Unsupervised reconstruction-based anomaly detection using a Variational Auto Encoder

Soumick Chatterjee[1], Alessandro Sciarra[1,2], Max Dünnwald[1,2], Shubham Kumar Agrawal[1], Pavan Tummala[1], Disha Setlur[1], Aman Kalra[1], Aishwarya Jauhari[1], Steffen Oeltze-Jafra[2,3,4], Oliver Speck[1,3,4,5], and Andreas Nürnberger[1,3]

[1]Otto von Guericke University, Magdeburg, Germany; [2]Medical Faculty, University Hospital, Magdeburg, Germany, [3]Center for Behavioral Sciences, Magdeburg, Germany, [4]German Center for Neurodegenerative Disease, Magdeburg, Germany; [5]Leibniz Institute for Neurobiology, Magdeburg, Germany

Introduction

Unsupervised pixel-precise segmentation of brain regions that appear anomalous or not can be a valuable assistance for radiologists. Most of the classification/segmentation models proposed use supervised training for a certain task and need large training data. Unsupervised anomaly detection (UAD)[1] systems can directly learn the data distribution from a large cohort of unannotated subjects and then be used to detect out of distribution samples and thus ultimately identify diseased or suspicious cases. The approach can be made independent of human input by decoupling reference annotations from abnormality detection. Such anomaly detection systems can be trained to detect any kind of anomaly present in the data without explicitly teaching them about any specific task. In this research, we have implemented an anomaly detection approach and tested the approach to segment artificial anomalies and brain tumours.

Methodology

We used a Variational Auto-Encoder (VAE), an auto-regressive model. The architecture of the network is a 5-layer fully convolutional encoder and decoder. During training, the model computes a Latent Space Representation (LSR) of the non-anomalous input image thus learning feature distribution of healthy dataset. The encodings obtained here are 2 vectors; means (μ) and standard deviations (σ). The loss function used is the evidence lower bound (ELBO), a combination of the error on the pixel-wise reconstructions, along with Kullback-Leibler (KL) Divergence. We trained the network by optimising the evidence lower bound (ELBO) using the MOOD [5] dataset. It comprises 800 handselected brain scans containing no anomalies. A total of 256 latent variables are used initially. The model is trained with Adam optimizer along with a learning rate of 1e-4 over a total of 100 epochs until it reaches convergence. To localize the anomalies, a novel post processing was performed i.e. morphological closing along with binary thresholding is performed on the mask obtained by subtracting the reconstructed image from the original and the location of maximum anomaly is detected. Evaluation were performed using artificial anomalies present in the MOOD dataset and using tumours (BraTS dataset [6]).

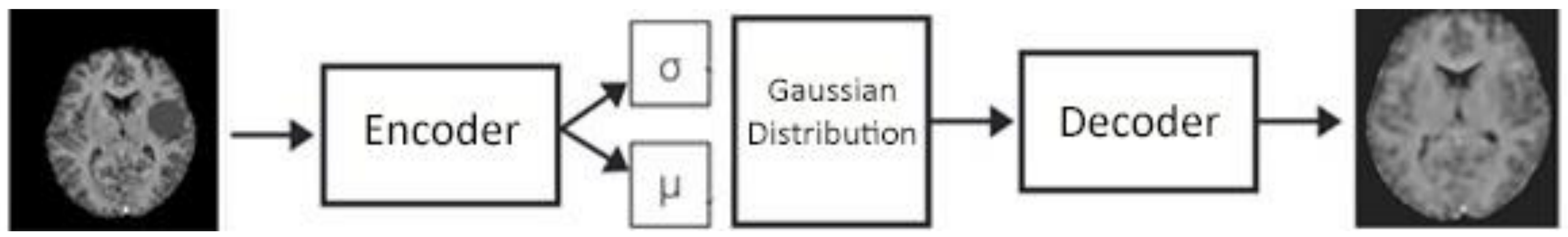

Fig 1: Block diagram of a Variational Autoencoder (VAE)

Results and Discussion

Fig 2 shows the results of two example reconstructions from the MOOD dataset, for anomaly free and anomalous data respectively. It can be observed that the model was able to identify the artificial anomaly present in the dataset correctly. Moreover, Fig 3 shows the results of an example reconstruction from the BraTS dataset, where the anomaly has been segmented but also generated some additional noise. We have benchmarked the results against the results obtained in [2], where model was trained and evaluated on the BraTS dataset, and obtained a Dice coefficient of 0.36. Our model has been trained on the MOOD dataset and evaluated on the BraTS dataset, obtaining a Dice Coefficient of 0.27.

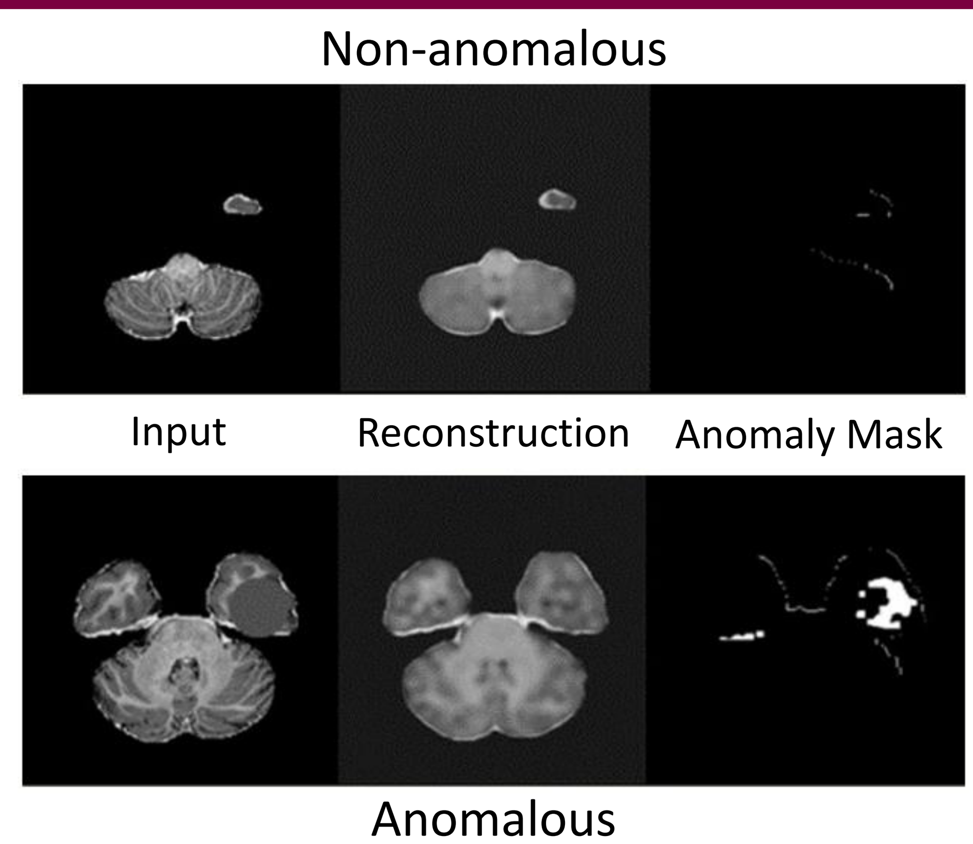

Fig 2: Results while Reconstructing MOOD dataset

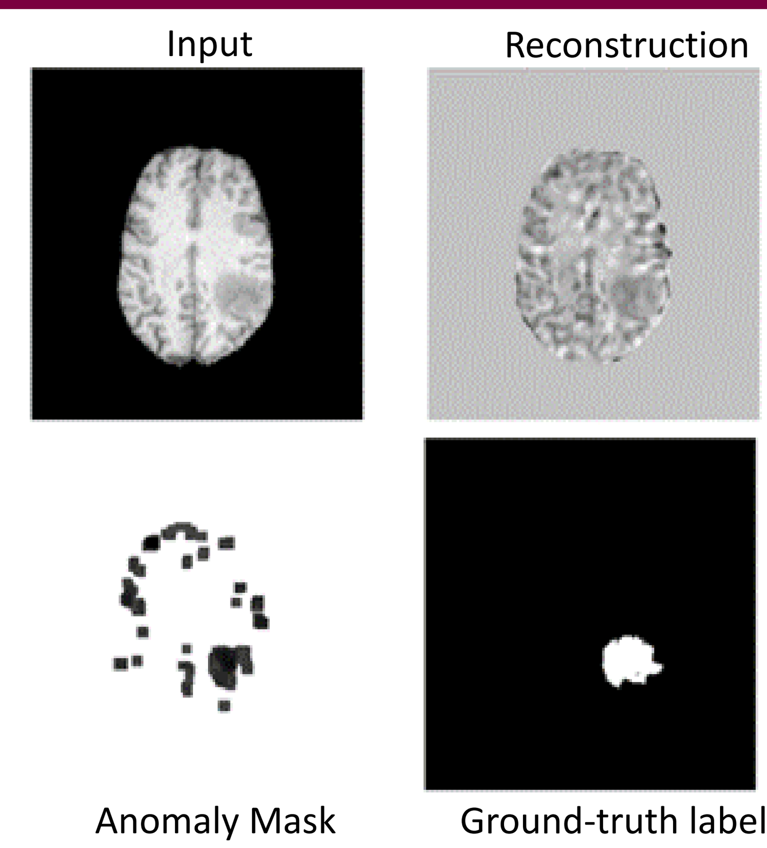

Fig 3: Results while Reconstructing BraTS dataset

Conclusion and future work

We demonstrated a proof-of-concept UAD that encodes the full context of brain MR slices. This approach was successful in detecting anomalous patterns in the MOOD dataset and provides opportunities for effective unsupervised training for anomaly detection, which can be used further for disease localization without explicit annotations. It was observed while localizing anomalies, the model generated many false positives. However, this kind of UAD might be used by clinicians for interactive decision making. As a future work, we plan to investigate the projection of healthy anatomy into a latent space that follows a Gaussian Mixture Model and intend to utilize 3D autoencoding models.

References

[1] Bergmann et al. MVTec AD – A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection. CVPR 2019.
[2] Zimmerer et al. (2019). Unsupervised Anomaly Localization using Variational Auto-Encoders
[3] An et al. "Variational Autoencoder based Anomaly Detection using Reconstruction Probability." (2015).
[4] Kiran et al. (2018). An Overview of Deep Learning Based Methods for Unsupervised and Semi-Supervised Anomaly Detection in Videos. Journal of Imaging.
[5] MOOD challenge dataset : https://zenodo.org/record/3961376
[8] BraTS dataset https://www.med.upenn.edu/sbia/brats2017/data.html

Acknowledgment

This work was in part conducted within the context of the International Graduate School MEMoRIAL at Otto von Guericke University (OVGU) Magdeburg, Germany, kindly supported by the European Structural and Investment Funds (ESF) under the programme Sachsen-Anhalt WISSENSCHAFT Internationalisierung (project no. ZS/2016/08/80646).

Email: soumick.chatterjee@ovgu.de
Phone: +49-391-6756119