Self-supervised vision transformers for pathology segmentation in medical images

Project Management and Software Development
for Medical Applications

General Info

Contact Person: Mehdi Astaraki

Contact Email: mehdi.astaraki@ki.se ; mehdi.astaraki@fysik.su.se ; mehast@kth.se ;

The project can be done in Stockholm (as an exchange or visiting student) or remotely.

Project Abstract

Pretraining models on large-scale image data play a vital role in the concept of representation learning-driven solutions. A well-designed pretrained model can potentially learn the local and global characteristics of image data. Such a model can be optimized to improve the performance of segmentation tasks in the absence of large-scale labeled images.

Background and Motivation

Background: Models pretrained on natural images such as image-net are not fundamentally appropriate for medical image analysis applications. On the other hand, applying effective strategies in the learning process of the pretrained models is quite important to capture both details and contextual characteristics. Recently, pretraining models based on the self-supervised strategy of masked image framework resulted in superior performance in different applications.

Motivation: To achieve accurate segmentation performance, large-scale pixel-level labeled datasets are required. However, due to the complexity of data sharing and tedious procedure of volumetric annotations the available datasets are usually limited. On the other hand, there exist quite a lot of unlabeled datasets that perfectly match the idea of self-supervised pretraining models.

Problem statement: In this project, a novel pretraining strategy based on vision transformer-based autoencoders will be employed as a backbone model for further optimizing and fine-tuning to the specific domain of pathology segmentation such as lung tumors.

Previous works:

Masked autoencoder (MAE) is a well-established method for representation learning in which the model learns to reconstruct original images while masking out a randomly selected subregion of that image. Although it is a powerful model for natural images, it is unable to preserve anatomical constraints that appear in medical images. More advanced versions of MAE include masked feature prediction and SimMM.

Student’s Tasks Description

In this project, the student is expected to:

1 – Review the literature to understand the methodology and theoretical background.

2 – Implementation and gaining experience of self-supervised models

3 – Implementation and gaining experience with vision transformers

4 – Practically learn how the pipeline of vision-transformer AE-based self-supervised model works.

5 – Implementing, optimizing, and fine-tuning the pretrained model for specific tasks such as lung tumor segmentation.

6 – The ultimate goal of the project is to deliver a dockerized package of the model for both the training and testing phases.

- All the implementations will be done in PyTorch and MONAI platforms.

From this project, the student will gain theoretical knowledge and practical experience in developing and packaging state-of-the-art self-supervised methods for volumetric medical image segmentation tasks.

Technical Prerequisites

The student is expected to already have proficiency in Python programming language. Theoretical knowledge and practical experience in the implementation of deep learning models for computer vision tasks and/or medical images with PyTorch will be helpful as well.

References

[1] High-resolution 3D abdominal segmentation with random patch network fusion; https://doi.org/10.1016/j.media.2020.101894

[2] Self-Supervised Pre-Training of Swin Transformers for 3D Medical Image Analysis;

https://doi.org/10.48550/arXiv.2111.14791

[3] Disruptive Autoencoders: Leveraging Low-level features for 3D Medical Image Pre-training ; https://doi.org/10.48550/arXiv.2307.16896