Journal of Neurotrauma

DOI: 10.1089/neu.2022.0413

A morphologically individualized deep learning brain injury model

Running title: Individualized deep learning brain injury model

Table of contents title: Individualized deep learning brain injury model

Nan Lin [1], Shaoju Wu [1], Songbai Ji [1,2]

Nan Lin (nlin@wpi.edu; Phone: 508-831-5447; Fax: 508-831-6362):

[1] Department of Biomedical Engineering, Worcester Polytechnic Institute, 60 Prescott Street, Worcester, MA 01605, USA

Shaoju Wu (swu3@wpi.edu; Phone: 774-701-0537; Fax: 508-831-6362):

[1] Department of Biomedical Engineering, Worcester Polytechnic Institute, 60 Prescott Street, Worcester, MA 01605, USA

Corresponding author: Songbai Ji (sji@wpi.edu; Phone: 508-831-4956; Fax: 508-831-6362):

[1] Department of Biomedical Engineering, Worcester Polytechnic Institute, 60 Prescott Street, Worcester, MA 01605, USA

[2] Department of Mechanical Engineering, Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA 01609, USA

Resubmitted to the *Journal of Neurotrauma*, March 24, 2023

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

Abstract

The brain injury modeling community has recommended improving model subject specificity and simulation efficiency. Here, we extend an instantaneous (<1 sec) convolutional neural network (CNN) brain model based on the anisotropic Worcester Head Injury Model (WHIM) V1.0 to account for strain differences due to individual morphological variations. Linear scaling factors relative to the generic WHIM along the three anatomical axes are used as additional CNN inputs. To generate training samples, the WHIM is randomly scaled to pair with augmented head impacts randomly generated from real-world data for simulation. An estimation of voxelized peak maximum principal strain of the whole brain is said to be successful when the linear regression slope and Pearson's correlation coefficient relative to directly simulated do not deviate from 1.0 (when identical) by more than 0.1. Despite a modest training dataset (N=1363 vs. ~5.7 k previously), the individualized CNN achieves a success rate of 86.2% in cross-validation for scaled model responses, and 92.1% for independent generic model testing for impacts considered as complete capture of kinematic events. Using 11 scaled subject-specific models (with scaling factors determined from pre-established regression models based on head dimensions and sex and age information, and notably, without neuroimages), the morphologically individualized CNN remains accurate for impacts that also yield successful estimations for the generic WHIM. The individualized CNN instantly estimates subject-specific and spatially detailed peak strains of the entire brain and thus, supersedes others that report a scalar peak strain value incapable of informing the location of occurrence. This tool could be especially useful for youths and females due to their anticipated greater morphological differences relative to the generic model, even without the need for individual neuroimages. It has potential for a wide range of applications for injury mitigation purposes and the design of head protective gears. The voxelized strains also allow for convenient data sharing and promote collaboration among research groups. The upgraded CNN is freely available at: https://github.com/Jilab-biomechanics/CNN-brain-strains.

Keywords: traumatic brain injury; brain model; convolutional neural network; deep learning; subject-specific model; Worcester Head Injury Model

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

Introduction

Traumatic brain injury (TBI) remains a leading contributor to mortality and morbidity[1] with an estimated 64 to 74 million incidents worldwide.[2] They often lead to debilitating negative neurological sequelae.[3] About three quarters of TBI incidents are mild TBI (mTBI),[4] which is particularly common in contact sports.[5,6] The incidents of sports-related mTBI are increasing in recent years,[7] especially for youth and female athletes due to their increased participation in sports.[8,9] Studies have suggest that females are generally at a greater risk of mTBI, often referred to as concussion in symptom, than males.[10] The number of concussion incidents is also likely underestimated because of the significant under-reporting issue.[11]

Indisputably, TBI, regardless of the severity, is initiated by brain deformation large or rapid enough from external head impact to cause tissue damage. Therefore, understanding the biomechanical mechanism of TBI is critical for designing improved head protective countermeasures to better protect the brain. Given that it is infeasible to directly measure live human brain deformation in an accident scenario, brain biomechanical models have been widely used to estimate tissue deformation.[12,13] There is general consensus that a physics-based and validated brain model has strong potential to improve the detection of injury and interpretation of subconcussive impact exposure over impact kinematics, alone.[14]

Nevertheless, several challenges remain for effective and practical use of a brain model. First, most brain models developed to date represent a generic, adult male head/brain.[12,14] They may suffer in accuracy when investigating brain injury mechanisms on a subject-specific basis,[15–17] especially for youth and female athletes because of the anticipated larger differences in brain morphology (shape and size) relative to the generic, adult male head/brain.[18] It is known that a larger (adult male) brain would lead to larger stress,[19] pressure,[20] and strain[16,17,21,22] than a smaller (youth and female) brain when applying the same head impact kinematics as model input. The second challenge is that brain injury models are notoriously poor in impact simulation efficiency, typically requiring hours[23–25] or even days[21,26] to simulate a single head impact and on a high-performance

computing platform. To address these challenges, the brain modeling community has recently recommended to enhance brain model subject specificity as well as to employ modern data science techniques to dramatically improve simulation efficiency.[14]

An early effort to improve model subject specificity was to scale a generic model to represent a 5th or a 95th percentile individual for a much smaller or larger brain, respectively.[27,28] More recently, subject-specific models are developed *via* mesh warping, by morphing a template brain model based on the deformation field obtained from co-registration between individual neuroimages and those from the template brain.[15,29,30] Alternatively, brain anatomy can also be meshed directly from neuroimages.[21,24,31] Most recently, another technique was developed to approximate subject-specific brain models by linearly scaling the brain along the three anatomical axes using pre-established statistical regression models between brain outer surface and head dimensions, along with subject's sex and age information.[17] Once the regression models are developed, this technique does not require individual neuroimages for scaling. It retains high accuracy compared to the presumably more accurate, "morphed model" developed from image warping. For example, for the smallest brain from a dataset of N=191 that is also anticipated to differ the greatest relative to the generic model, the scaled model had a strain difference of less than 1–3% for group-wise white and gray matter regions in terms of linear regression slope relative to those from the morphed subject-specific models.

On the other hand, deep learning has demonstrated remarkable success in dramatically improving model simulation efficiency without significant degradation in accuracy. It reduces hours of impact simulation on a high-end computing platform to under a second on an ordinary laptop. The first deep learning model in TBI biomechanics was a convolutional neural network (CNN) developed to instantly estimate regional peak strains.[32] It was then enhanced to estimate element-wise, spatially detailed peak strains of the entire brain[33] and those of automotive head impacts.[34] The CNN was further enhanced along with a separate transformer neural network (TNN) to rapidly estimate the complete spatiotemporal details of dynamic brain strain and strain rate, and with high accuracy (e.g., R^2>0.99 with normalized root mean squared error (NRMSE) of 2–3% for peak strains).[35] The earlier CNN studies have motivated others to use a U-Net[36] or explicit kinematic

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

features[37,38] to efficiently estimate stress and injury risk, whole-brain peak strains, or several peak strain/strain rate quantities, respectively. These deep learning brain models could allow for large-scale impact simulations, which are especially relevant to sports-related concussion and subconcussive exposure due to the large number of impacts typically sustained by each athlete.[14] A limitation, nonetheless, is that all existing deep learning brain models to-date are based on a generic model, representing a typical adult male brain.

Therefore, the purpose of this study is to enhance an existing CNN brain model to further account for strain differences due to individual variations in brain morphology (shape and size). Developing a separate CNN for each subject's brain is not optimal, as each CNN requires thousands of impact-response samples for training, and that each training sample requires a substantial simulation runtime.[33] Instead, here we take the advantage of approximating subject-specific brain models *via* simple linear scaling.[17] The three linear scaling factors are used as additional inputs to the existing, generic CNN to inform how strain values should be adjusted relative to those from the generic brain model.

The accuracy of the resulting individualized CNN is first evaluated for whole-brain strain prediction based on a generic model to ensure that it maintains the same level of accuracy relative to the generic CNN. Its accuracy for subject-specific brain models is further tested by comparing predicted strains with those directly simulated using scaled brain models from 11 subjects (with brain volume ranging from the smallest to largest in a dataset) and representative real-world impacts. The individualized CNN would lead to rapid whole-brain strain estimation on a subject-specific basis, and notably, without the need for individual neuroimages. Therefore, this tool could facilitate large-scale and individualized head impact simulations in the future. This may have important implications in designing improved head protective countermeasures and rapid estimation of concussion risks on a subject-specific basis. Finally, the individualized CNN purposefully outputs voxelized whole-brain peak strains[39] to allow for convenient data sharing and to promote collaboration among research groups.[14]

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

Methods

To train a neural network for predicting spatially detailed whole-brain peak strains for an arbitrary brain model subjecting to an arbitrary head impact, it is necessary to first generate appropriate training samples. This requires creating arbitrary brain models as well as impact kinematics as model simulation inputs. A recent study has demonstrated the feasibility of approximating subject-specific brain models by linearly scaling a generic counterpart along the three anatomical axes. The scaling factors are based on statistical linear regression models established between brain outer surface morphology (length, width, volume) and head dimensions (length, breadth, circumference, and tragion-to-top), along with the subject's sex and age information (see Eqns. A1–A3 in Appendix).[17] They are ideal for serving as additional inputs to the previously developed CNN model to inform how brain strains should be adjusted, without altering the neural network output data size. In the following sections, we describe the details of generating random scaled models and impact kinematics to prepare for training samples.

Randomly scaled models to generate training samples

The anisotropic Worcester Head Injury Model (WHIM) version 1.0 (V1.0)[40] served as the generic model (**Fig. 1a**). The WHIM was developed based on high-resolution T1-weighted MRI of an 18-year-old male athlete.[29] The head coordinate system was chosen such that the posterior-to-anterior, right-to-left, and inferior-to-superior directions corresponded to the *x*, *y*, and *z* directions, respectively. Compared with the previous isotropic version, the anisotropic V1.0 implements anisotropic material properties of the white matter based on whole brain tractography. Both isotropic and anisotropic WHIM V1.0 models have been successfully validated against high- and mid-rate cadaveric impact data in terms of relative brain-skull displacement and marker-based strain, as well as strain from low-rate *in vivo* volunteer rotations.[41]

Ideally, a statistical distribution model of the three scaling factors corresponding to brain length, width and height (in the *x*, *y*, and *z* directions, respectively) among real-world subjects could be used to generate random brain models for training samples. The three linear scaling factors are likely dependent, which would ensure the resulting training

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

samples to be "life-like". However, such a distribution model does not yet exist, and its establishment is outside the scope of this study. Therefore, we assumed a uniform distribution to randomly and independently produce the three linear scaling factors. They were empirically limited to be within a range of [0.8, 1.2], representing up to 20% decrease or increase in the corresponding brain dimension relative to that of the generic WHIM. Following a dimensional analysis, the scaling factors for the brain volume would be the multiplication of the three linear scaling factors.[17] **Figure 1** compares the generic WHIM with three representative scaled models.

Some scaled models may not yield a plausible real-world brain model (e.g., an extreme case would be to have a scaling factor of 0.8 along the *x* and *y* axes coupled with a scaling factor of 1.2 along the *z* direction). Nevertheless, all of the scaled models and the associated impact simulations would be useful to probe the impact-strain response hypersurface in neural network training. This was analogous to generating augmented impact kinematic profiles that may not be physical in the real world for producing neural network training data.[32]

[Figure 1]

Augmented impact profiles to generate training samples

Three real-world impact datasets were available for this study. They included laboratory reconstructed National Football League (NFL) impacts (N=53),[42] mouthguard-measured on-field impacts from various sports (SF; N=110)[43] and from American High School Football (HF; N=314).[44] The average length of temporal window of recorded non-zero rotational velocity profiles was 88±63 ms (range of 18–240 ms) for the NFL dataset. In comparison, the temporal length for the SF and HF datasets were fixed to 97 ms and 50 ms, respectively.[44] It has been shown that the deep brain corpus callosum has reached peak strains for a significantly more impacts in the first two datasets than in the third (79% and 88% vs. 34%). As the corpus callosum usually reaches peak strains the last, these observations suggest that significantly more head impact profiles in the first two datasets have sufficient durations to capture the complete kinematic events in modeling than the third. Therefore, we used the NFL and SF datasets usually of longer impact durations to

generate training samples while the HF dataset of a shorter duration was used for independent testing.

Similar to previous studies,[32,33] data augmentation was used to increase the number of training samples. Briefly, it involves permuting the *x*, *y*, and *z* components of the kinematic rotational profile, randomly rotating the rotational axis (assessed at the peak resultant rotational velocity magnitude) about the head center of gravity, and random scaling of the rotational velocity magnitude so that its peak was within a range of 2–40 rad/s. The range was determined from measured on-field impacts.[45] Each batch of data augmentation generated six times the number of profiles in the dataset (3! = 6).

For each NFL/SF dataset, two batches of augmented impacts were generated. This led to a total of 1956 impact rotational velocity profiles (53×6×2+110×6×2). Similarly,[33] "outliers" with out-of-range peak velocity (N=31), initial velocity greater than 10% of peak velocity (N=292), and peak acceleration below 10th percentile or over 90th percentile (N=365) were removed. They led to a total of 1363 impact cases (the same outlier could occur in multiple criteria). They were paired with the randomly scaled brain models for impact simulation in Abaqus/Explicit (Version 2018; Dassault Systèmes, France).

Data preprocessing

For kinematic input, the rotational velocity profile and the corresponding acceleration profile (determined by forward differentiation of velocity and scaled to 1%; both at a 1 ms temporal resolution) were concatenated. Explicitly combining velocity and acceleration as input was necessary (same for a fully connected neural network[37]), as the current CNN architecture may not be effective in deriving one from the other (but both are important to brain deformation).[33] The concatenated signals were then padded into 200 ms in length and were further shifted so that the resultant peak rotational velocity occurs at 100 ms.[33] Finally, the resulting 6-by-201 matrix data were uniformed scaled by a factor of 40, which led to a data range of approximately [-1 1]. This is typical for data normalization.[46]

Directly providing the three scaling factors as inputs to the CNN was not optimal as their values did not have the same data range relative to the kinematic data. Instead, they

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

were similarly transformed into a data range of [-1 1]. With trial and error to finetune parameters for the mapping of scaling factors, we found that transforming into a range of [-2, 2] further improved performance (e.g., by 5% in "success rate" as defined later in baseline training for the best fold in 10-fold cross-validation). Therefore, the following transformation was employed to each scaling factor, s, before providing to the CNN as input (s_{CNN}):

$$s_{CNN} = 10 \times (s - 1)\,, \tag{1}$$

Voxelized deformation field

In order to promote data sharing without the need to access a biomechanical model, itself,[14] we adopted a meshfree postprocessing technique[39] to generate voxelized peak maximum principal strain (MPS) for response representation. The resampling effectively applies an average filter among peak strains at neighboring element centroids, which would mitigate numerical artefacts associated with peak strains from model simulation (thus, avoiding the need to compute the 95th or 90th percentile peak strains commonly used). An isotropic spatial resolution of 4 mm was chosen for the generic model, which was consistent with our previous work.[35] **Fig. 2** compares FE elementwise MPS and the voxelized counterpart. The same deformation voxelization strategy was applied to the scaled brain models, with spatial resolutions along the three anatomical axes appropriately scaled to account for the difference in brain morphology. This approach retains the same image volume dimensions among various scaled or generic brain models, which is necessary to preserve the same CNN output size for training and prediction.

[Figure 2]

CNN architecture

The earlier generic CNN architecture[33] was modified to accept three morphological scaling factors as additional inputs. The CNN convolutional layers and pooling operations output a series of feature maps that are flattened into a 1D vector. They were expected to generate strain patterns for the generic brain model. Given that the scaling factors are anticipated to alter brain strain magnitude, but not strain pattern in a significant way,[17] the

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

three morphological scaling factors were provided to the flattening layer after the convolutional layers as inputs. In fact, providing the scaling factors as inputs to any of the three fully connected layers could also inform the neural network to adjust strain magnitudes resulting from brain morphological variations.

After trial and error, we identified that providing the scaling factors to the flattening layer led to the highest performance (e.g., SR increased by 5% compared to adding scaling factors to the second fully connected layer in baseline training for the best fold in a 10-fold cross-validation). Finally, the output size for the last fully connected layer was also modified to match the number of resampled voxel centroids corresponding to the brain parenchyma (N=20043 voxels in an image volume of dimension of 36-by-47-by-35, after removing non-brain regions). **Fig. 3** shows the final architecture of the resulting "individualized CNN".

[Figure 3]

CNN training

Compared to the previous work,[33,35] the number of training samples in this study was significantly fewer (~1.4 k vs. ~5.7 k). To mitigate the challenge in training, we adopted transfer learning, which has been successfully applied to a CNN designed for rapid brain strain estimation for automotive impacts.[34] Specifically, the previously simulated impacts based on the augmented NFL and SF datasets using the generic WHIM[33] (~5.7 k) were first employed to retrain the same earlier CNN architecture for outputting voxelwise peak strains at a 4 mm isotropic resolution. The resulting pretrained CNN provided initial weights for further training (all layers unfrozen) using the impact-response samples from the scaled brain models. To evaluate the effectiveness of transfer learning, a separate CNN with random initial weights was also trained using impacts from the scaled brain models ("baseline training"). Mean squared error (MSE) was used to define the following loss function:

$$loss = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{M}\sum_{j=1}^{M}\left(x_{ij} - y_{ij}\right)^2\right), \quad (2)$$

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

where x_{ij} and y_{ij} are the CNN estimated peak MPS value and the directly simulated counterpart for the j-th voxel of the i-th training sample, respectively. M is the number of brain voxels in the resampled image representation, and N is the number of training samples.

The same batch size of 256 was adopted.[33] The number of training epochs of 500 was determined from early stopping to avoid overfitting[34] (see **Fig. A1** in Appendix for a typical learning curve). With transfer learning, the learning rate was usually set to be lower than the pretrained model (e.g., 0.0001 or 10% of that in the pretrained model).

Performance evaluation: 10-fold cross-validation using training impact data

We adopted 10-fold cross-validation to assess the individualized CNN prediction performance. The impact-response samples based on the random scaled brain models and random impact kinematics were divided into 10 approximately equal subsets. Nine subsets were used for training and the remaining subset unseen by the training process was used for validation. The process was repeated 10 times until each subset was used for testing exactly once. After the cross-validation, all samples were combined to train a single individualized CNN for further accuracy assessment.

Linear regression slope (*k*) and Pearson's coefficient (*r*) between CNN-predicted voxelized MPS and those resampled from simulation results were calculated. They offer the convenience to intuitively indicate an overall over- or under-estimation in magnitude (*k* greater or less than 1.0) and the similarity in strain spatial pattern (*r*), respectively. A prediction was considered sufficiently accurate when both *k* and *r* did not deviate from their "perfect" score of 1.0 by more than 0.1, as adopted before.[33] A success rate (SR) was defined as the percentage of impact cases that the CNN estimation was sufficiently accurate. To provide a sense of error magnitude (e.g., relative to WHIM injury threshold of 0.2 previously established)[47], root mean squared error (RMSE) was also reported. Compared to *k*, RMSE loses the ability of informing over- or under-estimation, and does not effectively differentiate relative error for high vs. low strains. Additionally, stratified random sampling was also explored to confirm that the individualized CNN is stable across different datasets (**Table A1** in Appendix).

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

Performance evaluation: generic model responses using testing impact data

To ensure that the individualized CNN remains accurate when predicting responses for the generic brain model, three scaling factors of 1.0 (prior to transformation; Eqn. 1) were provided as inputs. The HF dataset was used for independent testing, where *k* and *r* were obtained for each impact, along with SR for all the impacts.

However, the HF impacts had a relatively shorter time window than the NFL/SF datasets, and significantly more impacts were considered inadequate to capture the complete kinematic events (66.2% vs. 20.8% and 11.8% for NFL and SF dataset, respectively).[44] Therefore, we also investigated whether this had a role in CNN prediction performance. Specifically, we excluded HF impacts if the peak resultant rotational velocity occurred within the last 5 ms relative to the temporal window right-handed boundary (as empirically used earlier).[44] They were considered to *not* have captured the complete kinematic event because all real-world head impacts are expected to include a deceleration phase to reach to a zero velocity. **Figure 4** illustrates sample HF impact profiles and their distribution of temporal peak locations.

[Figure 4]

Performance evaluation: scaled subject-specific model responses using testing impact data

For feasibility considerations, eleven subjects (same as before[17]) with brain volumes ranging from the smallest to the largest among 191 college and high school athletes (141 males aged 14–25 and 50 females aged 18–24; approved by the Institutional Review Board at Dartmouth College) were selected for evaluation.[17] The subject recruitment and neuroimage acquisitions (T1-weighted MR images; an isotropic resolution of $1.5\ \text{mm} \times 1.5\ \text{mm} \times 1.5\ \text{mm}$, with image dimension of $112 \times 171 \times 171$) were part of the previous effort to investigate the biomechanical basis of mild TBI.[29] In a previous study,[17] each subject's MR images were rigidly registered to the template MR volume of the subject used to develop the generic WHIM.[29] This ensured a consistent measurement of brain length and width. Brain volume was calculated after segmentation, which was then used to determine "brain height" through a dimensional analysis, where brain

volume is expected to be correlated to the product of brain length, width, and height.[17] The three scaling factors for the selected brain models are provided in **Table 1**.

[Table 1]

Similarly, it was not feasible or cost-effective to simulate hundreds of impacts for each scaled subject-specific brain model for performance evaluation due to the substantial simulation runtime required. We hypothesized that if the individualized CNN was successful in estimating brain strains for the generic model and for a given impact profile, likely it would be successful for a scaled model when using the same impact and the corresponding scaling factors as inputs. Therefore, we strategically selected ten impact profiles (N=10), where the individualized CNN either succeeded (N=4) or failed (N=3) to estimate strains with sufficient accuracy, or the performance was at the borderline in the *k*-*r* space (N=3). The selected impact profiles were then used as inputs for each of the 11 scaled models for direct simulation (N=110).

Significance of scaling factors: parametric investigation

Finally, the individualized CNN was used to parametrically investigate how, each scaling factor affected brain strains. A baseline response was first estimated by setting all scaling factors to 1.0. Each scaling factor was then parametrically swept across its value range while keeping the other two at 1.0. Two successfully estimated impacts (randomly chosen from the four) were used as kinematic inputs. The resulting *k* and *r* relative to the baseline responses were calculated.

Given the significant correlation between brain size/volume and peak brain MPS,[16,21,25] we further investigated how the scaling factor of brain volume (*sv*) was related to *k* and *r* for whole brain voxelwise strains. To limit the investigation to "life-like" scaled models, a convex hull in the parametric space of (*sx*, *sy*, and *sz*) was established (**Fig. 5**) based on the scaling factors from the 11 subjects, as well as those of the generic model (scaling factors uniformly of 1.0). Combinations of random *sx*, *sy*, *sz* were generated from their value ranges (N=10,000), and only those fell within the convex hull were retained (N=639), from which *sv* was calculated as the multiplication of the three factors. The

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

corresponding scaled models were used to determine the relationship between *sv* and *k*/*r*. For comparison, the convex hull formed by the 191 subjects is also shown.

[**Figure 5**]

Data analysis

All head impacts were simulated using the anisotropic WHIM V1.0, which took ~30 min for one impact of ~100 ms duration (double precision with 15 central processing units [CPUs]; Intel Xeon E5-2698 with 256 GB memory). Another ~30 min was necessary to generate voxelized peak MPS. All data analyses were conducted using MATLAB (R2022b; MathWorks, Natick, MA). Statistical significance was reached when the *p* value was <0.05. In the following section, we first present results of various performance evaluations and then report the significance of each scaling factor.

Results

Performance evaluation: 10-fold cross-validation using training impact data

Fig. 6 summarizes *k*, *r*, RMSE and SR for each impact in the 10-fold cross-validation, using either the baseline training or transfer learning. The baseline training had an SR of 69.3%, which was improved to 86.2% with transfer learning.

[Figure 6]

Performance evaluation: generic model responses using testing impact data

Fig. 7 summarizes *k*, *r*, RMSE and SR for the HF dataset when using the individualized CNN for estimating brain strains for the generic model (i.e., with scaling factors of 1.0). An SR of 72.9% was obtained for the whole HF dataset, which improved to 92.1% when limiting to impacts that were considered to have captured the complete kinematics events. This result suggests higher accuracy to follow when the testing and training datasets both capture brain peak strains in model simulation. The ten impacts selected for performance evaluations for the 11 scaled subject-specific models are

identified (**Fig. 7a**), with six of them labeled for further illustration (for visualization considerations). Their rotational velocity and acceleration profiles are reported in the Appendix **(Fig. A2–A7)**.

[Figure 7]

Performance evaluation: scaled subject-specific model responses using testing impact data

For all ten selected impacts with estimation either successful, failed, or at the borderline for the generic model (**Fig. 7a**), the same happened for the individualized CNN for the scaled models. **Fig. 8** summarizes results for predicting 11 scaled models (**Table 1**) using each of the six labeled impacts as kinematic input. Without scaling to account for morphological variation (**Fig. 8b**), the *k*-*r* coordinates of the scaled models along with that of the generic model were largely aligned along a line for each given impact, with *k* linearly increasing with the decrease of brain volume. However, no considerable variation of *r* happened, especially for "successful" impacts, suggesting that a similar strain pattern was retained when not accounting for brain morphological variations.

[Figure 8]

Fig. 9 shows the directly simulated peak MPS distribution for the odd-numbered scaled models using impact #1 in **Fig. 7a** as kinematic input (even-numbered subjects in **Fig. A8**). Their differences from estimation by the individualized CNN are also shown, when using either the appropriate scaling factors as additional inputs or uniformly of 1.0 without considering morphological variations. For the latter, over- or under-estimation was apparent for smaller or larger brains, respectively. This finding highlights the effectiveness of the individualized CNN to account for brain morphological differences.

[Figure 9]

Significance of scaling factors: parametric investigation

Fig. 10 reports how, each individual scaling factor affected the MPS prediction of the individualized CNN using a representative impact as kinematic input. The resulting *k*

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

and *r* were computed relative to the reference response obtained when setting all scaling factors uniformly to 1.0. In addition, the relationship between brain volume ratio relative to the generic WHIM (i.e., *sv*) and reference responses is also shown, for random scaling factors falling within the convex hull in the parametric space (**Fig. 5**). For all scaling factors including the volume ratio, increasing their values also increased *k*, or overall whole-brain strain magnitude. However, the strain pattern remained rather similar, with minimum differences in *r*, and for both impacts (see **Fig. A9** in Appendix for an additional example). Nevertheless, the relative significance among the scaling factors (i.e., amount of increase) depended on the given head impact.

[Figure 10]

Discussion

We have successfully enhanced a convolutional neural network (CNN) brain injury model to further account for strain differences due to individual brain morphological variations (shape and size) relative to the generic anisotropic WHIM V1.0.[40] The latter represents a 60th percentile adult male head.[29] The outcome is a morphologically individualized CNN with improved subject-specificity that remains instantaneous in estimating peak strain distributions of the whole brain.

The individualized CNN takes three scaling factors along the anatomical axes as additional inputs to adjust brain strains. A notable advantage with this approach is that no neuroimages are necessary to approximate individualized brain models, and without significant degradation of strain accuracy[17] relative to the "morphed models" created from a more sophisticated neuroimage-based mesh warping.[29] When subject-specific information such as head dimensions and age/sex is not available, the individualized CNN can still be used in place of the generic CNN, by setting the scaling factors uniformly to 1.0 (**Fig. 7**). Certainly, the model accuracy is expected to improve when neuroimages are otherwise available to directly determine scaling factors relative to WHIM, without relying on the regression models.

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

Training and testing datasets

Recent advancement in brain modeling has identified the importance of sufficient temporal duration for head impact profiles to ensure that brain strains reach peaks within the simulation time window.[25,44] The inconsistency among the three impact datasets used for training and testing appears important for the accuracy disparity between cross-validation (**Fig. 6**) and independent testing (**Fig. 7**). The majority of impacts from the NFL and SF datasets (~80–90%) were considered to have captured the kinematic events completely in model simulation, vs. ~34% in the HF dataset.[44] It was not surprising that after limiting HF impacts to those considered to have captured the complete kinematic events similarly to training data, the success rate (SR) considerably increased (from 72.9% to 92.1%; **Fig. 7**).

The importance of ensuring impact profiles to capture the complete kinematic event in terms of peak strains was more obvious when using the same CNN architecture to train and predict peak strain distributions separately for the surface layer brain voxels and for the corpus callosum deep in the brain (**Fig. A10** and **Fig. A11** in Appendix). The former achieved an SR of 91%, which was considerably higher than the latter of 63% in cross-validation. The brain surface and deep regions are usually the first and last, respectively, to reach peak strains in an impact simulation. Therefore, the performance difference was most likely because many impacts did not allow the corpus callosum region to reach peak strains in simulation. On the other hand, the two regions had a rather similar RMSE value overall (0.014 vs. 0.015); confirming that this accuracy metric may not be effective in differentiating performance between high vs. low strains.

It should be recognized that the discrepancy in estimation performance of peak strains is not likely to apply to the recent TNN and another CNN that predict the entire temporal evolution of brain deformation, including strain.[35] As long as impacts of sufficiently long durations capturing the complete kinematic events are used for training, strain estimation would likely still remain accurate for testing impacts, even of shorter durations. However, they may lead to strain time histories not reaching peak responses, which is the result of the impact data, themselves, independent of the TNN or CNN model.

The opposite (i.e., impacts of shorter durations for training but for impacts of longer durations for testing), nevertheless, may not be true, which was why the HF dataset was not used for training in this study. To summarize, longer duration impacts are recommended to serve as training data for developing deep learning brain models.

CNN accuracy performance

Transfer learning was effective at improving the accuracy of the individualized CNN, increasing the cross-validation SR from 69.3% (with random initial weights) to 86.2% for scaled model responses (**Fig. 6**). Nevertheless, the performance was notably poorer than the earlier generic CNN in cross-validation (vs. SR of 92–97%).[33] This was likely because of the much fewer training samples used here (~1.4 k vs. ~5.7 k). In terms of RMSE, all testing had a similar range of 0.012–0.017, representing 6–8.5% of the injury threshold of 0.2.

While more training samples could further improve accuracy, it is worth investigating how to generate the training samples more intelligently—using the fewest training samples to achieve a desirable accuracy.[48] This line of work may also explain why the CNN failed for some impacts, even though their peak velocity/acceleration magnitudes were unremarkable. In fact, the earlier CNN based on training data with peak rotational velocity capped at 40 rad/s remains accurate for testing data with peak rotational velocity up to ~80 rad/s.[33] In this study, for impacts that were successful predicted or failed to predict when treating them in a generic model, the individualized CNN also largely followed the same trend (**Fig. 8**). This suggests that the actual impact profile shapes,[49,50] or, the physical processes of head acceleration/deceleration event play a critical role in the CNN accuracy. Similarly, it is also important that impact profiles have sufficient durations to capture the complete kinematic events so that brain strains reach peak values in model simulation. The CNN performance would benefit when both the training and testing impacts are of sufficient durations (**Fig. 7**). These findings are not surprising, given the causal relationship between impact physical event and the resulting brain strains. It is, thus, recommended that future deep learning brain injury models fully consider impact biomechanical physics (vs. various kinematic peaks, alone) to maximize performance.

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

Morphological scaling factors on strain over- or under-estimation

Morphological scaling factors mostly affected regression slope, *k*, or overall strain magnitude (**Fig. 8**). Without them, the individualized CNN would over- or under-estimate peak MPS for smaller or larger brains relative to the generic WHIM, respectively, as expected (**Figs. 8** and **9**). Although estimations remained "successful" for most scaled models when using the two "successful" impacts as inputs, accounting for morphological variations considerably improved accuracy, even with a more stringent success criterion (i.e., both *k* and *r* not deviating from the perfect score of 1.0 by more than 0.05; **Fig. 8**). However, scaling factors had little effect on strain pattern as characterized by *r*, which was consistent with the parametric analysis (**Fig. 10** and **Fig. A9**).

Interestingly, *sx* and *sz* (along the anterior-posterior and inferior-to-superior direction, respectively) had the largest and least effect, respectively. This was likely because the corresponding brain length and height is usually the longest and shortest, respectively. Therefore, with the same percentage of increase/decrease, *sx* and *sz* would lead to larger or smaller magnitude increase or decrease of the corresponding brain dimension. Therefore, they would be expected to result in larger or smaller change in brain strain as well. Obviously, the magnitude of this change depends on the actual impact used as input, which explains why the selected two impacts had slightly different magnitudes of effects from the scaling factors as well as that from brain volume ratio (**Fig. 10** and **Fig. A9**).

Postprocessing of brain biomechanical responses with neuroimage awareness

Neuroimage-based brain model development and enhancement are common.[12,14] However, model response postprocessing has largely relied on strain output directly from model simulations, e.g., peak strain of the whole brain[51] and fiber/axonal strain through elementwise strain tensor projection.[29,52,53] The inherent mesh-image mismatch poses a challenge in translating biomechanical strains into a native neuroimage space for multimodal injury correlation and subsequent analysis.[40,54] The challenge is obliviated by processing brain deformation such as strain, stress, and strain rate directly in a

neuroimage space, which would then facilitate multimodal analysis between biomechanics and neuroimaging.

A voxelized brain displacement field allows easily deriving an image representation of strain, strain rate, and the complete strain tensor over time in a high-dimensional matrix form. By properly adjusting the spatial resolution, the same image dimension can be retained across subjects (generic vs. scaled brain models). When using a common neuroimage atlas such as the MNI (Montreal neurological imaging) for brain response resampling across different research groups, all brain models would conform to a common standard for response representation. This may significantly promote data sharing and collaboration among research groups, as no explicit access is necessary to the brain model, itself.[14]

Comparison of deep learning brain injury model architectures

To date, several deep learning brain injury models have been developed based on either a CNN[32–35], a TNN[35], a U-Net[36], or a fully connected neural network[37,38] architecture. Models that use CNN, TNN, or U-Net employ kinematic temporal profiles as input, emulating a two-dimensional image that these architectures were originally designed for. They retain the complete kinematic information required for direct FE model simulation and thus, may achieve a high estimation accuracy. However, their fixed input dimension requires kinematic input resampling or retraining the network when using data of a different temporal resolution. Their architectures are also relatively more complex and thus, more challenging to train than a fully connected neural network. The latter typically uses “engineered features” from kinematics, such as combinations of extrema magnitudes of accelerations, velocities, and other variants including integrations and differences. However, the extracted features could lose relative temporal information about the physical process of impact important to brain deformation. As a result, fully connected neural networks may suffer in accuracy. On the other hand, a TNN uses an attention mechanism to correlate history-dependent brain deformation, which is highly accurate in estimation (R^2 close to 1.0).[35] Nevertheless, this architecture requires significantly more computing resources for training. In comparison, a multi-task CNN using a one hot vector

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

for task representation achieves a comparable (though, slightly degraded) accuracy but is much lighter weight. More detailed comparisons of CNN/TNN models and with architectures used in other research fields can be found in Ref.[35]

Nevertheless, deep learning brain injury models have only emerged recently.[14] More efforts are necessary to compare their performances and optimize strategies for generating training samples, as each sample remains rather costly to produce from direct FE model simulation.

Limitations

Due to the fewer training samples compared to the previous work, the individualized CNN suffered in accurate compared to the generic CNN. Future work should investigate how best to design a minimum training dataset that is most effective in achieving a desirable accuracy.[48] For this reason, we also limited the scope of our study to estimating element-/voxel-wise peak strains of the whole brain, rather than the complete spatiotemporal details of brain strain now achievable with a TNN or a multi-task CNN.[35] Nevertheless, spatially detailed peak strains of the whole brain still supersede others that report a single peak strain value (e.g., 90th or 95th percentile peak strains), which is incapable of informing the location of occurrence.[14] Computing the latter from the former is trivial.

Second, the three scaling factors used to define head dimensions in training samples were generated randomly from a relatively large range of 0.8–1.2 (vs. within 0.9–1.1 for the 11 scaled models; **Table 1**). A more definitive statistical model among the scaling factors (vs. independent here), if known, "may" help establish a more effective training dataset in the future, with the caveat of costly re-simulations of many head impacts. Still, this hypothesis needs to be tested, as the earlier CNN model remains accurate even when testing on impacts quite different from the training samples (e.g., peak rotational velocity of ~80 rad/s, vs. peak magnitudes capped at 40 rad/s in training)[33] and the performance does not seem to depend on how peak kinematics are distributed in a parametric space.[34]

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Finally, we have only focused on peak MPS responses in this study. Individual variations in fiber/axonal strains will be tackled in the future, as this strain measure will necessarily require subject-specific diffusion tensor images to account for individual differences of white matter tracts at greater anatomical details.[55]

Conclusion

Despite a modest training dataset, the individualized CNN achieves reasonable accuracy for instantaneous predictions of voxelwise peak maximum principal strains. The individualized CNN improves model subject specificity by taking three morphological scaling factors as additional inputs. The scaling factors can be determined based on measurements of head dimensions, along with subject's sex and age information, and notably, without neuroimages necessary (certainly, accuracy anticipated to improve when neuroimages are otherwise available). Therefore, the individualized CNN has the potential to facilitate large-scale impact simulations, either for a generic brain or on a subject-specific basis. This could be especially useful for brain injury studies of youth and female athletes due to anticipated larger morphological variations relative to the generic, adult male model. The voxelized strain output also allows for convenient data sharing and promotes collaboration among research groups.

Acknowledgments

Funding is supported by the National Science Foundation (NSF) under Award Number 2114697.

Conflict of interest

No competing financial interests exist.

Author contribution

N. Lin: Conceptualization, Methodology, Software, Writing – original draft; S. Wu.: Conceptualization, Methodology, Reviewing; S. Ji: Conceptualization, Supervision, Writing, Reviewing and Editing, Funding acquisition.

References:

1. Carroll LJ, Cassidy JD, Cancelliere C, et al. Systematic Review of the Prognosis after Mild Traumatic Brain Injury in Adults: Cognitive, Psychiatric, and Mortality Outcomes: Results of the International Collaboration on Mild Traumatic Brain Injury Prognosis. Arch Phys Med Rehabil 2014;95(3 SUPPL); doi: 10.1016/j.apmr.2013.08.300.

2. Dewan MC, Rattani A, Gupta S, et al. Estimating the global incidence of traumatic brain injury. J Neurosurg 2019;130(4):1080–1097; doi: 10.3171/2017.10.JNS17352.

3. Jang, I. I, Chun Y, Brosch JR, et al. Every hit matters: white matter diffusivity changes in high school football athletes are correlated with repetitive head accel- eration event exposure. NeuroImage Clin 2019;24:101930.

4. Laker SR. Epidemiology of Concussion and Mild Traumatic Brain Injury. PM&R 2011;3(10 SUPPL. 2):S354–S358; doi: 10.1016/j.pmrj.2011.07.017.

5. Dompier TP, Kerr ZY, Marshall SW, et al. Incidence of concussion during practice and games in youth, high school, and collegiate American football players. JAMA Pediatr 2015;169(7):659–665; doi: 10.1001/jamapediatrics.2015.0210.

6. Graham R, Rivara FP, Ford MA, et al. Sports-Related Concussions in Youth. 2014.; doi: 10.1001/jama.2013.282985.

7. Coronado VG, Haileyesus T, Cheng TA, et al. Trends in sports-and recreation-related traumatic brain injuries treated in US emergency departments: The National Electronic Injury Surveillance System-All Injury Program (NEISS-AIP) 2001-2012. J Head Trauma Rehabil 2015;30(3):185–197; doi: 10.1097/HTR.0000000000000156.

8. CDC. Nonfatal Traumatic Brain Injuries Related to Sports and Recreation Activities Among Persons Aged ≤19 Years — United States, 2001–2009. 2011.

9. Tanaka MJ, LiBrizzi CL, Rivenburgh DW, et al. Changes in U.S. girls' participation in high school sports: implications for injury awareness. Phys Sportsmed 2021;49(4):450–454; doi: 10.1080/00913847.2020.1852861.

10. Iverson GL, Gardner AJ, Terry DP, et al. Predictors of clinical recovery from concussion: a systematic review. Br J Sports Med 2017;51(12):941–948; doi: 10.1136/bjsports-2017-097729.

11. Ferdinand Pennock K, McKenzie B, McClemont Steacy L, et al. Under-Reporting of Sport-Related Concussions by Adolescent Athletes: A Systematic Review. Int Rev Sport Exerc Psychol 2020;1–27; doi: 10.1080/1750984X.2020.1824243.

12. Madhukar A, Ostoja-Starzewski M. Finite Element Methods in Human Head Impact Simulations: A Review. Ann Biomed Eng 2019;1–23; doi: 10.1007/s10439-019-02205-4.

13. Yang KH, Hu J, White NA, 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 2006;50:429–490; doi: https://doi.org/10.4271/2006-22-0017.

14. Ji S, Ghajari M, Mao H, et al. Use of brain biomechanical models for monitoring impact exposure in contact sports. Ann Biomed Eng 2022;50:1389–1408; doi: 10.1007/s10439-022-02999-w.

15. Li X. Subject-Specific Head Model Generation by Mesh Morphing: A Personalization Framework and Its Applications. Front Bioeng Biotechnol 2021;9; doi: 10.3389/fbioe.2021.706566.

16. Liu JL, Jin J (Judy), Eckner JT, et al. Influence of Morphological Variation on Brain Impact Responses among Youth and Young Adults. J Biomech 2022;135:111036; doi: 10.1016/J.JBIOMECH.2022.111036.

17. Wu S, Zhao W, Wu Z, et al. Approximating subject-specific brain injury models via scaling based on head–brain morphological relationships. Biomech Model Mechanobiol 2022 2023;22(1):159–175; doi: 10.1007/S10237-022-01638-6.

18. Rollins JD, Collins JS, Holden KR. United States Head Circumference Growth Reference Charts: Birth to 21 Years. J Pediatr 2010;156(6):907-913.e2; doi: 10.1016/j.jpeds.2010.01.009.

19. Kleiven S, von Holst H. Consequences of head size following trauma to the human head. J Biomech 2002;35(2):153–60; doi: 10.1016/S0021-9290(01)00202-0.

20. Zhao W, Ruan S, Ji S. Brain pressure responses in translational head impact: a dimensional analysis and a further computational study. Biomech Model Mechanobiol 2015;14(4):753–766; doi: 10.1007/s10237-014-0634-0.

21. Li X, Zhou Z, Kleiven S. An anatomically accurate and personalizable head injury model: Significance of brain and white matter tract morphological variability on strain. Biomech Model Mechanobiol 2020;1–29; doi: 10.1101/2020.05.20.105635.

22. Liu Y, Domel AG, Yousefsani SA, et al. Validation and Comparison of Instrumented Mouthguards for Measuring Head Kinematics and Assessing Brain Deformation in Football Impacts. Ann Biomed Eng 2020;48(11):2580–2598; doi: 10.1007/s10439-020-02629-3.

23. Mao H, Zhang L, Jiang B, et al. Development of a finite element human head model partially validated with thirty five experimental cases. J Biomech Eng 2013;135(11):111002–15; doi: 10.1115/1.4025101.

24. Miller LE, Urban JE, Stitzel JD. Development and validation of an atlas-based finite element brain model model. Biomech Model 2016;15(5):1201–1214; doi: 10.1007/s10237-015-0754-1.

25. Liu Y, Domel AG, Cecchi NJ, et al. Time Window of Head Impact Kinematics Measurement for Calculation of Brain Strain and Strain Rate in American Football. Ann Biomed Eng 2021;1–14; doi: 10.1007/s10439-021-02821-z.

26. Lu YC, Daphalapurkar NP, Knutsen AK, et al. A 3D Computational Head Model Under Dynamic Head Rotation and Head Extension Validated Using Live Human Brain Data, Including the Falx and the Tentorium. Ann Biomed Eng 2019;47(9):1923–1940; doi: 10.1007/s10439-019-02226-z.

27. Combest JJ. Current Status and Future Plans of the GHBMC (Global Human Body Models Consortium). Hum Model Simul Automot Eng 2016;1–40.

28. Toyota Motor Corporation. TOTAL HUMAN MODEL FOR SAFETY (THUMS): REVOLUTIONIZING CRASH SIMULATION TO SUPPORT SAFE MOBILITY FOR ALL. 2021.

29. Ji S, Zhao W, Ford JC, et al. Group-wise evaluation and comparison of white matter fiber strain and maximum principal strain in sports-related concussion. J Neurotrauma 2015;32(7):441–454; doi: 10.1089/neu.2013.3268.

30. Giudice JS, Alshareef A, Wu T, et al. An Image Registration-Based Morphing Technique for Generating Subject-Specific Brain Finite Element Models. Ann Biomed Eng 2020;1–13; doi: 10.1007/s10439-020-02584-z.

31. Chen Y, Ostoja-Starzewski M. MRI-based finite element modeling of head trauma: spherically focusing shear waves. Acta Mech 2010;213(1–2):155–167; doi: 10.1007/s00707-009-0274-0.

32. Wu S, Zhao W, Ghazi K, et al. Convolutional neural network for efficient estimation of regional brain strains. Sci Rep 2019;9:17326; doi: https://doi.org/10.1038/s41598-019-53551-1.

33. Ghazi K, Wu S, Zhao W, et al. Instantaneous Whole-Brain Strain Estimation in Dynamic Head Impact. J Neurotrauma 2021;38(8):1023–1035; doi: 10.1089/neu.2020.7281.

34. Wu S, Zhao W, Barbat S, et al. Instantaneous brain strain estimation for automotive head impacts via deep learning. Stapp Car Crash J 2021;65:139–162.

35. Wu S, Zhao W, Ji S. Real-time dynamic simulation for highly accurate spatiotemporal brain deformation from impact. Comput Methods Appl Mech Eng 2022;394:114913; doi: 10.1016/J.CMA.2022.114913.

36. Bourdet N, Deck C, Trog A, et al. Deep Learning Methods Applied to the Assessment of Brain Injury Risk. In: Proceedings of International Research Conference on the Biomechanics of Impacts Online Virtual Conference; 2021; pp. 709–729.

37. Zhan X, Liu Y, Raymond SJ, et al. Rapid Estimation of Entire Brain Strain Using Deep Learning Models. IEEE Trans Biomed Eng 2021;9294(c):1–11; doi: 10.1109/TBME.2021.3073380.

38. Menichetti A, Bartsoen L, Depreitere B, et al. A Machine Learning Approach to Investigate the Uncertainty of Tissue-Level Injury Metrics for Cerebral Contusion. Front Bioeng Biotechnol 2021;9; doi: 10.3389/fbioe.2021.714128.

39. Ji S, Zhao W. Displacement voxelization to resolve mesh-image mismatch: application in deriving dense white matter fiber strains. Comput Methods Programs Biomed 2022;213:106528; doi: 10.1016/j.cmpb.2021.106528.

40. Zhao W, Ji S. White matter anisotropy for impact simulation and response sampling in traumatic brain injury. J Neurotrauma 2019;36(2):250–263; doi: 10.1089/neu.2018.5634.

41. Zhao W, Ji S. Displacement- and strain-based discrimination of head injury models across a wide range of blunt conditions. Ann Biomed Eng 2020;20(6):1661–1677; doi: 10.1007/s10439-020-02496-y.

42. Sanchez EJ, Gabler LF, Good AB, et al. A reanalysis of football impact reconstructions for head kinematics and finite element modeling. Clin Biomech 2018;64(February):82–89; doi: 10.1016/j.clinbiomech.2018.02.019.

43. Hernandez F, Wu LC, Yip MC, et al. Six Degree-of-Freedom Measurements of Human Mild Traumatic Brain Injury. Ann Biomed Eng 2015;43(8):1918–1934; doi: 10.1007/s10439-014-1212-4.

44. Ji S, Wu S, Zhao W. Dynamic characteristics of impact-induced brain strain in the corpus callosum. Brain Multiphysics 2022;3:100046; doi: 10.1016/j.brain.2022.100046.

45. Rowson S, Duma SM, Beckwith JG, et al. Rotational head kinematics in football impacts: an injury risk function for concussion. Ann Biomed Eng 2012;40(1):1–13; doi: 10.1007/s10439-011-0392-4.

46. Huang L, Qin J, Zhou Y, et al. Normalization Techniques in Training DNNs: Methodology, Analysis and Application. 2020;1–20.

47. Zhao W, Cai Y, Li Z, et al. Injury prediction and vulnerability assessment using strain and susceptibility measures of the deep white matter. Biomech Model Mechanobiol 2017;16(5):1709–1727; doi: 10.1007/s10237-017-0915-5.

48. Lin N, Wu S, Ji S. Stress-Test a Deep Learning Brain Injury Model Using Idealized Impact Kinematic Profiles. In: IRCOBI Porto, Portugal; 2022.

49. Zhao W, Ji S. Brain strain uncertainty due to shape variation in and simplification of head angular velocity profiles. Biomech Model Mechanobiol 2017;16(2):449–461; doi: 10.1007/s10237-016-0829-7.

50. Bian K, Mao H. Mechanisms and variances of rotation-induced brain injury: a parametric investigation between head kinematics and brain strain. Biomech Model Mechanobiol 2020;1–19; doi: 10.1007/s10237-020-01341-4.

51. Fahlstedt M, Meng S, Kleiven S. Influence of Strain post-processing on Brain Injury Prediction. J Biomech 2022;132:110940; doi: 10.1016/J.JBIOMECH.2021.110940.

52. Zhou Z, Li X, Liu Y, et al. Toward a Comprehensive Delineation of White Matter Tract-Related Deformation. J Neurotrauma 2021;38(23):3260–3278; doi: 10.1089/neu.2021.0195.

53. Sahoo D, Deck C, Willinger R. Development and validation of an advanced anisotropic visco-hyperelastic human brain FE model. J Mech Behav Biomed Mater 2014;33(1):24–42; doi: 10.1016/j.jmbbm.2013.08.022.

54. Zhou Z, Wang T, Jörgens D, et al. Fiber orientation downsampling compromises the computation of white matter tract-related deformation. J Mech Behav Biomed Mater 2022;132:105294; doi: 10.1016/J.JMBBM.2022.105294.

55. Giordano C, Zappalà S, Kleiven S. Anisotropic finite element models for brain injury prediction: the sensitivity of axonal strain to white matter tract inter-subject variability. Biomech Model Mechanobiol 2017;16(4):1269–1293; doi: 10.1007/s10237-017-0887-5.

Table 1. Morphological scaling factors, *sx*, *sy*, and *sz* along the three anatomical directions for 11 selected subjects whose brain volumes ranged from the smallest to the largest (from left to right) within the dataset. The scaling factors for brain volume (*sv*) are also shown.

Subject	1	2	3	4	5	6	7	8	9	10	11
Age/sex	23/F	21/F	18/F	18/F	22/M	22/M	21/M	19/M	20/M	19/M	19/M
sx	0.915	0.955	0.979	0.992	0.971	0.999	1.022	1.014	1.014	1.037	1.045
sy	0.949	0.995	0.979	1.016	1.071	1.053	1.021	1.059	1.059	1.049	1.056
sz	0.934	0.954	0.964	0.966	0.988	0.991	1.001	1.001	1.009	1.026	1.011
sv	0.811	0.907	0.923	0.973	1.027	1.043	1.047	1.081	1.084	1.116	1.117

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

Figure legends:

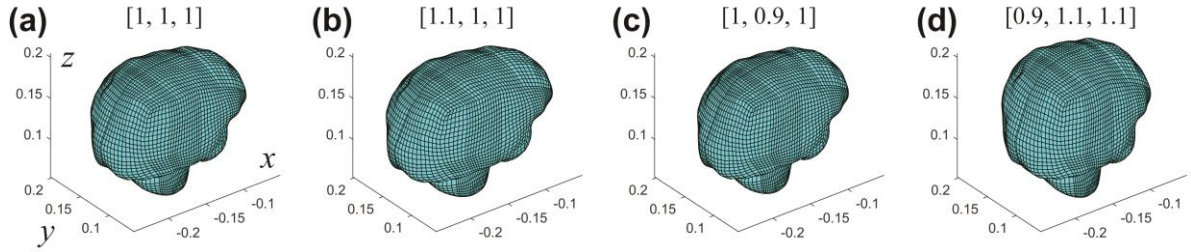

Figure 1. The generic anisotropic WHIM V1.0 (**a**) in comparison with three representative scaled models (**b**–**d**). Their corresponding scaling factors along the three anatomical axes are shown in brackets. Axis units are in meters.

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

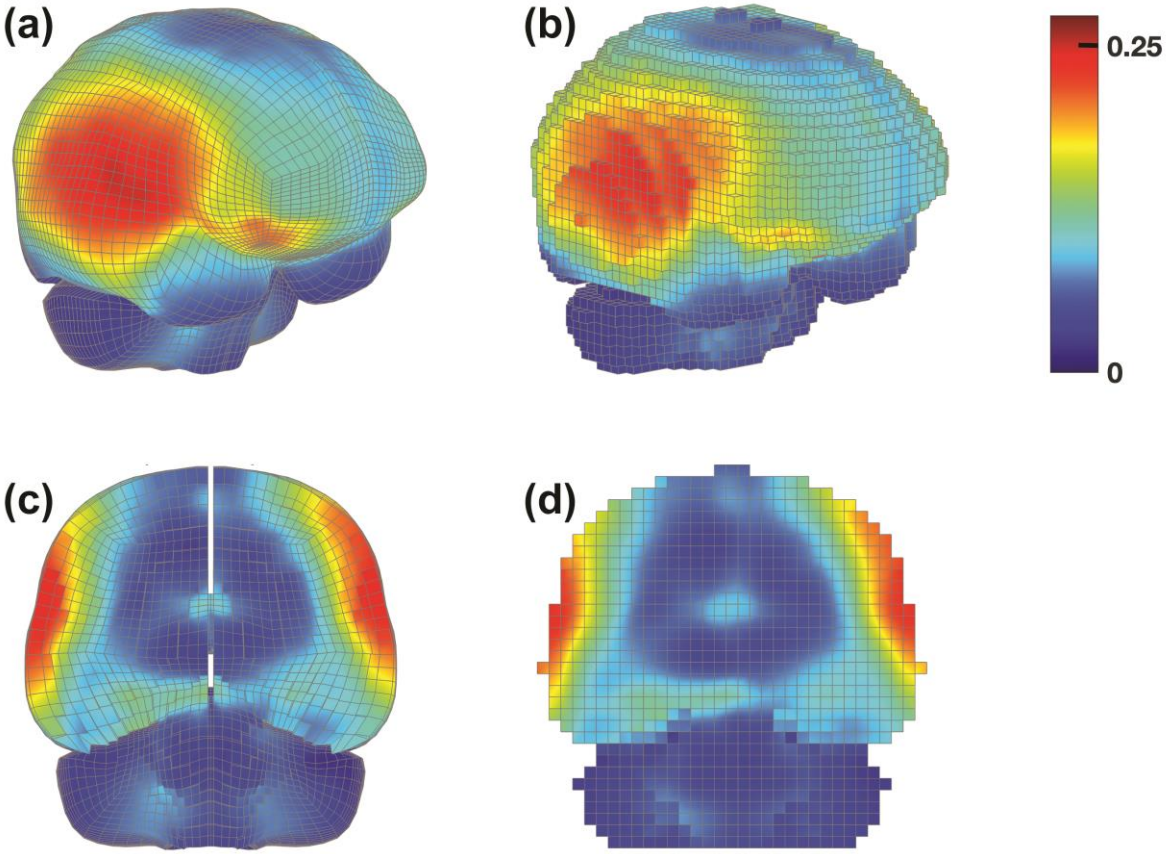

Figure 2. (**ac**) Elementwise MPS and its (**bd**) voxelized counterpart resampled at an isotropic spatial resolution of 4 mm for the generic WHIM, showing an oblique (top) and a coronal (bottom) view, respectively.

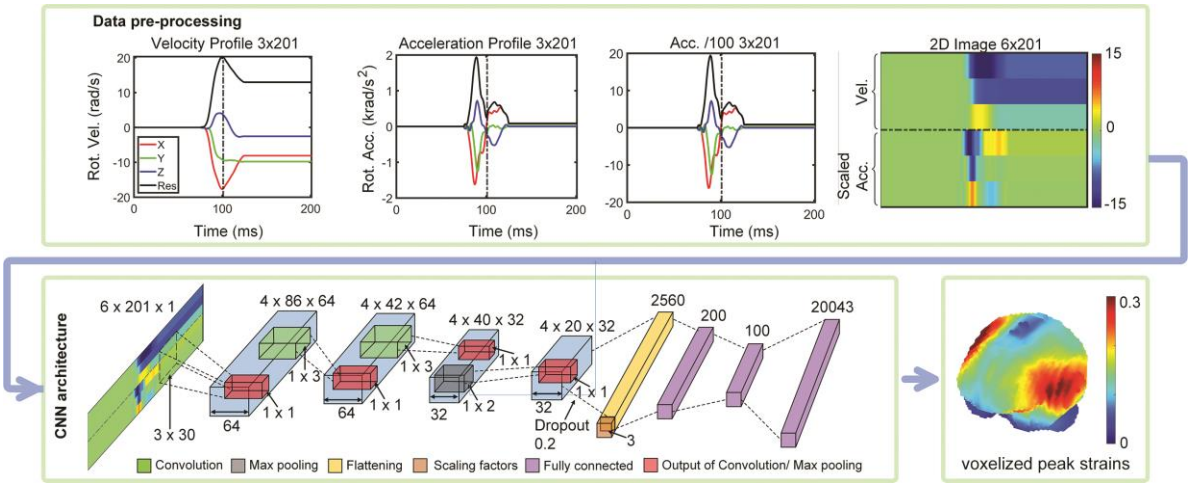

Figure 3. Architecture of the individualized CNN, where the three morphological scaling factors along the anatomical axes are provided as additional inputs to the flattening layer. The last output layer size is modified accordingly to produce voxelized peak MPS estimation of the brain. The scaling factors and the kinematic inputs are properly transformed to a data range of [-2, 2] and [-1 1], respectively, to facilitate CNN training and prediction.

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

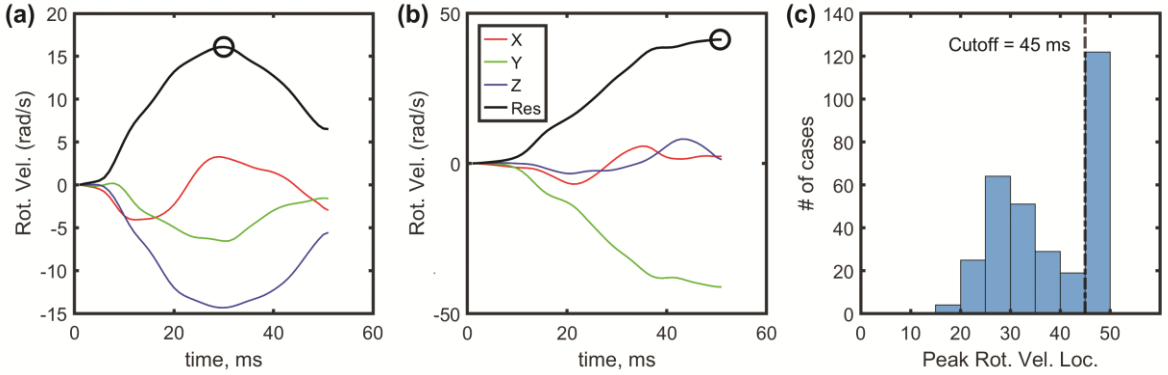

Figure 4. Illustration of HF impact profiles that are considered to have captured the complete kinematic event or not (**a** and **b**, respectively) based on the temporal peak location of the resultant rotational velocity profiles (circles). Histogram shows the distribution of the temporal peak locations (**c**).

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

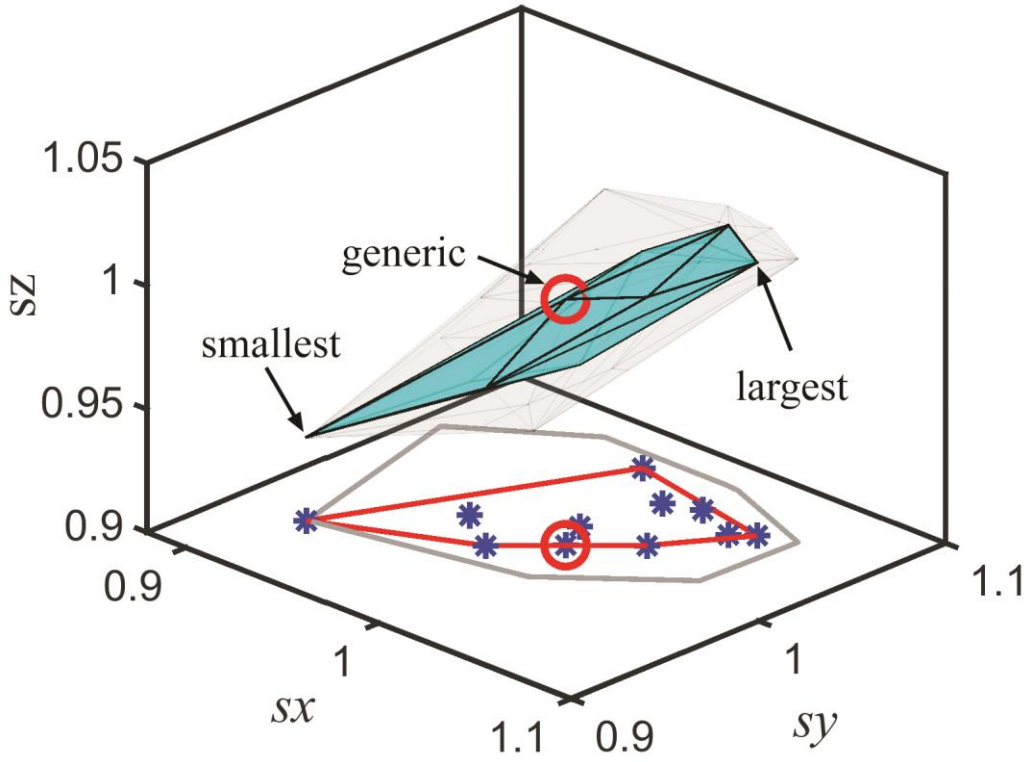

Figure 5. A convex hull in a 3D parametric space is formed by the three scaling factors from 11 subject-specific models (stars) and those of the generic model (circle). It is used to define an admissible range of or "life-like" scaled models for parametric investigations of the relationship between brain volume scaling factor, *sv*, and *k* and *r* (linear regression slope and Pearson correlation coefficient, respectively). The convex hull from the larger population of 191 subjects is also shown for comparison (gray boundaries). To improve visualization, 2D projections of the two hulls in the *sx*-*sy* plane are also shown.

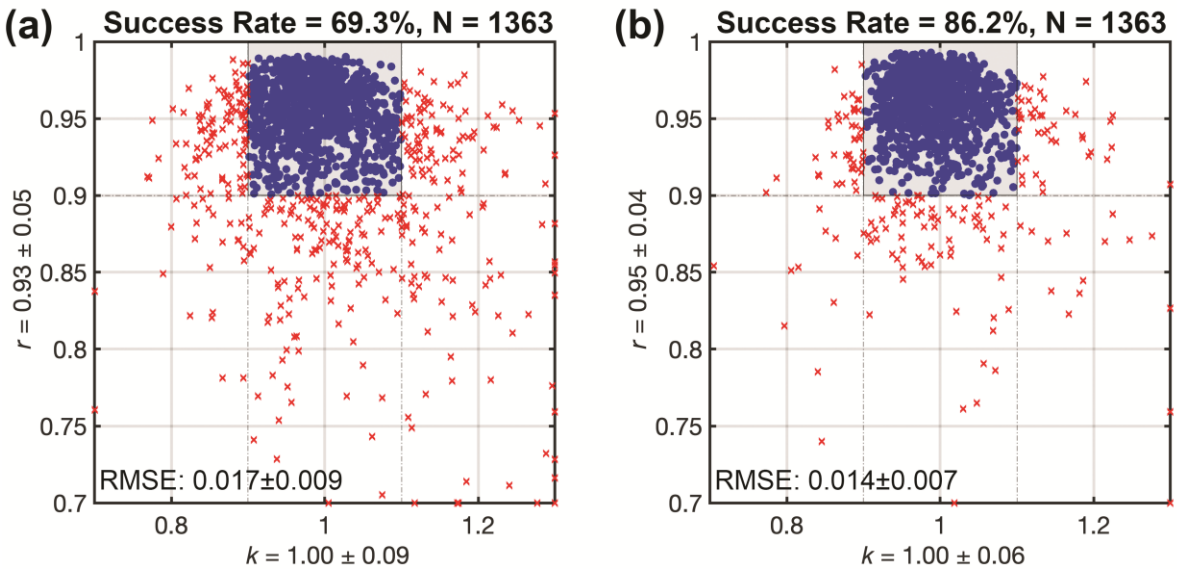

Figure 6. Summary of *k*, *r*, RMSE, and success rate (SR) for (**a**) baseline training and (**b**) transfer learning. To improve visualization, *k* and *r* values are capped. Points within the "shaded box" (i.e., when 0.9<*k*<1.1 and *r*>0.9) indicate successful estimations of sufficient accuracy.

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

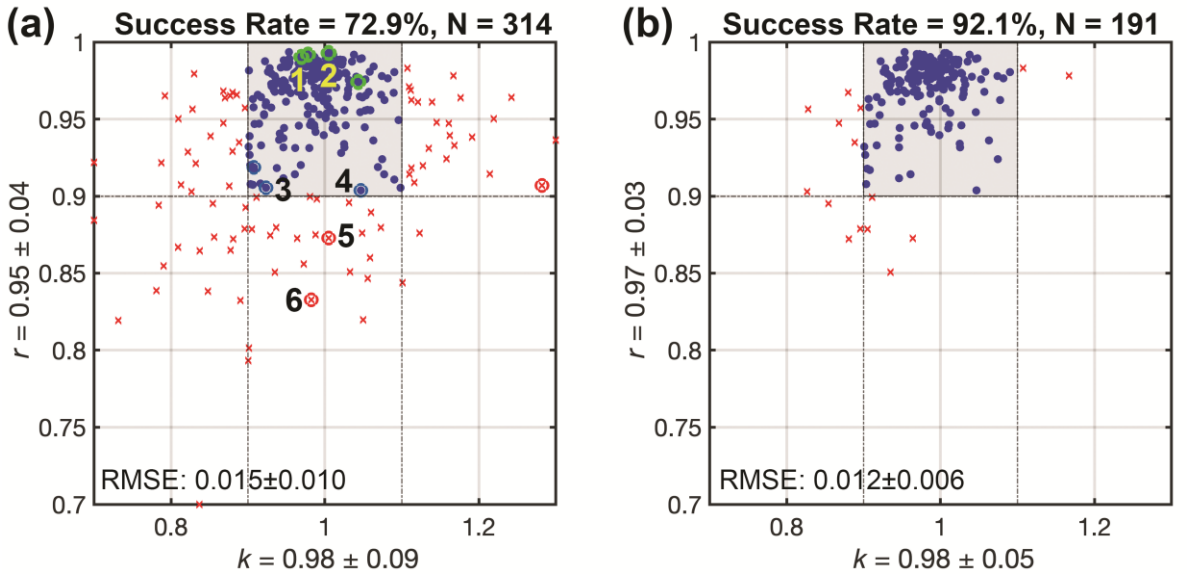

Figure 7. Summary of *k*, *r*, RMSE and SR when using the individualized CNN to predict peak strain distributions for (**a**) the entire HF dataset or (**b**) limited to impacts considered to have captured the complete kinematic events. The evaluations are based on generic model responses (i.e., scaling factors uniformly set to 1.0). Ten impact cases were selected (**a**; circles) to further assess accuracy for 11 scaled subject-specific models, from which six are labeled here for subsequent case illustration.

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

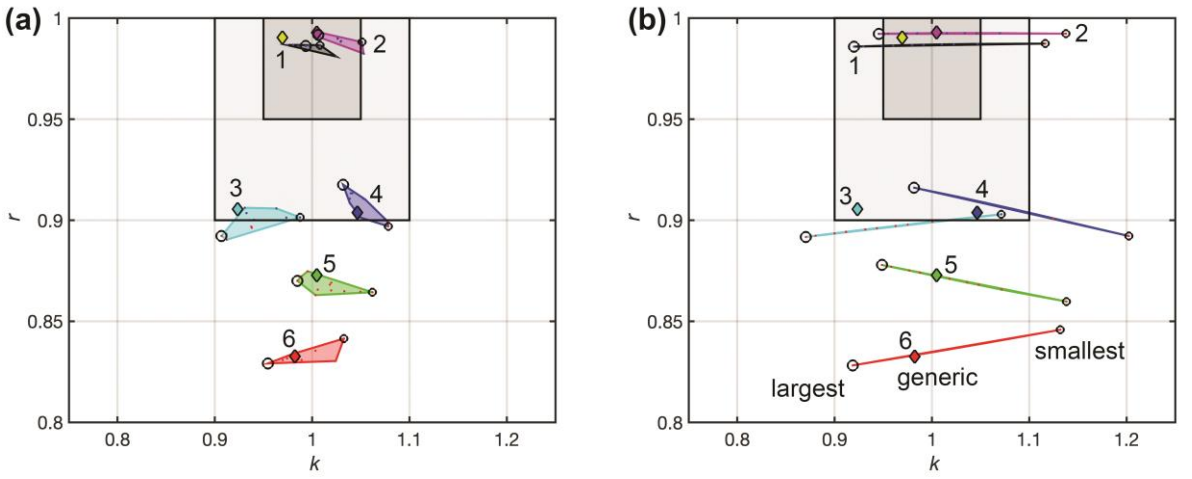

Figure 8. Comparison of *k*-*r* plots with the individualized CNN when (**a**) applying appropriate scaling factors for the 11 subject-specific models and the generic model or (**b**) using a uniform scaling factor of 1.0 as inputs, i.e., without accounting for morphological variations. With additional scaling factors as input, the estimated strains are considerably more accurate as the *k* values have a much smaller range and are clustered close to 1.0 ("perfect"). The shaded boxes indicate the criteria used to define estimation success: when both *k* and *r* do not deviate from the "perfect" value of 1.0 (when identical) by more than 0.1 or 0.05, respectively. The latter is a more stringent criterion for success. Markers indicating the largest, smallest, and the generic brain model are shown, along with kinematic impact numbers as identified in **Fig. 7**.

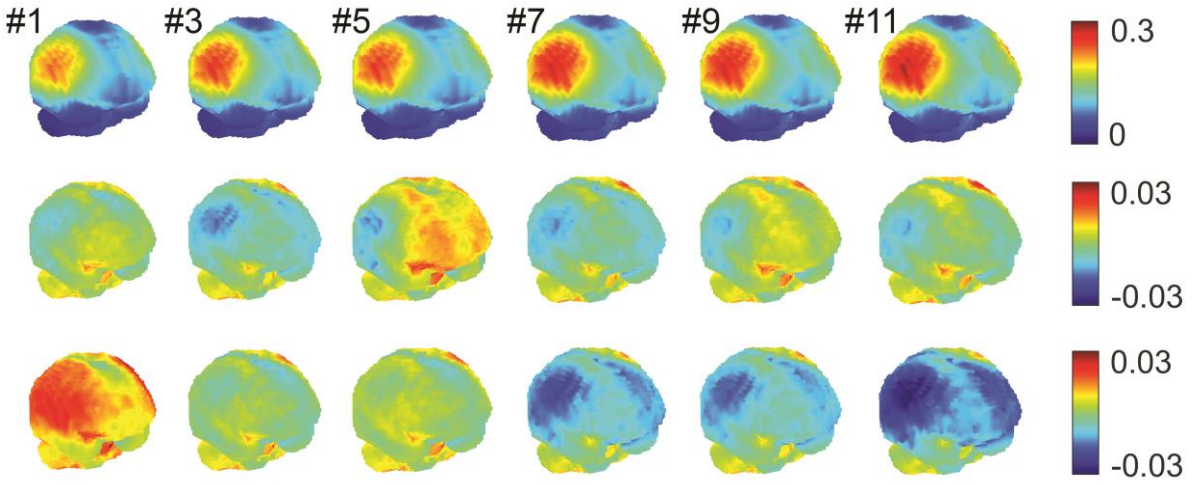

Figure 9. Directly simulated peak MPS of the whole brain using scaled subject-specific models (from left to right: odd-numbered subjects from smallest to largest brain in volume according to **Table 1**; top); difference of peak MPS using the individualized CNN with appropriate scaling factors as inputs relative to directly simulated counterparts (middle); the same difference when, instead, using scaling factors of 1.0 as inputs (bottom; i.e., without accounting for brain morphological differences), where over- or under-estimation is apparent from left to right. Comparisons for the even-numbered subjects are in Appendix (**Fig. A8**).

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)

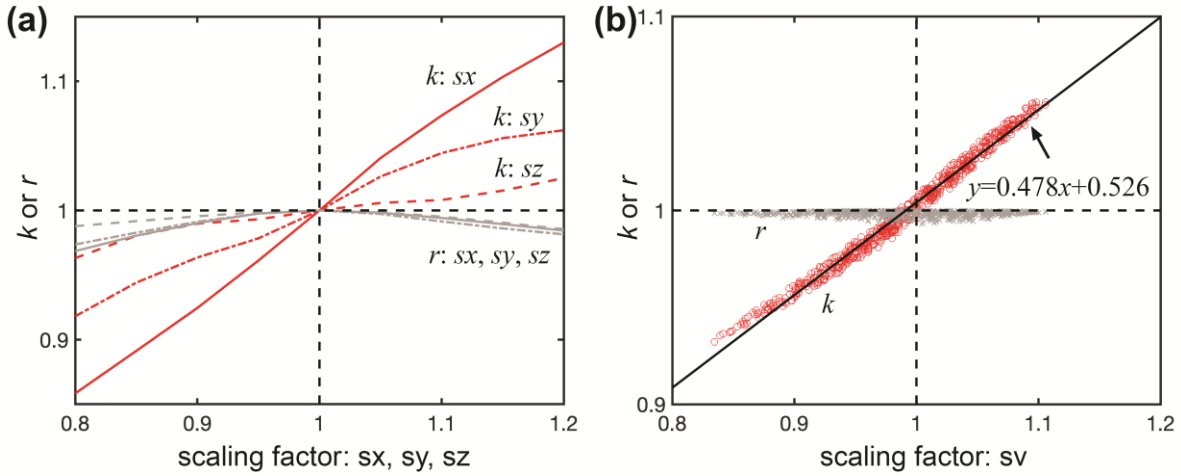

Figure 10. Significance of each anatomical scaling factor, *sx*, *sy*, and *sz* (left) and that for brain volume, *sv* (right) relative to the generic WHIM on peak MPS variations in terms of *k* and *r*, using a representative head impact as kinematic input (#1 in **Fig. 7a**). Example for an additional impact (#2 in **Fig. 7a**) is shown in Appendix **(Fig. A9)**.

Journal of Neurotrauma
A morphologically individualized deep learning brain injury model (DOI: 10.1089/neu.2022.0413)