Title: Transformers Make Strong Encoders for Medical Image Segmentation

Abstract: The process of extracting areas of interest (ROIs) from 3D image data, such as that from CT or MRI scans, is known as medical image segmentation. The major objective of segmenting this data is to locate the anatomical regions needed for a given study, such as diagnosing and planning treatment of a disease. The u-shaped architecture, often known as U-Net, has achieved great success and de facto status for many segmentation applications. However, U-Net often displays limits in clearly representing long-range reliance because of the inherent locality of



convolution processes. Transformers, built for sequence-to-sequence prediction, have arisen as alternative architectures with built-in global selfattention mechanisms, however they may have limited localization capabilities because they don't have enough low-level details. Therefore, this study will investigate the combined effects of u-net architectures and transforms for image segmentation tasks.

Data Availability: Data will be shared with the students who wish to conduct research in this domain.