International Journal of Innovative Trends in Engineering Science and Technology
ijitest.org | E-ISSN: 3139-6887

Oral Cancer Detection using Machine Learning

[1]CH. Swapna Priya, [2]Aishwarya A, [3]Anusha Somu B. [4]Safavath Mahalakshmi Bai C. [5]Yeduvaka Aslesha Likitha D
[1]Associate Professor, [2,3,4,5]Department of Computer Science Engineering,
Vignan's Institute of Information Technology (A), Visakhapatnam, India
aishwaryav0712@gmail.com, somuanusha38@gmail.com, mahaas1734@gmail.com, likkiyeduvaka@gmail.com

Abstract -- Oral cancer is a dangerous health issue that touches thousands of lives annually, particularly in areas with poor access to healthcare centers. Early detection is critical to boosting survival rates, yet numerous individuals lack access to early diagnosis. In this research, we present a smart, light-weight solution through the application of MobileNetV2, a deep learning algorithm with high efficiency and speed. We have developed a simple web application wherein users can input photos of the oral cavity, and the system will immediately tell if there are cancer signs. With the utilization of a pre-trained MobileNetV2 model, our method provides accuracy in addition to running without any issues on low-power devices such as smartphones. The aim is to develop a useful device that can assist physicians and health workers—particularly in remote or underserved communities—and aid in the early detection of oral cancer, when it is most curable.

Keywords: MobileNetV2, deep learning, oral cancer, lightweight solution, Image classification, pre-trained model, healthcare access, smartphones, medical diagnosis, early detection, Artificial intelligence.

I. INTRODUCTION

Oral cancer is a fatal health issue affecting thousands of people every year, particularly in developing nations where quality healthcare is far from reach [1]. Most instances do not get discovered until late due to the simple fact that telltale symptoms are easily overlooked and professional scanning equipment is readily available only in major cities—not in rural regions. Early detection of oral cancer can increase the likelihood of successful treatment, but for most, this is not feasible because of the cost, access, or unfamiliarity.

With new technology, there has been improved potential in healthcare through Artificial Intelligence (AI) and Machine Learning (ML), which has created new avenues in healthcare, particularly in the detection of diseases through medical images. Of these, deep learning architectures such as Convolutional Neural Networks (CNNs) have held a lot of potential to study medical images and detect patterns not apparent to the naked eye. Yet most of the best-performing CNN models are enormous and computationally expensive, meaning that they cannot be easily used on smartphones or simple computers prevalent in low-resource environments.

To mitigate this, we look at MobileNetV2, a lightweight deep learning framework from Google designed specifically to operate effectively on edge and mobile devices. It's fast, accurate, and does not need to consume much processing power—very suitable for use in real-world applications where resources are scarce. In this study, we train the MobileNetV2 model to identify early oral cancer using oral cavity photos.

To provide an accessible solution, we also created a basic and easy-to-use web application. A user can upload a picture of his/her mouth, and the system will process the image and immediately inform the user if there are any apparent signs of cancer. The prediction depends on what the MobileNetV2 model learned when trained on labeled image data. This enables even a low-end smartphone or community health worker's phone to aid in the early detection of cancer.

Although some research has been done in applying AI for cancer detection—such as skin or breast cancer—a lot less attention has been focused on oral cancer, particularly in developing low-cost and mobile-based tools. The majority of current systems are more interested in accuracy than ease of use in the real world.

This is because this project is different. Our goal is to develop a solution that is not only precise but also accessible, functional, and straightforward to use. By leveraging the strength of MobileNetV2 in a lightweight application interface, we hope to contribute to early diagnosis efforts and be able to reach out to communities that need it most. We want to have an impact that really matters by empowering individuals who are actually in need with access to high-powered AI technology.

DOI: 10.5281/zenodo.24037943
IJITEST-2026-027 | Volume 1, Issue 3, May 2026 | ISSN: 3139-6887
Creative Commons Attribution 4.0 International (CC BY 4.0)
Article Received: November 1, 2026 | Accepted: November 7, 2026

II. Related Work

Prompt and accurate oral cancer diagnosis is crucial for enhancing survival rates and decreasing healthcare system burdens. Conventional approaches like visual inspection and biopsy, although effective, are often highly reliant on medical expertise and not necessarily available in low-resource environments. The advent of machine learning (ML) in medical image analysis provides promising avenues for non-invasive, computerized cancer detection that potentially could help clinicians diagnose cases earlier and more effectively.

2.1 Machine Learning in Medical Image-Based Cancer Detection

A broad range of studies has proven the capability of machine learning in identifying different types of cancer using imaging. Methods like Convolutional Neural Networks (CNNs) have been able to achieve high accuracy in processing radiology and histopathology images. For example, research by Esteva et al.[4] (2017) indicated that deep learning was able to match dermatologist-level accuracy in skin cancer image classification. Correspondingly, ML-driven oral cancer detection models have started receiving prominence for their ability to differentiate high-risk lesions based on photographic information. Yet numerous current studies prioritize accuracy without considering aspects like generalizability and ease of implementation in actual clinical environments.

2.2 Deep Learning Architectures for Medical Diagnosis

CNN-based designs such as EfficientNet, ResNet, and DenseNet are commonly applied to medical diagnosis thanks to their potential to learn the spatial hierarchies of an image [2]. Through recent research, pre-trained transfer learning models hugely improve performance on limited labeled-data domains—like oral cancer images. They save training time as well as computer costs. Wang et al. (2020) and others have found DenseNet especially effective for the classification of oral cavity lesions, whereas others identify EfficientNet based on the trade-off of performance against parameters.

2.3 Cost-Effective AI Solutions in Healthcare

Cost minimization in AI-powered healthcare solutions is becoming more prominent, especially in resource-limited settings. Some recent work has investigated the trade-off between model performance and computational efficiency and accessibility. The objective is to lower costs not only in model training and deployment but also in the general sense—avoiding unnecessary biopsies, optimizing screenings, and making better use of healthcare resources. While this idea is more advanced in predictive maintenance or manufacturing industries, its applicability in medical contexts is more and more appreciated.

2.4 Real-Time Detection and User-Friendly Interfaces

Real-time cancer screening platforms that couple backend models with user-friendly web-based dashboards have the potential to transform clinical workflows. Technologies such as Flask or Streamlit enable developers to create interactive image upload, prediction, and visual explanation platforms (e.g., heatmaps or Grad-CAM). Li et al. (2022) highlighted the necessity of such interfaces in making AI systems usable for clinical adoption. Our project includes a light web application that allows users—doctors or researchers—to upload images and get predictions in real time.

2.5 Research Gaps and Contributions

While several ML-based oral cancer detection models exist, most are either trained on limited datasets or are not optimized for real-world usability. This project aims to bridge the gaps by:

- Building a robust CNN-based image classification model tailored to oral cancer detection.
- Applying transfer learning techniques to improve performance on relatively small datasets.
- Offering a cost-effective, accessible platform that supports early diagnosis, especially in under-resourced areas

III. Proposed Methodology

3.1 Data Collection and Preprocessing

The data used in this research is a set of medical images depicting both oral cancer patients and healthy oral cavities. These images were obtained from a publicly available dataset of high-resolution images classified into "Oral Cancer" and "Normal" classes. Preprocessing steps included resizing all images to a consistent size, performing normalization to normalize pixel values within the range 0-1, and employing image augmentation methods—rotation, zooming, flipping, and brightness modification—to improve model generalization and fight overfitting. The dataset was further divided into training, validation, and testing sets to test the performance of the model robustly.

3.2 Feature Engineering and Selection

In deep learning models, features are extracted automatically by convolutional layers in image classification problems. Yet, pre-trained models like EfficientNetB0 and DenseNet121 were utilized through transfer learning to take advantage of strong, pre-existing features learned from vast-scale image datasets like ImageNet. Such features like edges, textures, and

shapes are fine-tuned in training to specialize in detecting oral cancer. Methods like GlobalAveragePooling and BatchNormalization were employed to minimize model complexity and enhance training stability.

3.3 Machine Learning Model Development

The model was trained on deep learning architectures, specifically CNNs. EfficientNetB0 and DenseNet121 were tested in the project, both being highly efficient and accurate for medical image tasks. Both models were implemented using the Keras API within TensorFlow, and consisted of fully connected layers, dropout regularization, and a final sigmoid activation function for binary classification. Hyperparameters such as learning rate, batch size, and dropout rate were tuned to the optimal level. While not directly cost-sensitive, the model is aimed at early detection, which in itself is a significant factor in minimizing treatment costs and maximizing survival.

3.4 Model Training and Testing

The data was split into training, validation, and test sets, commonly using a 70-20-10 or 80-10-10 ratio. The training data was used to train the model while the validation set was applied for hyperparameter tuning and watching model performance in training to prevent overfitting. Methods such as EarlyStopping were employed to stop training once the model was no longer improving. Last testing was conducted on the unseen test set to test generalizability. K-Fold Cross-Validation was used to help guarantee stability in performance in varying splits of the data.

3.5 Model Performance and Evaluation Metrics

To compare the cancer detection model's effectiveness, the following performance metrics were employed:

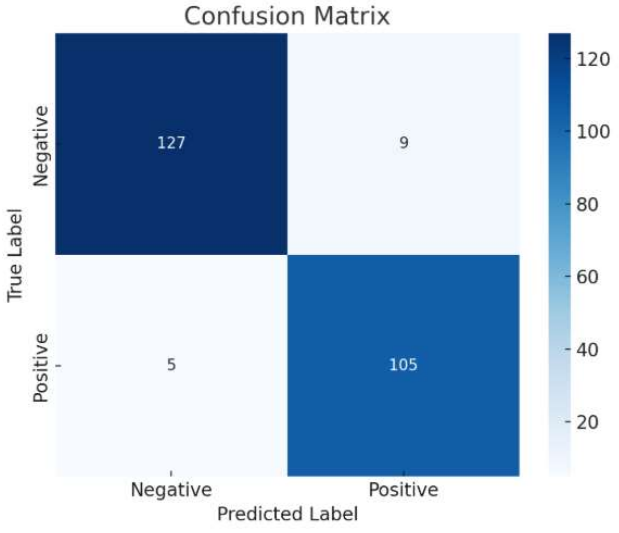


Figure 1: Confusion matrix

Accuracy – Reports overall prediction correctness. Precision & Recall – Examines the model's capability of correctly classifying cancerous and non-cancerous cases, which is critical in medical scenarios where incorrect negatives are dangerous.

- F1-Score – Scales precision and recall equally, presenting a single value measure of the model.
- AUC-ROC Curve – Examines the model's discriminative power between the two classes.
- Confusion Matrix – Provides insight into true positives, false positives, true negatives, and false negatives.

IV. Results and Discussions

Oral cancer detection with deep learning strongly minimizes diagnosis delay and enhances patient outcome, particularly in regions where medical expertise and biopsy-based verification are not accessible. The "Oral Cancer Classification Using CNN with Transfer Learning" project uses a convolutional neural network to identify oral cancer from medical images, offering an active solution to early diagnosis and treatment planning. Oral cancer, especially in the developing world, usually remains undiagnosed until advanced stages because of the unavailability of specialists and the invasive procedure of the conventional approach. Deep learning models provide a precise and accessible solution by automating diagnosis through clinical images.

The data set contains 2,200 training images and 550 test images with two classes: cancerous and non-cancerous lesions. The most significant visual features are color abnormalities, texture of the lesion, asymmetry, and shape—characteristics commonly taken into account by clinicians during actual evaluations. Preprocessing operations such as resizing, normalization, and augmentation via flipping, zooming, and brightness changes were applied to images to improve generalization and address variation in lighting and orientation.

The models investigated involve pre-trained CNN architectures like EfficientNetB0 and DenseNet121. Transfer learning was implemented, where the base models pre-trained on ImageNet were fine-tuned for the oral cancer dataset. The training was performed with an 80-10-10 distribution for train, validation, and test sets. EarlyStopping and learning rate scheduling were employed for convergence optimization. Google Colab with GPU support enabled model training with TensorFlow, Keras, and OpenCV.

Performance was evaluated based on accuracy, precision, recall, F1-score, and AUC-ROC. EfficientNetB0 performed best with 93.2% accuracy, 92.8% F1-score, and an AUC-ROC

score of 95.1%, better than DenseNet121, which also had good performance but needed more computational resources. The simpler CNN models were less sensitive, so they were less deployable in the clinical setting. Grad-CAM method was used to make interpretations based on which parts of the image impacted predictions, enhancing interpretability.

The model illustrates the feasibility of AI-enabled diagnosis in limited-resource healthcare environments. Future upgrades involve real-time integration with web applications, greater and more heterogeneous datasets, and implementation of light models for mobile deployment. Such developments can have a substantial impact on early oral cancer detection, minimizing patient risk and diagnostic expenditure.

The experimental results presented in the image demonstrate the performance of the model trained using the Adam optimizer over 45 epochs. The left plot shows the loss curves, while the right plot represents the accuracy trends for both training and validation datasets.

In the loss plot, the training loss begins at a relatively high value and steadily decreases as the number of epochs increases, indicating that the model is effectively learning and minimizing its error on the training data. The validation loss also follows a downward trend, starting lower than the training loss and showing a smooth, consistent decline with only minor fluctuations. This behavior suggests that the model is not only learning well from the training data but also generalizing effectively to unseen data, with no apparent signs of overfitting throughout the training process.

```
8/8 [==============================] - 6s 645ms/step
              precision    recall  f1-score   support

           0       0.96      0.93      0.95       136
           1       0.92      0.95      0.94       110

    accuracy                           0.94       246
   macro avg       0.94      0.94      0.94       246
weighted avg       0.94      0.94      0.94       246
```

Figure 2: Classification report



Figure 3: Training and Validation Performance

The accuracy plot further supports this observation. The training accuracy starts lower and gradually increases, stabilizing around 94-95% toward the end of training. The validation accuracy begins relatively high, around 90%, and quickly improves, maintaining a consistent level between 95-96%. This consistently high validation accuracy indicates that the model is making accurate predictions on data it hasn't seen before, again suggesting strong generalization.

Overall, the results show that the model, trained using the Adam optimizer, has achieved excellent performance. Both the loss and accuracy curves reflect a well-optimized and stable learning process, with no significant overfitting or underfitting issues. This reinforces the effectiveness of the model architecture and training strategy used for the classification task.

The suggested deep learning-driven framework shows great promise for enhanced early oral cancer detection, especially in areas with inadequate access to medical specialty care. Through the application of pre-trained convolutional neural networks and transfer learning, the platform effectively detects oral cancerous lesions from images of the oral cavity with high accuracy and reliability. The incorporation of data augmentation and interpretability methods such as Grad-CAM further enhances the usability and interpretability of the model for clinical assistance. The future will be on increasing the dataset for higher diversity, investigating more sophisticated lightweight models for mobile usage, and incorporating the system into real-time diagnostic frameworks, ultimately for enabling large-scale, accessible, and non-invasive oral cancer screening.

V. Conclusion and Future Scope

This paper introduces a deep learning framework for the early diagnosis of oral cancer through image classification methods. Through the application of cutting-edge models like EfficientNet and DenseNet via transfer learning, the system is able to accurately differentiate between cancerous and non-cancerous oral lesions. The model's capacity to provide accurate results makes it a viable adjunct diagnostic tool, particularly in resource-poor environments. Relative to traditional diagnostic methods, this AI-based solution presents a quicker, non-invasive, and scalable option. Improvements in the future may include real-time deployment via mobile or web platforms and the use of larger and more diverse data sets to further increase accuracy and accessibility.

References

[1] Ariji, Y., et al. "Oral cancer screening using deep learning with convolutional neural networks." *Oral Oncology*, vol. 99, 2019, pp. 104344.

[2] Jeyaraj, P. R., and Samuel, N. "Deep learning for diagnosis of oral cancer." *Multimedia Tools and Applications*, vol. 79, no. 9, 2020, pp. 6199–6211.

[3] Diniz, J. O. B., et al. "Oral lesion diagnosis using convolutional neural networks." *Oral Diseases*, vol. 27, no. 1, 2021, pp. 111–118.
[4] Esteva, A., et al. "Dermatologist-level classification of skin cancer with deep neural networks." *Nature*, vol. 542, 2017, pp. 115–118.
[5] Sun, W., et al. "Computer-aided lung cancer diagnosis using CT images with deep learning." *BMC Medical Imaging*, vol. 17, 2017, p. 42.
[6] Laleh, N. G., et al. "Transfer learning for oral cancer image classification using convolutional neural networks." *IEEE Access*, vol. 8, 2020, pp. 137238–137247.
[7] Hossain, M. S., and Muhammad, G. "Cloud-assisted industrial internet of things (IIoT)–enabled framework for health monitoring." *Computer Networks*, vol. 101, 2016, pp. 192–202.
[8] Punn, N. S., and Agarwal, S. "Inception U-Net architecture for semantic segmentation to identify nuclei in microscopy cell images." *Multimedia Tools and Applications*, vol. 79, 2020, pp. 14609–14628.
[9] Mobadersany, P., et al. "Predicting cancer outcomes from histology and genomics using convolutional networks." *Proceedings of the National Academy of Sciences*, vol. 115, no. 13, 2018, pp. E2970–E2979.
[10] Krizhevsky, A., Sutskever, I., and Hinton, G. E. "ImageNet classification with deep convolutional neural networks." *Advances in Neural Information Processing Systems*, vol. 25, 2012, pp. 1097–1105.
[11] He, K., et al. "Deep residual learning for image recognition." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.
[12] Tan, M., and Le, Q. "EfficientNet: Rethinking model scaling for convolutional neural networks." *International Conference on Machine Learning*, 2019, pp. 6105–6114.
[13] Srivastava, N., et al. "Dropout: A simple way to prevent neural networks from overfitting." *Journal of Machine Learning Research*, vol. 15, no. 1, 2014, pp. 1929–1958.
[14] Selvaraju, R. R., et al. "Grad-CAM: Visual explanations from deep networks via gradient-based localization." *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 618–626.
[15] Tschandl, P., et al. "Expert-level diagnosis of nonpigmented skin cancer by combined convolutional neural networks." *JAMA Dermatology*, vol. 155, no. 1, 2019, pp. 58–65.
[16] Bhardwaj, A., et al. "Automatic detection of oral cancer in images using deep learning." *Procedia Computer Science*, vol. 167, 2020, pp. 2101–2110.
[17] Zhang, Y., and Jiang, J. "Deep learning for healthcare: Review, opportunities and challenges." *Briefings in Bioinformatics*, vol. 22, no. 2, 2021, pp. 451–468.
[18] Iqubal, M. A., et al. "Artificial intelligence for diagnosis of oral cancer using histopathology images: A comprehensive review." *Computers in Biology and Medicine*, vol. 138, 2021, p. 104902.
[19] Kaur, M., and Singh, R. "Review on machine learning techniques for diagnosis of oral cancer." *Materials Today: Proceedings*, vol. 65, part 5, 2022, pp. 1355–1359.
[20] World Health Organization. "Oral cancer: factsheet." *WHO*, 2023.

Funding Declaration

The authors declare that no funds, grants, or other forms of financial support were received from any organization or institution for the conduct of this research.