

Internship Proposal

Contrastive Representation Learning in Medical Imaging

Keywords: Deep learning, Contrastive Analysis, Self-Supervised Learning, Medical Imaging, Generative models

Introduction Representation Learning has become a hot research topic in recent years. It deals with the automatic estimate of efficient data representations using deep learning architecture. Data are mapped to a latent space which should capture the important semantic information valuable for downstream tasks (classification, segmentation, etc). Lately, I have been mainly working on two related topics:

1. *Self-Supervised Learning for Medical Imaging*: the goal is to leverage non-annotated or weakly-annotated data to learn a meaningful representation for diagnostic (cancer, disease, etc). Recently, using a highly interpretable and versatile geometric approach [1], we have proposed new losses and regularization terms for a) including both continuous and discrete metadata into contrastive learning [3], b) removing known data biases [1] and 3) integrating generative model representations into contrastive learning [2].
2. *Contrastive Analysis for Medical Imaging*: given two datasets comprising healthy subjects and patients with a certain disease, the goal is the unsupervised identification and separation of the healthy latent patterns from the pathological ones using generative models (VAE, GAN, Diffusion Models) [4].

Objectives Depending on the background and interest of the student, several research directions are possible within each topic. Here, we list some of them:

1. *Self-Supervised Learning*: extend the proposed geometric approach to multi-modal data (e.g., text and images), longitudinal data or adapt it for a Federated Learning approach.
2. *Contrastive Analysis*: integrate new information-theoretic regularizations to better separate and estimate common and salient generative factors; study the identifiability of this kind of models; propose a way to spatially disentangle both spaces (common and salient) so that individual factors should only account for variations in a specific part of the input images.

When Deadline to candidate: 01/2025

Team This project will be carried out under the supervision of P. Gori (Télécom Paris, IPParis).

Salary ~600 euros/month.

Required background M2 in applied mathematics, statistics, computer science, engineering with a good knowledge of Python and deep learning.

How to apply Candidates are invited to send a CV to pietro.gori@telecom-paris.fr detailing their academic background, motivation and if they are willing to continue with a PhD.

References

- [1] BARBANO, C. A., DUFUMIER, B., TARTAGLIONE, E., GRANGETTO, M., AND GORI, P. Unbiased Supervised Contrastive Learning. In *The Eleventh International Conference on Learning Representations (ICLR)* (2023). 1
- [2] DUFUMIER, B., BARBANO, C. A., LOUSET, R., DUCHESNAY, E., AND GORI, P. Integrating Prior Knowledge in Contrastive Learning with Kernel. In *International Conference on Machine Learning (ICML)* (2023). 1
- [3] DUFUMIER, B., GORI, P., VICTOR, J., GRIGIS, A., AND DUCHESNAY, E. Contrastive Learning with Continuous Proxy Meta-data for 3D MRI Classification. In *Medical Image Computing and Computer Assisted Intervention - MICCAI* (2021). 1
- [4] LOUSET, R., DUCHESNAY, E., GRIGIS, A., AND GORI, P. Separating common from salient patterns with Contrastive Representation Learning. In *International Conference on Learning Representations (ICLR)* (2024). 1