International Journal of Multidisciplinary Studies and Innovative Technologies

e-ISSN: 2602-4888
dergipark.org.tr/en/pub/ijmsit
Research Article

2024, 8 (2), 151-158
DOI: 10.36287/ijmsit.8.2.16
Received: November 28, 2024; Accepted: December 25, 2024

AI-Powered Classification of Oral Lesions: Improving Early Detection and Diagnosis

Hakan Yılmaz[1*], Mehmet Özdem[2]

[1*] *Medical Engineering Dept., Karabuk University, Karabuk, Türkiye (hakanyilmaz@karabuk.edu.tr) (ORCID: 0000-0002-8553-388X)*
[2] *Türk Telekom, Ankara, Türkiye (mehmet.ozdem@turktelekom.com.tr) (ORCID: 0000-0002-2901-2342)*

Abstract – Oral malignancies pose significant global health challenges, with oral squamous cell carcinoma (OSCC) being the most prevalent form. Early detection of potentially malignant oral disorders (OPMDs) such as leukoplakia and oral submucous fibrosis is crucial for improving patient prognosis. Traditional diagnostic approaches often face limitations like subjective interpretation and potential delays. This study aimed to develop and evaluate a deep learning-based model for the classification of oral lesions as benign or malignant using publicly available image datasets. Utilizing a modified VGG16 architecture and optimized preprocessing techniques, the model was trained on 330 annotated intraoral images and achieved an overall accuracy of 94.79%. Key performance metrics included a precision of 95.11%, sensitivity and specificity of 94.58%, and an F1 score of 94.74%. The model's performance was comparable to or exceeded existing models with larger datasets, demonstrating its capability for effective feature extraction and reliable classification. The high area under the curve (AUC) value of 0.96 reinforced its potential for clinical application. While the model showed strong diagnostic capability, expanding the dataset size and incorporating a broader range of cases could further enhance generalizability. Future work should also consider integrating real-time image acquisition and optimizing computational processes for practical deployment. The findings underscore the promise of AI-driven diagnostic tools in supporting healthcare professionals by enabling timely, accurate, and scalable detection of oral malignancies, thereby contributing to improved patient care and outcomes. This study represents a significant step toward the practical application of AI in oral health diagnostics.

Keywords – *OSCC detection, Oral lesions, OPMD classification, VGG16 model, Malignancy prediction.*

Citation: Yılmaz, H., Özdem, M. (2024). AI-Powered Classification of Oral Lesions: Improving Early Detection and Diagnosis. International Journal of Multidisciplinary Studies and Innovative Technologies, 8(2): 151-158.

I. Introduction

Oral malignancies represent a significant health concern worldwide, encompassing a diverse range of neoplastic conditions that arise within the oral cavity. Among these, oral squamous cell carcinoma (OSCC) is the most prevalent, accounting for a substantial proportion of oral cancers. The World Health Organization has categorized various lesions as oral potentially malignant disorders (OPMDs), which include conditions such as leukoplakia, erythroplakia, and oral submucous fibrosis (OSMF) [1], [2]. These disorders are characterized by an increased risk of malignant transformation, with OSMF alone exhibiting a transformation rate of 3-13% into OSCC [3]. The recognition and management of these lesions are critical, as early detection can significantly improve patient outcomes and survival rates [4].

The clinical presentation of oral malignancies can be insidious, often leading to delayed diagnosis. For instance, oral malignant melanoma, although rare, is associated with a poor prognosis, with a 5-year survival rate ranging from 10% to 25% [5]. Symptoms may not manifest until the disease has progressed significantly, underscoring the importance of regular oral examinations and awareness of potential signs of malignancy [6]. Furthermore, the presence of oral lesions can severely impact patients' quality of life, affecting their ability to eat, speak, and maintain social interactions [7].

The etiology of oral malignancies is multifactorial, with risk factors including tobacco use, alcohol consumption, and chronic irritation from ill-fitting dentures or dental appliances [8].Additionally, systemic conditions such as hematological malignancies can present with oral manifestations, complicating the clinical picture [9]. The interplay between local and systemic factors necessitates a comprehensive approach to diagnosis and treatment, emphasizing the need for interdisciplinary collaboration among healthcare providers [10].

The application of artificial intelligence (AI) in the detection of oral malignancies has emerged as a pivotal advancement in modern dentistry and oncology. Traditional diagnostic methods, while valuable, often suffer from limitations such as reliance on subjective interpretation and the potential for human error, which can delay diagnosis and treatment [11].Recent studies have demonstrated that AI can significantly reduce the time required for diagnosis, thereby addressing critical delays that often occur in traditional diagnostic processes [12], [13]. Such AI-driven solutions could transform the current diagnostic landscape by enabling early detection in both clinical and community settings, thereby improving treatment outcomes and reducing mortality rates.

Warin et al. [14] evaluated deep convolutional neural network (CNN) algorithms for classifying and detecting OPMDs and OSCC using a dataset of 980 oral images. Various CNN architectures were used for image classification, with DenseNet-169 achieving the best performance. The model achieved high diagnostic accuracy metrics for detecting OSCC and OPMD in oral images. For OSCC, precision, sensitivity, and specificity were each 99%, with an F1 score of 98% and an area under the curve (AUC) of 1.0. For OPMD detection, the model recorded 95% precision, sensitivity, and F1 score, with 97% specificity and an AUC of 0.98. These results indicate that CNN models, particularly DenseNet-169, can perform at expert levels, making them promising tools for assisting general practitioners in early oral cancer detection.

Jubair et al. [15] aimed to develop a lightweight CNN for binary classification of oral lesions into benign or malignant/potentially malignant using real-time clinical images. The model utilized EfficientNet-B0, which is known for achieving state-of-the-art accuracy on large datasets while being smaller and faster than traditional CNNs, for transfer learning and was trained on 716 clinical images. The performance metrics included an accuracy of 85%, specificity of 84.50%, sensitivity of 86.70%, and an AUC of 0.93. These results suggest that CNN models can be effectively used to build cost-efficient, embedded AI devices with limited computational power for oral cancer screening and early detection, potentially expanding screening capabilities.

Huang et al. [16] presented a deep-learning model based on a metaheuristic approach for the accurate diagnosis of oral cancer, focusing on early detection to save lives. It used three preprocessing techniques—Gamma correction, noise reduction, and data augmentation—to enhance image quality and boost dataset size. Weights of the CNN were optimized using an improved version of the Squirrel Search Algorithm (ISSA) to increase accuracy. The model was tested on the "Oral Cancer (Lips and Tongue) Images" dataset from Kaggle, containing 131 images classified by ENT specialists. The dataset was split into 70% for training and 30% for testing. The proposed model achieved an accuracy of 97%, precision of 92.66%, sensitivity of 87.34%, and F1-score of 89.37%, demonstrating superior results compared to existing methods. However, the complex nature of both the CNN and the metaheuristic increases time complexity. Despite this, the model shows promise for adapting to different types of cancer diagnosis.

Fu et al. [17] aimed to develop a rapid, non-invasive, and cost-effective deep learning approach to identify oral cavity squamous cell carcinoma (OCSCC) using photographic images. The researchers employed cascaded convolutional neural networks, training them on 44,409 biopsy-proven OCSCC and normal control images from 11 hospitals in China collected over 13 years. The dataset was divided into development and internal validation sets, with an additional external validation set sourced from dental and oral surgery journals. It achieved an AUC of 0.98, a sensitivity of 94.90%, and a specificity of 88.70% on the internal validation dataset. The results suggest that this automated deep-learning approach is a viable clinical tool for fast screening, early detection, and assessment of therapeutic efficacy, demonstrating performance comparable to human specialists.

Bansal et al. [18] aimed to develop a new CNN model, termed "Oral_Cancer_Detection," to classify oral cancer images of lips and tongue into cancerous and non-cancerous categories. The model was trained using a small Kaggle dataset with 131 images (87 cancerous, 44 non-cancerous), incorporating data augmentation, feature extraction, and classification techniques. Implemented in MATLAB, the model achieved a 94% validation accuracy after 132 iterations. Key performance metrics showed a precision and specificity of 100%, sensitivity of 91%, and F1 score of 94%, indicating strong performance despite the dataset's limited size. The model is characterized by low computational requirements, making it effective for rapid cancer classification. While results are promising, the study suggests that increasing dataset size and training parameters could further improve accuracy, albeit with longer processing times.

Lin et al. [19] aimed to enhance the accuracy of smartphone-based deep learning methods for detecting oral diseases, focusing on improving diagnosis through systematic data collection and algorithm optimization. A centered image-capturing approach was developed to collect clear oral cavity images, leading to the creation of a medium-sized dataset with five disease categories: normal, ulcer, low-risk, high-risk, and cancer. A resampling method was also introduced to reduce variability from handheld smartphone cameras. The study employed the HRNet model, achieving a sensitivity of 83%, specificity of 96.60%, precision of 84.30%, and F1 score of 83.60% on 455 test images. The "center positioning" method improved the F1 score by about 8% over a simulated "random positioning" approach, while resampling added a further 6% performance boost. HRNet outperformed models like VGG16, ResNet50, and DenseNet169. The results highlight that smartphone-based imaging, when combined with targeted image capture, resampling, and HRNet, holds promise for primary oral cancer diagnosis.

Tanriver et al. [20] explored the potential of computer vision and deep learning for the automated detection of OPMDs using photographic images. With a two-stage pipeline, the model first detected lesions and then classified them as benign, OPMD, or carcinoma. Using EfficientNet-B4, the model achieved precision, sensitivity, specificity, and F1-score of 87%, 86%, 86%, and 86%, respectively, on the test set. The findings underscore the feasibility of this deep-learning approach as a low-cost, non-invasive tool that can support early screening processes and enhance OPMD detection, contributing to improved oral cancer outcomes. The model's real-time capabilities show potential for broader clinical use, aiding timely diagnosis and treatment.

In summary, oral malignancies encompass a spectrum of conditions that pose significant challenges in terms of diagnosis, treatment, and patient management. The incorporation of AI in the detection of oral malignancies represents a significant leap forward in the field of oral health. By improving diagnostic accuracy, facilitating early detection, and enabling personalized treatment approaches, AI technologies are poised to transform the landscape of oral cancer management.

This study aims to develop and evaluate a deep learning-based approach for the accurate classification of oral lesions as benign or malignant using publicly available image datasets. By employing a modified VGG16 architecture and optimized preprocessing techniques, the research seeks to improve diagnostic accuracy, sensitivity, and specificity compared to existing models. The overarching goal is to provide a reliable, efficient, and scalable tool that can assist healthcare professionals in early detection and diagnosis of oral malignancies, thereby enhancing patient outcomes and facilitating timely treatment interventions.

II. MATERIALS AND METHOD

A. Dataset Used and Programming Environment

In this study, the publicly available "Oral Images Dataset," published by Chandrashekar et al., was utilized (Chandrashekar et al., 2021). This dataset contains color images of oral lesions captured using mobile cameras and intraoral cameras. Lesion areas within the dataset were subsequently labeled as benign or malignant by experts using the VGG Image Annotator (VIA) tool. VGG is an open-source, JavaScript-based application commonly used for image annotation (Dutta & Zisserman, 2019). The labeled image regions were cropped from the original images and resized to a resolution of 224x224 pixels. This process resulted in a dataset comprising a total of 330 images, with 162 labeled as benign and 168 as malignant, to be used in the study. A train-test split with a ratio of 0.3 was applied to the dataset, which comprises a total of 330 images (162 labeled as benign and 168 as malignant). Following the split, 234 images were allocated for training, while 96 images were set aside for testing Sample images from the resulting dataset are presented in Figure 1.

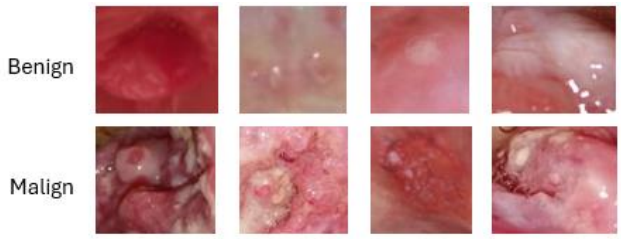

Fig. 1. Sample images from dataset.

The dataset was created, and preprocessing tasks on the images were carried out using a computer with an Intel i7 processor and 16 GB of RAM, alongside the Python programming language.

B. Convolutional Neural Network (CNN)

CNNs have emerged as a cornerstone of modern deep learning, particularly in the realm of image processing and computer vision. Their architecture is inspired by the human visual system, allowing them to effectively recognize patterns and features in visual data. CNNs utilize a hierarchical structure that includes convolutional layers, pooling layers, and fully connected layers, enabling them to learn increasingly abstract representations of input data as it progresses through the network (Bai & Li, 2023; O'Shea & Nash, 2015; Zakaria, 2023). This architecture is particularly adept at handling the spatial hierarchies inherent in images, making CNNs a preferred choice for tasks such as image classification, object detection, and medical image analysis (Kwiatkowska et al., 2021; Tian, 2020).

CNNs have emerged as a cornerstone of modern deep learning, particularly in the realm of image processing and computer vision. Their architecture is inspired by the human visual system, allowing them to effectively recognize patterns and features in visual data. CNNs utilize a hierarchical structure that includes convolutional layers, pooling layers, and fully connected layers, enabling them to learn increasingly abstract representations of input data as it progresses through the network [21], [22], [23]. This architecture is particularly adept at handling the spatial hierarchies inherent in images, making CNNs a preferred choice for tasks such as image classification, object detection, and medical image analysis [24], [25].

Among the various architectures developed for CNNs, the Visual Geometry Group (VGG) model, specifically VGG16, has garnered significant attention due to its depth and performance. The VGG16 architecture is a prominent model in the field of deep learning, particularly known for its application in image classification tasks. Developed by the Visual Geometry Group at the University of Oxford, VGG16 is characterized by its deep architecture, consisting of 16 layers with learnable parameters, which include 13 convolutional layers and 3 fully connected layers [23], [26]. The convolutional layers utilize small receptive fields of 3x3 pixels, which allows the network to capture fine-grained features in images while maintaining a manageable number of parameters [27]. This design choice is crucial as it enables the model to learn hierarchical representations of the input data, progressively extracting more complex features as the data passes through the layers [23]. VGG16's performance has been validated across numerous benchmarks, making it a popular choice for transfer learning in various applications, including medical imaging, where it has been employed for tasks. The application of VGG16 and similar CNN architectures has revolutionized fields, where automated systems can assist in the early detection, significantly improving diagnostic accuracy [24], [28].

C. Metrics Used

In the evaluation of machine learning models, several key performance metrics are commonly utilized to assess their effectiveness in classification tasks. These metrics include precision, accuracy, sensitivity (often referred to as recall), specificity, F1 score, confusion matrix, and the Receiver Operating Characteristic (ROC) curve. Each of these metrics provides unique insights into the model's performance [29].

Precision is defined as the ratio of true positive predictions to the total number of positive predictions made by the model. It indicates how many of the predicted positive cases were actually positive, thereby reflecting the model's ability to avoid false positives [30]. Accuracy, on the other hand, measures the overall correctness of the model by calculating the ratio of correctly predicted instances (both true positives and true negatives) to the total instances evaluated. While accuracy is a straightforward metric, it can be misleading in cases of imbalanced datasets where one class significantly outnumbers the other [31]. Sensitivity, or recall, quantifies the model's ability to correctly identify positive instances. It is calculated as the ratio of true positives to the sum of true positives and false negatives. High sensitivity is crucial in scenarios where missing a positive case (e.g., a disease diagnosis) could have severe consequences [32]. Conversely, specificity measures the proportion of true negatives correctly identified, providing insight into the model's ability to avoid false negatives. It is calculated as the ratio of true negatives to the sum of true negatives and false positives [33]. The F1 score is a harmonic mean of precision and recall, offering a single metric that balances the two. It is particularly useful in situations where there is a need to find an optimal balance between precision and recall, especially in imbalanced datasets [34]. The confusion matrix is a comprehensive tool that summarizes the performance of a classification model by displaying the counts of true positives, true negatives, false positives, and false negatives. This matrix allows for a detailed analysis of the model's performance and helps identify areas for improvement [35]. Lastly, the ROC curve is a graphical representation that illustrates the trade-off between sensitivity (true positive rate) and specificity (false positive rate) at various threshold settings. The area under the ROC curve (AUC) serves as a single scalar value to summarize the model's performance across all thresholds, with higher AUC values indicating better model performance [36]. Together, these metrics provide a robust framework for evaluating the effectiveness of machine learning models in classification tasks, enabling practitioners to make informed decisions based on the specific requirements of their applications.

The formulas for calculating accuracy, precision, sensitivity (recall), specificity, and the F1 score are presented in Equations 1, 2, 3, 4, and 5, respectively.

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

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

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

$$Specificity = \frac{TN}{TN+FP} \quad (4)$$

$$\text{F}_1 Score = \frac{2*TP}{2*TP+FP+FN} \quad (5)$$

III. RESULTS

A. Network Design and Training Information

In this study, modifications were made to the final two layers of the VGG16 model to incorporate specific parameters. A dense layer with 256 neurons and a ReLU activation function was added, followed by an output layer with a sigmoid activation function. The “Adam” optimizer was selected, and the model was configured to run for 200 epochs; however, early stopping was triggered after the 138th epoch, concluding the training. The batch size was set to 32. Initially, the loss value was 3.32, which decreased to 0.15 by the end of training. Similarly, the accuracy began at 0.43 and gradually increased, reaching 0.95. Graphs illustrating the changes in loss and accuracy are shown in Figure 2.

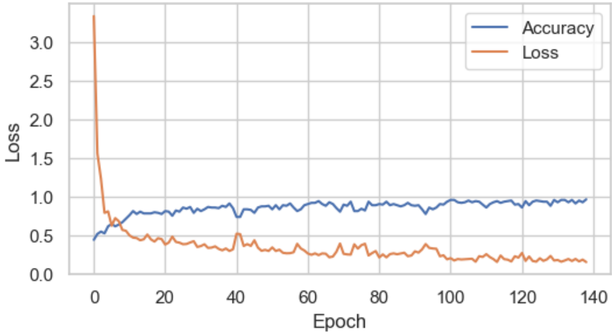

Fig. 2. Training information.

B. Evaluation of Training

In this study, a deep learning model was developed to classify intraoral lesion images captured by cameras as benign or malignant. Following the model evaluation, key performance metrics—including accuracy, precision, sensitivity, specificity, and F1 score—were calculated for each class. The findings are summarized below.

For benign lesions, the model achieved an accuracy of 94.79%, a precision of 97.62%, a sensitivity of 91.11%, and a specificity of 98.04%. The F1 score for the benign class was 94.25%. For malignant lesions, the model demonstrated similar robustness, with an accuracy of 94.79%, a precision of 92.59%, a sensitivity of 98.04%, and a specificity of 91.11%. The F1 score for the malignant class reached 95.23%.

These results indicate that the model performs well in distinguishing between benign and malignant lesions, showcasing strong precision and sensitivity across both classes. The obtained metrics are presented in Table 1. Following the model evaluation, a confusion matrix was generated to capture the classification performance for each class. This matrix enabled a detailed analysis of Type I and Type II errors across both classes, facilitating insights into the model’s misclassification tendencies. Figure 3 presents the confusion matrix, which provides a comprehensive view of the model's ability to correctly identify benign and malignant cases and highlights areas for potential improvement in accuracy.

Table 1. Evaluation metrics of results.

Class	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1 Score (%)
Benign	94.79	97.62	91.11	98.04	94.25
Malign	94.79	92.59	98.04	91.11	95.23
Average	94.79	95.11	94.58	94.58	94.74

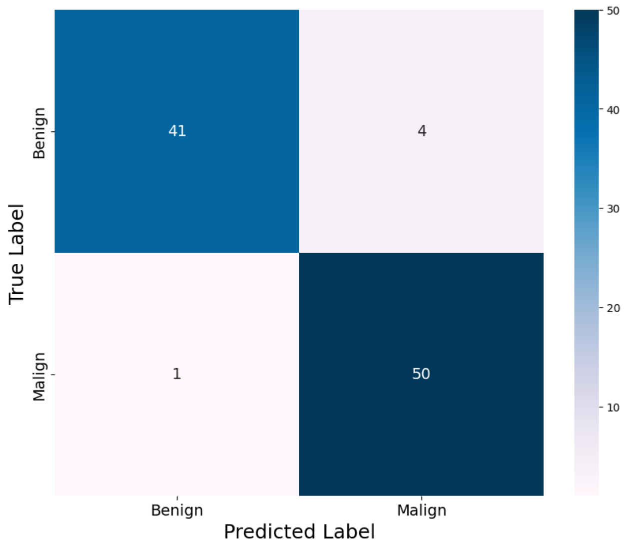

Fig. 3. Confusion matrix of results.

The True Positive Rate (TPR) matrix illustrates the proportion of actual positive cases accurately identified by the model in each class, providing insight into the model's effectiveness in detecting specific conditions. High TPR values across classes suggest that the model performs well in recognizing positive instances, thereby reducing the occurrence of false negatives and enhancing diagnostic reliability for each condition. The TPR matrix of the results is presented in Figure 4.

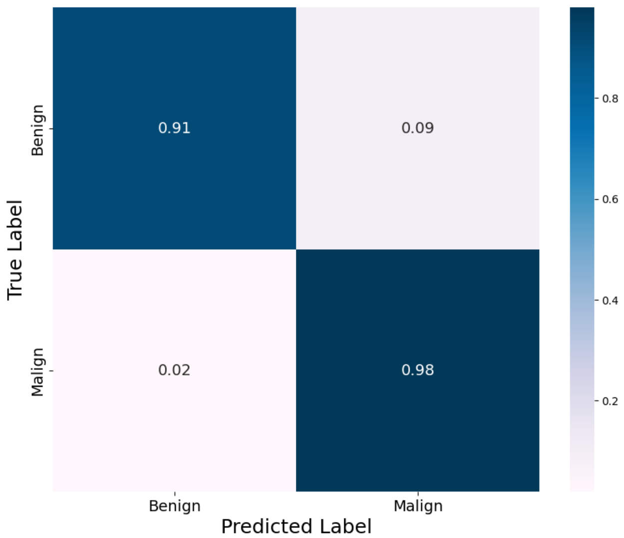

Fig. 4. TPR matrix of results.

Finally, the Area Under the Curve (AUC) value was calculated, and the Receiver Operating Characteristic (ROC) curve was generated to further evaluate the model's performance. This ROC curve, presented in Figure 5, visually represents the trade-off between sensitivity and specificity, providing an additional measure of the model's classification effectiveness.

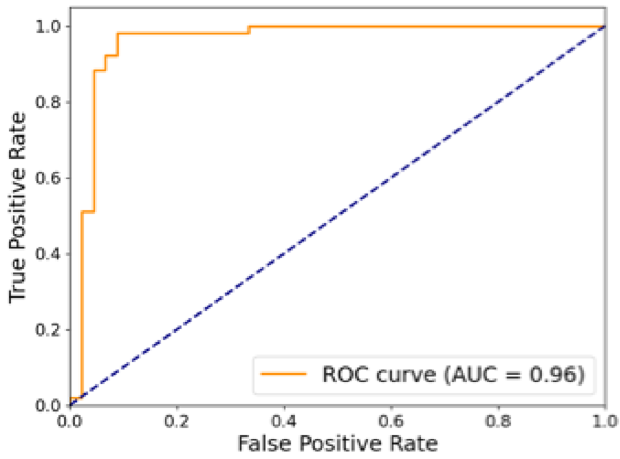

Figure 1. ROC curve and AUC value of results.

IV. DISCUSSION

The results from this study reveal a strong performance of the proposed deep learning model in classifying benign and malignant oral conditions from images. The model achieved an impressive accuracy of 94.79% on a dataset of 330 images, highlighting its capacity for reliable classification even with a relatively small dataset size. The relevant studies are given in Table 2 along with their performance metrics.

The precision of 95.11% signifies that the model maintains a low false positive rate, correctly identifying benign and malignant cases with high confidence. In terms of sensitivity, the model achieved 94.58%, indicating its effectiveness in detecting true positive cases of malignancy. This sensitivity is notable, especially when compared to models like EfficientNet-B4, which reached 86% sensitivity on a dataset of 684 images, suggesting that the proposed model better identifies malignant cases.

Moreover, the model's specificity was also 94.58%, reflecting a balanced capability to avoid false alarms by correctly identifying non-malignant cases. This level of specificity is comparable to the cascaded CNN's 88.7%, confirming that the model offers a significant reduction in false positives, making it suitable for clinical use where overdiagnosis can be problematic.

The F1 Score of 94.74% illustrates the model's balanced handling of precision and recall, showcasing its suitability for real-world application where both measures are critical. This score also suggests that the model performs well in maintaining a strong balance between false positives and false negatives, which is crucial in clinical diagnostics.

The AUC of 0.96 further reinforces the model's excellent discrimination capability between benign and malignant classes, approaching the maximum AUC value of 1.0. Compared to the AUC scores of other models, such as 0.983 for the cascaded CNN and 0.928 for EfficientNet-B0, the proposed model demonstrates highly competitive performance, despite the smaller dataset.

Table 2. Performance metrics of studies in the literature.

Study	Classification Type	CNN Model	Number of Images in Dataset	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1 Score (%)	AUC
Warin et al. (2022) [14]	Two-class	DenseNet-169	980		98% (OSCC) 95% (OPMD)	99% (OSCC) 95% (OPMD)	99% (OSCC) 97% (OPMD)	98% (OSCC) 95% (OPMD)	1.0 (OSCC) 0.98 (OPMD)
Jubair et al. (2020) [15]	Two-class	EfficientNet-B0	716	85.00%		86.70%	84.50%		0.93
Huang et al. (2023) [16]	Two-class	Unique Model	131	97.00%	92.66%	87.34%%		89.37%	
Fu et al. (2020) [17]	Two-class	Cascaded CNN Model	44409	92.40%		94.90%	88.70%		0.98
Bansal et al. (2023) [18]	Two-class	Unique Model	131	92.00%	100.00%	88.90%	100.00%	94.12%	
Tanriver et al. (2021) [20]	Multi-class	EfficientNet-b4	684		87.00%	86.00%		86.00%	
Lin et al. (2021) [19]	Multi-class	HRNet-W18	455		84.30%	83.00%	96.60%	83.60%	
Our study	Two-class	VGG16 based model	330	94.79%	95.11%	94.58%	94.58%	94.74%	0.96

While the model's performance metrics are promising, the relatively limited dataset size of 330 images may restrict generalizability. Expanding the dataset size, possibly incorporating a wider range of cases and image variations, could enhance the model's robustness and ensure its applicability across diverse populations.

Overall, the study confirms that AI-based image classification for oral diseases can achieve high accuracy and balance across key metrics, even with smaller datasets. The model's potential for clinical application is significant, as it provides accurate, efficient, and balanced diagnostic support, which could enhance early detection and patient outcomes.

V. CONCLUSION

In conclusion, this study demonstrates that deep learning models, particularly CNN architectures such as the modified VGG16, hold substantial potential for the classification of oral lesion images as benign or malignant. The proposed model, with a robust performance accuracy of 94.79%, precision of 95.11%, and sensitivity and specificity both at 94.58%, showcases its capability to effectively distinguish between different types of oral conditions. These results indicate that even with a dataset of moderate size, it is possible to develop a highly accurate model that can contribute significantly to early detection efforts in oral health diagnostics.

This study aimed to develop and evaluate a deep learning-based approach for the accurate classification of oral lesions using publicly available image datasets. By employing a modified VGG16 architecture and optimized preprocessing techniques, the research sought to improve diagnostic accuracy, sensitivity, and specificity compared to existing models. The overarching goal was to provide a reliable, efficient, and scalable tool that can assist healthcare professionals in early detection and diagnosis of oral malignancies, enhancing patient outcomes and facilitating timely treatment interventions.

The findings align well with existing literature, supporting the notion that AI-driven diagnostic tools can augment traditional methods and assist healthcare professionals by providing consistent, rapid, and reliable results. The high F1 score and AUC further underscore the model's balanced performance, indicating its readiness for potential integration into clinical workflows to support decision-making processes and improve patient care outcomes.

However, the study also recognizes that to achieve wider applicability and enhanced reliability, future research should focus on expanding dataset sizes and including more diverse and complex cases. This would aid in addressing limitations related to generalizability and ensure that the model can be effectively employed in various clinical settings. In addition, incorporating real-time image acquisition techniques and optimizing computational efficiency could further enhance the practical deployment of the model.

In summary, this research provides a promising step forward in leveraging deep learning for the early detection and classification of oral malignancies. By bridging the gap between traditional diagnostic methods and modern AI capabilities, this study contributes to the broader effort of enhancing oral health management and ultimately improving patient outcomes.

ACKNOWLEDGMENT

Authors' Contributions

The authors' contributions to the paper are equal.

Statement of Conflicts of Interest

There is no conflict of interest between the authors.

Statement of Research and Publication Ethics

The authors declare that this study complies with Research and Publication Ethics.

REFERENCES

[1] S. Warnakulasuriya *et al.*, 'Oral Potentially Malignant Disorders: A Consensus Report From an International Seminar on Nomenclature and Classification, Convened by the WHO Collaborating Centre for Oral Cancer', *Oral Diseases*, vol. 27, no. 8, pp. 1862–1880, 2020, doi: 10.1111/odi.13704.

[2] S. Yang, Y. Lee, L. Chang, C. Yang, C. Luo, and P. Wu, 'Oral Tongue Leukoplakia: Analysis of Clinicopathological Characteristics, Treatment Outcomes, and Factors Related to Recurrence and Malignant Transformation', *Clinical Oral Investigations*, vol. 25, no. 6, pp. 4045–4058, 2021, doi: 10.1007/s00784-020-03735-1.

[3] C. B. More and N. R. Rao, 'Proposed Clinical Definition for Oral Submucous Fibrosis', *Journal of Oral Biology and Craniofacial Research*, vol. 9, no. 4, pp. 311–314, 2019, doi: 10.1016/j.jobcr.2019.06.016.

[4] S. Abati, C. Bramati, S. Bondi, A. Lissoni, and M. Trimarchi, 'Oral Cancer and Precancer: A Narrative Review on the Relevance of Early Diagnosis', *International Journal of Environmental Research and Public Health*, vol. 17, no. 24, p. 9160, 2020, doi: 10.3390/ijerph17249160.

[5] K. Matsuoka, 'Oral Malignant Melanoma Detected After Resection of Amelanotic Pulmonary Metastasis', *International Journal of Surgery Case Reports*, vol. 4, no. 12, pp. 1169–1172, 2013, doi: 10.1016/j.ijscr.2013.10.004.

[6] L. Cigic, 'Increased Prevalence of Oral Potentially Malignant Lesions Among Croatian War Invalids, a Cross-Sectional Study', *Journal of Clinical and Experimental Dentistry*, pp. e734-741, 2023, doi: 10.4317/jced.60715.

[7] F. M. Ghanaei, F. Joukar, M. Rabiei, A. Dadashzadeh, and A. K. Valeshabad, 'Prevalence of Oral Mucosal Lesions in an Adult Iranian Population', *Iranian Red Crescent Medical Journal*, vol. 15, no. 7, pp. 600–604, 2013, doi: 10.5812/ircmj.4608.

[8] A. M. Kavarodi, M. Thomas, and J. Kannampilly, 'Prevalence of Oral Pre-Malignant Lesions and Its Risk Factors in an Indian Subcontinent Low Income Migrant Group in Qatar', *Asian Pacific Journal of Cancer Prevention*, vol. 15, no. 10, pp. 4325–4329, 2014, doi: 10.7314/apjcp.2014.15.10.4325.

[9] G. Guan and N. Firth, 'Oral Manifestations as an Early Clinical Sign of Acute Myeloid Leukaemia: A Case Report', *Australian Dental Journal*, vol. 60, no. 1, pp. 123–127, 2015, doi: 10.1111/adj.12272.

[10] H. Mawardi *et al.*, 'Oral Epithelial Dysplasia and Squamous Cell Carcinoma Following Allogeneic Hematopoietic Stem Cell Transplantation: Clinical Presentation and Treatment Outcomes', *Bone Marrow Transplantation*, vol. 46, no. 6, pp. 884–891, 2011, doi: 10.1038/bmt.2011.77.

[11] N. Al-Rawi *et al.*, 'The Effectiveness of Artificial Intelligence in Detection of Oral Cancer', *International Dental Journal*, vol. 72, no. 4, pp. 436–447, Aug. 2022, doi: 10.1016/j.identj.2022.03.001.

[12] M. García-Pola, E. Pons-Fuster, C. Suárez-Fernández, J. Seoane-Romero, A. Romero-Méndez, and P. López-Jornet, 'Role of Artificial Intelligence in the Early Diagnosis of Oral Cancer. A Scoping Review', *Cancers*, vol. 13, no. 18, p. 4600, Sep. 2021, doi: 10.3390/cancers13184600.

[13] S. Nath, R. Raveendran, and S. Perumbure, 'Artificial Intelligence and Its Application in the Early Detection of Oral Cancers', *Clin Cancer Investig J*, vol. 11, no. 1, pp. 5–9, 2022, doi: 10.51847/h7wa0UHoIF.

[14] K. Warin, W. Limprasert, S. Suebnukarn, S. Jinaporntham, P. Jantana, and S. Vicharueang, 'AI-based analysis of oral lesions using novel deep convolutional neural networks for early detection of oral cancer', *PLoS ONE*, vol. 17, no. 8, p. e0273508, Aug. 2022, doi: 10.1371/journal.pone.0273508.

[15] F. Jubair, O. Al-karadsheh, D. Malamos, S. Al Mahdi, Y. Saad, and Y. Hassona, 'A novel lightweight deep convolutional neural network for early detection of oral cancer', *Oral Diseases*, vol. 28, no. 4, pp. 1123–1130, May 2022, doi: 10.1111/odi.13825.

[16] Q. Huang, H. Ding, and N. Razmjooy, 'Optimal deep learning neural network using ISSA for diagnosing the oral cancer', *Biomedical Signal Processing and Control*, vol. 84, p. 104749, Jul. 2023, doi: 10.1016/j.bspc.2023.104749.

[17] Q. Fu *et al.*, 'A deep learning algorithm for detection of oral cavity squamous cell carcinoma from photographic images: A retrospective study', *EClinicalMedicine*, vol. 27, p. 100558, Oct. 2020, doi: 10.1016/j.eclinm.2020.100558.

[18] S. Bansal, R. S. Jadon, and S. K. Gupta, 'Lips and Tongue Cancer Classification Using Deep Learning Neural Network', in *2023 6th International Conference on Information Systems and Computer Networks (ISCON)*, Mathura, India: IEEE, Mar. 2023, pp. 1–3. doi: 10.1109/ISCON57294.2023.10112158.

[19] 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', *J. Biomed. Opt.*, vol. 26, no. 08, Aug. 2021, doi: 10.1117/1.JBO.26.8.086007.

[20] G. Tanriver, M. Soluk 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, 2021.

[21] M. Bai and M. Li, 'A Presentation of Structures and Applications of Convolutional Neural Networks', *Highlights in Science Engineering and Technology*, vol. 61, pp. 180–187, 2023, doi: 10.54097/hset.v61i.10291.

[22] K. O'Shea and R. R. Nash, 'An Introduction to Convolutional Neural Networks', 2015, doi: 10.48550/arxiv.1511.08458.

[23] N. Zakaria, 'Improved Image Classification Task Using Enhanced Visual Geometry Group of Convolution Neural Networks', *Joiv International Journal on Informatics*

Visualization, vol. 7, no. 4, p. 2498, 2023, doi: 10.30630/joiv.7.4.1752.

[24] D. Kwiatkowska, P. Kluska, and A. Reich, 'Convolutional Neural Networks for the Detection of Malignant Melanoma in Dermoscopy Images', *Advances in Dermatology and Allergology*, vol. 38, no. 3, pp. 412–420, 2021, doi: 10.5114/ada.2021.107927.

[25] Y. Tian, 'Artificial Intelligence Image Recognition Method Based on Convolutional Neural Network Algorithm', *Ieee Access*, vol. 8, pp. 125731–125744, 2020, doi: 10.1109/access.2020.3006097.

[26] K. Simonyan and A. Zisserman, 'Very Deep Convolutional Networks for Large-Scale Image Recognition', 2014, *arXiv*. doi: 10.48550/ARXIV.1409.1556.

[27] I. Fawwaz, T. Candra, D. A. M. Marpaung, A. Dinis, and M. R. Fachrozi, 'Classification of Beetle Type Using the Convolutional Neural Network Algorithm', *Sinkron*, vol. 7, no. 4, pp. 2340–2348, 2022, doi: 10.33395/sinkron.v7i4.11673.

[28] Akshitha and M. Veena, 'Melanoma Detection Using CNN', *International Research Journal of Modernization in Engineering Technology and Science*, 2023, doi: 10.56726/irjmets43733.

[29] H. Yılmaz, 'AI-Powered Healthcare Innovations: Rehabilitation, Education, And Early Diagnosis', Sep. 2024, *Serüven Yayınevi*. doi: 10.5281/ZENODO.13885904.

[30] M. Alehegn, 'Application of Machine Learning and Deep Learning for the Prediction of HIV/AIDS', *Hiv & Aids Review*, vol. 21, no. 1, pp. 17–23, 2022, doi: 10.5114/hivar.2022.112852.

[31] M. Sangeetha, 'Heart Disease Prediction Using ML', *International Journal of Innovative Science and Research Technology*, pp. 2630–2633, 2024, doi: 10.38124/ijisrt/ijisrt24mar2016.

[32] P. S. Mattas and I. Nadaan, 'Optimizing Cardiovascular Disease Diagnosis With Machine Learning: An Analysis', *International Journal of Research Publication and Reviews*, vol. 04, no. 02, pp. 430–434, 2023, doi: 10.55248/gengpi.2023.4217.

[33] C. S. Anita, P. Nagarajan, G. A. Sairam, P. Ganesh, and G. Deepakkumar, 'Fake Job Detection and Analysis Using Machine Learning and Deep Learning Algorithms', *Revista Gestão Inovação E Tecnologias*, vol. 11, no. 2, pp. 642–650, 2021, doi: 10.47059/revistageintec.v11i2.1701.

[34] U. Ramasamy and S. Santhoshkumar, 'Benchmark Datasets and Real-Time Autoimmune Disease Dataset Analysis Using Machine Learning Algorithms With Implementation, Analysis and Results', *Journal of Intelligent & Fuzzy Systems*, vol. 45, no. 2, pp. 2449–2463, 2023, doi: 10.3233/jifs-224115.

[35] H. T. Sihotang, M. K. Albert, F. Riandari, and L. A. Rendell, 'Efficient Optimization Algorithms for Various Machine Learning Tasks, Including Classification, Regression, and Clustering', *Idea*, vol. 1, no. 1, pp. 14–24, 2023, doi: 10.35335/idea.v1i1.3.

[36] S. A. Pane and F. M. Sihombing, 'Classification of Rock Mineral in Field X Based on Spectral Data (SWIR &Amp; TIR) Using Supervised Machine Learning Methods', *Iop Conference Series Earth and Environmental Science*, vol. 830, no. 1, p. 012042, 2021, doi: 10.1088/1755-1315/830/1/012042.