**IHUB Data IIIT HYD**

**AI-enabled Oral Cancer screening solution**

The focus of IHub-Data has been on putting together large-scale datasets to help develop solutions through applied research. One of the first initiatives in the healthcare domain of IHub-Data was on early screening and diagnosis of cancer based on image.

Oral Cancer is one of the most prevalent cancers afflicting Indians today according to the 2020 Report of the National Cancer registry program and its prognosis remains poor because most patients are diagnosed at advanced stages. About 70% of cancer afflicted appear for treatment in terminal stages, especially in rural areas. The AI enabled phone-image based solution, arising out of a collaborative effort with [Grace Cancer Foundation](https://www.gracecancerfoundation.org/) and [BioCon Foundation](https://www.bioconfoundation.org/%22%20%5Ct%20%22_blank), requires just a smart phone that flags abnormal oral lesions before they progress into advanced stages of cancer.



The NGO has been attempting an early screening and detection of cancers via its mobile camps. Sending a clinician along with a bus for every visit to remote locations is not always a viable option, due to irregular availability of highly skilled oncologists who can examine oral cavities visually or interpret x-ray test results instantaneously on the bus. Sometimes, redirecting patients with suspicious-looking lesions to the district or tertiary hospital for follow-ups, also induces delay in detecting the cancer. Therefore, we have brainstormed to arrive at a method where we could set up an AI based solution on bus that could screen oral cavity images taken by smart phone cameras by trained but semi-skilled health professionals, and flag them as malignant or benign and raise an alert immediately.



1. suspicious oral cavity image b. non-suspicious oral cavity image

Apart from the photographs of the oral cavity, our model could also take into consideration lifestyle, family and medical history, vitals and blood tests, clinical examination, as well as lifestyle information to improve the predictive capability of the AI algorithm.

The current model has been trained on Biocon Foundation dataset that contains 2176 oral cavity images. The weighted average F1-score is 0.85. Our model will be retrained on Grace Cancer Foundation dataset that contains more than 3000 oral cavity images that we have collected as part of our initiative using smart phones. The next plan would be to test on independent cohorts to prove the generalizability of the model.