overview of the proposed pipeline is illustrated in Figure 5.1.

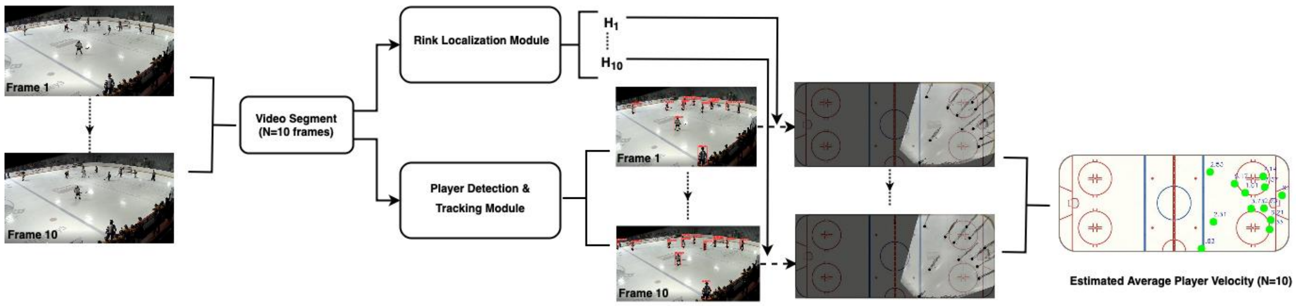

Figure 5-1. Overview of the Velocity Estimation Pipeline Using Side-View Video

5.1.4. Study Objective

The objective of this study was to validate the proposed video-based velocity estimation framework by comparing its values to ground-truth velocity measurements captured using a drone-mounted, top-down camera. Controlled skating trials were conducted in which a player performed a variety of movements, including forward, lateral, and diagonal skating, across four predefined regions of the rink. The pipeline's performance was evaluated under two commonly encountered video acquisition scenarios: (1) a panning side-view camera, simulating broadcast-style footage, and (2) a stationary side camera, representative of fixed arena installations. Additionally, in one region, zoom-in footage was analyzed to assess the impact of zoom level on the pipeline's performance.

5.2. Methodology

5.2.1. Experimental Design and Data Collection Setup

To validate the pipeline's player velocity estimations derived from side-camera footage, controlled on-ice skating trials were conducted at the ice hockey rink of the Minto Sports Complex Fitness Centre, University of Ottawa, Canada. A single skater performed a series of movement patterns specifically designed to test the system's ability to estimate velocity across various directions relative to the camera. These movements included forward motion toward the camera, lateral movements (left-to-right and right-to-left), and diagonal paths both approaching and moving away from the camera. The trials were conducted across four distinct zones (A–D), which were carefully selected to represent different distances and viewing angles relative to the side-view camera. These zones enabled evaluation of the system's performance under varying spatial conditions and provided insight into how distance from the camera influences velocity estimation accuracy. The layout of the analyzed zones and the directions of movement within each are illustrated in Figure 5.2. Example frames captured from the stationary side-view camera, showing the four movement types and zone-specific perspectives, are provided in Figure. 5.3 and 5.4.

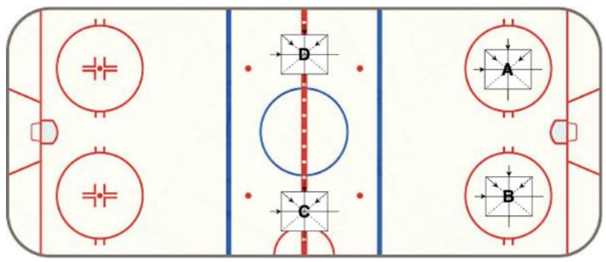

Figure 5-2. Layout of the analyzed regions (A, B, C, D) and illustration of the movement directions performed in each region.

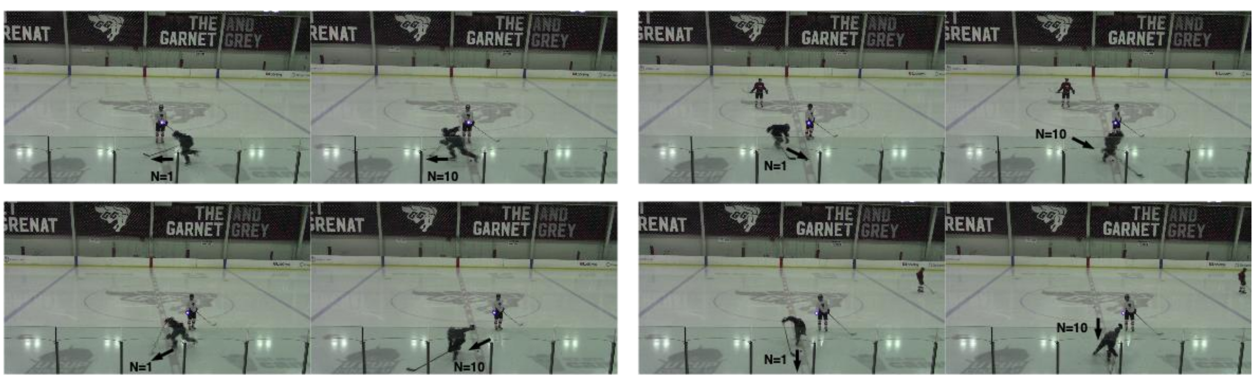

Figure 5-3. Sample movement types in Region C captured from the stationary side-view camera, showing player positions at n = 1 (frame 1) and n = 10 (frame 10) within a 10-frame velocity window. Top left: Lateral movement (right to left); Top right: Diagonal movement 1; Bottom left: Diagonal movement 2; Bottom right: Movement towards the camera.

Figure 5-4. Illustrating the four different regions (A, B, C, and D) for the "towards-camera" movement

For each zone, a rectangular region was marked on the ice. The skater was asked to start from rest and perform each movement while passing through the center of the rectangular region, which was positioned directly beneath a DJI Air 2S drone capturing synchronized top-down footage. This drone footage served as the ground truth for real-world velocity estimation, ensuring precise validation of the side-camera measurements. Simultaneously, two Canon VIXIA HF R700 video cameras positioned in the rink's broadcast area recorded side-view footage for the velocity estimation pipeline. One camera remained stationary throughout the trials, while the other was configured to pan, following the skater's movement. An additional set of trials was conducted in Zone D using a zoomed-in camera to investigate the impact of zoom level on velocity estimation accuracy. The experiment consisted of 20 movement sequences: 16 covering four movement types across the four zones, and four additional zoomed-in trials in Zone D. All videos were recorded at 30 frames per second with a resolution of 1920 × 1080 pixels. A 10-frame window was selected for velocity calculation as it captures key events, including head impacts, while covering approximately 0.33 seconds of video at 30 FPS. To evaluate the sensitivity of velocity estimates to the choice of temporal window, an additional analysis was performed using 5, 10, 20, and 30 frame intervals across all defined rink regions. A flashing light at the beginning of each trial was used to synchronize the three video streams. For consistency, the same temporal window, starting approximately 10 to 20 frames after the synchronization flash, was applied across all three cameras. Side-view footage was processed through the velocity estimation pipeline, while the drone footage served as ground truth for comparison.

5.2.2. Velocity Estimation Pipeline

The proposed pipeline consisted of three main stages. First, player detection and tracking were performed on the side-view footage using a bounding box, with the midpoint of the bottom edge used as the player's reference point. Second, the detected player coordinates were transformed into a standardized top-down view using a homography matrix generated by the rink localization module. This transformation projected the coordinates onto a reference plane of 1980 × 834 pixels, corresponding to the real-world dimensions of a North American ice rink, as illustrated in Figure 5. Third, for each movement within each region, a single velocity value, representing the average velocity over the selected 10-frame window, was calculated based on the displacement between the first and tenth frames. The velocity estimation process for a sample 10-frame window for a lateral movement in Region B is illustrated in Figure 5.6.

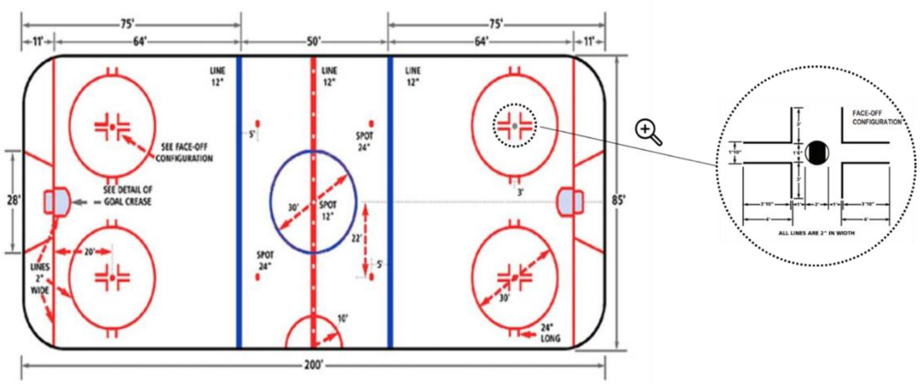

Figure 5-5. Dimensions of the ice rink used for camera calibration

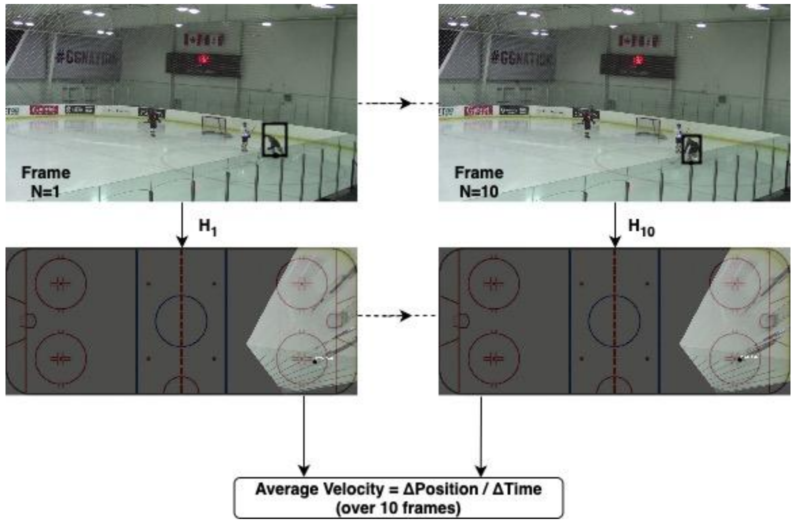

Figure 5-6. Example of velocity estimation for lateral movement performed in Region B. The top row shows the first (n = 1) and tenth (n = 10) frames captured from the stationary side-view camera. The middle row illustrates the transformation of these frames to the top-down view using the homography matrix calculated by the rink localization module. Velocity is then calculated based on the change in player coordinates over time.

5.2.3. Ground Truth Velocity Calculation

Ground truth velocities were established using the synchronized top-down drone footage. Helmet-mounted marker displacements were manually tracked frame-by-frame using Kinovea 0.9.5 [100] motion analysis software, as illustrated in Figure 5.7. Spatial calibration of the drone footage was performed using known rink features, such as the dimensions of face-off circles and hash mark spacing, as detailed in Figure 5.5 This calibration allowed for the calculation of real-world velocities by measuring coordinate displacements over the same 10-frame intervals synchronized across all camera views. These drone-derived velocities served as the ground truth reference for evaluating the accuracy of the side-view velocity estimates produced by the proposed pipeline.

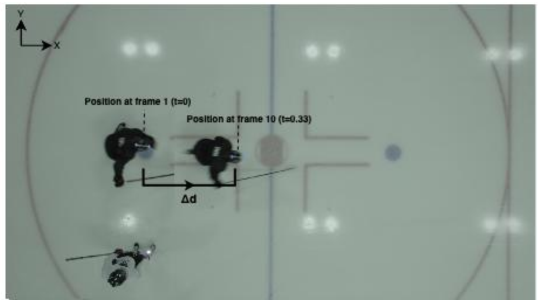

Figure 5-7. Frames captured by the drone showing the first (n = 1) and tenth (n = 10) frames of the movement sequence. The distance traveled between these frames was measured to calculate the ground truth velocity.

5.2.4. Pipeline Accuracy Assessment and Robustness Testing

To rigorously evaluate the performance of the proposed velocity estimation pipeline, three experimental comparisons were conducted: (1) panning side-view footage versus drone-derived velocity, (2) stationary side-view footage versus drone-derived velocity, and (3) zoomed-in side-view footage (Zone D) versus drone-derived velocity. Accuracy was assessed using mean absolute error (MAE), a standard metric for quantifying the average deviation between estimated and ground truth values, along with the percentage error relative to the ground truth velocity. In each trial, MAE was computed by comparing the single velocity value estimated over a 10-frame window to the corresponding drone-derived ground truth. Performance was evaluated across all trials, within each rink region, and for each movement type to ensure a comprehensive assessment under varying spatial and motion conditions.

Agreement between the proposed velocity estimation pipeline and the drone-derived ground truth was assessed using the Bland–Altman method. For each paired observation, the velocity difference was computed as the drone-derived value minus the pipeline estimate, and the mean velocity of the two methods was calculated. The bias (mean difference) and the 95% limits of agreement (LOA) were determined as bias ± 1.96 × the standard deviation of the differences, with all values expressed in meters per second (m/s). The normality of the differences was evaluated using the Shapiro–Wilk test to confirm the assumptions of the parametric Bland–Altman approach. In the absence of evidence for non-normality, the standard parametric method. was applied. Analyses were conducted separately for the fixed and panning camera configurations to characterize systematic bias and the expected range of differences relative to the high-fidelity drone reference.

Since player location was defined by the midpoint of the bottom edge of the bounding box, this reference point could vary due to changes in posture or movement. To assess the effect of this localization variability, and to examine the influence of the number of frames used for velocity calculation, a sensitivity analysis was conducted. Controlled perturbations of ±5, ±10, and ±15 pixels were applied to the player's coordinates in both the x (horizontal) and y (vertical) directions, while keeping the homography transformation fixed. This was repeated over frame windows of 5, 10, 20, and 30 frames. This analysis enabled a detailed quantification of the pipeline's sensitivity to small localization errors and to the choice of temporal window length. Together, the experimental comparisons and sensitivity analyses provided a robust validation framework for assessing the pipeline's performance across different camera configurations, viewing distances, and movement scenarios, with alignment to high-fidelity ground truth measurements.

5.3. Results

The data include 16 movements captured under fixed and panning camera conditions, along with the four movements recorded in zoom-in configuration (Region D). Across all trials, the participant's velocities ranged between 4 and 6 m/s. Quantitative results, including the absolute error and percentage error for each movement in each region, as well as the mean absolute error (MAE)[122] and mean percentage error for each camera configuration and overall, are summarized in Table 5.1. The panning side-view camera achieved the highest accuracy, with a MAE of 0.37 ± 0.26 m/s and a mean percentage error of 8.33%. The fixed camera configuration showed slightly lower performance, with a MAE of 0.43 ± 0.27 m/s and a mean error of 9.41%. The zoom-in configuration in Region D yielded the largest errors, with a MAE of 0.77 ± 0.62 m/s and a mean error of 19.74%.

When performance was analyzed by region, reflecting varying distances and camera viewing angles, velocity estimation errors were generally consistent, except in Region C, which showed the poorest performance. In this region, the panning camera produced a mean absolute error (MAE) of 0.49 m/s with a mean percentage error of 10.82%, while the fixed camera exhibited a higher MAE of 0.78 m/s and a mean percentage error of 17.54%. These elevated errors in Region C contributed substantially to the overall variability and higher standard deviations observed in the results. Further analysis identified lateral movements in Region C as the primary source of these increased errors. Both camera configurations yielded absolute errors of approximately 0.9 m/s and a mean percentage error of 22% for this movement type. A more detailed breakdown by movement type, summarized in Table 5.2, confirmed that lateral movements consistently produced the highest velocity estimation errors across all regions and camera configurations. This issue was particularly pronounced in the zoom-in setup, where lateral movements resulted in the largest deviations from ground truth.

Table 2.1. Absolute errors and mean absolute errors (MAE) for all regions (A–D) and four movement types across the three camera configurations: fixed, panning, and zoom-in.

Region	Movement	Panning Camera		Fixed Camera		Zoom-In Camera	
		Absolute Error (m/s)	% Error	Absolute Error (m/s)	% Error	Absolute Error (m/s)	% Error
A	Lateral Movement	0.01	0.2	0.72	16.74	N/A	
	Towards Camera	0.14	2.5	0.11	2.18	N/A	
	Diagonal Movement 1	0.6	13.86	0.31	6.76	N/A	
	Diagonal Movement 2	0.46	8.21	0.14	2.69	N/A	
	MAE ± SD	**0.3 ±0.27**	**6.19 ±6.12**	**0.32 ±0.28**	**7.09 ±6.75**		
B	Lateral Movement	0.4	8.47	0.33	6.86	N/A	
	Towards Camera	0.22	4.97	0.36	8.4	N/A	
	Diagonal Movement 1	0.29	6.33	0.45	10.11	N/A	
	Diagonal Movement 2	0.18	3.18	0.42	7.99	N/A	
	MAE ± SD	**0.27 ±0.1**	**5.74 ±2.23**	**0.39 ±0.05**	**8.34 ±1.35**		
C	Lateral Movement	0.99	24.69	0.9	21.95	N/A	
	Towards Camera	0.59	12.33	0.86	19.07	N/A	
	Diagonal Movement 1	0.36	6.25	0.55	11.34	N/A	
	Diagonal Movement 2	0	0	0.8	17.78	N/A	
	MAE ± SD	**0.49 ±0.42**	**10.82 ±10.53**	**0.78 ±0.16**	**17.54 ±4.48**		
D	Lateral Movement	0.36	7.44	0.4	8.33	1.5	42.86
	Towards Camera	0.7	19.79	0.15	3.51	0.93	17.83
	Diagonal Movement 1	0.3	7.5	0.1	2.38	0.61	16.94
	DiagonalMovement 2	0.3	7.5	0.2	4.44	0.06	1.31
	MAE ± SD	**0.42 ±0.19**	**10.56 ±6.16**	**0.21 ±0.13**	**4.67 ±2.58**		
	Total MAE ± SD	**0.37±0.26**	**8.33±6.65**	**0.43±0.27**	**9.41±6.34**	**0.78±0.6**	**19.74±17.18**

Table 5.2. Mean absolute errors (MAE) for each movement type across all regions (A–D), reported for the fixed and panning camera configurations

Movement	Panning Camera		Fixed Camera	
	Absolute Error (m/s)	% Error	Absolute Error (m/s)	% Error
Lateral Movement	0.44	10.2	0.59	13.47
Towards Camera	0.41	9.9	0.37	8.29
Diagonal Movement 1	0.39	8.49	0.35	7.65
Diagonal Movement 2	0.24	4.72	0.39	8.23

5.3.1. Bland–Altman Analysis

Bland–Altman analysis was conducted to evaluate agreement between the proposed velocity estimation pipeline and the drone-derived ground truth for both camera configurations. For the fixed camera (Figure 5.8), the mean difference (bias) was 0.4 m/s, indicating a slight systematic underestimation relative to the drone reference. The 95% limits of agreement ranged from –0.3 to 1 m/s, suggesting that most individual estimates fell within approximately ±1 m/s of the bias. For the panning camera (Figure 5.9), the bias was 0.2 m/s, also reflecting a small underestimation, with limits of agreement from –0.5 to 1.00 m/s. The Shapiro–Wilk test indicated no evidence of deviation from normality in the differences for either configuration ($p = 0.40$ for fixed; $p = 0.95$ for panning), supporting the validity of the parametric Bland–Altman method. No significant proportional bias was detected ($p > 0.05$), suggesting that the degree of agreement did not vary systematically with mean velocity.

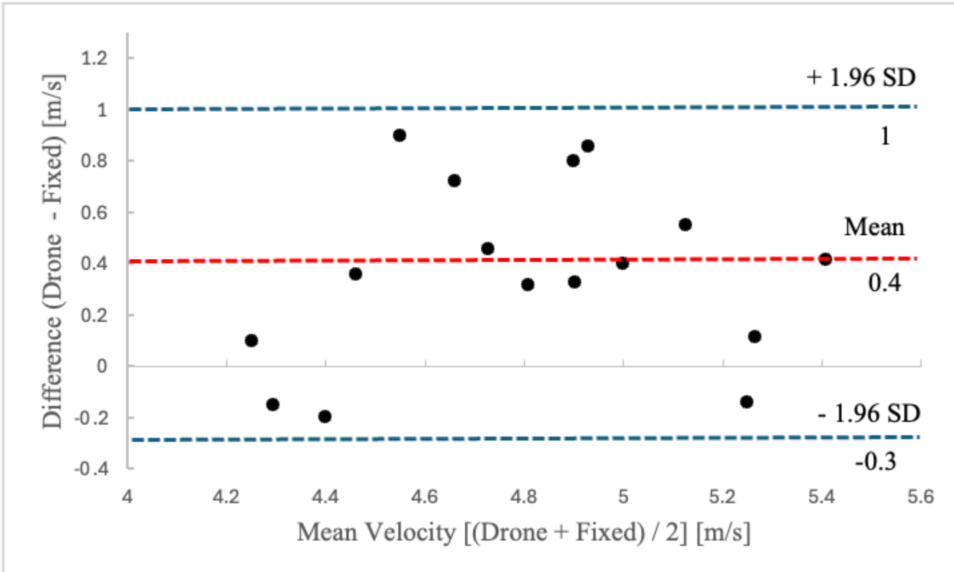

Figure 5-8. Bland-Altman plot of agreement between Drone and fixed camera

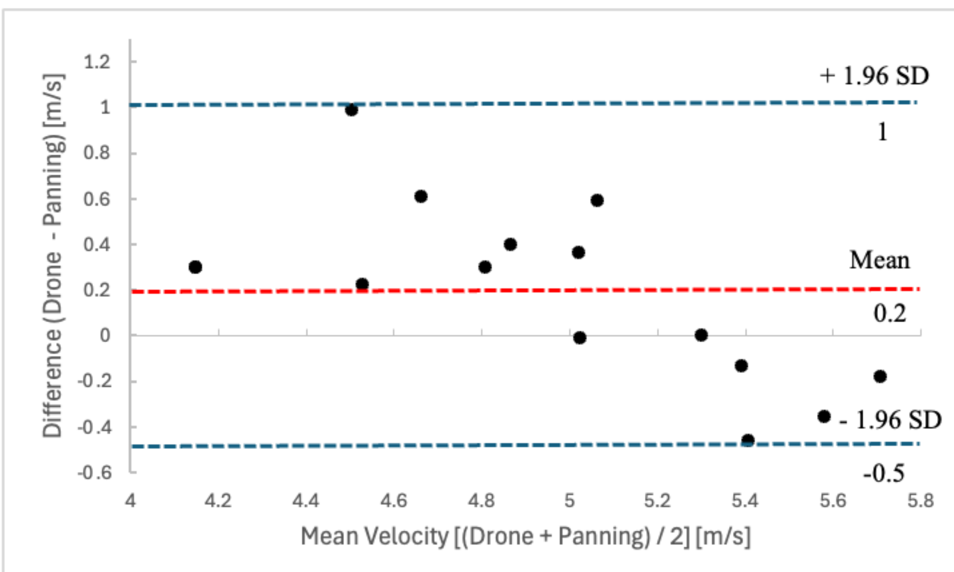

Figure 5-9. Bland-Altman plot of agreement between Drone and Panning camera

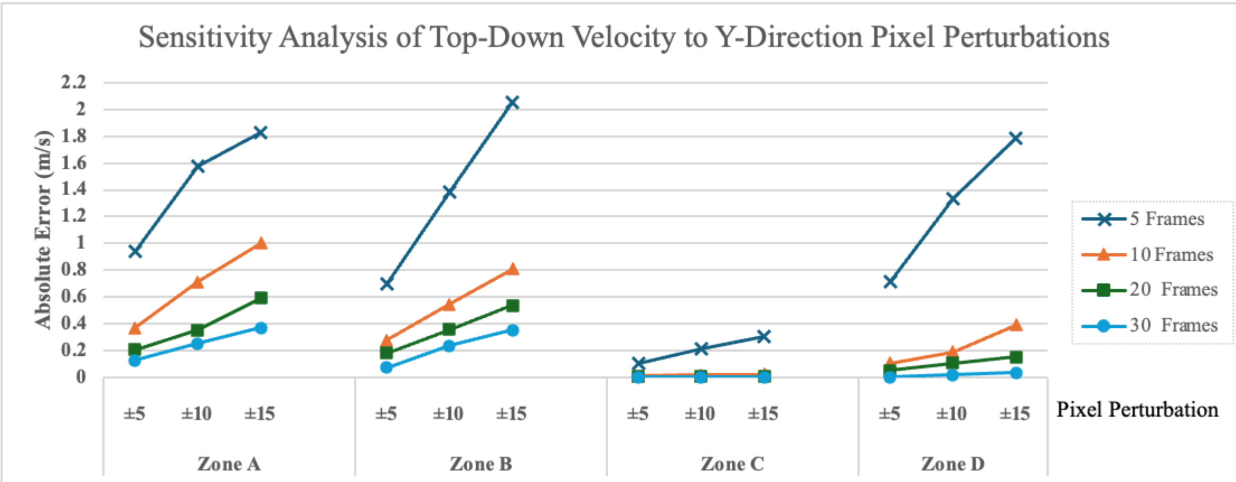

Figure 5-10. Analysis of velocity changes caused by perturbing the y-coordinate by ±5, ±10, ±15 pixels across different frame intervals (5, 10, 20, and 30 frames) in each rink region (Zones A–D).

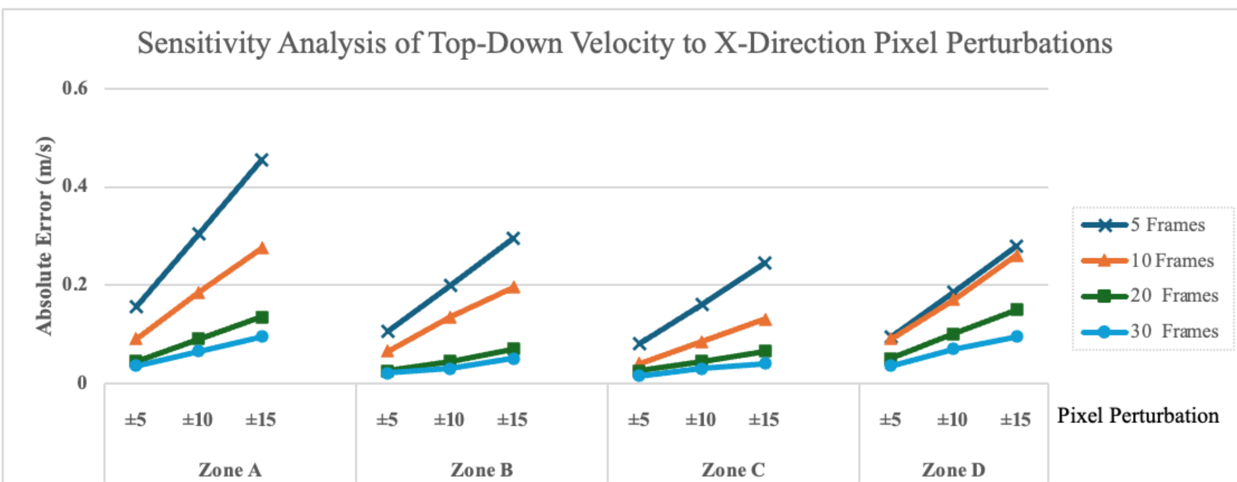

Figure 5-11. Analysis of velocity changes caused by perturbing the x-coordinate by ±5, ±10, ±15 pixels across different frame intervals (5, 10, 20, and 30 frames) in each rink region (Zones A–D).

5.3.2. Sensitivity Analysis

Due to the reliance on bounding boxes generated by the player detection network, variations in bounding box size and position, arising from changes in player posture or shifts in camera perspective, introduce localization errors that affect the accuracy of velocity estimates. These inconsistencies contribute to the high standard deviation observed in the velocity estimation error, as presented in Table 5.1 To systematically investigate the sensitivity of the proposed method to such localization errors, and to evaluate the influence of the frame window length on velocity accuracy, a perturbation analysis was conducted. In this analysis, the player's detected coordinates were artificially perturbed by ±5, ±10, and ±15 pixels along both the horizontal (X) and vertical (Y) axes. Perturbations were applied across all rink regions and evaluated under four different temporal window lengths: 5, 10, 20, and 30 frames. The results of this sensitivity analysis, illustrated in Figure 5.9 (Y-direction) and Figure 5.10 (X-direction), demonstrate that shorter temporal windows lead to greater velocity estimation errors under the same level of perturbation. As the frame window length increased, the magnitude of velocity error consistently decreased across all regions. For instance, in Zone A, a 5-pixel perturbation in the Y-direction using a 5-frame window resulted in an error of approximately 1 m/s, whereas increasing the window to 30 frames reduced the error to below 0.2 m/s. Moreover, for a fixed number of frames, perturbations in the Y-direction had a more pronounced impact on velocity estimation compared to those in the X-direction. For example, in Zone A, a 5-pixel perturbation over 5 frames led to a velocity error of less than 0.2 m/s in the X-direction, while the corresponding error in the Y-direction increased significantly to approximately 1 m/s. In addition, the results for both directions indicate that Zone C exhibited the lowest sensitivity across all regions. For example, under a 15-pixel perturbation with a 5-frame window, the velocity errors in both the X- and Y-directions remained below 0.4 m/s.

5.4. Discussion

This study evaluated the performance of a side-view, homography-based pipeline for estimating player velocity in youth ice hockey, using drone-derived measurements as ground truth. The results indicate that the proposed approach offers both promising accuracy and practical feasibility, particularly in environments where top-down or sensor-based tracking systems are not available.

As presented in Table 5.1, both panning and fixed camera configurations produced velocity estimates that are consistent with ground truth values, with mean percentage errors below 10%. Importantly, the achieved level of accuracy supports their application for head impact assessment and event reconstruction tasks commonly conducted in sports injury laboratories. In head impact biomechanics, a well-established approach for estimating trauma load involves categorizing player velocities into discrete ranges, such as very low (<2 m/s), low (2-4 m/s), medium (4-6 m/s), high (6-8 m/s), and very high (8+ m/s) [123], [124], [125], [126]. These velocity categories guide the reconstruction of head impacts, whose kinematics are then input into finite element (FE) brain models to compute Maximum Principal Strain (MPS), a biomechanical metric that quantifies brain tissue deformation and is commonly used to predict traumatic brain injury. The use of velocity ranges is supported by evidence that small variations in inbound velocity result in minimal changes in MPS for a given impact surface or type of event [124], [127]. For instance, youth football studies employing anatomically accurate FE models report graded MPS outcomes aligned with predefined acceleration magnitudes [22]. Similarly, research on soccer heading indicates that although kinematics varies, the 95th percentile MPS remains within specific ranges across impact scenarios [128]. This consistency suggests that range-based velocity classification is robust, even in the presence of minor velocity estimation errors. Minor velocity deviations typically do not shift MPS beyond its injury severity tier, making the velocity estimation errors observed in this study acceptable for predicting injury risk. In the current study, the participant's velocities predominantly fell within the medium range (4–6 m/s) as defined by the aforementioned classification. Although the analysis shown in Figures 5.8 and 5.9 revealed a consistent tendency for the system to slightly underestimate velocity compared to the drone-derived ground truth in most cases, the Bland–Altman analysis demonstrated that both fixed and panning camera configurations achieved limits of agreement within the acceptable range defined by the velocity bins used for reconstruction. This suggests that the proposed velocity estimation pipeline, when applied to a common broadcast-style youth hockey camera (i.e., a panning camera), shows good agreement with a top-down view camera for the purposes of this application. The present analysis was based on 16 paired observations, which is at the lower end of the recommended sample size for stable limits of agreement estimates. Consequently, the limits reported here should be interpreted with caution, and these findings should be considered preliminary validation.

The zoom-in camera configuration, which exhibited the lowest performance with a mean absolute error (MAE) of 0.78 m/s and a mean percentage error of approximately 20%, still produced estimation errors that remained within the acceptable bounds of the predefined velocity categories. However, this reduced accuracy was expected, as zooming limits the number of visible landmarks required for homography calculations, resulting in less precise transformations and larger velocity errors. To improve accuracy under zoomed-in conditions, it is recommended to increase the number of detectable landmarks within the rink, such as including every intersection of rink lines, to ensure sufficient reference points remain visible even when the view is zoomed in.

A fixed camera is generally more reliable for rink localization because the relative distances between landmarks and the camera remain constant, allowing a single homography matrix to be applied consistently across all frames. In contrast, a panning camera continuously alters its field of view, requiring frame-by-frame homography recalibration. Additionally, panning movements or occlusions can lead to temporary landmark loss, reducing the number of available reference points and potentially compromising accuracy. Despite these challenges, the present study found that the panning camera produced reasonable error levels and, in some cases, even more accurate velocity estimations than the fixed camera.

It is also important to note that the two camera configurations were not recorded under identical conditions, which likely influenced the observed discrepancies in their performance. Specifically, as shown in Figure 5.12, the panning and fixed camera views for Region A differed in zoom level, focus, and lighting. The fixed camera used a zoomed-in lens, which reduced the number of usable landmarks for homography calculations, and the darker lighting in its view likely affected the accuracy of player detection bounding boxes. These variations impacted the player coordinate extraction process and contributed to differences in the homography matrices and, consequently, in the velocity estimations.

These variations influenced the player coordinate extraction process and resulted in differences in the homography matrices and, consequently, in the velocity estimations. Since the player's position is extracted as the midpoint of the bounding box, even small variations in bounding box dimensions, often caused by changes in body posture, camera angle, or lighting conditions, can affect velocity estimation. As shown in Figures 5.9 and 5.10, these errors are more pronounced when using shorter temporal windows, such as 5 frames in a 30 fps video. This variability was particularly evident in Zones A, B, and D, where perspective effects amplified the impact of localization errors. Interestingly, while Zone C exhibited the lowest sensitivity to artificial perturbations (e.g., under a 15-pixel shift, velocity errors remained below 0.4 m/s), it had the highest overall velocity estimation error according to Table 5.1. This apparent contradiction is likely due to Zone C's proximity to the camera, which results in larger bounding boxes. In this region, even minor postural changes can shift bounding box coordinates by over 100 pixels, significantly affecting the extracted position and amplifying estimation errors. One way to reduce these errors is by increasing the number of frames used to compute average velocity, which, as demonstrated in the sensitivity analysis, consistently lowers the estimation error. Alternatively, maintaining the same segment length while increasing the recording frame rate to 60 FPS could improve velocity precision by providing finer temporal resolution and reducing the influence of transient noise. An additional contributing factor is the limited resolution of the top-down rink view (830 × 1920 pixels). With fewer pixels along the vertical (Y) axis, small shifts in this direction correspond to larger real-world displacements compared to the horizontal (X) axis. Improving the resolution and quality of the top-down view would likely enhance the robustness and accuracy of velocity estimates.

The analysis further revealed that lateral movements often resulted in the largest velocity estimation errors, particularly in Region C. These movements typically involve outstretched limbs (see Figure 5.13), which expand the bounding box and shift the calculated player coordinates, increasing localization variability. To mitigate these issues, velocity can be computed on a frame-by-frame basis and then smoothed using temporal filtering techniques. Averaging velocities over fixed-length segments (e.g., 5–10 frames) helps reduce the influence of outliers and abrupt fluctuations. A simple moving average filter can smooth short-term noise, while more advanced filters such as the Savitzky–Golay filter preserve motion trends. Alternatively, applying a low-pass Butterworth filter in the frequency domain can attenuate high-frequency noise caused by bounding box jitter. Implementing a continuous frame-by-frame velocity calculation followed by temporal filtering would help address common sources of error, including outlier spikes, systematic underestimation, and localization inaccuracies from bounding box shifts, as well as cases where the estimated velocity lies near the edge of a velocity bin. Such an approach would yield smoother and more consistent velocity estimates, particularly during complex movements and in regions prone to detection instability.

Figure 5-12. Comparison of panning (left) and fixed (right) camera views in Region A.

Figure 5-13. Side view of a player performing a lateral movement in Region C.

This study provided preliminary validation of the proposed pipeline and demonstrated that both panning and fixed camera setups could produce reliable and practical velocity estimates suitable for sports injury research. Importantly, the reported velocity estimates are derived from finite differences of player position over a fixed temporal window (10 frames), such that velocity accuracy is fundamentally governed by the accuracy of the underlying position estimates. In the validation presented in this Chapter, both the proposed method and the reference approach compute velocity using the same 10-frame differencing and the same time interval ($\Delta t = 10/\text{FPS}$); however, they rely on different sources of position information. Ground-truth player positions were obtained from a calibrated top-down camera view via manual digitization, whereas pipeline positions were generated by the full side-view workflow, including player detection and tracking, homography estimation, and projection into the top-down coordinate system. As a result, discrepancies between the estimated and reference velocities primarily reflect localization errors in the automated side-view pipeline propagated through the velocity computation. This framing is appropriate for the objectives of this thesis, as pre-impact velocity and derived closing velocity are the quantities required for downstream head-impact characterization, and the top-down camera provides an independent reference for evaluating dynamic trajectories under realistic motion.

The strength of this method lay in its accessibility and scalability. Unlike traditional motion capture systems, the proposed approach required only a standard 2D video, which could be captured using commonly available cameras, including handheld devices. This made it applicable in virtually any ice rink and across all age groups, including contexts where specialized tracking equipment was unavailable or impractical. By offering an objective, low-cost, and adaptable tool for player velocity estimation, this method had the potential to expand large-scale data collection efforts in head impact and brain trauma research. While the findings were promising, they were based on a limited sample size and should therefore be interpreted with caution. Future studies with larger datasets and broader testing conditions are warranted to strengthen the evidence for interchangeability and generalizability. Nevertheless, by demonstrating feasibility and robustness under diverse recording conditions, this work laid the groundwork for a scalable and reliable solution for estimating player velocity from standard side-view video recordings, enabling broader investigations into brain injury risk in resource-constrained sports environments.

5.4.1. Limitations and Future work

While the proposed pipeline showed high accuracy in rink localization and player tracking, several limitations remain. The system relies on visible rink landmarks to compute homography; when footage is overly zoomed in and landmarks are missing, velocity cannot be estimated. Similarly, if player detection fails, due to occlusion or crowding near the boards, velocity data is lost. The system also estimates the skate contact point using the bottom edge of the bounding box, which becomes inaccurate when only the upper body is visible. Since all components, rink localization, player detection, and skate visibility, must function correctly, any failure affects output reliability. Future work will focus on enhancing detection under partial visibility, using predictive models and adjacent clean frames to fill data gaps, and improving performance across varied lighting and camera angles.

In addition, the average skating speed in was approximately 5 m/s, which reflects realistic game-like motion but does not span the full range of velocities observed in ice hockey. Importantly, the velocity-estimation method itself is not inherently tied to the absolute skating speed: velocity is computed from frame-to-frame displacement divided by the inter-frame time interval (Δt), which is determined by the video frame rate (FPS). Consequently, estimation performance is primarily governed by temporal and spatial sampling (FPS and image resolution), which determine how well motion is captured and how precisely the player can be localized in the image. The method is therefore expected to generalize across realistic skating speeds, provided the frame rate is sufficient to avoid excessive inter-frame displacement and the resolution supports reliable localization. Future work will validate the pipeline across a broader range of skating speeds and video acquisition settings.

Chapter 6. General Discussion

Traumatic brain injury (TBI) is a significant public health concern, particularly in contact sports such as ice hockey [6]. The nature of the sport, characterized by high-speed skating, intense physical engagement, and frequent collisions, exposes athletes to an elevated risk of head impacts, which can result in both acute and chronic brain injuries. These injuries span a spectrum from concussions and diffuse axonal injury to long-term neurodegenerative conditions[6]. Although athletes at all levels are at risk, concern is greatest for youth athletes, who constitute the majority of registered sport participants and are particularly vulnerable due to ongoing neurodevelopment [23]. During adolescence, key neurological processes such as synaptic pruning, myelination, and cortical maturation are still in progress, rendering the brain more susceptible to biomechanical insult[6], [9]. This vulnerability is further exacerbated by the limited availability of specialized tools for detecting, monitoring, and managing head trauma in youth sports settings.

Existing literature suggests that a substantial proportion of concussions among youth athletes go unrecognized, unreported, or misdiagnosed[5], [23], [26], [123]. This underreporting is often due to limited symptom awareness, sociocultural pressures that discourage self-reporting, and the absence of standardized surveillance protocols. As a result, many head impacts, especially those that do not produce immediate or overt symptoms, go unnoticed, yet may accumulate over time and contribute to long-term neurological deficits [12]. To mitigate the risks associated with brain trauma effectively, there is an urgent need for accurate and scalable tools capable of objectively measuring and documenting head impact exposure. Brain trauma is not solely the result of isolated high-magnitude impacts; rather, it arises from a complex interplay of factors, including impact frequency, duration, recovery intervals, and the biomechanical properties of each event [27], [123]. The severity of an individual impact is primarily determined by its magnitude, which is influenced by several biomechanical parameters: impact velocity, surface compliance, anatomical location on the head, and the direction of applied force [123]. While high-magnitude impacts have long been recognized as a primary cause of traumatic brain injury (TBI), emerging evidence underscores the cumulative effects of repeated low-magnitude head impacts. Although often asymptomatic and difficult to detect in real time, these sub-concussive impacts have been shown to contribute to neurodegeneration and are increasingly linked to the development of chronic traumatic encephalopathy (CTE), a progressive neurodegenerative disorder.

Despite increased awareness of these biomechanical and physiological relationships, the field still lacks comprehensive, player-specific datasets that would allow for robust modeling of long-term injury outcomes. Conventional methods for collecting impact-related data, such as wearable accelerometers, helmet-mounted sensors, and motion capture systems, are typically expensive, labor-intensive, and unsuitable for routine use in youth sports environments. While manual video analysis remains a feasible alternative, it is time-consuming and lacks the scalability needed for widespread implementation.

To address these limitations, the present study developed an automated, scalable, and cost-effective system for capturing and documenting head impact characteristics on an individual player basis. A key focus of this system was the measurement of impact velocity; a critical determinant of the kinetic energy transmitted to the head during a collision and a principal factor in injury severity. High-resolution, accurate velocity data, when contextualized with additional impact variables, can provide deeper insights into injury mechanisms and support the development of predictive models for brain trauma.

To this end, the study introduced a novel deep learning-based pipeline for estimating player velocity from 2D side-view video footage, a widely accessible and non-invasive data source. The system was designed with youth hockey in mind, where access to advanced tracking infrastructure, such as multi-camera arrays, inertial measurement units (IMUs), or optical motion capture systems, is typically limited. By leveraging readily available video footage (e.g., broadcast or parent-recorded games) and automating the analysis workflow, the system enables the generation of large, annotated datasets at both individual and population levels. These datasets can facilitate detailed investigations into impact exposure, player-specific risk factors, and the effectiveness of injury prevention interventions across various levels of competition. The proposed pipeline estimates 2D planar velocity using a single-camera setup and consists of three interrelated components. In the first study (Rink Localization), a YOLOv5-based object detection model was trained to identify specific rink landmarks, including the intersections of the center line, blue line, and goal line with the boards, as well as all faceoff spots. The training dataset comprised more than 9,900 annotated images from youth hockey games at the novice, peewee, and atom levels. The model demonstrated strong performance, achieving an F1 score and a precision–recall of 0.99 at an 80% confidence threshold. The model also achieved mean average precision (mAP) scores of 98.5% at an Intersection over Union (IOU) threshold of 0.50, and 64.5% averaged across IOU thresholds from 0.50 to 0.95. Once at least four landmarks were detected within a frame, a custom script matched them to a standardized top-down rink template and computed the homography matrix to transform each side-view frame into a consistent top-down coordinate system. Transformation accuracy, evaluated using IOU, yielded an average IOU of 0.96. Accuracy was highest near the detected landmarks and decreased farther away due to extrapolation. The primary limitation of this approach was its dependency on the visibility of at least four landmarks per frame; in zoomed-in footage, where some markers were often obscured, the transformation could not be applied. However, because youth footage is typically captured as wide, zoomed-out shots, most frames contained a sufficient number of visible markers to enable reliable transformation. This study demonstrated that a uniform coordinate system can be established for each frame, enabling accurate mapping of player trajectories from side view to top-down view.

In the second study (Player Detection, Tracking, and Velocity Estimation), a YOLOv8 object detection model, initially trained on a dataset of 80,000 NHL images, was fine-tuned using 4,700 annotated images from U11, U13, and U18 youth hockey games. This fine-tuning addressed youth-specific challenges such as smaller player size, suboptimal lighting, and varying camera angles. The fine-tuned model achieved a precision of 0.96, a recall of 0.97, and an mAP at IOU 0.50 of 0.97. For tracking, the StrongSORT algorithm was enhanced by integrating an IOU-based cost function into its existing appearance-based matching matrix. This modification improved the tracker's ability to maintain consistent player identity, particularly in occlusion-prone environments. When tested on 100 video clips (three seconds each at 30 frames per second), the modified tracker reduced identity switches by 69.2% and improved Multiple Object Tracking Accuracy (MOTA) from 89.0% to 96.5%.

With reliable detection and tracking in place, player positions were extracted as the midpoints of the bottom edges of bounding boxes in the side-view footage. These coordinates were transformed into the top-down coordinate space using the homography matrices obtained in the first study. For each tracked player ID, displacement over five consecutive frames (0.2 seconds at 25 frames per second) was computed and converted into real-world distances based on standardized North American rink dimensions. The average velocity over each interval was then calculated. While the system is adaptable to other frame intervals and recording rates, its accuracy is reduced in cases where only the upper body is visible, typically near the boards, because the estimated position does not correspond to the actual skate contact point. Additionally, if a player was not detected or tracked for more than 15 consecutive frames, velocity could not be estimated. The system was validated to function effectively over 10 consecutive frames, which aligns with the short time window typical of head-impact events.

In the third study (Validation of Velocity Estimation), the estimated velocities from the pipeline were validated against drone-recorded top-down ground-truth data. A player performed lateral, diagonal, and head-on skating movements across four zones of the rink, each representing a different distance from the side-view camera. Validation was conducted using 30-frames-per-second, 1920 × 1080 resolution footage, with fixed, panning, and zoomed-in camera configurations. A 10-frame window (equivalent to 0.33 seconds) was selected to match the typical temporal duration of head impacts. Both fixed and panning camera setups produced mean percentage errors below 10%, while zoomed-in views, limited by reduced landmark visibility, yielded a mean absolute error (MAE) of 0.78 m/s, or approximately 20%. Despite these limitations, the velocity estimates remained within predefined biomechanical velocity bands commonly used in sports-injury laboratories for head-impact reconstruction. These categories are very low (less than 2 m/s), low (between 2 and 4 m/s), medium (between 4 and 6 m/s), high (between 6 and 8 m/s), and very high (more than 8 m/s). Prior research indicates that small variations in inbound velocity do not substantially alter Maximum Principal Strain (MPS) values, supporting the acceptability of the observed estimation errors for this purpose.

The validation results showed that the pipeline produced the largest errors in Zone C, the area closest to the camera. This was attributed to the high variability in bounding-box dimensions caused by perspective distortion and player postural changes. Because player coordinates were derived from bounding-box geometry, a sensitivity analysis was performed. This analysis introduced perturbations of 5, 10, and 15 pixels in both the x and y directions and assessed the impact of these shifts across temporal windows of 5, 10, 20, and 30 frames. The results revealed that shorter frame intervals (for example, five frames) resulted in significantly higher errors, especially in camera-proximal zones such as Zone C. In contrast, longer frame windows, such as 30 frames, substantially reduced the effect of localization errors; for instance, in Zone A, the velocity error dropped to below 0.4 m/s for all perturbation levels. Perturbations in the y direction introduced greater errors than those in the x direction, likely due to the lower vertical resolution in the top-down reference image. Furthermore, for an equivalent pixel shift, farther rink regions exhibited greater velocity errors due to spatial compression in the homography transformation.

To mitigate these variations, the application of temporal smoothing techniques, such as moving-average filters, Savitzky–Golay filters, and low-pass Butterworth filters, is recommended. These techniques reduce frame-to-frame noise resulting from bounding-box jitter while preserving the underlying motion trajectory, thereby improving the stability and reliability of velocity estimates.

6.1. Vision for implementation and development of a comprehensive, player-specific dataset

The pipeline developed in this thesis is intended as a foundational module within a broader automated framework for detecting and characterizing head impacts in youth ice hockey. The long-term objective is to enable the creation of a standardized, large-scale dataset that links video-derived impact characteristics to brain trauma risk and, when available, to clinical outcomes. This need is particularly acute in youth hockey, where objective monitoring tools are limited despite the high participation rate and the heightened vulnerability of developing athletes.

Brain trauma risk is influenced by both exposure variables (e.g., number of impacts, impacts per game/season, inter-impact intervals, and cumulative exposure duration) and severity-related variables (e.g., closing velocity, contact compliance, impact location, and loading direction). Within this framework, the present thesis focused on establishing an automated and validated method for estimating planar player velocity from single-camera side-view video. This capability provides a key severity descriptor, pre-impact velocity, and enables the derivation of closing velocity, which can be used alongside event context to better characterize impact severity. A major motivation for scalable exposure quantification is the growing evidence that long-term neurological outcomes are associated not only with clinically diagnosed concussions, but also with the cumulative burden of repetitive head impacts [18], [24], [129], including impacts that do not produce immediate symptoms. Dose–response relationships have been reported in contact sports [124], where greater cumulative exposure (e.g., years of play and total impact burden) is associated with increased risk and severity of chronic neurodegenerative conditions such as chronic traumatic encephalopathy (CTE). In American football [130], helmet-sensor datasets have demonstrated that combinations of impact frequency and head acceleration magnitude accumulated over an athletic career are related to CTE risk and pathological severity. Comparable longitudinal datasets are limited in youth hockey due to the cost and logistical barriers of deploying instrumented equipment. In this context, automated video-based analysis offers a complementary pathway for quantifying exposure at scale. Although camera-based methods do not directly measure head linear and rotational accelerations, they can provide pre-impact and closing velocities and, when integrated with event characteristics (e.g., contact type as a proxy for compliance, impact location, and direction), can support scalable estimation of relative impact severity. When applied across full seasons and linked to player identity, these measures can be aggregated into player-specific exposure profiles, such as total impacts, number of higher-velocity events, and distributions of collision types (e.g., head-to-boards versus head-to-shoulder). A key nuance is that external impact severity does not map perfectly onto clinical presentation. Symptom presence and severity vary widely even for apparently similar impacts, reflecting the importance of internal tissue deformation patterns and individual susceptibility. This further supports the need to capture multiple event descriptors rather than relying on a single metric. Accordingly, the broader project vision is to integrate velocity estimation with automated impact detection and event characterization, and to link these descriptors to brain deformation estimates such as maximum principal strain (MPS). This integration is also relevant for monitoring sub-concussive impacts that may be asymptomatic yet contribute to cumulative neurological burden over time.

In the intended end-to-end pipeline, impacts are not identified from velocity alone. Impact detection occurs upstream (through automated head-impact detection models that timestamp events and identify involved players). Following event identification, the velocity-estimation module is applied to the short pre-impact window to estimate each player’s absolute planar velocity and to compute closing velocity. From a mechanics standpoint, velocity is an upstream contributor to severity. For a given closing velocity, the resulting head accelerations and brain deformation depend strongly on impact duration and contact compliance (e.g., head-to-boards versus head-to-shoulder), as well as impact location and direction. Therefore, velocity estimates are most informative when interpreted in conjunction with these contextual variables.

Wearable sensors (helmet- or mouthguard-mounted) remain the most direct tools for measuring head accelerations at high sampling rates, but their use is often limited in youth hockey by cost and implementation barriers. Video-based approaches are therefore best viewed as complementary. They enable retrospective and prospective analysis of widely available footage and support large-scale dataset generation where direct instrumentation is not feasible. With continued development, such datasets can support practical injury-prevention applications, including player-level exposure monitoring across seasons, evaluation of high-risk collision scenarios that may motivate targeted coaching or rule enforcement, and empirical inputs for equipment design and testing protocols aligned with real-world youth hockey impacts.

6.2. Limitations and Future Work

While the proposed system represents a significant advancement in automated player velocity estimation, several limitations remain. The pipeline relies on visible rink landmarks to compute homography; when these landmarks are not visible, such as in overly zoomed-in footage, homography calculation is not feasible, and velocity estimation cannot be performed. Likewise, when players are not fully visible, the estimated player coordinates contain errors because the skate contact point, derived from the bottom edge of the bounding box, may be inaccurately positioned. Detection failures due to occlusion, crowding near the boards, or poor lighting further contribute to data loss. Since all components, rink localization, player detection, and accurate coordinate estimation, must function correctly, any single-point failure reduces output reliability.

Although validated for ice hockey in the third study, the underlying principles could be adapted for other sports with similar tracking challenges, such as soccer, basketball, or rugby. Future work will focus on enhancing detection under partial visibility, leveraging predictive models and adjacent clean frames to fill data gaps, improving robustness under varied lighting and camera conditions, and exploring cross-sport generalization to broaden applicability.

6.3. Conclusion

This research introduced and validated a deep learning-based pipeline for accurately estimating player velocity using 2D side-view video footage, specifically designed for youth ice hockey. The proposed system successfully addressed critical limitations of existing methods, particularly their reliance on expensive multi-camera setups or wearable sensors, by offering a scalable, cost-effective alternative suitable for deployment even in lower age categories. By integrating automated rink localization, robust player detection and tracking, and precise velocity estimation, the system demonstrated the capability to extract high-resolution biomechanical data from standard game recordings.

The pipeline's ability to provide frame-level velocity estimates with high accuracy across a range of video conditions, including variable lighting, player occlusions, and different camera configurations, represents a notable advancement in sports injury analytics. Through rigorous validation against drone-based top-down ground truth data and extensive sensitivity analyses, the study showed that the velocity estimates produced by the system fall within accepted error thresholds used in biomechanical modeling and concussion risk classification. These findings position the method as a practical tool for reconstructing head impact events and estimating injury severity via Maximum Principal Strain (MPS).

Furthermore, the system's scalability and automation enable large-scale dataset generation, which is essential for advancing data-driven research on brain trauma. By facilitating the systematic recording and annotation of impact characteristics, such as velocity, direction, and frequency, the pipeline provides a foundation for developing predictive models that can reveal patterns associated with traumatic brain injuries (TBI) in youth athletes. This capability is especially important for addressing a critical gap in youth sport safety: the lack of accessible, objective tools for monitoring head impacts and supporting concussion prevention strategies.

In the broader context of public health and sports science, the primary contribution of this work lies in enabling a new paradigm of injury monitoring that is both scientifically rigorous and practically feasible. Coaches, clinicians, and researchers may employ this system to assess individual and team-level exposure to potentially injurious impacts, adjust training protocols, evaluate rule modifications, and implement interventions aimed at reducing long-term neurological harm. Its applicability in resource-limited environments, such as community rinks and youth tournaments, helps democratize access to advanced biomechanical analysis, ensuring that player safety is not constrained by technological or financial barriers.

Ultimately, this study offers a proof-of-concept for a robust, automated, and accessible system for monitoring brain trauma risk using 2D video analysis. The proposed framework contributes to the scientific understanding of head impacts in youth hockey and offers a transferable model that can be adapted to other sports contexts. With continued development, the system holds strong potential to serve as a cornerstone in the broader ecosystem of injury prevention and athlete health surveillance, contributing to the long-term safety and well-being of young athletes globally.

Chapter 7. References

[1] P. Cheng, R. Li, D. C. Schwebel, M. Zhu, and G. Hu, “Traumatic brain injury mortality among U.S. children and adolescents ages 0–19 years, 1999–2017,” *J Safety Res*, vol. 72, pp. 93–100, Feb. 2020, doi: 10.1016/j.jsr.2019.12.013.

[2] C. A. Taylor, J. M. Bell, M. J. Breiding, and L. Xu, “Traumatic Brain Injury–Related Emergency Department Visits, Hospitalizations, and Deaths — United States, 2007 and 2013,” *MMWR. Surveillance Summaries*, vol. 66, no. 9, pp. 1–16, Mar. 2017, doi: 10.15585/mmwr.ss6609a1.

[3] A. A. Hyder, C. A. Wunderlich, P. Puvanachandra, G. Gururaj, and O. C. Kobusingye, “The impact of traumatic brain injuries: a global perspective.,” *NeuroRehabilitation*, vol. 22, no. 5, pp. 341–53, 2007.

[4] A. Gardner *et al.*, “A systematic review of concussion in rugby league,” *Br J Sports Med*, vol. 49, no. 8, pp. 495–498, Apr. 2015, doi: 10.1136/bjsports-2013-093102.

[5] T. M. Enniss *et al.*, “Primary prevention of contact sports-related concussions in amateur athletes: a systematic review from the Eastern Association for the Surgery of Trauma,” *Trauma Surg Acute Care Open*, vol. 3, no. 1, p. e000153, Jun. 2018, doi: 10.1136/tsaco-2017-000153.

[6] T. W. McAllister, “Neurobiological consequences of traumatic brain injury,” *Dialogues Clin Neurosci*, vol. 13, no. 3, pp. 287–300, Sep. 2011, doi: 10.31887/DCNS.2011.13.2/tmcallister.

[7] A. Gardner *et al.*, “A systematic review of concussion in rugby league,” *Br J Sports Med*, vol. 49, no. 8, pp. 495–498, Apr. 2015, doi: 10.1136/bjsports-2013-093102.

[8] A. E. Davis, “Mechanisms of Traumatic Brain Injury: Biomechanical, Structural and Cellular Considerations,” *Crit Care Nurs Q*, vol. 23, no. 3, pp. 1–13, Nov. 2000, doi: 10.1097/00002727-200011000-00002.

[9] A. C. McKee *et al.*, “The spectrum of disease in chronic traumatic encephalopathy,” *Brain*, vol. 136, no. 1, pp. 43–64, Jan. 2013, doi: 10.1093/brain/aws307.

[10] J. N. Ianof *et al.*, “Sport-related concussions,” *Dement Neuropsychol*, vol. 8, no. 1, pp. 14–19, Mar. 2014, doi: 10.1590/S1980-57642014DN81000003.

[11] E. Toledo *et al.*, “The young brain and concussion: Imaging as a biomarker for diagnosis and prognosis,” *Neurosci Biobehav Rev*, vol. 36, no. 6, pp. 1510–1531, Jul. 2012, doi: 10.1016/j.neubiorev.2012.03.007.

[12] A. P. Kontos *et al.*, “Incidence of Concussion in Youth Ice Hockey Players,” *Pediatrics*, vol. 137, no. 2, Feb. 2016, doi: 10.1542/peds.2015-1633.

[13] S. C. Pender, A. M. Smith, J. T. Finnoff, J. Huston, and M. J. Stuart, “Concussions in Ice Hockey — Moving Toward Objective Diagnoses and Point-of-care Treatment: A Review,” *Curr Sports Med Rep*, vol. 19, no. 9, pp. 380–386, Sep. 2020, doi: 10.1249/JSR.0000000000000752.

[14] J. Lloyd and F. Conidi, “Brain injury in sports,” *J Neurosurg*, vol. 124, no. 3, pp. 667–674, Mar. 2016, doi: 10.3171/2014.11.JNS141742.

[15] R. A. Stern *et al.*, “Tau Positron-Emission Tomography in Former National Football League Players,” *New England Journal of Medicine*, vol. 380, no. 18, pp. 1716–1725, May 2019, doi: 10.1056/NEJMoa1900757.

[16] B. W. Benson, A. S. McIntosh, D. Maddocks, S. A. Herring, M. Raftery, and J. Dvořák, "What are the most effective risk-reduction strategies in sport concussion?," *Br J Sports Med*, vol. 47, no. 5, pp. 321–326, Apr. 2013, doi: 10.1136/bjsports-2013-092216.
[17] J. Mez *et al.*, "Clinicopathological Evaluation of Chronic Traumatic Encephalopathy in Players of American Football," *JAMA*, vol. 318, no. 4, p. 360, Jul. 2017, doi: 10.1001/jama.2017.8334.
[18] R. A. Stern, D. O. Riley, D. H. Daneshvar, C. J. Nowinski, R. C. Cantu, and A. C. McKee, "Long-term Consequences of Repetitive Brain Trauma: Chronic Traumatic Encephalopathy," *PM&R*, vol. 3, no. 10S2, Oct. 2011, doi: 10.1016/j.pmrj.2011.08.008.
[19] C. LoBue *et al.*, "Traumatic Brain Injury and Risk of Long-Term Brain Changes, Accumulation of Pathological Markers, and Developing Dementia: A Review," *Journal of Alzheimer's Disease*, vol. 70, no. 3, pp. 629–654, Aug. 2019, doi: 10.3233/JAD-190028.
[20] C. S. Sahler and B. D. Greenwald, "Traumatic Brain Injury in Sports: A Review," *Rehabil Res Pract*, vol. 2012, pp. 1–10, 2012, doi: 10.1155/2012/659652.
[21] B. Sterling, "Future of Youth Hockey Culture.," 2024.
[22] L. E. Miller *et al.*, "Evaluation of Brain Response during Head Impact in Youth Athletes Using an Anatomically Accurate Finite Element Model," *J Neurotrauma*, vol. 36, no. 10, pp. 1561–1570, May 2019, doi: 10.1089/neu.2018.6037.
[23] M. L. Prins and C. C. Giza, "Repeat traumatic brain injury in the developing brain," *International Journal of Developmental Neuroscience*, vol. 30, no. 3, pp. 185–190, May 2012, doi: 10.1016/j.ijdevneu.2011.05.009.
[24] A. C. McKee, M. L. Alosco, and B. R. Huber, "Repetitive Head Impacts and Chronic Traumatic Encephalopathy," *Neurosurg Clin N Am*, vol. 27, no. 4, pp. 529–535, Oct. 2016, doi: 10.1016/j.nec.2016.05.009.
[25] R. A. Stern, D. O. Riley, D. H. Daneshvar, C. J. Nowinski, R. C. Cantu, and A. C. McKee, "Long-term Consequences of Repetitive Brain Trauma: Chronic Traumatic Encephalopathy," *PM&R*, vol. 3, no. 10S2, Oct. 2011, doi: 10.1016/j.pmrj.2011.08.008.
[26] C. (2019) Karton, "Profiling Brain Trauma in Professional American-style Football and the Implications to Developing Neurological Injury.," Université d'Ottawa/University of Ottawa), Ottawa.
[27] C. Karton and T. B. Hoshizaki, "Biomechanics of Sport-Related Neurological Injury," *Clin Sports Med*, vol. 40, no. 1, pp. 19–38, Jan. 2021, doi: 10.1016/j.csm.2020.09.001.
[28] J. Li, Z. Ye, and C. Zhang, "Study on the interaction between big data and artificial intelligence," *Syst Res Behav Sci*, vol. 39, no. 3, pp. 641–648, May 2022, doi: 10.1002/sres.2878.
[29] K. Marshall, A. Post, and M. D. Gilchrist, "A Comparison of dynamic impact response and brain deformation metrics within the cerebrum of head impact reconstructions representing three mechanisms of head injury in ice hockey," 2012, pp. 430–440.
[30] A. Azadi, P. Dehghan, C. Karton, M. D. Gilchrist, M. Fraser, and T. B. Hoshizaki, "Deep Learning Model for MaximumPrincipal Strain Prediction from Ice Hockey Video-Derived Impact Features: Velocity, Compliance, Location, and Elevation," 2025.
[31] G. Robinson and I. Robinson, "Radar speed gun true velocity measurements of sports-balls in flight: application to tennis," *Phys Scr*, vol. 91, no. 2, p. 023008, Feb. 2016, doi: 10.1088/0031-8949/91/2/023008.

[32] D. B. Dwyer and T. J. Gabbett, “Global Positioning System Data Analysis: Velocity Ranges and a New Definition of Sprinting for Field Sport Athletes,” *J Strength Cond Res*, vol. 26, no. 3, pp. 818–824, Mar. 2012, doi: 10.1519/JSC.0b013e3182276555.
[33] S. Noiumkar and S. Tirakoat, “Use of Optical Motion Capture in Sports Science: A Case Study of Golf Swing,” in *2013 International Conference on Informatics and Creative Multimedia*, IEEE, Sep. 2013, pp. 310–313. doi: 10.1109/ICICM.2013.58.
[34] N. Shishov, K. Elabd, V. Komisar, H. Chong, and S. N. Robinovitch, “Accuracy of Kinovea software in estimating body segment movements during falls captured on standard video: Effects of fall direction, camera perspective and video calibration technique,” *PLoS One*, vol. 16, no. 10, p. e0258923, Oct. 2021, doi: 10.1371/journal.pone.0258923.
[35] A. ’Azadi, P. ’Dehghan, R. ’Graham, and T. B. ’Hoshizaki, “A Deep Learning Network for Detecting Head Impacts in Ice Hockey from 2D Game Video,” in *IRCOBI conference*, 2024.
[36] E. ’Dubrofsky, “Homography estimation,” THE UNIVERSITY OF BRITISH COLUMBIA, vancouver, 2009.
[37] C. Liu, Y. Tao, J. Liang, K. Li, and Y. Chen, “Object Detection Based on YOLO Network,” in *2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC)*, IEEE, Dec. 2018, pp. 799–803. doi: 10.1109/ITOEC.2018.8740604.
[38] P. ’Dehghan, A. ’Azadi, A. ’Clouthier, and T. B. ’Hoshizaki, “Object Detection for Ice Surface Localisation in Youth Hockey,” in *IRCOBI*, 2024.
[39] Y. Du *et al.*, “StrongSORT: Make DeepSORT Great Again,” *IEEE Trans Multimedia*, vol. 25, pp. 8725–8737, 2023, doi: 10.1109/TMM.2023.3240881.
[40] M. Gupta, A. Vedaldi, and A. Zisserman, “ Synthetic data for text localisation in natural images. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2315-2320).,” in *Synthetic data for text localisation in natural images*, In Proceedings of the IEEE conference on computer vision and pattern recognition, 2011, pp. 2315–2320.
[41] H. Namdar, F. Sanja, and U. Raquel, “Sports Field Localization via Deep Structured Models,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5212–5220, 2017.
[42] W. Liu *et al.*, “SSD: Single Shot MultiBox Detector,” 2016, pp. 21–37. doi: 10.1007/978-3-319-46448-0_2.
[43] A. Vedaldi and A. Zisserman, “VGG Convolutional Neural Networks Practical,” oxford, 66.
[44] R. A. Sharma, B. Bhat, V. Gandhi, and C. V. Jawahar, “Automated Top View Registration of Broadcast Football Videos,” in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, IEEE, Mar. 2018, pp. 305–313. doi: 10.1109/WACV.2018.00040.
[45] S. Tarashima, “Sports Field Recognition Using Deep Multi-task Learning,” *Journal of Information Processing*, vol. 29, no. 0, pp. 328–335, 2021, doi: 10.2197/ipsjjip.29.328.
[46] N. Xiaohan, C. Shixing, and H. Raffay, “A Robust and Efficient Framework for Sports-Field Registration,” in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)},* 2021, pp. 1936–1944.

[47] W. Jiang, J. C. G. Higuera, B. Angles, W. Sun, M. Javan, and K. M. Yi, "Optimizing Through Learned Errors for Accurate Sports Field Registration," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 202AD.

[48] F. Erlik Nowruzi, R. Laganiere, and N. Japkowicz, "Homography Estimation From Image Pairs With Hierarchical Convolutional Networks," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV) Workshops*, 2017.

[49] T. Nguyen, S. W. Chen, S. S. Shivakumar, C. J. Taylor, and V. Kumar, "Unsupervised Deep Homography: A Fast and Robust Homography Estimation Model," *IEEE Robot Autom Lett*, vol. 3, no. 3, pp. 2346–2353, Jul. 2018, doi: 10.1109/LRA.2018.2809549.

[50] Hoang Le, Feng Liu, Shu Zhang, and Aseem Agarwala, "Deep Homography Estimation for Dynamic Scenes," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

[51] S. S. Beauchemin and J. L. Barron, "The computation of optical flow," *ACM Comput Surv*, vol. 27, no. 3, pp. 433–466, Sep. 1995, doi: 10.1145/212094.212141.

[52] Q. Li, R. Li, K. Ji, and W. Dai, "Kalman Filter and Its Application," in *2015 8th International Conference on Intelligent Networks and Intelligent Systems (ICINIS)*, IEEE, Nov. 2015, pp. 74–77. doi: 10.1109/ICINIS.2015.35.

[53] K. O. Arras, J. A. Castellanos, M. Schilt, and R. Siegwart, "Feature-based multi-hypothesis localization and tracking using geometric constraints," *Rob Auton Syst*, vol. 44, no. 1, pp. 41–53, Jul. 2003, doi: 10.1016/S0921-8890(03)00009-5.

[54] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Trans Pattern Anal Mach Intell*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.

[55] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A Review of Yolo Algorithm Developments," *Procedia Comput Sci*, vol. 199, pp. 1066–1073, 2022, doi: 10.1016/j.procs.2022.01.135.

[56] R. Pereira, G. Carvalho, L. Garrote, and U. J. Nunes, "Sort and Deep-SORT Based Multi-Object Tracking for Mobile Robotics: Evaluation with New Data Association Metrics," *Applied Sciences*, vol. 12, no. 3, p. 1319, Jan. 2022, doi: 10.3390/app12031319.

[57] X. Fu, Z. K., C. Wang, and F. C., " Multiple player tracking in basketball court videos. 17(6), 1811-1828.," *J Real Time Image Process*, vol. 17, no. 6, pp. 1811–1828, 202AD.

[58] B. T. Naik and Md. F. Hashmi, "YOLOv3-SORT: detection and tracking player/ball in soccer sport," *J Electron Imaging*, vol. 32, no. 01, Mar. 2022, doi: 10.1117/1.JEI.32.1.011003.

[59] A. Ruhe, A. Gänsslen, and W. Klein, "The incidence of concussion in professional and collegiate ice hockey: are we making progress? A systematic review of the literature," *Br J Sports Med*, vol. 48, no. 2, pp. 102–106, Jan. 2014, doi: 10.1136/bjsports-2012-091609.

[60] I. J. S. Williamson, "Converging evidence for the under-reporting of concussions in youth ice hockey * Commentary," *Br J Sports Med*, vol. 40, no. 2, pp. 128–132, Feb. 2006, doi: 10.1136/bjsm.2005.021832.

[61] A. P. Kontos *et al.*, "Incidence of Concussion in Youth Ice Hockey Players," *Pediatrics*, vol. 137, no. 2, Feb. 2016, doi: 10.1542/peds.2015-1633.

[62] V. A. Cubon, M. Putukian, C. Boyer, and A. Dettwiler, "A Diffusion Tensor Imaging Study on the White Matter Skeleton in Individuals with Sports-Related Concussion," *J Neurotrauma*, vol. 28, no. 2, pp. 189–201, Feb. 2011, doi: 10.1089/neu.2010.1430.

[63] J. J. Bazarian *et al.*, "Persistent, Long-term Cerebral White Matter Changes after Sports-Related Repetitive Head Impacts," *PLoS One*, vol. 9, no. 4, p. e94734, Apr. 2014, doi: 10.1371/journal.pone.0094734.
[64] T. A. Gennarelli, L. E. Thibault, J. H. Adams, D. I. Graham, C. J. Thompson, and R. P. Marcincin, "Diffuse axonal injury and traumatic coma in the primate," *Ann Neurol*, vol. 12, no. 6, pp. 564–574, Dec. 1982, doi: 10.1002/ana.410120611.
[65] S. Kleiven, "Influence of Impact Direction on the Human Head in Prediction of Subdural Hematoma," *J Neurotrauma*, vol. 20, no. 4, pp. 365–379, Apr. 2003, doi: 10.1089/089771503765172327.
[66] A. Post, T. B. Hoshizaki, M. D. Gilchrist, S. Brien, M. D. Cusimano, and S. Marshall, "The influence of dynamic response and brain deformation metrics on the occurrence of subdural hematoma in different regions of the brain," *J Neurosurg*, vol. 120, no. 2, pp. 453–461, Feb. 2014, doi: 10.3171/2013.10.JNS131101.
[67] T. Ignacy, A. Post, A. J. Gardner, M. D. Gilchrist, and T. B. Hoshizaki, "Comparison of dynamic response and maximum principal strain of diagnosed concussion in professional men's rugby league," *Proc Inst Mech Eng P J Sport Eng Technol*, vol. 236, no. 4, pp. 266–276, Dec. 2022, doi: 10.1177/17543371211016596.
[68] R. Duffield, M. Reid, J. Baker, and W. Spratford, "Accuracy and reliability of GPS devices for measurement of movement patterns in confined spaces for court-based sports," *J Sci Med Sport*, vol. 13, no. 5, pp. 523–525, Sep. 2010, doi: 10.1016/j.jsams.2009.07.003.
[69] R. J. Aughey, "Applications of GPS Technologies to Field Sports," *Int J Sports Physiol Perform*, vol. 6, no. 3, pp. 295–310, Sep. 2011, doi: 10.1123/ijspp.6.3.295.
[70] C. Castagna, M. Varley, S. C. A. Póvoas, and S. D'Ottavio, "Evaluation of the Match External Load in Soccer: Methods Comparison," *Int J Sports Physiol Perform*, vol. 12, no. 4, pp. 490–495, Apr. 2017, doi: 10.1123/ijspp.2016-0160.
[71] N. D'Apuzzo, "Surface measurement and tracking of human body parts from multi-image video sequences," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 56, no. 5–6, pp. 360–375, Aug. 2002, doi: 10.1016/S0924-2716(02)00069-2.
[72] J. P. Alexander, T. L. Hopkinson, D. W. T. Wundersitz, B. G. Serpell, J. K. Mara, and N. B. Ball, "Validity of a Wearable Accelerometer Device to Measure Average Acceleration Values During High-Speed Running," *J Strength Cond Res*, vol. 30, no. 11, pp. 3007–3013, Nov. 2016, doi: 10.1519/JSC.0000000000001396.
[73] G. J. Tierney, H. Joodaki, T. Krosshaug, J. L. Forman, J. R. Crandall, and C. K. Simms, "Assessment of model-based image-matching for future reconstruction of unhelmeted sport head impact kinematics," *Sports Biomech*, vol. 17, no. 1, pp. 33–47, Jan. 2018, doi: 10.1080/14763141.2016.1271905.
[74] K. Hader, A. Mendez-Villanueva, D. Palazzi, S. Ahmaidi, and M. Buchheit, "Metabolic Power Requirement of Change of Direction Speed in Young Soccer Players: Not All Is What It Seems," *PLoS One*, vol. 11, no. 3, p. e0149839, Mar. 2016, doi: 10.1371/journal.pone.0149839.
[75] B. T. Naik, M. F. Hashmi, and N. D. Bokde, "A Comprehensive Review of Computer Vision in Sports: Open Issues, Future Trends and Research Directions," *Applied Sciences*, vol. 12, no. 9, p. 4429, Apr. 2022, doi: 10.3390/app12094429.
[76] S. Barris and C. Button, "A Review of Vision-Based Motion Analysis in Sport," *Sports Medicine*, vol. 38, no. 12, pp. 1025–1043, 2008, doi: 10.2165/00007256-200838120-00006.

[77] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2016, pp. 779–788. doi: 10.1109/CVPR.2016.91.

[78] Elan Dubrofsky., “ Homography estimation.,” *Diplomova´ pra´ce. Vancouver: Univerzita Britske´ Kolumbie*, 2009.

[79] N. Homayounfar, S. Fidler, and R. Urtasun, “Sports Field Localization via Deep Structured Models,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jul. 2017, pp. 4012–4020. doi: 10.1109/CVPR.2017.427.

[80] K. Vats, M. Fani, D. A. Clausi, and J. Zelek, “Puck localization and multi-task event recognition in broadcast hockey videos,” in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, IEEE, Jun. 2021, pp. 4562–4570. doi: 10.1109/CVPRW53098.2021.00514.

[81] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2004. doi: 10.1017/CBO9780511811685.

[82] J. Chen, F. Zhu, and J. J. Little, “A Two-point Method for PTZ Camera Calibration in Sports,” Jan. 2018.

[83] S. Alexandrova, Z. Tatlock, and M. Cakmak, “RoboFlow: A flow-based visual programming language for mobile manipulation tasks,” in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, May 2015, pp. 5537–5544. doi: 10.1109/ICRA.2015.7139973.

[84] R. Girshick, “Fast R-CNN,” Apr. 2015.

[85] L. O. Chua, *CNN: A Paradigm for Complexity*, vol. 31. WORLD SCIENTIFIC, 1998. doi: 10.1142/3801.

[86] A. Paszke *et al.*, “PyTorch: An Imperative Style, High-Performance Deep Learning Library,” Dec. 2019.

[87] Pascale B. Walters, “Towards Efficient Ice Surface Localization From Hockey Broadcast Video,” Master of Applied Science, University of Waterloo, Waterloo, 2021.

[88] J. J. Crisco *et al.*, “Frequency and Location of Head Impact Exposures in Individual Collegiate Football Players,” *J Athl Train*, vol. 45, no. 6, pp. 549–559, Nov. 2010, doi: 10.4085/1062-6050-45.6.549.

[89] J. J. Crisco *et al.*, “Head impact exposure in collegiate football players,” *J Biomech*, vol. 44, no. 15, pp. 2673–2678, Oct. 2011, doi: 10.1016/j.jbiomech.2011.08.003.

[90] P. H. Montenigro *et al.*, “Cumulative Head Impact Exposure Predicts Later-Life Depression, Apathy, Executive Dysfunction, and Cognitive Impairment in Former High School and College Football Players,” *J Neurotrauma*, vol. 34, no. 2, pp. 328–340, Jan. 2017, doi: 10.1089/neu.2016.4413.

[91] M. ’Kendall, A. ’Post, P. ’Rousseau, A. ’Oeur, M. ’D. Gilchrist, and B. ’Hoshizaki, “A comparison of dynamic impact response and brain deformation metrics within the cerebrum of head impact reconstructions representing three mechanisms of head injury in ice hockey.,” International Research Council on the Biomechanics of Injury, 2012.

[92] C. A. Emery, B. Hagel, M. Decloe, and M. Carly, “Risk factors for injury and severe injury in youth ice hockey: a systematic review of the literature,” *Injury Prevention*, vol. 16, no. 2, pp. 113–118, Apr. 2010, doi: 10.1136/ip.2009.022764.

[93] A. C. Bretzin, M. Anderson, N. Bhandari, A. J. Schmitt, and E. Beidler, “Concussion Nondisclosure in Youth Sports,” *J Athl Train*, vol. 57, no. 7, pp. 688–695, Jul. 2022, doi: 10.4085/1062-6050-534-21.

[94] R. J. Aughey, "Applications of GPS Technologies to Field Sports," *Int J Sports Physiol Perform*, vol. 6, no. 3, pp. 295–310, Sep. 2011, doi: 10.1123/ijspp.6.3.295.
[95] R. Duffield, M. Reid, J. Baker, and W. Spratford, "Accuracy and reliability of GPS devices for measurement of movement patterns in confined spaces for court-based sports," *J Sci Med Sport*, vol. 13, no. 5, pp. 523–525, Sep. 2010, doi: 10.1016/j.jsams.2009.07.003.
[96] J. P. Alexander, T. L. Hopkinson, D. W. T. Wundersitz, B. G. Serpell, J. K. Mara, and N. B. Ball, "Validity of a Wearable Accelerometer Device to Measure Average Acceleration Values During High-Speed Running," *J Strength Cond Res*, vol. 30, no. 11, pp. 3007–3013, Nov. 2016, doi: 10.1519/JSC.0000000000001396.
[97] K. Hader, A. Mendez-Villanueva, D. Palazzi, S. Ahmaidi, and M. Buchheit, "Metabolic Power Requirement of Change of Direction Speed in Young Soccer Players: Not All Is What It Seems," *PLoS One*, vol. 11, no. 3, p. e0149839, Mar. 2016, doi: 10.1371/journal.pone.0149839.
[98] J. K. Mara, K. G. Thompson, K. L. Pumpa, and S. Morgan, "Quantifying the High-Speed Running and Sprinting Profiles of Elite Female Soccer Players During Competitive Matches Using an Optical Player Tracking System," *J Strength Cond Res*, vol. 31, no. 6, pp. 1500–1508, Jun. 2017, doi: 10.1519/JSC.0000000000001629.
[99] C. Castagna, M. Varley, S. C. A. Póvoas, and S. D'Ottavio, "Evaluation of the Match External Load in Soccer: Methods Comparison," *Int J Sports Physiol Perform*, vol. 12, no. 4, pp. 490–495, Apr. 2017, doi: 10.1123/ijspp.2016-0160.
[100] N. M. Nor Adnan, M. N. A. Ab Patar, H. Lee, S.-I. Yamamoto, L. Jong-Young, and J. Mahmud, "Biomechanical analysis using Kinovea for sports application," *IOP Conf Ser Mater Sci Eng*, vol. 342, p. 012097, Apr. 2018, doi: 10.1088/1757-899X/342/1/012097.
[101] J. J. Koenderink, "Optic flow," *Vision Res*, vol. 26, no. 1, pp. 161–179, Jan. 1986, doi: 10.1016/0042-6989(86)90078-7.
[102] G. 'Welch and G. 'Bishop, "An Introduction to the Kalman Filter," vol. 2, 1995.
[103] S. Denman, V. Chandran, and S. Sridharan, "An adaptive optical flow technique for person tracking systems," *Pattern Recognit Lett*, vol. 28, no. 10, pp. 1232–1239, Jul. 2007, doi: 10.1016/j.patrec.2007.02.008.
[104] L. Wu, "Multi-view hockey tracking with trajectory smoothing and camera selection," University of British Columbia, vancouver, 2008.
[105] P. M. Djuric *et al.*, "Particle Filtering," *IEEE Signal Process Mag*, vol. 20, no. 5, pp. 19–38, Sep. 2003, doi: 10.1109/MSP.2003.1236770.
[106] R. Chauhan, K. K. Ghanshala, and R. C. Joshi, "Convolutional Neural Network (CNN) for Image Detection and Recognition," in *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)*, IEEE, Dec. 2018, pp. 278–282. doi: 10.1109/ICSCCC.2018.8703316.
[107] K. Vats, P. Walters, M. Fani, D. A. Clausi, and J. S. Zelek, "Player tracking and identification in ice hockey," *Expert Syst Appl*, vol. 213, p. 119250, Mar. 2023, doi: 10.1016/j.eswa.2022.119250.
[108] Q. Chen and C. Poullis, "Tracking and Identification of Ice Hockey Players," 2023, pp. 3–16. doi: 10.1007/978-3-031-44137-0_1.
[109] M. Şah and C. Direkoğlu, "Review and evaluation of player detection methods in field sports," *Multimed Tools Appl*, vol. 82, no. 9, pp. 13141–13165, Apr. 2023, doi: 10.1007/s11042-021-11071-z.

[110] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, “Simple online and realtime tracking,” in *2016 IEEE International Conference on Image Processing (ICIP)*, IEEE, Sep. 2016, pp. 3464–3468. doi: 10.1109/ICIP.2016.7533003.
[111] L. Chen, H. Ai, Z. Zhuang, and C. Shang, “Real-Time Multiple People Tracking with Deeply Learned Candidate Selection and Person Re-Identification,” in *2018 IEEE International Conference on Multimedia and Expo (ICME)*, IEEE, Jul. 2018, pp. 1–6. doi: 10.1109/ICME.2018.8486597.
[112] N. Wojke, A. Bewley, and D. Paulus, “Simple online and realtime tracking with a deep association metric,” in *2017 IEEE International Conference on Image Processing (ICIP)*, IEEE, Sep. 2017, pp. 3645–3649. doi: 10.1109/ICIP.2017.8296962.
[113] K. Vats, P. Walters, M. Fani, D. A. Clausi, and J. S. Zelek, “Player tracking and identification in ice hockey,” *Expert Syst Appl*, vol. 213, p. 119250, Mar. 2023, doi: 10.1016/j.eswa.2022.119250.
[114] P. ’Bergmann, T. ’Meinhardt, and L. ’Leal-Taixe, “Tracking Without Bells and Whistles,” *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 941–951, 2019.
[115] Y. Zhang, C. Wang, X. Wang, W. Zeng, and W. Liu, “FairMOT: On the Fairness of Detection and Re-identification in Multiple Object Tracking,” *Int J Comput Vis*, vol. 129, no. 11, pp. 3069–3087, Nov. 2021, doi: 10.1007/s11263-021-01513-4.
[116] B. ’Guillem and L. ’Leal-Taixé, “Learning a neural solver for multiple object tracking,” *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020.
[117] Y. ‘Zhao, Z. ’Li, and K. ’Chen, “A method for tracking hockey players by exploiting multiple detections and omni-scale appearance features,” 2020.
[118] S. Alexandrova, Z. Tatlock, and M. Cakmak, “RoboFlow: A flow-based visual programming language for mobile manipulation tasks,” in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, May 2015, pp. 5537–5544. doi: 10.1109/ICRA.2015.7139973.
[119] R. Dick, J. Agel, and S. W. Marshall, “National Collegiate Athletic Association Injury Surveillance System commentaries: introduction and methods.,” *J Athl Train*, vol. 42, no. 2, pp. 173–82, 2007.
[120] A. Post, A. Oeur, B. Hoshizaki, and M. D. Gilchrist, “Examination of the relationship between peak linear and angular accelerations to brain deformation metrics in hockey helmet impacts,” *Comput Methods Biomech Biomed Engin*, vol. 16, no. 5, pp. 511–519, May 2013, doi: 10.1080/10255842.2011.627559.
[121] W. Chen, A. Post, C. Karton, M. D. Gilchrist, M. Robidoux, and T. B. Hoshizaki, “A comparison of frequency and magnitude of head impacts between Pee Wee And Bantam youth ice hockey,” *Sports Biomech*, vol. 22, no. 6, pp. 728–751, Jun. 2023, doi: 10.1080/14763141.2020.1754450.
[122] C. Willmott and K. Matsuura, “Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance,” *Clim Res*, vol. 30, pp. 79–82, 2005, doi: 10.3354/cr030079.
[123] C. Karton *et al.*, “Exposure to brain trauma in six age divisions of minor ice hockey,” *J Biomech*, vol. 116, p. 110203, Feb. 2021, doi: 10.1016/j.jbiomech.2020.110203.

[124] C. Karton, T. Blaine Hoshizaki, and M. D. Gilchrist, “A novel repetitive head impact exposure measurement tool differentiates player position in National Football League,” *Sci Rep*, vol. 10, no. 1, p. 1200, Jan. 2020, doi: 10.1038/s41598-019-54874-9.
[125] J. Butterfield, A. Post, C. Karton, M. A. Robidoux, M. Gilchrist, and T. B. Hoshizaki, “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 Biomech*, pp. 1–17, Mar. 2023, doi: 10.1080/14763141.2023.2186941.
[126] A. Post, C. Karton, O. Thevenot, T. B. Hoshizaki, M. Robidoux, and M. D. Gilchrist, “Comparison of frequency and magnitude of head impacts experienced by Peewee boys and girls in games of youth ice hockey,” *Comput Methods Biomech Biomed Engin*, vol. 24, no. 1, pp. 1–13, Jan. 2021, doi: 10.1080/10255842.2020.1805442.
[127] A. Post *et al.*, “The biomechanics of concussion for ice hockey head impact events,” *Comput Methods Biomech Biomed Engin*, vol. 22, no. 6, pp. 631–643, Apr. 2019, doi: 10.1080/10255842.2019.1577827.
[128] M. Fahlstedt, S. Meng, and S. Kleiven, “Influence of Strain post-processing on Brain Injury Prediction,” *J Biomech*, vol. 132, p. 110940, Feb. 2022, doi: 10.1016/j.jbiomech.2021.110940.
[129] A. C. McKee *et al.*, “Chronic Traumatic Encephalopathy in Athletes: Progressive Tauopathy After Repetitive Head Injury,” *J Neuropathol Exp Neurol*, vol. 68, no. 7, pp. 709–735, Jul. 2009, doi: 10.1097/NEN.0b013e3181a9d503.
[130] D. H. Daneshvar *et al.*, “Leveraging football accelerometer data to quantify associations between repetitive head impacts and chronic traumatic encephalopathy in males,” *Nat Commun*, vol. 14, no. 1, p. 3470, Jun. 2023, doi: 10.1038/s41467-023-39183-0.

Appendix A. Related PhD Projects (Completed/Submitted)

Published:

P. ’Dehghan, A. ’Azadi, A. ’Clouthier, and T. B. ’Hoshizaki, “Object Detection for Ice Surface Localisation in Youth Hockey,” in IRCOBI, 2024.

Azadi, A., Dehghan, P., Graham, R., & Hoshizaki, T. B. A Deep Learning Network for Detecting Head Impacts in Ice Hockey from 2D Game Video.

Karton, C., Dehghan, P., Azadi, A., Fahlstedt, M., & Hoshizaki, T. B. An Evaluation of the Protective Performance of a Motocross Helmet with and without Rotational Mitigating Technology Under Impacts Representing Real-world Impact Surface Conditions.

Demiannay, J. J., Rovt, J., Brannen, M., Xu, S., Kang, G., Yip, A., ... & Petel, O. (2024). Intracranial Displacements due to Blunt Force Impact in a Postmortem Human Surrogate Brain. SAE International Journal of Transportation Safety, 12(09-12-02-0011), 113-120.

Submitted:

P. Dehghan, A. Azadi, C. Karton, A. Clouthier, and B. Hoshizaki, “Development of an Automated System to Obtain the Planar Velocity of Individual Players From 2D Video of Youth Ice Hockey Games,” Submitted to Journal of Sports Engineering.

P. Dehghan et al., “Validation of Deep Learning-Based Velocity Estimation from Single, Side-View Ice Hockey Game Footage Using Top-Down Drone Measurements,” Submitted to Journal of Sports Biomechanics.

A. Azadi, P. Dehghan , R. Muhammad, C. Karton , M. Fraiser, R. Lagraniere, R, Graham, T.B. Hoshizaki, “Automated Detection of Physical Contact Events in Youth Ice Hockey: A Player-Focused Deep Learning Approach,” submitted to Journal of Biomechanics, 2025.

A. Azadi, P. Dehghan, C. Karton, M. D. Gilchrist, M. Fraser, and T. B. Hoshizaki, “Deep Learning Model for Maximum Principal Strain Prediction from Ice Hockey Video-Derived Impact Features: Velocity, Compliance, Location, and Elevation,” 2025.