A Deep Learning Pipeline for Impact Event Detection, Head Impact Classification, and Brain Strain Estimation in Youth Ice Hockey Videos

Amirhossein Azadi

Thesis submitted to the University of Ottawa
in partial Fulfillment of the requirements for the
Doctor of Philosophy

School of Human Kinetics
Faculty of Health Sciences
University of Ottawa

Declaration of Authorship

This thesis includes material that has been published or submitted for publication in peer-reviewed venues. I, Amirhossein Azadi, am the primary author of all manuscripts and confirm that I made the leading contribution to the conception and design of the studies, data collection and curation, model development, data analysis and interpretation, and drafting and revising of the manuscripts. The contributions of co-authors for each chapter are detailed below.

Chapter 2 – Automated Detection of Physical Contact Events in Youth Ice Hockey: A Player-Focused Deep Learning Approach (Nature Scientific Reports — Submitted)

Amirhossein Azadi: Led the conception and design of the study; developed the overall deep learning pipeline for contact event detection; implemented and tuned the detection, tracking, and classification models; performed data curation and statistical analyses; interpreted the results; and drafted and revised the manuscript.

Parisa Dehghan: Contributed to study design; coordinated and performed video review and event annotation; assisted with data curation, model implementation, and interpretation of results; and provided critical revisions to the manuscript.

Rowan Mohamed Amin Hefny Hussein: Contributed to the development and evaluation of tracking and baseline comparison methods; assisted with interpretation of results; and reviewed and edited the manuscript.

Clara Karton: Assisted with data collection and experimental design; provided expertise in injury biomechanics and interpretation of collision exposure metrics; and reviewed and edited the manuscript.

Maia Fraser: Provided expertise in machine learning methodology; contributed to study design and interpretation of model performance; and reviewed and edited the manuscript.

Robert Laganière: Provided expertise in computer vision and tracking methodology; contributed to method development and interpretation; and reviewed and edited the manuscript.

Ryan B. Graham: Contributed to study design and interpretation of the results in the context of human kinetics; and reviewed and edited the manuscript.

T. Blaine Hoshizaki: Supervising author; provided overall conceptual guidance; contributed to study design and interpretation; and provided critical revisions to the manuscript.

Chapter 3 – A Deep Learning Network for Detecting Head Impacts in Ice Hockey from 2D Game Video (IRCOBI 2024 — Published)

Amirhossein Azadi: Led the conception and design of the study; developed the head impact detection pipeline, including detection, tracking, and LRCN-based classification; performed data curation, analyses, and interpretation; and drafted and revised the manuscript.

Parisa Dehghan: Assisted with video review and head impact annotation; contributed to data curation, model implementation, and interpretation of results; and reviewed and edited the manuscript.

Ryan B. Graham: Contributed to study design and interpretation of the results in the context of human kinetics and sport applications; and reviewed and edited the manuscript.

T. Blaine Hoshizaki: Supervising author; provided overall conceptual guidance; contributed to study design and interpretation; and provided critical revisions to the manuscript.

Chapter 4 – Deep Learning Model for MPS Prediction from Ice Hockey Video-Derived Impact Features (Sports Engineering — Published)

Amirhossein Azadi: Led the conception and design of the study; designed the feature-extraction strategy and neural network architecture; performed FE-based ground truth generation, data preprocessing, model training, and statistical analysis; interpreted the results; and drafted and revised the manuscript.

Parisa Dehghan: Assisted with data collection and organization of impact datasets; contributed to model implementation and interpretation of results; and reviewed and edited the manuscript.

Clara Karton: Provided expertise in injury biomechanics and FE-based brain injury modelling; contributed to experimental design, interpretation of MPS-related findings, and manuscript revisions.

Michael D. Gilchrist: Provided expertise in FE modelling and material characterization; contributed to the interpretation of brain strain results; and reviewed and edited the manuscript.

Maia Fraser: Provided expertise in deep learning and model evaluation; contributed to methodological design, analysis of model performance, and manuscript revisions.

T. Blaine Hoshizaki: Supervising author; provided overall conceptual guidance; contributed to the design of the study and interpretation of results; and provided critical revisions to the manuscript.

I confirm that the above statements accurately reflect the nature and extent of the contributions of all authors for the manuscripts included in this thesis.

Abstract

Ice hockey poses a significant risk of head impacts, increasing the likelihood of traumatic brain injuries (TBIs), including concussions. Youth players constitute the majority of the ice hockey population, accounting for approximately 69% of all registered players in the United States alone. Notably, 15–25% of all hockey-related injuries in this age group involve head trauma, underscoring the high prevalence of TBIs among youth participants. Younger players face distinct vulnerabilities due to their larger head-to-body ratio, lower skill levels, weaker neck muscles, and less refined motor control, all of which contribute to a playing style that elevates their risk of head impacts. Furthermore, youth leagues often lack access to professional-grade monitoring systems, such as helmet-based sensors or on-site clinical staff, making early detection and response to brain injuries particularly challenging.

To address these limitations, this thesis introduces a video-based, artificial intelligence (AI)-powered system designed to automatically detect and analyze head impacts in youth sports, facilitating both large-scale dataset generation and injury assessment.

Chronologically, the first component of this work, referred to as Study 2 in the thesis, focused on head impact detection. This study was conducted using professional hockey footage, which offered consistent lighting, high video resolution, and clear player visibility, allowing for early testing under ideal visual conditions. In this phase, we manually annotated head impact moments by reviewing full-game recordings and cropping video segments to center on and isolate each player involved in the contact. To automate this process and enable future scalability, we also introduced a detection and tracking pipeline tailored to professional players, combining You Only Look Once version 8 (YOLOv8)x for player detection—achieving precision 0.97, recall 0.97, and mean average precision (mAP)50 0.99 (at a 50% intersection over union (IoU) threshold)—with StrongSORT for maintaining player identity across frames. The resulting player-centered clips were used to train a long-term recurrent convolutional network (LRCN), which achieved an accuracy of 0.87 in distinguishing head impacts from non-head impact clips. While this study demonstrated the feasibility of video-based head impact detection, it was limited by a small, manually constructed dataset, which restricted scalability and introduced potential inconsistencies in annotation.

Building on the head impact detection framework, the second phase focused on developing a scalable, youth-specific contact detection pipeline. This shift was motivated by a key limitation of the first phase: head impacts are relatively rare events, especially in youth games, making it difficult to build large training datasets for reliable head impact

classification. By broadening the scope to detect all physical contact events, not just head impacts, we were able to substantially increase the pool of relevant video clips and create a more scalable and efficient system for data generation and pre-filtering. A fine-tuned YOLOv8 model demonstrated strong player detection performance across diverse youth age groups and challenging visual environments, achieving precision 0.96, recall 0.97, average precision (AP)0.7 (at a 70% IoU threshold) of 0.94, mAP50 of 0.97, and mAP50–95 (averaged over IoU thresholds from 50% to 95%) of 0.82. For tracking, we evaluated the StrongSORT-based module on 100 benchmark clips characterized by high occlusion and rapid movement. The integration of an IoU-based cost term reduced identity switches from 172 to 53 (a 69% improvement) and improved Multiple Object Tracking Accuracy (MOTA) from 89.0% to 94.5%, confirming the system's robustness under realistic youth hockey conditions. A temporal shift module (TSM) network was then applied to classify physical contact events (e.g., player-to-player, player-to-board), achieving a recall of 0.94 and reducing manual review time from several hours to under 30 minutes per game.

The final study introduced a deep learning model to estimate maximum principal strain (MPS), a validated biomechanical indicator of brain tissue deformation, directly from video-derived features. Using 477 reconstructed youth hockey impacts involving players aged 5–17, four features were extracted from video for each event: impact velocity, surface compliance, impact location, and elevation, with player age group included as an additional non-video input. A fully connected neural network trained on this dataset achieved strong predictive performance, with a test-set coefficient of determination (R^2) = 0.89 and mean squared error (MSE) = 0.0015. Permutation importance analysis confirmed that impact velocity was the most influential predictor. These results demonstrate the feasibility of estimating brain strain from video features and basic contextual information, offering a scalable, non-invasive approach for monitoring injury risk in youth sports.

Together, these three studies form a comprehensive, video-based pipeline for detecting impactful events, identifying head impacts, and estimating head impact severity in youth ice hockey, advancing injury surveillance in settings where traditional tools are often unavailable.

Acknowledgements

I would like to express my deepest gratitude to my wife and colleague, Parisa Dehghan, for her unwavering love, encouragement, and intellectual support throughout this journey. Her understanding and patience have been invaluable in helping me overcome challenges and remain focused on my goals.

I am profoundly thankful to my parents, who, despite being far away, have always sent me their endless energy, prayers, and moral support. Their belief in me has been a constant source of strength.

I extend my sincere appreciation to my supervisor, Dr. Blaine Hoshizaki, for his mentorship, guidance, and expertise. His insightful feedback and dedication to advancing research have greatly shaped the direction and quality of this work.

This thesis would not have been possible without the contributions, support, and inspiration of these remarkable individuals.

Table of Contents

List of Tables

List of Figures

Abbreviations

R^2 coefficient of determination vi, xii, xiv, 49, 56, 58, 59, 61–64, 73, 75, 76

2D two-dimensional 6, 7, 74–76

AI artificial intelligence v, 5, 7, 9, 12, 13, 32, 41, 51, 63

AP average precision vi, 11, 14, 18

BrIC Brain Injury Criterion 4, 50

CNN convolutional neural network 31, 33, 42, 43

ConvLSTM convolutional long short-term memory 31, 33, 43

CTE chronic traumatic encephalopathy 7, 9

CV computer vision 7, 12, 24, 32–34, 42, 75

DTI diffusion tensor imaging 4

EMA exponential moving average 15

FE finite element 5, 6, 49, 51, 56, 63, 64, 73, 76

FEM finite element model 73

FPS frames per second 14

GAMBIT Generalized Acceleration Model for Brain Injury Threshold 4, 50

HAR human action recognition 32, 33

HIC Head Injury Criterion 4, 50

HIP Head Impact Power 4, 50

ID identity xiv, 11, 15, 18, 31, 35, 40–43, 71, 72, 76

IoU intersection over union v, vi, xiv, 11, 14, 15, 18, 26, 34–36, 40, 41, 76

LRCN long-term recurrent convolutional network v, xiii, 31, 33, 37, 39, 43, 70, 75

LSTM long short-term memory network 33, 37, 38

mAP mean average precision v, vi, 11, 14, 18, 31, 34, 38, 42, 43

MCC Matthews Correlation Coefficient xiii, 11, 17–19, 21, 26

MOTA Multiple Object Tracking Accuracy vi, 11, 15, 18, 76

MPS maximum principal strain vi, x, xiv, 4–8, 48–52, 54, 56–60, 63–65, 72, 73, 75–77

MSE mean squared error vi, xii, xiv, 49, 56, 58, 59, 61–64, 73, 75, 76

mTBI mild traumatic brain injury 1, 4, 31

NHL National Hockey League 34, 37

NMS non-maximum suppression 35

PLA peak linear acceleration 4, 50

PR-AUC precision–recall area under the curve 11, 17, 20, 22

PRA peak rotational acceleration 4, 50

PTSD post-traumatic stress disorder 7

RAFT Recurrent All-Pairs Field Transforms 17

RHI repetitive head impacts 1, 31, 32, 41

RIC Rotational Injury Criterion 4, 50

TBI traumatic brain injury v, 1, 4, 7, 11, 31, 49, 50

TSM temporal shift module vi, 11, 16, 17, 20, 24, 26, 33, 71, 72, 75, 76

TSN temporal segment network 33

ViT Vision Transformer 72

YOLOv8 You Only Look Once version 8 v, vi, 11, 14, 18, 26

Chapter 1

Introduction

Medical professionals have studied the relationship between impact severity and brain injury for over a century (Casper, 2018), seeking to understand the complexities of head injuries in sports. Brain injury mechanisms are inherently complex due to neuroanatomical variability, heterogeneous injury patterns, and multifaceted interactions between biomechanical forces (Karton and T Blaine Hoshizaki, 2018). Individual factors including age, sex, genetics, and pre-existing conditions further modulate injury responses (Giza and Hovda, 2014; Bazarian et al., 2010; McAllister, 2010).

TBI represents a heterogeneous condition arising from various physical forces, producing diverse neurological symptoms. Concussion, classified as mild traumatic brain injury (mTBI), carries significant health risks when improperly managed, particularly following repeated injuries. Unlike severe TBI, mTBI typically involves functional disturbances (e.g., dizziness, cognitive impairments) without visible structural damage, yet can yield acute and chronic consequences (Karton, T. Hoshizaki, and M. Gilchrist, 2014; Karton and T Blaine Hoshizaki, 2018).

In the United States, over 500,000 adolescents sustain mTBIs annually, many of which are subclinical. These contribute to cumulative brain damage from repetitive head impacts (RHI), elevating risks for neurodegenerative diseases (Nowinski et al., 2024). RHI refers to accumulated head impacts, symptomatic and asymptomatic, common in contact sports like ice hockey.

Ice hockey carries a significant risk of head impacts, increasing susceptibility to traumatic brain injuries, including concussions (Daneshvar et al., 2011). Youth athletes account for roughly 69 % of all registered U.S. players, and 15–25% of hockey-related injuries in this group involve head trauma (USA Hockey, 2024). Youth players are especially vulnerable to concussive impacts due to their larger head-to-body ratios, weaker neck muscles, and under-developed motor control (Cournoyer et al., 2021). These factors make head-injury prevention and monitoring particularly important in youth hockey.

Head-impact magnitude and frequency are two important variables for understanding brain-trauma risk. Several physical factors influence impact magnitude, including the velocity of the collision, the location and elevation of the head strike, and the type of event (e.g., shoulder check, fall to the ice). Repeated impacts can lead to cumulative neurological stress, emphasizing the need to examine both single-event severity and long-term exposure (Karton and T Blaine Hoshizaki, 2018; Karton, Blaine Hoshizaki, and M. D. Gilchrist, 2020; Karton and Thomas Blaine Hoshizaki, 2021; Karton, Andrew Post, et al., 2021; Andrew Post and T Blaine Hoshizaki, 2012; Andrew Post and Blaine Hoshizaki, 2015; Post et al., 2019; Andrew Post, Kendall, et al., 2015).

The magnitude of a head impact is important in determining the severity of the resulting brain injury **(see Figure 1.1)**. The greater the magnitude of the impact, the more severe the injury is likely to be (Andrew Post and T Blaine Hoshizaki, 2012; Andrew Post and Blaine Hoshizaki, 2015; Post et al., 2019; Andrew Post, Kendall, et al., 2015). The severity of a head impact is influenced by several biomechanical parameters **(Figure 1.1)**, including:

1. Impact Event Type and Surface Compliance:

The specific event type fundamentally dictates impact biomechanics and injury mechanisms. Different impact types influence not only the shape of the head-acceleration pulse but also its duration, which is a key determinant of brain-tissue deformation. Video analysis in youth ice hockey shows that body-to-body collisions account for approximately 62% of head-contact events, while board or glass impacts contribute 16%, and falls to the ice account for 11% (Robidoux et al., 2020). Laboratory reconstructions have shown the head-to-puck impacts produce sharp pulses lasting less than 5 ms; head-to-ice impacts have durations of approximately 10–15 ms; elbow-to-head and shoulder-to-head collisions fall within the 10–28 ms and 20–35 ms ranges, respectively; head-to-glass impacts last about 15–20 ms; and head-to-boards impacts have durations of approximately 22–26 ms (Andrew Post, T Blaine Hoshizaki, Karton, et al., 2019). These event-specific durations critically determine force distribution profiles: longer-duration impacts (shoulder/board/glass) distribute energy over extended periods, generating cumulative strain, while shorter, high-magnitude pulses (puck/ice) concentrate energy delivery, potentially exceeding instantaneous tissue tolerance thresholds. Consequently, accurate concussion risk assessment requires explicit consideration of event type, as each category represents a distinct biomechanical loading regime with unique pathological implications.

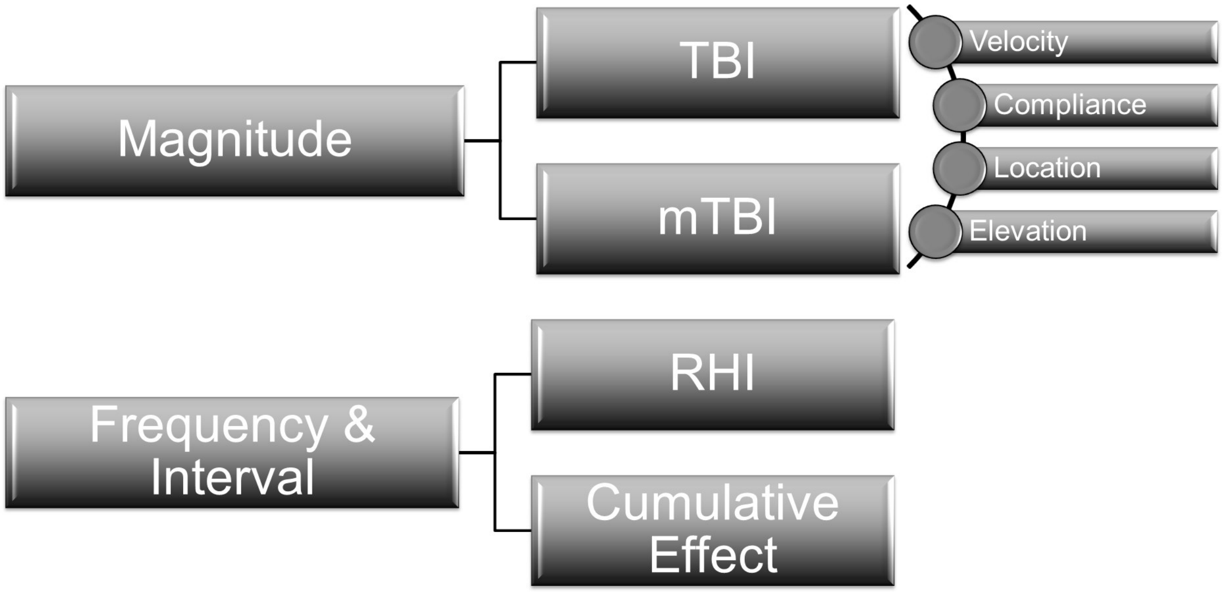


Figure 1.1: Understanding the head impact and Brain Injury Risks": Magnitude, Frequency, and interval

2. Location of Impact:

The specific area of the head that absorbs an impact influences both the mechanical load transmitted to the brain and the resulting injury pattern. Regional differences in skull curvature and underlying anatomy lead to distinct deformation pathways. In a helmeted–youth-hockey cohort, Mihalik et al. (Mihalik et al., 2012) found that crown (top-of-head) impacts produced the highest mean peak linear accelerations, significantly exceeding those recorded for frontal, lateral, or rear impacts. This location-dependent behaviour underscores the need for concussion-risk models to treat the head as a heterogeneous structure rather than a mechanically uniform shell.

3. Elevation of Impact:

The vertical position where impact occurs relative to the head's center of gravity, whether superior, aligned, or inferior, modulates rotational acceleration magnitude, a biomechanical determinant of brain trauma. This elevation-dependent effect occurs because impact height directly influences the lever arm through which forces act on the cranium, generating distinct rotational moments that affect deep brain structures. Understanding this mechanical relationship is therefore important for accurate injury risk assessment (Andrew Post and T Blaine Hoshizaki, 2012; Andrew Post, Kendall, et al., 2015).

4. Velocity of Impact:

Impact velocity influences the amount of energy transferred to the brain, resulting in varying degrees of deformation and stress on brain tissue during a collision. Accurately measuring velocity enables researchers to quantify the mechanical load applied to the head and estimate the potential risk of injury. Furthermore, velocity data play an important role in event reconstruction, an analytical process used to understand the dynamics of a head impact. Through reconstruction, scientists can evaluate the head's kinematic response, including both linear and rotational accelerations experienced by the brain. This information is essential for assessing the risk of concussions and other forms of TBI (Andrew Post and T Blaine Hoshizaki, 2012; Andrew Post and Blaine Hoshizaki, 2015; Post et al., 2019; Andrew Post, Kendall, et al., 2015).

In brain injury biomechanics, numerous metrics quantify impact severity and assess injury risk. Traditional measures, including peak linear acceleration (PLA), peak rotational acceleration (PRA), Head Injury Criterion (HIC) (McHenry, 2004), Brain Injury Criterion (BrIC) (Takhounts et al., 2013), Generalized Acceleration Model for Brain Injury Threshold (GAMBIT) (J. A. Newman, 1986), Head Impact Power (HIP) (J. Newman et al., 2000), and Rotational Injury Criterion (RIC) (Kimpara and Iwamoto, 2012), provide useful approximations of head kinematics but capture only partial aspects of the kinematic profile, typically focusing on peak values while neglecting the complete time-history profiles and three-dimensional vector characteristics. These metrics often fail to fully capture the complex tissue-level mechanical responses associated with concussions and mild traumatic brain injuries (mTBIs) (Rowson and Duma, 2013; Gabler, Crandall, and Panzer, 2016).

To overcome these limitations, MPS has emerged as an important biomechanical metric that directly quantifies three-dimensional, time-dependent brain tissue deformation during impacts (Kleiven, 2007). Unlike traditional kinematic measures, MPS integrates linear/rotational accelerations and impact duration to estimate peak tensile strain, thereby bridging external forces with internal injury mechanisms (Gabler, Crandall, and Panzer, 2019; Hernandez et al., 2015). Importantly, MPS accounts for neuroanatomical heterogeneity, including gray-white matter interfaces, sulcal geometry, and viscoelastic anisotropy (Chatelin, Constantinesco, and Rémy Willinger, 2010; Giordano and Kleiven, 2014), which governs regional vulnerability. This physiological relevance is further validated by diffusion tensor imaging (DTI) studies linking head impact exposure to altered white matter microstructure (DavenportElizabeth et al., 2014; Churchill et al., 2017). Clinically, MPS thresholds of 15–25% correlate with elevated concussion risk and structural damage for both animal models and human studies (Hernandez et al., 2015; Patton et al., 2012; Bain and Meaney, 2000), establishing MPS as a biologically grounded predictor of injury severity.

Finite element (FE) models enable high-fidelity computation of MPS by simulating brain responses to impacts using anatomically accurate geometries, viscoelastic tissue properties, and realistic skull-brain boundary conditions (Horgan and M. D. Gilchrist, 2003; Ho and Kleiven, 2007; Trotta et al., 2020). These simulations capture complex interactions between linear/rotational accelerations and neuroanatomical heterogeneity, generating detailed spatiotemporal strain distributions. However, the computational demands of FE modeling, requiring hours to days per simulation, and its reliance on specialized expertise limit its practicality for use in sporting environments like youth hockey (Kleiven, 2007; Raul et al., 2008). These constraints necessitate alternative approaches such as AI-driven models that provide rapid, accessible MPS estimation without full FE simulation.

While impact magnitude is an important indicator of brain injury risk, it does not fully capture the complexity of injury mechanisms. Researchers have increasingly examined how the interaction of magnitude, frequency, and interval contributes to injury severity (Greenwald et al., 2008; Karton, Blaine Hoshizaki, and M. D. Gilchrist, 2020; Karton, Andrew Post, et al., 2021), with frequency and interval being particularly important for cumulative risk assessment (Figure 1.1). Repeated exposure to even low-magnitude impacts may not produce immediate symptoms but can lead to long-term neurological damage. The accumulation of such sub-concussive impacts has been associated with structural brain changes and increased risk of neurodegenerative diseases (Karton and T Blaine Hoshizaki, 2018; Karton, Blaine Hoshizaki, and M. D. Gilchrist, 2020; Karton, Andrew Post, et al., 2021).

Understanding the impact frequency in specific sports or activities allows for the implementation of evidence-based policies and rules to promote player safety. This can involve modifying equipment, adjusting training techniques, or implementing guidelines to limit the number and intensity of head impacts (Chen et al., 2023; Broglio et al., 2011; Wilcox et al., 2014). These insights reinforce the need for scalable, objective systems to monitor head impact exposure and support long-term athlete health.

With the widespread adoption of routine video recording across sports leagues and age groups, an unprecedented volume of high-quality game footage is now readily available. This growing accessibility creates a unique opportunity to develop fully automated, video-based injury-surveillance systems that can scale efficiently from professional to youth settings without additional hardware or manual intervention.

Despite this potential, many of the tools currently used to monitor head-impact exposure remain limited in practicality and scalability, especially in real-world sports environments. Manual observation and player self-reporting continue to be the most commonly used methods for tracking head impacts across many sports, including youth ice hockey. Coaches and sports scientists typically rely on observation sheets or digital forms to document incidents during games or practices, while athletes provide additional information through interviews and symptom checklists. However, these methods are labor-intensive, prone to subjective bias, and vulnerable to human error. The rapid pace, frequent player collisions, and visual congestion inherent to ice hockey further increase the risk that observers will miss, misinterpret, or inconsistently record critical events. Additionally, athletes may underreport symptoms due to peer pressure, a desire to continue playing, or a lack of symptom awareness (Kroshus et al., 2015; Putukian et al., 2013).

While these manual approaches may be sufficient for immediate, on-field decision-making, they often fail to capture subclinical impacts and cumulative head loading, factors that may not produce acute symptoms but can contribute to long-term neurological damage (Kroshus et al., 2015; Putukian et al., 2013; Meehan, d'Hemecourt, and Dawn Comstock, 2010; Kamins et al., 2017). These limitations underscore the need for scalable, objective, and automated systems capable of accurately monitoring head-impact exposure in fast-paced sports like ice hockey.

Wearable sensors have been introduced as a more automated method for detecting head impacts during play, particularly in elite-level research settings. However, these devices present several limitations. They are costly, often inaccessible to youth and amateur teams, and have been shown to underestimate rotational acceleration, an important biomechanical factor in concussions. They frequently generate false positives that must be manually verified through video analysis (Jadischke et al., 2013; Rueda, Cui, and M. D. Gilchrist, 2011).

A third, research-focused approach involves the use of video analysis, physical reconstructions, and computational models. This method typically requires highly trained personnel to manually review entire games, identify potential head-impact events from two-dimensional (2D) video footage, and extract key impact features. These features are then used to reconstruct the impacts in laboratory settings, where detailed dynamic responses, such as linear and rotational velocities and accelerations, are captured. Finally, these kinematic inputs are applied to FE brain models to compute tissue-level injury metrics such as MPS and Von Mises stress (Remy Willinger and Baumgartner, 2003; Andrew Post, T Blaine Hoshizaki, M. D. Gilchrist, et al., 2014; Karton and T Blaine Hoshizaki, 2018; Karton, Andrew Post, et al., 2021). Although this approach yields highly informative, physiologically meaningful results, it is extremely labor-intensive, time-consuming, and heavily reliant on specialized equipment and expert knowledge, which limits its scalability for large datasets and precludes its use in real-time or near-real-time injury monitoring.

To address these limitations, AI and computer vision (CV) have emerged as powerful alternatives for head impact assessment. AI enables automated detection of head impacts directly from standard 2D video footage, extraction of relevant biomechanical features, and estimation of brain strain metrics such as MPS, which are directly linked to concussion risk. CV allows for efficient processing of large volumes of game footage, making it feasible to generate extensive, validated datasets of head impact events that would be impractical to compile manually. These datasets enable researchers to capture diverse impact scenarios and analyze the complex factors contributing to brain trauma. By integrating demographic and contextual information, such as age, sex, and skill level, AI-driven systems can further support investigations into long-term neurological outcomes, including chronic traumatic encephalopathy (CTE) and cognitive deficit (Fig. 1.2).

Recent studies have demonstrated the potential of AI-based models to improve brain injury monitoring and prediction. Machine learning algorithms have been applied to large datasets to identify biomarkers and risk factors for TBI, supporting more accurate injury assessment and early intervention (Rodger, 2015; Agoston and Langford, 2017). AI approaches have also shown promise in predicting secondary complications, such as post-traumatic stress disorder (PTSD) and other psychiatric conditions commonly associated with TBI (Yurgil et al., 2014). Collectively, these advances highlight the potential of AI-driven solutions to overcome the key limitations of traditional monitoring systems and enable more scalable, objective, and proactive brain injury assessment and care (Rowson, Duma, et al., 2012).

To address the lack of accessible and scalable monitoring tools in youth hockey, this thesis presents a novel video-based, AI-driven system for detecting head impacts using standard 2D game footage. The overall pipeline is organized into three main stages (hereafter referred to as Stage 1, Stage 2, and Stage 3), which correspond to the three technical chapters of the thesis.

Stage 1 introduces an automated contact event detection network that analyzes full-game footage to identify physical contact moments, substantially reducing manual review time and enabling scalable dataset creation for downstream analysis.

Stage 2 employs deep learning to classify head impacts within the clips identified in Stage 1, distinguishing them from non-head contact (e.g., body checks). This generates a

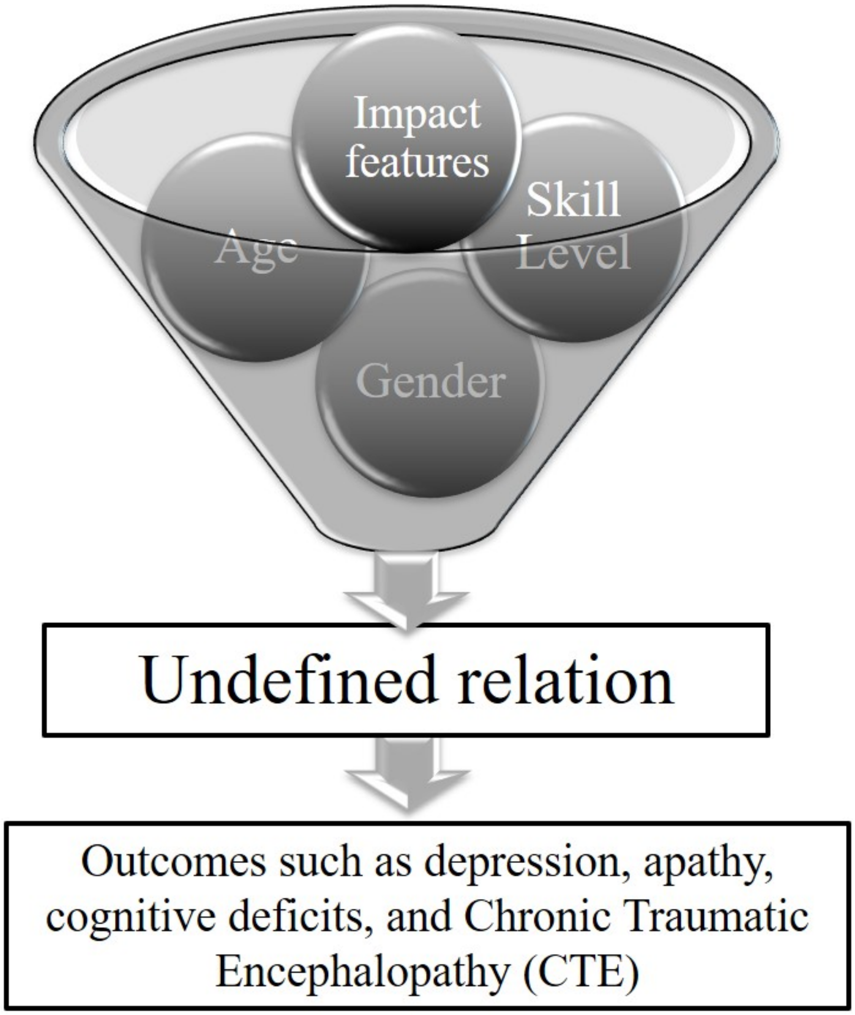


Figure 1.2: Illustration of the multifactorial relationships between head impact exposure and long-term neurological outcomes, which are still under investigation.

reliable dataset of video-verified head impact events.

Stage 3 advances the pipeline by training neural networks to estimate MPS using four video-derived features (impact velocity, surface compliance, location, and elevation) with player age group as supplementary input.

Although the models underlying Stage 2 were developed earlier in time than those in Stage 1, the stages are presented in the order in which they would be used in an operational pipeline, progressing from contact event detection to head-impact classification and finally to brain-strain estimation.

Together, these stages create a framework for detecting physical contact, isolating head impacts, and estimating injury severity directly from broadcast-quality footage, enabling continuous surveillance across professional and youth ice-hockey leagues. By generating a large, rigorously validated dataset, the pipeline supports biomechanical research into how impact frequency, type, and location influence acute concussion risk and cumulative brain trauma (Barnes et al., 2018; Anderson et al., 2005; Fann, Hart, and Schomer, 2009; Stulemeijer et al., 2010; Rowson, Duma, et al., 2012). Scaled, accurately labelled head-impact data allow detection of subtle patterns invisible in smaller samples, facilitate advanced AI and statistical modelling, and, when combined with demographic factors such as age and sex, can clarify links between repeated head trauma and outcomes including cognitive decline and CTE.

Chapter 2

Automated Detection of Physical Contact Events in Youth Ice Hockey: A Player-Focused Deep Learning Approach

Nature Scientific reports — Submitted

Amir Azadi[a,*], Parisa Dehghan[a], Rowan Mohamed Amin Hefny Hussein[b], Clara Karton[a], Maia Fraser[c], Robert Laganière[b], Ryan B. Graham[a], T. Blaine Hoshizaki[a]

[a] School of Human Kinetics, University of Ottawa, Ottawa, Ontario, Canada
[b] School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, Ontario, Canada
[c] Department of Mathematics and Statistics, University of Ottawa, Ottawa, Ontario, Canada

2.1 Abstract

Youth ice hockey lacks scalable, automated tools to quantify collision exposure and flag head-impact events from routine game video. This study presents a player-centric, video-based pipeline that detects physical contacts and uses them as a high-frequency proxy to accelerate head-impact dataset creation and support injury surveillance.

Broadcast footage was segmented into 30-frame clips. Players were detected using a youth-tuned YOLOv8 model and tracked with StrongSORT augmented by an IoU–based cost term. Contact events from 20 games (U11–U15) were manually annotated to generate contact-centered clips; non-contact clips were sampled to create datasets with contact:non-contact ratios of 1:4, 1:8, and 1:12. A TSM classifier operated on player crops, with crop scale and class balance systematically evaluated. Performance was assessed using precision, recall, F1, specificity, MCC, and precision–recall area under the curve (PR-AUC), and compared with a full-frame baseline.

The youth-specific YOLOv8 achieved precision 0.96, recall 0.97, AP0.7 0.94, mAP50 0.97, and mAP50–95 0.82. IoU-assisted tracking reduced ID switches by 69% (172→53) and improved MOTA from 89.0% to 94.5%. The best classifier (15 segments, shift-div = 4, 1.5× crop, 1:12 ratio) achieved precision 0.96, recall 0.94, F1 0.95, specificity ≈1.00, and MCC 0.95. On two full-game tests, PR-AUC was 0.51; at a 0.50 threshold, precision was 0.25 and recall was 0.79. As a head-impact pre-filter, the system detected 19/22 impacts (86.4% recall) and reduced review time from over three hours to under 30 minutes per game.

This pipeline combines youth-specific detection, IoU-enhanced tracking, and temporal context–aware classification to deliver a scalable, efficient solution for collision exposure monitoring and dataset generation, supporting safer, data-driven youth hockey environments.

Keywords: Action recognition; Deep learning; Injury surveillance; Physical contact detection; Temporal Shift Module; Youth ice hockey

2.2 Introduction

Ice hockey is a high-intensity contact sport in which athletes are routinely exposed to collisions that generate significant biomechanical forces. Among the most serious consequences of these impacts are TBIs, particularly concussions, which are a growing concern in both professional and youth sports (Daneshvar et al., 2011). Head impacts in ice hockey span a wide spectrum, ranging from repetitive, low-magnitude contacts to acute, high-energy collisions, each contributing variably to neurological trauma. While some consequences of these impacts are immediate, others may remain subclinical for extended periods, with evidence indicating that repeated head trauma can result in cumulative and potentially debilitating effects over time (Karton and Hoshizaki, 2018; Gysland et al., 2012; McKee et al., 2013; Nowinski et al., 2024).

The issue of head injury is especially pertinent in youth hockey, which comprises an estimated 60–70% of all organized hockey participants (USA Hockey, 2024). Youth athletes are believed to be at elevated risk due to several developmental factors, including incomplete neurodevelopment, decreased postural control, and increased susceptibility to falls (Johnson, 2011; Mihalik et al., 2012; Pfister et al., 2016). Despite this vulnerability, youth leagues generally lack the robust surveillance systems available in professional settings. While professional teams often utilize wearable sensors, video analytics, and clinical oversight to assess and manage head impacts, these resources are rarely accessible at the youth level (Echemendia et al., 2017; Johnson, 2011; Post and Hoshizaki, 2012). As a result, many injuries may go undetected or unmanaged, impeding both clinical care and the accumulation of high-quality injury data.

Given the resource constraints in youth sports settings, there is a pressing need for scalable, low-cost tools to monitor and analyze injury-related events. Video analysis may be a practical and cost-effective alternative to sensor-based monitoring, especially with recent advances in AI and CV. These technologies enable the automated extraction of relevant information from image and video data, removing the need for labor-intensive manual review (Allard et al., 2022; Butterfield et al., 2023; Shill et al., 2023). In sports, CV has been effectively used for player detection, tracking, and action recognition in both team and individual sports (Dehghan et al., 2024; Naik, Hashmi, and Bokde, 2022; F. Wu et al., 2022).

Recent studies have demonstrated the feasibility of using CV to detect head impacts directly from video footage. Systems such as DeepImpact, developed by Rezaei and L. C. Wu (2022), utilize deep learning models for object detection and temporal modeling to identify ball-to-head contacts in soccer, offering a scalable alternative to wearable sensors. However, ice hockey presents a more complex analytical landscape. Compared to soccer, hockey head impacts are infrequent—typically fewer than ten per game—and highly variable in appearance. The sport's fast tempo, frequent player occlusion, similar uniform attire, and lower camera perspectives compound the difficulty of accurate head impact detection (Azadi et al., 2024).

To address these challenges, this study proposes a broader strategy: automatically detecting all physical contact events, regardless of head involvement. This approach offers several advantages. First, physical contacts occur far more frequently—averaging over 200 per game—providing sufficient training data with fewer video analyses. Second, physical contacts are visually more distinct than non-contact scenes, making them easier for deep learning models to detect reliably. Most importantly, this method serves as an effective pre-filtering step to support the creation of head impact datasets. Manual review of entire game footage to locate head impacts is time-consuming, often requiring over three hours to find just ten head impacts. Automated contact detection can reduce this to under 30 minutes, significantly improving dataset curation efficiency for downstream head impact classification.

The ability to automatically detect and classify physical contact has broad implications for sports medicine, coaching, and injury prevention. Data on the frequency, location, and type of physical contact can inform player workload analysis, safety rule enforcement, and the design of training programs. This system may also assist in monitoring non-cranial injuries, such as those involving the shoulder, knee, or spine, which often receive less attention in surveillance efforts. In addition, the system holds significant value for head trauma research, providing an efficient tool to support the identification and analysis of head impacts in sports.

Despite the proliferation of deep learning applications in sports video analysis, to date no fully automated, video-based contact event detection system has been developed or validated specifically for youth ice hockey. This study aims to fill that gap by introducing a novel deep learning pipeline for automated physical contact detection in youth hockey video footage. By leveraging a high-frequency event category—general physical contacts—this system improves training data availability, reduces manual annotation time, and enables scalable injury surveillance under resource-constrained conditions.

The objectives of this study are twofold: (1) to develop and evaluate a deep learning-based system capable of detecting physical contacts in youth ice hockey videos; and (2) to demonstrate the system's potential to facilitate head impact dataset generation. By doing so, this work advances the use of AI for injury monitoring and lays the foundation for safer, data-informed youth sport environments.

2.3 Methods

2.3.1 Overview

We developed a deep learning-based pipeline tailored to detect physical contacts in youth ice hockey using broadcast video footage. The pipeline comprises five primary components: (1) video segmentation, (2) player detection and tracking, (3) manual event annotation, (4) dataset construction and post-processing, and (5) contact event classification. Our design prioritizes computational efficiency and robustness under the challenging visual conditions of youth hockey footage.

2.3.2 Video Segmentation

Full-length hockey games were segmented into clips of 30 consecutive frames (approximately 1 second at 30 frames per second (FPS)). This segment length balances temporal context with manageable computational cost. For source videos recorded at 60 FPS, we downsampled by selecting every other frame to maintain consistency in processing. Segment-based sampling was informed by prior work in action recognition (Wang et al., 2016).

2.3.3 Player Detection and Tracking

Player Detection. We developed a player detection system using YOLOv8 (Jocher et al., 2023), fine-tuned specifically for youth hockey. Our previous model, trained on professional NHL footage (Azadi et al., 2024), performed well on adult games but showed reduced accuracy on youth games due to differences in lighting, camera angles, and player appearance. To address this, we curated a dataset of 4,700 annotated images from 20 youth games across U11, U13, and U18 levels. Annotations covered five classes: *Player_Light*, *Player_Dark*, *Goalie_Light*, *Goalie_Dark*, and *Referee*. We fine-tuned the YOLOv8 model from our previous study using 1280×1280-pixel inputs for 250 epochs. Detection performance was evaluated using precision, recall, AP0.7, mAP50, and mAP50–95.

Player Tracking. For multi-object tracking, we developed an IoU-assisted StrongSORT algorithm that addresses the frequent occlusions and player overlaps characteristic of ice hockey (Azadi et al., 2024; Du et al., 2023; Wojke, Bewley, and Paulus, 2017). Standard tracking methods often struggle with the homogeneous appearance of same-team players and rapid direction changes. Our modified approach introduced: (1) temporary overlapping bounding boxes distinguished by jersey color, and (2) an IoU-based matching cost in StrongSORT's appearance matrix, reducing ID switches. An elevated IoU threshold of 0.85 was applied for redundant detection elimination. Tracking performance was evaluated using MOTA (Bernardin and Stiefelhagen, 2008) and ID switch counts on 100 three-second clips from youth games.

2.3.4 Event Annotation

Twenty youth hockey games (U11, U13, U15) were manually reviewed to annotate contact events. Annotation criteria:

- **Player-to-player collisions:** Included only if both players were fully visible and contact was forceful.
- **Falls to ice:** Included if torso or hips contacted the ice and head/upper body dropped visibly.
- **Board and glass impacts:** Included only if clearly visible and forceful.
- **Stick and puck impacts:** Included when clearly visible.

Each event was annotated with the frame number and coordinates of the involved player's head and body.

2.3.5 Dataset Construction and Post-Processing

Contact clips were constructed by extracting 30-frame sequences centered on each event. All tracked players within ± 0.50 seconds of the annotated frame were retained. Post-processing steps:

- **Gap filling:** Missing detections of ≤ 15 frames were filled using linear interpolation. Gaps at clip boundaries were addressed via IoU ≥ 0.50 or replication of the nearest bounding box.
- **Smoothing:** Applied an exponential moving average (EMA) to player trajectories.

- **Standardization:** Resized all bounding boxes to 224×224 pixels.

Negative (non-contact) clips were randomly sampled from segments without annotated contacts. For each event clip, ten non-event clips were selected to reflect the natural class imbalance.

2.3.6 Dataset Composition

The manual annotation process yielded 1,467 unique contact events. After post-processing, 1,609 event clips remained. Three datasets were created by pairing these with different amounts of non-contact clips (Table 2.1).

Table 2.1: Dataset composition for contact event detection.

Category	**Clips Count**
Contact Clips	1609 (from 1,467 unique events)
Non-Contact Clips (Set 1)	6000 (1:4 ratio)
Non-Contact Clips (Set 2)	12000 (1:8 ratio)
Non-Contact Clips (Set 3)	18000 (1:12 ratio)

2.3.7 Contact Event Detection Network

We trained a TSM architecture (J. Lin, Gan, and Han, 2019) with a ResNet-50 backbone pretrained on Kinetics-RGB (He et al., 2016; Kay et al., 2017). Segment counts (10, 15, 20, 30) and shift divisions (2, 4, 8) were tested, with optimal results from 15 segments and shift division = 4. Three crop scales (1.0×, 1.5×, 3.0×) were evaluated. Inverse-frequency class weighting (Buda, Maki, and Mazurowski, 2018) was used; focal loss, attention-in-attention, and hard negative mining were excluded due to inconsistent gains (Hao et al., 2022; T.-Y. Lin et al., 2017; Shrivastava, Gupta, and Girshick, 2016).

Training ran for up to 40 epochs with early stopping (patience = 10). Hyperparameters: learning rate = 0.01 (decayed at epochs 15, 30), momentum = 0.90, gradient clipping = 5, dropout = 0.50, batch size = 8, and 85/15 training-validation split.

2.3.8 Evaluation Metrics

We report precision, recall, F1-score, specificity, and MCC (Chicco and Jurman, 2020). PR-AUC was used for performance under class imbalance. Figure 2.1 shows the equations for these metrics.

$$\textbf{Precision} = \frac{TP}{TP + FP}$$

$$\textbf{Recall} = \frac{TP}{TP + FN}$$

$$\mathbf{F_1} = \frac{2\,(\text{Precision} \cdot \text{Recall})}{\text{Precision} + \text{Recall}}$$

$$\textbf{Specificity} = \frac{TN}{TN + FP}$$

$$\text{MCC} = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

Figure 2.1: Equations for key evaluation metrics: Precision, Recall, F_1, Specificity, and MCC.

2.3.9 Full-Frame Baseline Comparison

We compared the player-focused pipeline with a full-frame baseline that bypassed detection and tracking. Entire frames (resized to 600 pixels) and Recurrent All-Pairs Field Transforms (RAFT) optical-flow fields (Teed and Deng, 2020) were passed to a TSM network. Performance was poor (PR-AUC = 0.033, precision = 0.04, recall = 0.40, F1 = 0.07, specificity = 0.78, MCC = 0.06), confirming the necessity of player-focused cropping for reliable detection.

2.4 Results

2.4.1 Player Detection and Tracking

The youth-specific YOLOv8 model demonstrated high player detection performance across diverse age groups and visual environments. Evaluation metrics for the detector yielded a precision of 0.96, recall of 0.97, AP0.7 of 0.94, mAP50 of 0.97, and mAP50–95 of 0.82. These results confirm that the fine-tuned model reliably identifies players across standard youth hockey conditions, including variable lighting and overlapping players.

Tracking performance was evaluated on a benchmark set of 100 short clips characterized by high occlusion and rapid player motion. The integration of an IoU-based cost term into StrongSORT reduced ID switches from 172 to 53 (a 69% improvement) while improving MOTA from 89.0% to 94.5%. These gains underscore the effectiveness of the adapted tracker in preserving player identity over time in complex scenes.

2.4.2 Effect of Class Imbalance on Contact Event Detection Performance

Model performance varied with the class distribution in the training datasets. Table 2.2 presents precision, recall, F1-score, specificity, and MCC across three datasets: 1:4, 1:8, and 1:12 event-to-non-event clip ratios. Increasing the volume of non-event clips improved overall performance. The model trained on the 1:12 dataset achieved the best results: precision = 0.96, recall = 0.94, F1-score = 0.95, specificity ≈1.00, and MCC = 0.95. These findings suggest that higher class imbalance (with more negatives) enhances the model's ability to reject non-contact clips without sacrificing sensitivity.

Table 2.2: Contact event detection performance across three dataset versions with increasing event-to-non-event ratios (1:4, 1:8, and 1:12).

Dataset Version	Precision	Recall	F1-score	Specificity	MCC
Set 1 (1:4)	0.89	0.93	0.91	0.97	0.89
Set 2 (1:8)	0.94	0.89	0.91	0.99	0.90
Set 3 (1:12)	0.96	0.94	0.95	∼1.00	0.95

2.4.3 Effect of Bounding-Box Scale on Contact Event Detection Performance

Three bounding-box crop scales, 1.0×, 1.5×, and 3.0× were tested to assess the effect of spatial context on contact event classification (Table 2.3). The 1.5× scale yielded the highest overall performance with an F1-score and MCC of 0.95. The 1.0× crop showed lower performance (F1 = 0.87, MCC = 0.86), indicating insufficient context. Conversely, the 3.0× crop increased recall but reduced precision, likely due to background clutter. These results suggest that moderate contextual padding provides sufficient surrounding information to improve classification without introducing excessive background noise.

Table 2.3: Contact event detection performance across different bounding-box crop scales.

Scale Factor	Precision	Recall	F1-score	Specificity	MCC
1×	0.86	0.89	0.87	0.98	0.86
1.5×	0.96	0.94	0.95	∼1.00	0.95
3×	0.87	0.94	0.90	0.98	0.89

Confusion matrices for each scale confirm these trends, with the 1.5× configuration balancing true positive and false positive rates most effectively (Figure 2.2).

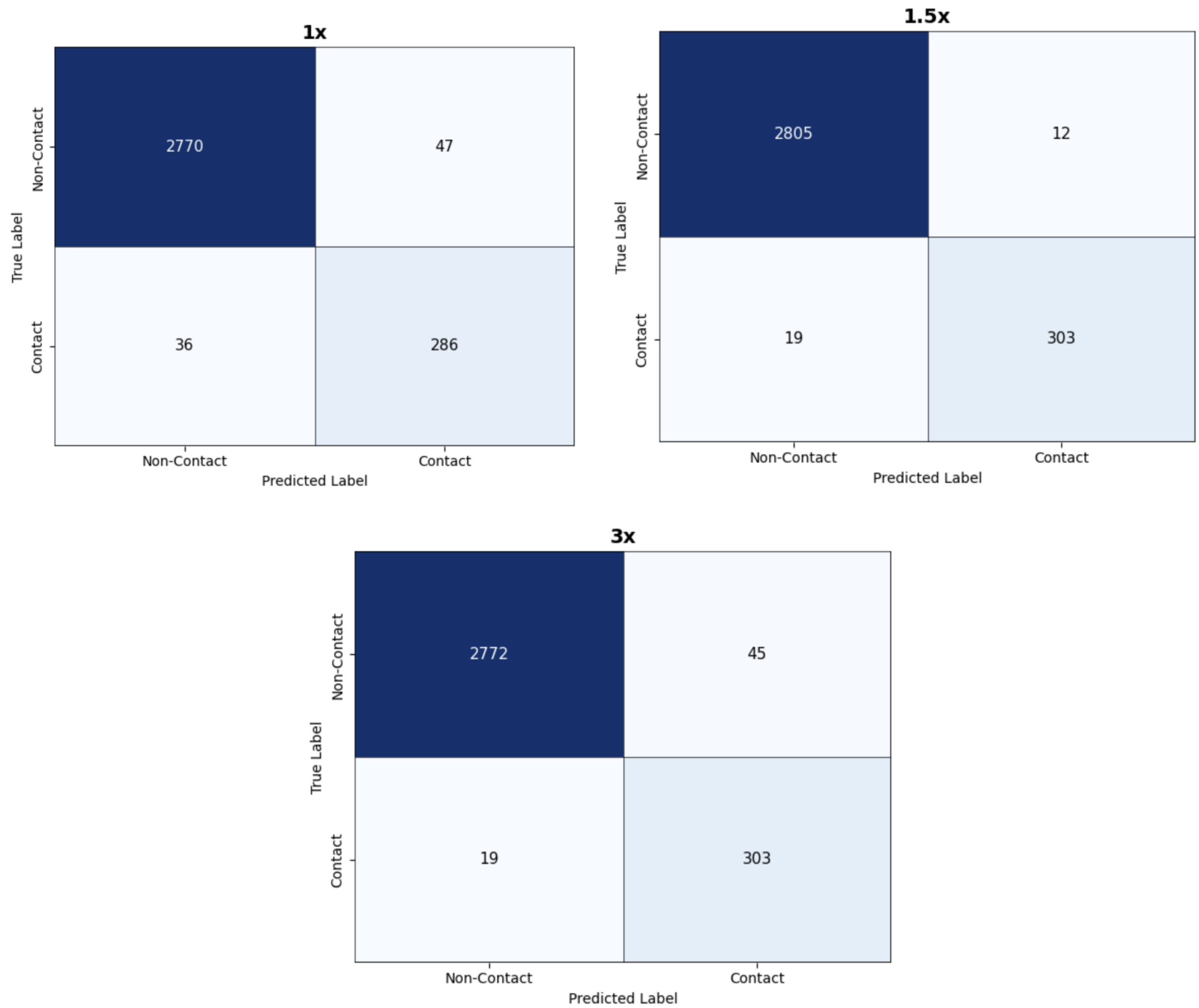
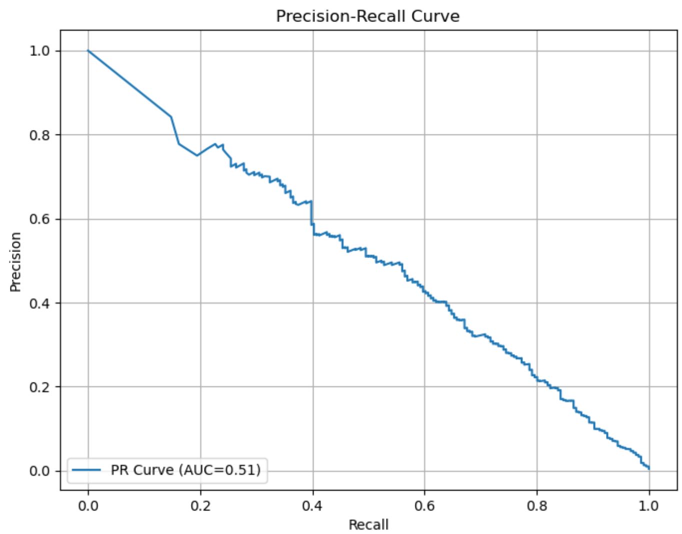


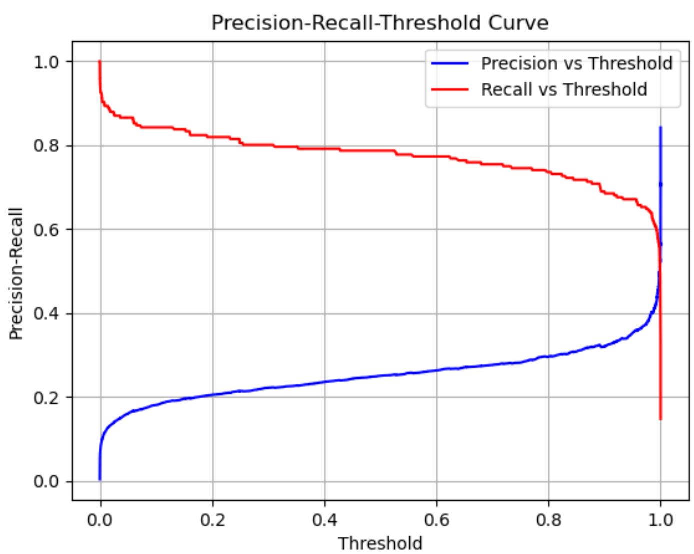
Figure 2.2: Confusion matrices for contact event detection at three bounding-box crop scales: 1×, 1.5×, and 3×.

2.4.4 Evaluation on Full-Game Videos

The final model, TSM with 15 segments, shift division 4, 1.5× crop, and 1:12 training ratio, was evaluated on two full-length U13 games. The model achieved a PR-AUC of 0.51 (Figure 2.3a). The precision–recall-versus-threshold curve (Figure 2.3b) illustrates typical trade-offs, with precision rising and recall falling as the threshold increases. At the default threshold (0.50), the confusion matrix (Figure 2.4) showed: 170 true positives, 46 false negatives, 507 false positives, and 49,193 true negatives. This corresponds to a precision of 0.25, recall of 0.79, F1-score of 0.38, specificity of 0.99, and MCC of 0.44. These results provide a realistic baseline for full-game inference and highlight remaining challenges in reducing false positives under real-world conditions.



(a) Precision–recall curve for the best model on two unseen games (PR-AUC = 0.51).



(b) Precision (red) and recall (blue) versus decision threshold, showing the typical trade-off.

Figure 2.3: (a) Precision–recall curve and (b) precision/recall versus threshold for the best model on two unseen games.

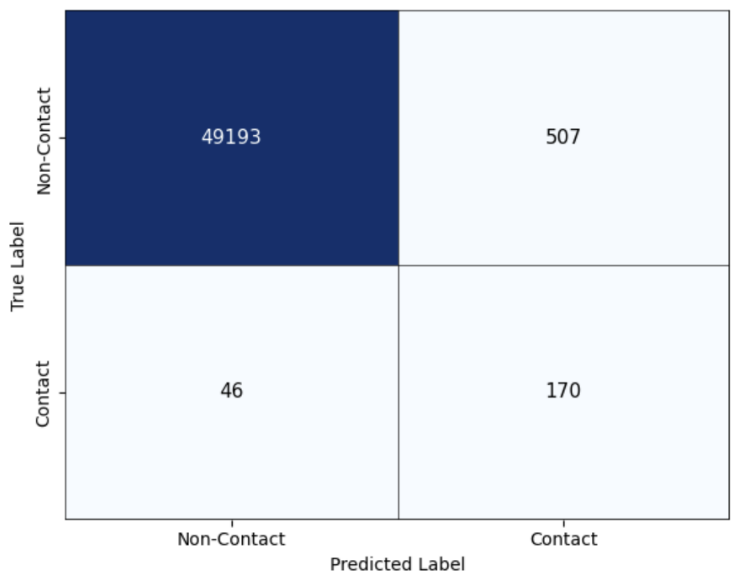


Figure 2.4: Confusion matrix for the full-game evaluation at the default 0.50 decision threshold.

2.4.5 Head Impact Detection Sensitivity

To assess the utility of the system as a filtering mechanism for head impact classification, we manually annotated all head impacts in the full-game test set. Out of 22 observed head impacts, 19 were successfully detected by the model, corresponding to a recall of 86.4%. This supports the system's applicability as a pre-filtering tool, capturing the majority of relevant contact events while reducing the burden of manual review.

2.5 Discussion

2.5.1 Bounding-Box Scale and Spatial Context

Our results highlight the critical influence of spatial context on contact event classification performance. Bounding-box size directly affected the system's ability to discriminate between true contact events and non-contact gameplay. With the smallest crop (1×), the model suffered from 47 false positives and 36 false negatives, suggesting that essential contextual cues were excluded, impairing the model's recognition of both contact and non-contact actions.

While expanding the crop to 3× reduced false negatives to 19, it failed to reduce false positives ($n = 45$), implying that excess background information can introduce visual noise and reduce precision. The optimal 1.5× crop provided the best balance, with only 12 false positives and 19 false negatives. This scale appears to include enough surrounding body motion and scene context to improve contact event recognition while avoiding the clutter that leads to misclassification. These findings emphasize the importance of carefully tuning spatial input size to optimize deep learning performance in action recognition tasks.

2.5.2 Threshold Selection and Trade-Offs

Threshold selection influences the model's precision and recall balance. In practical applications where missing potential contact events is a greater concern than over-flagging, operating at a lower threshold is advantageous. As shown in the evaluation, recall remained above 0.90 at thresholds below $\sim$0.20, albeit with increased false positives. In surveillance or pre-filtering applications, this trade-off is acceptable: false positives can be manually reviewed, whereas missed events may result in underreported injuries.

2.5.3 Error Analysis and Detection Robustness

A close examination of the 46 false negatives revealed that many were not true errors. Of these, 18 were duplicate tracklets of 9 unique contacts. Removing duplicates reduced the effective FN count to 37. Among the remaining, 10 were instances where the true contact was correctly detected via another player's tracklet, suggesting the model successfully localized the primary collision target. This leaves 27 true misses.

Of these 27 clips, six involved partial occlusion, a well-known challenge in CV. The remaining 21 involved low-intensity or brief contacts that resembled routine gameplay. These results indicate that the system is particularly effective at detecting moderate-to-high intensity events but may miss ambiguous or visually subtle incidents.

An additional limitation arises from the differences in clip generation between training and testing. Training clips were centered on the contact frame, whereas test clips were segmented at uniform 1-second intervals, potentially positioning the impact near the clip's edge. Because TSM models reduce temporal resolution via frame subsampling, off-center impacts may not receive sufficient representation. Overlapping clip generation during inference may mitigate this, though it increases computational costs (Lu et al., 2018).

Occlusion-related errors could also be reduced with multi-view camera setups, which would provide complementary perspectives and enhance visibility.

2.5.4 Operational Efficiency and Practical Utility

Despite these limitations, the system offers significant efficiency gains. On average, 85 true positives and 285 false positives were identified per game, meaning approximately 370 clips require manual review. At one second per clip, this equates to under 30 minutes of review time per game, compared to the 3 hours typically required for full manual annotation. This improvement supports the system's feasibility for real-world use in sports injury surveillance and dataset generation.

2.5.5 Head Impact Detection Reliability

The model successfully detected 19 of 22 manually annotated head impacts (86.4%), validating its application as a pre-filter for dedicated head injury classifiers. The three missed impacts illustrate important edge cases: (a) a partially occluded collision, (b) a soft contact during a face-off, and (c) a clear impact in a visually cluttered scene near the goal.

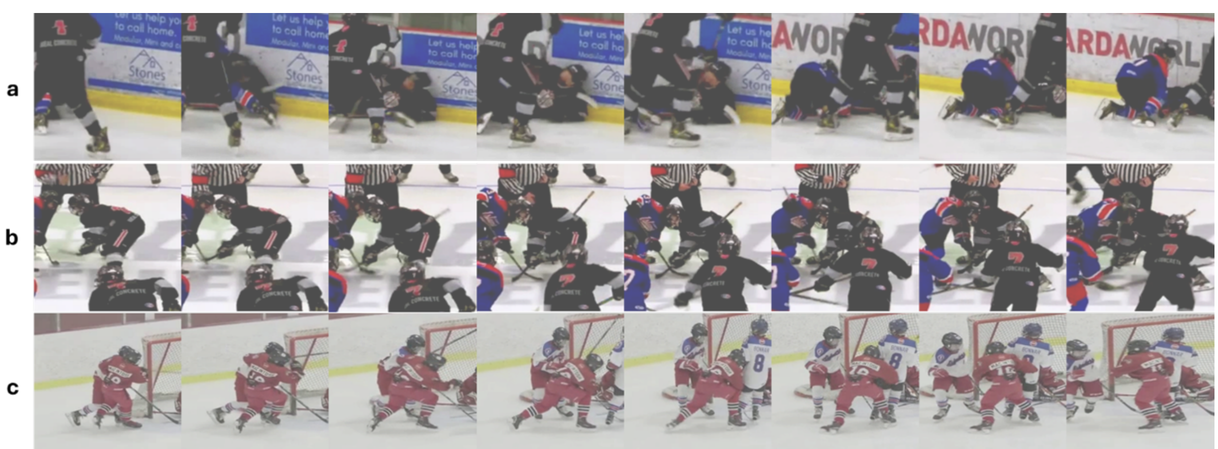


Figure 2.5: The three missed head impacts in the event detection stage: (a) a partially occluded impact, (b) a soft head contact during a face-off, and (c) a clear head impact in a cluttered scene near the goal.

Predicted probabilities for detected head impacts were high (mean ≈ 0.97), while missed head impacts had near-zero predicted probabilities (≤ 0.0043). This suggests that the model reliably detects clear head impacts but struggles with occluded or subtle events, pointing to key areas for future dataset and model improvements.

2.5.6 Future Directions in Architecture and Input Design

While the TSM architecture efficiently models temporal patterns, its reliance on a ResNet backbone may limit its sensitivity to fine-grained visual features. In ambiguous or low-motion situations, more expressive models may be needed. Transformer-based models such as the Vision Transformer (ViT) (Dosovitskiy et al., 2021), Swin Transformer (Ze Liu et al., 2021), or ConvNeXt (Zhuang Liu et al., 2022) offer promising alternatives. These architectures have shown improved performance in action recognition due to their ability to capture long-range dependencies and subtle spatial cues.

2.5.7 Limitations and Future Work

A limitation of this study is the variability in ground-truth annotations. Differences in interpretation among multiple video analysts may have introduced inconsistencies in what was labeled as a contact. This issue is particularly important when distinguishing between soft and forceful contacts, as analysts may have applied different thresholds for what constitutes a forceful contact. Such variability can affect model learning, especially when minor contacts are inconsistently labeled. Additionally, many of the labeled contacts in the current dataset are unlikely to result in injury. Focusing future datasets on more severe, clinically relevant contacts would better align the model's predictions with meaningful outcomes and reduce false positives from minor, non-injurious contacts.

2.6 Conclusion

This study introduces a fully automated, video-based pipeline for detecting physical contacts in youth ice hockey. By integrating a youth-specific YOLOv8 detector, IoU-enhanced StrongSORT tracker, and TSM classifier, the system achieved high precision (0.96) and recall (0.94), while reducing manual review time from over three hours to under 30 minutes per game. Key optimizations, such as using a 1.5× crop and incorporating a large number of non-contact clips, contributed to strong model performance (F1-score and MCC = 0.95).

Overall, the proposed system provides a scalable solution for rapid head impact dataset creation and objective collision exposure tracking. It offers a practical tool for researchers, coaches, and clinicians to enhance injury surveillance and promote safer youth hockey environments.

References

Allard, P. et al. (2022). "In-Season Session Training Load Relative to Match Load in Professional Ice Hockey". In: *The Journal of Strength & Conditioning Research* 36.2, pp. 486–492.

Azadi, Amir et al. (2024). *A Deep Learning Network for Detecting Head Impacts in Ice Hockey from 2D Game Video*. Manuscript.

Bernardin, Keni and Rainer Stiefelhagen (2008). "Evaluating Multiple Object Tracking Performance: The CLEAR MOT Metrics". In: *EURASIP Journal on Image and Video Processing* 2008, pp. 1–10.

Buda, Mateusz, Atsutoshi Maki, and Maciej A. Mazurowski (2018). "A Systematic Study of the Class Imbalance Problem in Convolutional Neural Networks". In: *Neural Networks* 106, pp. 249–259.

Butterfield, J. et al. (2023). "A Video Analysis Examination of the Frequency and Type of Head Impacts for Player Positions in Youth Ice Hockey and FE Estimation of Their Impact Severity". In: *Sports Biomechanics*, pp. 1–17.

Chicco, Davide and Giuseppe Jurman (2020). "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation". In: *BMC Genomics* 21.1, p. 6.

Daneshvar, David H. et al. (2011). "The Epidemiology of Sport-Related Concussion". In: *Clinics in Sports Medicine* 30.1, pp. 1–17. DOI: 10.1016/j.csm.2010.08.006.

Dehghan, Parisa et al. (2024). "Object Detection for Ice Surface Localisation in Youth Hockey". In: *Proceedings (venue TBA)*.

Dosovitskiy, Alexey et al. (2021). *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*. arXiv:2010.11929.

Du, Yihong et al. (2023). "StrongSORT: Make DeepSORT Great Again". In: *IEEE Transactions on Multimedia* 25, pp. 8725–8737.

Echemendia, Ruben J. et al. (2017). "The Sport Concussion Assessment Tool 5th Edition (SCAT5): Background and Rationale". In: *British Journal of Sports Medicine* 51.11, pp. 848–850.

Gysland, Shannon M. et al. (2012). "The Relationship Between Subconcussive Impacts and Concussion History on Clinical Measures of Neurologic Function in Collegiate Football Players". In: *Annals of Biomedical Engineering* 40, pp. 14–22.

Hao, Yutong et al. (2022). "Attention in Attention: Modeling Context Correlation for Efficient Video Classification". In: *IEEE Transactions on Circuits and Systems for Video Technology* 32.10, pp. 7120–7132.

He, Kaiming et al. (2016). "Deep Residual Learning for Image Recognition". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778.

Jocher, Glenn et al. (2023). *YOLOv8: The Next Generation of YOLO*. Ultralytics Documentation. URL: https://docs.ultralytics.com.

Johnson, L. S. M. (2011). "Concussion in Youth Ice Hockey: It's Time to Break the Cycle". In: *Canadian Medical Association Journal* 183.8, pp. 921–924. DOI: 10.1503/cmaj.110282.

Karton, Clara and T. Blaine Hoshizaki (2018). "Concussive and Subconcussive Brain Trauma: The Complexity of Impact Biomechanics and Injury Risk in Contact Sport". In: *Handbook of Clinical Neurology*, pp. 39–49. DOI: 10.1016/B978-0-444-63954-7.00005-7.

Kay, Will et al. (2017). *The Kinetics Human Action Video Dataset*. arXiv:1705.06950.

Lin, Ji, Chuang Gan, and Song Han (2019). "TSM: Temporal Shift Module for Efficient Video Understanding". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 7083–7093.

Lin, Tsung-Yi et al. (2017). "Focal Loss for Dense Object Detection". In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2980–2988.

Liu, Ze et al. (2021). "Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 10012–10022.

Liu, Zhuang et al. (2022). "A ConvNet for the 2020s". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11976–11986.

Lu, Nannan et al. (2018). "Deep Learning for Fall Detection: Three-Dimensional CNN Combined with LSTM on Video Kinematic Data". In: *IEEE Journal of Biomedical and Health Informatics* 23.1, pp. 314–323.

McKee, Ann C. et al. (2013). "The Spectrum of Disease in Chronic Traumatic Encephalopathy". In: *Brain* 136.1, pp. 43–64.

Mihalik, Jason P. et al. (2012). "Head Impact Biomechanics in Youth Hockey: Comparisons Across Playing Position, Event Types, and Impact Locations". In: *Annals of Biomedical Engineering* 40, pp. 141–149.

Naik, Bhushan T., M. F. Hashmi, and Neeraj D. Bokde (2022). "A Comprehensive Review of Computer Vision in Sports: Open Issues, Future Trends and Research Directions". In: *Applied Sciences* 12.9, p. 4429.

Nowinski, Christopher J. et al. (2024). "'Subconcussive' is a Dangerous Misnomer: Hits of Greater Magnitude Than Concussive Impacts May Not Cause Symptoms". In: *British Journal of Sports Medicine* 58.14, pp. 754–756.

Pfister, Thomas et al. (2016). "The Incidence of Concussion in Youth Sports: A Systematic Review and Meta-Analysis". In: *British Journal of Sports Medicine* 50.5, pp. 292–297.

Post, Andrew and T. Blaine Hoshizaki (2012). "Mechanisms of Brain Impact Injuries and Their Prediction: A Review". In: *Trauma* 14.4, pp. 327–349. DOI: 10.1177/1460408612446573.

Rezaei, Amir and Lawrence C. Wu (2022). "Automated Soccer Head Impact Exposure Tracking Using Video and Deep Learning". In: *Scientific Reports* 12.1, p. 9282.

Shill, Isaac J. et al. (2023). "How to Harness and Improve on Video Analysis for Youth Rugby Player Safety: A Narrative Review". In: *BMJ Open Sport & Exercise Medicine* 9.3, e001645.

Shrivastava, Abhinav, Abhinav Gupta, and Ross Girshick (2016). "Training Region-Based Object Detectors with Online Hard Example Mining". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 761–769.

Teed, Zachary and Jia Deng (2020). "RAFT: Recurrent All-Pairs Field Transforms for Optical Flow". In: *European Conference on Computer Vision (ECCV)*, pp. 402–419.

USA Hockey (2024). *2023-24 Membership Statistics*. https://www.usahockey.com/membershipstats. Retrieved from USA Hockey website.

Wang, Limin et al. (2016). "Temporal Segment Networks: Towards Good Practices for Deep Action Recognition". In: *European Conference on Computer Vision*, pp. 20–36.

Wojke, Nicolai, Alex Bewley, and Dietrich Paulus (2017). "Simple Online and Realtime Tracking with a Deep Association Metric". In: *2017 IEEE International Conference on Image Processing (ICIP)*, pp. 3645–3649.

Wu, Fan et al. (2022). "A Survey on Video Action Recognition in Sports: Datasets, Methods and Applications". In: *IEEE Transactions on Multimedia* 25, pp. 7943–7966.

Chapter 3

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

International Research Council on the Biomechanics of Injury (IRCOBI 2024) — Published

Amir Azadi[a,*], Parisa Dehghan[a], Ryan B. Graham[a], T. Blaine Hoshizaki[a]

[a] School of Human Kinetics, University of Ottawa, Ottawa, Ontario, Canada

3.1 Abstract

This paper presents a deep learning approach for detecting head impacts in ice hockey from 2D game videos. Traumatic brain injuries, including concussions and RHI, pose significant health risks in sports. Understanding the relationship between head impact features (magnitude and frequency) and outcomes such as mental health decline, cognitive deficits, and Chronic Traumatic Encephalopathy (CTE) requires extensive datasets. Tracking head impacts during games is challenging, and available tools are impractical for most leagues due to cost and equipment constraints. Utilising game videos for head impact detection offers a viable solution. A methodology combining player detection and tracking with a LRCN for head impact detection is proposed. Our player detection model achieved high precision and recall scores, facilitating accurate tracking. The YOLOv8x object detection model yielded precision, recall, mAP50, and mAP50–95 scores of 0.97, 0.97, 0.99, and 0.95, respectively. The StrongSORT tracking algorithm used for player tracking minimised ID switches, important for precise tracking in dynamic sports environments. The LRCN-based head impact detection model showed promising results, with an accuracy of 0.87 and a loss of 0.04. Future work involves refining dataset creation to address data imbalance and exploring alternative deep learning models like convolutional long short-term memory (ConvLSTM) and 3D convolutional neural network (CNN) for improved performance.

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

3.2 Introduction

Scientists continue to undertake research to better understand the complex relationship between head impact severity and brain injuries, particularly TBI such as concussions in sport (Casper, 2018). TBI is a heterogeneous injury resulting from various physical forces, leading to a spectrum of neurological and psychiatric symptoms. Concussion, despite its mTBI classification, can have severe consequences if not treated properly or if individuals experience repeated mTBI. Unlike more severe forms of TBI, mTBI often does not result in any visible brain abnormalities like bleeding. Instead, they manifest through functional disruptions, leading to symptoms such as dizziness, nausea, and cognitive disturbances, which can lead to a wide range of neurological and psychiatric symptoms, both acute and chronic (Karton, Blaine Hoshizaki, and Gilchrist, 2020; Karton and T Blaine Hoshizaki, 2018; Post and T Blaine Hoshizaki, 2012). Annually, over 500,000 adolescents in the United States alone experience mTBI, with many of these injuries subclinical and may contribute to accumulated brain damage from RHI, and risk for neurodegenerative disorders (Nowinski et al., 2024). RHI refers to the repeated exposure to head impacts (clinical and subclinical) over time, often experienced by athletes participating in contact sports like American football and ice hockey. Head impact magnitude and frequency play important roles in understanding and assessing the potential risks associated with brain trauma. Factors that affect the magnitude of impact include the velocity, head location, and the type of impact event (e.g., shoulder collision, fall to ice). Impact frequency can lead to cumulative effects on the brain, emphasising the importance of understanding both single and repeated head impacts on athletes' health (Karton, Blaine Hoshizaki, and Gilchrist, 2020; Karton and T Blaine Hoshizaki, 2018).

The need to understand the link between brain trauma, especially repetitive traumatic brain injury, and neurological health has led to the emphasis on large, validated datasets (B. Hoshizaki et al., 2013; Agoston and Langford, 2017). Objectively capturing both impact magnitude and frequency are challenging and the current available methods using head impact sensors or laboratory reconstructions both present with characteristics that limit their ability to create datasets large enough to disentangle intricate relationships (Liu et al., 2020; Daneshvar et al., 2011; Kuo et al., 2018; Larsen, 2022; W. Chen et al., 2023; B. Hoshizaki et al., 2013; Karton and Thomas Blaine Hoshizaki, 2021; Karton, Post, et al., 2021). With the advent of big data and the integration of AI in medical research, new possibilities have emerged for investigating these complex relationships in extensive, heterogeneous datasets (Rodger, 2015; Yurgil et al., 2014). CV, a branch of AI, enables machines to process and interpret visual data. Its utility in sports includes several applications including game analytics (Rodger, 2015; Yurgil et al., 2014; Duong et al., 2005; Blank et al., 2005; Ke, Sukthankar, and Hebert, 2007; Yan Du, F. Chen, and Xu, 2007). The application of CV in sports has led to most sports leagues and organizations filming their games for all age groups and competition levels leading to the wide accessibility of game footage. Automated methods using CV for measuring brain trauma and track head impacts is necessary for creating large data sets. In this regard, deep learning algorithms are key to automating the detection of head impacts from video footage.

Human action recognition (HAR) in sports is an important technological application that employs CV techniques to identify and understand player actions. This process is invaluable across various sporting events, including warm-ups, specific training, or competitions, as it assists coaches, medical staff, and journalists by enhancing performance analysis, preventing injuries, and accumulating match statistics (Host and Ivašić-Kos, 2022). The potential of HAR extends over a range of sports, from individual activities like skiing and swimming to team sports such as ice hockey, basketball, soccer, and volleyball (Host and Ivašić-Kos, 2022). One of the applications of HAR is in detecting a head impact event

(Rezaei and Wu, 2022). Recognising an event or action from 2D video is challenging due to the variations in pixel intensity, camera views, and complex patterns. Players might change their position, block each other from the camera, enter and exit the camera's field of vision, and operate at different angles (Host and Ivašić-Kos, 2022). These impacts in sports are brief and associated with distinct pre-impact reactions, player positions during the impact, and immediate post-impact reactions. Consequently, traditional methods of HAR often fail to capture the features due to differing illumination, scale, posture, perspectives, and complex human body movements during an activity like head impact (Rezaei and Wu, 2022; Ivašić-Kos, Host, and Pobar, 2021; Lu et al., 2018; Ngo, Ma, and Zhang, 2005). To address some of these challenges it would be advantageous to incorporate models that provide both spatial and temporal information to a HAR application of identifying head impacts from game footage. Spatial features are related to the visual attributes and pixel arrangements within an image, while temporal features pertain to time-based characteristics and changes over a sequence of frames.

Various deep learning architectures have been proposed to recognise human action from video. LRCN combine CNN with long short-term memory network (LSTM) for sequence learning tasks (Donahue et al., 2015). However, they can be computationally intensive and prone to overfitting (Pascanu, Mikolov, and Bengio, 2013). ConvLSTM use convolutional operations within LSTM units, making them suitable for spatial-temporal data but leading to increased computational complexity and larger model sizes (Shi et al., 2015). Temporal segment network (TSN) segment videos and sample brief clips, however they have been shown to miss critical actions and struggle to integrate temporal information effectively (Wang et al., 2018). The TSM optimises temporal modelling by shifting channels in a convolutional network but can have limitations in capturing longer temporal dependencies (Rezaei and Wu, 2022; Lin, Gan, and Han, 2019). Additionally, pose estimation enhances human activity recognition by providing skeletal representations. However, given the nature of ice hockey, where players frequently obscure each other from the camera's view and head impacts occur within the brief span of 1 or 2 frames, pose estimation is often ineffective in many parts of the game (Fang et al., 2017; Fani et al., 2017; Neher, 2018).

The purpose of this study was to develop, train and validate an activity detection network to detect head impacts from 2D videos in ice hockey games. This was accomplished by employing spatial and temporal techniques in CV. These sophisticated algorithms can be used to create big datasets to facilitate research to investigate the relationship between impact forces and brain injuries (B. Hoshizaki et al., 2013; Agoston and Langford, 2017). Additionally, this method could be used to improve the management of brain injuries in sports by enabling real-time monitoring and assessment (Rezaei and Wu, 2022).

3.3 Methods

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

3.3.1 Player Detection

To detect players in the visual field, methods in object detection were used to identify players on the ice. Specifically, the YOLOv8x model, the latest version of the YOLO (You Only Look Once) object detection algorithm, was compiled using the Adam optimiser and categorical cross-entropy loss function (Kingma and Ba, 2014; Mao, Mohri, and Zhong, 2023). This model was chosen as it treats object detection as a single regression problem, allowing it to make predictions with a single network evaluation, which significantly increases its speed compared to methods that require multiple evaluations for different image regions (Redmon et al., 2016). The dataset employed in this task consisted of 80,000 annotated images extracted from 25 National Hockey League (NHL) game videos (Prakash et al., n.d.). The dataset was split into a training set comprising 65,000 images and a validation set of 15,000 images. The validation set was solely utilised to evaluate the performance of the player detection model, and the output of the network on this dataset determined the results. Due to the frequent occlusion events in ice hockey, the object detection model was trained with multiple labels, including player, goal keeper, and referee, along with their corresponding jersey colour and occlusion statuses (either occluded or not-occluded), providing essential information for subsequent tracking analysis (details of label selection will be discussed in the tracking section). Annotating the images was accomplished by drawing bounding boxes around each object and assigning its appropriate label. The YOLOv8x object detection framework underwent training for 250 epochs, and the image resolution was 1280×720 pixels. The performance of player detection was evaluated using precision, recall and mAP of 50 and 50–90. The mAP was used to calculate the average precision (AP) for each class and then take the mean over all classes. 50–90 mAP refers to the mAP calculated at different IoU thresholds, specifically at 50% and 90%. An IoU, a commonly used metric in CV, calculated the overlap between a predicted bounding box and a ground truth bounding box. If there is no overlap between two bounding boxes, the IoU cost is 1, and if the two bounding boxes have 100% overlap, the IoU cost is 0.

3.3.2 Player Tracking

Our tracking algorithms employed the non-maximum suppression (NMS) method to eliminate redundant or overlapping bounding boxes of detections (Yihong Du et al., 2023). The player detection output consists of bounding boxes in each frame, some of which may overlap. This overlap can cause issues in player tracking, as the tracking algorithm may mistakenly treat overlapping boxes as the same player and delete duplicates. However, each box represents a distinct player. To address this, we propose changing the labels of the bounding boxes from a single label (Player) to different labels (P: Player, KP: Goalkeeper, Ref: Referee, OC: Occlusion, NOC: Non-Occlusion, and the jersey colour). This change allows the tracking method to treat each bounding box as a separate class, preventing the deletion of any box. The NMS process was employed to ensure that no bounding box is missed during occlusion events, thereby improving our player tracking accuracy. To continuously track the players throughout the video, the StrongSORT object tracking algorithm, that combines deep neural networks with traditional computer vision techniques, was employed in this study (Yihong Du et al., 2023). This was chosen as this algorithm improves tracking accuracy and robustness in challenging scenarios by using deep neural networks to extract features robust to appearance changes, such as illumination, viewpoint, and occlusion. Additionally, StrongSORT utilises a Kalman filter to model target object motion and estimate its position in subsequent frames (Yihong Du et al., 2023). At each frame, StrongSORT integrated both the detected players from the current frame and predictions of those players which was derived from the 2–3 previous frames. Utilising the feature embeddings of the detections and the predictions, along with considering motion costs, a cost matrix is constructed by StrongSORT (Yihong Du et al., 2023), meaning linear assignment techniques are employed to link detections to their corresponding predictions correctly. Feature embeddings are high-dimensional vectors that represent the detected objects' attributes, such as appearance, shape, and other visual characteristics. This function measured the similarity between the visual features of track bounding boxes and the current frame detection bounding boxes. However, in ice hockey, where players from each team often share similar appearances due to identical jerseys and helmets, the default cost function in StrongSORT is not robust enough to differentiate between multiple bounding boxes that are similar in appearance. For this reason, an adapted cost function using IoU was incorporated into this phase. The IoU cost was incorporated to provide additional information about the location of track and detection bounding boxes that helps distinguish between multiple similar bounding boxes, and decreasing the likelihood of ID switches between frames, especially in occlusion events. To evaluate player tracking, a series of 15 videos, each lasting 10 seconds were used, and the algorithm was run to test ID switches, referring to the number of times the tracking algorithm assigns a different ID to


(a) Whole-frame.

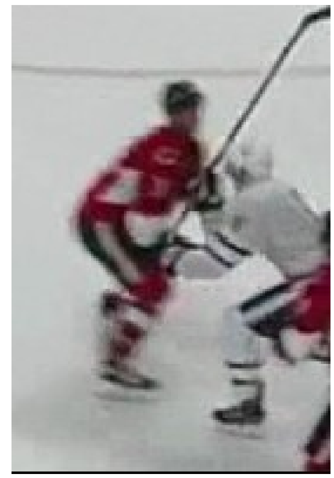

(b) Player-focused.

Figure 3.1: Sample of whole-frame (left) vs. player-focused (right) analysis.

the same object or player throughout the frames. The StrongSORT algorithm, employing its default cost function, was compared with those obtained using the adapted IoU cost function.

3.3.3 Video Segmentation

To prepare an input dataset for the activity detection network, a preprocessing phase of video segmentation was performed. A player analysis technique was chosen as it narrows its focus specifically to individual players as opposed to a whole frame analysis that uses all the pixels in the frames as input to the network (Figure 3.1). The advantage is that irrelevant pixels from other parts of the frame are excluded, increasing the accuracy of our network at identifying when the activity has occurred and specify the individual players involved (Karpathy et al., 2014). This technique was used to divide long video into smaller sets of frames of 50 frames, and analyse each set independently, allowing for a finer temporal resolution. This means that the network can more precisely determine the timing of head impacts. This approach was used to detect and track each player individually within every frame and segment. The result is that short video clips for each player during every video segment was created. Segmenting the video and using a player analysis approach provides information about which player experienced a head impact and at which frame it occurred.

3.3.4 Head Impact Dataset

A dataset was developed to train the head impact activity detection networks. Head impacts documented through video analysis from the NHL game videos from the 2016–17 season were utilised (Butterfield et al., 2023). The identified head impact events were cropped around the player area for 50 frames, ensuring that the head impact frame was the 25th frame. A total of 150 head impact events categorized by event type such as head-to-board, head-to-elbow, head-to-foot, head-to-glass, head-to-head, head-to-shoulder, head-to-puck/stick, head-to-ice, head-to-gloves, were included. To account for a small sample size in head impact events, we increased our data set using augmentation, where multiple videos can be created for each head impact event by extracting sequences of frames from various starting points. Two 25-frame videos were extracted from each video, from frames 5 to 30 and 20 to 45, with a frame size of 300×300. Additionally, each of these videos were flipped to augment the number of head impact events. For this study, 750 non-head impact videos were used. Non-head impact events were also classified into several categories, such as normal movement, impact with boards or glass, falls to the ice, body collisions, and occlusion scenes. Although some of these categories may resemble head impacts, they are not head impacts, and they served to prevent the network from overfitting to detect head impact events. To enable the network to differentiate between head and non-head impact events, our diverse data set was developed to ensure that the network would be exposed to a broad range of scenarios and conditions that occur during a game. In total, the dataset for the head impact detection was comprised of 1350 video clips, which were split into a training set and a validation set, with 75% and 25% of the clips, respectively. The validation dataset is used to assess the performance of the head impact detection dataset.

3.3.5 Head Impact Activity Detection Network

An activity detection network was utilised for the detection of head impacts within each video segment. For this study, an LRCN architecture was chosen because it combines convolutional layers for spatial feature extraction from the frames with LSTM layer(s) for temporal sequence modelling at each time step, making it well suited to capturing spatiotemporal patterns in head impact clips (Donahue et al., 2015). Although LRCN models can be computationally intensive and susceptible to overfitting, these drawbacks were mitigated through regularisation strategies such as dropout, early stopping, and class weighting. The code implements a sequential model architecture, incorporating Time-Distributed layers for applying convolutional operations and pooling across the temporal dimension of the input data. Dropout layers are included to prevent overfitting during training (Srivastava et al., 2014). Multiple convolutional layers with varying filter sizes and activation functions capture spatial features effectively, while LSTM layers enable the model to capture long-term dependencies in the data. The input data has a shape of $(\text{None}, 25, 300, 300, 3)$. None represents the batch size (which can vary during training or inference), 25 is the sequence length (number of time steps), 300 and 300 are the height and width of the input images, and the last number is the number of channels (e.g., RGB channels). The final output shape is $(\text{None}, 2)$, indicating two output classes. The model included the Adam optimiser and categorical cross-entropy loss function (Kingma and Ba, 2014; Mao, Mohri, and Zhong, 2023). Additionally, class weights were computed to address class imbalance during training, as there are more non-head impact than head impact datasets. The training loop iterated over 50 epochs, with early stopping set to 10 to prevent overfitting. Early stopping is a technique used to stop the training process if the model's performance on a separate validation dataset stops improving or starts deteriorating. The performance of the head impact activity detection neural network was evaluated using metrics of validation accuracy and validation loss. Validation accuracy was used to assess the percentage of correctly predicted labels on a separate validation dataset, which was not used for training. Validation loss was used to measure how well the model performed on the validation dataset, calculated as the difference between the predicted output and the actual target labels. A good model is indicated by both high validation accuracy and low validation loss, showing accurate and consistent predictions.

3.4 Results

3.4.1 Player Detection

After 250 epochs of training, the YOLOv8x model yielded results, with precision, recall, mAP50, and mAP50–95 scores of 0.97, 0.97, 0.99, and 0.95, respectively. These metrics reflect the model's ability to accurately identify and locate players, goal keepers and referees in the input images. A precision score of 0.97 indicates that the vast majority of predicted detections are correct (few false positives), while a recall score of 0.97 demonstrates the model's proficiency in capturing a substantial portion of the actual positive instances. The mAP50 and mAP50–95 scores of 0.99 and 0.95 signify the model's strong performance in detecting objects at varying confidence thresholds. The output of an unseen video is shown in Figure 3.3.

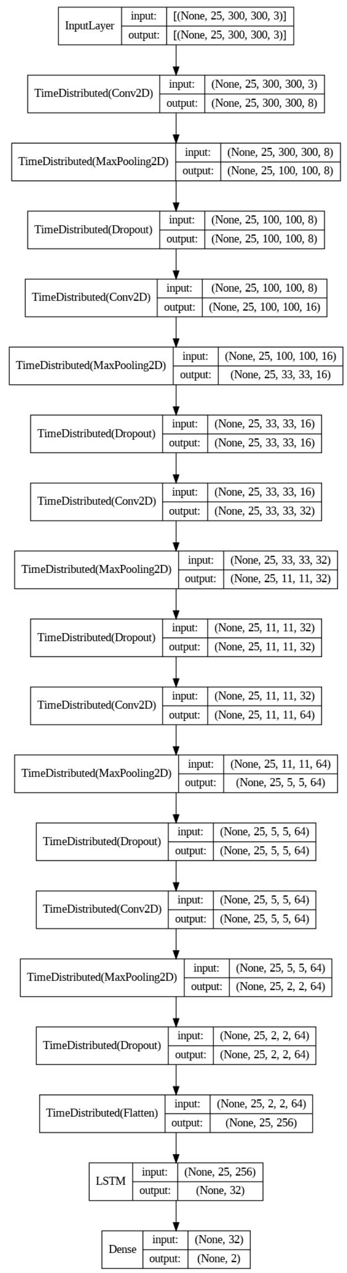


Figure 3.2: LRCN architecture for the head impact activity detection network.

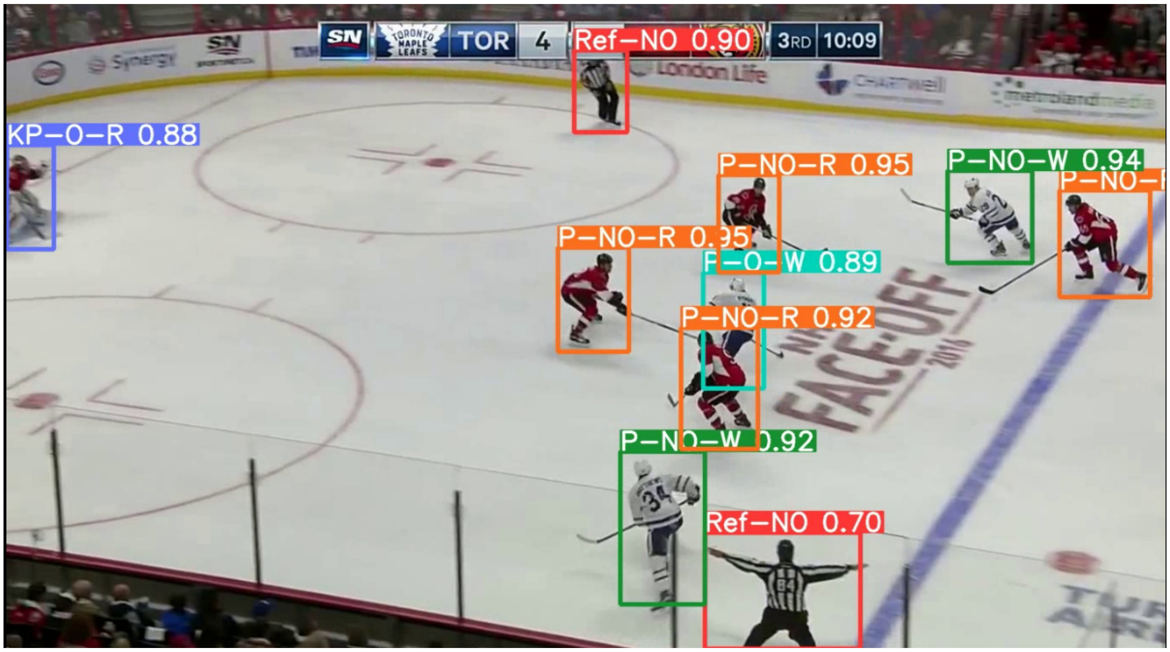


Figure 3.3: Output of the player detection network using a test video. The labels consist of three parts: the first letter corresponds to the category (P: Player, KP: Goal Keeper, Ref: Referee), the second letter indicates occlusion status (OC: Occlusion, NOC: Non-Occlusion), and the third letter denotes the jersey colour (R: Red, W: White).

3.4.2 Player Tracking

Player tracking was evaluated from a series of 15, 10-second videos, including scenarios with occlusions and head impact events. The StrongSORT algorithm, using the IoU-incorporated cost function, resulted in no ID switches observed in any of the videos, unlike the default version, which revealed an average of 9 ID switches during the tracking process. This difference emphasises the significance of our adapted cost function in mitigating ID switches, thereby contributing to the improved accuracy and consistency of the object (player) tracking system (Figure 3.4).

3.4.3 Head Impact Activity Detection Network

After 50 epochs of training, the head impact detection network achieved fair results. The validation accuracy reached 0.87, and the validation loss reached 0.04. This outcome

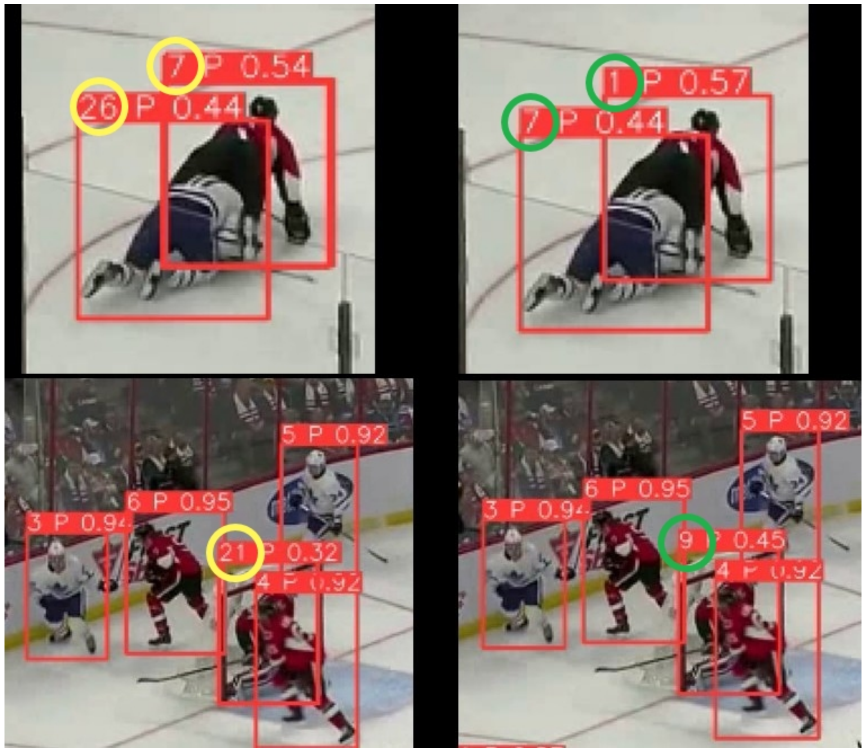


Figure 3.4: Comparison of player tracking IDs with StrongSORT. Left column: default cost function, where IDs switches/misassignments are highlighted (yellow circles). Right column: IoU-augmented cost function, which maintains consistent IDs for the same players (green circles).

highlights the model's ability to discern and classify head impact events effectively. The integration of early stopping contributes to model efficiency and prevents unnecessary computational expenditure beyond the point of optimal performance.

3.5 Discussion

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

The outcome of our player detection algorithm exhibits very good results, including high precision, recall, mAP50, and mAP50–95 scores. However, exploring alternative object detection algorithms, such as Faster R-CNN, could provide valuable insights into potential improvements in tracking accuracy. Faster R-CNN offers precise and high-quality object detection, leveraging CNN features effectively (Girshick, 2015; Ren et al., 2015). Hence, a comparative analysis of multiple object detection algorithms can offer a comprehensive understanding of their strengths and weaknesses, aiding informed decisions regarding their implementation for player detection in ice hockey games.

While results obtained in this analysis of unseen 10-second videos showed the absence of player ID switches, certain conditions pose challenges. These conditions include player substitutions, extended periods of players being out of view for over 40 frames, or specific occlusion events, which may lead to ID switches or incorrect ID assignments. To further enhance tracking performance, post-tracking ID refinement is deemed necessary in the future. Previous studies have highlighted the effectiveness of using jersey numbers for ID confirmation after tracking (Vats et al., 2023). Additionally, incorporating player trajectories and contextual information, such as team formations and player positions, have potential in providing cues for accurate ID assignment. Implementing these approaches could enhance tracking performance, minimising ID switches, and improving overall tracking accuracy.

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

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

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

3.6 Conclusions

Detecting head impacts in ice hockey from 2D game videos presents significant challenges including pixel intensity variations, camera views, and complex player movements. To address these challenges, a methodology incorporating spatial and temporal features in player detection, tracking, and head impact detection was proposed. The YOLOv8x model demonstrated high precision, recall, mAP50, and mAP50–95 scores for player detection. Object tracking using the StrongSORT algorithm with a modified cost function significantly reduced ID switches, improving player tracking accuracy. For head impact detection, LRCN architectures were used to achieve a validation accuracy of 0.87. However, challenges remain, including dataset size, model generalisation, and variability in head impact event categories, suggesting areas for future improvement. Future work will focus on exploring alternative object detection algorithms, refining tracking algorithms, and expanding the head impact detection dataset. Additionally, investigating different neural network architectures, such as ConvLSTM and 3D CNN, may further enhance head impact detection performance. Our findings describe a methodology for a head impact activity detection network, establishing a valid approach for automating the detection and tracking of head impacts from 2D videos. This initial step can be applied to a larger scaled automated system for collecting large brain trauma exposure datasets.

References

Agoston, Denes V and Dianne Langford (2017). "Big Data in traumatic brain injury; promise and challenges". In: *Concussion* 2.4, CNC44.

Blank, Moshe et al. (2005). "Actions as space-time shapes". In: *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*. Vol. 2. IEEE, pp. 1395–1402.

Butterfield, John et al. (2023). "A video analysis examination of the frequency and type of head impacts for player positions in youth ice hockey and FE estimation of their impact severity". In: *Sports Biomechanics*, pp. 1–17.

Casper, Stephen T (2018). "Concussion: A History of Science and Medicine, 1870-2005". In: *Headache: The Journal of Head and Face Pain* 58.6, pp. 795–810.

Chen, Wesley et al. (2023). "A comparison of frequency and magnitude of head impacts between Pee Wee and Bantam youth ice hockey". In: *Sports biomechanics* 22.6, pp. 728–751.

Daneshvar, David H. et al. (2011). "Helmets and mouth guards: the role of personal equipment in preventing sport-related concussions". In: *Clinics in Sports Medicine* 30.1, pp. 145–163.

Donahue, Jeff et al. (2015). "Long-term recurrent convolutional networks for visual recognition and description". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2625–2634.

Du, Yan, Feng Chen, and Wei Xu (2007). "Human interaction representation and recognition through motion decomposition". In: *IEEE Signal Processing Letters* 14.12, pp. 952–955.

Du, Yihong et al. (2023). "StrongSORT: Make DeepSORT Great Again". In: *IEEE Transactions on Multimedia*.

Duong, Tuan V. et al. (2005). "Activity recognition and abnormality detection with the switching hidden semi-Markov model". In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. Vol. 1. IEEE, pp. 838–845.

Fang, Haoshuo et al. (2017). "RMPE: Regional Multi-Person Pose Estimation". In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2334–2343.

Fani, Mahsa et al. (2017). "Hockey action recognition via integrated stacked hourglass network". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 29–37.

Girshick, Ross (2015). "Fast R-CNN". In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1440–1448.

Hoshizaki, Blaine et al. (2013). "The relationship between head impact characteristics and brain trauma". In: *J Neurol Neurophysiol* 5.1, pp. 1–8.

Host, Kristina and Marina Ivašić-Kos (2022). "An overview of Human Action Recognition in sports based on Computer Vision". In: *Heliyon* 8.6, e09756.

Ivašić-Kos, Marina, Kristina Host, and Miran Pobar (2021). "Application of deep learning methods for detection and tracking of players". In: *Deep Learning Applications*. IntechOpen.

Karpathy, Andrej et al. (2014). "Large-scale video classification with convolutional neural networks". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1725–1732.

Karton, Clara, T Blaine Hoshizaki, and Michael D Gilchrist (2020). "A novel repetitive head impact exposure measurement tool differentiates player position in National Football League". In: *Scientific reports* 10.1, p. 1200.

Karton, Clara and T Blaine Hoshizaki (2018). "Concussive and subconcussive brain trauma: the complexity of impact biomechanics and injury risk in contact sport". In: *Handbook of clinical neurology* 158, pp. 39–49.

Karton, Clara and Thomas Blaine Hoshizaki (2021). "Biomechanics of sport-related neurological injury". In: *Clinics in Sports Medicine* 40.1, pp. 19–38.

Karton, Clara, Andrew Post, et al. (2021). "Exposure to brain trauma in six age divisions of minor ice hockey". In: *Journal of biomechanics* 116, p. 110203.

Ke, Yan, Rahul Sukthankar, and Martial Hebert (2007). "Spatio-temporal shape and flow correlation for action recognition". In: *2007 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, pp. 1–8.

Kingma, Diederik P. and Jimmy Ba (2014). "Adam: A Method for Stochastic Optimization". In: *arXiv preprint arXiv:1412.6980*.

Kuo, Calvin et al. (2018). "Comparison of video-based and sensor-based head impact exposure". In: *PLOS ONE* 13.6, e0199238.

Larsen, Kyle (2022). "Comparison of Brain Strain Magnitudes Calculated Using Head Tracking Impact Parameters and Body Tracking Impact Parameters Obtained from 2D Video". PhD thesis. Université d'Ottawa / University of Ottawa.

Lin, Ji, Chuang Gan, and Song Han (2019). "Tsm: Temporal shift module for efficient video understanding". In: *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 7083–7093.

Liu, Yuzhe et al. (2020). "Validation and comparison of instrumented mouthguards for measuring head kinematics and assessing brain deformation in football impacts". In: *Annals of Biomedical Engineering* 48, pp. 2580–2598.

Lu, Nannan et al. (2018). “Deep learning for fall detection: Three-dimensional CNN combined with LSTM on video kinematic data”. In: *IEEE Journal of Biomedical and Health Informatics* 23.1, pp. 314–323.

Mao, Ansi, Mehryar Mohri, and Yutong Zhong (2023). “Cross-entropy loss functions: Theoretical analysis and applications”. In: *International Conference on Machine Learning*. PMLR, pp. 23803–23828.

Neher, Henry (2018). “Hockey pose estimation and action recognition using convolutional neural networks”. MA thesis. University of Waterloo.

Ngo, Chong-Wah, Yong-Fei Ma, and Hong-Jiang Zhang (2005). “Video summarization and scene detection by graph modeling”. In: *IEEE Transactions on Circuits and Systems for Video Technology* 15.2, pp. 296–305.

Nowinski, Christopher J et al. (2024). “‘Subconcussive’is a dangerous misnomer: hits of greater magnitude than concussive impacts may not cause symptoms”. In: *British Journal of Sports Medicine*.

Pascanu, Razvan, Tomas Mikolov, and Yoshua Bengio (2013). “On the difficulty of training recurrent neural networks”. In: *International Conference on Machine Learning*. PMLR, pp. 1310–1318.

Post, Andrew and T Blaine Hoshizaki (2012). “Mechanisms of brain impact injuries and their prediction: a review”. In: *Trauma* 14.4, pp. 327–349.

Prakash, H. et al. (n.d.). *VIP-HTD: A Public Benchmark for Multi-Player Tracking in Ice Hockey*. Dataset/benchmark description.

Redmon, Joseph et al. (2016). “You Only Look Once: Unified, real-time object detection”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779–788.

Ren, Shaoqing et al. (2015). “Faster R-CNN: Towards real-time object detection with region proposal networks”. In: *Advances in Neural Information Processing Systems*. Vol. 28.

Rezaei, Ahmad and Lyndia C Wu (2022). “Automated soccer head impact exposure tracking using video and deep learning”. In: *Scientific reports* 12.1, p. 9282.

Rodger, James A (2015). “Discovery of medical Big Data analytics: Improving the prediction of traumatic brain injury survival rates by data mining Patient Informatics Processing Software Hybrid Hadoop Hive”. In: *Informatics in Medicine Unlocked* 1, pp. 17–26.

Shi, Xingjian et al. (2015). “Convolutional LSTM network: A machine learning approach for precipitation nowcasting”. In: *Advances in neural information processing systems* 28.

Srivastava, Nitish et al. (2014). “Dropout: a simple way to prevent neural networks from overfitting”. In: *Journal of Machine Learning Research* 15.1, pp. 1929–1958.

Vats, Kunal et al. (2023). "Player tracking and identification in ice hockey". In: *Expert Systems with Applications* 213, p. 119250.

Wang, Limin et al. (2018). "Temporal segment networks for action recognition in videos". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 41.11, pp. 2740–2755.

Yurgil, Kate A et al. (2014). "Association between traumatic brain injury and risk of post-traumatic stress disorder in active-duty Marines". In: *JAMA psychiatry* 71.2, pp. 149–157.

Chapter 4

Deep Learning Model for MPS Prediction from Ice Hockey Video-Derived Impact Features

Sports Engineering — Published

Amir Azadi[a,*], Parisa Dehghan[a], Clara Karton[a], Michael D. Gilchrist[b], Maia Fraser[c], T. Blaine Hoshizaki[a]

[a] School of Human Kinetics, University of Ottawa, Ottawa, Ontario, Canada
[b] School of Mechanical and Materials Engineering, University College Dublin, Dublin, Ireland
[c] Department of Mathematics and Statistics, University of Ottawa, Ottawa, Ontario, Canada

4.1 Abstract

Accurate estimation of brain tissue strain in sports is often limited by the high cost of sensors and the computational demands of FE simulations. This study presents a deep-learning framework to predict MPS directly from video-derived biomechanical features and player age in youth ice hockey. We analysed 477 on-ice head impacts involving players aged 5–17 years, with ground-truth MPS values obtained from the UCDBTM v2.0 FE brain model. For each impact, four features—velocity, surface compliance, impact location, and elevation—plus player age group were extracted. A fully connected neural network was trained on 75% of the data, validated on 15%, and tested on the remaining 10%. The model achieved a test-set MSE of 0.0015 and an R^2 of 0.89, demonstrating strong accuracy and generalization across all age groups. Permutation-importance analysis identified impact velocity as the most influential predictor, followed by surface compliance. This video-based, non-invasive method provides a scalable solution for monitoring brain strain in youth hockey and can potentially be extended to other sports.

Keywords: Youth Ice Hockey; Deep Learning; Brain Strain Prediction; MPS; Finite Element Brain Model

4.2 Introduction

Ice hockey carries a significant risk of head impacts, increasing susceptibility to TBIs including concussions (Daneshvar et al., 2011). Youth players represent the majority of the ice hockey population, accounting for approximately 69% of all registered players in the United States (USA Hockey, 2024). Notably, 15–25% of all hockey-related injuries among youth players involve head trauma, highlighting the high prevalence of TBIs among youth participants (Johnson, 2011). Younger players face distinct risks due to their larger head-to-body ratio, lower skill level, weaker neck muscles, and less refined motor skills, all of which influence their style of play and increase their vulnerability to head impacts (Daneshvar et al., 2011; Cournoyer et al., 2021).

The severity of these injuries can be calculated using biomechanical features that determine how forces are transmitted to the brain (Post and T Blaine Hoshizaki, 2012; Karton and T Blaine Hoshizaki, 2018; T. Blaine Hoshizaki et al., 2014). These features include impact velocity, compliance, location, and elevation, each influencing tissue deformation and injury risk (Post and T Blaine Hoshizaki, 2012; Karton and T Blaine Hoshizaki, 2018; T. Blaine Hoshizaki et al., 2014). Impact velocity directly influences the energy transferred to the brain, with higher velocities generating greater forces and accelerations, thereby increasing the risk of concussion (Post, T. Blaine Hoshizaki, et al., 2017; Demiannay et al., 2024). Impact compliance, or the stiffness of the impacting surface, also plays an important role as rigid surfaces like ice or boards transfer more energy, whereas softer surfaces attenuate energy and potentially reduce injury risk; however, increased compliance can also prolong impact duration, which may elevate the risk of concussion (Post, T. Blaine Hoshizaki, et al., 2017). Moreover, impact location and elevation further modulate injury severity, with lateral impacts generating higher rotational accelerations, an important characteristic linked to concussion risk and subclinical brain trauma (Post, T. Blaine Hoshizaki, et al., 2017; Nowinski et al., 2024; Jason P. Mihalik et al., 2012b; Gysland et al., 2012; David F Meaney and D. H. Smith, 2011).

In addition to impact features, individual factors such as age also contribute to injury risk. Younger players exhibit different biomechanical responses to impacts due to variations in skull thickness, brain tissue properties, and neck strength, all of which can alter the way forces are distributed and absorbed (Sophie Chatelin et al., 2012; Jason P. Mihalik et al., 2012b). Understanding both the mechanical aspects of impacts and individual variability is important for accurately assessing injury risk across different player populations.

In brain injury biomechanics, several metrics have been developed to assess impact severity, including PLA, PRA, HIC, BrIC, GAMBIT, HIP, and RIC (McHenry, 2004; Takhounts et al., 2013; J. A. Newman, 1986; J. Newman et al., 2000; Kimpara and Iwamoto, 2012). Each metric evaluates specific aspects of head kinematics; however, none comprehensively account for all dynamic variables involved in TBI outcomes, which may limit their capacity to fully capture the complex mechanisms underlying brain trauma (Rowson and Duma, 2013; Lee F Gabler, Jeff R Crandall, and Matthew B Panzer, 2016; Kleiven, 2013; Albert I. King et al., 2003; Kleiven, 2007; Clark et al., 2015).

MPS has emerged as a more comprehensive biomechanical metric for evaluating brain injury risk. Unlike traditional measures that focus primarily on head kinematics, MPS provides a tissue-level assessment by quantifying the three-dimensional, time-dependent deformation of brain tissue during head impacts (David F Meaney and D. H. Smith, 2011; Kleiven, 2007) By integrating linear and rotational accelerations along with impact duration, MPS quantifies peak tensile strain, effectively capturing how mechanical forces induce tissue damage (Luke F. Gabler, Jeff R. Crandall, and Matthew B. Panzer, 2019; Hernandez et al., 2015). Additionally, MPS accounts for the anatomical complexity and heterogeneity of brain tissue, including gray and white matter differentiation, sulcal geometry, and viscoelastic anisotropy, which collectively determines regions of increased vulnerability to injury (Simon Chatelin, Constantinesco, and Willinger, 2010; Giordano and Kleiven, 2014).

Clinically, MPS provides a valuable indicator of injury severity, with strain thresholds between 15–25% often associated with heightened concussion risk and structural brain damage, as supported by both animal and human studies (Hernandez et al., 2015; Patton, McIntosh, et al., 2012; Bain and David F. Meaney, 2000). FE models offer a powerful tool for evaluating MPS by simulating brain responses to impacts using realistic viscoelastic properties and skull-brain boundary conditions (Horgan and Gilchrist, 2003; Johnson Ho and Kleiven, 2007; Antonia Trotta et al., 2020). However, while FE models provide highly detailed MPS calculations, their practical application in dynamic, real-time environments, such as youth hockey, remains limited. The computational intensity of these models, coupled with the need for specialized expertise, makes them challenging to implement in injury monitoring or decision-making scenarios (Kleiven, 2007; Raul et al., 2008).

These challenges underscore the need for alternative approaches, to enable rapid and accessible injury assessments. Recent studies have demonstrated the efficacy of AI techniques in predicting MPS without relying on computationally intensive FE models (Zhan et al., 2021; Wu, Zhao, Ghazi, et al., 2019; Wu, Zhao, Barbat, et al., 2022). Zhan et al. (2020) developed a deep learning head model using sensor-derived head kinematics (e.g., linear and rotational accelerations) as inputs, capable of estimating MPS across the entire brain, achieving an average root mean squared error of 0.025 (Zhan et al., 2021). Similarly, Wu et al. (2021) extended a convolutional neural network using accelerometer and gyroscopic data to instantly estimate peak MPS distributions throughout the brain, offering significant computational advantages over FE simulations (Wu, Zhao, Barbat, et al., 2022). However, existing AI models primarily depend on head kinematics measured from sensor data (e.g., accelerometers, instrumented mouthguards). While sensor-based methods provide valuable input for estimating MPS, studies show they may underestimate rotational acceleration measurements—one of the key biomechanical contributors to concussive brain injury (Liu et al., 2020; Greybe et al., 2020; Patton, Huber, McDonald, et al., 2020; Patton, Huber, Jain, et al., 2020). In a practical sense, sensors pose considerable limitations in youth sports as they are costly and therefore primarily used in elite or research settings. Additionally, with a high rate of false-positive impacts, many systems require manual verification through video review, which is time-consuming (Liu et al., 2020; Greybe et al., 2020).

To overcome these limitations, this study proposes a novel AI-based approach for predicting brain injury risk. By leveraging biomechanical impact features extracted directly from video analysis—including impact velocity, compliance, location, and elevation—specific to player age, our method eliminates the need for wearable sensors or FE models. This approach enhances accessibility, reduces processing time, and enables real-time, scalable assessments of injury risk, particularly suited for youth and amateur hockey environments.

4.3 Methods

4.3.1 Overview

The methodology involves a multi-step process, including data collection, laboratory reconstruction, finite element modelling, data preprocessing, neural network development, and evaluation. Each step is structured to ensure accurate and reliable predictions of MPS based on biomechanical impact features.

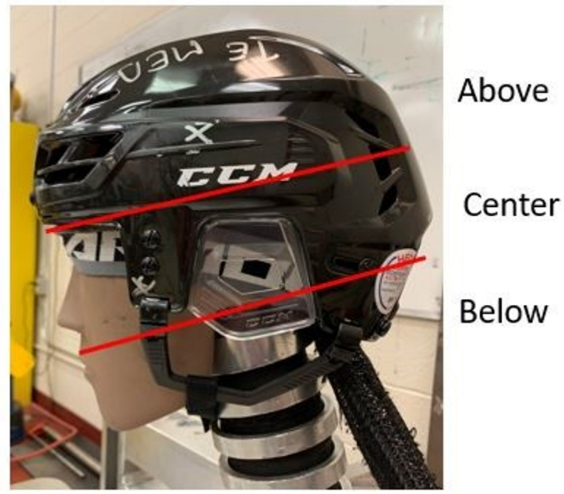
4.3.2 Data Collection and Preparation

Head impact events were identified from video recordings of 319 youth ice hockey games. Three trained, independent biomechanical injury specialists reviewed the footage, documenting 477 unique impact conditions (defined by each permutation of impact features) of common impact types across three participant age groups: 5–6 years (61 impacts), 7–14 years (280 impacts), and 15–17 years (136 impacts). Each impact event was characterized using a combination of video-derived features and player age category, including:

- **Surface of contact or compliance:** glass, boards, boards angled, shoulder, elbow, ice, glove, or head.
- **Impact location:** front, side, rear, rear boss, front boss, or crown.
- **Impact elevation:** above center of gravity, center of gravity, or below center of gravity.
- **Impact velocity:** categorized as Very Low, Low, Medium, or High based on frame-by-frame measurements (Post, Koncan, et al., 2018).
- **Player age category:** 5–6 years old, 7–14 years old, and 15–18 years old.

These detailed features served as inputs for subsequent analysis. Figure 4.1 illustrates categorized impact locations and elevation zones, providing a structured framework for analyzing head impacts (Post and T Blaine Hoshizaki, 2012; Karton and T Blaine Hoshizaki, 2018).

(a) Impact locations



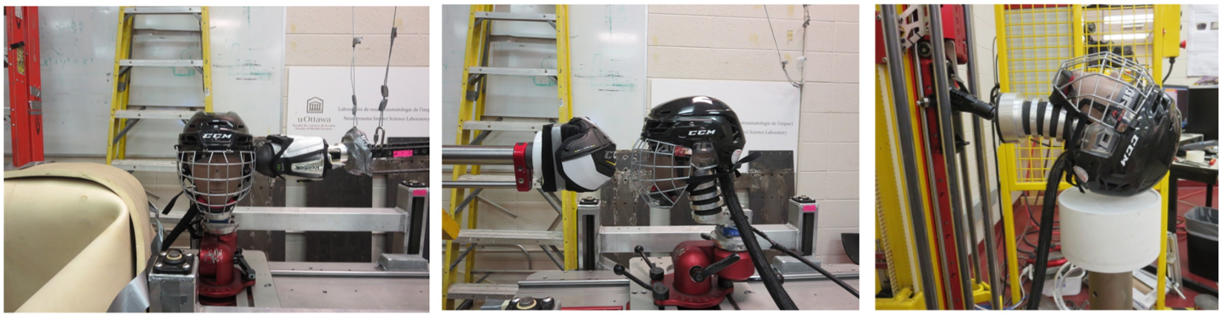
(b) Elevation zones

Figure 4.1: Impact locations and elevation zones used to categorize head impacts.

4.3.3 Ground Truth Data Generation for Neural Network Training and Validation

To generate ground truth dataset, the 477 unique head impacts were reconstructed in a controlled laboratory setting. The impact conditions observed in real-world ice hockey games were replicated using specialized apparatuses designed to simulate different impact mechanisms:

- Simple pendulum (Fig. 4.2a)
- Pneumatic linear impactor (Fig. 4.2b)
- Monorail drop rig (Fig. 4.2c)



a b c

Figure 4.2: Laboratory apparatus used to reconstruct head impacts: (a) pendulum, (b) pneumatic linear impactor, (c) monorail drop rig.

Instrumented Hybrid III headforms representing the size and mass of each demographic were used and equipped with a 3-2-2-2 accelerometer array to capture linear and rotational head accelerations for each impact (Padgaonkar, Krieger, and A. I. King, 1975). A detailed description of impact reconstruction methods for each age group and event condition is presented in Karton et al. (Karton, Post, et al., 2021).

The kinematic head acceleration time history data were input into the University College Dublin Brain Trauma Model version 2.0 (UCDBTMv2.0) (Antonia Trotta et al., 2020), to calculate the deformation, measured as MPS for each head impact event. The head geometry of this model was determined using adult male CT and MRI and was validated against pressure responses and brain motion cadaveric data (Nahum, R. Smith, and Ward, 1977; Hardy et al., 2001).

To account for the smaller head size and morphological differences of youth athletes, scaling adjustments were implemented. Specifically, this model was scaled to 90% for the 5–6 years old group, 95% for the 7–14 years old group, and full size for the 15–17 group, ensuring that age-appropriate head geometries were used in the simulations (Karton, Post, et al., 2021).

Consisting of approximately 184,000 hexahedral elements (Horgan and Gilchrist, 2003), the anatomical regions and material properties of UCDBTMv2.0 are presented in Table 4.1 (Horgan and Gilchrist, 2003; MacManus et al., 2017; Van Noort et al., 1981; Angelo Trotta and Ó Annaidh, 2019). Mesh integrity of the UCDBTMv2.0 was conducted to ensure no errors resulted from high aspect ratios (Yang, 2011; Johnny Ho, Holst, and Kleiven, 2009). In addition, the artificial energy of the simulation associated with hourglassing was monitored and found to be less than 3% of the total internal energy.

Table 4.1: UCDBTM v2.0 material models and properties.

Region	Model	Density (kg/m^3)	Poisson's ν	Key parameters (source)
Grey matter	Visco-hyperelastic	1060	~0.50	$\mu = 5715$ Pa, $g_1 = 0.534$, $\tau_1 = 0.020$ s $g_2 = 0.207$, $\tau_2 = 0.304$ s, $g_\infty = 0.258$ (Horgan and Gilchrist, 2003; MacManus et al., 2017)
White matter	Viscoelastic	1060	~0.5	$E = 37{,}500$ Pa $g_1 = 0.80$, $\tau_1 = 0.0125$ s (MacManus et al., 2017)
Cerebellum	Visco-hyperelastic	1060	~0.5	$\mu = 2611$ Pa, $g_1 = 0.515$, $\tau_1 = 0.020$ s $g_2 = 0.187$, $\tau_2 = 0.302$ s, $g_\infty = 0.298$ (Horgan and Gilchrist, 2003)
Brain stem	Visco-hyperelastic	1060	~0.5	$\mu = 4768$ Pa, $g_1 = 0.630$, $\tau_1 = 0.0185$ s $g_2 = 0.175$, $\tau_2 = 0.290$ s, $g_\infty = 0.290$ (Horgan and Gilchrist, 2003)
Pia	Linear elastic	1130	0.45	$E \approx 11.50$ MPa (Horgan and Gilchrist, 2003)
Dura, Falx, Tentorium	Hyperelastic (Ogden)	1130	~0.5	$\mu = 3.602$ MPa $\alpha = 13.73$ (Van Noort et al., 1981)
Ventricles (CSF space)	Visco-hyperelastic	1040	~0.5	$C_{10} = 3653.5$ Pa, $C_{01} = 4059.44$ Pa $g_1 = 0.527$, $\tau_1 = 0.008$ s, $g_2 = 0.303$, $\tau_2 = 0.145$ s (MacManus et al., 2017)
CSF	Linear elastic	1000	~0.5	$E \approx 0.15$ MPa (MacManus et al., 2017)
Trabecular bone	Linear elastic	1300	0.24	$E \approx 1000$ MPa (Horgan and Gilchrist, 2003)
Cortical bone	Linear elastic	2000	0.22	$E \approx 1500$ MPa (Horgan and Gilchrist, 2003)
Facial bone	Linear elastic	2100	0.22	$E \approx 5540$ MPa (Horgan and Gilchrist, 2003)
Scalp	Hyperelastic (Ogden)	1133	~0.5	$\mu = 1.48$ MPa $\alpha = 8.1$ (Angelo Trotta and Ó Annaidh, 2019)

4.3.4 Data Preprocessing for Neural Network Training

Each impact event was described by a set of biomechanical features—compliance (surface type), impact location, elevation, velocity—as well as player age category. The corresponding FE-derived MPS was used as the output target. All inputs were encoded using one-hot encoding to convert them into a numerical format compatible with neural network training. The dataset (477 samples) was split into training (75%, 358 samples), validation (15%, 72 samples), and test (10%, 47 samples) sets for robust model development and evaluation.

4.3.5 Neural Network Architecture and Training

A fully connected neural network (dense) was developed to capture the nonlinear relationship between impact features and MPS. The network was optimized using the Adam optimizer (learning rate = 0.001) (Kingma and Ba, 2014) with MSE as the loss function and R^2 as a key performance metric. The training was conducted over 100 epochs with a batch size of 16, using early stopping to prevent overfitting and retain the model weights corresponding to the lowest validation loss. The detailed neural network architecture employed in this study is illustrated in Figure 4.3.

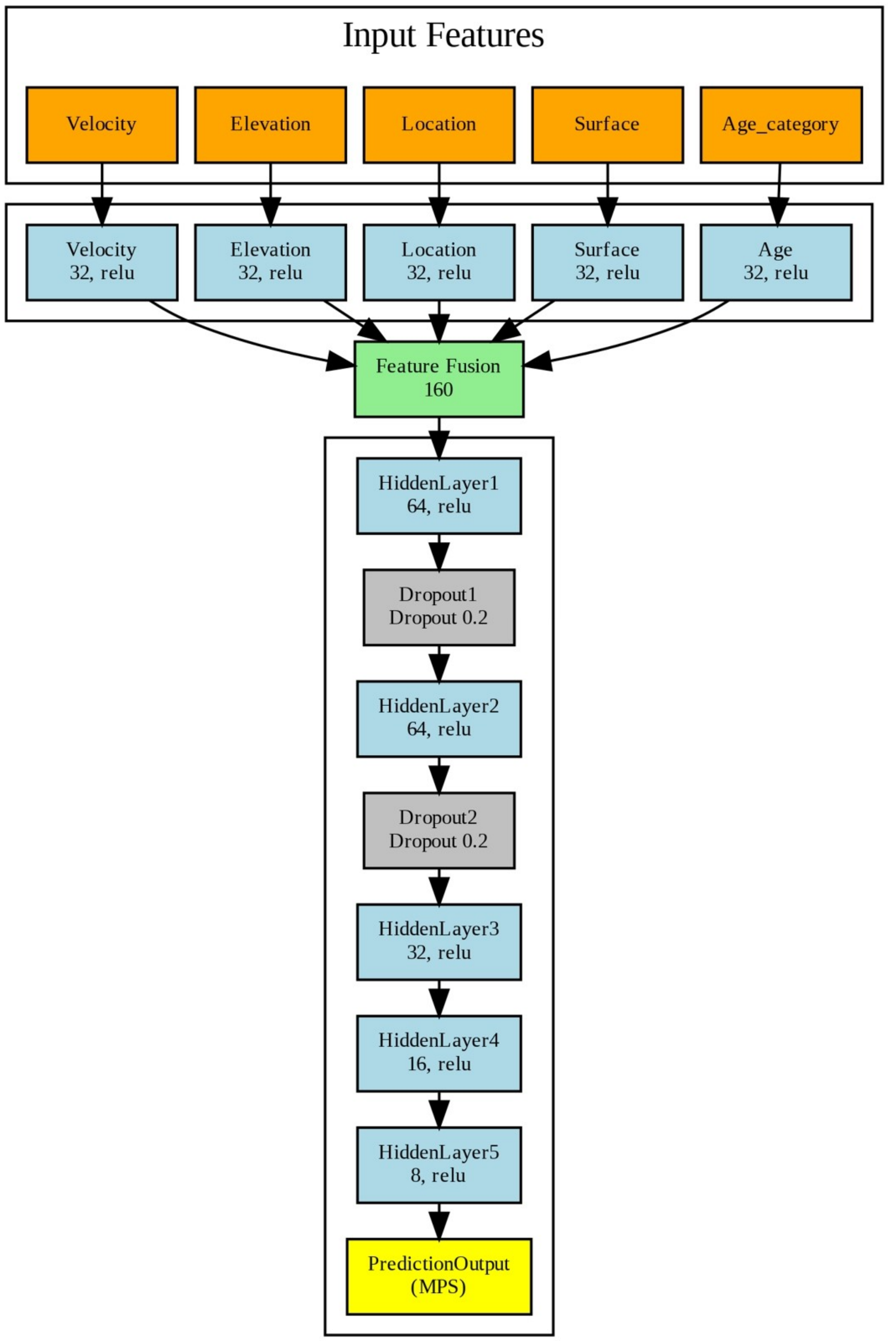


Figure 4.3: Architecture of the fully connected neural network used to predict MPS from impact features.

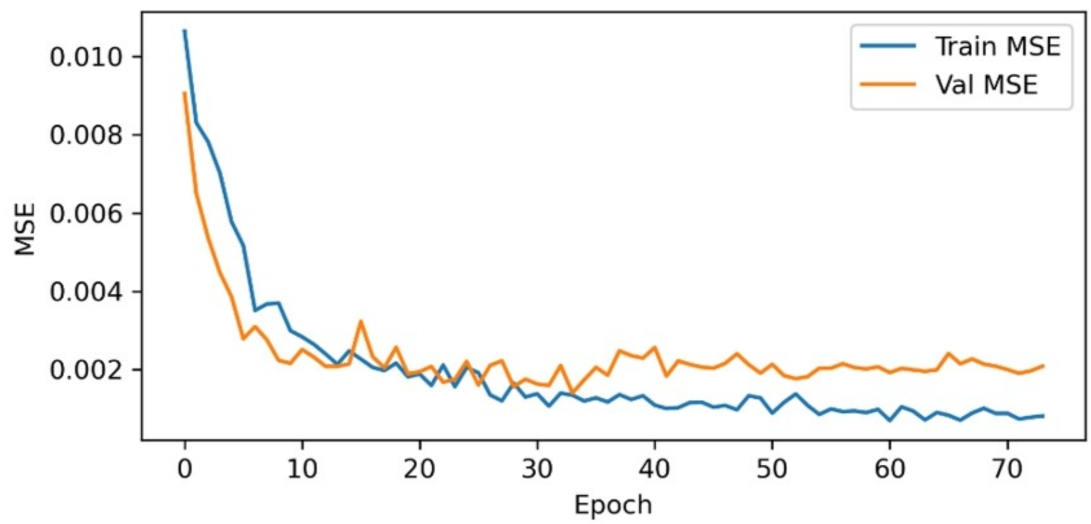
4.3.6 Evaluation and Feature Importance Analysis

The model's performance was evaluated using the test set, with final performance measured by MSE and R^2. To assess the contribution of each feature to the model's predictions, a permutation importance analysis was conducted. In this analysis, each feature was independently shuffled while keeping the others intact, and the resultant changes in MSE and R^2 were recorded. A larger increase in MSE (ΔMSE $\uparrow$) indicated that the feature had a greater influence on the model's prediction error. A larger drop in R^2 ($\Delta R^2 \downarrow$) suggested that the feature played an important role in explaining the variance in MPS. The analysis was repeated 15 times to average out variability, yielding mean importance scores and standard deviations for both metrics. This approach provided insights into the relative importance of each biomechanical feature in predicting MPS.

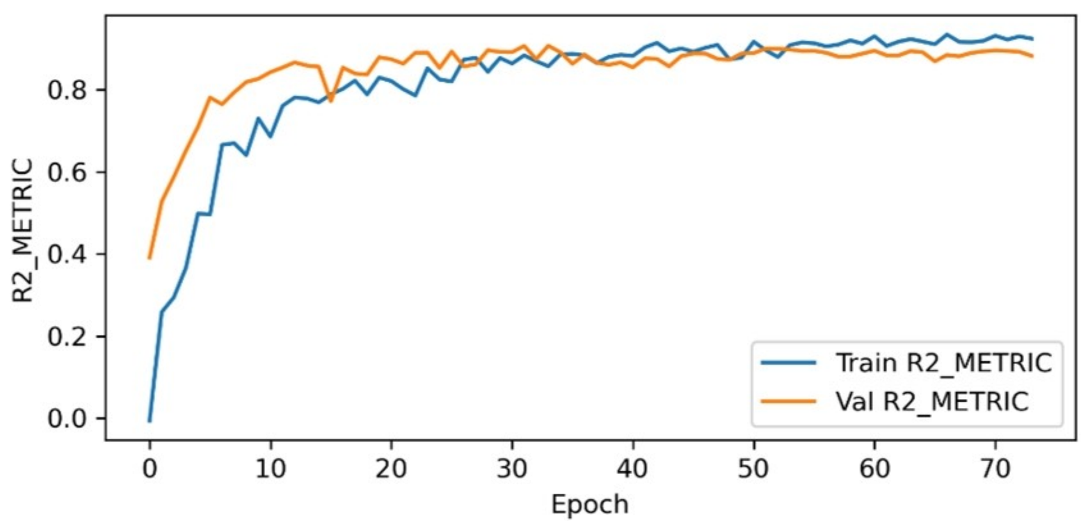
4.4 Results

4.4.1 Model Performance

To assess model effectiveness, the neural network was trained to predict MPS based on input features. The model demonstrated strong predictive performance on the test set, achieving a MSE of 0.0015 and an R^2 of 0.89. The training and validation curves for MSE and R^2, shown in Figure 4.4, illustrate smooth convergence across 100 epochs. The model exhibits stable learning dynamics, with no signs of overfitting. The close alignment between training and validation traces underscores the model's robustness and generalizability.



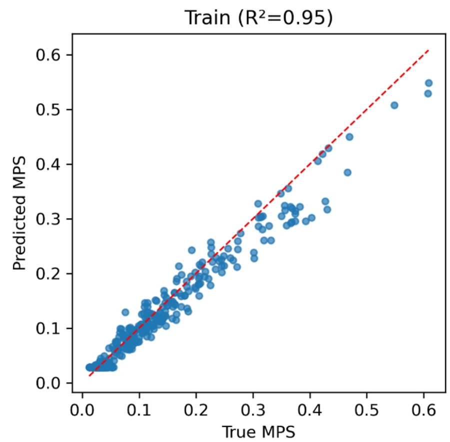
(a) MSE on train/validation sets over 100 epochs.



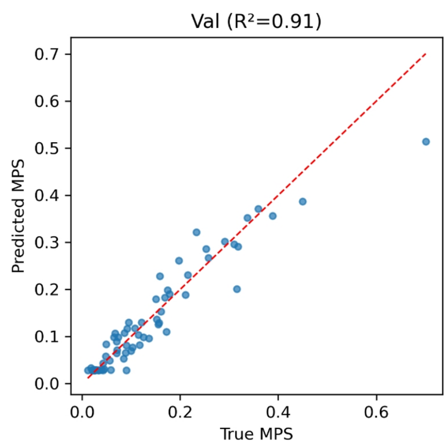
(b) R^2 on train/validation sets over 100 epochs.

Figure 4.4: Learning curves over 100 epochs for (a) MSE and (b) R^2 on train/validation sets.

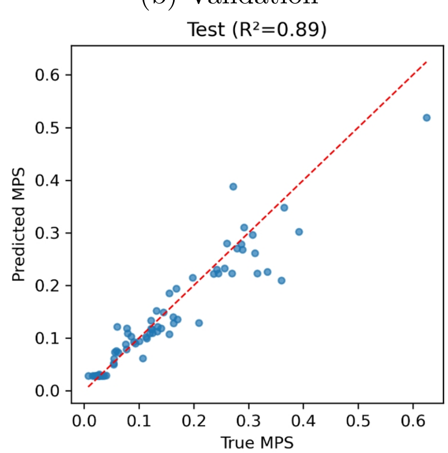
Predicted vs. true MPS scatter plots for the train, validation, and test sets are shown in Figure 4.5. Predictions were closely aligned with ground truth values, with data points tightly clustered around the identity line across all splits.



(a) Train



(b) Validation



(c) Test

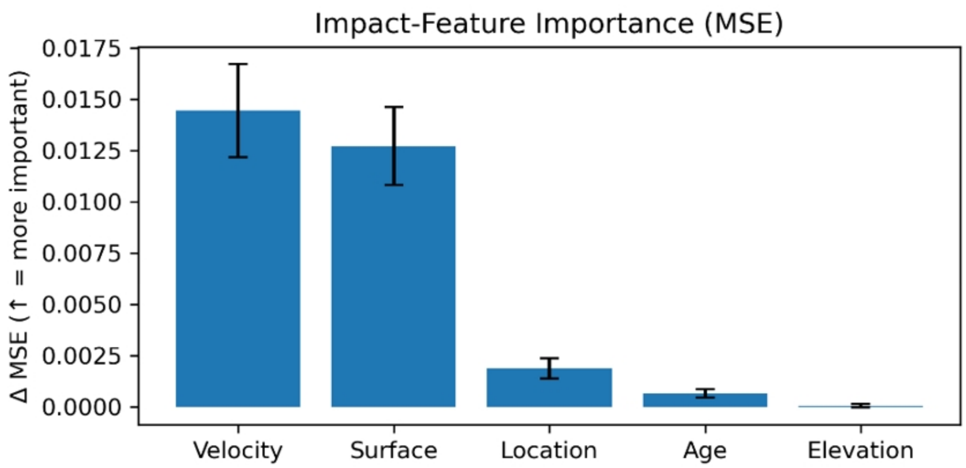
Figure 4.5: Predicted vs. true MPS for (a) training, (b) validation, and (c) test splits.

4.4.2 Impact-Feature Contributions

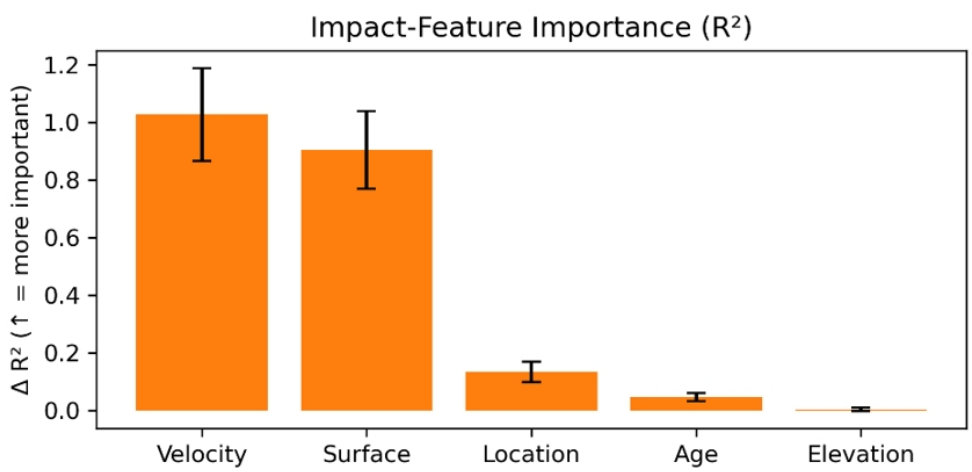
Permutation importance analysis was used to assess the contribution of each input feature to the model's predictive performance. Impact velocity showed the highest importance scores, followed by surface compliance. Other features, including impact location, age group, and impact elevation, ranked lower in comparison. Table 4.2 summarizes the ΔMSE and R^2 values (mean ± standard deviation) for each feature, and Figure 4.6 displays bar plots.

Table 4.2: Permutation-importance scores (mean ± sd over 15 shuffles). Larger ΔMSE / drop in R^2 indicates greater importance.

Feature	**ΔMSE**	**Δ$\mathbf{R}^2$**
Velocity	0.0144 ± 0.003	1.03 ± 0.12
Surface compliance	0.0047 ± 0.002	0.32 ± 0.07
Location	0.0015 ± 0.001	0.09 ± 0.03
Elevation	0.0012 ± 0.001	0.07 ± 0.03
Age group	0.0008 ± 0.001	0.05 ± 0.02



(a) Feature importance visualized by ΔMSE.



(b) Feature importance visualized by ΔR^2.

Figure 4.6: Feature importance visualized by ΔMSE (blue) and ΔR^2 (orange).

4.4.3 Generalization Across Player Categories

Model performance was also evaluated across three player age categories to assess generalization across different demographic profiles. As shown in Table 4.3, the model achieved MSE values below 0.0024 and R^2 values exceeding 0.85 across all age groups.

Table 4.3: Test performance across age categories.

Age group	n (test)	MSE	R^2
5–6 yrs	10	0.0012	0.90
7–14 yrs	24	0.0024	0.86
15–17 yrs	13	0.0018	0.87

4.5 Discussion

This study presents a novel AI-based framework for estimating MPS directly from video-extracted biomechanical features and player age in youth ice hockey. By training a neural network on strain outputs generated from FE simulations, the proposed method eliminates the need for wearable sensors or computationally intensive modeling, offering a scalable and accessible solution for head injury risk assessment. The results demonstrate the feasibility of MPS prediction using impact features, with key findings discussed below.

Model Performance The neural network model demonstrated strong predictive performance on the test set, achieving a MSE of 0.0015 and an R^2 score of 0.89. Scatter plots of predicted versus true MPS showed tight clustering around the identity line across training, validation, and test sets, indicating that the network effectively captured meaningful patterns in the data and generalized well to unseen samples.

Impact Feature Contributions Permutation importance analysis revealed that impact velocity was the most influential predictor of MPS. When velocity was shuffled, model performance deteriorated substantially—ΔMSE increased by 0.0144, and ΔR^2 dropped by approximately 1.03. This strong dependency highlights the biomechanical relevance of velocity in governing brain deformation responses and confirms its central role in impact severity estimation (Post and T Blaine Hoshizaki, 2012; Jason P Mihalik et al., 2012a; David F Meaney and D. H. Smith, 2011). Surface compliance ranked as the second most important feature, with notable effects on model accuracy. These results underscore the importance of surface material in modulating energy transfer during impacts. Rigid surfaces such as ice or boards tend to induce higher MPS due to minimal energy absorption, while more compliant surfaces help attenuate force transmission. This aligns with prior biomechanical research demonstrating that contact material properties significantly influence head kinematics and strain (Post and T Blaine Hoshizaki, 2012; Karton and T Blaine

Hoshizaki, 2018; Post, T. Blaine Hoshizaki, et al., 2017; Gysland et al., 2012). In contrast, impact location, age group, and elevation showed relatively low importance. Their limited contribution suggests that—within the context of this dataset—the variance in MPS was primarily driven by velocity and compliance. However, this should not be interpreted as a general conclusion. In datasets with less velocity variation or more controlled surface conditions, the relative importance of features such as location and elevation may increase (Cournoyer et al., 2021; Jason P. Mihalik et al., 2012b; Karton, Post, et al., 2021).

Generalization Across Player Categories The model demonstrated reliable generalization across all player categories, including 5–6 years, 7–14 years, and 15–17 years groups. MSE values ranged from 0.0012 to 0.0024, and R^2 exceeded 0.85 across all age groups, indicating stable performance across diverse demographic segments. This robustness suggests that the network captures generalizable patterns in head impact biomechanics, making it applicable across a wide range of youth hockey populations.

Limitations *Dataset Size*—Although the dataset was divided into training, validation, and test sets, the relatively small size of the validation and test subsets may limit the precision of performance evaluation. This could impact the model's ability to fully capture variability in impact features and player subgroups. Expanding the dataset would strengthen generalizability and allow for more robust assessments across diverse conditions.

Finite Element Model Assumptions—The ground truth strain values were derived from a FE brain model that uses fixed material properties for all individuals. To approximate age-related anatomical differences, a scaled-down version of the same model geometry was used for younger players. However, because the material properties remained constant, the model does not account for potential age-related variations in brain tissue biomechanics. This simplification, while practical, may introduce limitations in capturing the full biomechanical response across age groups.

4.6 Conclusion

This study introduced a deep learning framework for estimating MPS directly from video-extracted biomechanical features and player age in youth ice hockey. By training neural networks on ground truth strain values derived from FE simulations, the model offers a scalable and efficient alternative to sensor-based systems or computationally intensive injury risk assessments. Among the input features, impact velocity emerged as the most influential predictor of MPS, underscoring its central role in determining head impact severity. Surface compliance ranked second in importance, highlighting the biomechanical relevance of material properties at the point of contact. Other features—impact location, elevation, and player age category—contributed less to predictive accuracy within the current dataset. These results demonstrate the feasibility of using video-derived features and basic demographic information to accurately estimate brain strain responses. This approach enables real-time, non-invasive monitoring of head impact severity and presents a practical tool for enhancing concussion surveillance in youth and amateur sports.

References

Bain, Anthony C. and David F. Meaney (2000). "Tissue-level thresholds for axonal damage in an experimental model of central nervous system white matter injury". In: *Journal of Biomechanical Engineering* 122.6, pp. 615–622.

Chatelin, Simon, André Constantinesco, and Rémy Willinger (2010). "Fifty years of brain tissue mechanical testing: from in vitro to in vivo investigations". In: *Biorheology* 47.5-6, pp. 255–276.

Chatelin, Sophie et al. (2012). "Towards child versus adult brain mechanical properties". In: *Journal of the Mechanical Behavior of Biomedical Materials* 6.1, pp. 166–173.

Clark, J. M. et al. (2015). "Determining the relationship between linear and rotational acceleration and MPS for different magnitudes of classified brain injury risk in ice hockey". In: *IRCOBI Conference Proceedings*.

Cournoyer, J et al. (2021). "Brain trauma exposure for American tackle football players 5 to 9 and 9 to 14 years of age". In: *Journal of biomechanics* 127, p. 110689.

Daneshvar, Daniel H et al. (2011). "The epidemiology of sport-related concussion". In: *Clinics in sports medicine* 30.1, pp. 1–17.

Demiannay, Jean-Jacques et al. (2024). "Intracranial Displacements due to Blunt Force Impact in a Postmortem Human Surrogate Brain". In: *SAE International Journal of Transportation Safety* 12, pp. 113–120.

Gabler, Lee F, Jeff R Crandall, and Matthew B Panzer (2016). "Assessment of kinematic brain injury metrics for predicting strain responses in diverse automotive impact conditions". In: *Annals of biomedical engineering* 44, pp. 3705–3718.

Gabler, Luke F., Jeff R. Crandall, and Matthew B. Panzer (2019). "Development of a second-order system for rapid estimation of maximum brain strain". In: *Annals of Biomedical Engineering* 47.9, pp. 1971–1981.

Giordano, Chiara and Svein Kleiven (2014). "Evaluation of axonal strain as a predictor for mild traumatic brain injuries using finite element modeling". In: *SAE Technical Paper*.

Greybe, David G. et al. (2020). "Comparison of head impact measurements via an instrumented mouthguard and an anthropometric testing device". In: *Sports Engineering* 23, pp. 1–11.

Gysland, Sonia M et al. (2012). "The relationship between subconcussive impacts and concussion history on clinical measures of neurologic function in collegiate football players". In: *Annals of biomedical engineering* 40, pp. 14–22.

Hardy, Warren N. et al. (2001). "Investigation of head injury mechanisms using neutral density technology and high-speed biplanar X-ray". In: *SAE Technical Paper*.

Hernandez, Fidel et al. (2015). "Six degree-of-freedom measurements of human mild traumatic brain injury". In: *Annals of biomedical engineering* 43, pp. 1918–1934.

Ho, Johnny, Hans von Holst, and Svein Kleiven (2009). "Automatic generation and validation of patient-specific finite element head models suitable for crashworthiness analysis". In: *International Journal of Crashworthiness* 14.6, pp. 555–563.

Ho, Johnson and Svein Kleiven (2007). "Dynamic response of the brain with vasculature: a three-dimensional computational study". In: *Journal of biomechanics* 40.13, pp. 3006–3012.

Horgan, Timothy James and Michael D Gilchrist (2003). "The creation of three-dimensional finite element models for simulating head impact biomechanics". In: *International Journal of Crashworthiness* 8.4, pp. 353–366.

Hoshizaki, T. Blaine et al. (2014). "Current and Future Concepts in Helmet and Sports Injury Prevention". In: *Neurosurgery* 75.S4, S136–S148. DOI: 10.1227/NEU.0000000000000496.

Johnson, L. S. M. (2011). "Concussion in youth ice hockey: It's time to break the cycle". In: *CMAJ* 183.8, pp. 921–924. DOI: 10.1503/cmaj.110282.

Karton, Clara and T Blaine Hoshizaki (2018). "Concussive and subconcussive brain trauma: the complexity of impact biomechanics and injury risk in contact sport". In: *Handbook of clinical neurology* 158, pp. 39–49.

Karton, Clara, Andrew Post, et al. (2021). "Exposure to brain trauma in six age divisions of minor ice hockey". In: *Journal of biomechanics* 116, p. 110203.

Kimpara, Hideyuki and Masami Iwamoto (2012). "Mild traumatic brain injury predictors based on angular accelerations during impacts". In: *Annals of biomedical engineering* 40, pp. 114–126.

King, Albert I. et al. (2003). "Is head injury caused by linear or angular acceleration?" In: *IRCOBI Conference*. Lisbon, Portugal.

Kingma, Diederik P. and Jimmy Ba (2014). "Adam: A Method for Stochastic Optimization". In: *arXiv preprint arXiv:1412.6980*.

Kleiven, Svein (2007). “Predictors for traumatic brain injuries evaluated through accident reconstructions”. In: *SAE Technical Paper*.
— (2013). “Why most traumatic brain injuries are not caused by linear acceleration but skull fractures are”. In: *Frontiers in Bioengineering and Biotechnology* 1, p. 15.
Liu, Yuzhe et al. (2020). “Validation and comparison of instrumented mouthguards for measuring head kinematics and assessing brain deformation in football impacts”. In: *Annals of Biomedical Engineering* 48, pp. 2580–2598.
MacManus, Daniel B. et al. (2017). “Region and species dependent mechanical properties of adolescent and young adult brain tissue”. In: *Scientific Reports* 7, p. 13729.
McHenry, Brian G (2004). “Head injury criterion and the ATB”. In: *ATB Users' group* 29, pp. 5–8.
Meaney, David F and Douglas H Smith (2011). “Biomechanics of concussion”. In: *Clinics in sports medicine* 30.1, pp. 19–31.
Mihalik, Jason P et al. (2012a). “Head impact biomechanics in youth hockey: comparisons across playing position, event types, and impact locations”. In: *Annals of biomedical engineering* 40, pp. 141–149.
Mihalik, Jason P. et al. (2012b). “Head impact biomechanics in youth hockey: comparisons across playing position, event types, and impact locations”. In: *Annals of Biomedical Engineering* 40.1, pp. 141–149.
Nahum, Alan M., Robert Smith, and Charles C. Ward (1977). “Intracranial pressure dynamics during head impact”. In: *SAE Technical Paper*.
Newman, J et al. (2000). “A new biomechanical assessment of mild traumatic brain injury. Part 2: results and conclusions”. In: *Proceedings of the International Research Council on the Biomechanics of Injury conference*. Vol. 28. International Research Council on Biomechanics of Injury.
Newman, James A. (1986). “A generalized acceleration model for brain injury threshold (GAMBIT)”. In: *Proceedings of the International Conference on the Biomechanics of Impact*, pp. 121–131.
Nowinski, Christopher J et al. (2024). “‘Subconcussive’is a dangerous misnomer: hits of greater magnitude than concussive impacts may not cause symptoms”. In: *British Journal of Sports Medicine*.
Padgaonkar, A. J., K. W. Krieger, and A. I. King (1975). “Measurement of angular acceleration of a rigid body using linear accelerometers”. In: *ASME Journal of Applied Mechanics* 42.3, pp. 552–556.
Patton, Daniel A., Christine M. Huber, Divya Jain, et al. (2020). “Head impact sensor studies in sports: a systematic review of exposure confirmation methods”. In: *Annals of Biomedical Engineering* 48.11, pp. 2497–2507.

Patton, Daniel A., Christine M. Huber, Catherine C. McDonald, et al. (2020). "Video confirmation of head impact sensor data from high school soccer players". In: *American Journal of Sports Medicine* 48.5, pp. 1246–1253.

Patton, Daniel A., Andrew S. McIntosh, et al. (2012). "Injury data from unhelmeted football head impacts evaluated against critical strain tolerance curves". In: *Proceedings of the IMechE, Part P: Journal of Sports Engineering and Technology* 226.3-4, pp. 177–184.

Post, Andrew and T Blaine Hoshizaki (2012). "Mechanisms of brain impact injuries and their prediction: a review". In: *Trauma* 14.4, pp. 327–349.

Post, Andrew, T. Blaine Hoshizaki, et al. (2017). "Peak linear and rotational acceleration magnitude and duration effects on maximum principal strain in the corpus callosum for sport impacts". In: *Journal of Biomechanics* 61, pp. 183–192. DOI: 10.1016/j.jbiomech.2017.07.023.

Post, Andrew, David Koncan, et al. (2018). "Analysis of speed accuracy using video analysis software". In: *Sports Engineering* 21.3, pp. 235–241.

Raul, Jean-Sébastien et al. (2008). "Finite-element models of the human head and their applications in forensic practice". In: *International journal of legal medicine* 122, pp. 359–366.

Rowson, Steven and Stefan M Duma (2013). "Brain injury prediction: assessing the combined probability of concussion using linear and rotational head acceleration". In: *Annals of biomedical engineering* 41, pp. 873–882.

Takhounts, Edgar G. et al. (2013). "Development of brain injury criteria (BrIC)". In: *SAE Technical Paper*.

Trotta, Angelo and Aisling N. Ó Annaidh (2019). "Mechanical characterisation of human and porcine scalp tissue at dynamic strain rates". In: *Journal of the Mechanical Behavior of Biomedical Materials* 100, p. 103381.

Trotta, Antonia et al. (2020). "Biofidelic finite element modelling of brain trauma: Importance of the scalp in simulating head impact". In: *International Journal of Mechanical Sciences* 173, p. 105448.

USA Hockey (2024). *2023-24 Membership Statistics*. https://www.usahockey.com/membershipstats. Retrieved from USA Hockey website.

Van Noort, Ronald et al. (1981). "A study of the uniaxial mechanical properties of human dura mater preserved in glycerol". In: *Biomaterials* 2.1, pp. 41–45.

Wu, Siyuan, Wenshuai Zhao, Sorin Barbat, et al. (2022). "Instantaneous brain strain estimation for automotive head impacts via deep learning". In: *SAE Technical Paper*.

Wu, Siyuan, Wenshuai Zhao, Kianoush Ghazi, et al. (2019). "Convolutional neural network for efficient estimation of regional brain strains". In: *Scientific Reports* 9.1, p. 17326.

Yang, Jingwen (2011). “Investigation of brain trauma biomechanics in vehicle traffic accidents using human body computational models”. In: *Computational Biomechanics for Medicine: Soft Tissues and the Musculoskeletal System*. Springer, pp. 5–14.
Zhan, Xu et al. (2021). “Rapid estimation of entire brain strain using deep learning models”. In: *IEEE Transactions on Biomedical Engineering* 68.11, pp. 3424–3434.

Chapter 5

Discussion

This series of studies represents a progressive effort to develop a scalable, video-based pipeline for assessing brain injury risk in youth ice hockey. Together, they form a three-stage system that progresses from contact detection, to head-impact classification, to brain-strain prediction, systematically addressing challenges related to data scarcity, annotation efficiency, and biomechanical interpretation.

Chronologically, the earliest component of this work focused on the direct detection of head impacts using a LRCN (Donahue et al., 2015), presented in Chapter 3 as Stage 2 of the pipeline. The model demonstrated strong potential to differentiate head impacts from non-head impact events within game footage, suggesting that video-based head impact detection is a feasible and trainable task. However, due to the limited size of the available dataset, the model exhibited signs of overfitting, which highlighted a major challenge: head impacts are relatively rare events, especially in youth games, inherently restricting the number of training samples. Another limitation was the initial need for manual cropping to generate player-focused clips, which made the dataset creation process time-consuming, inconsistent, and difficult to scale. These observations motivated the development of an upstream contact-detection stage and a fully automated detect–track–crop pipeline to generate player-centred inputs more efficiently.

To address the data limitations observed in the head-impact detection task, the contact detection stage presented in Chapter 2 (Stage 1 of the pipeline) expanded the objective to include all physical contact events, not just head impacts. This broader scope significantly increased the number of candidate events from which head impacts could later be extracted and introduced a pre-filtering mechanism that reduced manual annotation time from several hours to under one hour per game. This advancement not only enhanced annotation efficiency but also enabled the rapid generation of large-scale datasets required for training more advanced deep learning models.

A central component of this approach was an automated cropping pipeline that generated player-centred clips by detecting, tracking, and cropping around individual players. In Stage 1, this pipeline was systematically adapted and evaluated for youth hockey, where contact event detection performance relied heavily on two foundational components: player detection and multi-object tracking. The original pipeline, developed using professional hockey footage characterized by consistent lighting, larger player size, and fewer occlusions, was challenged by the visual complexity of youth hockey, including smaller players, frequent occlusions, and erratic motion. To address these challenges, the detection and tracking models were retrained using youth-specific data and enhanced with improved identity-matching algorithms. These adaptations resulted in more accurate tracking, fewer ID switches, and more stable player trajectories, which were critical for reliably extracting player-centred clips for contact analysis.

To classify physical contact events from these generated clips, we employed a TSM (Lin, Gan, and Han, 2019), a lightweight yet effective architecture optimized for efficient video understanding. This model was trained to determine whether a given player-focused clip contained a physical contact event. We conducted a detailed evaluation of both the model and dataset, focusing particularly on the impact of spatial context and decision thresholds.

Our analysis revealed the critical role of bounding-box scale in classification performance. A 1.5× crop around each player yielded the best results, providing sufficient contextual cues such as surrounding opponents and motion without introducing excessive background noise. Compared to smaller (1×) or larger (3×) crops, the 1.5× configuration significantly reduced both false positives and false negatives.

Threshold selection was also instrumental in managing the trade-off between precision and recall. In injury surveillance applications, where missing a true contact is more concerning than over-flagging, lower thresholds were preferred. Operating at a threshold below 0.20 maintained recall above 90%, with a manageable increase in false positives, which could be efficiently reviewed.

Error analysis provided further insights into model performance. Many false negatives were not true classification failures but resulted from various factors, including:

- Duplicate tracklets for the same contact event,
- Successful detection through another player's trajectory,
- Low-intensity or brief contacts that visually resembled routine gameplay,

- Partial occlusions, or
- Misalignment between the timing of the contact and the center of the test clips caused the key moment of impact to appear near the edges of the clip or be only partially visible.

These findings suggest that future improvements could include overlapping clip segmentation to better capture off-center contacts and multi-view camera setups to reduce occlusion-related errors and enhance spatial visibility.

Nevertheless, several limitations remain that should be addressed in future work. First, despite adaptation for youth hockey, the detection and tracking modules continue to face challenges due to frequent occlusions, smaller player sizes, inconsistent lighting, and unpredictable movement patterns. ID switches and missed detections are still observed, especially during substitutions and crowded play. Leveraging multi-view video or applying post-processing techniques, such as jersey number recognition, re-identification, or trajectory smoothing, may further improve tracking reliability. Second, although the TSM architecture offers strong computational efficiency, its ResNet backbone may limit sensitivity to fine-grained features. More expressive architectures such as Transformers (e.g., Vision Transformer (ViT), Swin Transformer) or ConvNeXt could provide improved spatial and temporal modeling (Dosovitskiy et al., 2020; Ze Liu et al., 2021; Zhuang Liu et al., 2022). Third, the dataset's ground-truth annotations were based on manual interpretation, which may introduce inconsistencies in contact severity labeling. Moreover, many annotated events were low-intensity and not clinically relevant, potentially inflating false positives. Future datasets should focus on high-severity, injury-relevant events to align predictions with meaningful outcomes.

Despite these limitations, the pipeline significantly improved operational efficiency. On average, it flagged approximately 370 candidate clips per game, reducing manual review time for building the head impact dataset from three hours to under 30 minutes. Critically, the TSM-based system successfully detected 86.4% of manually annotated head impacts, validating its utility as a reliable pre-filter for downstream injury classification. Detected contacts consistently had high prediction scores, while missed cases showed near-zero confidence, confirming the system's strength in identifying clear contact events and highlighting areas for refinement in detecting more subtle or occluded contacts.

Building on this enriched contact event database, the final stage of the pipeline, presented in Chapter 4 (Stage 3), extended the system to predict the biomechanical consequence of head impacts by calculating MPS directly from video-derived features. This step was important in moving beyond head impact classification toward estimating the actual strain imposed on brain tissue, which is a more injury-relevant metric. By predicting MPS using accessible video features such as velocity, surface compliance, impact location, and elevation, plus player age group, Stage 3 provided a scalable alternative to traditional sensor-based or finite element model (FEM)-dependent methods. It demonstrated that with sufficient input features, it is possible to estimate tissue-level strain with reasonable precision from video alone.

The model showed strong predictive performance, achieving a MSE of 0.0015 and an R^2 score of 0.89 on the test set. Visual inspection of predicted versus ground-truth MPS values revealed tight clustering around the identity line, suggesting that the model captured meaningful biomechanical patterns and generalized well across unseen samples. A permutation importance analysis confirmed that impact velocity was the most influential predictor: when velocity values were shuffled, model performance dropped markedly, with a ΔMSE of +0.0144 and a decrease in ΔR^2 of approximately 1.03. This finding aligns with established biomechanical literature showing that velocity strongly governs brain deformation and strain (Meaney and Smith, 2011; Mihalik et al., 2012).

Surface compliance ranked as the second most important input. Impacts onto rigid surfaces like ice or boards resulted in higher predicted MPS values compared to those involving more compliant materials, highlighting the role of surface properties in modulating energy transfer (Post and Hoshizaki, 2012; Karton and Hoshizaki, 2018; Gysland et al., 2012). Impact location, player age group, and elevation had comparatively minor influence in this dataset, possibly due to limited variability across these dimensions; however, their contribution may be more significant in datasets with greater heterogeneity in impact scenarios.

Importantly, the model demonstrated reliable generalization across all youth player categories (ages 5–6, 7–14, and 15–17), with MSE ranging from 0.0012 to 0.0024 and R^2 exceeding 0.85 across all subgroups. This robustness suggests that the model captured age-independent biomechanical relationships, making it broadly applicable in youth hockey contexts.

Nonetheless, the study has limitations. First, the dataset, though split into training, validation, and test subsets, was relatively small, especially in the validation and test sets, which could limit the statistical precision of performance estimates. Second, the ground-truth MPS values were derived from a single FE brain model with fixed material properties. Although scaled-down geometries were used to approximate age differences in head size, the material properties remained constant across age groups. This simplification may overlook important developmental differences in brain tissue biomechanics and thus limit the model's ability to fully capture age-specific strain responses.

Importantly, each stage in this progression addressed a limitation of the previous work. The head-impact detection stage (Stage 2, Chapter 3) highlighted the challenges of limited training data and the rarity of true head impacts. The contact-detection stage (Stage 1, Chapter 2) introduced a scalable solution through automated contact detection and dataset expansion, providing high-recall pre-filtering for subsequent head-impact classification. Building on these foundations, the brain-strain prediction stage (Stage 3, Chapter 4) added a crucial biomechanical layer by linking observable video features to tissue-level brain strain. Together, these three stages present a cohesive pipeline that can potentially be deployed for field-based monitoring of head injury risks in youth sports.

5.0.1 Future Work

This work lays the foundation for a fully automated, video-based system capable of identifying head impacts and estimating their biomechanical consequences. Several directions can be pursued to further enhance and scale this pipeline.

First, the automated contact detection framework developed in the contact-detection stage (Stage 1, Chapter 2) enables the large-scale generation of physical contact datasets from youth hockey videos. These datasets can be used to train a more robust and generalizable head impact detection network that overcomes the limitations of rare event frequency and manual annotation.

Second, a logical next step is to develop dedicated deep-learning models that automatically classify both impact location (e.g., front, side, rear) and surface compliance (e.g., boards, glass, ice). These classifiers could be trained using features extracted from head-centric video clips generated by a head detection and human pose estimation pipeline. Pose estimation would improve localization of the head relative to the body and surroundings, enhancing the precision of classification tasks. State-of-the-art models such as HRNet (Sun et al., 2019) and OpenPose (Cao et al., 2019) offer accurate joint detection even under occlusions and could be integrated into this workflow to extract meaningful kinematic features, such as head orientation or approach angle, that inform both location and surface classification.

Third, a colleague in our lab has developed a velocity-estimation pipeline that computes player speed directly from 2D video by combining homography-based rink localization with robust tracking (Dehghan et al., 2024). Incorporating this velocity estimation system will eliminate the need for manually annotated velocity values or external sensor input, closing a critical gap in feature extraction.

Ultimately, integrating all these components, contact detection, head impact classification, location and compliance inference, and velocity estimation, will enable a fully automated pipeline for predicting MPS from standard 2D hockey game footage. This would mark a substantial advancement in youth sports safety, enabling continuous, non-invasive, and large-scale monitoring of injury risk across games, teams, and developmental levels.

5.0.2 Conclusion

This thesis presents a three-stage, video-based deep learning pipeline for the detection and biomechanical interpretation of head impacts in youth ice hockey. The system was designed to overcome key challenges in data scalability, annotation efficiency, and biomechanical assessment.

The first component of this work, developed earliest in the research timeline and presented in Chapter 3 (Stage 2), focused on directly detecting head impacts using a LRCN, demonstrating the feasibility of learning impact-specific spatiotemporal patterns from video clips. However, limited training data and the need for manual preprocessing hindered model generalization and scalability. To address these issues, the contact-detection stage presented in Chapter 2 (Stage 1) expanded the focus to include all physical contact events. A YOLOv8 + StrongSORT pipeline was combined with a TSM classifier to automatically generate player-centered clips likely to contain contact. This allowed the system to serve as a scalable pre-filter for building large annotated datasets. To evaluate its utility for head impact identification, all head impacts in the full-game test set were manually labeled. Of the 22 observed head impacts, 19 were successfully captured by the contact detection system, yielding a recall of 86.4%. This supports the system's role as an effective filtering mechanism for injury surveillance, significantly reducing manual review while retaining the majority of relevant events. The third and final stage, presented in Chapter 4 (Stage 3), extended the pipeline to estimate brain tissue deformation via MPS, a validated biomechanical metric. Using a fully connected neural network trained on 477 reconstructed youth impacts, the model achieved high predictive performance ($R^2 = 0.89$, MSE $= 0.0015$). Permutation analysis confirmed that impact velocity was the most influential feature, followed by surface compliance. Other features such as impact location and elevation had a more modest effect, possibly due to lower variability in the dataset.

This thesis contributes a comprehensive, automated video-analysis framework for detecting, categorizing, and interpreting head impacts in youth sports. By combining CV, temporal modeling, and biomechanical inference, the pipeline provides a scalable alternative to traditional sensor-based monitoring and lays the groundwork for future large-scale studies of brain injury risk in youth athletics.

Key Contributions

- An integrated, multi-stage, video-only pipeline for youth ice hockey that: (i) detects contact events, (ii) filters likely head impacts, and (iii) estimates tissue–level strain (MPS) from standard 2D broadcast footage—without wearable sensors or time–consuming FE simulations.
- Youth–specific detection & tracking framework (YOLOv8 + StrongSORT with IoU–augmented cost) that **reduces ID switches by 69%** and increases MOTA to **94.5%**, maintaining player identity under heavy occlusion and crowding.
- Scalable, player–centric dataset generation through an automated detect–track–crop pipeline, enabling reproducible large–scale creation of training and testing data without manual intervention.
- High–recall contact pre–filter: a TSM classifier with empirically tuned parameters that **captures 86.4%** of head impacts in full games. Design guidelines recommend a 1.5× crop and low decision thresholds (< 0.20) to prioritize recall in safety–critical applications.
- Video–to–strain regression model: a dense network trained on **477** lab–reconstructed youth impacts predicts MPS with test performance of **MSE = 0.0015** and R^2 = **0.89**, generalizing across age groups ($R^2 \geq 0.85$; MSE ≤ 0.0024).
- Biomechanically interpretable outputs: permutation importance identifies **impact velocity** as the dominant driver of MPS, followed by **surface compliance**, linking observable on–ice mechanics to tissue–level brain response.

Future Work

1. **Scale** head–impact datasets via fully automated contact detection (multi–game ingestion), with domain adaptation for new arenas/cameras and multi–view fusion to mitigate occlusion.

2. **Develop** pose/part–based heads for automatic impact–*location* and surface–*compliance* inference (e.g., ConvNeXt, Swin, or ViT backbones), coupled with robust head localization.

3. **Integrate** the lab's homography–based velocity pipeline for automatic speed extraction and richer kinematic features.

4. **Unify** modules into a single system that maps raw broadcast video directly to calibrated MPS predictions with uncertainty estimates for decision support.

5. **Validate** externally (additional leagues/age groups) and assess fairness across demographics; target real–time inference for rink–side deployment and prospective concussion surveillance.

References

Agoston, Denes V and Dianne Langford (2017). "Big Data in traumatic brain injury; promise and challenges". In: *Concussion* 2.4, CNC44.

Anderson, Vicki et al. (2005). "Functional plasticity or vulnerability after early brain injury?" In: *Pediatrics* 116.6, pp. 1374–1382.

Bain, A. C. and D. F. Meaney (2000). "Tissue-level thresholds for axonal damage in an experimental model of central nervous system white matter injury". In: *Journal of Biomechanical Engineering* 122, pp. 615–622.

Barnes, Deborah E et al. (2018). "Association of mild traumatic brain injury with and without loss of consciousness with dementia in US military veterans". In: *JAMA neurology* 75.9, pp. 1055–1061.

Bazarian, Jeffrey J et al. (2010). "Sex differences in outcome after mild traumatic brain injury". In: *Journal of neurotrauma* 27.3, pp. 527–539.

Broglio, Steven P et al. (2011). "Cumulative head impact burden in high school football". In: *Journal of neurotrauma* 28.10, pp. 2069–2078.

Casper, Stephen T (2018). "Concussion: A History of Science and Medicine, 1870-2005". In: *Headache: The Journal of Head and Face Pain* 58.6, pp. 795–810.

Chatelin, Simon, André Constantinesco, and Rémy Willinger (2010). "Fifty years of brain tissue mechanical testing: from in vitro to in vivo investigations". In: *Biorheology* 47.5-6, pp. 255–276.

Chen, Wesley et al. (2023). "A comparison of frequency and magnitude of head impacts between Pee Wee and Bantam youth ice hockey". In: *Sports biomechanics* 22.6, pp. 728–751.

Churchill, Nathan W et al. (2017). "Neuroimaging of sport concussion: persistent alterations in brain structure and function at medical clearance". In: *Scientific reports* 7.1, p. 8297.

Cournoyer, J et al. (2021). "Brain trauma exposure for American tackle football players 5 to 9 and 9 to 14 years of age". In: *Journal of biomechanics* 127, p. 110689.

Daneshvar, Daniel H et al. (2011). "The epidemiology of sport-related concussion". In: *Clinics in sports medicine* 30.1, pp. 1–17.

DavenportElizabeth, M et al. (2014). "Abnormal white matter integrity related to head impact exposure in a season of high school varsity football". In: *Journal of neurotrauma*.

Fann, Jesse R, Tessa Hart, and Katherine G Schomer (2009). "Treatment for depression after traumatic brain injury: a systematic review". In: *Journal of neurotrauma* 26.12, pp. 2383–2402.

Gabler, Lee F, Jeff R Crandall, and Matthew B Panzer (2016). "Assessment of kinematic brain injury metrics for predicting strain responses in diverse automotive impact conditions". In: *Annals of biomedical engineering* 44, pp. 3705–3718.

— (2019). "Development of a second-order system for rapid estimation of maximum brain strain". In: *Annals of biomedical engineering* 47, pp. 1971–1981.

Giordano, Chiara and Svein Kleiven (2014). *Evaluation of axonal strain as a predictor for mild traumatic brain injuries using finite element modeling.* Tech. rep. SAE Technical Paper.

Giza, Christopher C and David A Hovda (2014). "The new neurometabolic cascade of concussion". In: *Neurosurgery* 75, S24–S33.

Greenwald, Richard M et al. (2008). "Head impact severity measures for evaluating mild traumatic brain injury risk exposure". In: *Neurosurgery* 62.4, p. 789.

Hernandez, Fidel et al. (2015). "Six degree-of-freedom measurements of human mild traumatic brain injury". In: *Annals of biomedical engineering* 43, pp. 1918–1934.

Ho, Johnson and Svein Kleiven (2007). "Dynamic response of the brain with vasculature: a three-dimensional computational study". In: *Journal of biomechanics* 40.13, pp. 3006–3012.

Horgan, Timothy James and Michael D Gilchrist (2003). "The creation of three-dimensional finite element models for simulating head impact biomechanics". In: *International Journal of Crashworthiness* 8.4, pp. 353–366.

Jadischke, Ron et al. (2013). "On the accuracy of the Head Impact Telemetry (HIT) System used in football helmets". In: *Journal of biomechanics* 46.13, pp. 2310–2315.

Kamins, Joshua et al. (2017). "What is the physiological time to recovery after concussion? A systematic review". In: *British journal of sports medicine* 51.12, pp. 935–940.

Karton, Clara, T Blaine Hoshizaki, and Michael D Gilchrist (2020). "A novel repetitive head impact exposure measurement tool differentiates player position in National Football League". In: *Scientific reports* 10.1, p. 1200.

Karton, Clara and T Blaine Hoshizaki (2018). "Concussive and subconcussive brain trauma: the complexity of impact biomechanics and injury risk in contact sport". In: *Handbook of clinical neurology* 158, pp. 39–49.

Karton, Clara, Thomas Hoshizaki, and Michael Gilchrist (2014). “The influence of impactor mass on the dynamic response of the Hybrid III headform and brain tissue deformation”. In: *Mechanism of concussion in sports*. ASTM International.

Karton, Clara and Thomas Blaine Hoshizaki (2021). “Biomechanics of sport-related neurological injury”. In: *Clinics in Sports Medicine* 40.1, pp. 19–38.

Karton, Clara, Andrew Post, et al. (2021). “Exposure to brain trauma in six age divisions of minor ice hockey”. In: *Journal of biomechanics* 116, p. 110203.

Kimpara, Hideyuki and Masami Iwamoto (2012). “Mild traumatic brain injury predictors based on angular accelerations during impacts”. In: *Annals of biomedical engineering* 40, pp. 114–126.

Kleiven, Svein (2007). *Predictors for traumatic brain injuries evaluated through accident reconstructions*. Tech. rep. SAE Technical Paper.

Kroshus, Emily et al. (2015). “Concussion under-reporting and pressure from coaches, teammates, fans, and parents”. In: *Social science & medicine* 134, pp. 66–75.

McAllister, Thomas W (2010). “Genetic factors modulating outcome after neurotrauma”. In: *PM&R* 2.12, S241–S252.

McHenry, Brian G (2004). “Head injury criterion and the ATB”. In: *ATB Users' group* 29, pp. 5–8.

Meehan, William P, Pierre d'Hemecourt, and R Dawn Comstock (2010). “High school concussions in the 2008-2009 academic year: mechanism, symptoms, and management”. In: *The American journal of sports medicine* 38.12, pp. 2405–2409.

Mihalik, Jason P et al. (2012). “Head impact biomechanics in youth hockey: comparisons across playing position, event types, and impact locations”. In: *Annals of biomedical engineering* 40, pp. 141–149.

Newman, J et al. (2000). “A new biomechanical assessment of mild traumatic brain injury. Part 2: results and conclusions”. In: *Proceedings of the International Research Council on the Biomechanics of Injury conference*. Vol. 28. International Research Council on Biomechanics of Injury.

Newman, James A (1986). “A generalized acceleration model for brain injury threshold (GAMBIT)”. In: *Proceedings of International Conference on the Biomechanics of Impact, 1986*, pp. 121–131.

Nowinski, Christopher J et al. (2024). “'Subconcussive'is a dangerous misnomer: hits of greater magnitude than concussive impacts may not cause symptoms”. In: *British Journal of Sports Medicine*.

Patton, Declan A et al. (2012). “Injury data from unhelmeted football head impacts evaluated against critical strain tolerance curves”. In: *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology* 226.3-4, pp. 177–184.

Post, A et al. (2019). “An examination of the brain trauma in Novice and Midget ice hockey: Implications for helmet innovation”. In: *CMBES Proceedings* 42.

Post, Andrew and T Blaine Hoshizaki (2015). “Rotational acceleration, brain tissue strain, and the relationship to concussion”. In: *Journal of biomechanical engineering* 137.3, p. 030801.

Post, Andrew and T Blaine Hoshizaki (2012). “Mechanisms of brain impact injuries and their prediction: a review”. In: *Trauma* 14.4, pp. 327–349.

Post, Andrew, T Blaine Hoshizaki, Michael D Gilchrist, et al. (2014). “The influence of dynamic response and brain deformation metrics on the occurrence of subdural hematoma in different regions of the brain”. In: *Journal of neurosurgery* 120.2, pp. 453–461.

Post, Andrew, T Blaine Hoshizaki, Clara Karton, et al. (2019). “The biomechanics of concussion for ice hockey head impact events”. In: *Computer methods in biomechanics and biomedical engineering* 22.6, pp. 631–643.

Post, Andrew, Marshall Kendall, et al. (2015). “Characterization of persistent concussive syndrome using injury reconstruction and finite element modelling”. In: *Journal of the mechanical behavior of biomedical materials* 41, pp. 325–335.

Putukian, Margot et al. (2013). “Onfield assessment of concussion in the adult athlete”. In: *British journal of sports medicine* 47.5, pp. 285–288.

Raul, Jean-Sébastien et al. (2008). “Finite-element models of the human head and their applications in forensic practice”. In: *International journal of legal medicine* 122, pp. 359–366.

Robidoux, Michael A et al. (2020). “Comparing concussion rates as reported by hockey Canada with head contact events as observed across minor ice-hockey age categories”. In: *Journal of concussion* 4, p. 2059700220911285.

Rodger, James A (2015). “Discovery of medical Big Data analytics: Improving the prediction of traumatic brain injury survival rates by data mining Patient Informatics Processing Software Hybrid Hadoop Hive”. In: *Informatics in Medicine Unlocked* 1, pp. 17–26.

Rowson, Steven and Stefan M Duma (2013). “Brain injury prediction: assessing the combined probability of concussion using linear and rotational head acceleration”. In: *Annals of biomedical engineering* 41, pp. 873–882.

Rowson, Steven, Stefan M Duma, et al. (2012). “Rotational head kinematics in football impacts: an injury risk function for concussion”. In: *Annals of biomedical engineering* 40, pp. 1–13.

Rueda, MA Forero, Liang Cui, and Michael D Gilchrist (2011). “Finite element modelling of equestrian helmet impacts exposes the need to address rotational kinematics in future helmet designs.” In: *Comput Methods Biomech Biomed Engin* 14.12, pp. 1021–1031.

Stulemeijer, Maja et al. (2010). "How mild traumatic brain injury may affect declarative memory performance in the post-acute stage". In: *Journal of neurotrauma* 27.9, pp. 1585–1595.

Takhounts, Erik G et al. (2013). *Development of brain injury criteria (BrIC)*. Tech. rep. SAE Technical Paper.

Trotta, Antonia et al. (2020). "Biofidelic finite element modelling of brain trauma: Importance of the scalp in simulating head impact". In: *International Journal of Mechanical Sciences* 173, p. 105448.

USA Hockey (2024). *2023-24 Membership Statistics*. https://www.usahockey.com/membershipstats. Retrieved from USA Hockey website.

Wilcox, Bethany J et al. (2014). "Head impact exposure in male and female collegiate ice hockey players". In: *Journal of biomechanics* 47.1, pp. 109–114.

Willinger, Remy and Daniel Baumgartner (2003). "Human head tolerance limits to specific injury mechanisms". In: *International journal of Crashworthiness* 8.6, pp. 605–617.

Yurgil, Kate A et al. (2014). "Association between traumatic brain injury and risk of posttraumatic stress disorder in active-duty Marines". In: *JAMA psychiatry* 71.2, pp. 149–157.

Cao, Zhe et al. (2019). "Openpose: Realtime multi-person 2d pose estimation using part affinity fields". In: *IEEE transactions on pattern analysis and machine intelligence* 43.1, pp. 172–186.

Dehghan, Parisa et al. (2024). "Object Detection for Ice Surface Localisation in Youth Hockey". In.

Donahue, Jeffrey et al. (2015). "Long-term recurrent convolutional networks for visual recognition and description". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2625–2634.

Dosovitskiy, Alexey et al. (2020). "An image is worth 16x16 words: Transformers for image recognition at scale". In: *arXiv preprint arXiv:2010.11929*.

Gysland, Sonia M et al. (2012). "The relationship between subconcussive impacts and concussion history on clinical measures of neurologic function in collegiate football players". In: *Annals of biomedical engineering* 40, pp. 14–22.

Karton, Clara and T Blaine Hoshizaki (2018). "Concussive and subconcussive brain trauma: the complexity of impact biomechanics and injury risk in contact sport". In: *Handbook of clinical neurology* 158, pp. 39–49.

Lin, Ji, Chuang Gan, and Song Han (2019). "Tsm: Temporal shift module for efficient video understanding". In: *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 7083–7093.

Liu, Ze et al. (2021). “Swin transformer: Hierarchical vision transformer using shifted windows”. In: *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 10012–10022.

Liu, Zhuang et al. (2022). “A convnet for the 2020s”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 11976–11986.

Meaney, David F and Douglas H Smith (2011). “Biomechanics of concussion”. In: *Clinics in sports medicine* 30.1, pp. 19–31.

Mihalik, Jason P et al. (2012). “Head impact biomechanics in youth hockey: comparisons across playing position, event types, and impact locations”. In: *Annals of biomedical engineering* 40, pp. 141–149.

Post, Andrew and T Blaine Hoshizaki (2012). “Mechanisms of brain impact injuries and their prediction: a review”. In: *Trauma* 14.4, pp. 327–349.

Sun, Ke et al. (2019). “Deep high-resolution representation learning for human pose estimation”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5693–5703.