

# An Automated Oral Cancer Detection System Using Deep Learning Feature Extraction and Tree-Based Classification

Mr. Akash Kumar Bhagat<sup>1</sup>, Dr. Arvind Kumar Pandey<sup>2</sup>

<sup>1</sup>Research Scholar, Department of Computer Science, Arka Jain University, Jharkhand

<sup>2</sup>Dr. Arvind Kumar Pandey, Dean, School of Engineering and IT, Arka Jain University Jharkhand

## Abstract

Oral cancer represents a major global health threat because early detection with timely medical treatment improves patient outcomes. Diagnostic procedures face challenges which include both subjective evaluation and problems with identifying changes at advanced disease stages. This study introduces an innovative hybrid technique which employs deep learning and machine learning together with advanced computer vision methods to enhance oral cancer detection accuracy through biomedical imaging. The framework employs a pre-trained CNN model which functions as an automatic feature extractor because it can detect intricate hierarchical patterns within image datasets. The Random Forest classifier receives all extracted features for the process of final classification. Our two-tier system combines the deep network's powerful feature extraction capacity with the dependable performance of standard machine learning methods. Our hybrid method achieves superior performance metrics which include accuracy, precision, and recall when compared to traditional methods while delivering precise automatic detection of oral cancer at its initial stages. The research offers important potential to help doctors reach more accurate diagnostic evaluations which will lead to better patient treatment results.

**Keywords:** Oral Cancer Detection, Two-Stage Detection, Deep Learning, Machine Learning,, CNN, Random Forest etc.

## 1. Introduction

The Introduction section establishes early detection of oral cancer as an essential factor for achieving positive patient outcomes because this disease poses both high risk and widespread occurrence. Oral cancer occurs as a major global head and neck malignancy which occurs at higher rates in India due to tobacco consumption and public knowledge deficit about this disease. The current diagnosis process for diseases requires medical professionals to examine visual materials and perform biopsy procedures, which results in time-consuming and subjective evaluations that experience work delays. Medical delays result in severe negative effects because advanced stage diseases have a high likelihood of poor survival outcomes. The existing diagnostic techniques experience their current limitations which create an urgent demand for medical analysis methods that show greater accuracy and efficacy. Current deep learning and machine learning technology enables medical professionals to conduct medical image analysis at higher efficiency through automated systems. The deep learning models, and CNNs for the most part, tend to be strong enough to extract these intricate, slight and at times nearly impossible for human eyes to decipher

patterns from the high-dimensional image data. The medical image technologies utilize these systems to assist doctors with both tumor detection and disease classification tasks. Some deep learning models maintain their ability to learn complex image data features but they function as "black box" systems which make it impossible to understand decision-making processes. The medical level maintains this intransparency as a major problem because security requirements create accountability obligations which demand transparency. The two-stage framework for oral cancer detection which we present as our solution combines both paradigms through its integrated all-in-one system. Our method uses the pre-trained CNN as an automatic feature extractor because it enables us to detect and encode complex hierarchical patterns that exist in biomedical images. The learned features from CNN were used to create input for the Random Forest classifier, which became our final classification system. Our core approach operates through a two-tier framework system, which unites deep networks exceptional feature extraction abilities with traditional machine learning systems established reliability. Random Forest classifier functions as the final diagnostic decision tool because it gives explicit classification outcomes which handle high-dimensional datasets. The joint-set learning strategy provides superior results through its performance of essential indicators, which include memory capacity and accuracy and precision, compared to conventional methods that use single-model-based approaches. The developed work presents extensive possibilities to enhance healthcare tools, since it enables healthcare practitioners to obtain precise and prompt diagnosis of oral cancer through its powerful automated diagnostic assistance system. The healthcare system becomes supported through patient treatment and survival advantages provided by reliable diagnostic assistance system, which boosts diagnostic accuracy while enhancing patient treatment outcomes.

## **2. Literature Survey**

Clinical oncology faces difficulties in detecting oral cancer at its initial stage because most cases are identified during advanced disease progression. Oral cavity squamous cell cancer represents an etiologically complex neoplasm which develops metastases after its undetected initial stage [1]. The unification of these clinical elements results in better patient prognoses according to medical professionals. Medical professionals have relied on visual examination and physical examination to detect diseases throughout history. The evidence about organized screening programs remains unclear according to existing research. The existing research gap creates an immediate requirement for diagnostic tools which use technology to deliver precise test results. Researchers have shifted their research efforts to Computer-Assisted Medical Image Classification in order to solve the problems which arise during manual screening. The first deep learning algorithms demonstrated exceptional performance when they distinguished between malignant and benign lesions with extremely high accuracy [5]. Diagnostic screening processes become more reliable because the models provide second opinions which decrease both human error and diagnostic fatigue. The latest developments have created sophisticated automated detection systems. The system uses deep learning technology to classify images while automatically detecting lesion boundaries which enables earlier treatment than traditional methods permit [3]. The maintenance of operational consistency across various clinical settings requires automation because it plays an essential role in healthcare delivery. The advanced detection systems now use multi-layer neural networks to achieve sensitivity levels which match those of professional pathologists [4]. The AI-driven systems create an effective solution for resource-limited areas which lack access to specialists while enabling cost-effective early-stage monitoring [4,5]. The medical community will achieve a standardized yet widely available method for oral cancer prevention through technology integration.

### 3. Methodology

The study automates the image-based framework for oral cancer detection through our two-stage diagnostic solution. The system first learns to extract discriminative visual features from input images before it conducts classification in the second decision stage. Our methodology includes three core elements which involve dataset preprocessing and feature extraction through convolutional neural networks and two-stage oral cancer detection. The subsequent subsections provide detailed descriptions of each component.

#### 3.1 Description of the Dataset and Data Processing

The study utilized a dataset which researchers obtained from public Kaggle data sources. The dataset contains two types of images which show cancerous and non-cancerous oral conditions. The cancerous image class contains images of oral cancer patients who have been verified as candidates, while the non-cancerous healthy groups contain images of normal oral tissues images. The model training process required all images to undergo pre-processing before its commencement. The images undergo size conversion to  $224 \times 224$  pixels, which enables CNN model compatibility with the existing trained model. Our model required pixel value standardization from 0 to 1 range, which enabled better training performance. The dataset distribution process establishes training and testing sets at 80:20 ratio, while ensuring that all classes remain evenly distributed across both sets.

Sl. No	Image class	Count of images	Description of images
1	Cancerous Data	1000	Contains cancer images of patients.
2	Non-Cancerous Data	700	Contains non-cancerous healthy images.
3	Total	1700	Total no of images

**Table 1 shows how the oral cancer dataset’s picture classes are distributed.**



**Figure 1: Sample images representing each class category.**

#### 3.2 Using Convolutional Neural Networks to Extract Features

The Convolutional Neural Network (CNN) pre-trained models were used to extract features from the data.

MobileNetV2 is a pre-trained CNN that has lightweight characteristics. The model's categorization layer's upper part was removed to install the Global Average Pooling layer which produced fixed-size feature vectors for each image. The network achieved better performance because this method enabled it to discover important distinctive image features while reducing its power needs for processing. The feature extraction process created high dimensional representations which were used to train both the training set and the testing set before proceeding to the classification stage.

### 3.3 Two-Stage Oral Cancer Detection

The Random Forest Classifier receives extracted features which have gone through feature extraction process to create the final prediction. The Random Forest Classifier was selected as the most suitable method because it could analyze high-dimensional feature data and deliver understandable outcomes while stopping the training process from overfitting. The two-stage model workflow consists of multiple steps which include the following elements.

- Feature extraction: MobileNetV2 creates high level features vectors through its feature extraction process.
- Classification: The extracted feature goes into the random forest classifier which uses 100 decision trees to predict whether the labels indicate malignancy or non-malignancy.
- Evaluation: Our model performance is assessed through standard evaluation metrics which include F1-score and plotting confusion matrices and ROC curves and accuracy-recall curves and 5-fold validation.

## 4. Results and Conclusion

This section evaluates the functioning of the two-stage system which has been developed for oral cancer detection purposes. The model validation process tests an oral cancer dataset which follows a standard training-testing distribution of 80 percent for training and 20 percent for testing. The model assessment process measured its performance.

The assessment used ROC curve and Precision Recall curve and F1-Score and Accuracy and Precision and Recall metrics together with Confusion Matrix visualization as assessment methods.

### 4.1 Model Performance Metrics

The proposed two-stage model frameworks use convolutional neural networks (CNNs) to extract features and classify data through Random Forest classifiers. The assessment of our strategy effectiveness resulted in performance indicators which we gathered and compiled.

Sl No	Metrics	Value in %
1	Accuracy	92.67 %
2	Precision	92.38 %
3	Recall	97.00 %
4	F1-Score	94.63 %

**Table 2: Performance evaluation of the proposed two stage oral cancer detection model.**

The findings show that the suggested framework achieves strong classification performance across all evaluation measures. The high accuracy reflects the model's overall reliability, while the elevated recall value indicates its effectiveness in identifying oral cases of cancer, which is crucial in medical diagnosis where missed detections can have serious clinical consequences.

#### 4.1.1 Evaluation Metrics Formulation

Common methods for the classification metrics obtained from the results of our model are evaluated using the confusion matrix, which includes TP (True Positives), TN (True Negatives), FP (False Positives), and FN (False Negatives).

The percentage of all positive predictions that the model was accurate on is known as accuracy.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision measures the ratio of actual positive cases (images that depict a disease) to all positive cases:

$$Precision = \frac{TP}{TP + FP}$$

The percentage of anticipated positive cases that are actually positive is called recall, occasionally called sensitivity:

$$Recall = \frac{TP}{TP + FN}$$

Recall and precision are compared harmonically using F1Score, which generates the mean of the two metrics:

$$F1Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

#### 4.2 Matrix of Confusion

By comparing the classification results, the confusion matrix provides a complete view. It's predicted labels with actual ground truth values. It shows how correctly and incorrectly our model classified each class on the test data.

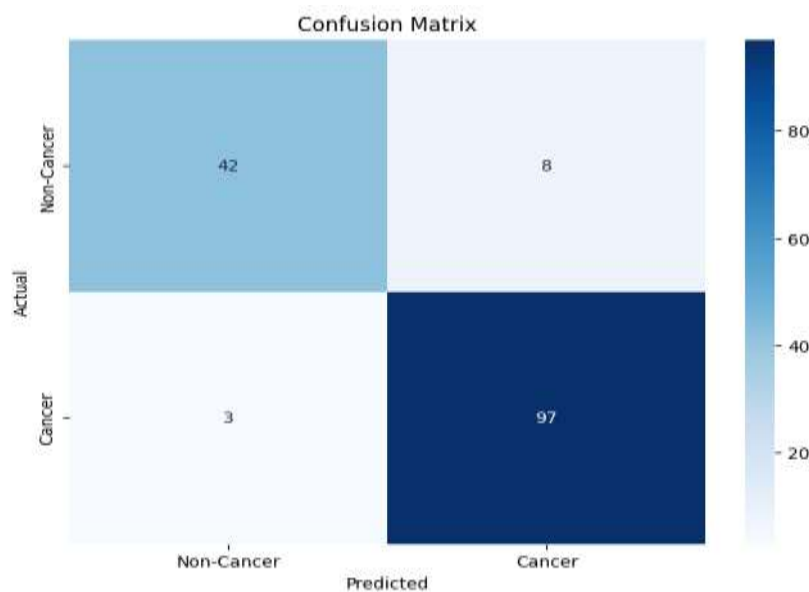
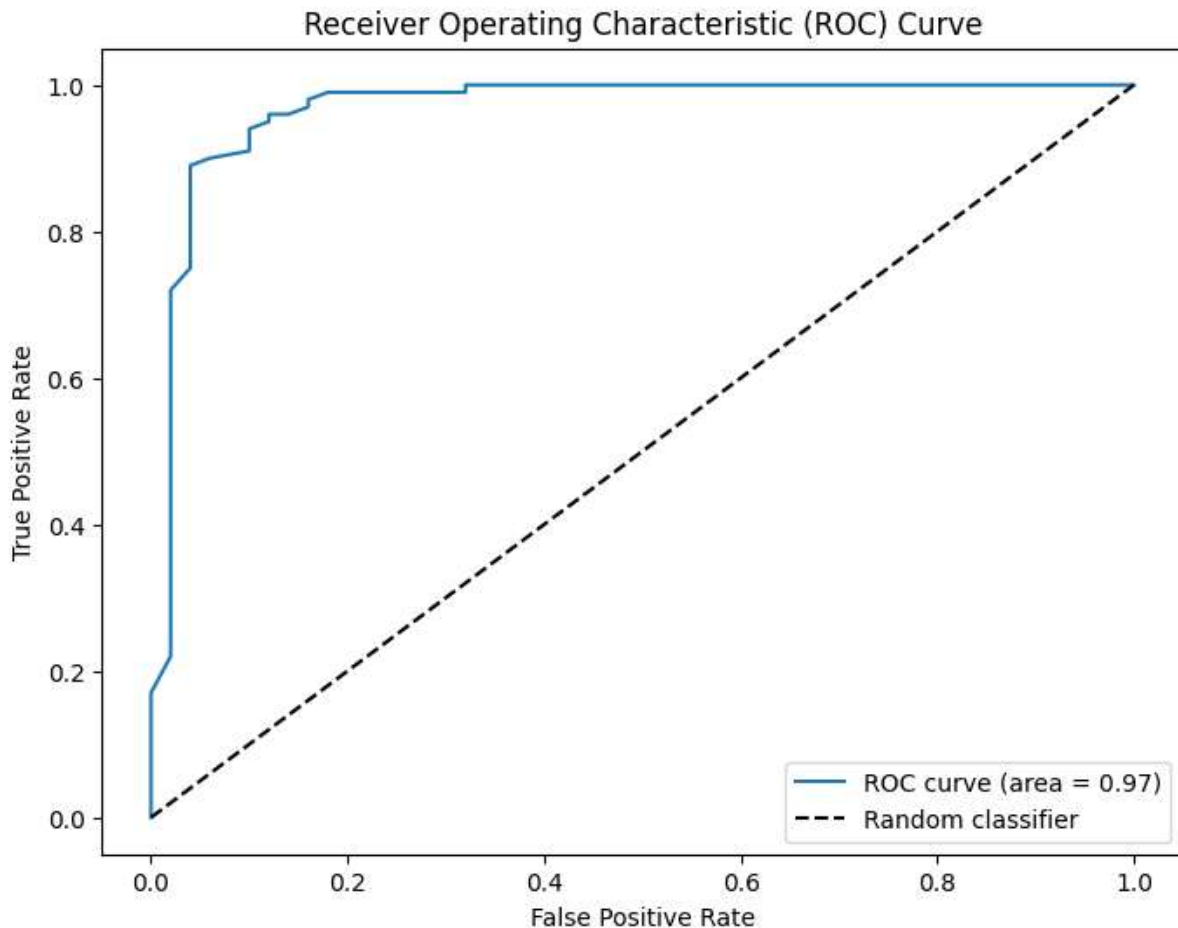


Figure 2: The two-stage oral cancer detection model's confusion matrix is depicted in this graph.

### 4.3 Operating Characteristic Curve of the Receiver (ROC)

To visualise by modifying the positive true rate (recall) and the false positive rate, these curves are mostly used to assess and compare how well your classifier or diagnostic test is accurate at various cutoff points. their trade-off between sensitivity & specificity by taking different threshold values.

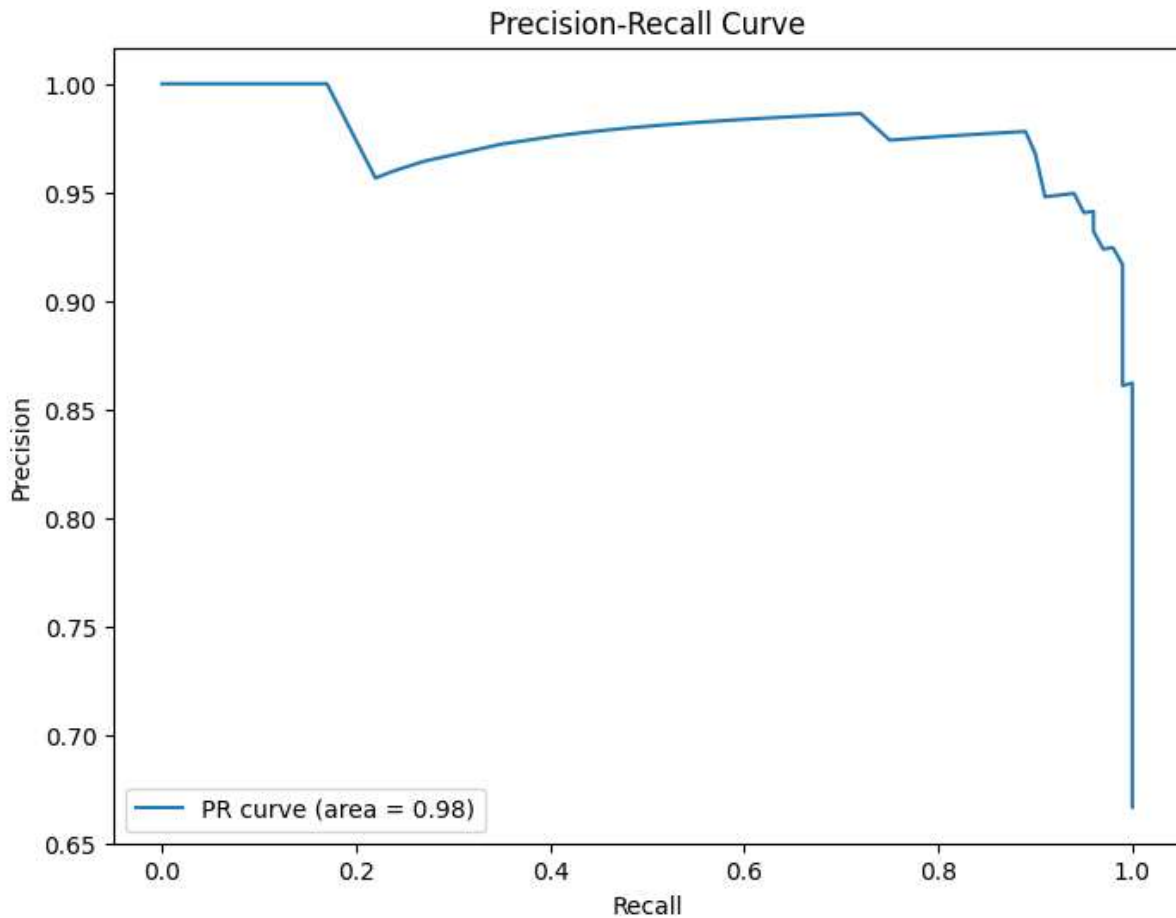


**Figure 3: The receiver operating characteristic curve for our two-stage oral cancer detection model.**

### 4.4 The Discussion

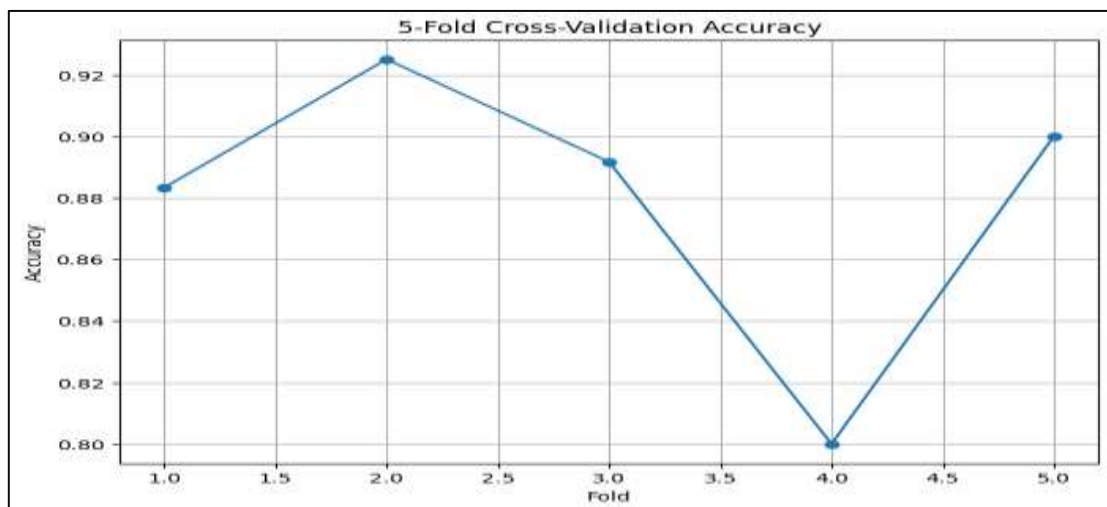
The experimental outcomes demonstrate that the suggested two-stage deep learning and machine learning framework outperforms traditional single-model methods for detecting oral cancer. The first stage of the system uses a CNN component to capture detailed visual features from oral images, while the second stage uses a Random Forest classifier to enhance classification performance through better generalization and resilience. The RF classifier enhances the classification process by improving its generalization capacity and resilience against various challenges.

- Recall (0.97): Ensuring that most of the malignant individuals are correctly diagnosed reduces the risk of a misdiagnosis.
- Precision (0.9238): Eliminates unnecessary clinical treatments by preventing false positives.
- F1-Score (0.9463): This indicates that recall and precision are performed in a balanced manner.



**Figure 5: Precision–recall analysis showing our model’s the capacity of a model to lower false-positive predictions.**

To assess the stability of our framework, Five-fold cross-validation was carried out. The results, shown in Fig. 5 which are indicating a consistent performance across different data partitions by confirming the generalizability and robustness of the model.



**Figure 6: 5-fold cross-validation accuracy showing that our model stability and consistent performance across different data partitions.**

Overall, the performance analysis confirms that the proposed two-stage framework represents a reliable and automated method for detecting mouth cancer early. Its strong predictive precision, memory, and accuracy, which makes it applicable to clinical decision support systems, facilitating timely diagnosis and improved patient outcomes.

## 5. Conclusion

The research developed a two-stage model framework which enables identification of oral cancer disease. We applied MobileNetV2 CNN as a pre trained feature extractor to process images together with Random Forest Classifier for precise classification tasks. The model was tested on public oral cancer dataset and achieved high performance results which included 92.67% accuracy and 92.38% cancer detection rate and 97% actual cancer case detection rate together with 94.63% F1-score.

The outcomes demonstrate how well our two-stage approach captures the complex pattern in biomedical images and gives clear, dependable results. Our model demonstrates high recall together with strong precision which shows that it detects actual cancer cases while it keeps healthy patients from being classified as having cancer. The model proves its value for medical applications because it enables doctors to discover cancer in its initial stages. The model enables users to boost performance through its three customization options which include adding extra images and selecting different feature extraction techniques and machine learning methods.

The proposed two-stage methodology provides an automated and efficient method for detecting early stage oral cancer with high accuracy. Doctors can use this system to complete their diagnosis work in a time efficient manner which results in improved patient treatment. The research team plans to analyze larger hospital datasets while they develop methods which will help doctors better understand our model.

## References

1. Rivera C. (2015). Essentials of oral cancer. *International journal of clinical and experimental pathology*, 8(9), 11884–11894
2. Brocklehurst P, Kujan O, O'Malley L, Ogden GR, Shepherd S, Glenny A-M. Screening programmes for the early detection and prevention of oral cancer. *Cochrane Database of Systematic Reviews* 2013, Issue 11. Art. No.: CD004150. DOI: 10.1002/14651858.CD004150.pub4
3. R. A. Welikala et al., "Automated Detection and Classification of Oral Lesions Using Deep Learning for Early Detection of Oral Cancer," in *IEEE Access*, vol. 8, pp. 132677-132693, 2020, doi: 10.1109/ACCESS.2020.3010180
4. Li, L., Pu, C., Tao, J. et al. Development of an oral cancer detection system through deep learning. *BMC Oral Health* 24, 1468 (2024). <https://doi.org/10.1186/s12903-024-05195-5>
5. Jeyaraj, P.R., Samuel Nadar, E.R. Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm. *J Cancer Res Clin Oncol* 145, 829–837 (2019). <https://doi.org/10.1007/s00432-018-02834-7>
6. S. Mehta and A. Aneja, "Deep Learning Meets Traditional AI: A Hybrid CNN-RF Model for Accurate and Explainable Skin Lesion Classification," 2025 International Conference on Automation and Computation (AUTOCOM), Dehradun, India, 2025, pp. 1-5, doi: 10.1109/AUTOCOM64127.2025.10956217.

7. R. A. Welikala et al., "Automated Detection and Classification of Oral Lesions Using Deep Learning for Early Detection of Oral Cancer," in *IEEE Access*, vol. 8, pp. 132677-132693, 2020, doi: 10.1109/ACCESS.2020.3010180
8. Janiesch, C., Zschech, P. & Heinrich, K. Machine learning and deep learning. *Electron Markets* 31, 685–695 (2021). <https://doi.org/10.1007/s12525-021-00475-2>
9. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. -C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018, pp. 4510-4520, doi: 10.1109/CVPR.2018.00474.
10. Mehrotra, R., & Gupta, D. K. (2011). Exciting new advances in oral cancer diagnosis: avenues to early detection. *Head & neck oncology*, 3, 33. <https://doi.org/10.1186/1758-3284-3-33>
11. Gunardi, I., Sufiawati, I., Goenawan, H., Herawati, D. M. D., Lesmana, R., & Abdullah, A. G. (2023). Research Trends in Molecular Biological Studies on Oral Squamous Cell Carcinoma: A Bibliometric Analysis. *Oncology reviews*, 17, 11585. <https://doi.org/10.3389/or.2023.11585>
12. Kouketsu, A., Doi, C., Tanaka, H., Araki, T., Nakayama, R., Toyooka, T., Hiyama, S., Iikubo, M., Osaka, K., Sasaki, K., Nagai, H., Sugiura, T., Yamauchi, K., Kuroda, K., Yanagisawa, Y., Miyashita, H., Kajita, T., Iwama, R., Kurobane, T., & Takahashi, T. (2024). Detection of oral cancer and oral potentially malignant disorders using artificial intelligence-based image analysis. *Head & neck*, 46(9), 2253–2260. <https://doi.org/10.1002/hed.27843>
13. Al-Rawi, N., Sultan, A., Rajai, B., Shuaeeb, H., Alnajjar, M., Alketbi, M., Mohammad, Y., Shetty, S. R., & Mashrah, M. A. (2022). The Effectiveness of Artificial Intelligence in Detection of Oral Cancer. *International dental journal*, 72(4), 436–447. <https://doi.org/10.1016/j.identj.2022.03.001>
14. Akl, A. A., Hosny, K. M., Fouda, M. M., & Salah, A. (2023). A hybrid CNN and ensemble model for COVID-19 lung infection detection on chest CT scans. *PloS one*, 18(3), e0282608. <https://doi.org/10.1371/journal.pone.0282608>
15. Ogundokun, R. O., Misra, S., Akinrotimi, A. O., & Ogul, H. (2023). MobileNet-SVM: A Lightweight Deep Transfer Learning Model to Diagnose BCH Scans for IoMT-Based Imaging Sensors. *Sensors (Basel, Switzerland)*, 23(2), 656. <https://doi.org/10.3390/s23020656>