Internship Proposal

Combining Knowledge Distillation and Generative AI

Keywords: Deep learning, Foundation models, Knowledge distillation, Medical Imaging, Generative models

Introduction In medical imaging, using multiple modalities to automatically segment anatomical or pathological structures has become a common practice. Since different image modalities can accentuate and better describe different tissues, their fusion can improve the segmentation accuracy. Although multi-modal models usually give the best results, it is often difficult to obtain multiple modalities in a clinical setting. In many cases, only one modality is acquired.

Related Works Three main strategies have been proposed in the literature to deal with problems where multiple modalities are available at training time but some or most of them are missing at inference time. The first one is to train a generative model to synthesize the missing modalities and then perform multi-modal segmentation [1]. The second strategy consists in learning a modality-invariant feature space that encodes the multi-modal information during training, and that allows for all possible combinations of modalities during inference [2]. The third one consists in transferring knowledge from a multi-modal network (Teacher) to a mono-modal one (Student) [3], based, for instance, on generalized knowledge distillation [4].
Recent preliminary results of our team have shown that these strategies, in particular the third one, obtain interesting results only when the number of training samples is low (N<1K), but when the number of samples increases (N>1k), their usefulness decreases.

Objectives The goal of this project will be to propose a new distillation method combining knowledge distillation and generative models. It will be tested using both foundation models (with a few-shot approach) and in-domain Teacher models. We will use the well-known, freely-available and open-access dataset BRATS (brain tumor segmentation). Depending on the background of the student, the project might take a more theoretical or practical path.

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] Ben-Cohen, A., Klang, E., Raskin, S., Soffer, S., Ben-Haim, S., Konen, E., Amitai, M., Greenspan, H.: Cross-modality synthesis from CT to PET using FCN and GAN networks for improved automated lesion detection. Engineering Applications of Artificial Intelligence 78, 186–194, 2018

[2] Dorent, R., Joutard, S., Modat, M., Ourselin, S., Vercauteren, T.: Hetero-Modal Variational Encoder-Decoder for Joint Modality Completion and Segmentation. In: MICCAI, 2019

[3] M. Hu, M. Maillard, Y. Zhang, T. Ciceri, G. Barbera, I. Bloch, P. Gori. Knowledge distillation from multi-modal to mono-modal segmentation network. In : MICCAI, 2020