Title: Oral Cancer Diagnosis Using Histopathology Images: An Explainable Hybrid Transformer Framework

Author names: Shahriar Sultan Ramit, Nayeem Ahmed, Md. Fatin Ishrak, Md. Ruhul Amin, Alaya Parven Alo & Md. Sadekur Rahman

Abastact: Oral Squamous Cell Carcinoma (OSCC) is among the most frequent cancer death causes, and early detection plays a vital role in improving patient survival. The traditional histopathological diagnosis is subjective and labor-intensive which necessitates an automated and standardized classification methods. This study has used a publicly available dataset comprising a total of 10,000 histopathological images. This paper evaluates four CNN architectures ResNet101, InceptionV3, MobileNetV2 and Xception to classify OSCC and normal cells. For better accuracy Hyperparameter Tuning was done on MobileNetV2. Tuned MobileNetV2 achieved the best performance with accuracy, recall, and F1-score of 0.99, demonstrating its efficacy in classifying malignant versus normal tissues. To further enhance interpretability Explainable AI techniques were employed, including LIME and Saliency Maps, enabling visual comprehension of model predictions. Our results demonstrate the importance of deep learning for OSCC detection overcoming the "black-box" issue of CNNs by explain ability. This study contributes to AI-driven diagnostic innovation through a more accurate and interpretable approach to OSCC classification.

Keywords: oral cancer; attention mechanism; swin transformer; explainable AI; vision transformer; histopathology image

https://doi.org/10.3390/technologies14010039

https://www.mdpi.com/2227-7080/14/1/39