

Application of neural networks for the detection of oral cancer: A systematic review

María del Pilar Beristain-Colorado^{1,A–F}, María Eugenia Marcela Castro-Gutiérrez^{1,A–F}, Rafael Torres-Rosas^{2,A,B,D–F}, Marciano Vargas-Treviño^{3,A,E,F}, Adriana Moreno-Rodríguez^{4,A,E,F}, Gisela Fuentes-Mascorro^{5,A,E,F}, Liliana Argueta-Figueroa^{6,7,A,C–F}

¹ Department of Biosciences, Postgraduate Division, Faculty of Medicine, Benito Juárez Autonomous University of Oaxaca, Oaxaca de Juárez, Mexico

² Center for Health and Disease Studies, Postgraduate Division, Faculty of Dentistry, Benito Juárez Autonomous University of Oaxaca, Oaxaca de Juárez, Mexico

³ Laboratory of Robotics, Bio-Inspired Systems and Artificial Intelligence, Faculty of Biological Systems and Technological Innovations, Benito Juárez Autonomous University of Oaxaca, Oaxaca de Juárez, Mexico

⁴ Laboratory of Epidemiological and Clinical Studies, Experimental Designs and Research, Faculty of Chemical Sciences, Benito Juárez Autonomous University of Oaxaca, Oaxaca de Juárez, Mexico

⁵ Laboratory of Animal Reproduction Research (LIRA), Benito Juárez Autonomous University of Oaxaca, Oaxaca de Juárez, Mexico

⁶ National Technological Institute of Mexico/Toluca Institute of Technology (Tecnológico Nacional de México (TecNM)/Instituto Tecnológico de Toluca), Metepec, México

⁷ National Council for Science and Technology (Consejo Nacional de Ciencia y Tecnología – CONACyT), Mexico City, México

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation;

D – writing the article; E – critical revision of the article; F – final approval of the article

Dental and Medical Problems, ISSN 1644-387X (print), ISSN 2300-9020 (online)

Dent Med Probl. 2024;61(1):121–128

Address for correspondence

Liliana Argueta-Figueroa

E-mail: L_argueta_figueroa@hotmail.com

Funding sources

None declared

Conflict of interest

None declared

Acknowledgements

The authors would like to thank the National Council for Science and Technology (CONACyT) for awarding the scholarships (No. 746072 and No. 755117), the Researchers for Mexico CONACyT program and the Postgraduate Division at the Faculty of Dentistry of the Benito Juárez Autonomous University of Oaxaca, Oaxaca de Juárez, Mexico, for their support, as well as the Academic Group for Health Research (UABJO-CA-63).

Received on December 14, 2022

Reviewed on January 5, 2023

Accepted on January 26, 2023

Published online on April 26, 2023

Cite as

Beristain-Colorado MdP, Castro-Gutiérrez MEM, Torres-Rosas R, et al.

Application of neural networks for the detection of oral cancer:

A systematic review. *Dent Med Probl.* 2024;61(1):121–128.

doi:10.17219/dmp/159871

DOI

10.17219/dmp/159871

Copyright

Copyright by Author(s)

This is an article distributed under the terms of the

Creative Commons Attribution 3.0 Unported License (CC BY 3.0)

(<https://creativecommons.org/licenses/by/3.0/>).

Abstract

One potential application of neural networks (NNs) is the early-stage detection of oral cancer. This systematic review aimed to determine the level of evidence on the sensitivity and specificity of NNs for the detection of oral cancer, following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and Cochrane guidelines. Literature sources included PubMed, ClinicalTrials, Scopus, Google Scholar, and Web of Science. In addition, the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool was used to assess the risk of bias and the quality of the studies. Only 9 studies fully met the eligibility criteria. In most studies, NNs showed accuracy greater than 85%, though 100% of the studies presented a high risk of bias, and 33% showed high applicability concerns. Nonetheless, the included studies demonstrated that NNs were useful in the detection of oral cancer. However, studies of higher quality, with an adequate methodology, a low risk of bias and no applicability concerns are required so that more robust conclusions could be reached.

Keywords: oral cancer, oral neoplasms, medical informatics applications, computer neural networks, cancer early detection

Introduction

Artificial intelligence (AI) has significantly impacted the field of medicine,¹ and much AI research focuses on the diagnosis and prognosis of cancer,² neurological disorders³ and cardiovascular diseases,⁴ among others.⁵ Neural networks (NNs) constitute an area of AI. They contain sets of artificial neurons organized in super-imposed layers – an input layer, n intermediate layers for data processing and a result layer.⁶ Deep learning (DL) is a combination of NN and machine learning; it enables the creation of computational models composed of multiple processing layers, able to learn the representations of data with multi-level abstraction.⁷ In DL, convolutional, recursive and recurrent NNs have been applied.⁸ Convolutional neural networks (CNNs) are a class of DL algorithms applied to medical image classification,⁹ including those used for cancer detection.^{10–13}

Oral cancer ranks sixth among the most common high-risk malignancies in middle-income countries globally.¹⁴ The most common type of oral cancer is oral squamous cell carcinoma (OSCC).¹⁵ Early diagnosis and treatment are crucial to improve patient survival. The histopathological examination of biopsy samples is the gold standard in diagnosing oral cancer. However, this approach is invasive and the samples require complex processing.¹⁶ The detection of oral cancer in situ results in survival rates as high as 82%, though these rates can decrease to 32% if metastases are detected.¹⁷ Therefore, an early diagnosis is essential, and recommendations state that any suspicious lesion that does not heal within 15 days after detection and the removal of the local causes of irritation should be biopsied.^{18,19} Although the histopathological examination of biopsy specimens is the current reference method,²⁰ there are still discrepancies (12%) between the initial diagnosis from the incisional biopsy and the final histopathology results following the excision of the lesion.²¹ However, many patients are reluctant to have a suspicious lesion biopsied by a clinician, for various reasons, including cost, fear of the procedure, concerns about healing, and esthetics. As a result, patients often postpone the biopsy to get a second opinion on its necessity. Therefore, research groups have proposed other diagnostic methods that are logistically more accessible. One of such approaches is the use of NNs for the early diagnosis of oral cancer through the analysis of risk factors, laboratory tests and the images of the lesion.^{22,23}

The process of detection of oral cancer through imaging has different phases. Ideally, during the training phase, a set of images classified into different types, such as the normal region, the cancerous region and the precancerous region, is introduced to NN. The classified images allow NN to learn the characteristics of each set of images. Subsequently, in the testing

phase, the image of a suspicious oral lesion is provided and the system outputs the predicted result.²⁴ Therefore, the NN diagnosis of oral cancer can be made by clinicians working in remote areas, where biopsy processing is complicated.

For the reasons outlined above, this systematic review aimed to determine the level of evidence on the sensitivity and specificity of NNs for detecting oral cancer.

Material and methods

Study protocol registration

This systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy Studies (PRISMA-DTA) statement²⁵ and the Cochrane guidelines.²⁶ The protocol was registered in the International Prospective Registry of Systematic Reviews (PROSPERO) (CRD42021256938). The articles included in the present systematic review were studies with an observational design (cohort and case-control studies). Case reports, case series, animal studies, pilot studies, short communications, and systematic reviews were excluded.

Eligibility criteria, information sources and search strategy

The eligibility of the studies was determined using the modified PICO strategy (Patient/Population, Intervention, Comparison, and Outcome). Searches with no restriction on the publication date were carried out in PubMed, ClinicalTrials, Scopus, Google Scholar, and Web of Science in April 2021, and were updated in July 2022. The search strategies used for each database are shown in Table 1. A manual search was performed by reading the reference sections of the included studies.

To meet the eligibility criteria, studies needed to use NN for the analysis of images for the detection of oral cancer in humans, and assess the sensitivity, specificity, precision, and accuracy of NN in comparison with the histopathological examination. All studies that used NN for the prognosis of oral cancer, to determine the efficacy of oral cancer treatment or to classify the stages of oral cancer, as well as studies that used other methods of data collection (not imaging), were excluded.

Study selection

For the selection of studies, the title and abstract of each paper were read. Those which answered the research question were reviewed in full text to determine if they met the eligibility criteria. If the eligibility criteria were not met, the articles were eliminated with reasons, as shown in Fig. 1.

Table 1. Keywords and algorithms used in the search strategy for each database

PICO strategy	Keywords
Population	patients with suspicious lesions or oral cancer
Intervention (diagnosis)	NN
Comparison	histopathological examination or clinical assessment
Outcome	sensitivity, specificity, precision, and accuracy, correlation coefficient, ROC curve, AUC
Study design	observational
Restrictions	in English or Spanish
Electronic database	PubMed, Clinical Trials, Scopus, Google Scholar, Web of Science
Focus question	What is the evidence on the use of NN for oral cancer detection?
Databases and registries	Algorithm
Google Scholar	("detection") + ("mouth neoplasm" OR "neoplasm, mouth" OR "malignant oral lesions" OR "tongue squamous cell carcinoma" OR "oral cancer" OR "cancer, oral" OR "submucous fibrosis" OR "oral submucous fibrosis") + ("deep learning" OR "neural network")
PubMed	("mouth neoplasm" OR "neoplasm, mouth" OR "malignant oral lesions" OR "tongue squamous cell carcinoma" OR "oral cancer" OR "cancer, oral" OR "submucous fibrosis" OR "oral submucous fibrosis") AND ("hierarchical learning" OR "deep learning" OR "neural network" OR "computer neural network" OR "computer neural networks" OR "network, computer neural" OR "networks, computer neural" OR "neural network, computer")
Clinical Trials	oral cancer + neural network
Scopus	("mouth neoplasm" OR "neoplasm, mouth" OR "malignant oral lesions" OR "tongue squamous cell carcinoma" OR "oral cancer" OR "cancer, oral" OR "submucous fibrosis" OR "oral submucous fibrosis") AND ("hierarchical learning" OR "deep learning" OR "neural network" OR "computer neural network" OR "computer neural networks" OR "network, computer neural" OR "networks, computer neural" OR "neural network, computer")
Web of Science	("mouth neoplasm" OR "neoplasm, mouth" OR "malignant oral lesions" OR "tongue squamous cell carcinoma" OR "oral cancer" OR "cancer, oral" OR "submucous fibrosis" OR "oral submucous fibrosis") AND ("hierarchical learning" OR "deep learning" OR "neural network" OR "computer neural network" OR "computer neural networks" OR "network, computer neural" OR "networks, computer neural" OR "neural network, computer")

NN – neural network; ROC curve – receiver operating characteristic curve; AUC – area under the ROC curve.

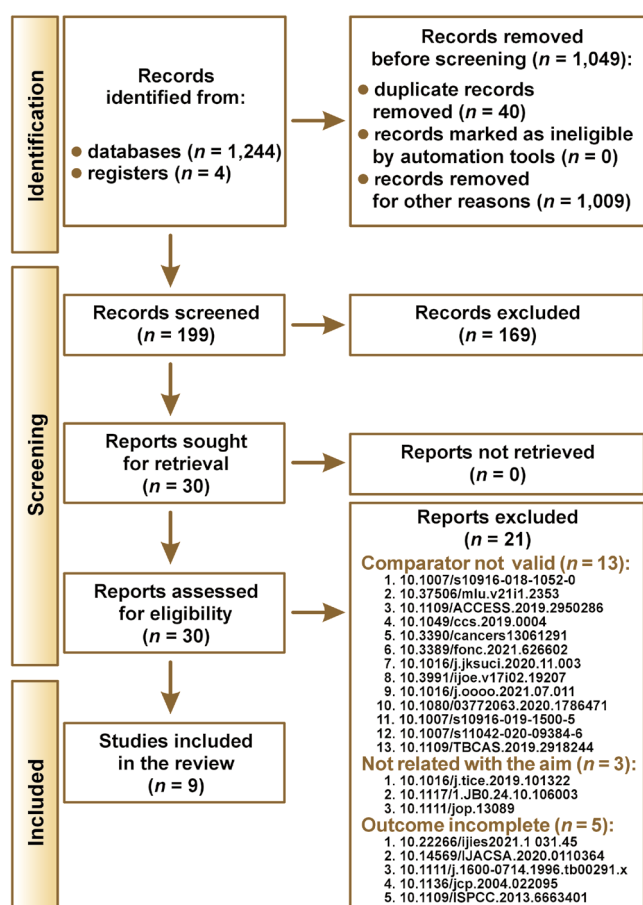


Fig. 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow chart

Data collection and data extraction

The relevant data from the selected articles was extracted, processed and tabulated using a Microsoft Excel spreadsheet. Data extraction was performed independently by 2 reviewers (M.P.B.-C. and M.E.M.C.-G.).

Data synthesis

The results were formally synthesized by grouping the data according to the type of images used for cancer detection, which included photographic images, confocal laser endomicroscopy (CLE), hyperspectral imaging (HSI), optical coherence tomography (OCT), and high-resolution microendoscopy (HRME). The summary of the individual studies with the details of the relevant data, such as the type of images, the NN computing technique, comparators, and outcomes, are presented in the result tables.

Synthesis of the results

If the results of the studies showed high heterogeneity in methodological or population characteristics, a synthesis without a meta-analysis (SWiM) was performed using the qualitative synthesis²⁷ and a representative graph.

Risk of bias and applicability

Two reviewers (R.T.-R. and L.A.-F.) assessed the risk of bias and the applicability of each study, using

the modified Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool,²⁸ which includes the patient selection, index test, reference standard, and flow and timing domains (Table 2). Any disagreement in the assessment of the risk of bias was resolved by the consensus of the research group.

Results

Study selection and study characteristics

The search of electronic databases and registries generated 1,248 records, of which 40 duplicate records were eliminated. After reading the titles and abstracts, it turned out that 30 articles answered the research question, and thus their full texts were retrieved. Subsequently, it was determined if they met the eligibility criteria, which resulted in the reasonable exclusion of 21 articles. For the qualitative analysis, 9 articles were included (Fig. 1). The characteristics of each of the studies and the extracted data are shown in Table 3. The synthesis of the results without a meta-analysis is shown in Fig. 2.

Synthesis of the results

Studies detecting oral cancer used various image types, including photographic images, CLE, HSI, OCT, and HRME.

Photographic images were used most frequently for the detection of oral cancer. Welikala et al. acquired images with a cell phone camera at the primary clinical care level²²; Jubair et al. used various types of digital cameras and smartphones.²⁹ Welikala et al., who used ResNet-101 for image classification and Faster R-CNN for object detection, reported a precision of 84.77% and a recall

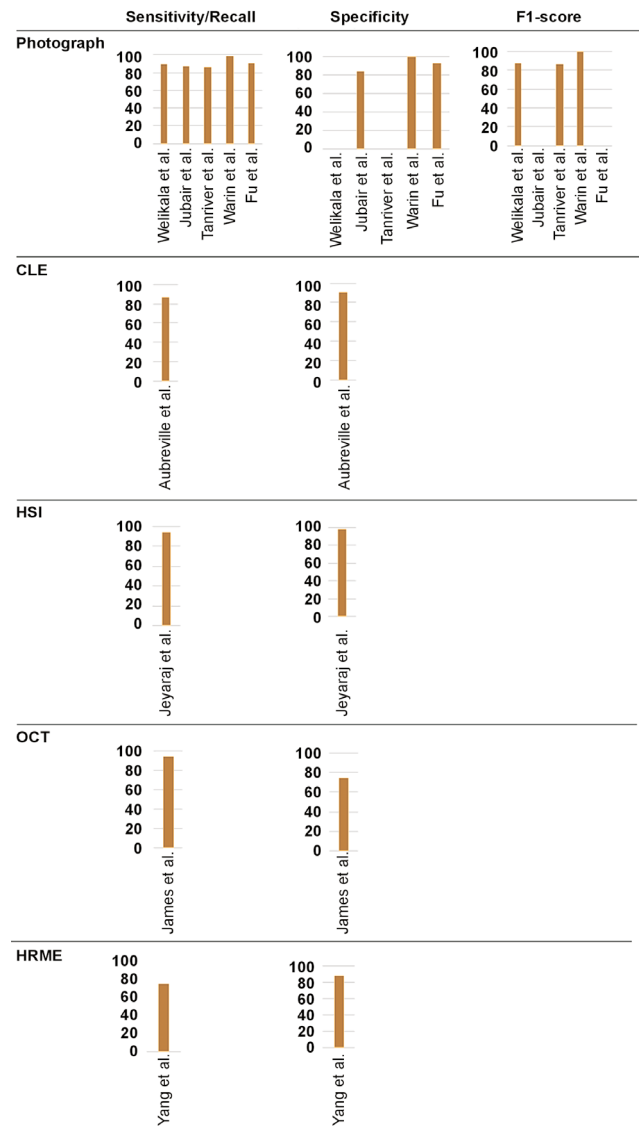


Fig. 2. Graphical representation of the synthesis of the results without a meta-analysis

Table 2. QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies 2) modified for the review

Domain	Questions
Risk of bias	Was a consecutive or random sample of patients enrolled?
	1. Patient selection
	Was a case-control design avoided?
	Did the study avoid inappropriate exclusions?
	2. Index test
	Do the datasets for training and testing contain at least 100 images for each classification to be evaluated?*
	Is the ratio of images used for training to patients no greater than 3:1*
	Is the testing dataset separate from the training dataset?*
	3. Reference standard
	Is the reference standard likely to classify the target condition correctly?
	Were the reference standard results interpreted without the knowledge of the results of the index test?
	4. Flow and timing
	Was there an appropriate interval between the index test and the reference standard?
Applicability	Did all patients receive the same reference standard?
	Were all patients included in the analysis?
	1. Patient selection
	Are there concerns that the included patients and the setting do not match the review question?
	2. Index test
	Are there concerns that the index test, its conduct or interpretation do not match the review question?
	3. Reference standard
	Are there concerns that the target condition as defined by the reference standard does not match the review question?

* element modified according to the review topic.

Table 3. Characteristics of the included studies and their results

Type of images	Study	Data type or population	NN	Results
Photographic images	Welikala et al. 2020 ²²	oral cavity images from cell phone camera testing: 204 images	ResNet-101 Faster R-CNN	multi-class image classification results lesion: precision – 84.77% recall – 89.51% F1 score – 87.07%
	Jubair et al. 2022 ²⁹	photographic images of various tongue lesions testing: 100 images	EfficientNet-B0	AUC = 0.928 sensitivity – 86.7% specificity – 84.5% accuracy – 85.0%
	Tanriver et al. 2021 ³⁰	photographic images of oral lesions testing: 69 images	CNN architecture: YOLOv5 EfficientNet-B4	benign class: precision – 89% recall – 86% F1-score – 88% support – 29 OPMD class: precision – 74% recall – 87% F1-score – 90% support – 23 carcinoma class: precision – 100% recall – 82% F1-score – 90% support – 17 weighted average: precision – 87% recall – 86% F1-score – 86% support – 69
	Warin et al. 2021 ³¹	clinical oral photographs collected retrospectively testing: 140 images	CNN with DenseNet-121 and Faster R-CNN	classification with DenseNet-121: precision – 99.00% sensitivity – 98.75% specificity – 100% F1-score – 99.00% detection accuracy with Faster R-CNN: precision – 76.67% recall – 82.14% F1-score – 79.31%
	Fu et al. 2020 ³²	photographs of biopsy-proven OSCC and normal controls clinical validation: 666 images external validation: 402 images	automated DL algorithm using cascaded CNNs	clinical validation: AUC = 0.970 sensitivity – 91.0% specificity – 93.5% accuracy – 92.3%
CLE	Aubreville et al. 2017 ³⁴	116 video sequences of a suspicious carcinogenic region 12 patients	CNN-based approaches	AUC = 0.960 sensitivity – 86.6% specificity – 90.0% accuracy – 88.3%
HSI	Jeyaraj et al. 2019 ³⁶	HSI images of oral cancer training: 500 images	CNN with 2 partitioned layers for labeling and classifying the region of interest in multidimensional HSI	for 500 training patterns: sensitivity – 94% specificity – 98% accuracy – 94.5%
OCT	James et al. 2021 ²³	OCT images validation: 1,078 images	14 pre-trained NN best results with DenseNet-201 and NASNetMobile	delineating cancer: sensitivity – 93% specificity – 74%
HRME	Yang et al. 2020 ⁴¹	HRME images testing: 253 images	U-Net	sensitivity – 75% specificity – 89% accuracy – 86%

CLE – confocal laser endomicroscopy; HSI – hyperspectral imaging; OCT – optical coherence tomography; HRME – high-resolution microendoscopy; OSCC – oral squamous cell carcinoma; CNN – convolutional neural network; DL – deep learning; OPMD – oral potentially malignant disorder.
F1 score = $2 \times (\text{precision} \times \text{recall}) / (\text{precision} + \text{recall})$

of 89.51%.²² Jubair et al. used a pre-trained EfficientNet-B0 as a lightweight transfer learning model for oral cancer detection, and reported a sensitivity of 86.7% and a specificity of 84.5%.²⁹ On the other hand, Tanriver et al. collected photographic images of oral lesions with histopathological results from the archive of the Department of Tumor Pathology of the Oncology Institute at Istanbul University, Turkey; the rest of the images were collected from publicly available sources by using search engines (<https://images.google.com> and <https://yandex.com/images>).³⁰ The dataset comprised a diverse set of lesions coming from a wide range of oral diseases and anatomical regions. The authors reported a precision of 87% and a recall of 86%.³⁰ Likewise, Warin et al. retrospectively collected clinical oral photographs obtained between 2009 and 2018 at an oral and maxillofacial surgery center.³¹ They used DenseNet-121 for classification, and reported a sensitivity of 98.75% and a specificity of 100%.³¹ Finally, Fu et al. used biopsy-confirmed OSCC photographs from 11 hospitals in China, and reported a sensitivity of 91.0% and specificity of 93.5%.³²

Confocal laser endomicroscopy is an adaptation of the conventional optical microscopy technique, in which the light from a laser source directed at a pinhole geometrically removes information from the outside of the focal plane and generates an optical plane at a specific depth from the surface.³³ Aubreville et al. used 16-bit grayscale CLE images to analyze 4 regions of interest, including the inner lower lip, the upper alveolar ridge and the hard palate.³⁴ The images acquired from suspicious lesions and 3 other areas that were assumed to be healthy resulted in a sensitivity of 86.6% and a specificity of 90.0%.³⁴

Hyperspectral imaging acquires a three-dimensional (3D) data set called a hypercube, formed by 2 spatial dimensions and 1 spectral dimension. Using HSI provides information on tissue physiology, morphology and composition. One field of application for HSI is image classification for detecting tissues at risk of cancer.³⁵ Jeyaraj et al. applied a novel CNN with 2 partitioned layers to label and classify the region of interest in multidimensional HSI, and reported a sensitivity of 94% and a specificity of 98%.³⁶

Optical coherence tomography is a non-invasive high-resolution optical imaging technology that produces real-time cross-sectional images in two-dimensional (2D) space (a lateral coordinate and an axial coordinate).³⁷ It is analogous to ultrasound imaging, except it uses light instead of sound, and is a powerful imaging technology for medical diagnosis, acting as a type of optical biopsy. However, unlike the conventional histopathological examination, which requires the extraction and processing of a tissue sample for microscopic evaluation, OCT can generate real-time images of the tissue.³⁸ James et al. used OCT images to classify non-dysplasia, dysplasia and malignancy through artificial NN/machine learning, and reported a sensitivity of 93% and a specificity of 74% for OSCC identification.²³

High-resolution microendoscopy enables real-time epithelial imaging with subcellular resolution. Numerous

research studies on gastrointestinal neoplasms has indicated that HRME is a modality that provides high specificity and precision for diagnosing different diseases.^{39,40} Yang et al. developed an algorithm to determine whether HRME images show enough oral epithelial nuclei to differentiate between oral cancer and benign tissue.⁴¹ Their study used 811 HRME images from 169 patients and demonstrated that HRME images were suitable for classifying oral cancer. The researchers reported a sensitivity of 75% and a specificity of 89%.⁴¹

Assessment of the risk of bias and applicability

Regarding domain 1 (patient selection), all studies exhibited a high risk of bias, with the main issues being an inadequate selection of patients and a lack of investigator blinding. Furthermore, 1 study (11%) had a high risk of bias with regard to domain 2 (index test), as it used a small number of images.³⁴ As many as 33% of the articles showed high applicability concerns (Fig. 3).

Discussion

In this systematic review, the type of NN applied for the detection of oral cancer was analyzed. All studies used CNN, probably due to the ease of working with images. The comprehensive search aimed to identify whether the studies used an additional type of NN to support the oral cancer detection process. In this regard, Sharma and Om developed a probabilistic NN and general regression model for the early detection and prevention of oral cancer, using various indicators, such as clinical symptoms, medical history and personal history.⁴² This review identified an area of opportunity, which involves using CNN and other types of NN in the analysis of risk factors to provide a more reliable diagnosis of oral cancer, as this combination of data has not been assessed so far.

When verifying the accuracy of the algorithms used for oral cancer diagnosis, the images used for training must come from patients with the diagnosis confirmed through the histopathological examination. Studies were excluded if they did not report the gold standard for the validation of diagnosis, since the absence of an adequate comparator invalidates the results of such studies.

In the study by Tanriver et al., the training, validation and testing dataset was inadequate.³⁰ They obtained some of the images from a hospital (validated by a histopathologist), but as the sample was insufficient, they sourced other images through searching publicly accessible repositories.³⁰ However, such images do not provide the certainty of histopathological diagnostic validation.

Several studies tested the effect of the sample size during the training phase. Narayana et al. determined that a sample size of at least 50 was necessary.⁴³ Fang et al. conducted

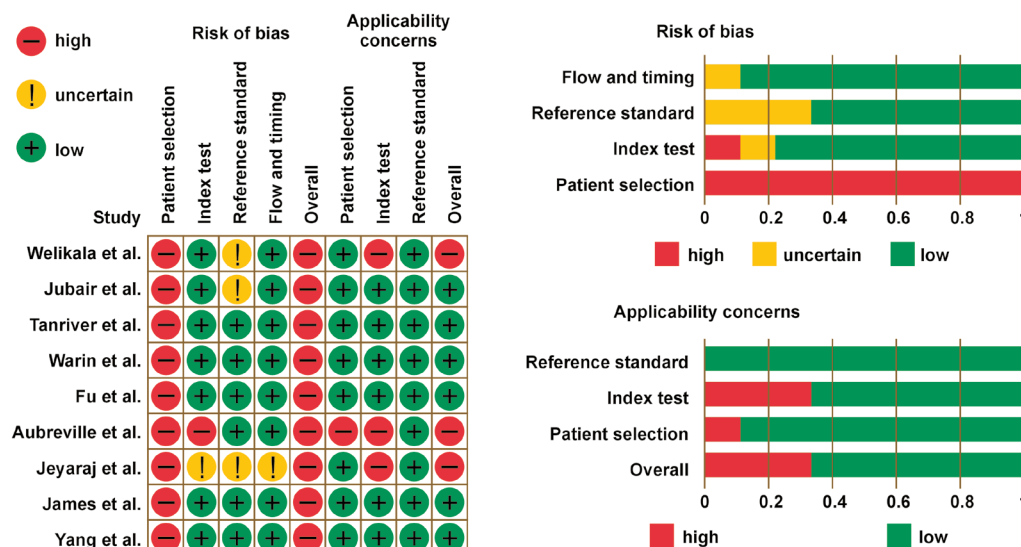


Fig. 3. Graphical representation of the risk of bias and the applicability of the included studies, obtained by means of the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies 2) evaluation

a study that aimed to investigate the impact of the training sample size on the performance of organ self-segmentation (Eye L, Eye R, Lens L, Lens R, Optic nerve L, Optic nerve R, Parotid L, Parotid R, Spinal cord, Larynx, and Body) in computed tomography (CT) based on DL for head and neck cancer patients.⁴⁴ They found that 200 samples were required to obtain a 90% yield for lenses and optic nerves, whereas the remaining organs needed at least 40 images for their detection.⁴⁴ However, according to Narayana et al., the minimum training sample size depends on a number of factors, such as the acquisition protocol, the type of tissue to be segmented, and others.⁴³ The results are not only associated with the dataset, but also with the specific CNN configuration.⁴³ According to Samala et al., assessing the precision and accuracy of CNN architecture by using a test set may be overly optimistic.⁴⁵ Therefore, validating the training process with unknown and independent cases derived from actual clinical practice is crucial. So far, no studies have tested the algorithms developed in this way.

Artificial intelligence can support the detection of cancer in its early stages. The evidence on the efficacy of CNN in image-based oral cancer detection demonstrated that NN could be used in daily clinical practice using photographs. This could be particularly helpful for clinicians in remote locations, where access to specialist oral pathology advice is limited.

Conclusions

Convolutional neural networks can potentially detect oral cancer in its early stages, though the results need to be verified by the corresponding histopathological examination. Most of the analyzed studies showed an accuracy greater than 85%. However, several studies encountered training problems due to the reduced number of images or because

the testing process was performed on the same samples and not in clinical practice. In addition, the analysis of patient-specific risk factors and habits should complement these applications to formulate a more accurate diagnosis.

Ethics approval and consent to participate

Not applicable.

Data availability

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Consent for publication

Not applicable.

ORCID iDs

María del Pilar Beristain-Colorado

<https://orcid.org/0000-0002-0526-1408>

María Eugenia Marcela Castro-Gutiérrez

<https://orcid.org/0000-0003-0631-4812>

Rafael Torres-Rosas <https://orcid.org/0000-0002-5934-003X>

Marciano Vargas-Treviño <https://orcid.org/0000-0002-2090-9727>

Adriana Moreno-Rodríguez <https://orcid.org/0000-0003-4100-8255>

Gisela Fuentes-Mascorro <https://orcid.org/0000-0002-3909-3342>

Liliana Argueta-Figueroa <https://orcid.org/0000-0002-1044-6757>

References

- Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: Past, present and future. *Stroke Vasc Neurol.* 2017;2(4):230–243. doi:10.1136/svn-2017-000101
- Huang S, Yang J, Fong S, Zhao Q. Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges. *Cancer Lett.* 2020;471:61–71. doi:10.1016/j.canlet.2019.12.007
- Senders JT, Staples PC, Karhade AV, et al. Machine learning and neurosurgical outcome prediction: A systematic review. *World Neurosurg.* 2018;109:476–486.e1. doi:10.1016/j.wneu.2017.09.149

4. Kilic A. Artificial intelligence and machine learning in cardiovascular health care. *Ann Thorac Surg.* 2020;109(5):1323–1329. doi:10.1016/j.athoracsurg.2019.09.042
5. Ambros-Antemate JF, Reyes-Flores A, Argueta-Figueroa L, et al. Accuracy of machine learning algorithms for the assessment of upper-limb motor impairments in patients with post-stroke hemiparesis: A systematic review and meta-analysis. *Adv Clin Exp Med.* 2022;31(12):1309–1318. doi:10.17219/acem/152596
6. Fourcade A, Khonsari RH. Deep learning in medical image analysis: A third eye for doctors. *J Stomatol Oral Maxillofac Surg.* 2019;120(4):279–288. doi:10.1016/j.jormas.2019.06.002
7. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521(7553):436–444. doi:10.1038/nature14539
8. Abiodun Ol, Jantan A, Omolara AE, Dada KV, Mohamed NA, Arshad H. State-of-the-art in artificial neural network applications: A survey. *Heliyon.* 2018;4(11):e00938. doi:10.1016/j.heliyon.2018.e00938
9. Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK. Medical image analysis using convolutional neural networks: A review. *J Med Syst.* 2018;42(11):226. doi:10.1007/s10916-018-1088-1
10. Yamlome P, Akwaboah AD, Marz A, Deo M. Convolutional neural network based breast cancer histopathology image classification. *Annu Int Conf IEEE Eng Med Biol Soc.* 2020;2020:1144–1147. doi:10.1109/EMBC44109.2020.9176594
11. Munir K, Elahi H, Ayub A, Frezza F, Rizzi A. Cancer diagnosis using deep learning: A bibliographic review. *Cancers (Basel).* 2019;11(9):1235. doi:10.3390/cancers11091235
12. Rama Krishnan MM, Venkatraghavan V, Acharya UR, et al. Automated oral cancer identification using histopathological images: A hybrid feature extraction paradigm. *Micron.* 2012;43(2–3):352–364. doi:10.1016/j.micron.2011.09.016
13. Das DK, Bose S, Maiti AK, Mitra B, Mukherjee G, Dutta PK. Automatic identification of clinically relevant regions from oral tissue histological images for oral squamous cell carcinoma diagnosis. *Tissue Cell.* 2018;53:111–119. doi:10.1016/j.tice.2018.06.004
14. Sung H, Ferlay J, Siegel RL, 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(3):209–249. doi:10.3322/caac.21660
15. Torres-Rosas R, Torres-Gómez N, Hernández-Juárez J, Pérez-Cervera Y, Hernández-Antonio A, Argueta-Figueroa L. Reported epidemiology of cancer of the lip, oral cavity and oropharynx in Mexico. *Rev Med Inst Mex Seguro Soc.* 2020;58(4):494–507. doi:10.24875/RMIMSS.M20000075
16. Zini A, Czerninski R, Sgan-Cohen HD. Oral cancer over four decades: Epidemiology, trends, histology, and survival by anatomical sites. *J Oral Pathol Med.* 2010;39(4):299–305. doi:10.1111/j.1600-0714.2009.00845.x
17. Bagan J, Sarrion G, Jimenez Y. Oral cancer: Clinical features. *Oral Oncol.* 2010;46(6):414–417. doi:10.1016/j.oraloncology.2010.03.009
18. Gómez I, Seoane J, Varela-Centelles P, Diz P, Takkouche B. Is diagnostic delay related to advanced-stage oral cancer? A meta-analysis. *Eur J Oral Sci.* 2009;117(5):541–546. doi:10.1111/j.1600-0722.2009.00672.x
19. Abati S, Bramati C, Bondi S, Lissani A, Trimarchi M. Oral cancer and precancer: A narrative review on the relevance of early diagnosis. *Int J Environ Res Public Health.* 2020;17(24):9160. doi:10.3390/ijerph17249160
20. Carreras-Torras C, Gay-Escoda C. Techniques for early diagnosis of oral squamous cell carcinoma: Systematic review. *Med Oral Patol Oral Cir Bucal.* 2015;20(3):e305–e315. doi:10.4317/medoral.20347
21. Chen S, Forman M, Sadow PM, August M. The diagnostic accuracy of incisional biopsy in the oral cavity. *J Oral Maxillofac Surg.* 2016;74(5):959–964. doi:10.1016/j.joms.2015.11.006
22. Welikala RA, Remagnino P, Lim JH, et al. Automated detection and classification of oral lesions using deep learning for early detection of oral cancer. *IEEE Access.* 2020;8:132677–132693. doi:10.1109/ACCESS.2020.3010180
23. James BL, Sunny SP, Heidari AE, et al. Validation of a point-of-care optical coherence tomography device with machine learning algorithm for detection of oral potentially malignant and malignant lesions. *Cancers (Basel).* 2021;13(14):3583. doi:10.3390/cancers13143583
24. Sunny S, Baby A, James BL, et al. A smart tele-cytology point-of-care platform for oral cancer screening. *PLoS One.* 2019;14(11):e0224885. doi:10.1371/journal.pone.0224885
25. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ.* 2021;372:n71. doi:10.1136/bmj.n71
26. Higgins JP, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions.* 2nd ed. Hoboken, NY: John Wiley & Sons; 2019. doi:10.1002/9781119536604
27. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: Reporting guideline. *BMJ.* 2020;368:l6890. doi:10.1136/bmj.l6890
28. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155(8):529–536. doi:10.7326/0003-4819-155-8-201110180-00009
29. Jubair F, Al-Karadsheh O, Malamos D, Al Mahdi S, Saad Y, Hassona Y. A novel lightweight deep convolutional neural network for early detection of oral cancer. *Oral Dis.* 2022;28(4):1123–1130. doi:10.1111/odi.13825
30. Tanriver G, Tekkesin MS, Ergen O. Automated detection and classification of oral lesions using deep learning to detect oral potentially malignant disorders. *Cancers (Basel).* 2021;13(11):2766. doi:10.3390/cancers13112766
31. 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 Med.* 2021;50(9):911–918. doi:10.1111/jop.13227
32. Fu Q, Chen Y, Li Z, et al. A deep learning algorithm for detection of oral cavity squamous cell carcinoma from photographic images: A retrospective study. *EclinicalMedicine.* 2020;27:100558. doi:10.1016/j.eclinm.2020.100558
33. Abbaci M, Breuskin I, Casiraghi O, et al. Confocal laser endomicroscopy for non-invasive head and neck cancer imaging: A comprehensive review. *Oral Oncol.* 2014;50(8):711–716. doi:10.1016/j.oraloncology.2014.05.002
34. Aubreville M, Knipfer C, Oetter N, et al. Automatic classification of cancerous tissue in laserendomicroscopy images of the oral cavity using deep learning. *Sci Rep.* 2017;7(1):11979. doi:10.1038/s41598-017-12320-8
35. Lu G, Fei B. Medical hyperspectral imaging: A review. *J Biomed Opt.* 2014;19(1):10901. doi:10.1117/1.JBO.19.1.010901
36. 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(4):829–837. doi:10.1007/s00432-018-02834-7
37. Podoleanu AG. Optical coherence tomography. *J Microsc.* 2012;247(3):209–219. doi:10.1111/j.1365-2818.2012.03619.x
38. Fujimoto JG, Pitris C, Boppart SA, Brezinski ME. Optical coherence tomography: An emerging technology for biomedical imaging and optical biopsy. *Neoplasia.* 2000;2(1–2):9–25. doi:10.1038/sj.neo.7900071
39. Tang Y, Anandasabapathy S, Richards-Kortum R. Advances in optical gastrointestinal endoscopy: A technical review. *Mol Oncol.* 2021;15(10): 2580–2599. doi:10.1002/1878-0261.12792
40. Yang YJ, Bang CS. Application of artificial intelligence in gastroenterology. *World J Gastroenterol.* 2019;25(14):1666–1683. doi:10.3748/wjg.v25.i14.1666
41. Yang EC, Brenes DR, Vohra IS, et al. Algorithm to quantify nuclear features and confidence intervals for classification of oral neoplasia from high-resolution optical images. *J Med Imaging (Bellingham).* 2020;7(5):054502. doi:10.1117/1.JMI.7.5.054502
42. Sharma N, Om H. Usage of probabilistic and general regression neural network for early detection and prevention of oral cancer. *ScientificWorldJournal.* 2015;2015:234191. doi:10.1155/2015/234191
43. Narayana PA, Coronado I, Sujit SJ, Wolinsky JS, Lublin FD, Gabr RE. Deep-learning-based neural tissue segmentation of MRI in multiple sclerosis: Effect of training set size. *J Magn Reson Imaging.* 2020;51(5):1487–1496. doi:10.1002/jmri.26959
44. Fang Y, Wang J, Ou X, et al. The impact of training sample size on deep learning-based organ auto-segmentation for head-and-neck patients. *Phys Med Biol.* 2021;66(18). doi:10.1088/1361-6560/ac2206
45. Samala RK, Chan HP, Hadjiiski L, Helvie MA, Richter CD, Cha KH. Breast cancer diagnosis in digital breast tomosynthesis: Effects of training sample size on multi-stage transfer learning using deep neural nets. *IEEE Trans Med Imaging.* 2019;38(3):686–696. doi:10.1109/TMI.2018.2870343