Can a novel computer vision-based framework detect head-on-head impacts during a rugby league tackle?

Manish Mohan[1], Dan Weaving[1,2,3], Andrew J Gardner[1,4], Sharief Hendricks[1,5], Keith Stokes[6,7], Gemma Phillips[1,8,9] Matt Cross[1,8], Cameron Owen[1,8], Ben Jones[1,5,8,10,11]

[1] Carnegie Applied Rugby Research (CARR) centre, Carnegie School of Sport, Leeds Beckett University, Leeds, UK

[2] Department of Physical Activity and Sport, Faculty of Arts and Sciences, Edge Hill University, Ormskirk, United Kingdom

[3] Applied Sports Science and Exercise Testing Laboratory, The University of Newcastle, Ourimbah, NSW, Australia

[4] Sydney School of Health Sciences, Faculty of Medicine and Health, The University of Sydney, Camperdown, NSW, 2006, Australia.

[5] Division of Physiological Sciences and Health through Physical Activity, Lifestyle and Sport Research Centre, Department of Human Biology, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa

[6] Centre for Health, and Injury & Illness Prevention in Sport, University of Bath, Bath, UK

[7] Rugby Football Union, Twickenham, UK

[8] England Performance Unit, Rugby Football League, Manchester, UK

[9] Hull Kingston Rovers, Hull, United Kingdom

[10] Premiership Rugby, London, United Kingdom

[11] School of Behavioural and Health Sciences, Faculty of Health Sciences, Australian Catholic University, Brisbane, QLD, Australia

Corresponding author:
Manish Mohan
Carnegie Applied Rugby Research (CARR) Centre,
Carnegie School of Sport,
Leeds Beckett University,
Leeds,
United Kingdom
m.mohan@leedsbeckett.ac.uk

Word count 3209; Tables 1; Figures 1; Supplementary Tables 0; Supplementary Figures 0

Can a novel computer vision-based framework detect head-on-head impacts during a rugby league tackle?

Abstract

Background: Head-on-head impacts are a risk factor for concussion, which is a concern for sports. Computer vision frameworks may provide an automated process to identify head-on-head impacts, although this has not been applied or evaluated in rugby.

Methods: This study developed and evaluated a novel computer vision framework to automatically classify head-on-head and non-head-on-head impacts. Tackle events from professional rugby league matches were coded as either head-on-head or non-head-on-head impacts. These included non-televised standard-definition and televised high-definition video clips to train (n=341) and test (n=670) the framework. A computer vision framework consisting of two deep learning networks, an object detection algorithm and 3D Convolutional Networks, was employed and compared to the analyst coded criterion. Sensitivity, specificity, and positive predictive value were reported.

Results: The overall performance evaluation of the framework to classify head-on-head impacts against manual coding had a sensitivity, specificity, and positive predictive value (95% confidence intervals) of 68% (58% to 78%), 84% (78% to 88%) and 0.61 (0.54 to 0.69) in standard-definition clips, and 65% (55% to 75%), 84% (79% to 89%) and 0.61 (0.53 to 0.68) in high-definition clips.

Conclusion: The study introduces a novel computer vision framework for head-on-head impact detection. Governing bodies may also use the framework in real-time, or for retrospective analysis of historical videos, to establish head-on-head rates and evaluate prevention strategies. Future work should explore the application of the framework to other head-contact mechanisms, and also the utility in real-time to identify potential events for clinical assessment.

What is already known on this topic

Head contact and collisions are a risk for concussion and head injuries. Computer vision techniques exist in other disciplines to automatically identify heads. Applying this to rugby, could as a practical solution, enabling stakeholders like governing bodies to assess the frequency, timestamp, and incidence of such events during matches using video-based data.

What this study adds

The study introduces and validates a novel computer vision framework for detecting head-on-head impacts in rugby league. Despite a sensitivity of 68%, the framework shows competitive sensitivity compared to wearable sensors currently being used in collision sports.

How this study might affect research, practice, or policy

The framework enables non-intrusive monitoring of potential head injuries, eliminating the need for player cooperation, unlike wearable sensors. Given the lack of alternative methods for retrospective analysis aside from manual video analysis, the framework can be applied to create datasets related to head-to-head impact exposure rates. These datasets can inform research organizations and governing bodies in shaping policies and evaluation ongoing interventions.

Introduction

Head impact exposures are a concern for many sports (e.g., rugby, soccer, American football) due to their potential to increase the susceptibility to negative clinical outcomes such as concussion[1] and/or neurodegenerative disease in elite-level athletes.[2] Rugby League is an international full-contact sport[3] with relatively high rates of concussion and head impacts during match play.[4] During match play, head impacts with the potential for concussion are identified by a side-line or an offsite video review process, or via direct observation. As part of the in-game evaluation process players suspected of having sustained a concussion are removed from play for a Head Injury Assessment (HIA).[5] Tackle-related HIAs and medically diagnosed concussions occur most commonly from head-on-head impacts, followed by head-to-shoulder, and head-to-hip impacts.[6, 7] Following the retrospective review of tackle events, there are some occasions where players may have required an HIA.[8] Therefore, establishing valid, reliable, and practically feasible methods for identifying head impact exposure from video footage is likely beneficial both for later HIA review and longer-term head impact exposure monitoring.

An approach to identifying head contact events that have been increasingly adopted is the use of wearable sensors[9] which are typically embedded within helmets (e.g., Head Impact Telemetry System [HITS])[10]) or instrumented mouthguards.[11] These can measure both the kinematics and frequency of head impacts or head accelerations.[11] However, there are currently several barriers to the adoption of wearable sensors for real-time identification in sports. Manual video verification is often required to confirm sensor-recorded head impacts.[12] Wearable sensors are also dependent on financial and logistical resources, limiting their feasibility. Significant research advancements have been made in the application of head detection and head tracking using computer vision in surveillance systems.[13] To overcome the challenges of wearable sensors and given their recent advancements, deep learning and computer vision techniques can be developed for sports application. This may provide an additional, accessible, and scalable method to identify the frequency of head impacts. This has previously been demonstrated in the identification of soccer headers during match play, achieving 21% precision.[14]

Rugby league has not investigated the capability of computer vision to automatically identify when a head-on-head impact occurs between two players. This is important as head-on-head impacts have the highest risk for HIAs.[6] The capability to automatically detect these events poses a possible practical solution for rugby league and other sports to allow stakeholders (e.g., match officials, doctors) to evaluate the frequency, timestamp, and incidence of head-on-head events that occur during a match using only video-based data. Therefore, this study aims to introduce a novel computer vision framework for automatic head-on-head impact classification. It also aims to establish the validity of the framework to correctly classify head-on-head impacts against a video analyst criterion using single-view broadcast video from professional rugby league matches.

Methods

In this study, a framework to detect head-on-head impacts from videos is proposed and evaluated against video analysis. The proposed framework consists of two deep-learning models: the head detection model and the impact classification model (Figure 1c). The framework takes video as input, uses a head detection model to detect players' head frame-by-frame and then classifies an impact. In the current study, this framework was applied to both standard definition (SDTV) and high definition (HDTV). The details of the two deep-learning models that compose the framework and datasets used to train and evaluate each model and framework are described below. Ethics approval was gained from the institutions' review board for the use of publicly available video data (REF: 100411). Patients and/or the public were not involved in the design, conduct, reporting, or dissemination plans of this research.

******Insert Figure 1 here******

Tackle events from rugby league matches (2021 European Super League season and 2020-2022 Australian National Rugby League [NRL] season) using non-televised standard-definition (SDTV) and televised high-definition (HDTV) were coded by an experienced (>10 years) rugby league video analyst as either head-on-head or non-head-on-head impact events and randomly selected to train and test the framework (Figure 1b). The analyst's coding of head-on-head impacts was deemed as the criterion measure to evaluate the validity of the automated framework allowing the determination of true-positive, false-positive, true-negative, and false-negative performance of the framework. A head-on-head impact was defined as direct contact by one player's head on another player's head (opponent or teammate), resulting in a visible change in the trajectory of the head. A non-head-on-head impact was defined as a tackle event where the players do not have a head-on-head collision. To establish the intra-rater reliability of the analyst to code a head-on-head impact or a non-head-on-head impact, a random sample of 50 videos were re-coded following a 30-day washout period. Reliability statistics of 0.94 Kappa with *almost perfect* agreement were observed.[15]

The required sample size and test-sets allocation were computed for a positive predicted value (PPV) with power assumed as 80% at the 5% level.[16] The anticipated sensitivity and specificity for the calculation of sample size were assumed as 0.70 and 0.75 respectively.[17] The prevalence in the population of interest was estimated by randomly sampling 13 matches from the 2021 Super League season and determining the frequency of head-on-head impacts (n = 90) and the total number of tackles (n = 4191) to determine their ratio (prevalence: 90/4191 = 0.02). Consequently, to achieve the power of 80% with a prevalence of 0.02 and an anticipated sensitivity and specificity of 0.70 and 0.75 respectively, the total number of samples required is ≥335 with the number of positives ≥92 and the number of negatives ≥243.[16]

Two samples were created to train and test the validity of the framework (Figure 1b). The SDTV sample included videos recorded at a frame size of 1280x720 pixels and a frame rate of 25 frames-per-second (fps). The HDTV sample included videos recorded at a higher frame size of 1920x1080 and a frame rate of 25 fps. These were created to explore the potential effects of video quality on the performance of the framework. The number of total videos with their number of samples in the train and test dataset split are also shown in Figure 1b.

The head detection model takes an input image and outputs the bounding box coordinates of an object (head) by detecting its instance within the image. Dataset for the training of Head detection model – CrowdHuman is an annotated and large human and head detection dataset consisting of 19,370 images (15,000, 4,370 for training and validation respectively). [18] The CrowdHuman has ~340k heads annotated in the 19,370 images of the dataset which is considerably larger than other available datasets. Hence CrowdHuman dataset was chosen for training the head-detection model.

In this study, the You only look once version 7 (YOLOv7) object detector[19] which has the highest accuracy among all known real-time object detectors[20] was used to detect players' heads using bounding boxes. The model was trained on the CrowdHuman dataset for 300 epochs with a batch-size of 8 and each image was resized to 640x480 pixels using OpenCV package[21] in Python. The number of classes was also changed to 1 (head) from the 80 in the original model. The calculation of new anchor boxes was done using k-means.[22] The performance of the YOLOv7 model in CrowdHuman was evaluated by using mean average precision (mAP)[23] and the performance of YOLOv7 on the SDTV and HDTV samples was also evaluated by using the mean average precision against random 1000 images from the total images annotated manually using Microsoft Visual Object Tagging Tool (VoTT)[24] to establish the true bounding box coordinates of players' head. A mean average precision calculated the average precision value for recall value range over 0 to 1 with values closer to 1 considered excellent.

In the current study 88,000 images from 352 training videos across both samples were annotated automatically to establish the bounding box coordinates of the player's head in each image using the above-trained YOLOv7 model. These images encompassed a wide variety of scenarios, ranging from various stadiums, lighting situations, camera angles, player-head sizes, positions, and skin tones to instances of players wearing head protective gear, as well as images featuring partially occluded player-heads.

The tackle videos with head-on-head impacts (219 in SDTV and 41 in HDTV [Figure 1b]) were used to find the impact image. From the 260 selected videos, an array of 9 images (the image frame containing the impact plus 4 frames before the impact frame, and 4 frames ahead of the impact frame) were filtered resulting in 2,300 images. Finally, the images were split randomly into training and validation sets and used to train and validate the impact detection model. The input data for the classification model was head detection from 9 images and head detections were cropped to 128x128 size to improve processing speed. Several methods of image-data augmentation such as horizontal flip, coarse dropout, grid mask, shift scale, rotate, blur and random brightness contrast were used to prevent the possibility of overfitting. Lastly to classify head-on-head or non-head-on-head impacts, a 3D Convolutional Neural Network (CNN)[25] was used because of the movement of the kernel in three dimensions to capture the depth. It is crucial as it enables the model to extract spatiotemporal features effectively, capturing the dynamic nature of head movements and enhancing the accuracy of head-on-head impact classification.

The dataset comprises cases, specifically head-on-head impacts, and controls, which are non-head-on-head impacts. The framework's evaluation involves calculating sensitivity, specificity, and positive predictive value (PPV) based on true positives, false positives, true negatives, and false negatives. These values are determined by comparing the classifier's output against the video analyst's coding of head-on-head and non-head-on-head impacts. The sensitivity, specificity, and positive predictive value (PPV) range over 0 to 1 with values closer to 1 considered excellent.

Results

The performance of the head detection model in the CrowdHuman dataset used to train the model for automatic annotation of a player's head had a mean average precision of 0.752.

Secondly, the performance of the model in the SDTV and HDTV samples against the random 500 images was similar with mean average precision at 0.751 and 0.771, respectively.

Table 1 presents the results of the validity analysis. Sensitivity, specificity, and PPV was 68% (58% to 78%), 84% (78% to 88%) and 0.61 (0.54 to 0.69) in SDTV, and 65% (55% to 75%), 84% (79% to 89%) and 0.61 (0.53 to 0.68) in HDTV video clips using the Impact Detection Framework.

******Insert Table 1 here******

******Insert Figure 2 here******

Discussion

In this study, a novel computer vision framework for the detection of head-on-head impacts was applied for the first time and validated using rugby league tackle videos. The proposed framework uses only single view video and deep learning models to identify head-on-head impacts at a sensitivity of 68% (66% to 71%). The most comparable previous computer vision study to head-on-head identification carried out in rugby is the automatic detection of high-risk tackles (e.g., a tackle that resulted in an HIA) using a computer vision framework performed at 62.5%[26] and 50%[27] respectively. The head-on-head impact detection framework could have outperformed the rugby studies because it did not rely on pose estimation models. Avoiding pose estimation models eliminates the potential inaccuracies with estimating player positions in tackle scenes due to occlusions, leading to more accurate and efficient detection of impacts. However, the drawback is that this approach is limited to detecting only head-on-head impacts and cannot identify all types of impacts. In soccer, the automated detection of headers using a computer vision framework performed at a 21.1% precision.[14] The framework exhibits superior performance than other studies using computer vision in sports. Therefore, the current framework has promise as a low-cost practical solution to identify head-on-head impacts that researchers and governing bodies could use for head-on-head impact monitoring.

In contrast to computer vision, instrumented mouthguards have a sensitivity of 40 to 75% for head acceleration events during rugby league tackle events.[11] The head-on-head Impact Detection framework in the current study had better sensitivity performance than two instrumented mouthguard products (40% and 51%).[11] However, the sensitivity in the instrumented mouthguard study was for all rugby league impact events (i.e., not only head-on-head). All mouthguards would likely have higher sensitivity than computer vision when only head-on-head impacts are considered. This is because the magnitude of head accelerations during head-on-head events will likely be higher than when considering all rugby league tackle events. These findings suggest that further development is required for the framework to replace wearable sensors such as instrumented mouthguards. However, it has applications where wearable sensors cannot be used such as for retrospective analysis of historical video data by research organizations and governing bodies. To address the false positives and negatives in resource-limited longitudinal monitoring, statistical models like Bayesian methods can be utilized.[29] These models can incorporate uncertainty and error rates into their calculations, providing more robust head-on-head rate changes over time.

The false positive rate (SDTV 0.16 and HDTV 0.17; Table 1)and false negative rate (SDTV 0.32 and HDTV 0.35; Table 1) shown in this study suggest that the current framework will underestimate rather than overestimate the number of head-on-head impacts. The current false-positive and false-negative rates highlight that the model cannot be used as a standalone method for detecting head-on-head impacts in rugby league. However, with the availability of pitch side review in rugby codes, the current model could be implemented to support the identification of potential head-on-head events that may be missed due to human error. The necessity for further refinement and rigorous validation to ensure reliable performance in real-world scenarios is required for its use as a standalone method to identify possible injury. Possible explanations for the false negative performance of the current framework can be attributed to the single-view broadcast footage that changes constantly to follow the ball. This causes blurry images and provides limited video footage at an individual player's level. This can be further influenced by the focus of the players in the video recording (e.g., camera does not appropriately zoom in on the tackle event). Therefore, it is recommended that higher frames per second (e.g., 60 fps vs the 25 fps used in this study) are used for video recordings to reduce blurry images, and also the video capture uses an appropriate zoom on each tackle event which would likely improve false positive performance.[30] A possible solution is also integrating the framework live with commercially available higher rate video analysis software for pitch side analysis. For example, World Rugby uses higher definition and greater frames per second (FPS) than standard broadcast footage and holds the potential to significantly enhance performance. The improved quality and granularity of images could lead to more accurate detection and classification of head impacts, thus further reducing false positives and negatives.

Although the introduction of HIA check or video review was important for player welfare, there is a need to continually enhance systems and procedures by providing solutions for current deficiencies such as a doctor or pitch-side reviewer missing a head-on-head impact and potential HIA event. It is important to highlight that the false-positive identification of head-on-head impacts by the framework could be easily addressed by the practitioners through a quick secondary review. While the existing framework may not be a completely automated solution for stakeholders in its current form, it would have the utility to flag events for possible HIAs retrospectively to minimize the probability of overlooking an HIA incident through human review alone. However, the frequency that the framework identifies such events above that identified by human review requires future investigation. Future work should also explore how the framework can be used to flag events for HIA check in real-time. There are also situations whereby there is no sensor data available, and the only method of generating head impact exposure data within-matches and between-matches is to manually video analyse their presence/absence during each tackle event which is extremely labour intensive. Equally, there is often a lack of lack of resources in governing bodies for retrospective video analysis. This can limit data generation to cross-sectional evaluations of exposure. Therefore, the framework allows practitioners to oversee possible head injuries in a non-intrusive manor, eliminating the need for player cooperation unlike wearable sensors. Given that historical video data is stored, it is possible that the framework could provide an approach to generating retrospective and longitudinal datasets of head-on-head exposure rates. These datasets can then be utilized by research organizations and governing bodies to inform policies and provide insights into head-on-head exposure rates and how they might change over time. However, future work is needed to consider the false positive/negative rates identified in the current study, and how these can be considered in the analysis using statistical models like Bayesian methods[29] and interpretation of changes in exposure over time using this framework.

Given this is the first to apply computer vision in rugby league for detection of the head-on-head impacts, there are several limitations and considerations for future development to consider. In this study, the YOLOv7 head detection model outperforms another published object detector by ~0.15[31] with a mean average precision of 0.752 in the CrowdHuman sample, as well as SDTV and HDTV samples. The study findings agrees that the current leading object detector for head detection is YOLOv7 and can be used to automate the annotations of players' heads. The use of YOLOv7, which provides automated head detection, can also save time by avoiding manual annotation of players' heads in videos to train the head detection model. However, the detection accuracy of the model could be further improved by using convolutional filters of larger receptive size to model context.[32] This would help in crowded scenes such as during a rugby league tackle, as players' heads often resemble inconsequential objects (with no clear boundaries of players' heads) and therefore appear homogeneous. In addition, context-sensitive prediction module (CPM) augmentation on top of each Feature Pyramid Network could also be used to improve head detection.[33] This is because the detection of players' heads in a rugby league tackle is considered as the detection of many small-sized adjacently placed objects and the use of transpose convolution on features[34] has shown promise in the broader head detection literature. Such approaches could assist in further improving the performance of the head detection model and therefore improve the performance of the head-on-head impact detection model. Occlusion of players' heads in the tackle scene is also a major limitation, as it impacts the performance of head detection. Hence, increasing the number of camera angles used to record game play will reduce occlusion of players and improve the performance of head detection.

Another inherent limitation of this framework is its inability to yield impact kinematics. A potential solution with promise involves combining wearable sensors like instrumented mouthguards with the suggested video-based framework. Comprehending head impact exposure requires a grasp of both frequency and kinematic details alongside video context (e.g., role of player, impacting object) that are not measurable by instrumented mouthguards. The framework could also be used by researchers and video analysts to streamline sensor verification and sensors can measure the head kinematics of verified impacts. The current solution also falls short in delivering exposure at the level of individual players due to its incapacity to recognize the players participating in each tackling event. Hence, aside from eliminating false-positive and false-negative identifications, additional video analysis would be required to calculate exposure for individual players. In the present study, the framework solely captures head-on-head impacts that take place within the context of tackle events. Despite this event being the primary source of HIAs,[6] impacts caused by other areas (such as the shoulder or knee) and impacts outside the camera's field of view will remain undetected, possibly resulting in an under-reporting of head impact frequency.

Conclusion

The study introduces a novel computer vision framework for detecting head-on-head impacts in rugby league videos. The framework exhibits superior performance at 68% sensitivity compared to other computer vision studies in sports. The framework enables non-intrusive monitoring for potential head injuries by eliminating the requirement for player cooperation, unlike wearable sensors. It provides efficient head detections, saving manual annotation time. In the absence of alternative methods for retrospective analysis beyond manual video analysis, the framework offers a cost-effective solution for generating datasets on head-to-head impact exposure rates. These datasets will offer valuable information for governing bodies to shape policies and gain insights into head-on-head exposure rates. In conclusion, the study provides data on the strengths and limitations of a computer vision framework for head-on-head impact detection during a rugby league tackle.

References

1. Eliason PH, Galarneau J-M, Kolstad AT, Pankow MP, West SW, Bailey S, et al. Prevention strategies and modifiable risk factors for sport-related concussions and head impacts: a systematic review and meta-analysis. British journal of sports medicine. 2023;57(12):749-61.
2. Iverson GL, Castellani RJ, Cassidy JD, Schneider GM, Schneider KJ, Echemendia RJ, et al. Examining later-in-life health risks associated with sport-related concussion and repetitive head impacts: a systematic review of case-control and cohort studies. British journal of sports medicine. 2023;57(12):810.
3. Rennie G, Hart B, Dalton-Barron N, Weaving D, Williams S, Jones B. Longitudinal changes in Super League match locomotor and event characteristics: A league-wide investigation over three seasons in rugby league. Plos one. 2021;16(12):e0260711.
4. Fitzpatrick AC, Naylor AS, Myler P, Robertson C. A three-year epidemiological prospective cohort study of rugby league match injuries from the European Super League. Journal of science and medicine in sport. 2018;21(2):160-5.
5. Gardner AJ, Iverson GL, Quinn TN, Makdissi M, Levi CR, Shultz SR, et al. A preliminary video analysis of concussion in the National Rugby League. Brain injury. 2015;29(10):1182-5.
6. Gardner AJ, Iverson GL, Edwards S, Tucker R. A case–control study of tackle-based head injury assessment (HIA) risk factors in the National Rugby League. Sports Medicine-Open. 2021;7(1):1-11.
7. Spiegelhalter M, Scantlebury S, Heyward O, Hendricks S, Cummins C, Gardner AJ, et al. The propensity of non-concussive and concussive head contacts during elite-level women's rugby league matches: A prospective analysis of over 14,000 tackle events. Journal of science and medicine in sport. 2023;26(3):195-201.
8. Gardner AJ, Kohler R, McDonald W, Fuller GW, Tucker R, Makdissi M. The use of sideline video review to facilitate management decisions following head trauma in super rugby. Sports medicine-open. 2018;4:1-8.
9. Patton DA, Huber CM, Jain D, Myers RK, McDonald CC, Margulies SS, et al. Head impact sensor studies in sports: a systematic review of exposure confirmation methods. Annals of biomedical engineering. 2020;48(11):2497-507.
10. Beckwith JG, Greenwald RM, Chu JJ. Measuring head kinematics in football: correlation between the head impact telemetry system and Hybrid III headform. Annals of biomedical engineering. 2012;40(1):237-48.
11. Jones B, Tooby J, Weaving D, Till K, Owen C, Begonia M, et al. Ready for impact? A validity and feasibility study of instrumented mouthguards (iMGs). British journal of sports medicine. 2022.
12. Cortes N, Lincoln AE, Myer GD, Hepburn L, Higgins M, Putukian M, et al. Video analysis verification of head impact events measured by wearable sensors. The American journal of sports medicine. 2017;45(10):2379-87.
13. Kang S-K, Chung K-Y, Lee J-H. Development of head detection and tracking systems for visual surveillance. Personal and ubiquitous computing. 2014;18:515-22.
14. Rezaei A, Wu LC. Automated soccer head impact exposure tracking using video and deep learning. Scientific reports. 2022;12(1):9282.
15. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Physical therapy. 2005;85(3):257-68.
16. Steinberg DM, Fine J, Chappell R. Sample size for positive and negative predictive value in diagnostic research using case–control designs. Biostatistics. 2009;10(1):94-105.
17. Campbell KR, Marshall SW, Luck JF, Pinton GF, Stitzel JD, Boone JS, et al. Head impact telemetry system's video-based impact detection and location accuracy. Medicine and science in sports and exercise. 2020;52(10):2198.
18. Shao S, Zhao Z, Li B, Xiao T, Yu G, Zhang X, et al. Crowdhuman: A benchmark for detecting human in a crowd. arXiv preprint arXiv:180500123. 2018.

19. Wang C-Y, Bochkovskiy A, Liao H-YM. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv preprint arXiv:220702696. 2022.
20. Wang C-Y, Bochkovskiy A, Liao H-YM, editors. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2023.
21. Gollapudi S, Gollapudi S. OpenCV with Python. Learn Computer Vision Using OpenCV: With Deep Learning CNNs and RNNs. 2019:31-50.
22. Chen P, Li Y, Zhou H, Liu B, Liu P. Detection of small ship objects using anchor boxes cluster and feature pyramid network model for SAR imagery. Journal of Marine Science and Engineering. 2020;8(2):112.
23. Wang X, Yang M, Zhu S, Lin Y, editors. Regionlets for generic object detection. Proceedings of the IEEE international conference on computer vision; 2013.
24. Ansari S. Industrial Application: Real-Time Defect Detection in Industrial Manufacturing. Building Computer Vision Applications Using Artificial Neural Networks: Springer; 2020. p. 361-87.
25. Ji S, Zhang C, Xu A, Shi Y, Duan Y. 3D convolutional neural networks for crop classification with multi-temporal remote sensing images. Remote Sensing. 2018;10(1):75.
26. Martin Z, Hendricks S, Patel A, editors. Automated tackle injury risk assessment in contact-based sports-a rugby union example. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2021.
27. Nonaka N, Fujihira R, Nishio M, Murakami H, Tajima T, Yamada M, et al., editors. End-to-End High-Risk Tackle Detection System for Rugby. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2022.
28. Carey L, Stanwell P, Terry DP, McIntosh AS, Caswell SV, Iverson GL, et al. Verifying head impacts recorded by a wearable sensor using video footage in rugby league: a preliminary study. Sports medicine-open. 2019;5(1):1-11.
29. Qian SS, Refsnider JM, Moore JA, Kramer GR, Streby HM. All tests are imperfect: Accounting for false positives and false negatives using Bayesian statistics. Heliyon. 2020;6(3).
30. Handa A, Newcombe RA, Angeli A, Davison AJ, editors. Real-time camera tracking: When is high frame-rate best? European Conference on Computer Vision; 2012: Springer.
31. Wang J, Song L, Li Z, Sun H, Sun J, Zheng N, editors. End-to-end object detection with fully convolutional network. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2021.
32. Najibi M, Samangouei P, Chellappa R, Davis LS, editors. Ssh: Single stage headless face detector. Proceedings of the IEEE international conference on computer vision; 2017.
33. Tang X, Du DK, He Z, Liu J, editors. Pyramidbox: A context-assisted single shot face detector. Proceedings of the European conference on computer vision (ECCV); 2018.
34. Sundararaman R, De Almeida Braga C, Marchand E, Pettre J, editors. Tracking pedestrian heads in dense crowd. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2021.

Competing Interests Statement:

Andrew Gardner serves as a scientific advisor for HitIQ, Ltd. He has a clinical practice in neuropsychology involving individuals who have sustained sport-related concussions (including current and former athletes). He has been a contracted concussion consultant to Rugby Australia. He has received travel funding or been reimbursed by professional sporting bodies, and commercial organisations for discussing or presenting sport-related concussion research at meetings, scientific conferences, workshops, and symposiums. Previous grant funding includes the NSW Sporting Injuries Committee, the Brain Foundation (Australia), an Australian-American Fulbright Commission Postdoctoral Award, a Hunter New England Local Health District, Research, Innovation and Partnerships Health Research & Translation Centre and Clinical Research Fellowship Scheme, and the Hunter Medical Research Institute (HMRI), supported by Jennie Thomas, and the HMRI, supported by Anne Greaves. He has current philanthropic support from the Nick Tooth Foundation. He acknowledges unrestricted philanthropic support from the National Rugby League (NRL). Keith Stokes is employed by the Rugby Football Union. Gemma Phillips is employed in a consultancy capacity by the Rugby Football League and Hull Kingston Rovers. Matt Cross is employed by the Premiership Rugby. Ben Jones is employed in a consultancy capacity by Premiership Rugby, and Rugby Football League. Cameron Owen's research fellowship is funded by the Rugby Football League.

Contributorship Statement:

Manish Mohan: Conceptualization, Methodology, Investigation, Data Curation, Writing Original Draft. Dan Weaving: Conceptualization, Data Curation, Writing – Review & Editing. Andrew Gardner: Writing – Review & Editing. Sharief Hendricks: Writing – Review & Editing. Keith Stokes: Writing – Review & Editing. Gemma Phillips: Writing – Review & Editing. Matt Cross: Writing – Review & Editing. Cameron Owen: Conceptualization, Data Curation, Writing – Review & Editing. Ben Jones: Conceptualization, Data Curation, Writing – Review & Editing, Project Administration. Manish Mohan is responsible for the overall content as guarantor.

Funding: No funding was received to undertake this study.

Ethics Approval Statement: Ethics approval was gained from the institutions' review board for the use of publicly available video data (REF: 100411).

Patient and public involvement: Patients and/or the public were not involved in the design, conduct, reporting, or dissemination plans of this research.

Acknowledgements: The authors would like to thank club medical staff and members of the Rugby Football League, specifically Matt Halkier, Laura Fairbank, and Julie Turton for their involvement in this study.

Table 1. Results of the Impact Detection Framework for 335 Standard-definition (SDTV) and High-definition (HDTV) test dataset videos.

	SDTV		HDTV	
	True Head-On-Head (positive)	True non-head-on-head (negative)	True Head-On-Head (positive)	True non-head-on-head (negative)
Test Head-on-head (positive)	63 True Positive	40 False Positive	60 True Positive	38 False Positive
Test non-head-on-head (negative)	29 False Negative	203 True Negative	32 False Negative	205 True Negative
Total	92	243	92	243