Head Model for Entire Brain Deformation Calculation in Real-time with Machine Learning

Xianghao Zhan, Yuzhe Liu, Sam J. Raymond, Hossein Vahid Alizadeh, August G. Domel,
Olivier Gevaert, Michael Zeineh, Gerald Grant, David B. Camarillo

I. INTRODUCTION

Mild traumatic injury (mTBI), often caused by head impacts, can result in acute health consequences, such as unconsciousness, in post-concussive symptoms, including seizures, cognitive deficits and emotional disabilities [1-2], and even in long-term neurodegenerative diseases [3]. In sport-related mTBI, brain damage can accumulate with repetitive injuries if the mTBI goes undetected [4-5], necessitating quick and accurate evaluation of brain damage. The fast diagnosis and early warning approaches of mTBI are crucial in helping to prevent repetitive sport-related mTBI since the in-time intervention after early detection can attenuate the injury to a significant extent [6].

Previous work has shown that brain strain, particularly maximum principal strain (MPS), is a mechanical parameter that can accurately predict the risk of mTBI [7]. Currently, finite element (FE) modelling is used to calculate MPS [8-9], but such modelling requires expertise in biomechanics and at least several hours to run. Due to the amount of time involved and the lack of access to computational resources, the brain strain parameter is not widely used by medical professionals in diagnosing mTBI. For practical clinical applications, a fast and easy-to-use brain strain calculation method is needed.

To address the costly computation of FE simulations and calculate the entire brain MPS, in this study a head model was developed to calculate the peak MPS of every element of the brain with deep learning. We used kinematic data (angular acceleration and angular velocity) of 2,130 impacts generated by FE simulations of Hybrid III anthropomorphic test dummy (ATD) head impacts and 381 impacts collected from on-field football and MMA games with a previously validated instrumented mouthguard. The KTH head model was used, which is a validated FE head model [10], to obtain the true brain MPS for comparison with the deep learning model. Engineered features with the temporal information of the kinematics were extracted, and a deep neural network (DNN) was developed to predict the MPS at each brain element.

II. METHODS

Data description

The data used in this study included: 2,130 head impacts simulated by a validated FE model of the Hybrid III ATD [11] (dataset HM); 184 on-field college football head impacts [12] collected by the original version of the Stanford instrumented mouthguard [13] (dataset CF1); 118 on-field college football head impacts collected by the updated Stanford instrumented mouthguard [14] (dataset CF2); and 79 mixed martial arts (MMA) head impacts [5] collected by the updated Stanford instrumented mouthguard (dataset MMA). Based on the measured kinematics, the KTH head model [15], a validated FE model for mTBI/TBI modelling, was used to calculate the true MPS.

Feature Engineering

For each impact, we extracted features to aggregate the temporal information from the rotational kinematics for the deep learning model. First, angular accelerations $(\alpha_x, \alpha_y, \alpha_z)$ were calculated by 5-point stencil derivative on angular velocities $(\omega_x, \omega_y, \omega_z)$. Then, the magnitudes and the components in each direction (x, y, z) of angular accelerations and angular velocities were defined as 8 channels (denoted as $c_i(t), i = 1, 2, 3, \ldots, 8$). For each channel, six time-domain features were extracted based on the rationale of including the information of

X. Zhan is a PhD student in Bioengineering at Stanford University in USA (+1-650-788-9066, xzhan96@stanford.edu). Y. Li is a MSc student in Biomedical Informatics at Stanford University in USA. Y. Liu, S. J. Raymond, H. Vahid Alizadeh, A. G. Domel are post-doc researchers in Bioengineering at Stanford University in USA. O. Gevaert is a professor in Biomedical Informatics at Stanford University in USA. M. Zeineh is a professor in Radiology at Stanford University. G. Grant is a professor in Neurosurgery at Stanford University. D. B. Camarillo is a professor in Bioengineering at Stanford University.

the signal intensity and time history:

1) Maximum value (8 features): max$(c_i(t)), i = 1, 2, 3, \ldots, 8$;
2) Minimum value (8 features): min$(c_i(t)), i = 1, 2, 3, \ldots, 8$;
3) Integral of the time-signal (8 features): $\int c_i(t)\, dt, i = 1, 2, 3, \ldots, 8$;
4) Integral of the absolute values of time-signal (8 features): $\int |c_i(t)|\, dt, i = 1, 2, 3, \ldots, 8$;
5) Maximum and minimum of the exponential moving average of the signal derivative (48 features) with three smoothing factors [16];
6) Information of extrema except max. and min. (80 features): the total number of positive extrema, the total number of negative extrema, the values of the second-largest to the fifth-largest positive extrema, the values of the second-smallest to the fifth-smallest negative extrema. This type of feature was extracted because repetitive impacts might have an accumulating effect on the brain deformation. Also note that the first largest and the smallest extrema were not selected since they were previously included already.

Upon feature engineering there was standardisation on the features and logarithmic transformations, as well as data whitening on the labels (MPS) to deal with feature unit mismatch, to avoid negative MPS predictions and to improve prediction accuracy.

Model Development

The deep neural network, as a proxy to conventional FE modelling, consisted of five layers besides the input layer, with 160 units, and the output layer, with 4,124 units: (1) Hidden layer 1: 300 neurons with the rectified linear unit (ReLU) as the activation; (2) Dropout layer 1 with a dropout rate of 0.5 and no activation; (3) Hidden layer 2: 100 neurons with ReLU as the activation; (4) Dropout layer 2 with a dropout rate of 0.5 and no activation; (5) Hidden layer 3: 20 neurons with ReLU as the activation. Dropout layers and L2 penalty were used to avoid overfitting. Mean squared error was used as the loss function, and adaptive moment estimation (Adam) optimizer was used as the optimizer.

Model Evaluation

The model accuracy was evaluated in three tasks: (1) basis (HM dataset); (2) on-field (CF1/CF2/MMA dataset); (3) mixture (all datasets). The datasets were partitioned into the training, validation and test sets according to Fig. 1, for model training, hyperparameters tuning (training epochs, learning rate, strength of L2 penalty), and model performance testing. To test model reproducibility and robustness, in basis and mixture tasks, we partitioned the dataset with randomness in 20 repeats; in the on-field tasks, we followed a 5-fold cross-validation testing pattern shown in Fig. 1. Data augmentation, illustrated in Fig. 1, was carried out with the addition of noise abiding by normal distributions with a mean of zero and standard deviations of 0.01 and 0.02 times the standard deviation of the original data.

To quantify the accuracy, mean absolute error (MAE), root mean squared error (RMSE), coefficient of determination R2 and Spearman coefficient of correlation (Spearman Corr.) were used as the metrics. Furthermore, to represent the overall brain strain and avoid numerical error in simulation, R2 and RMSE of 95% MPS were calculated over the impacts in the test set. The metrics (mean values of MAE, RMSE, R2, Spearman coefficient of correlation, 95% MPS R2 and 95% MPS RMSE) were evaluated by comparing the predicted and the true MPS at every element over all test impacts. A group of metrics quantifying the model accuracy in each repeat (basis/mixture task) or fold (on-field task) was calculated and the summary statistics (mean, median, STD, 95% CI) of the metrics over repeats and folds were calculated.

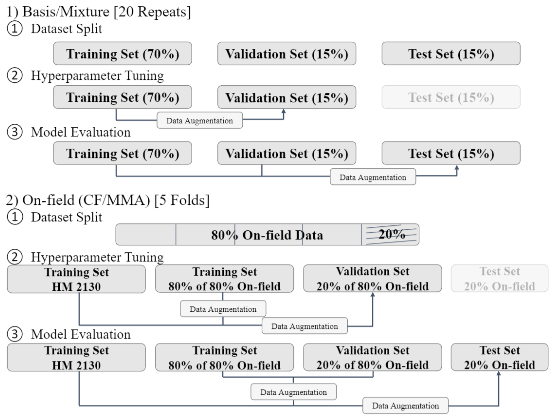
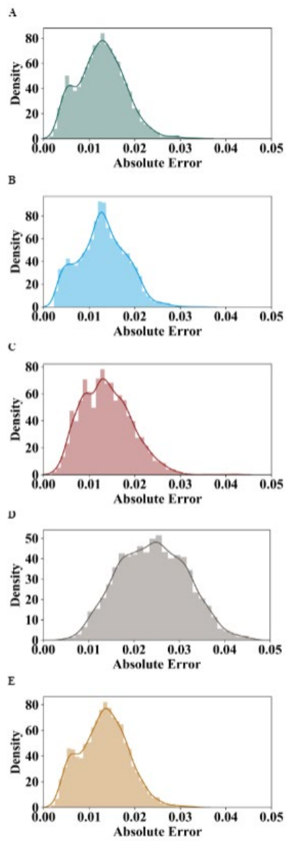


Fig. 1. The description of dataset partition process and the function of each dataset.

III. INITIAL FINDINGS

The hyperparameters and model accuracy metrics for each task are listed in Fig. 2. The absolute error of MPS at every element was averaged over impacts and repeats/folds and the density distribution of the absolute error is plotted in Fig. 2.



Basis (HM 2130)					
Dataset Details		Accuracy Metrics over 20 Experiments			
Total Cases	2130	MAE Mean	0.013	RMSE Mean	0.020
Training Cases	1490	MAE Median	0.013	RMSE Median	0.020
Validation Cases	320	MAE STD	0.001	RMSE STD	0.002
Testing Cases	320	MAE 95%CI	0.001	RMSE 95%CI	0.001
Optimal Hyperparameters		R2 Mean	0.896	Spearman Corr. Mean	0.933
Learning Rate	0.0002	R2 Median	0.900	Spearman Corr. Median	0.933
Epoch	2000	R2 STD	0.018	95%MPS R2 Mean	0.956
L2 Regularization	0.01	R2 95%CI	0.008	95% MPS RMSE Mean	0.021

On-field (CF1 184)					
Dataset Details		Accuracy Metrics over 5-fold Experiments			
Total Cases	2314	MAE Mean	0.013	RMSE Mean	0.022
Training Cases	2247	MAE Median	0.012	RMSE Median	0.021
Validation Cases	30	MAE STD	0.002	RMSE STD	0.003
Testing Cases	37	MAE 95%CI	0.001	RMSE 95%CI	0.003
Optimal Hyperparameters		R2 Mean	0.769	Spearman Corr. Mean	0.848
Learning Rate	0.0002	R2 Median	0.801	Spearman Corr. Median	0.852
Epoch	1000	R2 STD	0.054	95%MPS R2 Mean	0.880
L2 Regularization	0.01	R2 95%CI	0.047	95% MPS RMSE Mean	0.026

On-field (CF2 118)					
Dataset Details		Accuracy Metrics over 5-fold Experiments			
Total Cases	2248	MAE Mean	0.014	RMSE Mean	0.021
Training Cases	2205	MAE Median	0.014	RMSE Median	0.021
Validation Cases	19	MAE STD	0.001	RMSE STD	0.004
Testing Cases	24	MAE 95%CI	0.001	RMSE 95%CI	0.003
Optimal Hyperparameters		R2 Mean	0.664	Spearman Corr. Mean	0.838
Learning Rate	0.0001	R2 Median	0.711	Spearman Corr. Median	0.850
Epoch	1000	R2 STD	0.113	95%MPS R2 Mean	0.751
L2 Regularization	0.01	R2 95%CI	0.099	95% MPS RMSE Mean	0.028

On-field (MMA 79)					
Dataset Details		Accuracy Metrics over 5-fold Experiments			
Total Cases	2209	MAE Mean	0.024	RMSE Mean	0.040
Training Cases	2180	MAE Median	0.023	RMSE Median	0.037
Validation Cases	13	MAE STD	0.008	RMSE STD	0.015
Testing Cases	16	MAE 95%CI	0.005	RMSE 95%CI	0.009
Optimal Hyperparameters		R2 Mean	0.601	Spearman Corr. Mean	0.783
Learning Rate	0.0001	R2 Median	0.622	Spearman Corr. Median	0.795
Epoch	3000	R2 STD	0.105	95%MPS R2 Mean	0.813
L2 Regularization	0.01	R2 95%CI	0.065	95% MPS RMSE Mean	0.046

Mixture (Alldata 2511)					
Dataset Details		Accuracy Metrics over 20 Experiments			
Total Cases	2511	MAE Mean	0.013	RMSE Mean	0.022
Training Cases	1757	MAE Median	0.013	RMSE Median	0.022
Validation Cases	377	MAE STD	0.001	RMSE STD	0.001
Testing Cases	377	MAE 95%CI	0.001	RMSE 95%CI	0.001
Optimal Hyperparameters		R2 Mean	0.871	Spearman Corr. Mean	0.915
Learning Rate	0.0002	R2 Median	0.870	Spearman Corr. Median	0.914
Epoch	2000	R2 STD	0.011	95%MPS R2 Mean	0.940
L2 Regularization	0.01	R2 95%CI	0.005	95% MPS RMSE Mean	0.024

Fig. 2. The distribution of absolute error between predicted MPS (from our deep learning approach) and true MPS (from the KTH head model). The model information and the accuracy metrics on different prediction tasks are shown.

To visualize the prediction accuracy, three cases were selected as the impacts with 30th, 60th and 90th percentile highest true 95% MPS over all test impacts of all repeats. Figure 3 shows that the brain regions with high strain were very similar in the predicted and true results, which indicates the model's feasibility in diagnosing the region of injury with fast and accurate entire brain strain calculation. The MPS of each element plotted in Fig. 3 A-F were compared in Fig. 3 G-I. All scatters indicating MPS lie close to the reference line y=x, which shows that the predictions of MPS were accurate at every element of the brain.

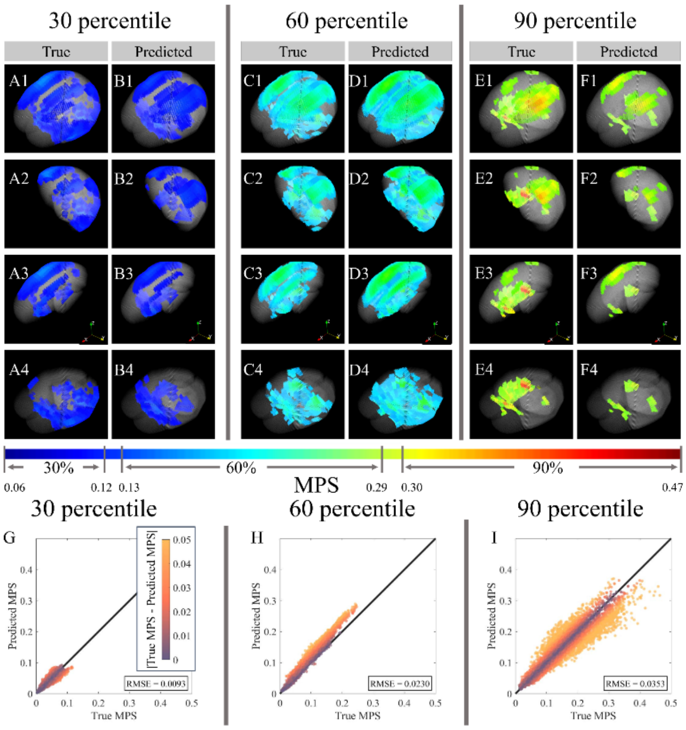


Fig. 3. The visualization of three typical cases of predicted and true 95% MPS in the mixture task.

IV. DISCUSSION

Compared to the time-consuming FE modelling, this study provides a new approach to calculating brain strain as a substitute for the conventional FE modelling after head impact. The deep learning head model is much faster (<0.001 s as compared with 7–8 hours for FE modelling for one head impact) and can give accurate brain strain. In addition to providing researchers with a fast and accurate method of calculating brain strain that includes the brain strain of the entire brain, the deep learning head model makes it easier for non-experts in biomechanics to interpret brain damage in an understandable manner. This method will also allow players suffering from head impacts to monitor in real time those dangerous impacts that could lead to a mTBI.

There are two limitations to this study. First, the feature extracted may not provide the best model accuracy. It is possible that without artificially engineering the features, other algorithms, such as convolutional neural network and recurrent neural network, may be better in terms of model accuracy. Secondly, the current model may not be accurate on those impacts outside of the domains under which the model was trained. More data from diverse sources can be used to improve the model accuracy.

V. REFERENCES

[1] Wallace T., Morris J., *Arc Phy Med & Rehab*, 2019.
[2] Shlosberg D., *et al*., *Nat Rev Neurology*, 2010.
[3] Doherty C. P., *et al*., *J Neuropathology & Experimental Neurology*, 2016.
[4] Schneider D. K., *et al*., *J Neurotrauma*, 2019.
[5] O'Keeffe E., *et al*., *J Neurotrauma*, 2020.
[6] Ponsford J., *et al*., *Pediatrics*, 2001.
[7] Wu S., *et al*., *Biomech & Modeling in Mechanobiol*, 2019.
[8] Kang H. S. , *et al*., *SAE Trans*, 1997.
[9] Zhao W., *et al*., *J Neurotrauma*, 2019.
[10] Wu S., *et al*., *Sci Rep*, 2020.
[11] Kleiven S., *et al*., SAE Tech. Paper, 2007.
[12] Giudice J. S., *et al*., *Ann Biomed Eng*, 2019.
[13] Hernandez F., *et al*., *Ann Biomed Eng*, 2015.
[14] Camarillo D. B., *et al*., *Ann Biomed Eng*, 2013.
[15] Liu Y., *et al*., *Ann Biomed Eng*, 2020.
[16] Zhan X., *et al*., *Measurement*, 2020.