ACS 2024 Surgeons and Engineers: A Dialogue on Surgical Simulation Meeting

O-08

Research Abstracts

AdaptiveSAM: Towards Efficient Tuning of SAM for Surgical Scene Segmentation

Jay Nitin Paranjape; Nithin Gopalakrishnan Nair; Shameema Sikder, MD; S. Swaroop Vedula; and Vishal M. Patel

Johns Hopkins University, Baltimore, MD

Introduction: The task of segmentation in surgical scene analysis using AI is of paramount importance. However, the inherent data scarcity within this domain presents a formidable challenge in adapting traditional segmentation techniques. To address this challenge, contemporary research endeavors involve the utilization of pre-trained models, which are subsequently fine-tuned using the available data. Nonetheless, this approach necessitates the recurrent training of deep neural networks, each containing millions of parameters, whenever new data is introduced. A recent foundational model known as Segment-Anything (SAM) has demonstrated impressive generalization capabilities across a wide array of natural images, offering a notable degree of resolution to this challenge. However, SAM's applicability within the medical domain is hindered by its inability to generalize effectively without substantial computational resources for fine-tuning, along with the need for task-specific prompts. These prompts are often in the form of bounding boxes or foreground/background points, mandating explicit annotation for every image. As data volume increases, this solution becomes progressively laborious.

Methods: In this research, we introduce AdaptiveSAM—an adaptive refinement of SAM—designed to efficiently assimilate new datasets while enabling segmentation guided by textual prompts. To facilitate the fine-tuning of AdaptiveSAM, we present a technique named bias-tuning, which remarkably reduces the number of trainable parameters in comparison to SAM (less than 2%), while significantly improving performance. Notably, AdaptiveSAM's integration does not require expert intervention, as it leverages free-form textual prompts, enabling precise object segmentation using merely the label name as guidance.

Results: We observe around 60% improvement over SAM as well as significant improvements over SOTA segmentation methods.

Conclusions: Our empirical investigations validate the superiority of AdaptiveSAM over contemporary state-of-the-art methodologies across diverse medical imaging datasets encompassing surgical scenarios, ultrasounds, and X-rays. This innovation holds the potential to significantly enhance segmentation performance and curtail expert involvement while adapting SAM for niche applications.