

# Classification of Oral Cancer into Precancerous Stages from White images Using Light GBM Algorithm

Ganesana Sreeja, Yodupati Hemalatha, Botlagunta Vijitha, Ganedi Harsha Vardhan, Ragipati Chandu

Department of Electronics and Communication Engineering, PBR Visvodaya Institute of Technology and Science, Kavali, India

Guide: Mr. M.Sreehari

\*\*\*

**Abstract** - Oral cancer is a significant global health concern, requiring precise and early diagnosis to improve treatment outcomes. This project enhances oral cancer classification accuracy using CNNs and K-means clustering-based image segmentation. The methodology involves preprocessing images, segmenting using K-means, and classifying into leukoplakia, erythroleukoplakia, and erythroplakias. Traditional approaches rely on visual examination and biopsy, which may lead to delays and misdiagnosis. The proposed system uses image processing and deep learning techniques to classify oral cancer into precancerous stages. The system achieves reliable and accurate classification, assisting medical professionals in early diagnosis and improving patient care.

## 1. INTRODUCTION

Oral cancer is a significant global health issue, especially in developing countries like India, often linked to tobacco use, alcohol consumption, and poor oral hygiene. Early detection is crucial for successful treatment and reducing mortality rates. Precancerous conditions like leukoplakia, erythroplakia, and erythroleukoplakia are early indicators, but identifying them is challenging with traditional methods. This project uses image processing and machine learning, specifically CNNs, to analyze medical images and detect patterns not easily visible to the human eye, improving diagnosis accuracy. The motivation stems from rising oral cancer cases and the need for automated diagnostic systems.

## 2. Body of Paper

### 2.1 Literature Survey

Oral cancer is a significant global health issue, linked to tobacco use and poor oral hygiene. Early detection is crucial for successful treatment. Precancerous conditions

like leukoplakia and erythroplakia are early indicators. Epigenetic changes drive carcinogenesis. Oral cancer often goes undetected until late stages. OSCC comprises 90% of oral cancers, starting asymptotically. Advanced diagnostic methods are being developed. Research focuses on improving quality of life and treatment outcomes. Early detection can significantly improve survival rates. Treatment involves surgery, radiation, and chemotherapy

### 2.2 Existing Method

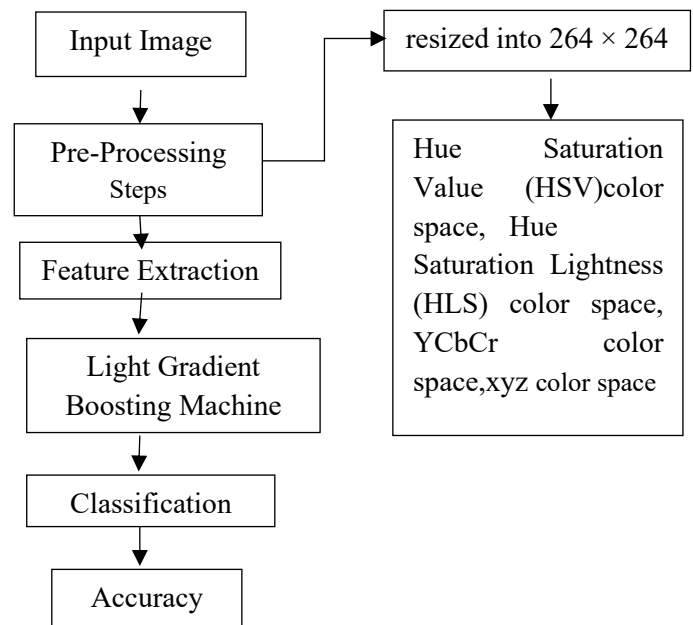


Figure 1: Existing Method Block Diagram

Existing methods for oral cancer classification have limitations like single-color space reliance and high complexity. The proposed method uses multiple color spaces and extracts texture features with GLCM and GLRLM.

It employs LightGBM for efficient classification. This improves accuracy and supports early detection. It's more robust for medical image analysis. Early identification of precancerous conditions helps in timely treatment. The system aims to reduce disease progression risk effectively.

### 2.3 Proposed Method

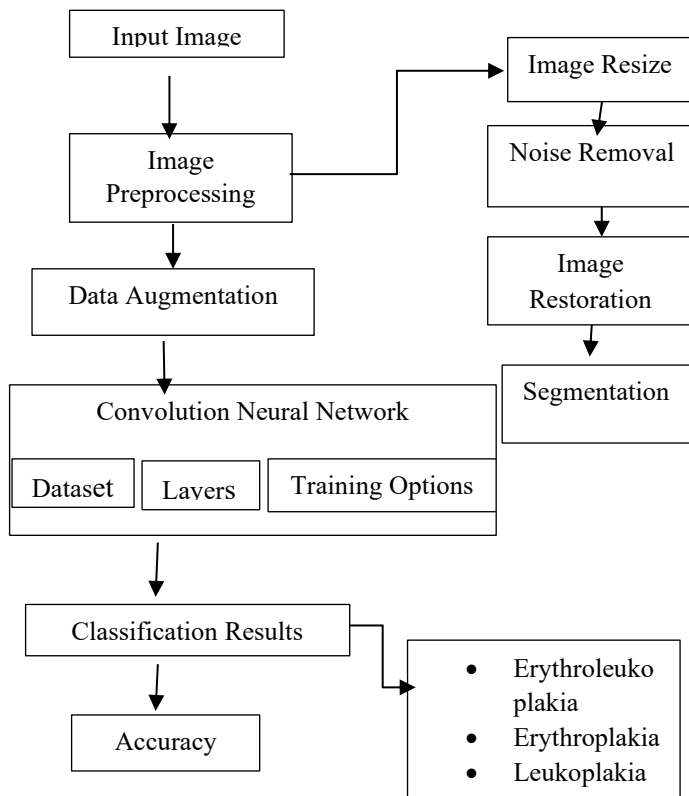


Figure 2: Block Diagram of Proposed Model

The proposed methodology uses CNNs with K-means clustering for oral cancer classification. Images are pre-processed and segmented into regions. CNN extracts spatial features and classifies into three categories. Data augmentation improves generalization. The system aims for early and effective diagnosis.

### 2.4 Implementation and Results



Figure 3:Input image

The images show processing steps for oral cancer detection, starting with a gum lesion input.



Figure 4:Resized RGB image



Figure 5:Noise removed image



Figure 6:Labeled image

The image undergoes preprocessing steps: resizing to 264×264 pixels for uniformity and reduced processing time, noise removal using filters like Median or Gaussian to improve clarity, and labeling to differentiate healthy and abnormal regions for model training.

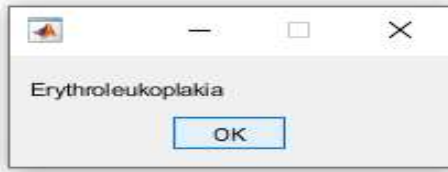


Figure 7: Classification Results Image

Table 1: Accuracy Comparison

S.No	Existing Method	Proposed Method
1	85.71	91.09
2	78.21	90.09
3	82.21	92.08
4	78.21	91.99
5	78.21	90.58

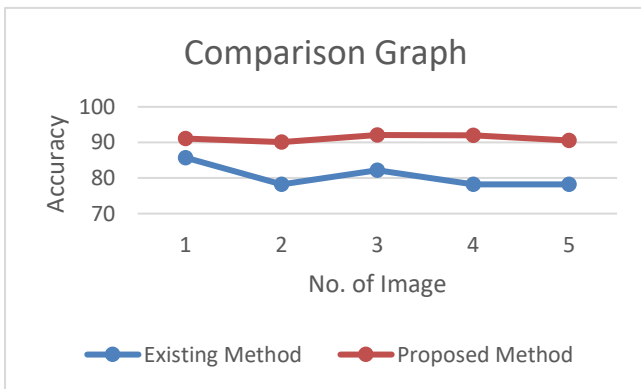


Figure 8: Accuracy Comparison Graph

The proposed model shows higher accuracy in oral cancer classification, outperforming existing techniques. It demonstrates improved effectiveness in identifying cancer stages. The model's better performance is clear. It enhances oral cancer detection capabilities.

### 2.5 Advantages and Applications

The proposed system has advantages like early oral cancer detection, high accuracy using CNN and LightGBM, automated diagnosis, and time efficiency. It's cost-effective for rural areas and scalable for large datasets. However, it depends on image quality, requires large datasets, and has computational complexity. The

system can be used in medical diagnosis, telemedicine, and training medical professionals. It's useful for rural healthcare and large-scale screening programs.

### 3. CONCLUSIONS

This project uses CNNs and K-means clustering to improve oral cancer classification accuracy. It preprocesses images, segments regions, and classifies them into leukoplakia, erythroleukoplakia, and erythroplakia. The system shows improved performance over traditional methods and supports early detection, improving survival rates. It's scalable and adaptable for clinical use, making it practical for areas with limited healthcare access. The approach demonstrates AI's role in transforming medical diagnostics and patient care, reducing human dependency and diagnostic errors. It enables timely medical intervention, preventing disease progression. Future enhancements can integrate it into telemedicine and mobile platforms. This solution contributes to intelligent healthcare systems and better patient outcomes.

### ACKNOWLEDGEMENT

We would like to express their sincere thanks to their project guide, Mr. MSeehari, for his valuable guidance and support throughout the project work. We also thank the Department of Electronics and Communication Engineering, PBR Visvodaya Institute of Technology and Science, Kavali, for providing the necessary facilities to complete this project successfully.

### REFERENCES

- [1] S. Sarkar, G. Horn, K. Moulton, A. Oza, S. Byler, S. Kokolus, and M. Longacre, "Cancer development, progression, and therapy: An epigenetic overview," *Int. J. Mol. Sci.*, vol. 14, no. 10, pp. 21087–21113, Oct. 2013.
- [2] S. Warnakulasuriya, N. W. Johnson, and I. Van Der Waal, "Nomenclature and classification of potentially malignant disorders of the oral mucosa," *J. Oral Pathol. Med.*, vol. 36, no. 10, pp. 575–580, Nov. 2007.
- [3] B. W. Neville and T. A. Day, "Oral cancer and precancerous lesions," *CA, A Cancer J. Clinicians*, vol. 52, no. 4, pp. 195–215, Jul. 2002.

- [4] I. van der Waal, R. de Bree, R. Brakenhoff, and J. Coebergh, “Earlydiagnosis in primary oral cancer: Is it possible?” *Medicina Oral Patología Oral y CirugiaBucal*, vol. 16, no. 3, pp. e300–e305, 2011.
- [5] A. Humeau-Heurtier, “Texture feature extraction methods: A survey,” *IEEE Access*, vol. 7, pp. 8975–9000, 2019.