

OPEN

Convolutional neural network for efficient estimation of regional brain strains

Shaohu Wu^{1,3}, Wei Zhao^{1,3}, Kianoosh Ghazi¹ & Songbai Ji^{1,2*}

Head injury models are important tools to study concussion biomechanics but are impractical for real-world use because they are too slow. Here, we develop a convolutional neural network (CNN) to estimate regional brain strains instantly and accurately by conceptualizing head rotational velocity profiles as two-dimensional images for input. We use two impact datasets with augmentation to investigate the CNN prediction performances with a variety of training-testing configurations. Three strain measures are considered, including maximum principal strain (MPS) of the whole brain, MPS of the corpus callosum, and fiber strain of the corpus callosum. The CNN is further tested using an independent impact dataset (N = 314) measured in American football. Based on 2592 training samples, it achieves a testing R^2 of 0.916 and root mean squared error (RMSE) of 0.014 for MPS of the whole brain. Combining all impact-strain response data available (N = 3069), the CNN achieves an R^2 of 0.966 and RMSE of 0.013 in a 10-fold cross-validation. This technique may enable a clinical diagnostic capability to a sophisticated head injury model, such as facilitating head impact sensors in concussion detection *via* a mobile device. In addition, it may transform current acceleration-based injury studies into focusing on regional brain strains. The trained CNN is publicly available along with associated code and examples at <https://github.com/Jilab-biomechanics/CNN-brain-strains>. They will be updated as needed in the future.

Traumatic brain injury (TBI) remains a major public health problem in the world^{1,2}. According to the World Health Organization, more than 40 million people worldwide suffer from a mild TBI (mTBI) each year³. In the United States alone, the number of concussion incidents could reach 1.6–3.8 million annually, and is particularly common in athletes playing contact sports^{4,5}. Although mild in nature, about 300,000 of the incidents involve loss of consciousness, with the majority occurring in American football⁶.

To mitigate the risk of concussion, head impact sensors such as Head Impact Telemetry (HIT) System⁷ and mouthguards^{8,9} are deployed in many contact sports. They record impact kinematics upon head collision and have been extensively used to measure head impact exposure¹⁰. However, only peak linear and/or rotational accelerations such as “g-forces” are often used that do not inform impact-induced brain strains thought responsible for brain injury^{11,12}. Consequently, there is question about their effectiveness in concussion detection¹³.

Using measured kinematics as input, a sophisticated computational head injury model can estimate detailed brain strains. They are generally believed to be more effective than impact kinematics in detecting brain injury, including concussion¹⁴. However, a significant challenge preventing injury models from real-world use such as facilitating head impact sensors for concussion detection on the sports field is that they are too slow—typically requiring hours to simulate even a single head impact on a high-end workstation^{15–17}. As a result, the use of head injury models has been largely restricted to retrospective research efforts to date with no obvious clinical diagnostic value, and head impact sensors are also significantly underutilized.

There exist two competing strategies to mitigate the computational cost in model simulation. They share similarities in conceptualizing a head injury model as a nonlinear, high-dimensional, but smooth and continuous mapping function between impact kinematics and brain responses^{15,18}. One strategy is to simplify the model and response output. For example, several reduced-order models have been proposed, including a one-degree-of-freedom (DOF) dynamic model based on modal analysis¹⁹ or equation of motion²⁰, and a second-order model²¹. By fitting parameters of a reduced-order model against directly simulated responses obtained from a finite element (FE) model of the human head, hybrid brain injury metrics were developed to

¹Department of Biomedical Engineering, Worcester Polytechnic Institute, Worcester, MA, USA. ²Department of Mechanical Engineering, Worcester Polytechnic Institute, Worcester, MA, USA. ³These authors contributed equally: Shaohu Wu and Wei Zhao. *email: sji@wpi.edu

correlate with peak maximum principal strain (MPS) of the whole brain. These metrics have shown promise over other conventional injury metrics derived solely from rotational kinematics when correlating against MPS of the whole brain over a large spectrum of impact severities and in diverse injury scenarios on a group-wise basis^{21,22}.

However, MPS estimation accuracy from reduced models degrade for large strain impacts^{21,22}, presumably as a result of failing to capture the significant nonlinearities when the impact severity is high. Unfortunately, large strains from more severe impacts (vs. low severity blows) likely would be the most important to consider when assessing the risk of brain injury on an individual basis. Further, reduced models relying on MPS of the whole brain lose critical information on brain strain distribution, because the strain variable is overly simplified to a single scalar value for the entire brain and is not region-specific. It is also unclear how reduced models could estimate directionally informed “axonal”^{23–25} or “fiber”²⁶ strains that characterize stretches along white matter fiber tracts, or strain rate²⁷ thought important to brain injury as well. These observations indicate some inherent limitations with reduced-order models in practical applications, despite their potential for gaining physical insight into the induced brain strains.

To preserve the nonlinearity and spatial distribution of brain strains, a pre-computed brain response atlas was also introduced¹⁵. Instead of simplifying the model and output, this approach idealizes impact kinematic profiles serving as model input. Brain strains are pre-computed for a large library of impacts by discretizing peak rotational acceleration/velocity, and azimuth and elevation angles of head rotation²⁸. Element-wise MPS values (vs. peak magnitude of the whole brain^{21,22}) for an arbitrary impact are then interpolated/extrapolated instantly. The pre-computed response atlas was shown to be effective using dummy head impacts simulating American football²⁸. However, the idealized rotational kinematic profiles are limited to triangular shapes of acceleration impulses that do not include deceleration. For more complex kinematic profiles involving deceleration and velocity reversal, the atlas may need to be expanded to include additional characteristic impacts. Unfortunately, this requires an explicit analysis of rotational velocity profile shapes, which is not trivial^{28,29}.

Instead of simplifying the impact kinematic input^{15,28}, head injury model, or response output^{19–21}, here we develop a convolutional neural network (CNN) to learn the nonlinear impact-strain relationship without any simplification. The CNN is a typical data-driven approach where the network weights are determined by the given training data iteratively *via* backpropagation. By “implicitly” capturing important features of head rotational velocity profiles, regional brain strains can be estimated instantly with sufficient accuracy while maintaining the fidelity of impact kinematics, the sophistication of a head injury model, and the detailed response output. Such a capability is critical for effective real-world applications such as concussion detection on the sports field.

CNN is a class of deep learning neural network that has been extensively used in medical imaging and computer vision^{30,31}. Object detection and pattern recognition are achieved *via* local filter convolution to capture structural information among neighboring pixels/voxels^{32,33}. This is analogous to detecting local shape variations in head rotational kinematic temporal profiles that serve as model input to determine brain strains. This inspired us to conceptualize time-varying biomechanical signals of head rotation, specifically, rotational velocity profiles along the three anatomical directions, as two-dimensional (2D) images to apply CNN for response regression. The image representation preserves the temporal locality of head rotational velocity as the three components along the temporal dimension are given at the same time. In contrast, concatenating the three velocity profiles into a one-dimensional (1D) vector may not work well with existing CNN architectures because it destroys the temporal locality required for data convolution.

We organized the study as the following. We first used two real-world impact datasets to generate sufficient training samples through augmentation. Next, we empirically optimized a CNN architecture and probed its testing performance behaviors under a variety of training-testing configurations. All training samples were then combined to conduct a 10-fold cross-validation³⁴ and to further test on a third, independent impact dataset measured in American football³⁵. Finally, an additional 10-fold cross-validation was conducted using all response samples available. The CNN technique developed here allows instantly converting a head impact measured from impact sensors into regional brain strains. Therefore, it may facilitate head impact sensors in monitoring neural health and detecting concussion more effectively; thus, transforming kinematics-based TBI studies into brain strain-related investigations in the future.

Materials and Methods

Data augmentation. We used two impact datasets to generate CNN training samples: (1) video-confirmed impacts measured in American college football, boxing, and mixed martial arts from Stanford University (SF; $N = 110$)⁸; and (2) lab-reconstructed impacts from National Football League (NFL; $N = 53$)³⁶. The SF dataset was measured using instrumented mouthguards⁸. The latter were reconstructed and recently reanalyzed³⁶ in the laboratory using dummy heads by matching the location, direction, and speed of the impacts approximated from video analysis³⁷. As contribution of linear acceleration to brain strains was negligible³⁸, only isolated 3-DOF rotational velocity profiles in a ground-fixed coordinate²⁸ were used. The data sizes were small compared to typical CNN applications (e.g., thousands or even millions³³). Therefore, data augmentation was necessary to increase the variation in head rotational kinematics for training.

To do so, components along the x , y , and z directions were permuted to construct 6 ($N = 3!$) v_{rot} profiles (i.e., xyz , xzy , yxz , yzx , zxy , and zyx ; Step 1 in Fig. 1). Each profile was further rotated about a random axis passing through the head center of gravity with a random magnitude (within $0–90^\circ$; Step 2 in Fig. 1)²⁸. The azimuth and elevation angles (θ and α , respectively) of the rotational axis, $\Omega(\theta, \alpha)$, were then determined based on peak magnitude of rotational velocity²⁸. Due to head symmetry about the mid-sagittal plane, only half of the Ω sampling space was necessary (shaded area in Fig. 1)¹⁵. Therefore, for Ω with $\|\theta\| > 90^\circ$, its corresponding “conjugate rotational axis”, $\Omega''(180^\circ - \theta, -\alpha)$, was used to maximize the use of v_{rot} profiles for generating unique brain responses (optional Step 3 in Fig. 1).

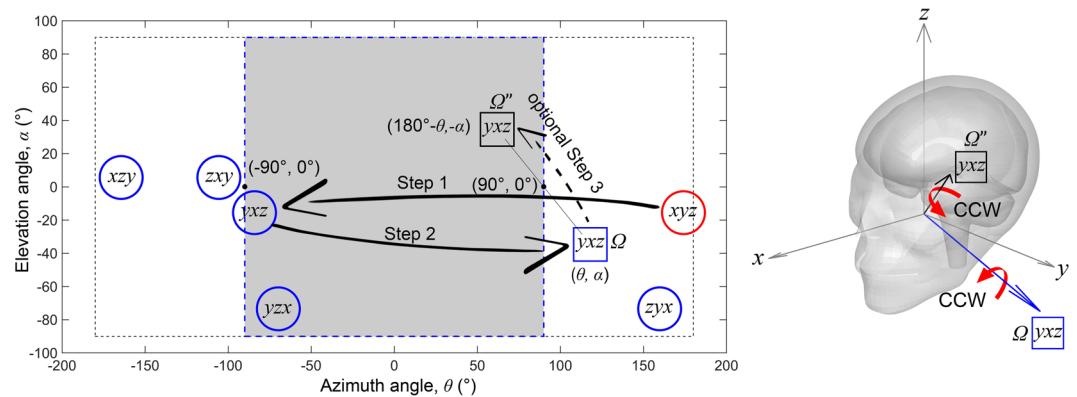


Figure 1. Illustration of data augmentation for a representative head impact rotational profile in terms of rotational axis, $\Omega(\theta, \alpha)$. Rotational velocity components along the three anatomical axes were permuted (Step 1), randomly rotated (Step 2), and then converted to its conjugate rotational axis, Ω'' (optional Step 3), if the azimuth angle, θ , is outside the sampling range (shaded area). Counter clockwise (CCW) rotation about Ω'' generates a mirroring response about the mid-sagittal plane relative to CCW rotation about Ω ¹⁵.

Finally, to focus on more “severe” head impacts relevant to potential “injury”, all profile magnitudes were randomly scaled so that the peak resultant velocity magnitude was above the median concussive value of 21.9 rad/s found in football while below 40 rad/s, sufficient to capture the 95th percentile of 34.1 rad/s³⁹. For a given impact dataset, each permutation and random perturbation/scaling constituted one batch of training data.

The augmented SF dataset was used to optimize a CNN architecture, including the number of layers and their associated parameters (details below). We found two batches of v_{rot} profiles ($N = 1320, 110 \times 6 \times 2$) were necessary to yield a coefficient of determination (R^2) above 0.90 (deemed successful) in a 10-fold cross-validation³⁴. Similarly, we generated two batches of v_{rot} profiles for the NFL impacts ($N = 636, 53 \times 6 \times 2$). Some impacts from the augmented datasets may not be physically admissible or rare to occur; still, they uniquely probed the impact-strain response hypersurface and were useful for CNN training.

Data preprocessing. The CNN requires a fixed input size. Therefore, all v_{rot} profiles were reformatted into a 3×201 matrix. The first dimension was fixed to represent time-varying v_{rot} components along the three anatomical directions. The second dimension corresponded to 200 ms in temporal length at a resolution of 1 ms (from 0 to 200 ms; sufficient for all impacts used). The second dimension could be adjusted based on the impact temporal resolution, which may require adjusting the CNN filter and stride sizes accordingly (details below). As peak rotational velocity was known to be important for brain strains, we synchronously shifted the three v_{rot} components so that the resultant peak occurred at a fixed time point of 100 ms (Fig. 2). At both ends of the rotational velocity profile, replicated padding was used to maintain a zero acceleration, where values at the two velocity profile borders were replicated along the temporal axis. Controlling the temporal location of the resultant velocity peak reduced kinematic data variation, which was expected to decrease the number of training samples required because the same shifting and padding were applied to all testing v_{rot} profiles. Repositioning v_{rot} profiles in time did not affect output, as the strain responses were accumulated maximum values, regardless of the time of occurrence.

Impact simulation. All impacts were simulated using the recent Worcester Head Injury Model (WHIM) that incorporates anisotropic material properties of the white matter based on whole-brain tractography⁴⁰. The anisotropic WHIM employs the same mesh and brain-skull boundary conditions as in the previous isotropic version²⁶. It was successfully validated against six cadaveric impacts and an *in vivo* head rotation⁴⁰.

Three strains from model simulations were separately used for training and testing: MPS of the whole brain, MPS of the corpus callosum (CC; a particularly vulnerable region^{41,42}), and fiber strain of the CC, all assessed at the 95th percentile level. Their increasing level of sophistication (whole brain vs. region-specific; direction invariant MPS vs. directionally informed fiber strain) was designed to stress test the CNN prediction capability.

CNN architecture. Application of CNN in TBI investigation is limited to image-based tasks so far, such as mimicking neuronal behaviors in cognitive deficits⁴³ and contusion image segmentation⁴⁴. Its application in biomechanics has also emerged, for example, to predict musculoskeletal forces⁴⁵ and to perform mobile gait analysis⁴⁶. However, application of CNN in TBI biomechanics does not yet appear to exist.

Therefore, we first empirically optimized a CNN architecture⁴⁵ using whole-brain MPS obtained from the augmented SF dataset. The number of CNN filters and their sizes and stride sizes were iteratively and empirically updated until a 10-fold cross-validation performance was maximized in terms of R^2 between the predicted and directly simulated responses. Figure 2 shows the optimized CNN, which led to a maximized R^2 of 0.937 with root mean squared error (RMSE) of 0.018. The 32 filters had sizes of 3×10 , 1×10 and 1×5 with stride sizes of 1×2 , 1×2 and 1×1 for the three convolutional layers, respectively. They were followed by a flattening layer (with a dropout rate of 0.2⁴⁷) and two fully connected layers. Rectified linear unit (ReLU) activation functions⁴⁸ were used.

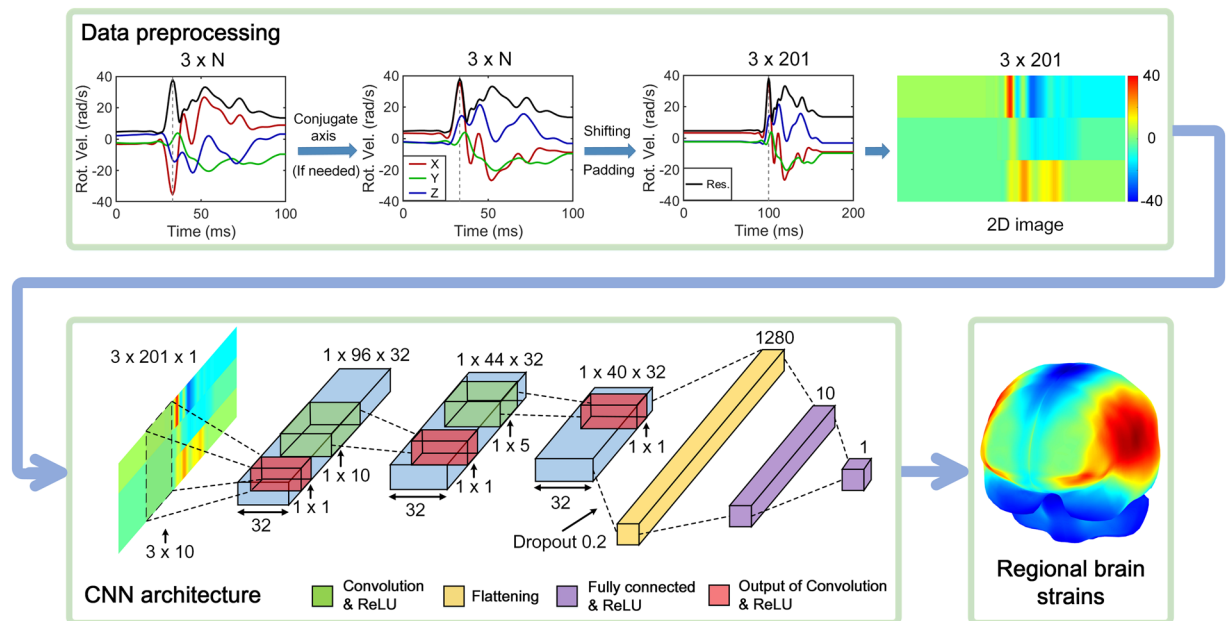


Figure 2. Overview of impact data preprocessing (top) and an empirically optimized CNN architecture (bottom) for training and testing. A typical rotational velocity temporal profile is conceptualized as a 2D image for CNN input. Note the difference in velocity profile temporal axes, as no curves are “squeezed”. The technique is illustrated for three strain variables in this study but can be easily extended to the whole-brain, element-wise responses. CNN: convolutional neural network.

The same optimized CNN architecture was employed for all subsequent trainings with 250 epochs, an initial learning rate of 10^{-6} , and a batch size of 64. Mean squared error (MSE) between the predicted and directly simulated strain measure of interest served as the loss function for minimization *via* an adaptive moment estimation (Adam) optimizer⁴⁹ implemented in Keras (Version 2.08)⁵⁰. Validation-based early stopping⁵¹ was used to avoid overfitting.

Performance evaluations. All testing v_{rot} profiles followed the same preprocessing steps (Fig. 2). We first used the augmented SF dataset to train and test on the reconstructed NFL dataset, and conversely, used the augmented NFL dataset to train and test on the measured SF dataset. To probe the importance of training sample size on testing performances, two additional batches were generated for the augmented NFL dataset, leading to a total of 1272 v_{rot} impacts ($53 \times 6 \times 4$; comparable to the augmented SF data size). We then repeated the same training and testing on the measured SF dataset. To further investigate the importance of v_{rot} profile shape characteristics in training (measured on-field vs. lab-reconstructed), we also used the augmented SF dataset to train and estimate strains in the measured SF dataset, as they were from the same data source and shared the same v_{rot} resultant profile shapes on a group-wise basis. For completeness, the augmented NFL dataset was also used to train and test on the reconstructed NFL impacts.

Next, we combined the augmented SF and NFL datasets ($N = 2592$) and conducted a 10-fold cross-validation. To demonstrate real-world use of our technique, we used the combined dataset to train and test on a third, independent impact dataset measured and video-confirmed using an impact monitoring mouthguard⁹ in American high school football (HF; $N = 314$ ³⁵). Finally, all impact-strain response samples available in this study were combined ($N = 3069$; $1320 + 1272 + 110 + 53 + 314$) to conduct a final 10-fold cross-validation.

For each strain measure, we evaluated testing accuracy using R^2 and RMSE. Because the augmented datasets intentionally focused on more severe head impacts most relevant to injury, we reported the testing performances for impacts within the focused rotational velocity peak range, in addition to the full dataset.

Statistical tests. For all tests, 30 trials with random CNN initialization seeds were conducted ($n = 30$). For 10-fold cross-validations, their performances were compared using a corrected one-tailed t -test to avoid high Type I error⁵². For other tests, a Welch’s one-tailed t -test was used to account for unequal variances⁵³. Significance level was set at 0.05.

Data analysis. Although all velocity profiles were reformatted to 200 ms as CNN input, only impact profiles prior to shifting and padding were necessary for simulation (Fig. 2), as the padded zero accelerations had no effect on peak brain strains. Each impact of 100 ms required ~ 30 min for simulation with Abaqus/Explicit (double precision with 15 CPUs and GPU acceleration; Intel Xeon E5-2698 with 256 GB memory, and 4 NVidia Tesla K80 GPUs with 12 GB memory) and another ~ 30 min for post-processing to calculate regional strains. In total, 3069 impacts were simulated, typically with 5–10 jobs running simultaneously. Training a CNN required ~ 3 min per fold on an NVIDIA Titan X Pascal GPU with 12 GB memory, while predicting on a testing profile was instant (< 0.1 sec) on a low-end laptop. All data analyses were conducted in MATLAB (R2018b; MathWorks, Natick, MA) and Python (Version 2.7.0).

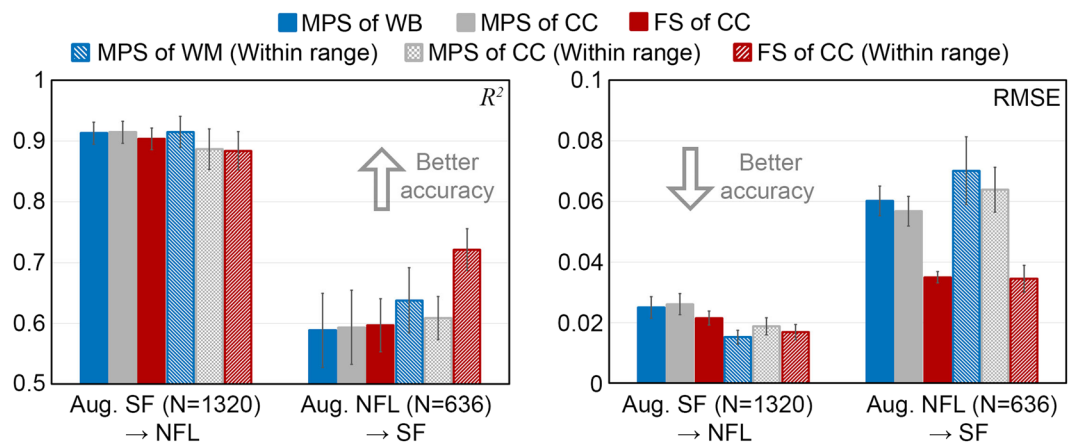


Figure 3. Average R^2 (left) and RMSE (right) from 30 random trials (with bars indicating standard deviation) using augmented SF dataset to predict responses in the reconstructed NFL impacts, and vice versa. For each strain measure, performances are reported using either impacts from the entire dataset or those within the targeted resultant rotational velocity peak range (within-range). MPS: maximum principal strain; FS: fiber strain; WB: whole brain; CC: corpus callosum.

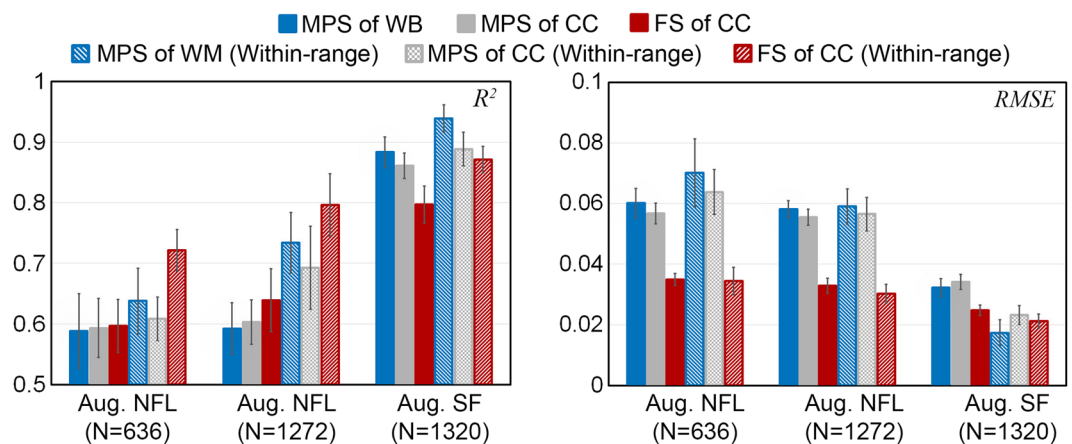


Figure 4. R^2 (left) and RMSE (right) when using three training datasets to test on the same measured SF impact dataset. Increasing the training samples in the augmented NFL dataset improved performances but were still outperformed by those using the augmented SF dataset for training, even when the training sample sizes were comparable. MPS: maximum principal strain; FS: fiber strain; WB: whole brain; CC: corpus callosum.

Results

Performance evaluation: between the two datasets. Using the augmented SF dataset for training and testing on the reconstructed NFL impacts consistently achieved significantly higher R^2 and lower RMSE than the other way around switching the two datasets for training/testing for the same strain measure ($p < 0.001$; Fig. 3; range of R^2 and RMSE: 0.884–0.915 and 0.015–0.026, respectively, vs. 0.588–0.721 and 0.035–0.07 for the latter). The variance across 30 trials was also consistently smaller in the former (e.g., standard deviation for R^2 range of 0.018–0.033 vs. 0.034–0.061).

Figure 4 reports R^2 and RMSE using three augmented datasets for training but the same measured SF dataset for testing. Increasing the augmented NFL training sample size consistently increased R^2 and lowered RMSE ($p < 0.001$), especially for impacts within the targeted rotational velocity peak range. Still, using the augmented SF dataset for training significantly outperformed those using the augmented NFL for training ($p < 0.001$), even when the sample sizes were comparable (e.g., R^2 of 0.871 vs. 0.796, and RMSE of 0.021 vs. 0.03, for within-range fiber strain in the CC).

Figure 5 compares the predictions with the directly simulated MPS of the whole brain for four training-testing configurations. When the training and testing datasets were from the same source, (Fig. 5c,d), the testing performance was always satisfactory (e.g., $R^2 > 0.937$ and RMSE < 0.018 for within-range impacts). Using the augmented SF dataset also successfully predicted strains for the reconstructed NFL dataset ($R^2 = 0.921$ and RMSE = 0.014 for within-range impacts; Fig. 5a). However, when using the augmented NFL dataset to predict

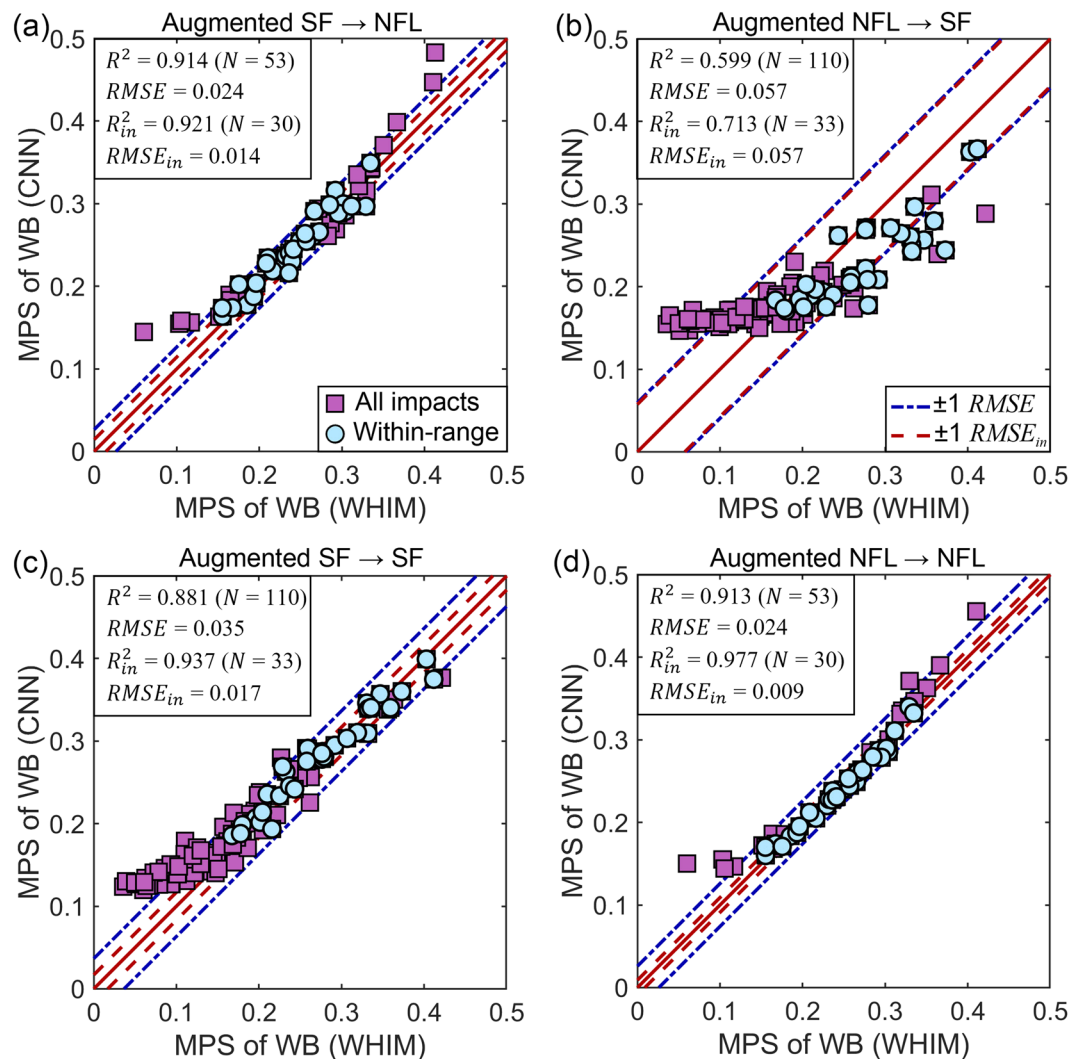


Figure 5. Typical prediction performances for four training-testing configurations (a–d; augmented NFL dataset used all four batches). Each selected plot reflects a typical trial, with its performance in terms of R^2 closest to the average value from the 30 random trials. Subscript, *in*, refers to impacts within the targeted resultant rotational velocity peak range. MPS: maximum principal strain; WB: whole brain.

measured SF dataset, the performance was notably poorer (e.g., $R^2 = 0.599$ and $RMSE = 0.057$ for all impacts; Fig. 5b). For the three “successful” predictions, CNN predictions for out-of-range impacts remained either within the ± 1 RMSE range for many (overall $R^2 > 0.880$) or somewhat overestimated.

Performance evaluation: combining the training datasets. Figure 6 reports the testing performances on the three measured impact datasets by combining the augmented SF and NFL datasets for training. Testing on the reconstructed NFL dataset achieved the highest R^2 of 0.978 with RMSE of 0.008 for within-range MPS of the whole brain. Compared to those when using the augmented SF dataset alone for training (Fig. 3), they were increased by 7–9% and decreased by 38–45%, respectively ($p < 0.01$). A smaller performance gain was also observed for the within-range impacts in the measured SF dataset (R^2 increased by 2–4% and RMSE decreased by 1–14% as compared to those when using the augmented SF dataset alone for training; Fig. 3).

When tested on the third HF dataset, the testing R^2 for MPS of the whole brain achieved the highest value of 0.916, with RMSE consistently < 0.02 for within-range impacts. Figure 7 compares the predicted three strain measures with the directly simulated counterparts for the three impact datasets in a typical trial using the augmented datasets combined for training.

Finally, Fig. 8 compares R^2 and RMSE for the three strain measures in 10-fold cross-validations using either the augmented SF and NFL datasets combined ($N = 2592$) or all impact response data available ($N = 3069$). Increasing the training samples significantly improved R^2 (to values of 0.966 ± 0.001 , 0.942 ± 0.002 and 0.930 ± 0.002 for MPS of the whole brain, MPS of the CC, and fiber strain of the CC, respectively; $p < 0.01$), but not statistically significant for decreasing RMSE (p of 0.051–0.110; RMSE values consistently < 0.018).

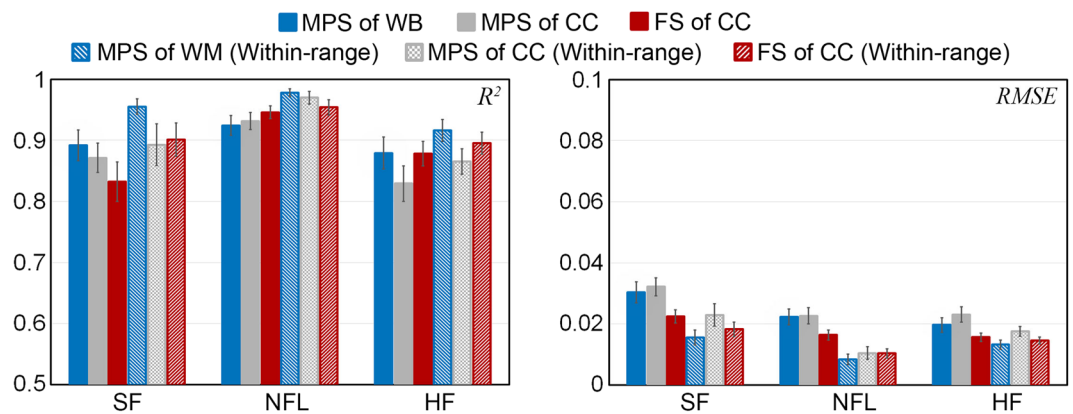


Figure 6. R^2 (left) and RMSE (right) when using the augmented training datasets combined ($N = 2592$ impacts) to predict brain strains for the measured SF, reconstructed NFL, and measured HF datasets based on 30 random trials, respectively. MPS: maximum principal strain; FS: fiber strain; WB: whole brain; CC: corpus callosum.

Discussion

Improving the computational efficiency while maintaining the sophistication of a head injury model and accuracy in response output is critical for prospective, real-world applications such as facilitating impact sensors in concussion detection on the sports field. This has long been desired^{16,17} but has remained infeasible to date. In this study, we introduced a data-driven approach using a convolutional neural network (CNN) to learn the non-linear impact-strain relationship without the need to simplify impact kinematic input^{15,28}, head injury model, or response output^{19,21,22}. The substantial increase in computational efficiency—from hours on a high-end workstation to under a second on a laptop—may enable a sophisticated head injury model for practical real-world use.

For all training-testing configurations regardless of the strain measure, the trained CNN instantly estimated regional brain strains with sufficient accuracy, especially for within-range impacts. The only exception might be when using the augmented NFL dataset to test on measured SF dataset (Fig. 5b; further discussed below). MPS of the whole brain generally achieved the best performance, with some degradation for more sophisticated strain measures characterizing region-specific MPS and directionally informed fiber strain of the corpus callosum (Figs 7 and 8). Although the regions of interest were limited to the whole brain and corpus callosum for illustration in this study, most likely the technique can be extended to other regions, including specific gray matter regions and their white matter interconnections. In contrast, reduced-order methods are limited to estimating peak MPS of the whole brain^{19–22} but not element- or region-wise MPS, or directionally informed fiber strain in specific regions. The pre-computation technique estimates element-wise MPS of the entire brain^{15,28}; however, it remains unclear how it performs when estimating fiber strains.

Nevertheless, directly benchmarking against these competing methods was not feasible in this study because they did not use the same impact data to report performance. Accuracy assessment in this study was limited to those occurring in contact sports; albeit, they were scaled to levels most relevant to injury. In contrast, earlier reduced-order models used impacts from a broad range of head impact conditions including sports, automotive crashes, sled tests and human volunteers^{21,22}. However, in addition to degraded accuracy for more severe, but perhaps the most important, impacts, inclusion of low-severity impacts could have also artificially skewed the correlation to a more favorable score. Performance with the pre-computation technique was limited to dummy head impacts, but not real-world recorded impacts yet²⁸.

Both the CNN technique and the previous pre-computed brain response atlas approach¹⁵ require a large training dataset. However, the CNN technique is significantly more advantageous because it is scalable—when model-simulated responses from new or “unseen” impacts become available (e.g., to challenge the CNN estimation accuracy), they can be assimilated into the existing training dataset to re-train. This iterative updating process is expected to continually improve the estimation accuracy, as illustrated (Fig. 8). In contrast, the pre-computation approach is unable to assimilate fresh impact-response samples¹⁵.

Even for impacts whose resultant peak velocity fell outside of the targeted peak resultant velocity range, many of their CNN predictions were still within ± 1 RMSE relative to their directly simulated counterparts (Figs 5 and 7). This indicated some impressive generalizability and robustness of the CNN technique. The ReLU activation functions were very effective in characterizing the inherently nonlinear brain-skull dynamic system, as it introduced sparsity effect⁵⁴ on the neural network to improve information representation and to avoid overfitting³³. ReLU was also desirable here since its output and strains of interest were all non-negative. Nevertheless, further fine-tuning the CNN architecture may be desirable in the future using all of the impact-strain response samples available ($N = 3069$).

The data augmentation scheme *via* permutation and random perturbation (Fig. 1) was important to generate sufficient data variations to the head rotational axis and velocity magnitude. However, the temporal “shapes” of the head rotational velocity profiles were limited to what the measured/reconstructed impact datasets offered. Unfortunately, a parameterized descriptor of head rotational velocity temporal shape does not exist to allow generating its variations for CNN training. However, potentially this could be somewhat augmented by linearly

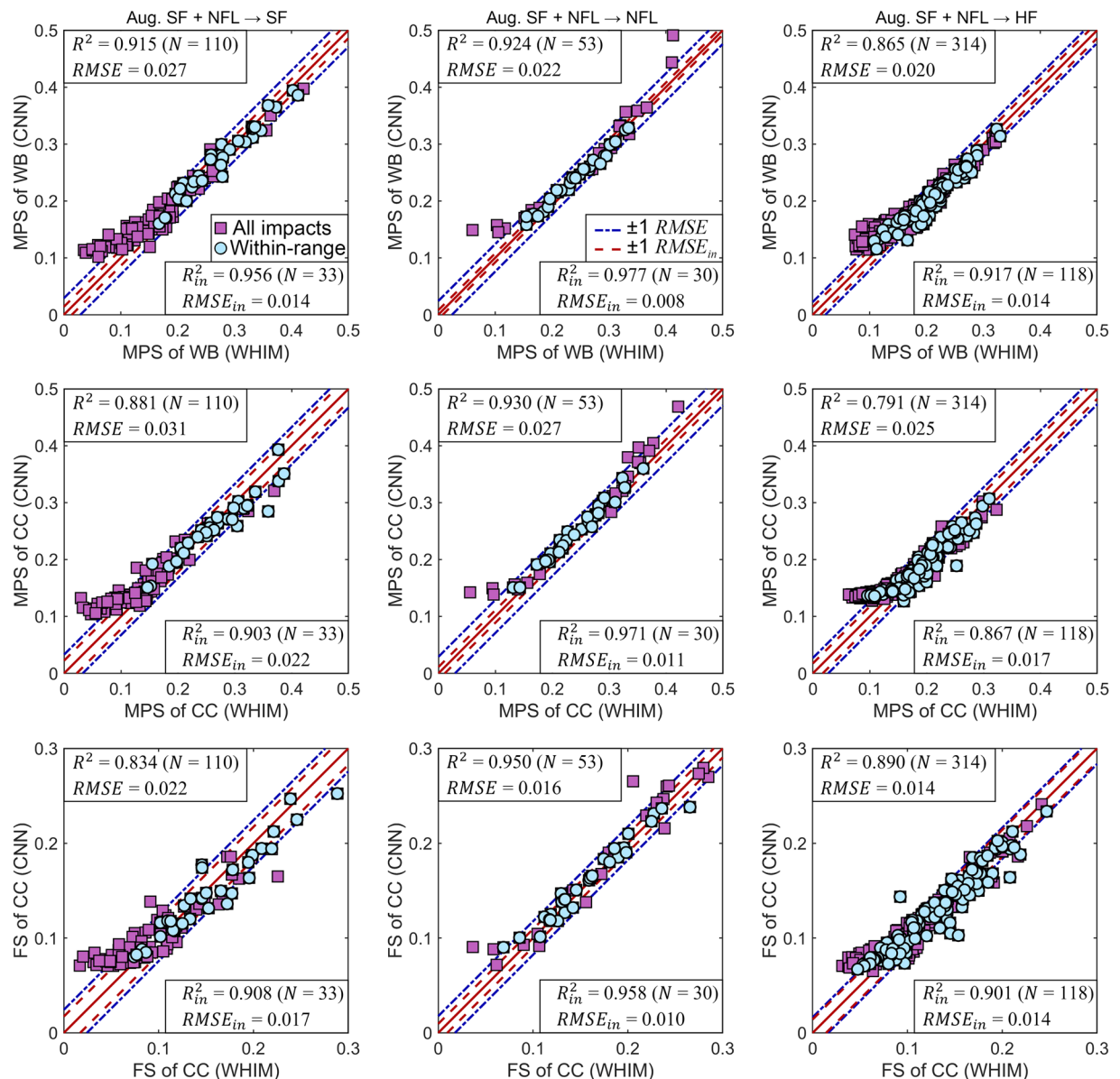


Figure 7. Typical CNN prediction performances for maximum principal strain (MPS) of the whole brain (WB; top), MPS of the corpus callosum (CC; middle row) and fiber strain (FS) of the CC (bottom) for three impact datasets: measured SF (left), reconstructed NFL (middle column), and measured HF (right).

scaling along the temporal direction, which can be explored in the future. Nevertheless, shifting the rotational velocity profiles so that the resultant peak velocity occurred at a fixed temporal location indeed slightly improved the testing performance; otherwise, R^2 typically dropped by 0.01 for within-range evaluations.

We chose to use v_{rot} profiles for training because v_{rot} is known to be important to brain strain¹⁵. The corresponding rotational acceleration profiles or the combination of rotational velocity and acceleration profiles could also be used for training. They are equivalent in prescribing head motion, with the caveat of a possible non-zero initial velocity in v_{rot} profiles, which was found to be negligible for brain strain. Therefore, they were found to lead to virtually identical CNN performances (confirmed but not shown). The CNN input matrix was fixed to 3 in the first dimension to correspond to the three anatomical directions, but the second temporal dimension could be adjusted. We found that increasing the temporal resolution did not improve CNN estimation accuracy, but decreasing the temporal resolution degraded the accuracy. This suggested that the latter led to some loss of information, as expected.

Comparing performances across different training-testing configurations, we found that training using the augmented SF dataset to test on the reconstructed NFL dataset outperformed that when instead, using the augmented NFL dataset for training to test on the measured SF dataset, even when the sample sizes were comparable (Fig. 5). We suspected that this was because the measured impacts on the field contained more “feature” variations in the rotational kinematics profiles that included acceleration, deceleration, as well as reversal in rotational velocity^{8,28}.

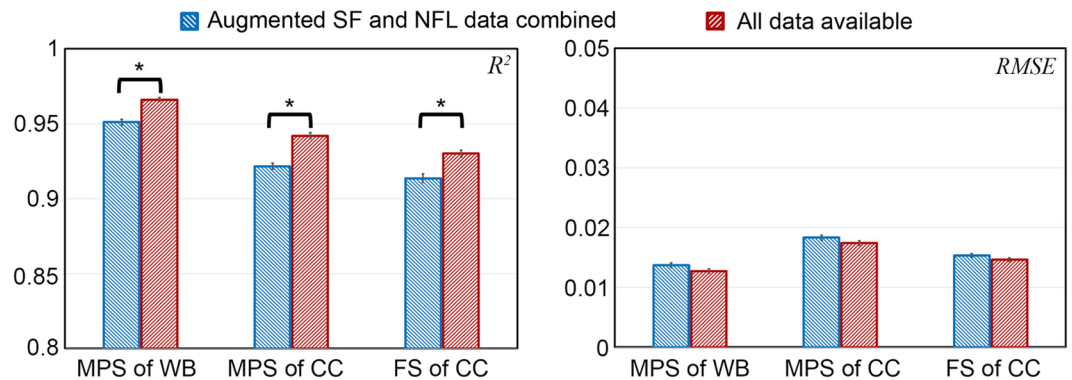


Figure 8. R^2 (left) and RMSE (right) in 10-fold cross-validations using either the augmented training datasets combined ($N = 2592$) or all impact response data available ($N = 3069$) based on 30 random trials. MPS: maximum principal strain; FS: fiber strain; WB: whole brain; CC: corpus callosum; “*” indicates $P < 0.01$.

In contrast, lab-reconstructed head impacts mainly focused on matching impact location, direction, and speed³⁶ but may not include more complicated shape variations. These findings suggested the importance of using “feature-rich” impact kinematics data to maximize variation in the training samples. Nevertheless, the addition of the augmented NFL dataset did improve testing performance (Figs 3 vs. 6), suggesting that they also contained information useful for CNN training. On the other hand, prediction on relatively “simple” impacts such as those in the reconstructed NFL dataset always yielded the best performance (Figs 5–7).

Finally, although sophisticated computational hardware is desirable/necessary for model simulations⁴⁰ to generate training samples and to train the CNN, this is not necessary for using an already trained CNN for prediction. In fact, even a portable mobile device may be sufficient for applying a trained CNN for prediction, which is desirable for potential real-world deployment in the future. To potentially better serve the research community and other interested parties, we have made the trained CNN publicly available, along with code and examples at <https://github.com/Jilab-biomechanics/CNN-brain-strains>. It is anticipated that the link will be updated as needed in the future.

Limitations. For brevity, we only reported accuracies of strains measured at the 95th percentile for illustration. The same approach was also tested at the 100th and 50th percentile levels (both used in injury prediction). A lower percentile would effectively serve as a smoothing filter to the impact-response hypersurface, which would slightly improve the prediction accuracy compared to the 100th percentile responses (confirmed but not shown).

We focused on a relatively narrow range of peak resultant velocity magnitude because they were most relevant to injury. However, expanding the coverage range to lower impact severities to allow considering the cumulative effects of repeated sub-concussive impacts is straightforward. For out-of-range impacts, the CNN generally over-estimated strains to some degree, but with many still within ± 1 RMSE relative to the directly simulated counterparts. This highlighted the generalizability and robustness of the technique, which was important for real-world applications where unexpected, “out-of-range” impacts may occur.

Further, the data-driven CNN technique does not address any physics behind brain biomechanical responses. However, as a fast and accurate brain strain response generator, the trained CNN may allow other researchers to efficiently produce impact-response data to explore the physics behind brain strains and concussion in reduced-order models⁵⁵.

Nevertheless, the CNN was only tested using impacts in contact sports in this study because our focus was to enable head injury models for facilitating concussion detection on the sports field. It merits further investigation into whether the application can be expanded more broadly to other impact scenarios such as automotive crashes^{21,22}. If found to be similarly effective, the technique may allow transforming state-of-the-art impact kinematics-based studies of brain injury into focusing more on brain strains in the future.

Finally, the trained CNN is model-dependent, and regional strain estimates are subject to all limitations related to the WHIM used for generating training samples. Nonetheless, the CNN can be easily re-trained to accommodate another model or a future, upgraded WHIM. In this case, a large amount of impacts may need to be simulated again. However, existing training data already simulated can still be reused to set an appropriate initial starting point for CNN training, which would reduce the number of impacts required to re-simulate.

Conclusion

In this study, we developed a deep convolutional neural network (CNN) to train and instantly estimate impact-induced regional brain strains with sufficient accuracy. The technique is significantly more advantageous than other alternative methods, because it does not need to simplify impact kinematic input, head injury model, or response output, and is effective for estimating more sophisticated, region-specific and directionally informed strains. The trained neural network is uniquely capable of assimilating fresh impact-response samples to iteratively improve accuracy. Together with sensors that measure impact kinematics upon head collision, this technique may enable a sophisticated head injury model to produce region-specific brain responses, instantly, potentially even on a portable mobile device. Therefore, this technique may offer clinical diagnostic values to

a sophisticated head injury model, e.g., to facilitate head impact sensors in concussion detection *via* a mobile device. This is important to mitigate the millions of concussion incidents worldwide every year. In addition, the technique may transform current acceleration-based injury studies into focusing on regional brain strains.

Received: 31 July 2019; Accepted: 24 October 2019;

Published online: 22 November 2019

References

1. Peden, M. *et al.* *World report on road traffic injury prevention* (2004).
2. CDC. *Report to Congress on Traumatic Brain Injury in the United States: Epidemiology and Rehabilitation*, <https://doi.org/10.1161/HYPERTENSIONAHA.111.186106> (2015).
3. Cassidy, J. D. *et al.* Incidence, risk factors and prevention of mild traumatic brain injury: results of the WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury. *J Rehabil Med* **43**, 28–60 (2004).
4. Dompier, T. P. *et al.* Incidence of concussion during practice and games in youth, high school, and collegiate American football players. *JAMA Pediatr.* **169**, 659–665 (2015).
5. Graham, R., Rivara, F. P., Ford, M. A., Spicer, C. M. & Graham, R. *Sports-Related Concussions in Youth*. *Jama* **311** (2014).
6. Thurman, D., Branche, C. & Sniezek, J. The Epidemiology of Sports-Related Traumatic Brain Injuries in the United States: Recent Developments. *J. Head Trauma Rehabil.* **13**, 1–8 (1998).
7. Greenwald, R. M., Gwin, J. T., Chu, J. J. & Crisco, J. J. Head Impact Severity Measures for Evaluating Mild Traumatic Brain Injury Risk Exposure. *Neurosurgery* **62**, 789–798 (2008).
8. Hernandez, F. *et al.* Six Degree-of-Freedom Measurements of Human Mild Traumatic Brain Injury. *Ann. Biomed. Eng.* **43**, 1918–1934 (2015).
9. Bartsch, A., Samorezov, S., Benzel, E., Miele, V. & Brett, D. Validation of an ‘Intelligent Mouthguard’ Single Event Head Impact Dosimeter. *Stapp Car Crash J.* **58**, 1–27 (2014).
10. Beckwith, J. G. *et al.* Head Impact Exposure Sustained by Football Players on Days of Diagnosed Concussion. *Med. Sci. Sports Exerc.* **45**, 737–746 (2013).
11. King, A. I. A. I., Yang, K. H. K. H., Zhang, L., Hardy, W. N. W. N. & Viano, D. C. D. C. Is head injury caused by linear or angular acceleration? in *IRCOBI Conference* 1–12 (2003).
12. Kleiven, S. Predictors for Traumatic Brain Injuries Evaluated through Accident Reconstructions. *Stapp Car Crash J.* **51**, 81–114 (2007).
13. Mihalik, J. J. P., Lynall, R. R. C., Wasserman, E. E. B., Guskiewicz, K. M. K. & Marshall, S. W. S. Evaluating the ‘Threshold Theory’: Can Head Impact Indicators Help? *Med. Sci. Sports Exerc.* **49**, 247–253 (2017).
14. Yang, K. H. *et al.* Development of numerical models for injury biomechanics research: a review of 50 years of publications in the Stapp Car Crash Conference. *Stapp Car Crash J.* **50**, 429–490 (2006).
15. Ji, S. & Zhao, W. A Pre-computed Brain Response Atlas for Instantaneous Strain Estimation in Contact Sports. *Ann. Biomed. Eng.* **43**, 1877–1895 (2015).
16. Franklyn, M., Fildes, B., Zhang, L., King, Y. & Sparke, L. Analysis of finite element models for head injury investigation: reconstruction of four real-world impacts. *Stapp Car Crash J.* **49**, 1–32 (2005).
17. Takhousts, E. G. *et al.* Investigation of traumatic brain injuries using the next generation of simulated injury monitor (SIMon) finite element head model. *Stapp Car Crash J.* **52**, 1–31 (2008).
18. Zhao, W., Choate, B. & Ji, S. Material properties of the brain in injury-relevant conditions – Experiments and computational modeling. *J. Mech. Behav. Biomed. Mater.* **80**, 222–234 (2018).
19. Laksari, K. *et al.* Resonance of human brain under head acceleration. *J. R. Soc. Interface* **12**, 20150331 (2015).
20. Gabler, L. F., Joodaki, H., Crandall, J. R. & Panzer, M. B. Development of a Single-Degree-of-Freedom Mechanical Model for Predicting Strain-Based Brain Injury Responses. *J. Biomech. Eng.* **140**, 031002 (2018).
21. Gabler, L. F., Crandall, J. R. & Panzer, M. B. Development of a Second-Order System for Rapid Estimation of Maximum Brain Strain. *Ann. Biomed. Eng.* 1–11, <https://doi.org/10.1007/s10439-018-02179-9> (2018).
22. Gabler, L. F., Crandall, J. R. & Panzer, M. B. Development of a Metric for Predicting Brain Strain Responses Using Head Kinematics. *Ann. Biomed. Eng.* <https://doi.org/10.1007/s10439-018-2015-9> (2018).
23. Giordano, C. & Kleiven, S. Evaluation of Axonal Strain as a Predictor for Mild Traumatic Brain Injuries Using Finite Element Modeling. *Stapp Car Crash J.* **November**, 29–61 (2014).
24. Wright, R. M. & Ramesh, K. T. An axonal strain injury criterion for traumatic brain injury. *Biomech. Model. Mechanobiol.* **11**, 245–60 (2012).
25. Cloots, R. J. H. H., van Dommelen, J. A. W. W., Nyberg, T., Kleiven, S. & Geers, M. G. D. D. Micromechanics of diffuse axonal injury: influence of axonal orientation and anisotropy. *Biomech. Model. Mechanobiol.* **10**, 413–22 (2011).
26. Ji, S. *et al.* Group-wise evaluation and comparison of white matter fiber strain and maximum principal strain in sports-related concussion. *J. Neurotrauma* **32**, 441–454 (2015).
27. King, A. I., Yang, K. H., Zhang, L., Hardy, W. N. & Viano, D. C. Is head injury caused by linear or angular acceleration? in *Proc. IRCOBI Conf* (2003).
28. Zhao, W., Kuo, C., Wu, L., Camarillo, D. B. & Ji, S. Performance evaluation of a pre-computed brain response atlas in dummy head impacts. *Ann. Biomed. Eng.* **45**, 2437–2450 (2017).
29. Zhao, W. & Ji, S. Brain strain uncertainty due to shape variation in and simplification of head angular velocity profiles. *Biomech. Model. Mechanobiol.* **16**, 449–461 (2017).
30. Yamashita, R., Nishio, M., Do, R. K. G. & Togashi, K. Convolutional neural networks: an overview and application in radiology. *Insights Imaging* **9**, 611–629 (2018).
31. Vouloimos, A., Doulamis, N., Doulamis, A. & Protopapadakis, E. Deep Learning for Computer Vision: A Brief Review. *Comput. Intell. Neurosci.* **2018**, 1–13 (2018).
32. Ren, S., He, K., Girshick, R. & Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal. *Networks. IEEE Trans. Pattern Anal. Mach. Intell.* **39**, 1137–1149 (2017).
33. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.* 1–9, <https://doi.org/10.1016/j.protcy.2014.09.007> (2012).
34. Hastie, T., Tibshirani, R. & Friedman, J. *The Elements of Statistical Learning, Data Mining, Inference, and Prediction*. (Springer, 2008).
35. Zhao, W. *et al.* Regional Brain Injury Vulnerability in Football from Two Finite Element Models of the Human Head. In *IRCOBI* **1**, 619–621 (2019).
36. Sanchez, E. J. *et al.* A reanalysis of football impact reconstructions for head kinematics and finite element modeling. *Clin. Biomech.* <https://doi.org/10.1016/j.clinbiomech.2018.02.019> (2018).
37. Pellman, E. J. *et al.* Concussion in professional football: reconstruction of game impacts and injuries. *Neurosurgery* **53**, 799–814 (2003).

38. Ji, S., Zhao, W., Li, Z. & McAllister, T. W. Head impact accelerations for brain strain-related responses in contact sports: a model-based investigation. *Biomech. Model. Mechanobiol.* **13**, 1121–36 (2014).
39. Rowson, S. *et al.* Rotational head kinematics in football impacts: an injury risk function for concussion. *Ann. Biomed. Eng.* **40**, 1–13 (2012).
40. Zhao, W. & Ji, S. White matter anisotropy for impact simulation and response sampling in traumatic brain injury. *J. Neurotrauma* **36**, 250–263 (2019).
41. Zhao, W., Cai, Y., Li, Z. & Ji, S. Injury prediction and vulnerability assessment using strain and susceptibility measures of the deep white matter. *Biomech. Model. Mechanobiol.* **16**, 1709–1727 (2017).
42. Hernandez, F. *et al.* Lateral impacts correlate with falx cerebri displacement and corpus callosum trauma in sports-related concussions. *Biomech. Model. Mechanobiol.* 1–19, <https://doi.org/10.1007/s10237-018-01106-0> (2019).
43. Lusch, B., Weholt, J., Maia, P. D. & Kutz, J. N. Modeling cognitive deficits following neurodegenerative diseases and traumatic brain injuries with deep convolutional neural networks. *Brain Cogn.* **123**, 154–164 (2018).
44. Roy, S., Butman, J. A., Chan, L. & Pham, D. L. TBI contusion segmentation from MRI using convolutional neural networks. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* 158–162 (IEEE, 2018), <https://doi.org/10.1109/ISBI.2018.8363545>
45. Rane, L., Ding, Z., McGregor, A. H. & Bull, A. M. J. Deep Learning for Musculoskeletal Force Prediction. *Ann. Biomed. Eng.* **47**, 778–789 (2019).
46. Hannink, J. *et al.* Sensor-Based Gait Parameter Extraction With Deep Convolutional. *Neural Networks. IEEE J. Biomed. Heal. Informatics* **21**, 85–93 (2017).
47. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.* **15**, 1929–1958 (2014).
48. Nair, V. & Hinton, G. E. Rectified Linear Units Improve Restricted Boltzmann Machines. *Proc. 27th Int. Conf. Mach. Learn.* 807–814, 10.1.1.165.6419 (2010).
49. Kingma, D. P. & Ba, J. L. Adam: a Method for Stochastic Optimization. *Int. Conf. Learn. Represent. 2015* 1–15, <https://doi.org/10.1145/1830483.1830503> (2015).
50. Chollet, F. & others. Keras (2015).
51. Prechelt, L. In *Neural Networks: Tricks of the Trade - Second Edition (2012)* 53–67, https://doi.org/10.1007/3-540-49430-8_3 (Springer, Berlin, Heidelberg, 1998).
52. Bouckaert, R. R. & Frank, E. In 3–12, https://doi.org/10.1007/978-3-540-24775-3_3 (2010).
53. Welch, B. L. The Generalization of ‘Student’s’ Problem when Several Different Population Variances are Involved. *Biometrika* **34**, 28 (1947).
54. Glorot, X., Bordes, A. & Bengio, Y. Deep Sparse Rectifier Neural Networks. in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics* 315–323 (2011).
55. Abderezaei, J. *et al.* Nonlinear Dynamical Behavior of the Deep White Matter During Head Impact. *Phys. Rev. Appl.* **12**, 014058 (2019).

Acknowledgements

Funding is provided by the NIH Grant R01 NS092853 and the Ford University Research Program. The authors are grateful to the National Football League (NFL) Committee on Mild Traumatic Brain Injury (MTBI) and Biokinetics and Associates Ltd. for providing the reconstructed head impact data. They also thank Dr. David Camarillo at Stanford University and Dr. Adam Bartsch at Prevent Biometrics for sharing head impact data. The Titan X Pascal GPU used in this work was donated by the NVIDIA Corporation.

Author contributions

S.W. conception, data analysis, writing, editing; W.Z. conception, data preparation, writing, editing; K.G. data preparation; S.J. conception, funding, writing, editing, final approval.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.J.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019