UNIVERSITÉ DE LYON

CREATIS

Master Internship 2021 - 2022

Self supervised learning for anomaly detection in MRI neuroimaging. Application to early diagnosis of Parkinson disease.

Host laboratory : Laboratoire CREATIS, 69 Villeurbanne- MYRIAD Team
Supervisors : Carole Lartizien -Nicolas Pinon
Key words : Neuroimaging, Deep learning, Self-supervised learning, Latent representation learning
Duration : 6 months
Starting date : Winter- Spring 2021-22 (flexible)
Gratuity : $\sim$ 560 euros/month

Scientific context

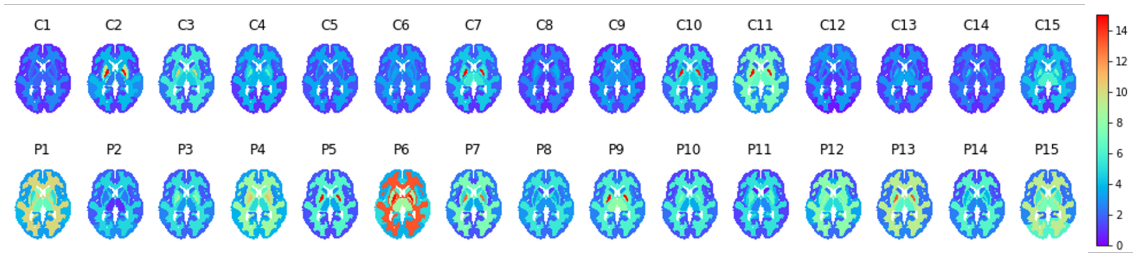

Figure 1: Example parametric maps of percentage of abnormal voxels found by the SAE per anatomical region of interest. Top: Test control healthy subjects; Bottom: 15 randomly selected PD patients.[1]

Recent advances in machine learning have led to very promising results in medical imaging for segmentation, registration or reconstruction as well as for the design of automated diagnosis or prognosis models.

The vast majority of deep architecture for medical image analysis are based on supervised method requiring the collection of large datasets of annotated examples. Building such annotated datasets is hardly achievable, especially for some specific tasks, including the detection of small and subtle lesions, which are sometimes impossible to visually detect and thus manually outline. This is the case for various brain pathologies including microbleeds, epilepsy or multiple sclerosis lesions as well as Parkinson's disease (PD) [2].

An alternative methodological framework is that of anomaly detection in an *unsupervised* context (also called *self-supervised*). It consists in learning a model of representation of normality from the healthy data only, and then to consider as anomalies the test samples that deviate too much from normality. This last step is usually performed by calculating the error between the original and the reconstructed data from their projection in the latent representation space.

We have developed an expertise in the field of anomaly detection methods for the analysis of multi-modality brain images. Our approach is based on representation learning by Siamese auto-encoding networks (SAEs) as illustrated in the bottom of the figure 2. This type of model gives good performances for epileptogenic lesion detection in multiparametric MRI [3]. In collaboration with the Grenoble Institute of Neurosciences (GIN) and INRIA, we recently applied it to the detection of early forms of Parkinson's disease in Parkinson's disease in multiparametric MRI and compared its performance with that of a simple auto-encoder (AE) illustrated on the top part of figure 2. The results of this study illustrated on figure 1 were recetnly presented at the 2021 MICCAI MLCN (Machine learning clinical neuroimaging) workshop [1].

Figure 2 illustrates the two architectures used for this project. One of the particularities of the SAE (bottom figure) is to force the colocalized patches in the brain to be close in the latent space according to a certain metric. The influence of the choice of this metric as well as that of the network architecture on the spatial distribution of the encoded patches in the latent space is still little studied and could be an aid to the choice of the former.

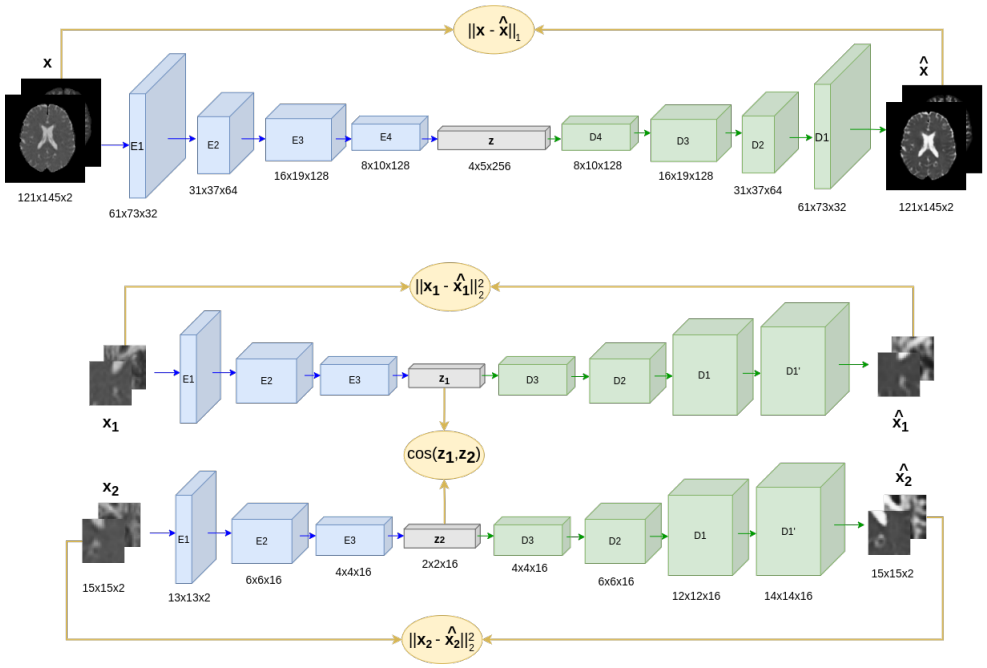

Figure 2: Classical auto-encoder (AE) on top, Siamese auto-encoder (SAE) at the bottom for latent representation learning of healthy subjects.

Objectives

The purpose of this master project is to improve performance achieved with the current model architectures by exploring two methodological research axes :

- Comparison of patch versus whole image and SAE versus AE approaches :
  The current versions of the SAE and AE models consider as input 2D patches (15x15) extracted from the original images for the former, and 2D cross-sections of the brain or a hemisphere for the latter. Compared to the AE model, the SAE model increases the size of the learning base (extraction of tens of thousands of patches per 3D image volume) at the expense of a loss of information on the spatial context. Moreover, the SAE model constrains the colocalized patches of different patients to be close in the latent space, the AE model does not impose constraints on the colocalized slices.
  The first objective of the internship will be to implement the complementary SAE and AE models (SAE with whole image/ AE with patches).

- Development of analysis tools of the latent space :
  In order to better understand the mechanisms of these models, we wish to explore the characteristics of their representation space. The objective will be to take in hand methods of visualization of the latent space (t-SNE, UMAP, spectral embedding, etc.) and adapt them to our problem. It could also be interesting to study the influence of the chosen cost functions or the variational regularization on these representations.

The successful candidate will have access to the PPMI database (https://www.ppmi-info.org/accessdata-specimens/download-data) containing multiple images of controls and parkinsonian (PD) patients in different modalities and as well as to computing resources (CREATIS and/or CNRS supercomputer).

Skills

Candidate should have a background either in machine learning and/or deep learning or image processing as well as good programming skills. Experience with deep learning libraries such (TensorFlow, Pytorch, scikit-learn) would be apreciated.

We are looking for an enthusiastic and autonomous student with strong motivation and interest in multidisciplinary research (image processing and machine learning in a medical context). The candidate will also have strong interaction with a PhD student, Nicolas Pinon, working on this project.

Application

Interested applicants are required to send a cover letter, CV and any other relevant documents (reference letter, recent transcripts of marks,. . .) to:
carole.lartizien@creatis.insa-lyon.fr et **nicolas.pinon@creatis.insa-lyon.fr**

References

[1] Verónica Muñoz-Ramírez, Nicolas Pinon, Florence Forbes, Carole Lartizen, and Michel Dojat. Patch vs. global image-based unsupervised anomaly detection in mr brain scans of early parkinsonian patients. In Ahmed Abdulkadir, Seyed Mostafa Kia, Mohamad Habes, Vinod Kumar, Jane Maryam Rondina, Chantal Tax, and Thomas Wolfers, editors, *Machine Learning in Clinical Neuroimaging*, pages 34–43, Cham, 2021. Springer International Publishing.

[2] Verónica Muñoz Ramírez, Virgilio Kmetzsch, Florence Forbes, and Michel Dojat. Deep learning models to study the early stages of parkinson's disease. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pages 1534–1537, 2020.

[3] Zaruhi A., J. Jung, R. Bouet, and C. Lartizien. Regularized siamese neural network for unsupervised outlier detection on brain multiparametric magnetic resonance imaging: Application to epilepsy lesion screening. *Medical Image Analysis*, 60:101618, 2020.