Unsupervised Lesion Detection in Brain CT using Bayesian Convolutional Autoencoders

Nick Pawlowski[1], Matthew C.H. Lee[1,2], Martin Rajchl[1], Steven McDonagh[1], Enzo Ferrante[3], Konstantinos Kamnitsas[1], Sam Cooke[4], Susan Stevenson[4], Aneesh Khetani[4], Tom Newman[4], Fred Zeiler[4], Richard Digby[4], Jonathan P. Coles[4], Daniel Rueckert[1], David K. Menon[4], Virginia F.J. Newcombe[4*], and Ben Glocker[1*]

[1]Biomedical Image Analysis Group, Imperial College London, UK
[2]HeartFlow, California, CA 94063, USA
[3]Universidad Nacional del Litoral / CONICET, Argentina
[4]Division of Anaesthesia, Department of Medicine, University of Cambridge, UK
[*]joint senior authors
n.pawlowski16@imperial.ac.uk

Abstract

Normally, lesions are detected using supervised learning techniques that require labelled training data. We explore the use of Bayesian autoencoders to learn the variability of healthy tissue and detect lesions as unlikely events under the normative model. As a proof-of-concept, we test our method on registered 2D mid-axial slices from CT imaging data. Our results indicate that our method achieves best performance in detecting lesions caused by bleeding compared to baselines.

1 Introduction

Deep learning is arguably now one of the most widely used machine learning methods for medical imaging [7]. A common task is the segmentation of lesions and other pathologies. However, most methods are based on *supervised* learning, which means they require large amounts of carefully annotated training data. Our work here focuses on *unsupervised* lesion detection that resembles pixel-wise outlier detection related to [13, 2, 1, 11]. Most of those methods build on generative models that capture the normal distribution and detect outliers by checking their likelihood.

We introduce the use of Bayesian autoencoders to model the data distribution and interpret the reconstruction error as a measure of abnormality. Applying this method to CT mid-axial slices we show that our approach achieves superior performance to various baselines.

2 Detecting Outlier Regions using Autoencoders

We are interested in building an autoencoder AE for healthy data points $x \in \mathcal{D}_{healthy}$ so that the probability of the data point given the autoencoder reconstruction $\mathcal{N}(x|AE(x), 1)$ is maximised. Rather than only learning point estimates of the weights w of the autoencoder we use dropout [12] and an uninformative prior to model the weight uncertainty of the autoencoder with MC-Dropout [3]. This gives a Bayesian autoencoder so we estimate $AE(x)$ as the Monte-Carlo (MC) estimate

$$AE(x) = \int AE(x|w)p(w|\mathcal{D})\mathrm{d}w \approx \frac{1}{N}\sum_{i=1}^{N} AE(x|w_i) \mid w_i \sim p(w|\mathcal{D}) \tag{1}$$

In general, autoencoders learn to compress the data and thus find a lower dimensional manifold that the training data lies on. An optimal autoencoder would therefore have a zero reconstruction error $\delta_{rec}(x) = |x - AE(x)| = 0$ for any $x \in \mathcal{D}_{training}$. For data samples different from the training

1st Conference on Medical Imaging with Deep Learning (MIDL 2018), Amsterdam, The Netherlands.

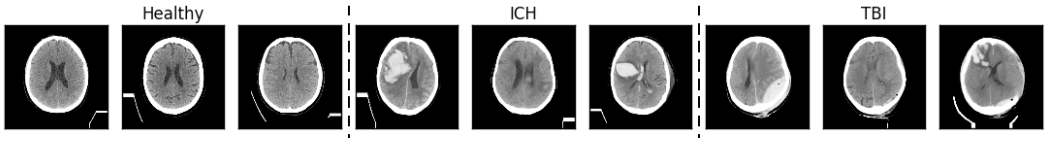


Figure 1: Examples of mid-axial slices of the data used. The healthy data is used for learning a normative model to detect lesions in ICH and TBI.

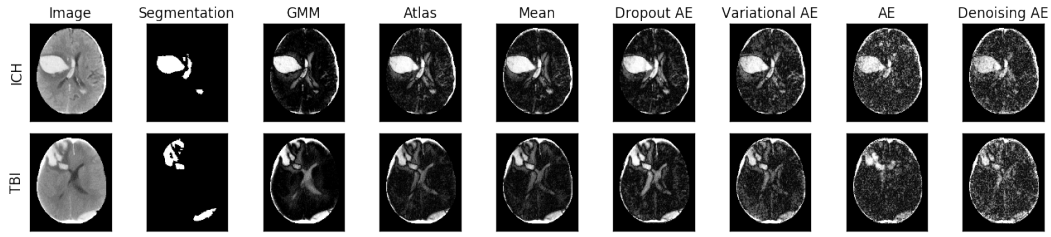


Figure 2: Comparison of the difference maps generated by the different methods. Brighter spots correspond to a higher difference.

manifold $x' \notin \mathcal{D}_{training}$ we argue that the autoencoder generates a reconstruction that projects the data towards the manifold. Therefore, we interpret $\delta_{rec}(x)$ as a distance to the found manifold that is related to the probability of the new sample being part of the same manifold $p(x' \in \mathcal{D}|\mathcal{D}) \propto \delta_{rec}(x')$. Here, $p(x' \in \mathcal{D}|\mathcal{D})$ describes the inverse of the probability of x' being an outlier. We use this estimate $p(x' \notin \mathcal{D}|\mathcal{D}) \propto |x - AE(x)|$ to find localised outliers by thresholding.

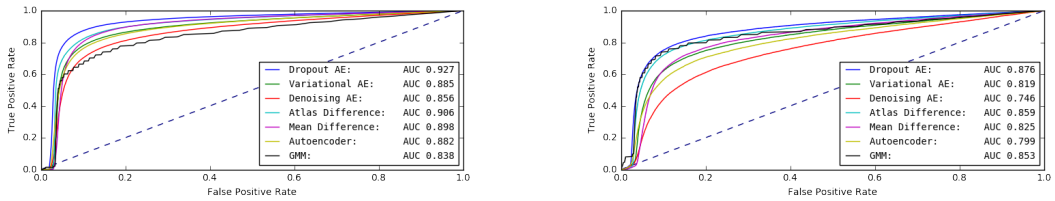
3 Experiments & Results

We test our Bayesian autoencoder on mid-axial slices of registered CT images. The 3D images are registered with an affine transformation to a CT atlas in canonical MNI space. We use 102 healthy cases to train the autoencoder and 107 cases with intracranial haemorrhages (ICH) and 98 cases with traumatic brain injuries (TBI) to test. We only evaluate the performance of outlier detection within a brain mask that was derived from the atlas used for pre-registration. We set the pixel intensities outside of that mask to -20, clip the intensities to the HU range of $[-20, 300]$ and rescale to $[-1, 1]$. Examples of the raw mid-axial slices before masking are shown in Figure 1. The blood lesions are clearly visible as bright spots in the images, whereas oedema are not trivial to identify due to the low contrast differences compared to normal tissue.

We train convolutional autoencoders based on the implementation in DLTK [9] using Adam [5].We compare our approach to regular autoencoders, variational autoencoders [6] and denoising autoencoders as well as simple baselines such as the difference to the mean training image and the difference to the CT atlas. Further, we trained a Gaussian mixture model (GMM) on the intensities within the brain mask and used the fitted model to score the probability of unseen pixel values within the brain mask. For the denoising, variational and Bayesian autoencoders we use 100 MC estimates. Here, the denoising autoencoder requires a MC estimate as we also apply additional noise during testing.

Figure 2 shows an overview of the difference maps generated by the different methods for an example from ICH (first row) and TBI (second row). All methods have issues with imperfect skull stripping as the intensity differences are quite high. The GMM provides the least noisy difference maps, however fails to capture some lesion regions. The atlas and mean image based differences achieve good performances, but have difficulties dealing with structural information. The autoencoder-based methods should be able to capture the structural information better. However, the regular and denoising autoencoder exhibit noisy reconstruction errors. The dropout and variational autoencoder smooth this noise as they combine multiple samples from the weights or latent code.

We show ROC curves as well as their AUC for quantitative results in Figure 3. Those curves show the true and false positive rates across all pixels given the values from the concatenated masked difference maps. The MC-Dropout-based Bayesian autoencoder achieves the best performance on the task of detecting blood in the masked CT mid-axial slice.



(a) ICH (b) TBI

Figure 3: ROC curves for the segmentation of blood using thresholding of difference maps.

4 Discussion & Conclusion

We show that autoencoders are able to perform outlier detection because they fit a lower dimensional manifold for data compression and project unseen data onto this manifold. Therefore, outliers will be lost and are visible as reconstruction error. The Bayesian variant of this approach enables the autoencoder to smooth out uncertainties in the weight space and outperforms the baselines.

GANs [4] are shown to be able to perform similar tasks [11] but have training instabilities. Because of this, we were not able to train a GAN with sufficient fidelity to the data at hand. Future work should evaluate the method on full 3D volumes and use improved approximations for Bayesian deep learning[8, 10]. Lastly, better tuned GAN training should be used for another baseline that might fit the manifold better.

Acknowledgments: NP is supported by Microsoft Research PhD Scholarship and the EPSRC Centre for Doctoral Training in High Performance Embedded and Distributed Systems (HiPEDS, Grant Reference EP/L016796/1). MR is supported by an Imperial College Research Fellowship. This work is partially funded by a European Union FP7 grant (CENTER-TBI; Agreement No: 60215). We gratefully acknowledge the support of NVIDIA with the donation of one Titan X GPU.

References

[1] Jinwon An and Sungzoon Cho. “Variational autoencoder based anomaly detection using reconstruction probability”. In: *SNU Data Mining Center, Tech. Rep.* (2015).

[2] Diego Carrera, Giacomo Boracchi, et al. “Detecting anomalous structures by convolutional sparse models”. In: *IJCNN*. 2015.

[3] Yarin Gal and Zoubin Ghahramani. “Bayesian convolutional neural networks with Bernoulli approximate variational inference”. In: *arXiv preprint arXiv:1506.02158* (2015).

[4] Ian Goodfellow, Jean Pouget-Abadie, et al. “Generative adversarial nets”. In: *NIPS*. 2014.

[5] Diederik Kingma and Jimmy Ba. “Adam: A method for stochastic optimization”. In: *arXiv preprint arXiv:1412.6980* (2014).

[6] Diederik P Kingma and Max Welling. “Auto-Encoding Variational Bayes”. In: *arXiv preprint arXiv:1312.6114* (2013).

[7] Geert Litjens, Thijs Kooi, et al. “A survey on deep learning in medical image analysis”. In: *Medical image analysis* 42 (2017), pp. 60–88.

[8] Christos Louizos and Max Welling. “Multiplicative Normalizing Flows for Variational Bayesian Neural Networks”. In: *arXiv preprint arXiv:1703.01961* (2017).

[9] Nick Pawlowski, Sofia Ira Ktena, et al. “DLTK: State of the Art Reference Implementations for Deep Learning on Medical Images”. In: *arXiv preprint arXiv:1711.06853* (2017).

[10] Nick Pawlowski, Martin Rajchl, and Ben Glocker. “Implicit Weight Uncertainty in Neural Networks”. In: *arXiv preprint arXiv:1711.01297* (2017).

[11] Thomas Schlegl, Philipp Seeböck, et al. “Unsupervised anomaly detection with generative adversarial networks to guide marker discovery”. In: *IPMI*. 2017.

[12] Nitish Srivastava, Geoffrey E Hinton, et al. “Dropout: a simple way to prevent neural networks from overfitting.” In: *JMLR* (2014).

[13] Koen Van Leemput, Frederik Maes, et al. “Automated segmentation of multiple sclerosis lesions by model outlier detection”. In: *IEEE transactions on medical imaging* (2001).