Problem Statement: Self-Supervised Learning for Medical Image Analysis

Medical image analysis plays a crucial role in diagnosing diseases and guiding medical interventions. However, labelled medical data is often scarce and expensive to obtain due to the need for expert annotation. In this assignment, your task is to develop a self-supervised learning framework for medical image analysis.

Participants are requested to use a dataset of unlabelled medical images from various modalities such as MRI, CT scans, and X-rays (Images can be Nifti or dicom) . The goal is to leverage the inherent structures and relationships within the data to learn meaningful representations without relying on manual annotations.

You must design and implement a self-supervised learning algorithm that can:

Extract relevant features from medical images in an unsupervised manner.

Utilize these features to perform one or more downstream tasks, such as organ segmentation, tumor detection/segmentation.

Demonstrate robustness and generalization on a separate test set with annotated ground truth. The evaluation criteria will focus on the accuracy, efficiency, and scalability of the proposed self-supervised learning approach. Additionally, you will be judged on the ability to effectively communicate the methodology, experimental results, and insights gained throughout this assignment.

It is encouraged to explore innovative techniques from the latest research in self-supervised learning, as well as adapt existing methodologies to the unique challenges posed by medical imaging data.

Example datasets you can refer: -

Brain Tumor Segmentation (BraTS) dataset (https://www.kaggle.com/code/rastislav/3d-mri-brain-tumor-segmentation-u-net/input) Different versions of BraTS are available on the internet.

Breast Ultrasound Images Dataset (https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset)

COVID-19 CT scan lesion segmentation dataset (https://www.kaggle.com/datasets/maedemaftouni/covid19-ct-scan-lesion-segmentation-dataset?select=frames)

Left Atrial Segmentation (https://www.kaggle.com/code/brsdincer/left-atrial-segmentation-all-process-nii-doc/input?select=imagesTr)

Liver Tumor Segmentation (https://www.kaggle.com/datasets/andrewmvd/liver-tumor-segmentation/data?select=volume_pt1)

These are some of the datasets you can refer (However, people using a dataset of NifTi/Dicom/Tif files will be given more priority than normal segmentations in jpeg/png images)