FishTaxa - Journal of Fish Taxonomy
Vol 36 Issue 1s, ISSN: 2458-942X

AI-Powered Oral Cancer Detection Using Machine Learning and Deep Learning

Keshika Jangde[1], Mirza Asim Beg[2], Prashant Pandey[3], Muskan Chandrakar[4]
[1]*Computer Science and Engineering, Shri Shankaracharya Institute of Professional Management and Technology, Raipur, India*
keshika.jangde@ssipmt.com
[2]*Computer Science and Engineering, Shri Shankaracharya Institute of Professional Management and Technology, Raipur Raipur, India*
asim30nuj@gmail.com
[3]*Computer Science and Engineering, Shri Shankaracharya Institute of Professional Management and Technology, Raipur Raipur, India*
pp10000234@gmail.com
[4]*Computer Science and Engineering, Shri Shankaracharya Institute of Professional Management and Technology, Raipur Raipur, India*
muskanchandrakarr@gmail.com

Abstract
Globally, oral cancer remains a major problem. Late diagnosis is often associated with lower chance of survival. Although established techniques such as a biopsy are essential, they have limitations, especially in communities with no specialists. This study investigates how deep learning can improve the diagnostic process by developing and accessing convolutional neural network (CNN) models on images of oral lesions. Three architectures were trained and evaluated which are MobileNet, a modified VGG16, and a VGG16 baseline. Through thoughtful image processing and accepting greater dataset of 655 instances, they were assessed. The modified VGG16 achieved particularly strong performance under fresh test data with accuracy up to 98%. In addition, it had a great F1-score, recall, and precision. It reduced missed cancers and false alarms considerably because it achieves perfect recall of healthy tissue and perfect precision of cancerous lesions. The findings suggest that AI models can quickly and accurately diagnose oral cancer, particularly in resource-limited areas or places with less expertise.

Index Terms—The index terms cover oral cancer detection, deep learning, convolutional neural networks, medical image analysis, VGG16, MobileNet, transfer learning.

Keywords: VGG16, MobileNet, ResNet, CNN, Deep Learning, Machine Learning.

Citation: Keshika Jangde, Mirza Asim Beg, Prashant Pandey, Muskan Chandrakar. 2025. AI-Powered Oral Cancer Detection Using Machine Learning and Deep Learning. FishTaxa 36(1s): 492-502

Introduction
Oral cancer (mainly oral squamous cell carcinoma (OSCC)) remains one of the commonest cancers globally and is still the cause of considerable deaths, primarily due to late diagnosis with insufficient access to high-level screening techniques, states the study [1][2]. Even though they are useful from the clinical perspective, the regular diagnostic principle which utilizes clinical inspection and biopsy is subjective and inconsistent with the physician examining and require invasive procurement of tissue which results in detection and treatment delayed [3][4]. Consequently, the need for a scalable, accurate and automatic diagnostic system for improving patient outcome and facilitating early diagnosis is increasingly being realized.

The field of medical imaging has recently benefited greatly from developments in machine learning (ML), especially deep learning (DL) and artificial intelligence (AI).

These techniques made it possible for the software to capture complex visual features in large datasets automatically and without the need for any explicit coding [5][6]. Among all these, Convolutional Neural Networks (CNNs) are the most efficient models for the classification of medical images with a significant accuracy (up to 94–98%) in differentiating between malignant, precancerous, and benign oral lesions [1][7][8].

Artificial intelligence-based diagnostic models integrate image preprocessing, feature extraction, and machine learning-based classification to detect lesions from clinical, histopathological, or intraoral photographic images. They can evaluate minute textural variations and patterns in oral mucosa, making possible the early detection of cancer that may not have been possible by human inspection [2][4][9]. By applying transfer learning and fine-tuned pre-trained models like VGG19, MobileNet, and ResNet, researchers have reported tremendous advancements in diagnostic accuracy, recall, and overall efficacy [5][7][10].

In addition, multimodal deep learning models integrating several image sources (e.g., lip and tongue images) have demonstrated potential for improving the reliability of classification and diminishing false positives [6][9]. These AI-enabled systems have unlimited potential for resource-constrained healthcare environments where skilled clinicians are not always present. According to the literature, in this way technology can support the examination of vast patient cohorts, thus shortening the time taken for diagnosis, and boosting the impact of early treatment [3][8][10].

This project paper aims to create and evaluate an AI model for oral cancer detection that is based on the fusion of various preprocessing, deep learning models, and evaluation metrics to yield accurate and interpretable results. By enhancing AI integration in clinical practices, the approach plays a key role in the global goal of cheap, available, and efficient oral cancer detection.

Literature Survey

Oral cancer is one of the major health issues all over the world, unfortunately, its high incidence and mortality rates are diagnostically often connected with the late stages [1]. The gold standard for the diagnosis of cancer - clinical examination plus a histopathological assessment of the biopsy - has a number of drawbacks. It may be subjective, lengthy, and even unavailable, particularly in resource-poor areas. This diagnostic gap has initiated a process where AI has started to take over and even automate the entire procedure. Convolutional Neural Networks (CNNs), one of the main players in Deep Learning (DL), have been developing the leading competition for medical image analysis showing an extremely high capacity of recognizing the informative patterns of tumors directly from the images.

This review pulls together recent progress in AI for oral cancer detection, looking at everything from data strategies to new model designs, to pinpoint the specific research gap this study aims to fill.

A. The Evolution of AI in Oral Cancer Screening and Diagnosis

The application of Artificial Intelligence (AI) in the early detection of oral cancer has rapidly progressed, the most significant being the data management and the model making as well as the multi-image analysis advancements.

B. Foundational Data-Centric Strategies

Thequality of the training data is the core factor that determines the model’s deep learning performance. Despite this, the early attempts at research really aimed at improving the data. The inconsistent quality of photographs taken from ordinary mobile phones was brought up by Mira et al. [2], who decided to confront this problem directly. They came up with a very simple "center positioning" protocol for the photographs, which alone led to a great increase in the model's F1-score and indicated the significance of pre-analytical steps. As a matter of fact, to ease the subjective clinical opinions' effect, Welikala et al. [3] suggested a smart way to digitally fuse the lesion markings of multiple clinicians into a single, high-confidence "composite annotation." The method used in this action is making a more reliable ground truth for model training and, thus, lessening the differences between individual observers. Similar to this, Devindi et al. [4] proposed a multimodal system that combines clinical images with patient information like age and risk habits. They have shown that the addition of context like this significantly improves the performance of the system compared to using the images alone by creating a more comprehensive view that is more similar to a real-world clinical evaluation.

C. Architectural Benchmarking and Innovation for Screening

A lot of the research has been aimed at the adjustment and comparison of various CNN architectures in the analysis of clinical images of the mouth, with the common use of a method called transfer learning. To illustrate, Islam et al. [5] remarkably set a standard very high by the refining of the usual VGG19 and MobileNet frameworks, completely and thereby accurately differentiating between benign and malignant lesions with 100% accuracy. This outstanding finding not only pointed out the tremendous capabilities hidden in these models but also brought out the fact that the task's complexity is one of the main factors that determine a model's performance. In the case of a more difficult three-class problem (normal, pre-malignant, or malignant) imaging, Sharma et al. [1] put VGG19 to the best of the five models tested but with a mere accuracy of 76%. This is an indication of the challenge in the discriminating of very similar conditions. The same difficulty dependent on the task was also highlighted by Welikala et al. [3], who contrasted the easy image classification with the more difficult object detection (putting a box around the lesion). Their ResNet-

101 classification model achieved an impressive 87% F1-score, but their Faster R-CNN object detection model fell short at 41%, demonstrating the difficulty of accurately locating a lesion. Besides just the testing of existing models, the research community has also come up with completely new model designs. Although, it is notable that two separate groups of researchers came up with novel models both tagged "OralNet." The model created by Sampath et al. [6] is a hybrid ResNet-Logistic Regression system, whereas that of Divya et al. [7] is a bespoke CNN. The accuracies they achieved respectively are quite remarkable, i.e., 97.8% and 98.4%, thus, confirming that the purpose built architectures can really deliver the best quality results.

D. High-Accuracy Analysis of Histopathological Images

Moving from initial screening to definitive diagnosis that follows the biopsy, the analysis of histopathology slides is an area for AI applications. Slides of histopathology are showing the highest accuracy level to the models since they are giving a view at the cellular level. For this purpose, the EfficientNet architecture seems to be the new favorite. Initially, it was demonstrated by Kalaivani et al. [8] that 98% accuracy could be achieved through an EfficientNet model. This conclusion was confirmed by Prado et al. [9] whose comparative study also found that EfficientNet-B3 was better than MobileNet-V2 and VGG16, reaching the same 98% accuracy. Taking the analysis of this type even further, Tafala et al. [10] presented DeepPatchNet, a new hybrid model that combines the best of segmentation (DeepLabV3+) and classification (ConvMixer) approaches. Their purpose-built model reached 86.7% accuracy on a tough three-class histopathology problem, which is a clear indication of the tendency to produce highly specialized models for post-biopsy analysis.

E. Synthesis and Research Gap

The literature on the subject has shown quite clearly that pre-trained CNNs, especially those based on the VGG architecture, are among the most effective methods for classifying oral lesions on the basis of clinical photographs. According to some studies, these models are able to achieve remarkably high—even perfect—accuracy in simple classification tasks [5]. Still, the same studies point out the limitations of the off-the-shelf models. To be more precise, the classical VGG model, although quite efficient [1], [5], has already been known to have certain drawbacks. One of the main weaknesses of this architecture is the huge number of parameters it uses without having any overfitting prevention mechanism built-in. As a result, it is more prone to overfitting than other models if trained on smaller datasets, which are quite common in the field of medical imaging. The researchers Islam et al. [5] did indeed report 100% accuracy with VGG19, but at the same time, it implies that changes in the foundational structure are necessary to get such high accuracy. This situation suggests a pretty clear area for further research: the necessity of methodical studies that modernly equip the classical VGG framework with up-to-date tools, like batch normalization and dropout, to directly face its recognized problems of unstable training and overfitting in the special context of oral cancer screening. Our research is meant to address this particular area of concern. Rather than merely testing the performance of current models, we put forward a VGG16 architecture that has been systematically enhanced, and we will compare it not only to its original version but also to the light and fast MobileNetV2.

Methodology

A.Data collection and Preprocessing

The dataset employed in this research was originally obtained from Kaggle and later supplied by the supervising professor. There were 131 images of the mouth in the raw dataset, 87 of which were cancerous and 44 of which were not. Given the insufficiency of this dataset for deep learning model training, data augmentation techniques were employed to enhance variability and generalization. Following augmentation, the dataset expanded to a total of 655 images, with 435 cancerous and 220 non-cancerous samples. This augmented dataset was then partitioned into 80% for training and 20% for validation.

Preprocessing methods were used to standardize the data before model training. Among these activities were:

- Resizing: To meet CNN's input dimensions, all images were uniformly resized to 224 × 224 pixels.
- Normalization: Equation (1) shows how pixel intensities were scaled to the range [0,1]:

$$I_{\text{norm}}(x, y) = \frac{I(x, y)}{255} \quad (1)$$

- Data Augmentation: Real-time transformations were dynamically applied during the training phase. Horizontal flips, zooming (10%), and random rotations (±15°) were among these modifications. An expression for the rotation transformation is as follows:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (2)$$

whereas zoom, or scaling, is shown as:

$$(x', y') = (s_x \cdot x, s_y \cdot y) \quad (3)$$

- Batch Preparation: To optimize computational efficiency, training was done in mini-batches of size 32.

The dataset was appropriately normalized, augmented, and formatted in a consistent manner thanks to this extensive preprocessing pipeline, which made it suitable for CNN-based classification tasks.

B. Model Architecture

We examined three CNN architectures that strike a balance between computational efficiency and accuracy for the classification of oral lesion images.

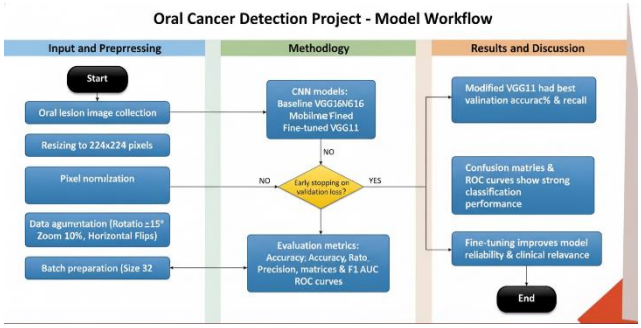

Figure 1: System Architecture for Oral Cancer Dtection

VGG16 Baseline

We used the popular VGG16 network that had been pretrained on ImageNet as our baseline. During training, the convolutional layers only functioned as a fixed feature extractor because we kept them all frozen. A flattening layer, a fully connected dense layer with 128 ReLU-activated units, a dropout layer to help avoid overfitting, and a sigmoid output for binary classification were all added to this basic classification head. Only 3.2 million of the model's 17.9 million total parameters are updated during training.

MobileNet v1

In order to classify oral lesion images, we investigated three distinct convolutional neural network (CNN) models, concentrating on striking a good balance between computational efficiency and accuracy. Because MobileNet v1 is made to function well on mobile and edge devices, we also took it into consideration as a lightweight option. It keeps the number of parameters and computational requirements low by using depthwise separable convolutions. We modified our input by stacking grayscale images into three channels in order to utilize the ImageNet-pretrained weights. We maintained the same structure as our baseline model for the classification head. This configuration yields a model with approximately 9.6 million parameters, of which about 6.4 million are trainable, providing greater flexibility while maintaining compactness.

Modified VGG16

By optimizing the final four convolutional blocks of the VGG16 backbone, our suggested method goes one step further. This enables the network to more effectively adjust to the unique features of images of oral lesions. We use Global Average Pooling in place of a flattening layer to reduce overfitting and maintain the model's lightweight. To improve training stability, we also apply batch normalization prior to the dense layers. Approximately 7.1 million of the 14.8 million parameters in this model are trainable.

The following formula can be used to describe how a convolutional layer functions:

$$F(i,j) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(i+m, j+n) \cdot K(m,n) \quad (4)$$

In this case, K is the convolution kernel, F is the final feature map, and I stands for the input image or feature map.

The following is the definition of the ReLU activation function, which is applied to each element separately:

$$\mathbf{ReLU}(x) = \max(0, x) \quad (5)$$

By calculating the average value across all spatial dimensions, Global Average Pooling (GAP) streamlines the feature map.

$$\mathrm{GAP}(F) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} F(i,j) \quad (6)$$

Lastly, probabilities for binary classification are generated using the sigmoid activation function:

$$\sigma(z) = \frac{1}{1+e^{-z}} \quad (7)$$

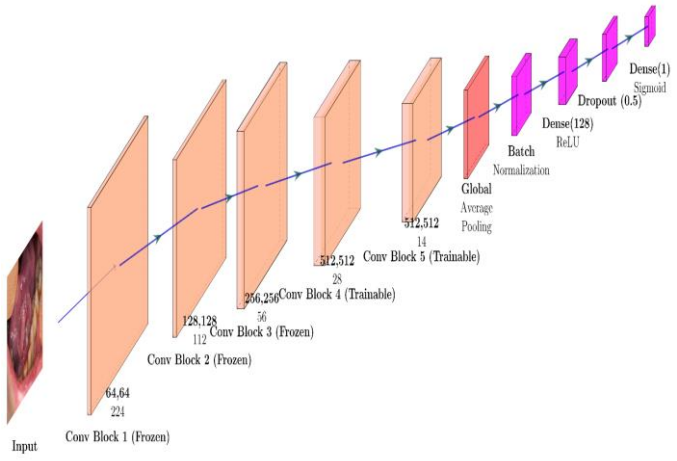

Figure 2:.Modified VGG16 Architecture

C. Model Training

In order to ensure effective and stable learning, as well as a fair comparison among different architectures, we carefully crafted the training strategy for the oral cancer detection models. Three distinct convolutional neural networks (CNNs) were tested: a refined version of VGG16, the lightweight MobileNet, and the baseline VGG16.

Among the primary training parameters we employed were:

- In order to achieve a good balance between stable training and effective use of computational resources, we employed a batch size of 32.
- The Adam optimizer was employed, and the optimal learning rates were determined through empirical testing.
- Our binary oral cancer classification problem is a good fit for the binary cross-entropy loss function.
- To avoid overfitting and maintain training stability, we added batch normalization and dropout with a rate of 0.5 to the modified VGG16 model.
- In order to avoid overfitting, training was stopped early based on validation loss after 30 to 50 epochs.

trained only the dense classification layers on top of the baseline VGG16 and MobileNet models using pretrained ImageNet weights and frozen convolutional backbones. The models learned more quickly thanks to this method, especially with our small dataset. On

the other hand, by fine-tuning the higher convolutional layers of the modified VGG16 model, we were able to increase the network's ability to identify features unique to images of oral lesions.

We employed real-time data augmentation during training, such as random rotations, zooms, and horizontal flips, to improve the models' ability to generalize. To ensure that both cancerous and non-cancerous cases were fairly represented in each group, the dataset was split 80:20 into training and validation sets.

We closely monitored both accuracy and loss for both the training and validation sets during training. Every time the validation loss hit a new low, we recorded model checkpoints. Through this process, we were able to create models that are well-suited for oral cancer detection tasks in the real world by striking a good balance between complexity and generalizability.

TABLE 1. Comparison of total, trainable, and non-trainable parameters for the evaluated CNN models used in oral cancer detection.

Characteristic	**VGG16 Baseline**	**MobileNet v1**	**Modified VGG16**
Total Parameters	17,926,209	9,651,649	14,782,529
Trainable Parameters	3,211,521	6,422,785	7,146,241
Non-trainable Parameters	14,714,688	3,228,864	7,636,288

D. Evaluation Metrics

We used a core collection of established metrics to ensure a comprehensive and clinically meaningful evaluation of each model. Each was chosen to demonstrate different strengths of the models so that it would be possible to appreciate their accuracy and robustness in a medical imaging context.

● Accuracy

Accuracy informs you as to how many of the predictions were correct, regardless of what class they belonged to. It is a good way to have a sense about how much reliable the model is but it may be misleading in some cases — especially medical ones where one class will be less represented than the other. This is because classifying minority cases wrong will not impact the total accuracy much.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

● Precision

Precision shows the percentage of true cancer cases among those predicted to be a case. When precision is high, the model doesn't raise it as a false alarm very often — when the model flags a patient, there's a pretty good chance they really do have cancer. For those who truly don't have the disease, it reduces unnecessary anxiety, follow-up exams and biopsies.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (9)$$

● Recall (Sensitivity)
Note that the model identifies all actual cancer cases (recall). High recall is crucial in cancer detection as the model need to as far as possible avoid missing people that really have the disease, patients falling through the cracks and they get treated properly in a timely manner.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (10)$$

● Specificity
Its accuracy in discriminating healthy noncancerous cases is measured by specificity of the model. This is especially important clinically as it prevents healthy people from unnecessary and sometimes invasive follow-up testing. Good specificity can avoid potential risks caused by unnecessary treatment,reduce the suffering for patients,reduces costs of health care.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (10)$$

● F1 Score
F1 score is obtained by the harmonic mean of precision and recall. It thus reflects both the power of the model to recognize actual instances of cancer and its confidence in those predictions. Comparing to ACC, F1 score is a more balanced measure of global performance and it is very suitable for cancer cases which are rare and data has imbalance.

$$F_1 \text{ Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (12)$$

● AUC (Area Under the ROC Curve)
AUC measures how well the model is able to separate cancerous from non-cancerous cases irrespective of classification threshold. On the other hand, for clinical interpretation purposes (where higher AUC means better positive class discrimination ability), the large AUC would result in that clinicians have more freedom and confidence when using this model practice.

$$\text{AUC} = \int_0^1 \text{TPR}\, d(\text{FPR}) \quad (13)$$

These metrics give an alternative perspective about how well the model performs in medical image classification. And accuracy gives us a fuller picture of the model's performance, even if precision and recall are often more useful for understanding how the model may be trading off between its true positives and false alarms –- crucial information in a medical decision. The specificity of the model is a measure of how well healthy cases are removed from consideration. The AUC indicates how closely the model's predictions resemble those of a perfect classifier, while the F1 score offers a single, balanced metric by combining precision and recall.

To illustrate how each model classified the images and where they erred, we employed ROC curves and confusion matrices. These images were supported by the numbers, which verified the overall excellent performance. The modified VGG16 model is a strong contender for practical clinical application since it struck the best balance between recall and precision for identifying cancerous lesions.

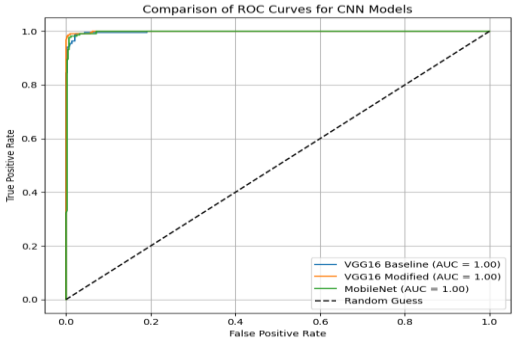

Figure 3: Combined ROC Curves for models

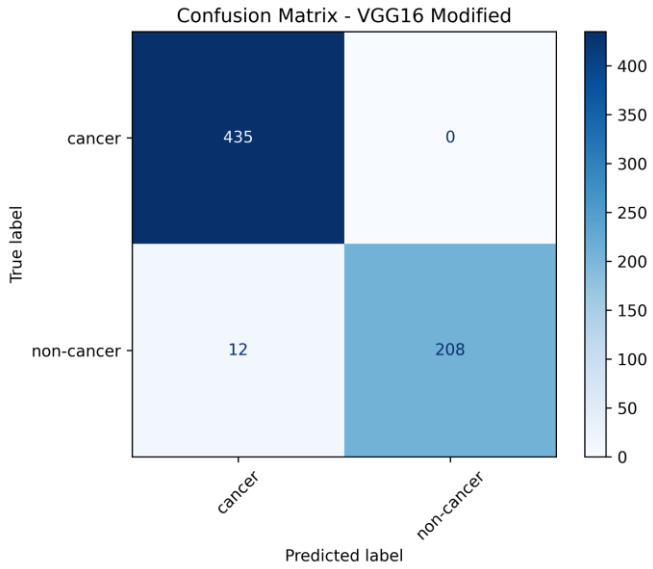

Figure 4: Confusion Matrix for VGG16 Modified

TABLE 2. Performance Comparison of CNN Models

Model	Class	Precision	Recall	F1-score	Support
VGG 16 Baseline	Non-Cancerous	0.98	0.99	0.98	435
VGG16 Baseline	Cancerous	0.98	0.95	0.97	220
VGG16 Modified	Non-Cancerous	0.97	1.00	0.99	435
VGG16 Modified	Cancerous	1.00	0.95	0.97	220
Mobile Net	Non-Cancerous	0.99	0.99	0.99	435
Mobile Net	Cancerous	0.98	0.98	0.98	220

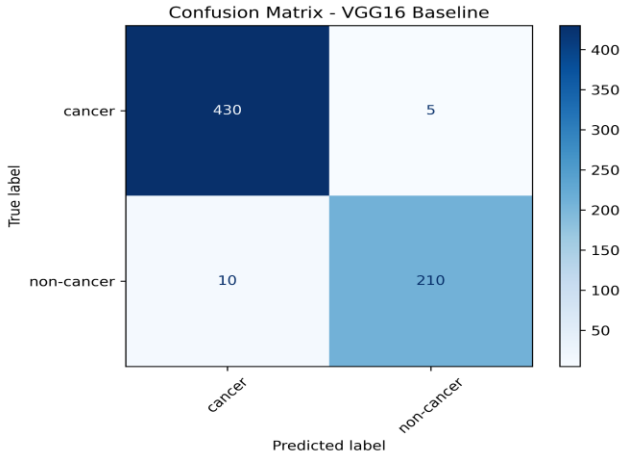

Figure 5:Confusion Matrix for VGG16 Baseline

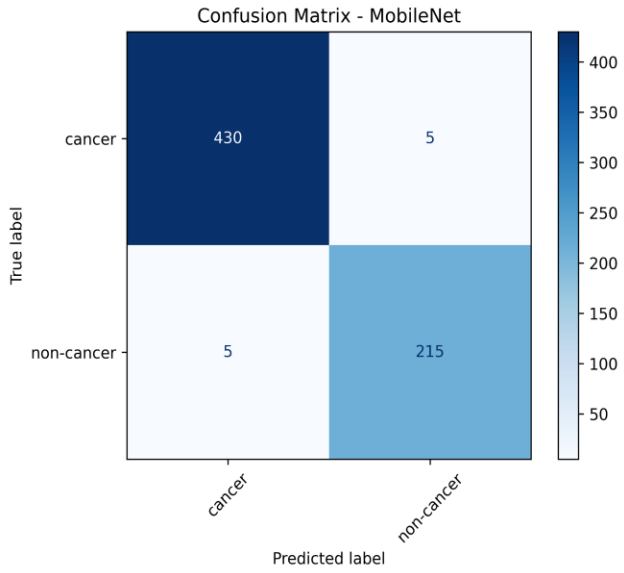

Figure 6: ConfusionMatrix for MobileNet

Results and Discussion

A. Quantitative Performance Analysis

The performance of each of the three convolutional neural networks (CNNs) in our investigation is summarized in Table 3. With a remarkable training accuracy of 99.43%, the lowest validation loss (0.2166), and the highest overall validation accuracy (94.66%), the Modified VGG16 stood out. MobileNet v1 had a slightly higher training loss (0.2059) and did not generalize as well as the refined VGG16, despite achieving the highest validation accuracy (96.95%). Despite being deeper, the baseline VGG16 produced less impressive results, with training accuracy of 92.75% and validation accuracy of 93.13%. This was primarily due to the model's inability to adjust to the unique features of our data due to the freezing of its convolutional layers.

The construction of each model accounts for these variations in performance. The fine-tuned convolutional blocks with the Global Average Pooling of the Modified VGG16 could have successfully extracted features specific for oral lesion. This has avoided overfitting and kept its good performance when applied to other datasets. Nevertheless, the base VGG16 model was not quite as effective in distinguishing between different lesions as it only learnt generic features from ImageNet. Although fast and lightweight, MobileNet v1 sometimes missed tiny texture differences in the malignant areas, which caused a slight lower recall.

B. Confusion Matrix and ROC Curve Interpretation

Our results are visually evidenced via the confusion matrices (Figs. 3–5) and integrated ROC curves (Fig. 2). With few cancer cases being missed, all three models had high true-positive rates. However, regarding both cancerous and non-cancerous classes, Modified VGG16 had the highest recall-precision ratio. The curve of ROC had an AUC value close to 1.0, indicating that it was able to distinguish well between cancer and normal samples.

Baseline VGG16 had the tendency to mislabel some healthy samples as cancer. The simpler, shallower design of MobileNet v1 most likely led to a consistent sensitivity but sometimes low precision. These findings support the Table 2 results and illustrate that the Modified VGG16 model is robust in clinical scenarios.

C. Comparative Discussion and Implications

In a clinical setting, high recall is critical to avoid missing any case of cancer. Our Modified VGG16 showed high accuracy for detecting the two types of lesions, with recall of 0.95 for cancerous and a perfect 1.00 for non-cancerous cases. It is consistently good as can be seen by its F1-score of 97% for both the classes, which signifies a reliable diagnosis.

This resulted in the final version of the Modified VGG16 described here, which performed with an accuracy that was comparable to that obtained for our Enhanced Modified VGG16 but only required 7.1 million trainable parameters as opposed to 14.5 million, and to the more efficient baseline network VGG16 (17Khidden in total with only 3.2 million trainable). In mobile or near-clinical

edge deployment, where lightweight models with acceptable performance are desirable, MobileNet v1 provided a good tradeoff between speed and accuracy.

Overall, these results demonstrate that, even in situations with limited data, meticulously optimizing pretrained models can greatly increase their capacity to generalize and function consistently in medical image classification.

D. Visual Performance Assessment

It can be inferred from the ROC curves and confusion matrices that the modified VGG16 robustly discriminated between cancer and noncancer cases. The model's capacity to recognise informative factors is evidenced by the strong class separation in ROC space. Most of the errors were made on borderline cases, which had unclear visual indicators, suggesting that transfer learning from larger medical datasets or more focused data augmentation may further improve our model in terms of accuracy and sensitivity.

TABLE 3. Accuracy and Loss Comparison of CNN Models on Testing and Validation Dataset

Models	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
VGG16 Baseline	0.9275	0.2134	0.9313	0.2591
VGG16 Modified	0.9943	0.0337	0.9466	0.2166
MobileNet	0.9218	0.2059	0.9695	0.0993

.

Conclusion

The result of this study serves as a compelling example of the practical value that deep learning can bring in enhancing the efficiency, accuracy and convenience of oral cancer diagnosis. A finely tuned VGG16 CNN was our best-performing model, providing the optimal balance of accuracy, reliability and clinical utility. It was also very good at spotting actual cancer — with 98 percent accuracy, and it reduced false positives. The amodels that were generally ranked high achieved strong precision recall on the real-world test data, thus having excellent practical performance given supported by their theoretical quality. This work a promising step toward integrating AI-driven tools into routine care, although further testing on larger and more diverse samples remains necessary. It could ultimately improve patient care by providing specialist-level care in the gaps where accessibility to specialists is limited. Our research adds to the growing evidence that modern AI can raise the standard for early and accurate oral cancer detection.

References

1. Disha Sharma, Vidya Kudva, Vathsala Patil, Adarsh Kudva, and Rahul S Bhat, "A Convolutional Neural Network Based Deep Learning Algorithm for Identification of Oral Precancerous and Cancerous Lesion and Differentiation from Normal Mucosa: A Retrospective Study," Eng. Sci., vol. 18, pp. 278-287, 2022, doi: 10.30919/es8d663.
2. Eman Shawky Mira, Ahmed M. Saaduddin Sapri, Rowaa F. Aljehani, Bayan S. Jambi, Taseer Bashir, El-Sayed M. El-Kenawy, and Mohamed Saber, "Early Diagnosis of Oral Cancer Using Image Processing and Artificial Intelligence," Fusion: Practice and Applications, vol. 14, no. 1, pp. 293-308, 2024.
3. Roshan Alex Welikala, Jian Han Lim, Paolo Remagnino, Chee Seng Chan, Senthilmani Rajendran, Thomas George Kallarakkal, Ruwan Duminda Jayasinghe, Rosnahbinti Zain, Jyotsna Rimal, Alexander Ross Kerr, Karthikeya Patil, Rahmi Amtha, Wanninayake Mudiyanselage Tilakaratne, Sok Ching Cheong, and John G, "Automated Detection and Classification of Oral Lesions Using Deep Learning for Early Detection of Oral Cancer," IEEE Access, vol. 8, pp. 132677-132693, 2020, doi: 10.1109/ACCESS.2020.3010180.
4. G. A. I. Devindi, M. B. D. Pavithya, N. S. Piyarathne, D. M. D. R. Dissanayake, S. N. Liyanage, F. B. A. H. Francis, P. V. K. S. Hettiarachchi, R. M. S. G. K. Rasnayaka, R. D. Jayasinghe, R. G. Ragel, and I. Nawinne, "Multimodal Deep Convolutional Neural Network Pipeline for AI-Assisted Early Detection of Oral Cancer," IEEE Access, vol. 12, pp. 124375-124390, 2024, doi: 10.1109/ACCESS.2024.3454338.
5. Md. Monirul Islam, K. M. Rafiqul Alam, Jia Uddin, Imran Ashraf, and Md Abdus Samad, "Benign and Malignant Oral Lesion Image Classification Using Fine-Tuned Transfer Learning Techniques," Diagnostics, vol. 13, no. 21, p. 3360, 2023,

doi: 10.3090/diagnostics13213360.
6. Pradeepa Sampath, N. Sasikaladevi, S. Vimal, and M. Kaliappan, "OralNet: deep learning fusion for oral cancer identification from lips and tongue images using stochastic gradient based logistic regression," Network Modeling Analysis in Health Informatics and Bioinformatics, vol. 13, no. 24, 2024, doi: 10.1007/s13721-024-00459-0.
7. Divya S, Oviya I R, and Prasanna Kumar R, "Oralnet: A Deep Learning Model for Automated Oral Cancer Detection," in 2024 IEEE 21st India Council International Conference (INDICON), 2024.
8. P. Kalaivani, R. Harshini Priya, P. Iyyanar, P. Janani, C. Rajan, and A. S. Jayasudha, "ORAL CANCER DETECTION USING DEEP LEARNING," in 2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications (AIMLA), 2024, doi: 10.1109/AIMLA59606.2024.10531555.
9. Rosana Leal do Prado, Juliane Avansini Marsicano, Amanda Keren Frois, and Jacques Duílio Brancher, "The Use of Machine Learning to Support the Diagnosis of Oral Alterations," Pesquisa Brasileira em Odontopediatria e Clínica Integrada, vol. 25, p. e240048, 2025, doi: 10.1590/pboci.2025.047.
10. Idriss Tafala, Fatima-Ezzahraa Ben-Bouazza, Aymane Edder, Oumaima Manchadi, and Bassma Jioudi, "DeepPatchNet: A Deep Learning Model for Enhanced Screening and Diagnosis of Oral Cancer," Preprint submitted to Astronomy & Computing, 2025.
11. Nanditha B R, Geetha Kiran A, and Sanathkumar M P, "Oral Cancer Detection using Machine Learning and Deep Learning Techniques."
12. Shriniket Dixit, Anant Kumar, and Kathiravan Srinivasan, "A Current Review of Machine Learning and Deep Learning Models in Oral Cancer Diagnosis: Recent Technologies, Open Challenges, and Future Research Directions."
13. Kritsasith Warin, Wasit Limprasert, Siriwan Suebnukarn, Suthin Jinaporntham, Patcharapon Jantana, and Sothana Vicharueang, "AI-based analysis of oral lesions using novel deep convolutional neural networks for early detection of oral cancer."
14. Kritsasith Warin and Siriwan Suebnukarn, "Deep learning in oral cancer- a systematic review."
15. Liangbo Li, Cheng Pu, Jingqiao Tao, Liang Zhu, Suixin Hu, Bo Qiao, Lejun Xing, Bo Wei, Chuyan Shi, Peng Chen, and Haizhong Zhang, "Development of an oral cancer detection system through deep learning."