"Metode de optimizare Riemanniene pentru învățare profundă"
Proiect cofinanțat din Fondul European de Dezvoltare Regională prin
Programul Operațional Competitivitate 2014-2020

Detection of Tumours in Brain MRIs with Variational AutoEncoders

Alexandra-Ioana Albu Alina Enescu Luigi Malagò

Romanian Institute of Science and Technology

Machine Learning for Pharma and Healthcare Applications, ECML PKDD 2020 Workshop, Ghent September 14

Semi-supervised Anomaly Detection

Standard approach in the literature: learn the distribution of healthy images and identify anomalies as outliers

Our approach employs in training an additional unlabelled dataset of healthy and unhealthy individuals (semi-supervised negative/unlabelled learning)

Additionally, to improve the quality of the reconstructions, we have employed a perceptual loss regularizer during the training of `VAE-H`.

Anomaly detection algorithm (Albu et al., 2020)

Train 2 VAEs: `VAE-H` on data without anomalies, `VAE` on mixed data

Obtain the reconstruction x_h of the original image x through `VAE-H`

Compute $d_x = d(\texttt{VAE.enc}(x), \texttt{VAE.enc}(x_h))$ for a given distance / dissimilarity measure in the latent space

Classify anomalies based on a threshold $d_x^\star$ learned by cross-validation

VAE-H: healthy images in training

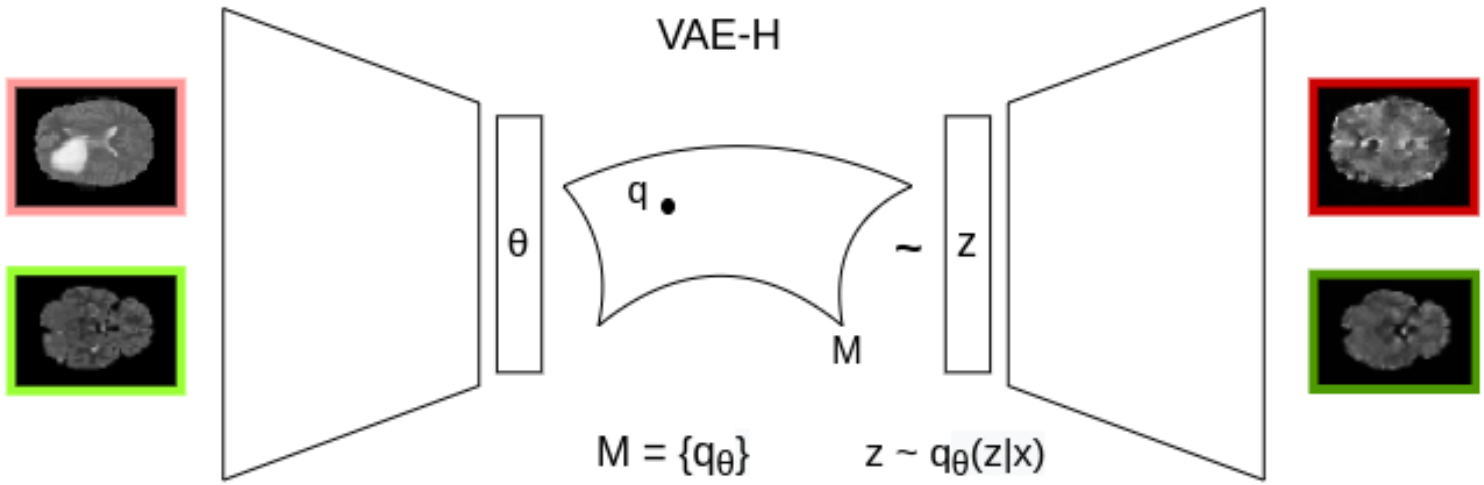


Figure: Architecture of VAE trained on healthy.

VAE: healthy and unhealthy images in training

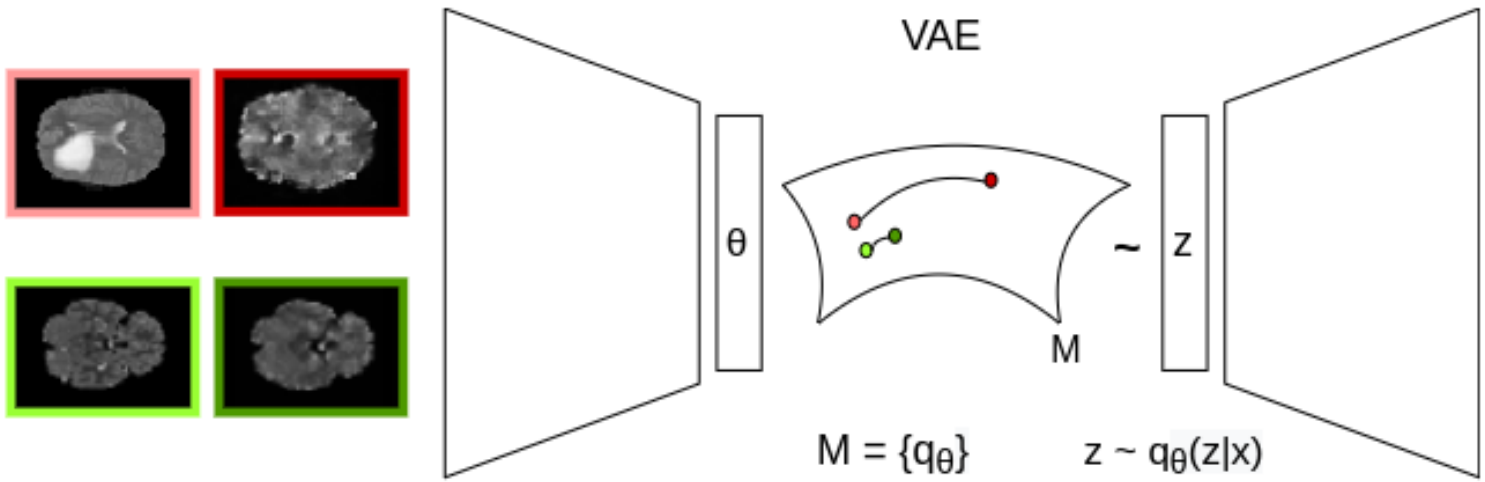


Figure: Workflow of proposed methodology.

Datasets

- **HCP dataset** [5] (the dataset of healthy individuals)
- **BRATS-2018 dataset** [3, 2] (the dataset of both normal and tumoural tissues)
- Data preprocessing:
  - removed the black slices from both datasets
  - applied bias field correction to the images using the N4ITK algorithm [4]
  - normalized in the $[-1, +1]$ range
  - matched the histogram of each individual
  - cropped and resized the images to 200×200 and down sampled them to 128×128
  - Gaussian noise and data augmentation were applied

Comparison with a baseline method that uses only the VAE trained on healthy images

Table: Area Under the ROC Curve, accuracy and F1 score for BRATS-2018 (test set, averaged over 5 folds) for L^2 distances computed in the input space versus L^2 of the means in the latent space. PL stands for Perceptual Loss. Intervals (±) correspond to standard deviations.

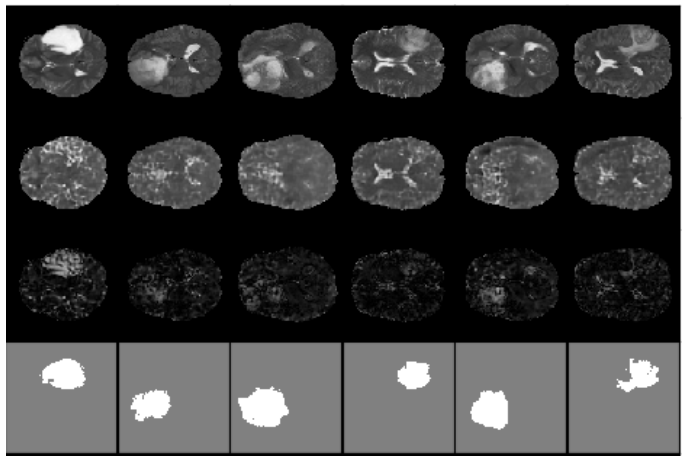
	ROC-AUC			Accuracy			F1 score		
Tumour size	L^2 input	Without PL	PL	L^2 space	Without PL	PL	L^2 input	Without PL	PL
0	0.848 ± 0.004	0.884 ± 0.003	0.890 ± 0.002	0.652 ± 0.008	0.805 ± 0.002	0.804 ± 0.004	0.507 ± 0.006	0.803 ± 0.004	0.809 ± 0.007
20	0.851 ± 0.004	0.890 ± 0.004	0.895 ± 0.003	0.673 ± 0.007	0.809 ± 0.002	0.809 ± 0.008	0.520 ± 0.005	0.803 ± 0.005	0.803 ± 0.010
50	0.852 ± 0.004	0.892 ± 0.003	0.897 ± 0.002	0.684 ± 0.006	0.811 ± 0.002	0.811 ± 0.006	0.526 ± 0.005	0.802 ± 0.004	0.801 ± 0.009
150	0.856 ± 0.004	0.897 ± 0.004	0.901 ± 0.003	0.707 ± 0.005	0.814 ± 0.002	0.818 ± 0.007	0.541 ± 0.005	0.795 ± 0.005	0.800 ± 0.004

Comparison with the state-of-the-art anomaly detection methods

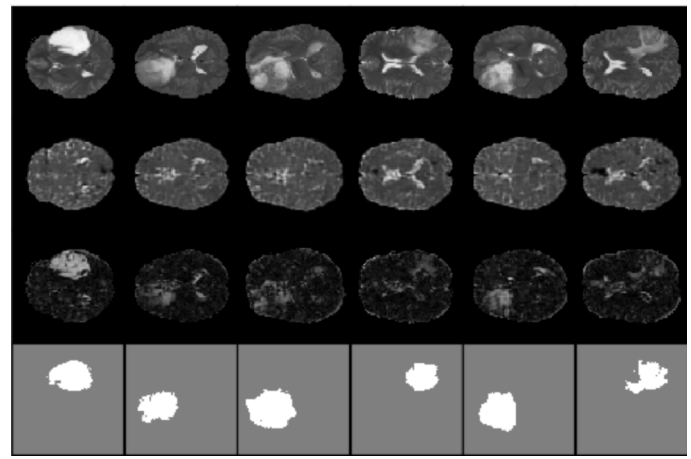
Table: ROC-AUC values. The table includes the resolution of images after rescaling (Res.) and the threshold (Th.) expressed in annotated pixels (before rescaling) used to determine if a slice contains anomalies.

Method	ROC-AUC	Dataset	# patients test set	Res.	Th.
ADAE [6]	0.892	BRATS2017	285	32×32	?
ceVAE [7]	0.867	BRATS2017	266	64×64	20
Proposed model without PL	0.890	BRATS2018	69	128×128	20
Proposed model with PL	0.895	BRATS2018	69	128×128	20

Examples of reconstructions through VAE-H of scans containg tumours



(a) Model without PL.



(b) Model without PL.

Figure: First row: original BRATS images from the test set. Second row: reconstruction through VAE-H. Third row: residual between original and reconstructed images. Fourth row: ground truth segmentation of the tumour.

Comclusions and future work

We have presented improved results for anomaly detection compared to the preliminary experiments presented in [1].

As expected, by using more powerful models, by maximizing the likelihood only for the pixels in the brain based on precomputed segmentation masks, as well as by using the perceptual loss, we were able to obtain sharper reconstructions and, in turn, increase the performance of the slice-wise tumour detection algorithm.

Future work will be carried out in the direction of obtaining more informative features to be used for the perceptual loss, such as features extracted from segmentation models.

Moreover, we plan to study the possibility of combining the components of out framework in a single model, trained end-to-end.

References


Alexandra Albu, Alina Enescu, and Luigi Malagò.
Tumor detection in brain mris by computing dissimilarities in the latent space of a variational autoencoder.
In *Proceedings of the Northern Lights Deep Learning Workshop*, volume 1, pages 6–6, 2020.


Michael Kistler, Serena Bonaretti, Marcel Pfahrer, Roman Niklaus, and Philippe Büchler.
The virtual skeleton database: an open access repository for biomedical research and collaboration.
Journal of medical Internet research, 15(11), 2013.


Bjoern H Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, Roland Wiest, et al.
The multimodal brain tumor image segmentation benchmark (BRATS).
IEEE transactions on medical imaging, 34(10):1993–2024, 2014.