CREATIS

Internship position Fall 2024/2025

Self-supervised deep learning representation learning for 3D brain MRI analysis

Scientific context

The Master student will join a research project gathering specialists with complementary expertise in image processing and machine learning (CREATIS) and medicine (HCL). The project is funded by the ANR research grant SEIZURE. The main objective is to **investigate popular self-supervised deep learning based representation learning methods and evaluate how they perform on medical images**. This internship project is part of a broader study to detect epileptogenic zones from heterogeneous data, including MRI, PET, magnetoencephalography (MEG) and clinical data.

Many deep learning methods for image classification, detection or segmentation rely on the extraction of relevant features from input images (Figure 1). In traditional computer vision, a large field of research is dedicated to finding the models to extract those features. However, most of them are designed for and trained on 2D RGB real-world images and consequently do not transfer well to the medical image field. Indeed, medical images are semantically different from natural images and have inherent particularities such as the type of modality (Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Positron Emission Tomography (PET), etc.). Being able to compute meaningful features from medical images could lead to significant improve in performances in many tasks in medical image processing.
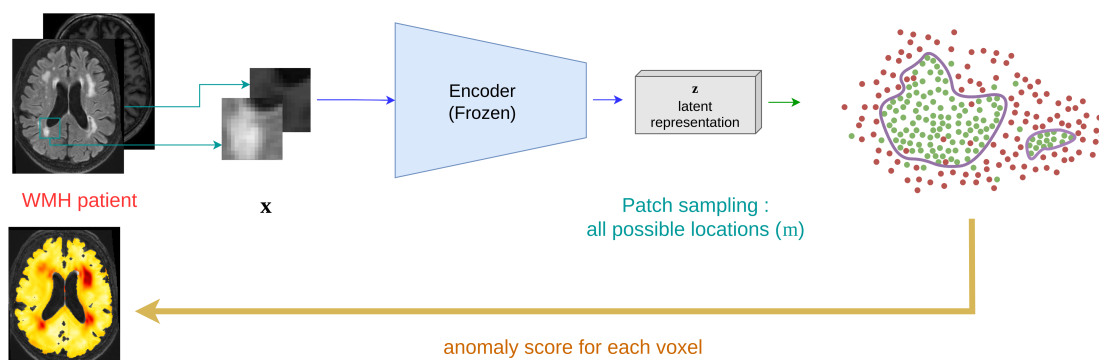

Figure 1: Many tasks in medical image analysis, such as this work [1] in anomaly detection, require extracting meaningful semantic information from input images. Such networks are usually trained on large databases that contain natural images, and are hence not particularly suited for medical images. Having powerful feature extractors could help improve the performance of downstream deep learning algorithms on medical images.

Objectives of the internship

The purpose of this master project is to improve the performance achieved with current models by investigating strategies to extract features from 3D brain MRI. The intern will explore the following methodological axes :

- A first objective is to select and implement several methods from the literature. The candidate will use classic convolutional neural networks (ResNet, WideNet, DINO[2]) but will also have the opportunity to manipulate the most recent deep learning models, among which vision transformers (ViT[3], Swin-ViT) and foundation models (CLIP[4], Segment Anything[5] and variants designed to be better suited to medical images). Open-source collections of standard models, like HuggingFace, can be used to facilitate the implementation of the models. He/she will also explore self-supervised techniques such as reconstruction-based methods (CNN auto-encoders, MAE[6]), contrastive learning (SimCLR[7]) and other pretext tasks (*e.g.* solving Jigsaw puzzle).

- The second research axis is to design strategies to adapt all these models to 3D brain MRI. In general, all the models are tailored for fixed and standardized 2D image size (typically 224×224 on ImageNet dataset), which may not be suited for medical applications. The adaptation strategies may differ from one model to another, in particular depending on if the model is a CNN or a Transformer.

- The last objective is to design and conduct an extensive and rigorous evaluation procedure for all the selected methods. The benchmark will include typical metrics for evaluating feature exaction methods (accuracy after k-NN clustering and linear probing) as well as metrics on more complex downstream tasks (*e.g.* segmentation, anomaly detection) and visual analysis (UMAP, t-SNE). The results will permit to draw conclusions on which method(s) can get the most informative features from medical images, and quantify the benefits of leveraging pretrained model weights on large natural image databases.

The candidate will work closely with other students involved in the project, in particular with Robin Trombetta, a PhD student who will prepare the general pipeline analysis (data formatting, image preprocessing, etc.) before the arrival of the student and co-supervise the master project. The work carried out during this internship could lead to a publication in a national or international conference.

Skills

The 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 as PyTorch would be appreciated. We are looking for an enthusiastic, autonomous and rigorous student with strong motivation and interest in multidisciplinary research (image processing and machine learning in a medical context).
He/she will have access to computing resources (CREATIS and/or CNRS supercomputer) as well as to public datasets with 1000+ subjects (ADNI, OASIS and BraTS) and to the private database of the SEIZURE project gathering multimodality exams of epilepsy patients. The candidate will benefit from a stimulating research environment, as he/she will have the opportunity to interact with clinicians and members of the MYRIAD team working in the field of deep machine learning for medical image analysis.

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** and **robin.trombetta@creatis.insa-lyon.fr**

References

[1] Nicolas Pinon, Robin Trombetta, and Carole Lartizien. "One-Class SVM on siamese neural network latent space for Unsupervised Anomaly Detection on brain MRI White Matter Hyperintensities". In: *Medical Imaging with Deep Learning*. 2023. URL: https://openreview.net/forum?id=_c9r6-HCEaN.

[2] Mathilde Caron et al. "Emerging Properties in Self-Supervised Vision Transformers". In: *Proceedings of the International Conference on Computer Vision (ICCV)*. 2021.

[3] Alexey Dosovitskiy et al. "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale". In: *International Conference on Learning Representations*. 2021. URL: https://openreview.net/forum?id=YicbFdNTTy.

[4]   Alec Radford et al. "Learning Transferable Visual Models From Natural Language Supervision". In: *International Conference on Machine Learning.* 2021.

[5]   Alexander Kirillov et al. "Segment Anything". In: *arXiv:2304.02643* (2023).

[6]   Kaiming He et al. "Masked Autoencoders Are Scalable Vision Learners". In: *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).* 2022, pp. 15979–15988.

[7]   Ting Chen et al. "A Simple Framework for Contrastive Learning of Visual Representations". In: *Proceedings of the 37th International Conference on Machine Learning.* Vol. 119. Proceedings of Machine Learning Research. PMLR, 13–18 Jul 2020, pp. 1597–1607.