Prediction of Posttraumatic Epilepsy using Machine Learning

Haleh Akrami[a], Andrei Irimia[a], Wenhui Cui[a], Anand A. Joshi[a], and Richard M. Leahy[a]

[a]University of Southern California, Los Angeles, USA

ABSTRACT

Post-traumatic Epilepsy is one of the common aftereffects of brain injury. This neurological disorder can persist throughout the lifetime of patients and impacts their quality of life significantly. Identification of markers that indicate the likelihood of developing PTE can help develop preventive care for subjects identified as at risk. Despite the relatively high prevalence of PTE, brain imaging-based biomarkers for the diagnosis of PTE are lacking. This is due in part to the heterogeneity of injury in traumatic brain injury patients. Here we investigate the use of structural and functional imaging features for training machine learning models. We compare the performance of four popular machine learning methods in predicting development of PTE after brain injury: (i) support vector machines, (ii) random forests, (iii) fully connected neural networks, and (iv) graph convolutional networks. Our results demonstrate the advantage of using a combination of connectivity features (functional) and lesion volume (structural) in conjunction with a Kernel SVM approach in predicting PTE. We also demonstrate that using a feature reduction method such as principal component analysis (PCA) can be more effective than penalizing classifiers. This might be due to the limitation of penalized models for a framework where features are correlated.

Keywords: fMRI, lesion detection , PTE, machine learning

1. PROBLEM STATEMENT

Traumatic brain injury (TBI) results in neurological disorders in a significant fraction of patients.[1] Post-traumatic epilepsy is one of the common consequences of TBI occurring in 15-20% of patients with brain trauma.[2] Prediction of PTE can help in the development of preventive care for subjects who would be identified at risk for PTE. Brain injury severity has been recognized as the primary risk factor for PTE development in several studies.[3,4] However, only a few studies have investigated imaging bio-markers for establishing risk factors. The heterogeneous nature of TBI injury types, pathology, and lesions present additional challenges to biomarker discovery. Because the locations, spatial extent, and content of lesions vary considerably within and between patients with and without PTE, there is no simple spatial overlap of injury profiles across the two groups. Some studies considered this heterogeneity in their analysis, and reported lesions in the temporal lobes are associated with an increased PTE incidence.[5,6] This conclusion is also supported by studies utilizing inverse localization of EEG potentials recorded from acute TBI patients.[7–9]

Resting fMRI (rfMRI) has recently attracted considerable attention in studies of brain connectivity and functional organization.[10] Resting fMRI in TBI subjects is affected through the impact of lesions, both locally and potentially throughout networks in which the lesioned regions are involved.[11] It has been noted that focal epilepsy affects common neuronal networks beyond the seizure onset zone.[12] More specifically, increased segregation of fMRI brain networks was reported in focal epilepsy.[13] A recent study[14] used machine learning approaches to reveal functional brain changes that may serve as PTE biomarkers. However, the high dimensionality of connectivity features can quickly lead to overfitting of these models. Regularized models with a ridge[15] or Lasso penalty[16] can select a subset of features and prevent overfitting by enforcing penalizing the weights with large norms.

In the case of connectivity features, the feature space can suffer from ill-conditioning, which results from the existence of groups of highly correlated features and a limited number of subjects available for study. Methods using simple penalties that discard most of the correlated features can become unstable.[17] Here we illustrate the advantage of using a feature reduction method such as principal component analysis (PCA)[18] for reducing the highly correlated high dimensional features to a smaller number.

Medical Imaging 2021: Biomedical Applications in Molecular, Structural, and Functional Imaging, edited by Barjor S. Gimi, Andrzej Krol, Proc. of SPIE Vol. 11600, 116001Q · © 2021 SPIE
CCC code: 1605-7422/21/$21 · doi: 10.1117/12.2580953

Several factors such as size, quality, and nature of features affect the choice of appropriate machine learning methods.[19] We trained four machine learning methods: (i) Random Forests (RF), (ii) Support Vector Machines (SVM), (iii) a fully connected Neural Network (NN), and (iv) a Graph Convolutional Network (GCN), to predict PTE. We use the connectivity matrix (Pearson correlation coefficients) between 16 regions of interest (ROIs) defined in the USCLobes atlas (http://brainsuite.org/usclobes-description)[20] as the functional features for the classifier. This atlas consists of lobar delineations (white matter, left and right frontal, parietal, temporal, and occipital lobes, cingulate gyrus, as well as the bilateral insulae, brainstem, corpus callosum, and cerebellum). We also use lesion volume in 13 ROIs (we removed white matter, brainstem, and corpus callosum) as structural features. Lesion volumes were computed using an automated lesion delineation method based on a variational autoencoder described in Akrami et al (2020).[21] We compared the performance of the classifiers using subsets of features as follows: functional only, structural only, and both. To prevent overfitting we investigated both regularization and reducing data dimensionality with PCA.

2. METHODOLOGY

We trained four machine learning methods to predict PTE as follows:

Random Forests:[22] RFs fit a collection of decision tree classifiers on subsamples of the dataset and uses averaging decisions of the weak learners to improve predictive accuracy. Decision trees are prone to overfitting, especially when a tree is particularly deep. The maximum depth of the tree can control the trade-off between underfitting and overfitting.

Support Vector Machines and Kernel-Support Vector Machines: SVMs[23] find the solution for the optimal hyperplane with the largest margins for separating data between two groups. In cases of data that are not linearly separable, the features can be mapped to a higher dimensional reproducing kernel Hilbert space (RKHS) for classification using a kernel version of SVM (KSVM). The dot products necessary in the computations can be directly computed without an explicit mapping to RKHS using the kernel trick. One of the most popular kernels for this purpose, and that used here, is the radial basis function (RBF).

Neural Networks: Our NN consists of multi-layer perceptron blocks[24] trained using back-propagation via stochastic gradient descent. Neural networks can achieve state of the art performance with enough training data. However, due to the limited size of available datasets, their use is still often challenging in medical imaging settings.

Graph Convolutional Network (GCN): Graph Convolutional Networks (GCNs)[25] are powerful tools when learning from graph-structured data.[24,26] GCNs are extensions of convolutional neural networks; the classical definition of a convolution operation cannot be easily generalised to the graph setting. Spectral graph theory is the key to this generalisation by defining filters in the graph spectral domain. Recently, GCNs have been used for learning useful representations of brain functional networks while considering topological information.

For implementing RF and SVM/KSVM we used sklearn libraries.[27] For implementing the NN and GCN we used Keras[28] and Pytorch[29] respectively.

2.1 Data

We used two datasets in this study:[6] (i) The Maryland TBI MagNeTs dataset,[30] of which we used 72 subjects for classification testing. The remaining 40 subjects were used for training the neural network for lesion delineation; (ii) The TRACK-TBI Pilot[31] dataset, which includes 97 subjects, was also used for training the lesion delineation neural network. The imaging data are available from FITBIR (https://fitbir.nih.gov), with FLAIR, T_1, T_2, diffusion, and other modalities available for download. For classification, we used 36 TBI subjects with epilepsy (25M/11F) from the Maryland TBI MagNeTs dataset and 36 randomly selected TBI subjects without epilepsy (distinct from the set used to train the lesion detection algorithm; 26M/10F) from the same dataset.[30] The age range for the epilepsy group was 19-65 years (yrs) and 18-70 yrs for the non-epilepsy group.

2.2 Analysis of connectivity

The functional features we used for SVM, KSVM, RF and NN were the upper triangular part of the connectivity matrix computed as the set of pair-wise Pearson correlation coefficients. We first used the connectivity matrix and SVM/KSVM[32] to classify PTE vs. non-PTE groups. We used leave one out cross-validation and the area under the receiver operating characteristic curve (AUC) metric to tune the hyperparameters of SVM/KSVM (regularization, gamma). As noted above each group (PTE/nonPTE) had 36 subjects. Randomized training/test allocation of subjects was repeated 100 times to obtain the standard deviation of the AUC. In addition to the SVM/KSVM, we also trained a Random Forest (RF) classifier.[33] For tuning the maximum depth of the RF, we used leave one out cross-validation, and the AUC metric. Then, randomized training/test allocation of subjects was repeated 100 times to obtain the standard deviation of the AUC. We also used a shallow, fully connected neural network. The neural network architecture consists of 3 hidden layers of size (32,16,16). Since the connectivity matrix has a high dimension, For the SVM, KSVM and RF methods we compared two different approaches: regularization using ridge regression (L2 penalty), and regularization with application of PCA for feature dimensionality reduction. We found that, generally, PCA was more effective for increasing the AUC, probably due to the high degree of correlation between features. The number the components used from PCA was treated as a hyper-parameter and tuned similarly to the regularization parameters via cross-validation.

Since the connectivity matrix (Fig. 1) can be viewed as a graph, we applied the GCN approach for predicting PTE using the connectivity matrix. We constructed a GCN model consisting of a graph convolution block and a Multi-layer Perceptron (MLP) block.[24] The two blocks have three graph convolution layers (L_1, L_2, L_3) and two fully-connected layers (L_4, L_5), respectively. The edge weights are the Pearson correlation coefficients and we discarded the edges with correlation coefficients lower than 0.2 to preserve only the stronger connections in the graph. Inspired by Li et al (2019),[24] we concatenated hand-crafted features as input node features at each node, namely the mean, standard deviation of fMRI time-series, node degree, and ROI center coordinates. We normalized each feature to zero mean, unit variance. As suggested in Li et al.(2019) [24] the output feature vector F of the graph convolution block was summarized using:

$$F = (\frac{1}{N}\sum_{i=1}^{N} v_i) \parallel \max(\{v_i : i = 1, \ldots, N\}) \tag{1}$$

where N is the number of nodes in the graph, v_i denotes the feature vector of node i, and $\parallel$ denotes the concatenation operation. Note that max operates element-wise. We then input F into the MLP block to produce the final classification results for each graph.

2.3 Analysis of Lesion volume

We repeated the above analysis for each of the learning methods, with the exception of the GCN, using lesion volume rather than connectivity data as the feature vector. We used multimodal MR images (T_1, T_2, FLAIR) and machine learning (ML) to automatically identify and delineate lesions. We used a variational auto-encoder[6,34,35] which is an unsupervised approach. By training the VAE using nominally normal imaging data, the network learns to encode only normal images. Lesions can then be identified from the difference between original and VAE-decoded images (Fig. 2). In practice, VAEs exhibit some degree of robustness to outliers (in our case, lesions) in the training data. To identify lesion-based PTE biomarkers, we trained the VAE using the T_1-weighted, T_2-weighted, and FLAIR MRIs in the Maryland TBI MagNeTs dataset.[30] The lesions were delineated based on VAE reconstruction error in the FLAIR images. A detailed description and validation of our method is available elsewhere.[6,21] We generated reconstruction error maps for the Maryland MagNeTs test set (36 TBI subjects with epilepsy and 36 TBI subjects without epilepsy) for lesion mapping.

To identify the lesions as binary masks a one-class SVM[36] was applied to the VAE lesion maps at each voxel and across subjects to identify subjects with abnormally large errors at each voxel.[33,36] We defined the outliers marked by the one-class SVM as lesions and computed lesion volumes per ROI by counting the number of outlier voxels in each ROI for each subject. The lesion volume in each ROI (lobe) was then used as the feature vector to train and test the SVM, KSVM, RF and NN classifiers.

Finally, we re-ran each of these classifiers using a combination of the functional and structural features.

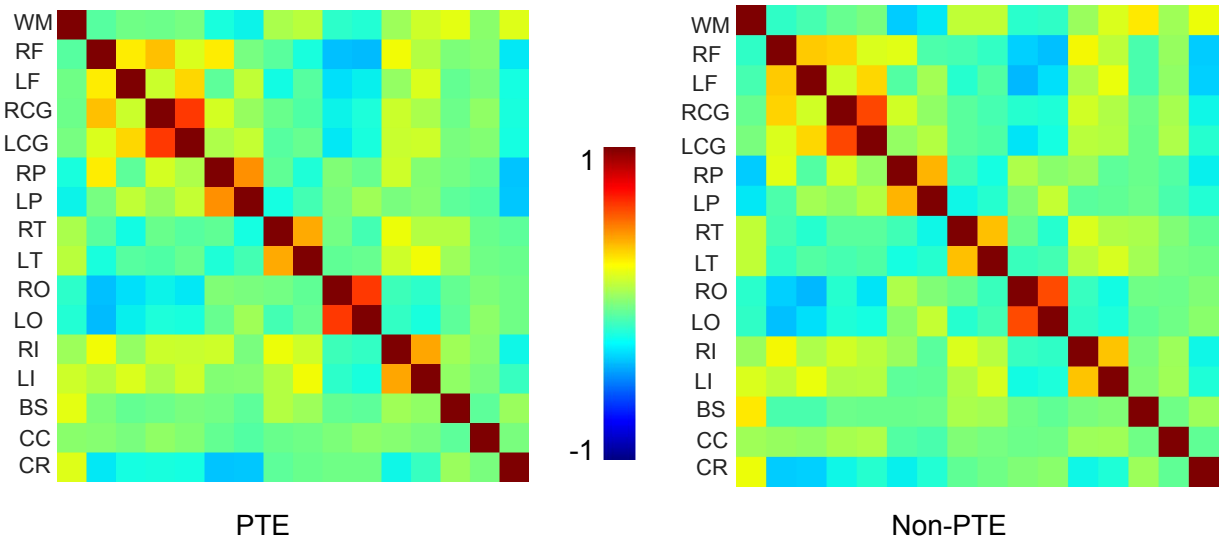

Figure 1. Connectivity matrices in PTE versus not-PTE subjects, WM: white matter, RF: R. frontal lobe, LF: L. frontal lobe, RCG: R. cingulate gyrus, LCG: L. cingulate gyrus, RL: R. parietal lobe, LP: L. parietal lobe, RT: R. temporal lobe, LT: L. temporal lobe, RO :R. occipital lobe, LO: L. occipital lobe, RI: R. insula, LI: L. insula, BS: brainstem, CC: Corpus callosum, CR: cerebellum.

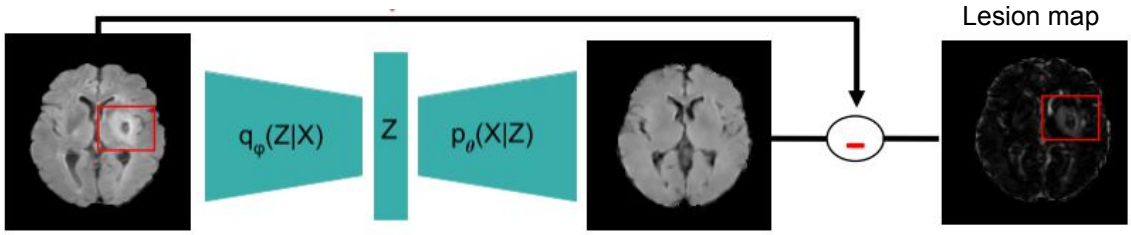

Figure 2. A VAE was used to identify lesions in each lobe to generate a feature vector (lesion volume) for PTE prediction. The VAE in inference mode will reconstruct a lesion-free version of the image when trained on lesion-free data. Subtraction of the input and reconstructed image was used to identify lesion locations and volumes.

3. RESULTS

Classification results obtained with the machine learning models (RF SVM/KSVM, NN, and) are reported in Table 1 for lesion volume, connectivity features, and both. The AUC was 0.59 ($\sigma = 0.05$) using SVM with lobe lesion volume as the feature vector. This values increased to 0.64 ($\sigma = 0.05$) when we used PCA to reduce feature dimensionality. KSVM, RF, and NN all showed lower AUC compared to linear SVM. These, results indicate that lesion volume alone is not a useful biomarker at the individual level. At the group level, a clear difference with respect to lesion volume features was reported for PTE vs non-PTE groups.[6] The AUCs we report are in a similar range to a previous study that used morphological features extracted from FreeSurfer to predict PTE.[14]

Results based on connectivity features were on average slightly higher than for lesion features. The AUC of the ROC curve was 0.54 ($\sigma = 0.05$) using SVM. This increased to 0.67 (σ=0.05) using a combination of PCA dimensionality reduction and the kernel SVM (KSVM). All other methods performed, on average, worse than KSVM with PCA. Using a combination of both features increased the AUC to 0.74 ($\sigma = 0.05$) for KSVM with PCA, which is significantly larger than either structural or functional features separately. Due to the relatively small size of the training set and high dimensionality of the features, the NN did not improve the results.

The mean and standard deviation after 100 iterations of 36-fold cross-validation for the GCN was 0.59 (0.02) which is below that of KSVM with PCA. We believe that GCN has substantial potential here since it exploits

Table 1. AUC of classification methods with standard deviation in the parenthesis .

Method	Lesion volume (lobes)	connectivity(lobes)	both features
SVM	0.59 (0.05)	0.54 (0.05)	0.49 (0.05)
SVM (with PCA)	0.64 (0.05)	0.58 (0.05)	0.50 (0.05)
KSVM	0.56 (0.05)	0.61 (0.05)	0.61 (0.05)
KSVM (with PCA)	0.55 (0.06)	0.67 (0.06)	0.74 (0.05)
RF	0.48 (0.05)	0.60 (0.06)	0.59 (0.05)
RF (with PCA)	0.58 (0.06)	0.52 (0.05)	0.57 (0.07)
NN	0.53 (0.06)	0.54 (0.05)	0.57 (0.05)
GCN	N/A	0.59 (0.02)	N/A

the inherent graph structure of these data. The method tested here should be viewed as a baseline and may improve from a different choice of features and modifications of the network structure and training approach. .

4. CONCLUSION

Our results show that combining lesion volume data and fMRI based connectivity analysis has potential for predicting PTE onset in TBI patients. Further, the KSVM approach appears better suited to this problem than the random forest, SVM, NN and GNN approaches. We also showed using PCA for feature reduction is more effective to prevent over-fitting than using a regularizer. Preliminary results using a graph neural network are promising but need refinement to meet, or potentially exceed, the performance of the best of the other methods evaluated.

ACKNOWLEDGMENTS

This work is supported by the following grants: R01 NS074980, W81XWH-18-1-0614, R01 NS089212, and R01 EB026299.

REFERENCES

[1] Englander, J., Bushnik, T., Duong, T. T., Cifu, D. X., Zafonte, R., Wright, J., Hughes, R., and Bergman, W., "Analyzing risk factors for late posttraumatic seizures: a prospective, multicenter investigation," *Archives of physical medicine and rehabilitation* **84**(3), 365–373 (2003).

[2] Diaz-Arrastia, R., Agostini, M. A., Madden, C. J., and Van Ness, P. C., "Posttraumatic epilepsy: the endophenotypes of a human model of epileptogenesis," *Epilepsia* **50**, 14–20 (2009).

[3] Christensen, J., Pedersen, M. G., Pedersen, C. B., Sidenius, P., Olsen, J., and Vestergaard, M., "Long-term risk of epilepsy after traumatic brain injury in children and young adults: a population-based cohort study," *The Lancet* **373**(9669), 1105–1110 (2009).

[4] Xu, T., Yu, X., Ou, S., Liu, X., Yuan, J., Huang, H., Yang, J., He, L., and Chen, Y., "Risk factors for posttraumatic epilepsy: A systematic review and meta-analysis," *Epilepsy & Behavior* **67**, 1–6 (2017).

[5] Tubi, M. A., Lutkenhoff, E., Blanco, M. B., McArthur, D., Villablanca, P., Ellingson, B., Diaz-Arrastia, R., Van Ness, P., Real, C., Shrestha, V., et al., "Early seizures and temporal lobe trauma predict post-traumatic epilepsy: A longitudinal study," *Neurobiology of disease* **123**, 115–121 (2019).

[6] Akrami, H., Leahy, R. M., Irimia, A., Kim, P. E., Heck, C., and Joshi, A., "Neuroanatomic markers of post-traumatic epilepsy based on magnetic resonance imaging and machine learning," *medRxiv* (2020).

[7] Irimia, A. and Van Horn, J., "Epileptogenic focus localization in treatment-resistant post-traumatic epilepsy," *Journal of Clinical Neuroscience* **22**(4), 627–631 (2015).

[8] Irimia, A., Goh, S., Torgerson, C., Stein, N., Chambers, M., Vespa, P., and Van Horn, J., "Electroencephalographic inverse localization of brain activity in acute traumatic brain injury as a guide to surgery, monitoring and treatment," *Clinical Neurology and Neurosurgery* **115**(10), 2159–2165 (2013).

[9] Irimia, A., Goh, S., Torgerson, C., Chambers, M., Kikinis, R., and Van Horn, J., "Forward and inverse electroencephalographic modeling in health and in acute traumatic brain injury," *Clinical Neurophysiology* **124**(11), 2129–2145 (2013).

[10] Arslan, S., Ktena, S. I., Makropoulos, A., Robinson, E. C., Rueckert, D., and Parisot, S., "Human brain mapping: A systematic comparison of parcellation methods for the human cerebral cortex," *Neuroimage* **170**, 5–30 (2018).

[11] Palacios, E. M., Sala-Llonch, R., Junque, C., Roig, T., Tormos, J. M., Bargallo, N., and Vendrell, P., "Resting-state functional magnetic resonance imaging activity and connectivity and cognitive outcome in traumatic brain injury," *JAMA neurology* **70**(7), 845–851 (2013).

[12] Yamazoe, T., von Ellenrieder, N., Khoo, H. M., Huang, Y.-H., Zazubovits, N., Dubeau, F., and Gotman, J., "Widespread interictal epileptic discharge more likely than focal discharges to unveil the seizure onset zone in eeg-fmri," *Clinical Neurophysiology* **130**(4), 429–438 (2019).

[13] Pedersen, M., Omidvarnia, A. H., Walz, J. M., and Jackson, G. D., "Increased segregation of brain networks in focal epilepsy: an fmri graph theory finding," *NeuroImage: Clinical* **8**, 536–542 (2015).

[14] La Rocca, M., Garner, R., Jann, K., Kim, H., Vespa, P., Toga, A. W., and Duncan, D., "Machine learning of multimodal mri to predict the development of epileptic seizures after traumatic brain injury," in [*International Conference on Medical Imaging with Deep Learning–Extended Abstract Track*], (2019).

[15] Friedman, J., Hastie, T., Tibshirani, R., et al., [*The elements of statistical learning*], vol. 1, Springer series in statistics New York (2001).

[16] Tibshirani, R., "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)* **58**(1), 267–288 (1996).

[17] Toloşi, L. and Lengauer, T., "Classification with correlated features: unreliability of feature ranking and solutions," *Bioinformatics* **27**(14), 1986–1994 (2011).

[18] Wold, S., Esbensen, K., and Geladi, P., "Principal component analysis," *Chemometrics and intelligent laboratory systems* **2**(1-3), 37–52 (1987).

[19] Saccà, V., Sarica, A., Novellino, F., Barone, S., Tallarico, T., Filippelli, E., Granata, A., Chiriaco, C., Bossio, R. B., Valentino, P., et al., "Evaluation of machine learning algorithms performance for the prediction of early multiple sclerosis from resting-state fmri connectivity data," *Brain imaging and behavior* **13**(4), 1103–1114 (2019).

[20] Joshi, A. A., Choi, S., Sonkar, G., Chong, M., Gonzalez-Martinez, J., Nair, D., Shattuck, D. W., Damasio, H., and Leahy, R. M., "A whole brain atlas with sub-parcellation of cortical gyri using resting fMRI," in [*Medical Imaging 2017: Image Processing*], Styner, M. A. and Angelini, E. D., eds., **10133**, 186 – 194, International Society for Optics and Photonics, SPIE (2017).

[21] Akrami, H., Joshi, A. A., Li, J., Aydore, S., and Leahy, R. M., "Brain lesion detection using a robust variational autoencoder and transfer learning," in [*2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*], 786–790, IEEE (2020).

[22] Breiman, L., "Random forests," *Machine learning* **45**(1), 5–32 (2001).

[23] Cortes, C. and Vapnik, V., "Support-vector networks," *Machine learning* **20**(3), 273–297 (1995).

[24] Li, X., Dvornek, N. C., Zhou, Y., Zhuang, J., Ventola, P., and Duncan, J. S., "Graph neural network for interpreting task-fmri biomarkers," (2019).

[25] Kipf, T. N. and Welling, M., "Semi-supervised classification with graph convolutional networks," (2017).

[26] Gadgil, S., Zhao, Q., Pfefferbaum, A., Sullivan, E. V., Adeli, E., and Pohl, K. M., "Spatio-temporal graph convolution for functional mri analysis," (2020).

[27] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research* **12**, 2825–2830 (2011).

[28] Chollet, F. et al., "Keras." https://github.com/fchollet/keras (2015).

[29] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al., "Pytorch: An imperative style, high-performance deep learning library," *arXiv preprint arXiv:1912.01703* (2019).

[30] Gullapalli, R. P., "Investigation of prognostic ability of novel imaging markers for traumatic brain injury (tbi)," tech. rep., BALTIMORE UNIV MD (2011).

[31] Yue, J. K., Vassar, M. J., Lingsma, H. F., Cooper, S. R., Okonkwo, D. O., Valadka, A. B., Gordon, W. A., Maas, A. I., Mukherjee, P., Yuh, E. L., et al., "Transforming research and clinical knowledge in traumatic brain injury pilot: multicenter implementation of the common data elements for traumatic brain injury," *Journal of neurotrauma* **30**(22), 1831–1844 (2013).

[32] Suykens, J. A. and Vandewalle, J., "Least squares support vector machine classifiers," *Neural processing letters* **9**(3), 293–300 (1999).

[33] Duda, R. O., Hart, P. E., and Stork, D. G., [*Pattern classification*], John Wiley & Sons (2012).

[34] Chen, X. and Konukoglu, E., "Unsupervised detection of lesions in brain mri using constrained adversarial auto-encoders," *arXiv preprint arXiv:1806.04972* (2018).

[35] Kingma, D. P. and Welling, M., "Auto-encoding variational Bayes," *arXiv preprint arXiv:1312.6114* (2013).

[36] Zhang, J., Ma, K.-K., Er, M.-H., and Chong, V., "Tumor segmentation from magnetic resonance imaging by learning via one-class support vector machine," (2004).