Original Article

Snap Oral Cancer: An Artificial Intelligence-Based Application to Assist in the Diagnosis of Oral Cancer

Ivisson Alexandre Pereira da Silva [1], DDS, MSc; Carlos Alberto Correia Lessa Filho [2], MSc; Anne Caroline dos Santos Barbosa [3], DMD Student; José Marcos dos Santos Oliveira [4], MSc, PhD; Matheus Henrique Alves de Lima [4], DDS, MSc; Marcelo de Castro Meneghim [5], DDS, MSc, PhD; Sonia Maria Soares Ferreira [4], DDS, MSc, PhD;

[1] Dept. of Public Health, Pediatric Dentistry and Orthodontics, Piracicaba Dental School, Universidade Estadual de Campinas, Piracicaba, Brazil.
[2] Dept. of Information Systems, Centro Universitário Cesmac, Maceió, Brazil.
[3] Undergraduate Student, School of Dentistry, Centro Universitário Cesmac, Maceió, Brazil.
[4] Postgraduate Program in Health Research, Centro Universitário Cesmac, Maceió, Brazil.
[5] Dept. of Public Health, Pediatric Dentistry and Orthodontics, Piracicaba Dental School, Universidade Estadual de Campinas, Piracicaba, Brazil.

KEY WORDS
Oral Cancer;
Artificial Intelligence;
Early Diagnosis;
Deep Learning;

Received: 13 October 2025;
Revised: 15 November 2025;
Accepted: 21 Feburary 2026;

ABSTRACT

Background: Oral cancer is one of the most common head and neck malignancies, often diagnosed at advanced stages, which compromises prognosis and increases treatment-related morbidity. Artificial intelligence (AI) has emerged as a promising tool to support early cancer diagnosis.

Purpose: To develop and validate an AI-based model incorporated into the *Snap Oral Cancer* mobile application to assist in the screening of oral cancer and potentially malignant lesions.

Materials and Method: This was an experimental technological study approved by the Research Ethics Committee (CAAE: 45329721.0.0000.0039). A total of 1,523 clinical images of SCC and potentially malignant disorders (leukoplakia, erythroplakia, and actinic cheilitis), collected between 2001 and 2022, were used. The model was built with MobileNet architecture, trained in Python with Keras/TensorFlow, and validated using performance metrics, including sensitivity, specificity, accuracy, and area under the curve (AUC).

Results: The oral cancer model achieved excellent performance, with sensitivity of 98.4%, specificity of 87.7%, accuracy of 95.0%, and AUC of 0.93. Leukoplakia and erythroplakia models showed high sensitivity (100% and 96.3%, respectively), but low specificity (29.5% and 47.6%), resulting in higher false-positive rates. The actinic cheilitis model presented intermediate performance, with sensitivity 77.2%, specificity 68.4% and AUC 0.73.

Conclusion: The AI model demonstrated high efficacy in detecting oral cancer, highlighting its clinical potential through the *Snap Oral Cancer* application. Despite limitations in precursor lesions, the findings reinforce the relevance of AI as an innovative and scalable tool for early diagnosis in dentistry. Future studies should expand the dataset and optimize algorithms to improve performance in precursor conditions and contribute to reducing inequalities in access to cancer diagnosis.

Corresponding Author: Silva IAP, Dept. of Public Health, Pediatric Dentistry and Orthodontics, Piracicaba Dental School, Universidade Estadual de Campinas, Piracicaba, Brazil. Email: iapereira29@gmail.com

Cite this article as:

Introduction

Oral cancer, also known as squamous cell carcinoma (SCC), is one of the most significant head and neck malignancies, with an estimated 377,713 new cases annually worldwide [1]. In Brazil, the National Cancer Institute (INCA) projects 10,900 new cases among men

and 4,200 among women for the 2023–2025 triennium [2]. Tobacco use and alcohol consumption are recognized as the main risk factors associated with the development of this disease [3-5].

One of the major challenges in managing oral cancer is late diagnosis, which drastically reduces five-year survival rates to 30–50% for patients at stages III and IV. In contrast, early detection at stages I and II can increase survival rates to up to 80%, highlighting the critical role of timely diagnosis in improving prognosis [6-7].

Delayed diagnosis of oral cancer remains a recurrent and multifactorial problem, associated with patient-related delays in seeking care, difficulties in accessing health services, and limitations in professional training for the recognition of suspicious lesions [3, 8]. In addition to reducing survival rates, late diagnosis leads to longer, more invasive, and costly treatments. Dentists play a fundamental role in prevention and early detection; however, uncertainties in clinical recognition and limited access to complementary diagnostic procedures still represent important barriers [3].

Artificial intelligence (AI) has emerged as a promising tool to enhance diagnostic accuracy in oncology. Recent studies have demonstrated the potential of AI-based systems to identify complex visual patterns in medical images, surpassing some limitations of human visual assessment [9-11]. In oral oncology, the application of convolutional neural networks (CNNs) to clinical images has shown encouraging results in the detection of malignant and potentially malignant oral lesions [10-11].

Ilhan *et al.* [8] emphasized that AI-based imaging tools may contribute to reducing diagnostic delays in oral cancer, while Kar *et al.* [10] highlighted the potential of Deep Learning models to improve screening accuracy and expand access to early diagnosis. Despite these advances, validated and accessible AI-based solutions remain limited, particularly within the Brazilian public health context.

The use of AI models applied to clinical images offers a promising alternative to reduce diagnostic delays, enhance screening capacity, and support clinical decision-making. This study describes the development and initial validation of *Snap Oral Cancer*, an AI-based application designed to support early diagnosis.

Materials and Method

This was an experimental, technology-based study involving the development and validation of a smartphone application, approved by the Research Ethics Committee (CAAE: 45329721.0.0000.0039; opinion: 4.952.990).

For the construction of the AI model, only images with confirmed histopathological diagnosis were included, while those of low quality, blurred, or lacking diagnostic confirmation were excluded. Only images of biopsied lesions with a confirmed histopathological diagnosis were included. In cases where the clinical hypothesis differed from the microscopic diagnosis, the final classification was based solely on the histopathological report, which was adopted as the gold standard.

A total of 1,523 clinical photographs were used, comprising malignant lesions, specifically SCC, and potentially malignant oral disorders, including leukoplakia, erythroplakia, and actinic cheilitis. Images were obtained from patients users of the Brazilian Unified Health System (SUS), who authorized the use of photographs through informed consent. Additionally, records provided by specialists from other states were incorporated, covering the period from 2001 to 2022. The inclusion criteria adopted in the study exclusively comprised clinical images of oral lesions biopsied between 2001 and 2022, all accompanied by a confirmed histopathological diagnosis, encompassing SCC, leukoplakia, erythroplakia, or actinic cheilitis. Furthermore, only images that presented adequate quality for analysis were considered, including good resolution, focus, and illumination, and which were linked to duly signed informed consent from the patients.

Conversely, all images lacking histopathological confirmation were excluded, as were those that were out-of-focus, low-resolution, or poorly illuminated. Records pertaining to non-neoplastic lesions or those outside the scope of potentially malignant disorders, including inflammatory, infectious, or traumatic lesions, were also excluded. Finally, duplicate or redundant images were disregarded to ensure the integrity and accuracy of the database used for training and validating the AI.

Data collection for algorithm training took place between February 2022 and February 2023. Complementary images from the Digital Atlas of Clinical-Pathological Correlations of Oral Cancer, developed by the research team during PPSUS (Programa de Pesquisa para o Sistema Único de Saúde) 02/2013, were also

included. The software was trained to recognize lesion patterns through a Deep Learning approach, resulting in a model capable of predicting the presence of pathology from clinical photographs.

In the first stage, the Image Data Augmentation algorithm was applied to expand the dataset by generating new images from the originals. Subsequen-tly, the development of the AI model for cancer image screening was initiated. The model was built on a convolutional neural network (CNN) with a MobileNet architecture, composed of four layers (1,024, 1,024, 512, and 2 final nodes), as shown in Figure 1. Training was performed in Python using the Keras and TensorF-low libraries. Keras supports neural network development and operates in conjunction with TensorFlow, an open-source library designed for machine learning and deep learning. The resulting data were stored in HDF5 format (Figure 2).

Considering the differences between Android and iOS operating systems, which rely on distinct coding structures, the application was developed using the React Native platform, ensuring cross-platform compatibility. For the web service, Python was employed in combination with the Flask framework, while the web manager was developed in JavaScript using the NEXT framework. This structure enabled the administration of the image authentication system used for AI training, with access restricted to previously authorized users integrated into the back end.

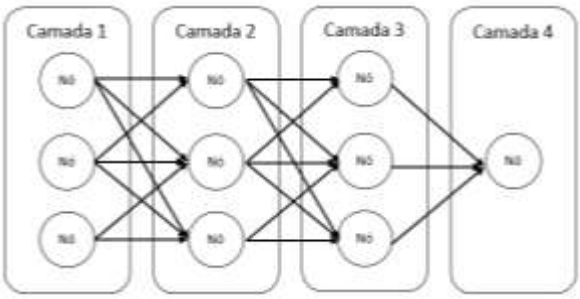

Figure 1: Representation of a convolutional neural network (CNN)

```
#Modelo base, usando MobileNet
base_model=MobileNet(weights='imagenet',include_top=False)

#Estrutura da Rede Neural
x=base_model.output
x=GlobalAveragePooling2D()(x)
x=Dense(1024,activation='relu')(x)
x=Dense(1024,activation='relu')(x)
x=Dense(512,activation='relu')(x)
preds=Dense(2,activation='softmax')(x)

#Definição do modelo
model= Model(inputs=base_model.input,outputs=preds)
```

Figure 2: Structure of the convolutional neural network (CNN) model

These technological choices were made to ensure an efficient, functional, and scalable implementation of the application, leveraging the specific capabilities of each programming language and library employed.

Organization and Validation of the AI Model

The database was segmented into four separate files, with each file corresponding to a specific binary classification task for the CNN used in the application. Images were organized into four distinct categories: oral cancer (positive/negative), leukoplakia (positive/ negative), erythroplakia (positive/ negative), and actinic cheilitis (positive/negative). This strategy enabled the individualized training of the model for each lesion type, as illustrated in Figure 3.

Model validation was a critical step to ensure the AI's ability to correctly identify malignant and potentialtially malignant lesions. The performance metrics employed were sensitivity, specificity, and accuracy, following the criteria proposed by Welikala *et al*. [12].

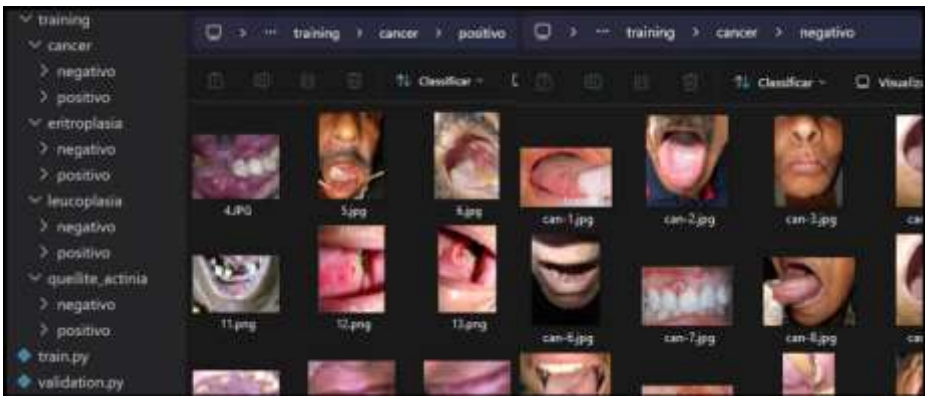

Figure 3: A demonstration of clinical image examples from the artificial intelligence (AI) training dataset, showing the directory structure and the diverse visual patterns utilized in the learning process

These metrics were essential for evaluating the model's effectiveness in detecting and correctly classifying lesions: sensitivity measured the ability to correctly identify positive cases, specificity assessed the ability to recognize negative cases, and accuracy represented the overall proportion of correct classifications.

Performance Metrics and Statistical Analysis

To ensure transparency and reproducibility, the performance metrics and statistical procedures used for model validation were explicitly defined. For each binary classification task, the database was randomly split into training (80%) and test (20%) sets, as shown in Figure 4.

Model performance was evaluated using the metrics of sensitivity, specificity, and accuracy, calculated from the components of the confusion matrix: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The metrics were defined as shown in Table 1 [13].

Table 1: Calculation of performance metrics. True positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Sensitivity: true positive rate; specificity: true negative rate; accuracy: proportion of correct classifications

Performance Metric	Calculation
Sensitivity (True Positive Rate)	TP/ TP+FN
Specificity (True Negative Rate)	TN/ TN+FP
Accuracy	TP+TN/ TP+TN+FP+FN

To assess the reliability of the estimates, 95% confidence intervals (CI) for sensitivity, specificity, and accuracy were calculated using the Wilson method [14], which is widely recommended for binary proportions in diagnostic studies.

Additionally, receiver operating characteristic (ROC) curves were constructed for each classification model, and the area under the curve (AUC) was calculated using the non-parametric DeLong method, widely employed in the validation of AI-based diagnostic models. The ROC curves and AUC values were obtained using the scikit-learn library (Python, version 3.10.4). All statistical analyses were performed in Python, using TensorFlow/Keras for model training and scikit-learn for calculating metrics and ROC curves. The adopted significance level was $\alpha = 0.05$.

Results

The confusion matrices indicated that the model developed for oral cancer achieved excellent performance, characterized by few false negatives and a high accuracy rate. In contrast, the models for erythroplaki and leukoplakia demonstrated high sensitivity but low specificity, resulting in a greater number of false positives. For actinic cheilitis, performance was

Figure 4: Image training process used for the artificial intelligence (AI) model based on convolutional neural network (CNN)

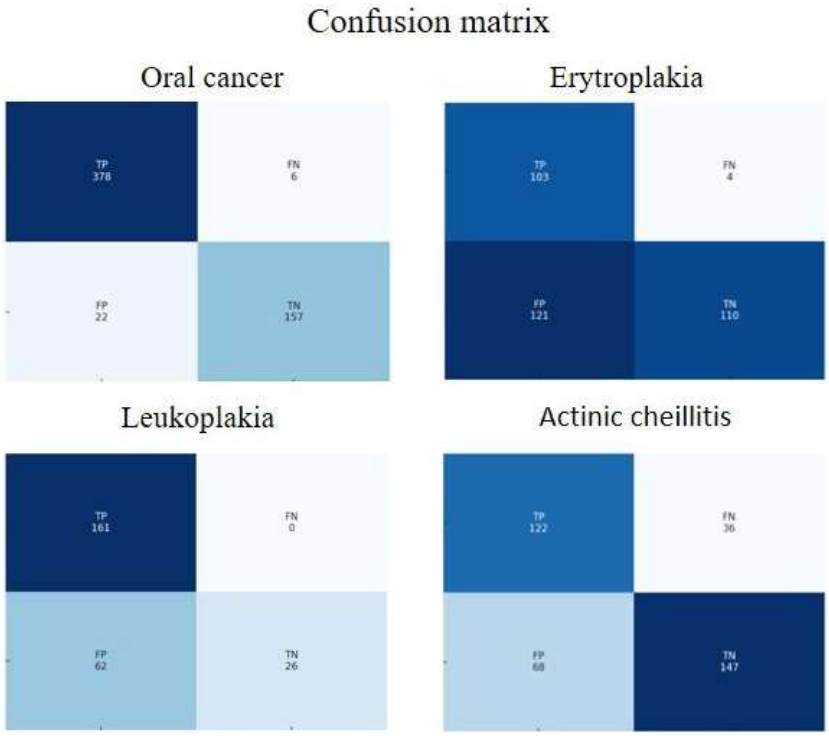

Figure 5: Confusion matrices of the artificial intelligence (AI) models used in the Snap Oral Cancer application for the diagnosis of oral lesions. True positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) are presented. AI: artificial intelligence

intermediate, with a more balanced distribution of correct and incorrect classifications (Figure 5).

The values presented correspond to the number of cases classified as TP, FN, FP, and TN. This model showed excellent performance, with high sensitivity and a low number of false negatives. Conversely, the erythroplakia and leukoplakia models exhibited high sensitivity but low specificity, reflecting a higher number of false positives. The actinic cheilitis model showed intermediate performance, with a partially balanced distribution between correct and incorrect classifications.

For data analysis, mean sensitivity, specificity, and accuracy were calculated to assess model performance. Prediction errors were first measured for each class separately, and then the final scores were reported by calculating the mean of these results, weighted equally. The metrics were defined as follows: sensitivity represents the correct detection rate of true positives; and specificity corresponds to the proportion of true negatives correctly identified.

Table 2 presents the performance metrics of the models trained for the four evaluated conditions. The oral cancer model showed high sensitivity (98.4%; 95% CI: 96.7–99.3) and accuracy of 95.0% (95% CI: 92.1–96.9). In contrast, the leukoplakia and erythroplakia models showed low specificity (29.5% and 47.6%, respectively), resulting in a greater number of false positives.

To complement the performance analysis, ROC curves were constructed for each classification model. This type of analysis enabled assessment of the algorithm's discriminative capacity across different cutoff points, highlighting the relationship between sensitivity and specificity. In addition, the calculation of the AUC provided a global indicator of model accuracy, widely used in AI-based medical diagnostic studies (Figure 6).

The best performance was observed for the oral cancer model (AUC=0.93), whereas leukoplakia (AUC

Table 2: Performance metrics of the models trained for the four evaluated conditions. CI indicates confidence interval

Model	Sensitivity	Specificity	Accuracy
Oral cancer	98.4% (96.6–99.3)	87.7% 82.1–91.7)	95.0% (92.9–96.5)
Erytroplakia	96.3% (90.8–98.5)	47.6% (41.3–54.0)	63.0% (57.8–68.0)
Leukoplakia	100.0% (97.7-100.0)	29.5% (21.0–39.8)	75.1% (69.4–80.1)
Actinic cheilitis	77.2% (70.1–83.1)	68.4% (61.9–74.2)	72.1% (67.4–76.4)

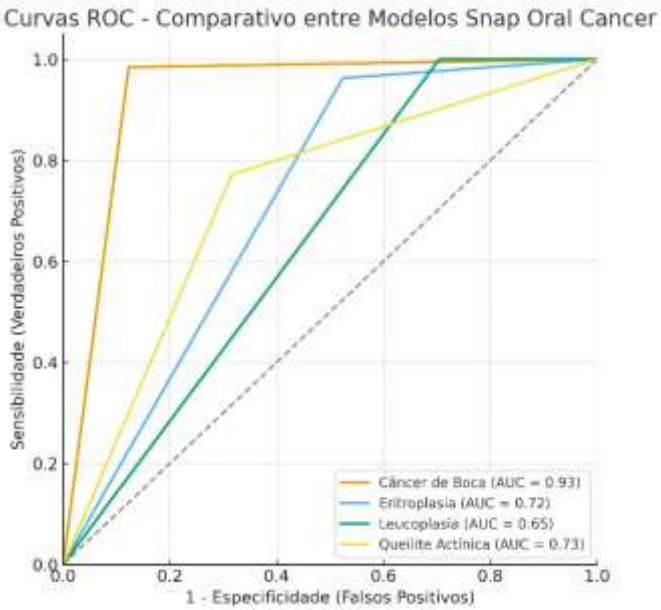

Figure 6: Receiver Operating Characteristic (ROC) curves of the artificial intelligence (AI) models in the Snap Oral Cancer application for the evaluated conditions. ROC: receiver operating characteristic; AUC: area under the curve; AI: artificial intelligence

=0.65) and erythroplakia (AUC= 0.72) showed lower specificity, reflecting higher false-positive rates. The actinic cheilitis model demonstrated intermediate performance (AUC=0.73).

Overall, the oral cancer model achieved excellent results, with high sensitivity, satisfactory specificity, and an AUC of 0.93, confirming its clinical potential. In contrast, the leukoplakia and erythroplakia models achieved high sensitivity but low specificity, leading to an increased number of false positives. The actinic cheilitis model obtained intermediate outcomes, with partial balance between correct and incorrect classifications. These findings suggest that although AI is highly promising for oral cancer diagnosis, further refinements are needed to optimize its application in precursor lesions.

Discussion

The AI system developed in this study demonstrated promising performance for the diagnosis of oral cancer, achieving an accuracy of 95.0%, specificity of 87.7%, and sensitivity of 98.4%. The Deep Learning-based model showed high accuracy in image analysis, particularly for oral cancer, reinforcing its potential as a supportive tool for early detection. In contrast, the models applied to leukoplakia and erythroplakia, although achieving high sensitivity, demonstrated low specificity, resulting in a greater number of false positives. For actinic cheilitis, the outcomes were intermediate, with a more balanced distribution of correct and incorrect classifications.

The results obtained in this study outperformed previous findings, such as those reported by Lin *et al.* [15] and Jeyaraj and Nadar [16], who achieved sensitivities of 94% and 83%, respectively. These data highlight progress compared with models previously tested for the early diagnosis of this condition. Moreover, earlier studies have reinforced the potential of AI in dentistry and oncology. Fu *et al.* [17] demonstrated that algorithms can identify detailed visual patterns in complex oral images with performance surpassing that of specialists, while Kar *et al.* [10] showed that Deep Learning models detected potentially malignant lesions with greater accuracy than conventional methods. These findings are consistent with the present study and strengthen the role of AI as a reliable clinical support tool.

Recently, Bajpai [18] reinforced the potential of AI-based tools in the screening and detection of malignant and potentially malignant lesions of the oral cavity. These findings corroborate the results of the present study, which demonstrate the model's robust performance in recognizing SCC and moderate performance for precursor lesions.

When analyzing specific lesion types, leukoplakia showed 100% sensitivity but low specificity (29.5%),

resulting in a high false-positive rate. This contrasts with the study of González and Quintero- Rojas [19], who reported 75% sensitivity and 98% specificity, highlighting different strategies to balance these two parameters. For erythroplakia, sensitivity of 96.2% and specificity of 47.6% were similar to the results reported by Liyanage *et al.* [20], reiterating the challenges inherent in diagnosing this lesion type. Regarding actinic cheilitis, sensitivity of 77.2% and specificity of 68.3% were comparable to the findings of Spyridonos *et al.* [21], indicating satisfactory performance but with room for improvement.

The high sensitivity achieved for oral cancer suggests that the model could play a meaningful role in the screening of patients in the early stages of the disease, thereby improving the likelihood of early detection. An innovative aspect of this study lies in the translation of results into clinical practice through an application accessible to primary care professionals. Such a tool could enhance diagnostic accuracy, reduce inequalities in access to healthcare, and streamline referral pathways in settings with limited specialist coverage. Additionally, the use of this application could contribute to the development of real-time epidemiological databases, supporting more effective public health policies for the prevention and control of oral cancer.

Among the limitations of this study, the low specificity of the models applied to precursor lesions, particularly leukoplakia and erythroplakia should be highlighted, as this resulted in a high number of false positives. This limitation may be related to the clinical variability of the lesions, the heterogeneous quality of the images, and the imbalance within the dataset.

The integration of histopathological images and complementary examinations may contribute to the development of more robust models. Multicenter clinical trials will be essential to validate the application across different dental practice scenarios and to consolidate its clinical applicability. In this way, AI could be effectively incorporated as a large-scale diagnostic support tool, strengthening its contribution to both oral health and oncology.

Conclusion

The AI model developed in this study demonstrated excellent performance for oral cancer detection, with high sensitivity, satisfactory specificity, and strong overall accuracy, confirming its clinical potential. Although challenges remain for precursor lesions, the integration of this model into a mobile application represents an innovative and scalable strategy to support early diagnosis, particularly in primary care settings. By enhancing diagnostic capacity and reducing disparities in access to specialized services, this approach offers a promising contribution to oral health and cancer control.

Acknowledgments

We thank the Research Support Foundation of Alagoas (FAPEAL) for the financial support from Public Call FAPEAL 06/ 2020 – PPSUS - Research Program for the SUS: Shared Health Management Decit-SCTIE-MS/ CNPq/FAPEAL/ SESAU-AL.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021; 71: 209-249.

[2] Instituto Nacional de Câncer (BR). Diagnóstico precoce do câncer de boca [Internet]. Rio de Janeiro: INCA; 2022 [cited 2022 Aug 13]. Available at: https://www.inca.gov.br/publicacoes/livros/diagnostico-precoce-do-cancer-de-boca.

[3] Mu G, Wang J, Liu Z, Zhang H, Zhou S, Xiang Q, et al. Association between smokeless tobacco use and oral cavity cancer risk in women compared with men: a systematic review and meta-analysis. BMC Cancer. 2021; 21: 960.

[4] Neville BW. Patologia oral e maxilofacial. 4th ed. Rio de Janeiro: Elsevier; 2016. p. 912.

[5] Wong TSC, Wiesenfeld D. Oral cancer. Aust Dent J. 2018; 63: 91-99.

[6] Bosetti C, Carioli G, Santucci C, Bertuccio P, Gallus S, Garavello W, et al. Global trends in oral and pharyngeal cancer incidence and mortality. Int J Cancer. 2020; 147: 1040-1049.

[7] Ilhan B, Guneri P, Wilder-Smith P. The contribution of artificial intelligence to reducing the diagnostic delay in

oral cancer. Oral Oncol. 2021; 116: 105254.

[8] Ilhan B, Lin K, Guneri P, Wilder-Smith P. Improving Oral Cancer Outcomes with Imaging and Artificial Intelligence. J Dent Res. 2020; 99: 241-248.

[9] García-Pola M, Pons-Fuster E, Suárez-Fernández C, Seoane-Romero J, Romero-Méndez A, López-Jornet P. Role of artificial intelligence in the early Diagnosis of oral cancer. A scoping review. Cancers (Basel). 2021; 13: 4600.

[10] Kar A, Wreesmann VB, Shwetha V, Thakur S, Rao VUS, Arakeri G, et al. Improvement of oral cancer screening quality and reach: The promise of artificial intelligence. J Oral Pathol Med. 2020; 49: 727-730.

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

[12] Welikala RA, Remagnino P, Lim JH, Chan CS, Rajendran S, Kallarakkal TG, et al. Automated detection and classification of oral lesions using deep learning for early detection of oral cancer. IEEE Access.2020; 8: 132677-93.

[13] Parikh R, Mathai A, Parikh S, Chandra Sekhar G, Thomas R. Understanding and using sensitivity, specificity and predictive values. Indian J Ophthalmol. 2008; 56: 45-50.

[14] Wilson EB. Probable inference, the law of succession, and statistical inference. J Am Statis Assoc. 1927; 22: 209–212.

[15] Lin H, Chen H, Weng L, Shao J, Lin J. Automatic detection of oral cancer in smartphone-based images using deep learning for early diagnosis. J Biomed Opt. 2021; 26: 086007.

[16] Jeyaraj PR, Samuel Nadar ER. Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm. J Cancer Res Clin Oncol. 2019; 145: 829-837.

[17] Fu Q, Chen Y, Li Z, Jing Q, Hu C, Liu H, Bao J, et al. A deep learning algorithm for detection of oral cavity squamous cell carcinoma from photographic images: A retrospective study. E Clin Med. 2020; 27: 100558.

[18] Bajpai M. Use of Artificial Intelligence in the Early Diagnosis of Oral Cancer: Current Update, Limitations and Future Perspectives. J Dent Shiraz Univ Med Sci (In press).

[19] González JD, Quintero-Rojas J. Use of convolutional neural networks in smartphones for the identification of oral diseases using a small dataset. Rev Fac Ing Univ Antioquia. 2021; 30: 1-16.

[20] Liyanage V, Tao M, Park JS, Wang KN, Azimi S. Malignant and non-malignant oral lesions classification and diagnosis with deep neural networks. J Dent. 2023; 137: 104657.

[21] Spyridonos P, Gaitanis G, Bassukas ID, Tzaphlidou M. Evaluation of vermillion border descriptors and relevance vector machines discrimination model for making probabilistic predictions of solar cheilosis on digital lip photographs. Comput Biol Med. 2015; 63: 11-18.