

Traumatic brain injury lesion detection using a variational autoencoder

Haleh Akrami, Anand A. Joshi, Jian Li, Richard M. Leahy

Signal and Image Processing Institute, University of Southern California, Los Angeles, CA, USA

Introduction

The FITBIR (Federal Interagency Traumatic Brain Injury Research) informatics systems makes available a unique large collection of traumatic brain injury (TBI) patients with a sub-population of post-traumatic epilepsy (PTE) cases. This offers a unique opportunity to study the relationship between TBI and epilepsy, and potentially to discover biomarkers for the location and types of lesions that have a higher likelihood of leading to PTE. To tackle this problem, we ultimately need to investigate not only lesions but also the brain networks that they affect, identified either from diffusion or functional MRI. As a step towards automated analysis of the large databases such as those in the FITBIR system, here we describe a method for automatic delineation of brain lesions from multispectral (T1, T2, FLAIR) MRI.

Methods

We used MRI datasets from 110 subjects from the Maryland MagNeTs study of neurotrauma (<https://fitbir.nih.gov>) for training where no ground truth was given. We split this dataset into 100 subjects for training and 10 subjects for validation. We used another 15 subjects from the ISLES (The Ischemic Stroke Lesion Segmentation) database (Maier et al. 2017) for testing and performance evaluation. While these later images are not from TBI patients, they present a similar challenge and have the advantage that all lesions have been hand segmented for validation purposes. Both datasets consist of images with sparse lesions in three modalities (T1, T2, FLAIR). The three modality images were co-registered to the MNI atlas and re-sampled to 1mm isotropic resolution. Skull and other non-brain tissue were removed using BrainSuite¹. All the images were then normalized to a range between 0 and 1. To detect lesions in MR images without knowing the ground truth, we trained a variational autoencoder (VAE)² model to efficiently represent normal brain structures in a low dimensional space. For our purposes, the optimal VAE when trained will be able to accurately encode (and decode) normal brain structures but not lesions. Specifically, here we used a deep VAE due to the robustness of its latent space representations. The deep VAE consists of four consecutive blocks of convolutional layers, max-pooling layers and rectified linear unit activation function (ReLU) for the encoder and four deconvolutional layers, up-sampling layers and ReLU for the decoder. Finally, at the output, a median filter of size 7x7 is applied to remove small errors. An error map between the reconstructed and the original image was generated for each image to segment lesions. These lesion error maps were clipped between (0,1) to identify hyperintensities in FLAIR and then binarized through thresholding. Varying the threshold used to identify lesions was used to generate a Receiver Operating Characteristic (ROC) curve to quantitatively evaluate performance. We also measured the area under the curve (AUC) in the ROC plot. For comparison purposes, we also applied a (non-variational) regular autoencoder with similar architecture to the same datasets.

Results

Fig. 1 (A) shows example slices from the validation set of the original FLAIR image in the 1st row. The reconstructed images and the error maps using the deep VAE are shown in the 2nd and 3rd rows and their counterparts for the regular autoencoder are shown in the 4th and 5th rows. In contrast to the regular autoencoder where lesions were also encoded into the latent space, our deep VAE reconstructed brain images without encoding lesions as illustrated in the error maps. Fig. 1 (B) shows results for the test dataset

where the additional row (last row) shows the ground truth (hand segmentation). Similar to (A), the lesions were not encoded by our deep VAE model, thus were successfully identified in the error maps. The detection results matched well with the manual segmentation shown in the ground truth images. Fig. 2 shows the ROC curves when applying different thresholds to the error maps and a 0.89 AUC was achieved using our deep VAE model as opposed to 0.70 using the regular autoencoder. The average Dice coefficient across subjects for the deep VAE was 0.55 and for the regular autoencoder was 0.06 at a threshold of 0.1.

Conclusion

Delineation of brain lesions from MR images can be difficult and time consuming, particularly for research studies of large datasets. Our deep VAE network was able to represent normal brain structures without encoding lesions and generate error maps that facilitate lesion detection.

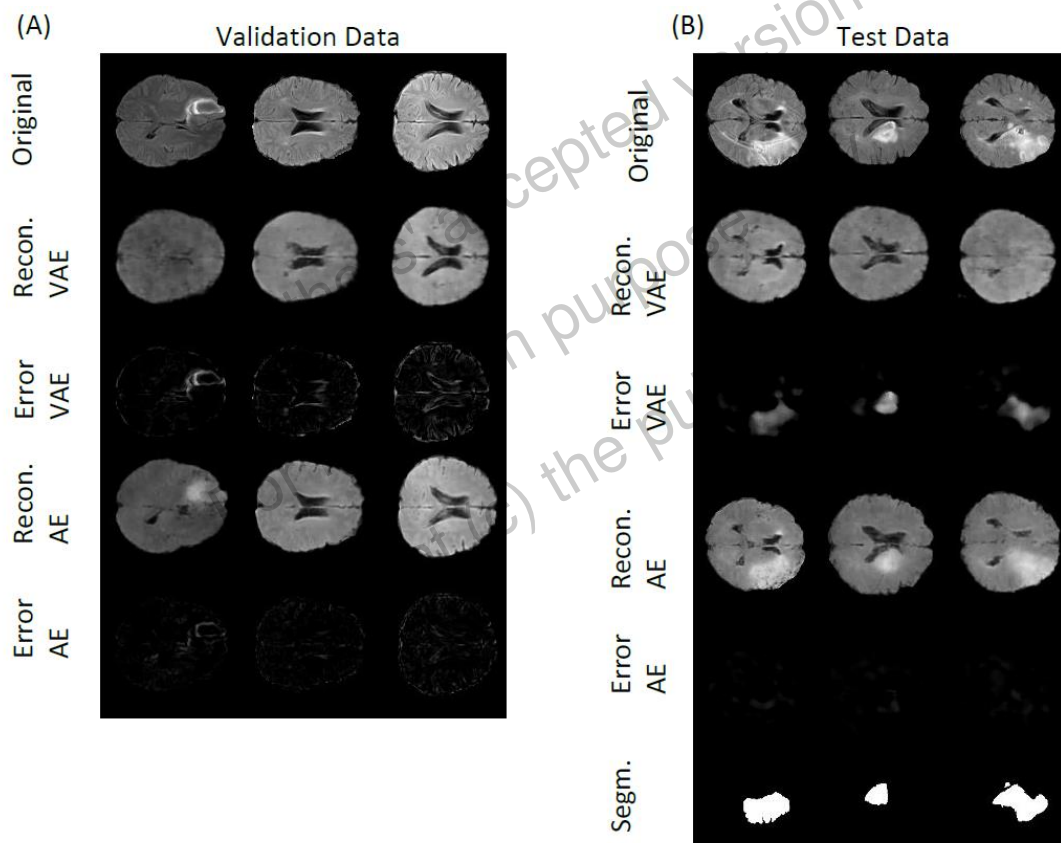


Figure 1: A) Original FLAIR image, reconstruction result and Error map of VAE and AE For validation data-set. B) Original FLAIR image, reconstruction result and Error map of VAE and AE (after a median filter with size of 7x7 is applied) for the test data-set

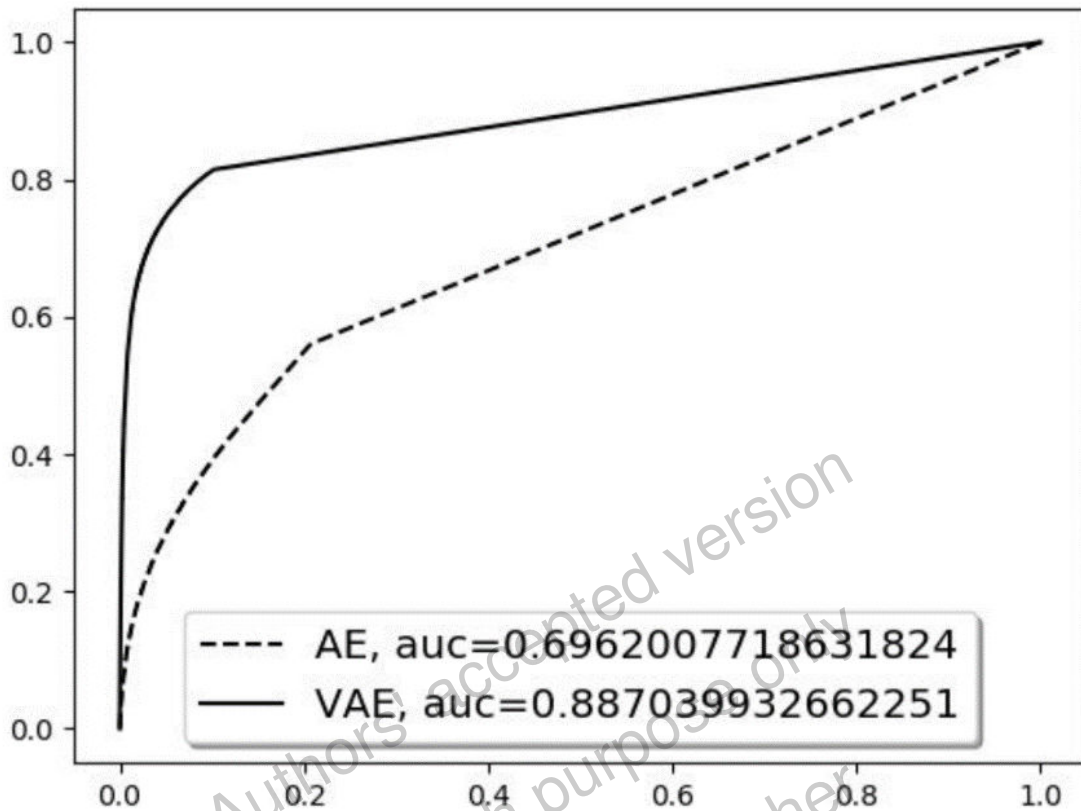


Figure 2: ROC curve of VAE and AE. Error map is clipped between (0,1) and a median filter with size of 7x7 is applied, then ROC curve is calculated for all pixels.

References

1. D. W. Shattuck, R. M. Leahy, "BrainSuite: An automated cortical surface identification tool", *Medical Image Analysis*, vol. 6, no. 2, pp. 129–142, 2002. DOI: 10.1016/S1361-8415(02)00054-3.
2. D. P. Kingma, M. Welling, "Auto-encoding variational bayes", *arXiv:1312.6114 [cs, stat]*, 2014.