Capturing Inter-Slice Dependencies of 3D Brain MRI-Scans for Unsupervised Anomaly Detection

Finn Behrendt[1] FINN.BEHRENDT@TUHH.DE
Marcel Bengs[1] MARCEL.BENGS@TUHH.DE
Debayan Bhattacharya[1] DEBAYAN.BHATTACHARYA@TUHH.DE
Julia Krüger[2] JULIA.KRUEGER@JUNG-DIAGNOSTICS.DE
Roland Opfer[2] ROLAND.OPFER@JUNG-DIAGNOSTICS.DE
Alexander Schlaefer[1] SCHLAEFER@TUHH.DE
[1] *Institute of Medical Technology and Intelligent Systems, Hamburg University of Technology*
[2] *Jung Diagnostics GmbH*

Editors: Under Review for MIDL 2022

Abstract

The increasing workloads for radiologists in clinical practice lead to the need for an automatic support tool for anomaly detection in brain MRI-scans. While supervised learning methods can detect and localize lesions in brain MRI-scans, the need for large, balanced data sets with pixel-level annotations limits their use. In contrast, unsupervised anomaly detection (UAD) models only require healthy brain data for training.
Despite the inherent 3D structure of brain MRI-scans, most UAD studies focus on slice-wise processing. In this work, we capture the inter-slice dependencies of the human brain using recurrent neural networks (RNN) and transformer-based self-attention mechanisms together with variational autoencoders (VAE). We show that by this we can improve both reconstruction quality and UAD performance while the number of parameters remain similar to the 2D approach where the slices are processed individually.

Keywords: Deep Learning, Unsupervised Anomaly Detection, 3D Brain MRI

1. Introduction

Magnetic Resonance Imaging (MRI) is commonly used for the detection and diagnosis of lesions in human brains. Radiologists manually annotate the MRIs which is time-consuming and error-prone (Kim and Mansfield, 2014). Therefore, an automated decision support tool that can guide the radiologists' assessment would be beneficial. To overcome the need for large, pixel-wise annotated data sets, unsupervised anomaly detection (UAD) methods, which learn a representation of healthy brain anatomy, have shown promising results. Unhealthy brains can then be discriminated from healthy brains by their distance from the learnt latent healthy brain representation. By this, the labelling effort solely involves keeping anomalous brain images out of the training set. This substantially reduces the labelling effort. Numerous 2D approaches (Baur et al., 2021) have been proposed where the 3D volumes are processed in a slice-wise fashion. However, including 3D information by using 3D-convolutions has been shown to improve UAD performance (Bengs et al., 2021). One reason for the dominance of the slice-wise approach is the increased computational cost of processing 3D volumes by 3D-convolutions.
In this paper, we capture the inter-slice dependencies of the MRI-scans in the latent space using RNNs and transformers. By this, we can model the inter-slice relationships efficiently without the need for costly 3D-convolutions, keeping the number of parameters similar to a 2D approach.

2. Methods

Data Sets:
We use T1- and T2-weighted MRI-scans of subjects without known anomalies from the IXI[1] data set. 387 samples are used for training, 43 samples as healthy validation set and 160 samples as healthy test set. For evaluation, we use MRI-scans from two publicly available data sets, namely the Brain Tumor Segmentation Challenge 2019 data set (Menze et al., 2014) (BraTS19) with 335 T2-weighted samples and the Anatomical Tracings of Lesions After Stroke data set (Liew et al., 2018) (ATLAS) with 304 T1-weighted samples. We construct validation sets of 166 and 175 samples from the ATLAS and BraTS19 data sets, respectively. The remaining samples are used as unhealthy test sets. We register all skull-stripped scans to the SRI24-Atlas, cut the volumes to the brain size and pad the resulting volume to a resolution of $160 \times 190 \times 160$ voxels before resizing each dimension by a factor of 0.5. As post-processing, we apply a median filter, similar to (Baur et al., 2021).
Deep Learning Methods:
The input to our models is a 3D MRI-scan $V \in \mathbb{R}^{H \times W \times D}$ where H, W and D are the height, width and depth of the volume, respectively. In the 2D-VAE framework (Baur et al., 2021), individual slices $S \in \mathbb{R}^{H \times W}$ are encoded to a latent vector $z \in \mathbb{R}^{d_z}$. From the latent vector, the input slice is reconstructed by a decoder network. In contrast to that, in the 3D-VAE framework (Bengs et al., 2021) the entire volume $V \in \mathbb{R}^{H \times W \times D}$ is encoded to and reconstructed from $z \in \mathbb{R}^{d_z}$. In our approach, we individually encode all D slices $S \in \mathbb{R}^{H \times W}$ of $V \in \mathbb{R}^{H \times W \times D}$ into a stack of latent vectors $z \in \mathbb{R}^{d_z \times D}$. This stack can be considered as a sequence and is fed into a sequential model which outputs a sequence of latent vectors $\tilde{z} \in \mathbb{R}^{d_z \times D}$. The individual slices are then reconstructed from their corresponding latent vector in the sequence $\tilde{z}$ and stacked together to the reconstructed volume $\hat{V} \in \mathbb{R}^{H \times W \times D}$. With this sequential VAE framework, we capture inter-slice dependencies which are ignored in the 2D-VAE framework by a sequential model. We evaluate bidirectional Gated Recurrent Units (Chung et al.) (S-VAE-GRU), Long-Short-Term-Memory (Hochreiter and Schmidhuber, 1997) (S-VAE-LSTM) networks and multi-headed self-attention layers from transformer models (Vaswani et al., 2017) (S-VAE-Trans) as sequential models. We train our models for 800 epochs, use Adam as optimizer, a learning rate of $lr = 5 \times 10^{-4}$ and a latent vector size of $d_z = 32$ for 2D-VAEs and $d_Z = 512$ for 3D-VAEs.

3. Results & Discussion

To evaluate our methods, we consider the Dice score (DICE) and utilize our validation set to determine the binarization threshold. Furthermore, the AUPRC and the reconstruction error ($l1$-error) of healthy brain anatomy is reported. We train the models with 5 random seeds and report the mean with twice the standard deviation. Our results from Table 1 show that the S-VAE approaches outperform the 2D-VAE across both data sets. This indicates that 3D information is beneficial and can be captured with our sequential approach. Moreover, in contrast to a previous study (Bengs et al., 2021) we find no clear improvements for the 3D-VAE. This could be explained by our different and smaller training set together with the large number of parameters of the 3D-VAE.

1. https://brain-development.org/ixi-dataset/

Table 1: Averaged performance metrics. The $l1$-error is evaluated on the healthy test sets. The DICE and AUPRC values are provided in percent.

	Brats (T2)		ATLAS (T1)		$l1$-error (10^{-3})		param.
Model	DICE	AUPRC	DICE	AUPRC	IXI (T2)	IXI (T1)	(10^6)
2D-VAE	41.26±1.27	36.62±1.48	27.51±0.25	18.47±0.42	31.74±0.02	38.41±0.02	1.412
3D-VAE	42.25±1.17	37.28±1.17	26.20±0.55	18.43±0.67	29.26±0.03	35.35±0.03	10.486
S-VAE-GRU	**46.85±1.01**	**43.85±1.46**	28.37±2.53	19.27±4.57	**27.74±0.01**	**33.38±0.01**	1.417
S-VAE-LSTM	46.00±1.27	43.07±2.45	27.82±1.52	18.76±2.55	29.49±0.06	34.73±0.02	1.418
S-VAE-Trans	45.39±1.58	42.36±2.20	**30.41±0.29**	**23.06±0.29**	30.17±0.04	35.12±0.00	1.422

Overall, we show that S-VAEs can efficiently capture inter-slice relationships of 3D brain MRI-volumes while preserving a number of parameters similar to 2D-VAEs. Our results indicate that this leads to improved reconstruction quality and segmentation performance. While our findings support that 3D information can be beneficial for UAD, systematically comparing different approaches for 3D processing is an interesting direction for future work.
Funding: This work was partially funded by Grant Number ZF4026303TS9.

References

Christoph Baur, Stefan Denner, Benedikt Wiestler, Nassir Navab, and Shadi Albarqouni. Autoencoders for unsupervised anomaly segmentation in brain mr images: a comparative study. *Med. Image Anal.*, page 101952, 2021.

Marcel Bengs, Finn Behrendt, Julia Krüger, Roland Opfer, and Alexander Schlaefer. Three-dimensional deep learning with spatial erasing for unsupervised anomaly segmentation in brain mri. *CARS*, pages 1–11, 2021.

Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.

Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

Young W. Kim and Liem T. Mansfield. Fool me twice: Delayed diagnoses in radiology with emphasis on perpetuated errors. *AJR*, 202(3):465–470, 2014.

Sook-Lei Liew, J. M. Anglin, Nick W. Banks, and Matt Sondag et al. A large, open source dataset of stroke anatomical brain images and manual lesion segmentations. *Scientific data*, 5:180011, 2018.

Bjoern Menze, András Jakab, Stefan Bauer, and Jayashree et al. Kalpathy-Cramer. The multimodal brain tumor image segmentation benchmark (brats). *IEEE Transactions on Medical Imaging*, 99, 2014.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *NIPS*, volume 30, 2017.