OralScan: A Multi-Stage Multi-Modal Architecture for Early Diagnosis of Oral Cancer and Potentially Malignant Disorders via Intra-Oral Images Captured on a Mobile App

Pai, Ronak (School: Bentonville West High School)
Unnam, Veera Sai Joshik (School: Bentonville West High School)

Oral cancer affects over 300,000 people worldwide each year and has a mortality rate of approximately 47%. Early diagnosis can significantly improve survival rates by up to 56%, but less than 33% of cases are diagnosed early due to limitations in cost, accessibility, and accuracy of current diagnosis methods. An early sign of oral cancer is abnormal lesions in the oral cavity, which can be captured using intraoral images. These images are affordable and accessible but contain excessive background noise, making diagnosis difficult. To address this, we developed OralScan, a 4-stage, multi-modal computer vision pipeline aimed at improving oral cancer diagnosis. The pipeline, trained on a dataset of over 6,000 intraoral images, segments the oral cavity and suspicious lesions with F1-scores of 97% and 93% respectively. During the classification stages, our model incorporates patient risk factors such as age, sex, and lifestyle habits into the final diagnosis, which are processed using a fully connected layer. The pipeline's multi-modal classifiers identify lesions as malignant or benign with a 91% F1-score and further classify cancer as early or late stage with an 84% F1-score. We have developed a cross-platform mobile app that ensures affordability and ease of use. To improve model accuracy, we designed a 3D-printable alignment chassis that attaches to a phone, ensuring that patients take well-lit, properly framed intraoral images. By integrating multimodal data, our pipeline achieved unprecedented segmentation and classification accuracy. OralScan's affordability makes it an accessible diagnostic tool for those in need of immediate oral care worldwide.

Awards Won: