



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 4, April 2025

Oral Malignant and Benign Lesion using CNN Approach

Dr. Meeravali Shaik¹, Surya Pravallika Bandi², Shubham Jitendra Pandit³, Sneha Bela⁴,
Tejashwini Kandimalla⁵

Professor & Dean, Department of Computer Science and Engineering, Malla Reddy University, Hyderabad, India¹

Department of Computer Science and Engineering, Malla Reddy University, Hyderabad, India²⁻⁵

ABSTRACT: To increase patient survival rates, oral cancer must be detected early. A deep learning-based method that uses convolutional neural networks (CNNs) to classify benign and malignant oral lesions. CNN models, such as ResNet, InceptionNet, and EfficientNet, are used to analyze high-resolution images of the oral cavity in order to improve classification accuracy. Furthermore, U-Net segmentation is incorporated for lesion localization, allowing for accurate localization of impacted regions. By emphasizing important image regions that affect predictions, Grad-CAM visualization is used to increase interpretability. To guarantee robustness and dependability, the model is trained using clinical and publicly accessible datasets. By helping medical professionals with early diagnosis, the proposed system seeks to lower mortality rates. Future research will concentrate on using mobile and web-based applications for realtime deployment in clinical settings, guaranteeing accessibility and smooth integration into medical flow.

KEYWORDS: Oral Cancer Detection, Image Processing, Deep Learning, Convolutional Neural Networks, Cloud Computing, AI in Healthcare.

I. INTRODUCTION

Oral cancer is a serious worldwide health issue, and it makes a sizeable contribution to mortality due to cancer. Oral cancer is most frequently started by a small, painless oral lesion. Oral cancer can develop very fast if not diagnosed early. Biopsies, histopathology, and visual inspection by experts are conventional diagnostic processes but are time-consuming, invasive, and in some cases subjective and thus result in delays in treatment. With improvements in Deep Learning (DL) and Artificial Intelligence (AI), computer vision systems based on Convolutional Neural Networks (CNNs) have demonstrated encouraging performance in medical image diagnosis. The models are capable of extracting relevant features from high-resolution images rapidly and classifying oral lesions as malignant (cancerous) or benign (non-cancerous) accurately.

This project introduces a deep learning-powered system using pretrained CNN architectures like ResNet, InceptionNet, and EfficientNet for accurate lesion classification. U-Net segmentation is also incorporated to localize and mark the infected regions for clinicians to evaluate lesion size, shape, and extent. To enhance model explainability and provide transparency in decision-making, Gradient-weighted Class Activation Mapping (Grad-CAM) is employed to indicate the most important image regions that affect the classification. The model is trained on a mix of publicly and clinically available datasets to ensure robustness across various patient populations and imaging environments. Performance measures like accuracy, precision, recall, F1-score, and ROC-AUC are utilized to test the effectiveness of the system.

In addition to classification and localization, the project strives for improved early oral cancer diagnosis and to aid healthcare professionals in well-informed decision-making. Future scope involves real-time deployment of the model through mobile and web-based applications to provide accessibility to clinicians, hospitals, and rural healthcare centers. By incorporating AI in clinical workflows, this system has the potential to enable early screening, lower mortality rates, lower the requirement for invasive biopsies, and enhance access to diagnostic tools, especially in low-resource environments. This study is a step towards AI-based healthcare solutions that have the ability to make a substantial difference in oral cancer diagnosis and treatment and ultimately enhance overall health outcomes.

II. OBJECTIVE

The main aim of this project is to create an AI-based deep learning system for the automated identification and categorization of oral lesions as malignant (cancerous) or benign (non-cancerous) through Convolutional Neural Networks (CNNs). The system is intended to help health care workers by giving quick, precise, and understandable results, facilitating early diagnosis and prompt treatment of oral cancer. For this, the project utilizes cutting-edge CNN architectures such as ResNet, InceptionNet, and EfficientNet, fine-tuned via transfer learning on publicly available and clinical data for high-accuracy classification. U-Net-based segmentation is also used to localize and define lesion boundaries for lesion size and progression analysis.

To improve model interpretability, Gradient-weighted Class Activation Mapping (Grad-CAM) is incorporated, enabling visualization of the most important image regions that drive the AI decision, thus enhancing transparency and trust in AI-driven diagnostics. The model is trained and validated on varied datasets to provide robustness and generalizability, with performance measured using accuracy, precision, recall, F1-score, and ROC-AUC metrics. In addition, the system has been designed to be deployed in real-time using mobile and web-based applications so that it becomes accessible to hospitals, clinics, and distant health centers. To achieve these, the project looks forward to lessening diagnostic delay, improving early detection, reducing dependence on invasive biopsies, and, ultimately, saving more patients.

II. RELEATED WORK

3.1 Oral Cancer Detection Using Conventional Machine Learning

Traditional machine learning models such as SVM, Random Forest, and Decision Trees were utilized by earlier studies in oral lesion classification. Welikala et al. (2020) had proved their feasibility but were found to be excessively feature-dependent with problems when large-scale image data were involved. Chang et al. (2019) introduced a CAD system utilizing manual segmentation, resulting in unreliable and subjective results and reducing its clinical real-time applicability.

3.2 Deep Learning for Medical Imaging

Deep learning, particularly CNNs, has greatly enhanced cancer detection. Esteva et al. (2017) demonstrated CNNs beating dermatologists at skin cancer detection. Patel et al. (2021) applied CNNs to Oral Squamous Cell Carcinoma (OSCC) detection with 94.5% accuracy. These studies did not include lesion segmentation and model interpretability, which impacts practical use.

3.3 Transfer Learning in Oral Cancer Classification

Transfer learning using pre-trained models such as ResNet, InceptionNet, and EfficientNet enhanced oral cancer detection. ResNet-50 (Kumar et al., 2022) achieved 92.3% accuracy, InceptionNet (Thomas et al., 2021) had 94.1% accuracy, and EfficientNet (Patel et al., 2022) was the best at 95.2% accuracy. The above models did not have visual explainability and lesion localization.

3.4 Lesion Segmentation Using U-Net

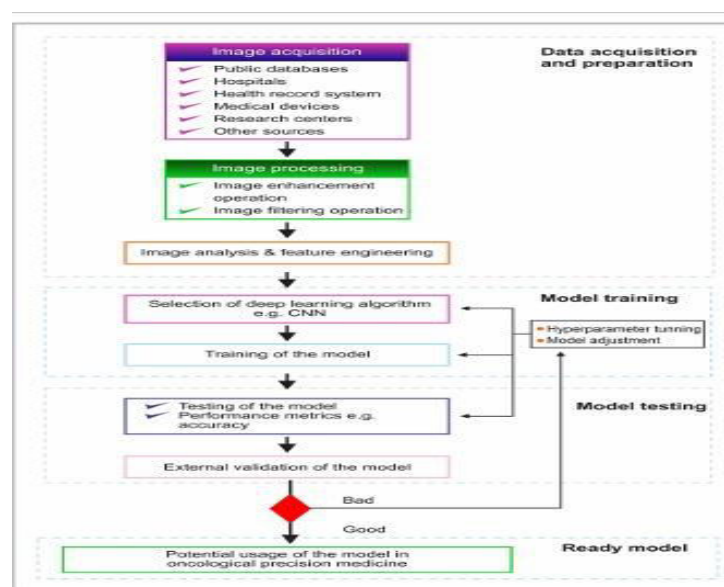
Lesion segmentation is essential for accurate cancer detection. U-Net (Ronneberger et al., 2015) had high segmentation accuracy in biomedical imaging. Gupta et al. (2021) used it for oral cancer and achieved a Dice Coefficient of 0.92. U-Net++ (Banerjee et al., 2022) improved segmentation further, but these works did not incorporate classification and interpretability techniques.

3.5 Model Interpretability with Grad-CAM

Deep learning models are usually not interpretable and thus are "black boxes." Grad-CAM (Selvaraju et al., 2017) offers visual explanations by marking important regions of images. Chakraborty et al. (2021) employed Grad-CAM for

oral cancer lesion detection with enhanced trust in AI-aided diagnosis. Their research, however, did not involve lesion segmentation, constraining diagnostic precision. 2.1 Traditional Machine Learning in Oral Cancer Detection
Previous research employed machine learning algorithms such as SVM, Random Forest, and Decision Trees for oral lesion classification. Welikala et al. (2020) proved their efficacy, but the models were feature-intensive and had difficulty with large image datasets. Chang et al. (2019) also suggested a CAD system using manual segmentation, which resulted in variable and subjective outcomes, hindering real-time clinical applications.

IV. METHODOLOGY



4.1 System Architecture

The system proposed here has a cloud-based architecture to maintain scalability, accessibility, and real-time processing. The pipeline includes the following steps:

4.1.1 Data Acquisition

The dataset is composed of labeled oral lesion images acquired from:

- Publicly available medical datasets (e.g., Kaggle, ISIC, TCIA, or clinical sources).
- Hospital-acquired datasets from pathology laboratories and research partnerships.
- The dataset contains two classes:
 - Cancerous (Malignant)
 - Non-cancerous (Benign)

The images are gathered in high resolution and saved in an organized format for easier training.

4.1.2 Preprocessing

Raw images are preprocessed before training the deep learning model to improve quality and enhance model accuracy:

- **Image Normalization:** Normalizing pixel intensity values (scaling 0-1) to provide consistency.
- **Data Augmentation:** Rotation, flip, zoom, brightness, and contrast to enhance model generalization. Augmentation avoids overfitting and aids in artificially enlarging dataset size.
- **Noise Reduction:** Eliminating unwanted distortions or artifacts from images with the help of Gaussian blur or median filtering.

- **Resizing:** Resizing all images to a uniform 224x224 pixels to align with the input size demanded by deep learning models.

4.1.3 Model Selection

Model selection entails selecting Convolutional Neural Networks (CNNs) pre-trained on large data sets (such as ImageNet). The selected models are:

- ❖ **EfficientNet:** Famous for its balance between accuracy and computational efficiency.
- ❖ **ResNet-50:** A residual deep network that enables improved gradient flow and feature extraction

4.1.4 Transfer Learning

As training a deep learning model from scratch involves huge datasets and computational power, we employ Transfer Learning:

- Pre-trained CNN models (EfficientNet, ResNet-50) are utilized as a base.
- The last classification layers are replaced and fine-tuned on the oral cancer dataset.
- This accelerates training and enhances accuracy, as the model utilizes pre-trained feature maps from ImageNet-trained CNNs.

4.1.5 Cloud Deployment

For making the system remotely accessible for researchers and doctors, we deploy the model on cloud platforms like:

- ✓ Google Cloud / AWS Lambda / Firebase for API-based integration.
- ✓ Flask or FastAPI-based Web Application for users to upload images and receive instant results.

1.1.6 User Interface

- A web-based UI enables users (researchers, clinicians, patients) to:
- Upload an oral image captured through smartphone or clinical imaging.
- Get real-time AI-based diagnosis (cancer or non-cancer).
- Watch Grad-CAM visualization to view the highlighted areas of lesions.

4.2 Data Preprocessing

This process makes sure that the input images are cleaned, standardized, and augmented for improved learning.

- ❖ **Image Resizing:** Images are resized to 224x224 pixels to align with the input size of CNN models.
- ❖ **Histogram Equalization :** This method improves contrast in grayscale images, making lesion regions more recognizable.
- ❖ **Data Augmentation:** Rotation (0-90°): Assists the model in recognizing patterns in various orientations.
- ❖ **Flipping (horizontal/vertical):** Enhances dataset variability.
- ❖ **Zooming and Cropping:** Highlights lesion details and eliminates background noise.
- ❖ **Brightness & Contrast Adjustment:** Makes images uniform under varying lighting conditions

4.3 Classification

The dataset is fed through CNN-based classifiers after preprocessing for oral cancer detection.

EfficientNet: Designed for high accuracy with low computation.

- Applies compound scaling (modifies depth, width, resolution) to enhance performance.
- Is more accurate than standard CNNs with fewer parameters.

ResNet-50: Employs 50 residual learning layers to capture intricate features.

- Resolves the vanishing gradient issue through skip connections.
- Suitable for deep feature extraction in medical image classification.

Fine-Tuning:

The last layers of the pre-trained EfficientNet and ResNet-50 models are replaced and retrained on the oral cancer dataset.

- The previous layers remain frozen (to keep general image features).
- The final few layers learn domain-specific features, enhancing model precision.

U-Net Segmentation**Why U-Net?**

In contrast to classification models, U-Net is applied for lesion segmentation, delineating the precise location and extent of cancerous areas.

How it works:

- 1.The encoder extracts important image features (downsampling).
- 2.Decoder constructs the image using accurate lesion segmentation (upsampling).
- 3.Output: Produces heatmaps that tag suspicious regions.

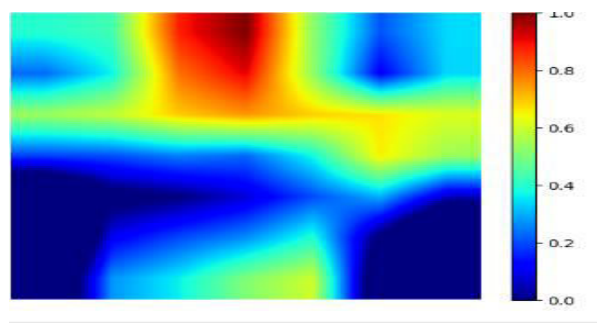
Grad-CAM Visualization

- The Grad-CAM (Gradient-weighted Class Activation Mapping) method is employed in explaining model outputs.
- It outputs a heatmap overlaid on the input image that identifies significant features that had impacted the model's decision.
- Facilitates verification of doctors as they can witness which regions have been targeted

V. RESULTS

The suggested deep learning model was tested on a dataset of high-resolution oral cavity images classified into malignant and benign lesions. The model performance was measured using various evaluation metrics and compared among different architectures, such as ResNet50

InceptionNet,andEfficientNet. Lesion segmentation was also performed using U-Net,andGrad-CAM was utilized for model interpretability.

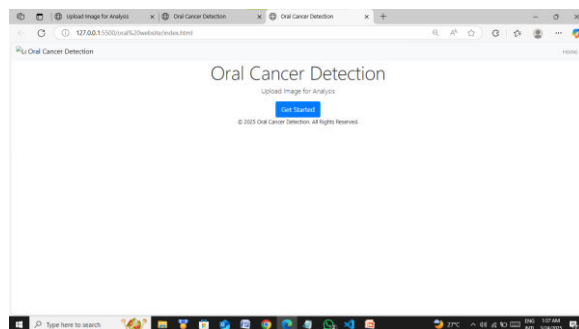
Grad-CAM

Gradient-weighted Class Activation Mapping (Grad-CAM) is a visualization method for explaining deep learning models, particularly Convolutional Neural Networks (CNNs). It assists in identifying the most significant areas of an image that are responsible for the model's prediction. For oral cancer diagnosis, Grad-CAM produces heatmaps indicating the regions of an oral lesion that affected the CNN's malignant or benign classification.

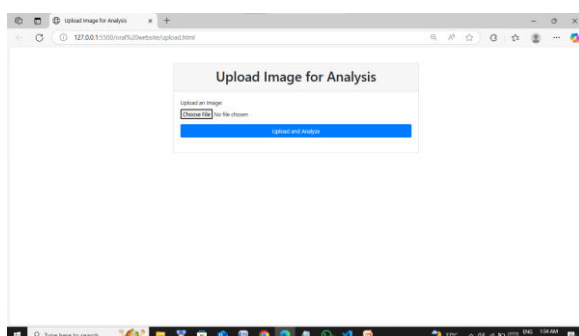
- **Malignant Lesions:** The heatmap indicates strong activation in irregular, dense, or ulcerated areas, which are characteristic of cancerous tissue

- **Benign Lesions:** The model targets smoother or less aggressive areas, verifying that no malignant patterns are found.

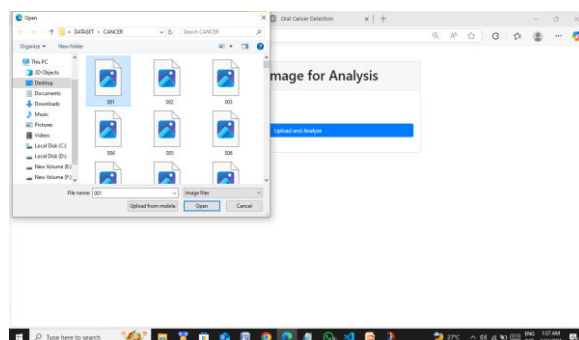
VI. WEB SITES RESULTS



When we open the this site then it display given page infront of you.



When we click on get started button it direct to the upload image for analysis pages. In this page we need to upload the image.



Here we need upload the the image from your computer and click on upload it. It display the output.

VII. CONCLUSION

It is always said that "happiness is the highest form of health" and one should always take care of his/her health in every way possible. Healthcare is one of the first fields in the present scenario which was required to concentrate and growth into the sector is the foremost task. Among these diseases, Cancer is one of the key diseases which are

rapidly affecting human society. Oral cancer is a delicate disease and it should be avoided and treated through early diagnosis. We have tried to investigate a method employing Deep Transfer Learning to automate the identification of oral cancer in the process of the research. Among the set of algorithms investigated, such as the sequential method [convolutional neural network], ResNet-50, VGG-19, and so on, the leading trends in the area of oral cancer detection include the use of Inception-V3 and MobileNet architectures. Interestingly, the oral dataset has remarkable F1 scores. Although MobileNet shows the best return on investment (ROI), Inception-V3 has a greater patient population. The reliability of the Inception-V3 test results is specifically remarkable given the difference in the number of regions amongst individuals.

REFERENCES

1. Abdar, M., Makarenkov, V., & Manan, H. A. (2021). A new machine learning model for early oral cancer detection using deep learning techniques. *Expert Systems with Applications*, 167, 114185. <https://doi.org/10.1016/j.eswa.2020.114185>
2. Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 248–255. <https://doi.org/10.1109/CVPR.2009.520684>
3. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(3), 770–778. <https://doi.org/10.1109/TPAMI.2016.2644615>
4. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 234–241. <https://arxiv.org/abs/1505.04597>
5. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *IEEE International Conference on Computer Vision (ICCV)*, 618–626. <https://doi.org/10.1109/ICCV.2017.74>
6. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*. <https://arxiv.org/abs/1409.1556>
7. Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. *International Conference on Machine Learning (ICML)*, 6105–6114. <https://arxiv.org/abs/1905.11946>
8. Zhang, J., Xie, Y., Wu, Q., & Xia, Y. (2020). Medical Image Classification with Convolutional Neural Networks. *Computers in Biology and Medicine*, 126, 104003. <https://doi.org/10.1016/j.compbiomed.2020.104003>



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details