Black Sea Journal of Engineering and Science
doi: 10.34248/bsengineering.1528581

Open Access Journal
e-ISSN: 2619 – 8991

Research Article
Volume 8 - Issue 1: 94-101 / January 2025

ORAL CANCER CLASSIFICATION WITH CNN BASED STATE-OF-THE-ART TRANSFER LEARNING METHODS

Kaan GÜMELE[1], Muhammet Sinan BAŞARSLAN[1*]

[1]*İstanbul Medeniyet University, Faculty of Science, Department of Computer Engineering, 34700, İstanbul, Türkiye*

Abstract: The importance of oral and dental health closely affects other vital organs. In this study, CNN-based transfer learning models are built on histopathologic and intraoral images with benign and malignant lesions. Histopathologic and intraoral images from two different sources have benign or malignant classes of lesions in the mouth. EfficientNetB7, ResNet50, VGG16, and VGG19, Xception, ConvNextBase, and MobileNetV2 were used as transfer learning methods. Model training was performed with 80%-20% train test separation and 20% validation separation on the train set. Accuracy (Acc), Precision (Prec), Recall (Rec), and F1-score (F1) metrics were used to evaluate the model. In histopathologocial images, ResNet50 was ahead with 0.8125 Acc and 0.8525 F1. In intraoral images, ConvNextBase with 0.84 Acc, and 0.80 F1 was found to be more accurate.

Keywords: Oral cancer, Image processing, Convolutional neural network, Transfer learning

Corresponding author: İstanbul Medeniyet University, Faculty of Science, Department of Computer Engineering, 34700, İstanbul, Türkiye
E mail: msinanbasarslan@gmail.com (M. S. BAŞARSLAN)
Kaan GÜMELE https://orcid.org/0009-0002-4262-0585
Muhammet Sinan BAŞARSLAN https://orcid.org/0000-0002-7996-9169

Received: August 05, 2024
Accepted: November 25, 2024
Published: January 15, 2025

Cite as: Gümele K, Başarslan MS. 2025. Oral cancer classification with CNN based state-of-the-art transfer learning methods. BSJ Eng Sci, 8(1): 94-101.

1. Introduction

Early detection of oral cancer is crucial for successful treatment and patient outcomes. Oral cancer, which includes cancers of the lips, tongue, gums, floor of the mouth and other sites in the oral cavity, can often be asymptomatic in its early stages, leading to late diagnosis and poor prognosis. Therefore, effective and accurate diagnostic methods are essential in the fight against this disease Goswami et al. (2024).

Histopathologic examination, which involves taking tissue samples from the oral cavity for analysis in a laboratory setting, is the primary method for diagnosing oral cancer. However, this process is time-consuming as it requires meticulous examination of tissue samples under a microscope at different magnifications, typically 40x, 100x, and 200x. Moreover, the Acc of the diagnosis relies heavily on the expertise of the pathologists, leading to potential discrepancies and delays in initiating treatment. On the other hand, obtaining intraoral images offers a non-invasive and relatively quick way to screen for oral abnormalities. Unlike histopathologic examination, intraoral imaging can be performed by individuals without specialized training, making it accessible for routine screening and self-assessment. This accessibility not only enables individuals to monitor their oral health, but also facilitates early detection of suspicious lesions, thereby increasing the chances of successful treatment outcomes. Intraoral images are obtained from digital images acquired during routine examinations, while histopathologic images are obtained by examining tissue samples under a microscope (de Lima et al., 2023).

It has become very important for health that imaging devices produce well-defined visual outputs. With artificial intelligence (AI) technologies, especially images on health will be highly efficient for human health.

In this study, transfer learning methods were used to build models to identify oral cancer and potentially malignant lesions on an open-source dataset with two different sources and data types (histopathologic, intraoral).

The dataset used in the study is an important resource for training and testing deep learning algorithms, while also providing a guide for researchers to improve oral cancer diagnosis. However, the collection, processing and evaluation of these images require more expertise and time than the detection of a photograph taken with a camera. This process involves the careful work of one or more pathologists and the accurate analysis of each image. Therefore, the collection and processing of histopathologic images usually takes longer and requires more expertise.

The contribution of this study is listed below:

- Open source Two different types of images (intraoral, histopathologic) of oral cancer patients will be analyzed to see which state-of-the-art Transfer learning method gives the best results in the ImageNet competition.
- The performance of transfer learning models with different parameters will be investigated in the

classification of images that can be taken with all kinds of cameras that can be accessed as well as images that require expertise to determine whether the lesion in the mouth is benign or malignant.

The block diagram of the study is given in Figure 1.

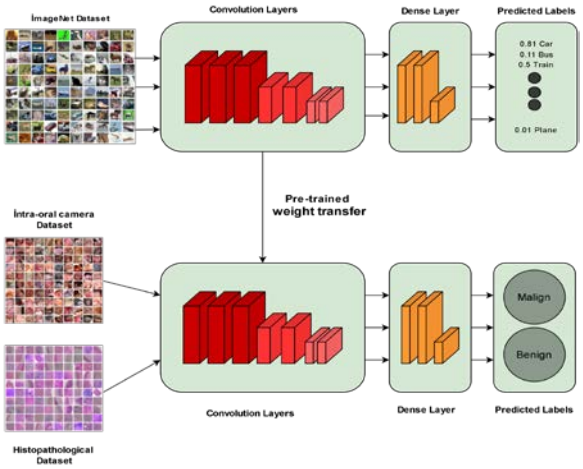


Figure 1. Block diagram of study.

The remainder of this paper is structured as follows: Section 2 reviews the literature. Section 3 details the data sources and discusses the deep learning techniques, transfer learning methods, and CNN-based approaches utilized in this study. Section 4 outlines the experimental settings and presents the results. Section 5 provides an evaluation and discussion of the study's overall findings.

2. Literature Review

In this section, we review the studies on the detection of good or bad lesions in the human mouth. Babu et al. proposed a CNN-based transfer learning method for early detection and diagnosis of oral cancer after various preprocessing and categorizing images by clustering with Fuzzy C means. They used UTI medical data sets in the study. Image categorization was developed using various Artificial Neural Network (ANN) topologies. VGG16, VGG19, DenseNet121, and DenseNet169 were used, as well as EfficientNetB0, EfficientNetB1, and EfficientNetB2, InceptionV3, and ResNet101. These include ResNetV3, MobileNet, Exception and ANN. Prominent trends in the field of oral cancer detection include the use of Inception-V3 and MobileNet architectures Babu et al. In the models they created on intraoral images; they obtained results of 0.7059 Acc with Support Vector Machine (SVM) (Chu et al., 2020) and 0.856 Acc with VG19 (Song et al., 2021).

In this work, we propose a method that can effectively distinguish between benign and malignant oral lesions and also classify their precancerous stages. After preprocessing the color spaces, they created a classifier model with Light Gradient Boosting Machine. The overall performance is promising with 0.9925 Acc, 0.9918 Prec, 0.9931 Rec, 0.9924 F1, and 0.9931 specificity for binary classification and 0.9888 tests Acc, 0.9886 Prec, 0.9792 Rec, 0.9838 F1, and 0.9903 specificity for multiclass classification, outperforming the latest methods for the task of oral cancer classification Goswami et al. (2024). Welikala et al. (2020) proposed architecture for detecting malignant or benign lesions that may lead to oral cancer in the form of image classification with ResNet-101 after Faster R-CNN object Rec. The classification achieved an F1 of 0.8707 for the identification of images containing lesions. Jeyaraj and Samuel (2019) proposed CNN architecture for segmentation and classification of multidimensional hyperspectral images in a regression-based segmented deep learning method. They obtained 0.914 Acc on facial cancer images and 0.945 Acc on 500 images. After extracting features such as energy and entropy from the color components of histopathological images of 1224 people with suspected oral cancer, models were created with SVM and K-Nearest Neighbor to classify them as normal and abnormal. They obtained better results from these images with 0.98 Acc with SVM (Bakare and Kumarasamy, 2021). Warin et al. created models with DenseNet121 and Faster R-CNN on images of 350 oral carcinoma and 350 normal oral mucosa images from retrospective images. In the study, they obtained 0.99 Acc and 0.99 F1 with DenseNet121. In the Faster R-CNN model, they obtained an F1 of 0.7931 (Warin et al., 2021).

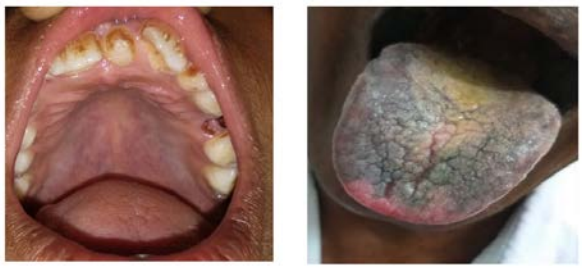
3. Material and Methods

This section describes the dataset, the classification algorithm (CNN) used in the study and the VGG16, VGG19, EfficientNetB7, MobileNetV2, ResNet50, Xception, ConvNextBase transfer learning architectures developed based on this algorithm.

3.1. Data Source

The intraoral dataset used in the study includes color images of oral lesions captured using mobile cameras and intraoral cameras. These images can be used to identify potential oral malignancies through image analysis. The dataset was collected through consultations with doctors from various hospitals and colleges in the Indian state of Karnataka. This dataset contains images of 165 benign lesions and 158 malignant lesions. This dataset is an important resource for training and testing deep learning algorithms and can be used for early detection and diagnosis of oral cancer (Chandrashekar et al., 2021).

Intraoral images consist of digital photographs that can be easily obtained during routine examinations. Considering that almost everyone nowadays has a mobile device, the collection of such images is an extremely easy and accessible process. The creation of this data set requires only a camera and no special medical training. Therefore, the collection of intraoral images is much less time-consuming and requires less expertise compared to the examination of histopathologic images under a microscope. A benign and malign sample of the intraoral images in the study is given in Figure 2.

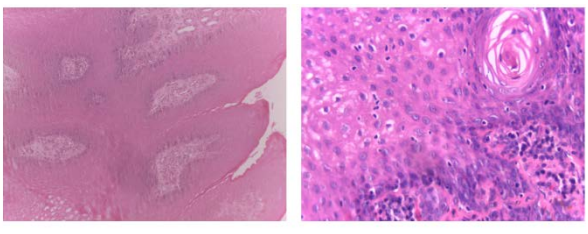


a) Bening b) Malign

Figure 2. An example of intraoral images from the dataset.

Microscopic analysis is conducted by two or three oral pathologists who reach a consensus on the histopathological diagnosis, considering sociodemographic, clinical, and imaging data in conjunction with histopathological findings. The dataset comprises a total of 237 samples (images and metadata) from 77 lesions in 69 patients. The primary objective of this dataset is to provide open access to histopathologic images and metadata of oral potentially malignant disorders and oral cancer, facilitating the testing of machine learning and deep learning models. It also serves educational purposes, such as training dental students and standardizing the diagnostic criteria for oral epithelial dysplasia and squamous cell carcinoma among specialists at the same center. This dataset complements (Ribeiro-de-Assis et al., 2023) which demonstrates that curated demographic and clinical data enhance the performance of AI models in the automated classification of oral cancer.

The histopathologic dataset contains images obtained by examining tissues under a microscope. The images were acquired using high-resolution microscopes to enable precise assessment of oral lesions. These images are labeled to include diseased and healthy tissues. Of particular note, histopathologic images are obtained during the examination of tissue samples under a microscope. This process is performed by a specialized pathologist and is often time-consuming and specialized. Meticulous work is required to correctly assess and label each image. Therefore, these images show in detail the cellular structures and morphologic features of oral lesions (Ribeiro-de-Assis et al., 2023). A benign and malign sample of the histopathologic images in the study is given in Figure 3.



a) Bening b) Malign

Figure 3. An example of histopathologic images from the dataset.

In this study, models for benign and malignant image classification were developed using transfer learning methods. When building the models with the images belonging to the two datasets, 80-20% training and 20% test separation was made. In addition, 20% validation set separation was made in the training set.

3.2. Convolution Neural Network

Deep learning is a multi-layered method that performs traditional processes like feature extraction during the learning phase without requiring additional cost and yields better results, especially with large datasets. Deep learning networks consist of interconnected nodes (neurons) that process and transform data. Each layer extracts increasingly abstract features from the input, enabling the learning of complex patterns (Zavrak and Yilmaz, 2023). Deep neural networks, often referred to as deep architectures, contain multiple hidden layers through which forward and backpropagation occurs to perform the learning process (Başarslan and Kayaalp, 2023).

CNN is a specialized type of neural network designed to process data with a grid-like structure. Convolution is a specific type of linear operation. In simpler terms, CNNs are neural networks that apply convolution instead of a general matrix multiplication in at least one of their layers. A typical CNN architecture consists of five main layers: the input layer, convolutional layer, pooling layer, fully connected layer, and output layer. Figure 4 illustrates a CNN architecture.

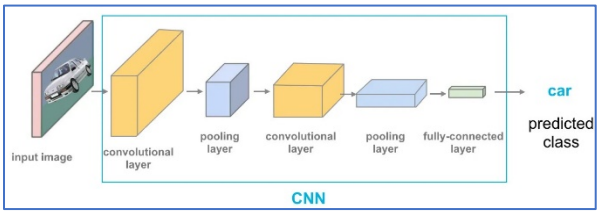


Figure 4. CNN architecture (Bal and Kayaalp, 2023).

In this study, a CNN-based transfer learning method was utilized. Transfer learning reduces the need for labeled data by leveraging pre-trained models that have already learned robust features from large datasets. This is particularly important for medical image analysis, where labeled data can be scarce, especially for rare conditions.

3.3. Transfer Learning

Transfer learning is the process of taking the knowledge gained from training a deep learning model, such as a CNN, and applying it to a different or similar domain. It involves taking a pre-trained model, usually large in size and trained with different datasets, and fine-tuning it to a new dataset or task (Dawud et al., 2019).

Transfer learning can be performed in various ways. A common approach is to freeze some layers in the pre-trained model and fine-tune only the upper layers for the new task. This helps to preserve the general features learned by the model and adapt them to the specific nuances of the new task. Another approach is to use the

pre-trained model as a feature extractor, extracting its upper layers and using the activations in the lower layers as input to a new model specifically designed for the new task (Kabakus and Erdogmus, 2022). The following image illustrates transfer learning (Gilik et al, 2022).

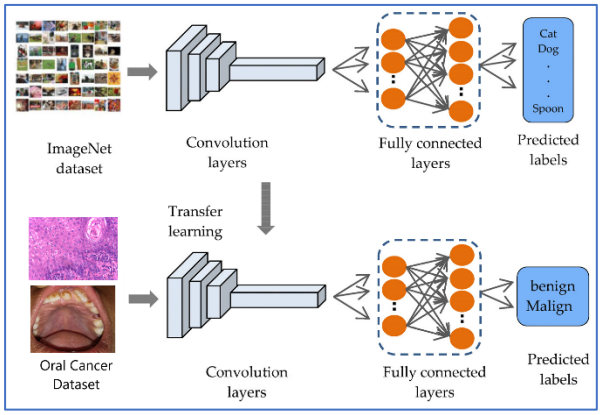


Figure 5: Transfer learning (Gilik et.al., 2022)

The study applied architectures such as VGG16, EfficientNetB7, and ResNet50 to identify the model that performs best in distinguishing between benign and malignant lesions.

3.3.1. EfficientNet

EfficientNet is a family of models developed by Google that offers an innovative approach to scaling CNN. EfficientNet uses Compound Scaling, a simple but effective method to solve the model scaling problem. This method optimizes the trade-off between the width, depth and resolution of the model. EfficientNet-B7 is one of the largest and most powerful models in the EfficientNet family, and the EfficientNet-B7 model consists of 66 layers in total (Koonce, 2021a). EfficientNetB7 was used in this study.

3.3.2. MobileNet

MobileNet is a family of CNN developed by Google that provides high Acc rates in resource-limited environments such as mobile and embedded devices. MobileNet uses depthwise separable convolutions to reduce the number of parameters and computational cost. MobileNetV2 is an improved version of MobileNet and further improves its performance by adding some innovative components. While MobileNet consists of 28 layers, MobileNetV2 consists of 58 layers. MobileNetV2 is used in this study (Dong et al., 2020).

3.3.3. ResNet

ResNet is a CNN architecture developed by He et al. (2016) facilitates the training of deep ANN. ResNet50 is a popular variation of the ResNet family and is a 50-layer deep ANN. ResNet50 is used in this study.

3.3.4. VGGNet

VGG refers to a set of CNN models developed by the Visual Geometry Group from Oxford University in 2014. VGG models build deep layered networks using predominantly small 3x3 convolution filters. This approach increases the depth of the model while at the same time ensuring computational efficiency. VGG19 is a CNN model with a total of 19 layers, while VGG16 has 16 layers (Koonce, 2021b). VGG16 and VGG19 were used in this study.

3.3.5. Xception

Xception is a CNN architecture used in deep learning and image recognition. It was developed by Chollet (2017) and introduced in 2017. Xception is a model based on and extending the Inception architecture. Its name is derived from the phrase "Extreme Inception".

3.3.6. ConvNextBase

ConvNextBase is a CNN architecture developed to improve the performance and efficiency of existing convolutional networks. ConvNextBase has deeper and wider layers and is equipped with advanced data processing capabilities. This model is optimized to achieve high Acc rates, especially on large datasets. The ConvNextBase model is usually characterized by a 12-layer structure. These layers can include convolutional filters of different sizes and types, pooling layers and normalization layers. Through these layers, ConvNextBase can learn deeper and more complex data representations. ConvNextBase was used in this study (Woo et al., 2023).

3.4. Performance metric

Performance evaluation methods such as Acc, Pre, Rec, and F1 are used to assess models developed for classification tasks, including image processing. These metrics are derived from the confusion matrix. The confusion matrix is presented in Table 1 (Öztürk et al., 2022).

Table 1. Confusion Matrix

		Actual Value	
Estimate Value	Positive	Positive	Negative
		TP	FP
	Negative	FN	TN

The Acc of a model is determined by the ratio of correctly classified samples to the total number of samples. A high Acc indicates that the model performs well in making accurate predictions, whereas a low Acc implies that there is room for improvement in the model's performance. The method for computing the Acc value is provided in equation 1 (Öztürk et al., 2022).

$$Acc = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

The Acc performance metric is calculated as in equation 5). Prec refers to the proportion of true positive predictions out of all the samples that the model predicts as positive. A high precision indicates that the number of false positive predictions is low, meaning most of the samples classified as positive are actually positive. Prec is calculated as shown in equation 2 (Kayaalp et al., 2018).

$$Prec = \frac{TP}{TP + FP} \quad (2)$$

Rec is a performance measure used in classification problems. It refers to the rate at which all true positive examples are correctly predicted as positive. Sensitivity, another term for Rec, is crucial for reducing the number of false negatives and minimizing the omission of true positive examples. Rec is calculated as shown in equation 3 (Öztürk et al., 2022).

$$Rec = \frac{TP}{TP + FN} \quad (3)$$

F1 is a measure often used in classification tasks such as information retrieval and machine learning. It provides a balance for evaluating the performance of a model, taking into account both Prec and Rec. Mathematically, F1 is defined as the harmonic mean of Prec and Rec. The F1 is given in equation 4 (Kayaalp et al., 2018).

$$F1 = 2 * \frac{Prec * Rec}{Prec + Rec} \quad (4)$$

Prec measures the ratio of correct positive predictions to total predicted positives. It aims to quantify how accurate positive predictions are. Prec, also known as sensitivity or true positive rate. It measures the ratio of correct positive predictions to total true positives. It shows how well the model is able to detect all positive samples.

The F1 balances these two metrics, considering Prec and Rec, and provides a single value. It is particularly useful when the class distribution is unbalanced or when false positives and false negatives lead to different results. The F1 is especially important when both Prec and Rec are important when evaluating tasks such as medical diagnostics or information retrieval systems (Kayaalp et al., 2018).

4. Results and Discussions

In this study, models were created separately on histopathologic and intaoral images using transfer learning methods. The models were run on Google Colab Pro with Python language and Tensorflow. During the model building phase for the two separate datasets, 80%-20% training and 20% test separation was made. In addition, 20% validation set separation was made in the train set. Callbacks such as early stopping and reduceonLr were used to ensure that all models were not over fit. In Table 2, detailed parameters of all models are given in the order of model creation.

The Acc, F1, Rec, Prec, AUC results of the performance of the models created on histopathological images are shown in Table 3. According to Table 3, in terms of Acc, the model created with ResNet50 is well ahead of the other models. ResNet50 is followed by EfficientNetB7, VGG19, Xception, ConvNextBase, MobilNetV2 and VGG16. While F1 and AUC also yielded similar results, EfficientNetB7 came first in terms of Prec. Similarly, ResNet50 is ahead in the Rec metric.

Table 2. Detailed Model's Parameters

Parameters	Values
Input Shape	64x64
Transfer Learning layer	EfficientNetB7, ResNet50, VGG16, VGG19, Xception, ConvNExtBase, MobileNetV2
Pooling Layer	Global Averege Pooling 2D
Dense Layer 1 Unit	256
Dense Layer 1 Activaton	ReLU
Dense Layer 2 Unit	2

Table 3. Histopatalogical Dataset Results

Models	Acc	F1	Rec	Prec	AUC
ConvNextBase	0.7083	0.7742	0.7742	0.7742	0.80
EfficientNetB7	0.7500	0.8333	0.7317	0.9677	0.84
MobileNetV2	0.6875	0.7761	0.7222	0.8387	0.73
ResNet50	0.8125	0.8525	0.8667	0.8387	0.85
VGG16	0.6875	0.7826	0.7105	0.8710	0.71
VGG19	0.7500	0.8182	0.7714	0.8710	0.71
Xception	0.7083	0.7742	0.7742	0.7742	0.71

Figure 6 shows the relative ROC of the ResNet50 transfer learning model that performs best on histopathological images. When evaluated with the ROC curve and the AUC value (0.85) shown in Figure 6, it can be seen that the classification success is quite good. The fact that the AUC is close to 1 indicates that the model successfully discriminates between positive and negative classes. However, considering the F1 value and the accuracy metrics, we can say that the overall performance of the model is balanced and satisfactory.

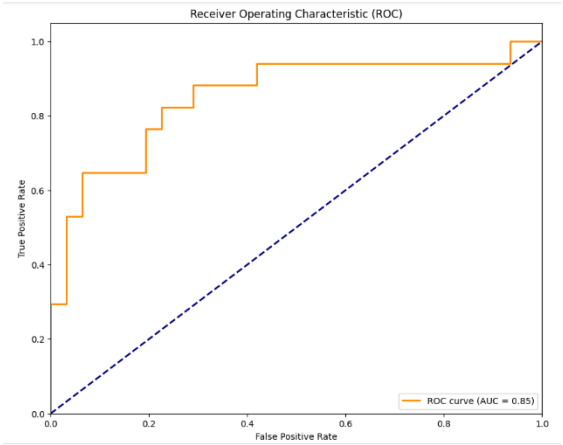


Figure 6. ROC of the ResNet50, which gives the best result in the histopathologic images.

The values of the confusion matrices of the models created with histopathological data are tabulated in Table 4. Acc, F1, Rec, Prec, AUC results of the performance of the models created on intraoral images are shown in Table 5. According to Table 5, in terms of Acc, the model created with ConvNextBase is well ahead of the other models. The next order is VGG16, ResNet50,

VGG-19, Xception, MobilNetV2. Similarly, ConvNextBase is ahead in F1 and AUC, while ResNet50 is ahead in Prec and VGG-16 is ahead in Rec. Figure 7 shows the relative ROC of the ConvNextBase transfer learning model that performs best on intraoral images.

Table 4. Confusion Matrix Values of Histopathological images

	TN	FP	FN	TP
ConvNextBase	24	7	7	10
MobileNetV2	26	5	10	7
EfficientNetB7	30	1	11	6
XCEPTION	24	7	7	10
ResNet50	28	3	4	15
VGG19	24	7	8	9
VGG16	27	4	11	8

Table 5. Intraoral Dataset Results

Models	Acc	F1	Rec	Prec	AUC
ConvNextBase	0.84	0.80	0.80	0.8387	0.7647
EfficientNetB7	0.82	0.7077	0.7246	0.7143	0.7353
MobileNetV2	0.68	0.6308	0.6364	0.6563	0.6176
ResNet50	0.80	0.7538	0.7778	0.7368	0.8235
VGG16	0.80	0.7937	0.8621	0.7353	0.82
VGG19	0.6769	0.6866	0.6970	0.6765	0.70
Xception	0.6769	0.6441	0.7600	0.5588	0.65

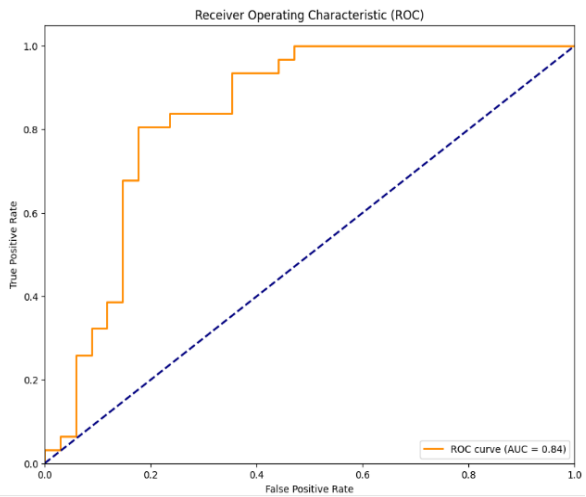


Figure 7. ROC of the ConvNextBase, which gives the best result in the histopathologic images.

In Figure 7, the ROC curve and the AUC value of the model are shown as 0.84. An AUC close to 1 indicates that the classification performance of the model is good. An AUC value of 0.84 indicates that the model is successful in balancing the true positive and false positive rates. In terms of F1 value and accuracy, it can be said that the model successfully discriminates between both positive and negative classes and misclassifications are kept to a minimum. These results indicate that the overall performance of the model is balanced and its ability to correctly identify positive classes is satisfactory.

The values of the confusion matrices of the models created with intraoral data are tabulated in Table 6.

Table 6. Confusion Matrix Values of Intraoral Images

	TN	FP	FN	TP
ConvNextBase	26	8	5	26
MobileNetV2	14	16	10	25
EfficientNetB7	26	5	5	25
XCEPTION	19	15	6	25
ResNet50	25	9	10	21
VGG19	25	9	4	27
VGG16	23	11	10	21

The performance of the models based on intraoral and histopathological image data was calculated using the TN, FP, FN, TP values shown in Table 5. Detailed information about the calculations is described under the title Performance Metrics.

This study demonstrated that CNN-based transfer learning models can be effectively applied to both histopathological and intraoral images for disease detection, even without the involvement of a pathologist. This finding highlights the potential of AI-based decision support systems to assist medical professionals in diagnosing diseases with high accuracy. By utilizing CNN-based transfer learning methods, clinicians and physicians can gain reliable insights from diverse image sources, highlighting the adaptability and scalability of AI in medical imaging. Furthermore, this study highlights the importance of selecting appropriate transfer learning models for different image types and sources. The differences in performance between histopathological and intraoral images indicate that there is no one-size-fits-all solution when it comes to applying AI in healthcare. Instead, the choice of model should be carefully tailored to the specific characteristics of the data.

Comparison of the models that gave the best results in the study with similar data sets of histopathologic and intraoral data of oral cancer is given in Table 7 and Table 8, respectively.

According to Table 7, the best model in our experiments on histopathologic image data competes with the literature with similar content. Bayesian and quadratic discriminant analysis gave better results than our model. According to Table 8, our model built on the intraoral dataset competes with the literature with similar content. The Roboflow model differs from our work as it performs object detection.

Table 7. Compare to Similar Histopathologic Oral Cancer Data Set

References	Model	Acc	F1
(Welikala et al., 2020).	ResNet101	0.7500	0.7830
(Muthu et al., 2012)	Bayesian	0.9075	None
(Chang S.W. et al.2013)	SVM	0.75	None
(Lu et al. 2017)	quadratic discriminant analysis	0.88	None
The present	ResNet50	0.8125	0.8525

Table 8. Compare to Similar Intraoral Oral Cancer Data Set

References	Model	Acc	F1
(Chu et al., 2020).	SVM	0.7059	None
(Dinesh, et.al.,2023)	Roboflow	0.8947	None
(Song, et.al, 2021)	VGG19	0.856	None
The present	ConvNextBase	0.84	0.80

Looking ahead, the potential for developing hybrid models that integrate different transfer learning techniques offers exciting opportunities. By combining methods designed for different image types or sources, it may be possible to achieve even more robust and accurate diagnostic support systems. These hybrid models could increase the flexibility of AI applications, making them more adaptable to a variety of medical scenarios. With further work, this approach could help refine the role of AI in healthcare, providing more comprehensive tools for early diagnosis and treatment, particularly in settings with limited access to specialist expertise.

Author Contributions

The percentages of the authors' contributions are presented below. The authors reviewed and approved the final version of the manuscript.

	K.G.	M.S.B.
C	50	50
D	50	50
S	50	50
DCP	50	50
DAI	50	50
L	50	50
W	50	50
CR	50	50
SR	50	50

C=Concept, D= design, S= supervision, DCP= data collection and/or processing, DAI= data analysis and/or interpretation, L= literature search, W= writing, CR= critical review, SR= submission and revision.

Conflict of Interest

The authors declared that there is no conflict of interest.

Ethical Consideration

Ethics committee approval was not required for this study because of there was no study on animals or humans.

References

Babu PA, Rai AK, Ramesh JVN, Nithyasri A, Sangeetha S, Kshirsagar PR, Rajendran A, Rajaram A, Dilipkumar S. 2024. An explainable deep learning approach for oral cancer detection. J Electr Eng Technol., 19: 1837–1848.

Bakare YB, Kumarasamy M, 2021. Histopathologıcal image analysis for oral cancer classification by support vector machine. Int J Adv Signal Image Sci, 7: 1–10.

Bal F, Kayaalp F, 2023. A novel deep learning-based hybrid method for the determination of productivity of agricultural products: apple case study. IEEE access, 11:7808–7821.

Başarslan MS, Kayaalp F, 2023. MBi-GRUMCONV: A novel Multi Bi-GRU and Multi CNN-Based deep learning model for social media sentiment analysis. J Cloud Comput, 12: 1-16.

Chandrashekar HS. Geetha A, Kiran S, Murali MS, Dinesh BR, Nanditha, 2021. Oral images dataset, URL: https://data.mendeley.com/datasets/mhjyrn35p4/2 (accessed date 13 April, 2024)

Chang SW, Abdul-Kareem S, Merican AF, Zain RB. 2013. Oral cancer prognosis based on clinicopathologic and genomic markers using a hybrid of feature selection and machine learning methods. BMC Bioinform, 14: 1-15.

Chollet F. 2017. Xception: Deep Learning with depthwise separable convolutions. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, pp: 1800–1807.

Chu CS, Lee NP, Adeoye J, Thomson P, Choi SW. 2020. Machine learning and treatment outcome prediction for oral cancer. J Oral Pathol Med, 49: 977-985.

Dawud AM, Yurtkan K, Oztoprak H, 2019. Application of deep learning in neuroradiology: brain haemorrhage classification using transfer learning. Comput Intell Neurosci, 2019:

4629859

de Lima LM, de Assis MCFR, Soares JP, Grão-Velloso TR, de Barros LAP, Camisasca DR, Krohling RA, 2023. Importance of complementary data to histopathological image analysis of oral leukoplakia and carcinoma using deep neural networks. Intell Med 3: 258–266.

Dinesh Y, Ramalingam K, Ramani P, Deepak RM, 2023. Machine learning in the detection of oral lesions with clinical intraoral images. Cureus 15:e44018

Dong K, Zhou C, Ruan Y, Li Y, 2020. MobileNetV2 model for image classification. 2nd International Conference on Information Technology and Computer Application (ITCA), December 18-20, Guangzhou, China, pp: 476–480.

Gilik A, Ogrenci AS, Ozmen A. Air quality prediction using CNN+LSTM-based hybrid deep learning architecture. Environ Sci Pollut Res 29: 11920–11938 2022.

Goswami B, Bhuyan MK, Alfarhood S, Safran M, 2024. Classification of oral cancer into pre-cancerous stages from white light images using LightGBM algorithm. IEEE Access, 12: 31626–31639.

He K, Zhang X, Ren S, Sun J. 2016. Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, 17-20 March, NV, USA, pp: 770-778.

Jeyaraj PR, Samuel N, 2019. Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm. J Cancer Res Clin Oncol, 145: 829–837.

Kabakus AT, Erdogmus P, 2022. An experimental comparison of the widely used pre-trained deep neural networks for image classification tasks towards revealing the promise of transfer-learning. Concurr Comput Pract Ex, 34: e7216.

Kayaalp F, Basarslan MS, Polat K, 2018. TSCBAS: A novel correlation based attribute selection method and application on telecommunications churn analysis. International Conference on Artificial Intelligence and Data Processing (IDAP), 28-30 September, Malatya, Tükiye, pp: 1–5.

Koonce B. 2021a EfficientNet. In convolutional neural networks with swift for tensorflow. Apress Berkeley, CA, USA, 1st ed., pp: 109–123

Koonce B. 2021b. VGG Network. convolutional neural networks with swift for tensorflow. Apress Berkeley, CA, USA, 1st ed., pp: 35-50

Lu C, Lewis Jr, JS, Dupont WD, Plummer Jr, WD, Janowczyk A, Madabhushi, A. 2017. An oral cavity squamous cell carcinoma quantitative histomorphometric-based image classifier of nuclear morphology can risk stratify patients for disease-specific survival. Mod Pathol, 30: 1655-1665.

Muthu Rama Krishnan M, Shah P, Chakraborty C. 2012 Statistical analysis of textural features for improved classification of oral histopathological images. J Med Syst 36: 865–881.

Öztürk T, Turgut Z, Akgün G, Köse C. 2022. Machine learning-based intrusion detection for SCADA systems in healthcare. Netw Model Anal Health Inform Bioinform, 11: 47.

Ribeiro-de-Assis MCF, Soares JP, de Lima LM, de Barros LAP, Grão-Velloso TR, Krohling RA, Camisasca DR. 2023 NDB-UFES: An oral cancer and leukoplakia dataset composed of histopathological images and patient data. Data Brief, 48: 109128.

Song B, Sunny S, Li S, Gurushanth K., Mendonca P, Mukhia N, Liang R, 2021. Bayesian deep learning for reliable oral cancer image classification. Biomed Opt Express, 12: 6422-6430.

Warin K, Limprasert W, Suebnukarn S, Jinaporntham S, Jantana P, 2021. Automatic classification and detection of oral cancer in photographic images using deep learning algorithms. J Oral Pathol Med, 50: 911–918.

Welikala, RA, Remagnino P, Lim JH, Chan CS, Rajendran S, Kallarakkal TG, Zain RB, Jayasinghe RD, Rimal J, Kerr AR, Amtha R, Patil K, Tilakaratne WM, Gibson J, Cheong SC, Barman SA. 2020. Automated detection and classification of oral lesions using deep learning for early detection of oral cancer. IEEE Access, 8: 132677–132693.

Woo S, Debnath S, Hu R, Chen X, Liu Z, Kweon IS, Xie, S. 2023. ConvNeXt V2: Co-Designing and scaling convnets with masked autoencoders. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 11-15 Vancouver, BC, Canada, pp: 16133–16142.

Zavrak S, Yilmaz S, 2023. Email spam detection using hierarchical attention hybrid deep learning method. Expert Syst Appl, 233: 120977.