Comprehensive Benchmarking Large Vision Language Models (LVLM) in Oral Cancer Classification and Segmentation

Project Description:

Recently, Large Vision-Language Models (LVLMs) such as Flamingo and BLIP2 have exhibited remarkable advancements across various domains, including embodied AI, autonomous driving, and remote sensing. Encouraged by their achievements, a growing number of LVLMs tailored for medical applications have emerged, claiming impressive performance across a wide spectrum of medical challenges. However, despite the growing attention these models have obtained, there has been a noticeable lack of comprehensive evaluations, particularly when it comes to real medical images, which strongly hinders a thorough understanding of their applicability and performance in medical contexts. Therefore, in the project we aim to evaluate the performance of state-of-the-art LVLMs in detecting, classifying, and segmenting oral cancer images in standard fashion.

Duties/Activities:

The Intern will literature review in order to understand the scope of the project. Also, he/she will work existing LVLMs along with our collected datasets. Then, each intern will work on a specific task such as prompting LVLM, evaluation metrics, etc.

Required Skills:

Python programming skills is required. Understand the basics of deep leaning models and previous experience in dealing with LLMs is a plus.

Learning Opportunities:
The intern will have opportunity to work with scientists and software engineers on promising research problems. In addition, it is great opportunity to participate in solving real-world healthcare problems.

Expected Team Size: *2-3*

Mentors

Name: Dr. Mohamed Elshrif   email: melshrif@hbku.edu.qa