

Improved Outcome prediction in mild Traumatic Brain Injury using Latent Feature Extraction from Volumetric MRI

Sanjay Purushotham¹, Ashwathy Samivel Sureshkumar¹, Li Jiang², Shiyu Tang², Steven Roys², Chandler Sours Rhodes^{2,3}, Rao P. Gullapalli², and Jiachen Zhuo²

¹Department of Information System, University of Maryland, Baltimore County, Baltimore, MD, United States, ²Diagnostic Radiology and Nuclear Medicine, University of Maryland School of Medicine, Baltimore, MD, United States, ³National Intrepid Center of Excellence, Walter Reed National Military Medical Center, Bethesda, MD, United States

Synopsis

Mild traumatic brain injury (mTBI) patients account for more than 70% of all TBI, with a subset experiencing persistent post concussive symptoms more than 3 months post injury. Patients clinical presentation and conventional imaging findings acutely post injury often lack the ability to predict chronic outcome. In this study, we present a novel method for latent feature extraction from volumetric MRI. We show that with machine learning methods, these acute latent volumetric MRI features are able to improve our symptom prediction in mTBI patients at 18 months post injury.

Introduction

Mild traumatic brain injury (mTBI) patients account for more than 70% of all TBI [1], with a subset experiencing persistent post concussive symptoms more than 3 months post injury. Patients clinical presentation and conventional imaging findings acutely post injury often lack the ability to predict chronic outcome. In this study, we present a novel method for latent feature extraction from volumetric MRI. We investigate how machine learning combined with these latent volumetric features may improve prediction of chronic symptoms.

Methods

Participants: 59 prospectively recruited mTBI patients were included in this study (Age: 40±17yrs, Gender: 40M/19F). All underwent MRI and behavioral assessment within 10 days and at approximately 18 months post injury.

Patient assessment: Self-reported symptoms, global functioning, and cognitive assessments were performed at each visit, which included: Modified Rivermead Post-Concussion Symptoms Questionnaire (RPQ), the Automated Neuropsychological Assessment Metrics (ANAM) battery, Military Acute Concussion Evaluation (MACE), and Mini-Mental Status Exam (MMSE). Poor outcome, chronic post-concussive syndrome (PCS) was defined as reporting 3 or more of the following symptoms lasting for greater than three months [2] (at 18 months): headaches, dizziness, sleep disturbances, trouble concentrating, fatigue, memory problems, and irritability.

Imaging and processing: MRIs performed on the 3T Siemens Tim Trio scanner included T1MPRAGE, T2w, FLAIR, DTI and SWI. MRI findings were defined as positive or negative from review of all sequences. Sub-cortical volume segmentation was performed with Freesurfer on 3D T1MPRAGE sequence (TE/T1/TR = 2.91ms/900ms/2300ms, Flip angle: 9°, FOV = 256 mm, resolution = 1 × 1 × 1 mm³). Regional volumes from all 123 regions were extracted. **Data Analysis:** We extracted three different feature sets (each set with different number of covariates) from the RPQ, ANAM, MMSE, MACE, demographics, MRI findings, and regional volumetric data at patients' acute visit. Continuous valued covariates were standardized, and one-hot encoding was done for categorical variables. Table 1 shows the summary of these feature sets. We used a probabilistic matrix factorization approach [3] to learn low-dimensional latent feature from the brain region volumetric data. Latent Feature Extraction using Probabilistic Matrix Factorization: As the number of features in the brain region volumetric data is large, we can project this data to a low-dimensional latent space using principal component analysis or a matrix factorization approach. The spatial (and functional) relations of the brain regions could be beneficial to learn patient-specific low-dimensional latent volumetric features. In other words, the key idea is to use a matrix factorization to derive a good low-dimensional volumetric feature representation for the brain regions by analyzing the brain adjacency matrix (which encodes the spatial and functional properties of the brain regions). Inspired by the success of matrix factorization in recommendation systems domain [4, 5], in this work, we propose to model the brain region volumetric matrix factorization using the graphical model shown in Figure 1, which fuses both the brain region graph and the patient-brain volumetric observations into a consistent and compact feature representation. To obtain the latent features including the patient-specific and brain-region specific latent features, we maximize the log-posterior distribution which is same as minimizing a sum-of-errors objective function with quadratic regularization terms. We then use a gradient descent algorithm to find the local minimum of the objective function and learn the model parameters.

PCS label prediction: We employed ten machine learning techniques for predicting the PCS labels based on the different feature sets as inputs. Area under receiver operating curve (AUROC) and accuracy were used as the evaluation metrics and stratified 5-fold cross validation was used to avoid overfitting.

Results and Discussion

Figure 2 (a) and (b) shows the empirical results for the ten machine learning techniques for the PCS prediction using different feature sets. From these figures, we note that most of all the machine learning techniques obtain similar or better accuracy and AUROC scores for feature set C i.e., using latent volumetric features. For Random Forests (with 500 trees) and SVM (with linear and RBF kernels) ML classifiers, we see using latent volumetric features (Set C) obtain at least 10% and 15% improvement compared to not incorporating the latent volumetric features (Set A). Using raw volumetric features did not provide good results as the total number of features (Set B) is larger than the number of patients, so the standard machine learning methods fail in the high-dimensional regimen. It is likely that the better performance would result after fine tuning the hyperparameters of the ML algorithms. Although deep neural networks could potentially provide better performance we did not apply this technique due to the limited number of patients.

Conclusion

In this work, we presented a novel approach to learn latent brain volumetric data by fusing patient-brain region observations with brain-region network using a probabilistic matrix factorization approach. We show in this study that we long-term PCS prediction in mTBI patients is improved using our feature extraction and latent representation of the brain volumetric MRI data from acute time post injury.

Acknowledgements

The study was conducted at University of Maryland School of Medicine Center for Innovative Biomedical Resources, Translational Research in Imaging @ Maryland (CTRIM) – Baltimore, Maryland. The work is supported by NIH under grant 5R01NS105503; DOD under award W81XWH-08-1-0725, and grant CRII (IIS--1948399) from the National Science Foundation. The views expressed in this manuscript are those of the authors and do not reflect the official policy of the Department of the Army/Navy/Air Force, Department of Defense, or U.S. Government.

References

- [1] Carroll LJ, Cassidy JD, Peloso PM et al. Prognosis for mild traumatic brain injury: Results of the WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury. *J Rehabil Med Suppl.* 2004;(43):84–105.
- [2] World Health Organization. *ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research.* Geneva: World Health Organization; 1992.
- [3] Ma, H., Yang, H., Lyu, M.R. and King, I., 2008, October. Sorec: social recommendation using probabilistic matrix factorization. In *Proceedings of the 17th ACM conference on Information and knowledge management* (pp. 931-940).
- [4] Mnih, A. and Salakhutdinov, R.R., 2008. Probabilistic matrix factorization. In *Advances in neural information processing systems* (pp. 1257-1264).
- [5] Purushotham, S., Liu, Y. and Kuo, C.C.J., 2012. Collaborative topic regression with social matrix factorization for recommendation systems. *ICML*

Figures

Feature Sets	Features/Covariates	Number of features
A	ANAM, MMSE, MACE, RPQ, MRI finding, Demographics	40
B	ANAM, MMSE, MACE, RPQ, MRI finding, Demographics, Raw Volumetric features	163
C	ANAM, MMSE, MACE, RPQ, MRI finding, Demographics, Latent volumetric features	45

Table 1: Summary of feature sets

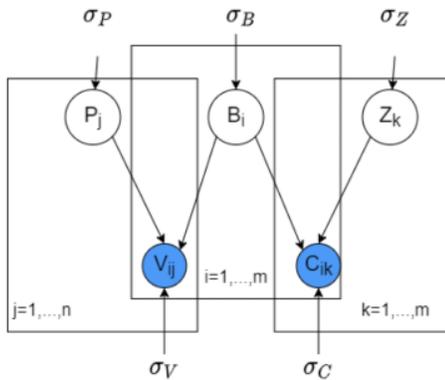


Figure 1: Graphical Model for the brain region volumetric matrix factorization. $C = \{C_{ik}\}$ is brain region adjacency matrix, B and Z are latent brain region and factor feature matrices with B_i and Z_k representing brain-region specific and factor-specific latent feature vectors. P_j represents patient latent vector for patient j . R_{ij} represents the volumetric observation (value) of brain region i for patient j .

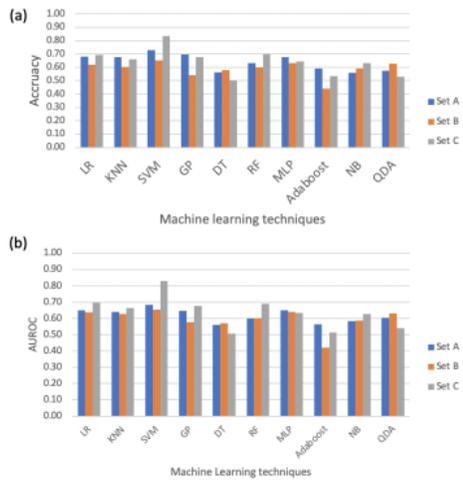


Figure 2: AUROC and Accuracy plots for predicting patient long-term outcome (PCS labels) using different feature sets.