International Journal of Emerging Research in Engineering, Science, and Management
Vol. 4, Issue 2, pp.07-13, Apr-Jun 2025.
www.ijeresm.com eISSN – 2583-4894

Oral Lesions Classification using EfficientNet Transfer Learning Model

[1]Devika G, [2]Asha Gowda Karegowda

[1]Associate Professor, Dept. of CSE, Government Engineering College, Kushalnagar, India. sgdevika@gmail.com, ORCID: 0000-0002-2509-2867
[2]Associate Professor, Dept. of MCA, Sigganga Institute of Technology, Tumkur, Karnataka. ashagk@sit.ac.in

Abstract: Due to their wide variety of diseases, oral lesions present a substantial diagnostic problem. This research uses deep learning techniques, particularly the EfficientNetB7 model, to present an automated categorisation analysis of oral lesions. The study divides lesions into benign and malignant categories using the Oral Lesions: Cancer Detection Dataset, comprising 2270 high-resolution pictures. Known for its effectiveness in processing large-scale image collections, the EfficientNetB7 architecture is employed in this work. The model successfully distinguishes between benign and malignant tumors with an exceptional accuracy rate of 99.12%. The study highlights the diagnostic dependability of the model by analyzing its performance, including metrics for sensitivity, specificity, and accuracy. Moreover, the study investigates how interpretable the model's predictions are, emphasizing essential aspects that support its decision-making process.

Keywords: Benign, EfficientNet-B7, Malignant, Oral Lesions, Transfer Learning

1 Introduction

The field of oral pathology is marked by the intricate and often perplexing nature of lesions that manifest within the oral cavity. Oral lesions encompass a broad spectrum of abnormalities, ranging from benign growths to potentially malignant neoplasms, presenting a diagnostic challenge for healthcare professionals. The timely and accurate identification of these lesions is pivotal for effective treatment planning, prognosis, and overall patient care. In recent years, advancements in artificial intelligence and machine learning techniques have revolutionized medical image analysis, offering the potential to augment traditional diagnostic approaches. This paper delves into the realm of automated classification of oral lesions, leveraging the cutting-edge capabilities of the EfficientNetB7 model. The endeavour is rooted in recognising the critical need for robust and efficient diagnostic tools that can assist healthcare professionals in navigating the complexity of oral pathology. The Oral Lesions: Malignancy Detection Dataset under scrutiny emerges as a cornerstone of this research.

Comprising 2270 meticulously curated images, this dataset encapsulates the diverse morphological variations in oral lesions. From innocuous benign growths to potentially life-altering malignant entities, the dataset mirrors clinicians' nuanced challenges in distinguishing between these two classes. The images encapsulate the rich heterogeneity inherent in oral lesions, reflecting the amalgamation of histopathological variations and clinical presentations encountered in routine clinical practice [1]. At the heart of the investigation lies the EfficientNetB7 model, a convolutional neural network renowned for handling large-scale and diverse image datasets. The choice of this model is deliberate, aiming to harness its efficiency in extracting intricate features and patterns from the complex oral lesion images present in the dataset. The goal is to achieve high accuracy and ensure the model's robustness in generalising to novel cases, a critical aspect for any diagnostic tool seeking real-world applicability [2][3].

2 Literature Survey

This section briefly reviews prior research on the Classification of oral lesions and compares the initial investigations. The study by G. Tanriver et al. investigates computer vision and deep learning in diagnosing oral cancer, demonstrating the potential for automated systems to provide real-time, cost-effective, and non-invasive diagnostic options [4]. The author used the EfficientNet-B4 model and achieved an accuracy rate of 86%. The study by D. K. Das et al. presented a two-step method for synthesising pictures of oral histology: segmentation using a 12-layered deep convolutional neural network and keratin pearl recognition using texture-based feature-trained random forests [5]. The algorithm's detection accuracy is 96.88%. The study by Y. Liu et al. developed a convolutional neural network (CNN) model to identify suspicious regions in OED whole-slide pathology images [6]. The model outperformed UNet++ in accuracy, precision, and segmentation metrics, with the best-performing model achieving 93.3% and 90.9% accuracy.

R. Gomes et al. carried out this study to turn six clinical representation types of oral lesion images into elementary lesions through automatic identification methods by constructing a CDC-based model and defining four styles: ResNet-50, VGG16, InceptionV3, Xception [7]. Then there's evaluation for that model. After optimizing the parameters, the model's average accuracy for all six types of lesions was 95.09%. One of the following goals is to examine trained layers to distinguish between benign, maybe malignant, and malignant tumours.

The study by K. Warin et al. develops an automated classification and detection model for oral cancer screening using convolutional neural network (CNN) deep learning methods [8]. The 700 clinical oral photos in the model are split between 350 pictures showing oral squamous cell carcinoma and 350 images showing the normal oral mucosa. 99% accuracy, recall, F1 score, sensitivity, specificity, and area under the receiver operating characteristic curve were attained with the DenseNet121 model. The accuracy, recall, F1 score, and area under the precision-recall curve of the quicker R-CNN model were all 76.67%. The study by H. Lin et al. presented a deep learning algorithm-based smartphone-based image diagnostic technique for the automated identification of oral illnesses [9]. The method proposes a deep learning network (HRNet) for oral cancer diagnosis, builds a medium-sized oral dataset, and employs a basic centered rule image-capturing technique. On 455 test photos, the method yields sensitivity of 83.0%, specificity of 96.6%, accuracy of 84.3%, and F1 of 83.6%. The technique may be used to diagnose primary oral cancer.

The study by M. Aubreville et al. presented a novel automated technique for diagnosing OSCC using deep learning technology, using CLE pictures, which is shown and assessed [10]. With an AUC of 0.96 and a mean accuracy of 88.3%, the system beats the state of the art in CLE image identification compared to existing textural feature-based machine learning algorithms. The study by D. F. D. dos Santos et al. provides a technique to localise and fine-tune the segmentation of tumour patches originating from the mouth cavity in H&E-stained histological whole slide pictures using a fully convolutional neural network [11]. The method identifies tissue areas using colour attributes from the HSV colour model. These regions are then converted into the CIE L*a*b* colour model and divided into image-patches. Utilising a WSI dataset including sentinel lymph nodes with breast cancer metastases and tissue samples from oral squamous cell carcinoma, the technique was used. According to experimental assessments, the procedure attained a sensitivity of 92.9%, a specificity of 98.4%, and an accuracy of 97.6%.

3 Proposed Methodology

The Descriptive Statistics section introduces the research object and contents of this chapter. This section provides the dataset on which the model was trained and discusses the preparatory methods to obtain that data in great depth. Fig. 1 shows a simplified flow diagram of the proposed research methodology. The accompanying visual guide aims to broadly cover the sequential methodologies and procedures employed in the study. It comprehensively describes the data post-processor methods, including experimental designs and processing flowcharts.

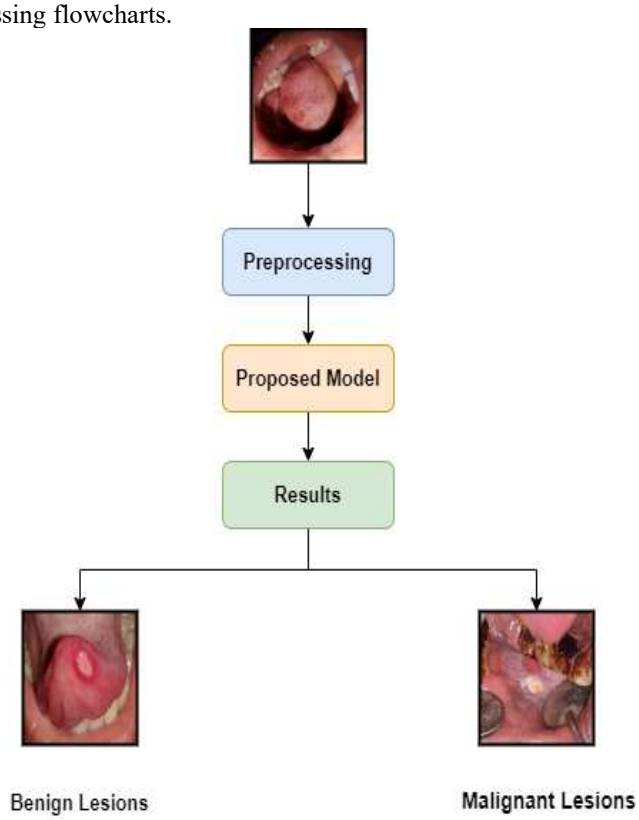

Fig. 1. Oral Lesions Classification Using an Efficient Net Transfer Learning Model Architecture

This study relies entirely on the comprehensive Oral Lesions: Malignancy Detection Dataset [12] to advance the automated classification of oral lesions. Everything in the research depends on that meticulously curated data set, which provides a wide variety of photos for any oral lesion, benign or malignant.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 4, Issue 2, pp.07-13, Apr-Jun 2025.
www.ijeresm.com eISSN – 2583-4894

Initially, the dataset contains 165 images of benign lesions and 158 images of malignant ones. It recognises that training strong models depends on dataset size, so a process of expansion is used to augment the original dataset, until there are some 2270 images in all (Fig. 2). The augmentation method involves varying the original images through operations such as rotation, flip, and zoom. It adds another dimension to the dataset, supplying a greater range of viewpoints for the model to capture more features over a broader spectrum. To ensure the model's effectiveness and prevent it from overtraining, the dataset is carefully divided. An 80-20 split is used: 80% of the data will be used for training, while 20% will be set aside for validation. This separates things, so a large portion of the training set is left to be learned between training sessions. Still, at least one complete subset remains independent to watch the performance and prevent overfitting its auto head.

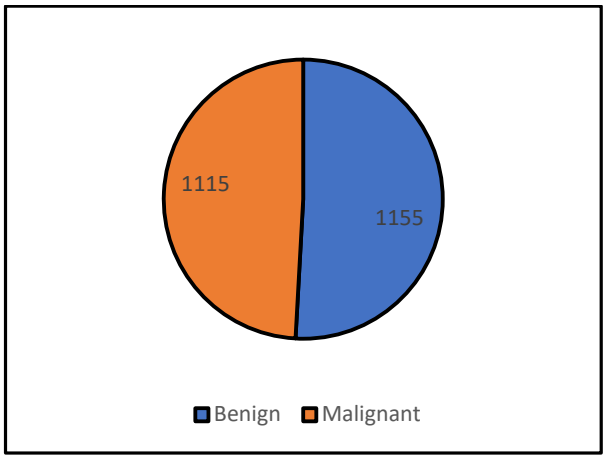

Fig. 2. Dataset Distribution

After the rigorous training phase, a special testing subset is used to test the model's ability. In this study, 30% of the augmented data set or 681 images, is used for testing purposes (Fig. 3). This is the most crucial part for assessing how well the model works in new circumstances, indicating its generalisation abilities and accuracy when encountering real-world sets of events. The incorporation and systematic separation of the benign, malignant, and data and the painterly augmentation process all contribute to making the system robust.

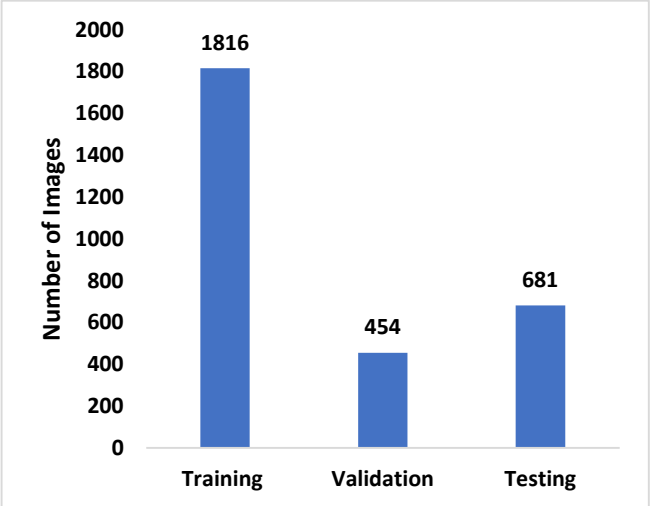

Fig. 3. Dataset Splitting

3.1. Preprocessing

The pre-processing steps for the EfficientNetB7 model are described in this study to enhance its ability to discriminate between benign and malignant oral lesions. The parameter "target_size" is set to (244,244). This ensures that input dimensions are uniform, benefiting the model's performance. The parameter "color_mode" is set to 'rgb'. This means that any image is considered in the RGB colour space, allowing for complex information to be included in the model. "Categorical" as the parameter for "class_mode." Therefore, with estimated probability distributions across defined categories, the model can assist in the differentiation between benign and malignant lesions in training. The "batch_size" was fixed at 4 to make memory usage efficient and converge faster. "To the parameter is set False shuffle." Proceeding in this manner, the order of the images in the dataset is held fixed during training.

This choice may be motivated by a desire to keep the dataset's structure intact and to ensure a uniform presentation of consecutive images. Whether this shuffling is done or not can affect the model's ability to learn temporal dependencies or spatial patterns in the dataset [13]. Overall, the preprocessing steps in this study enhance the model's capacity to differentiate effectively between benign and malignant oral lesions.

3.2. Model

The study builds on the neural network model EfficientNetB7, starting from the base model and adding more layers to distinguish between benign and malignant oral lesions. The model's shape begins with the input and is configured at (224, 224, 3). This corresponds to images of size 224 x 224 pixels and three-colour channels (Red, Green, Blue). The base model is initiated through the Keras API of TensorFlow, excluding the fully connected layer and providing flexibility in customising model outputs. The model uses pre-trained weights from the "ImageNet" dataset, which may help it generalise better to the oral lesion dataset. The final model is created by stacking the base EfficientNetB7 model with a Sequential model, acting as a feature extractor and extracting meaningful features from input images [14]. A Dense layer with two neurons and a sigmoid activation function is added. The output layer is thus configured for binary classification (Fig. 4). This configuration allows the model to return the probability that the input image belongs to either the benign or malignant class. Reversing the layers in the base model enables the retention of pre-trained weights during training and enhances the ability of the architecture to perform transfer learning.

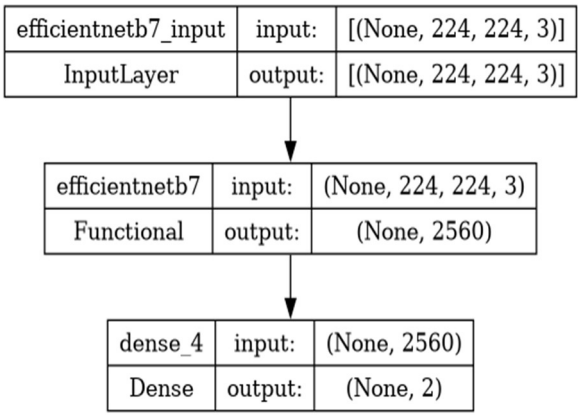

Fig. 4. Model Architecture

4 Results

This study centres on assessing the performance and efficacy of a deep learning model. The paper thoroughly examines the process involved in training and evaluating the model. The training phase of a model is crucial in realising its potential for accurate classification of oral lesions. The training regimen of this model in the experiment is composed of 5 epochs. That means, within the epoch, the whole training dataset is exposed five times, and so on. This is an iterative process, as the model gradually refines internal parameters. It learns to differentiate subtle features in the images, improving its prediction ability. For a batch size set at 4, each epoch has 454 batches, representing how partitioned samples are handled efficiently in the augmented dataset. The relatively small batch size benefits the model by allowing it to update its weights more often. This might bring about faster convergence and enhance model responsiveness towards training data. However, small batch sizes require caution since they can introduce some jitter into weight updates. Thus, factors must be considered to balance the process's efficiency and stability.

Over the course of five epochs of training, the model delivers substantial achievements in terms of accuracy and loss metrics The training set amasses an impressive accuracy of 98 35% On average, during training, the model correctly classifies nearly 98.35% of the images within this dataset This high training accuracy means that the model effectively captures patterns and features in training images Simultaneously, the training loss is reported as 0.0480: A lower index reflects more consistency with real-world statistics, whereas a high number suggests theoretical contradictions in those predictions In this context, low training loss signifies that the model is good at reducing errors through its training process This way it is free to adjust and fine-tine internal parameters During training phase, validation indicators offer an essential gauge of model generalization to unseen instances of data The validation metrics show an impressive accuracy of 98.68% (Fig. 5) This seems to indicate an ability on the model 's part for generalization to other unseen instances of data outside its training set This high validation accuracy implies that the model generalizes well to new data instances, helping to validate its commercial potential At the same time, the validation loss is documented as 0.0604 (Fig. 6) Although slightly higher than the training loss this is still an indicator that in validation set, disparity between model predictions and objective truths is relatively small At no time during either the training or validation phases is there appreciable difference between loss values, and they remain low throughout This indicates the robustness of the model as it effectively captures those subtle features which set benign and malignant growths apart.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 4, Issue 2, pp.07-13, Apr-Jun 2025.
www.ijeresm.com eISSN – 2583-4894

Once the training phase is complete, the model's performance is usually evaluated in detail. The model surpasses previous training and validation runs and reaches 99.12% accuracy during testing. Loss during testing is 0.0422, suggesting little difference was induced between model predictions and genuine values. These complex measurements help build confidence in the model's ability to distinguish between benign and malignant oral lesions and its process robustness. Table I presents the summarised performance parameters, vividly displaying this model's powerful capability for oral lesion classification.

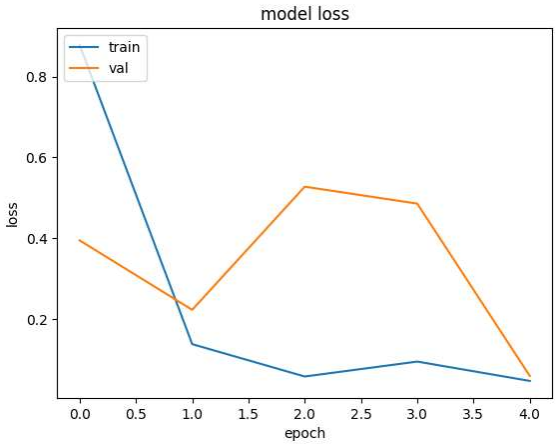

Fig. 5. Loss in Training and Validation Stages

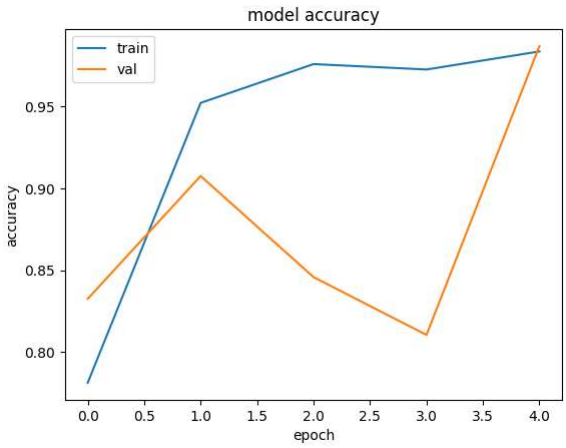

Fig. 6. Accuracy in Training and Validation Stages

Table 1. Performance Matrix

Metric	Value
Accuracy	0.9912
Precision	0.9914
Recall (Sensitivity)	0.9914
F1 Score	0.9914
Specificity	0.9909
False Positive Rate	0.0091
False Negative Rate	0.0086

The comparison table (Table 2) summarises prior studies alongside the proposed model, highlighting the differences in dataset types, model architectures, and performance metrics. It reveals that while several earlier approaches, such as EfficientNet-B4, ResNet-50, and DenseNet121, achieved promising results, the proposed EfficientNet-B7 model outperforms them regarding accuracy, precision, recall, F1 score, and specificity. Notably, the proposed method attains a high testing accuracy of 99.12% and balanced precision and recall values, indicating strong generalisation ability and reliable classification of benign and malignant lesions. This comparison demonstrates the effectiveness and superiority of the proposed approach in the domain of oral lesion classification.

Table 2. Performance Analysis

Study / Model	Dataset Size / Type	Model Used	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	Specificity (%)
Tanriver et al. [4]	Not specified	EfficientNet-B4	86	-	-	-	-
Das et al. [5]	Histology images	DCNN + Random Forest	96.88	-	-	-	-
Liu et al. [6]	OED WSI images	CNN	93.3	-	-	-	-
Gomes et al. [7]	6 types of oral lesion images	ResNet-50, etc.	95.09	-	-	-	-
Warin et al. [8]	700 clinical photos	DenseNet121	99	-	-	-	-
Lin et al. [9]	455 smartphone images	HRNet	84.3	-	83	83.6	96.6
Aubreville et al. [10]	CLE images	Deep Learning Model	88.3	-	-	-	-
Dos Santos et al. [11]	WSI images	FCNN	97.6	-	92.9	-	98.4
Proposed Model	2270 augmented images	EfficientNet-B7	99.12	99.14	99.14	99.14	99.09

5 CONCLUSIONS

This study is an advance in automated classification of oral lesions, and it takes a carefully screened data set and the power of EfficientNetB7 to do so. Trained over five epochs, achieved remarkable success in learning, with a batch size of 4, the training accuracy was 98.35%, and validation accuracy 98.68%. These figures reflect how the model is good at learning from both training and validation sets. It indicates its potential for clinical application. Moreover, the model's performance was further verified using a special test set, where it was found that it could achieve an astonishing accuracy rate of 99.12% and a low loss of 0.0422. This system is robust since the testing accuracy was significantly higher than the training accuracy. This is an essential requirement for deploying such systems in clinical practice.

Finally, the study was successful because the dataset itself had been thoroughly pre-processed: image resizing, colour representation and category division all align with the requirements for the EfficientNetB7 model. Checking every part of the pre-trained EfficientNetB7 model transferred to the ImageNet dataset from which it was trained previously made it more able to extract detailed information from oral lesion images. This study's findings offer promising prospects for applying automatic classification systems to medical image analysis. An automated classification system as accurate and robust as demonstrated here will be of great value in helping medical workers identify the nature of oral lesions early and accurately. Integrating such technology into clinical workflows can streamline diagnostic processes, expedite treatment decisions, and improve patient outcomes. While the study showcases impressive results, ongoing research efforts should continue to refine and expand upon these findings. Exploring larger and more diverse datasets, investigating the interpretability of the model's decisions, and conducting prospective clinical validations can further solidify the practical applicability of the presented automated classification system. This study lays a foundation for future endeavours in artificial intelligence and oral pathology, paving the way for enhanced diagnostic precision and patient care.

FUNDING INFORMATION

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

ETHICS STATEMENT

This study did not involve human or animal subjects and, therefore, did not require ethical approval.

STATEMENT OF CONFLICT OF INTERESTS

The authors declare no conflicts of interest related to this study.

LICENSING

REFERENCES

[1] D. C. G. De Veld, M. Skurichina, M. J. H. Witjes, R. P. W. Duin, H. J. C. M. Sterenborg, and J. L. N. Roodenburg, "Clinical study for classification of benign, dysplastic, and malignant oral lesions using autofluorescence spectroscopy," *Journal of Biomedical Optics*, vol. 9, no. 5, p. 940, Jan. 2004, doi: 10.1117/1.1782611.

[2] 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.

[3] S. Warnakulasuriya, Newell. W. Johnson, and I. Van Der Waal, "Nomenclature and classification of potentially malignant disorders of the oral mucosa," *Journal of Oral Pathology and Medicine*, vol. 36, no. 10, pp. 575–580, Jul. 2007, doi: 10.1111/j.1600-0714.2007.00582.x.

[4] G. Tanriver, M. S. Tekkesin, and O. Ergen, "Automated detection and classification of oral lesions using deep learning to detect oral potentially malignant disorders," *Cancers*, vol. 13, no. 11, p. 2766, Jun. 2021, doi: 10.3390/cancers13112766.

[5] D. K. Das, S. Bose, A. K. Maiti, B. Mitra, G. Mukherjee, and P. K. Dutta, "Automatic identification of clinically relevant regions from oral tissue histological images for oral squamous cell carcinoma diagnosis," *Tissue and Cell*, vol. 53, pp. 111–119, Jun. 2018, doi: 10.1016/j.tice.2018.06.004.

[6] Y. Liu, E. Bilodeau, B. Pollack, and K. Batmanghelich, "Automated detection of premalignant oral lesions on whole slide images using convolutional neural networks," *Oral Oncology*, vol. 134, p. 106109, Sep. 2022, doi: 10.1016/j.oraloncology.2022.106109.

[7] R. Gomes *et al.*, "Use of Artificial Intelligence in the Classification of Elementary Oral Lesions from Clinical Images," *International Journal of Environmental Research and Public Health*, vol. 20, no. 5, p. 3894, Feb. 2023, doi: 10.3390/ijerph20053894.

[8] K. Warin, W. Limprasert, S. Suebnukarn, S. Jinaporntham, and P. Jantana, "Automatic classification and detection of oral cancer in photographic images using deep learning algorithms," *Journal of Oral Pathology and Medicine*, vol. 50, no. 9, pp. 911–918, Aug. 2021, doi: 10.1111/jop.13227.

[9] H. Lin, H. Chen, L. Weng, J. Shao, and J. Lin, "Automatic detection of oral cancer in smartphone-based images using deep learning for early diagnosis," *Journal of Biomedical Optics*, vol. 26, no. 08, Aug. 2021, doi: 10.1117/1.jbo.26.8.086007.

[10] H. Lin, H. Chen, L. Weng, J. Shao, and J. Lin, "Automatic detection of oral cancer in smartphone-based images using deep learning for early diagnosis," *Journal of Biomedical Optics*, vol. 26, no. 08, Aug. 2021, doi: 10.1117/1.jbo.26.8.086007.

[11] D. F. D. D. Santos, P. R. De Faria, B. A. N. Travençolo, and M. Z. D. Nascimento, "Automated detection of tumor regions from oral histological whole slide images using fully convolutional neural networks," *Biomedical Signal Processing and Control*, vol. 69, p. 102921, Jul. 2021, doi: 10.1016/j.bspc.2021.102921.

[12] M. Gobara, "Oral lesions: Malignancy Detection Dataset," Dec. 31, 2023. https://www.kaggle.com/datasets/mohamedgobara/oral-lesions-malignancy-detection-dataset

[13] T. Izumo, "Oral premalignant lesions: from the pathological viewpoint," *International Journal of Clinical Oncology*, vol. 16, no. 1, pp. 15–26, Jan. 2011, doi: 10.1007/s10147-010-0169-z.

[14] V. Liyanage, M. Tao, J. S. Park, K. N. Wang, and S. Azimi, "Malignant and non-malignant oral lesions classification and diagnosis with deep neural networks," *Journal of Dentistry*, vol. 137, p. 104657, Aug. 2023, doi: 10.1016/j.jdent.2023.104657.