Review Article

Deep Learning in the Diagnosis and Prognosis of Oral Potentially Malignant Disorders

Xin-Lei Li[1] and Gang Zhou[1,2]*

[1]*The State Key Laboratory Breeding Base of Basic Science of Stomatology (Hubei-MOST) & Key Laboratory of Oral Biomedicine Ministry of Education, School and Hospital of Stomatology, Wuhan University, Wuhan, Hubei, China;* [2]*Department of Oral Medicine, School and Hospital of Stomatology, Wuhan University, Wuhan, Hubei, China*

Abstract

Oral potentially malignant disorders (OPMDs), characterized by a wide variety of types and diverse clinical manifestations, have always been difficult to diagnose and differentiate. All of them carry a risk of malignant transformation. In addition to pathological examination, which remains the gold standard, various auxiliary diagnostic tests are used in clinical practice. Deep learning, a branch of artificial intelligence, has been applied to medical image analysis. Among deep learning techniques, convolutional neural networks are commonly used for image segmentation, detection, classification, and computer-aided diagnosis. We reviewed several image analysis methods based on deep learning neural networks for the diagnosis and prognosis of OPMDs, including photographic images, autofluorescence images, exfoliative cytology images, histopathological images, and optical coherence tomography images. Additionally, we assessed the current limitations and challenges in applying deep learning to the diagnosis of OPMDs.

Introduction

Oral cancer is one of the most dangerous forms of cancer.[1] Many patients with oral cancer do not receive treatment until the disease has reached an advanced stage, resulting in a poor prognosis. Early screening and diagnosis of oral cancer can improve survival rates and reduce unnecessary costs.[2] The development of oral cancer is a long process that passes through a "precancerous" period. Oral potentially malignant disorders (OPMDs), as defined by the World Health Organization in 2020, are any oral mucosal abnormalities associated with a statistically increased risk of developing oral cancer.[3] These include oral leukoplakia, proliferative verrucous leukoplakia (PVL), erythroplakia, oral submucosal fibrosis, oral lichen planus (OLP), actinic keratosis, palatal changes in reverse smokers, discoidal lupus erythematosus, dyskeratosis congenital, oral lichenoid lesions, and oral chronic graft-versus-host disease. OPMDs is considered a general concept because all 11 diseases carry a risk of progression to oral cancer, although not all cases

eventually become malignant.[3] Therefore, it is crucial to identify lesions that are likely to undergo malignant transformation. Visual examination is the most common and intuitive method for this purpose. Additionally, special staining and fluorescence examination are also used to identify OPMDs.[4–6] Pathological examination remains the gold standard for diagnosing OPMDs, but it has the drawbacks of invasiveness and operational difficulty.[7] Furthermore pathologistsmust spend considerable time and effort analyzing images from pathological slides, and the influence of the examiners subjective judgment cannot be ignored.[8]

The concept of artificial intelligence (AI) was first proposed in the 1950s. AI, which originates from computer science, refers to a set of theories, methods, technologies, and application systems that simulate, extend, and enhance human intelligence. AI attempt to develop computer systems that imitate human work and thought processes. The strength of AI lies in its ability to learn and identify patterns and relationships from large, multidimensional, and multimodal data sets. However, due to limitations in computing power and data availability, the application of AI has faced many challenges.[9] Machine learning is one of the derivative technologies of AI and is an essential condition for machines to achieve intelligence.[10] Deep learning (DL) enables the processing of more complex data by increasing the number of hidden layers in artificial neural network algorithm. It represents a more advanced stage of machine learning. DL specializes in discovering complex structures in multi-dimensional data and extracting features, overcoming the data size limitations inherent in traditional ML.[11] DL tech-
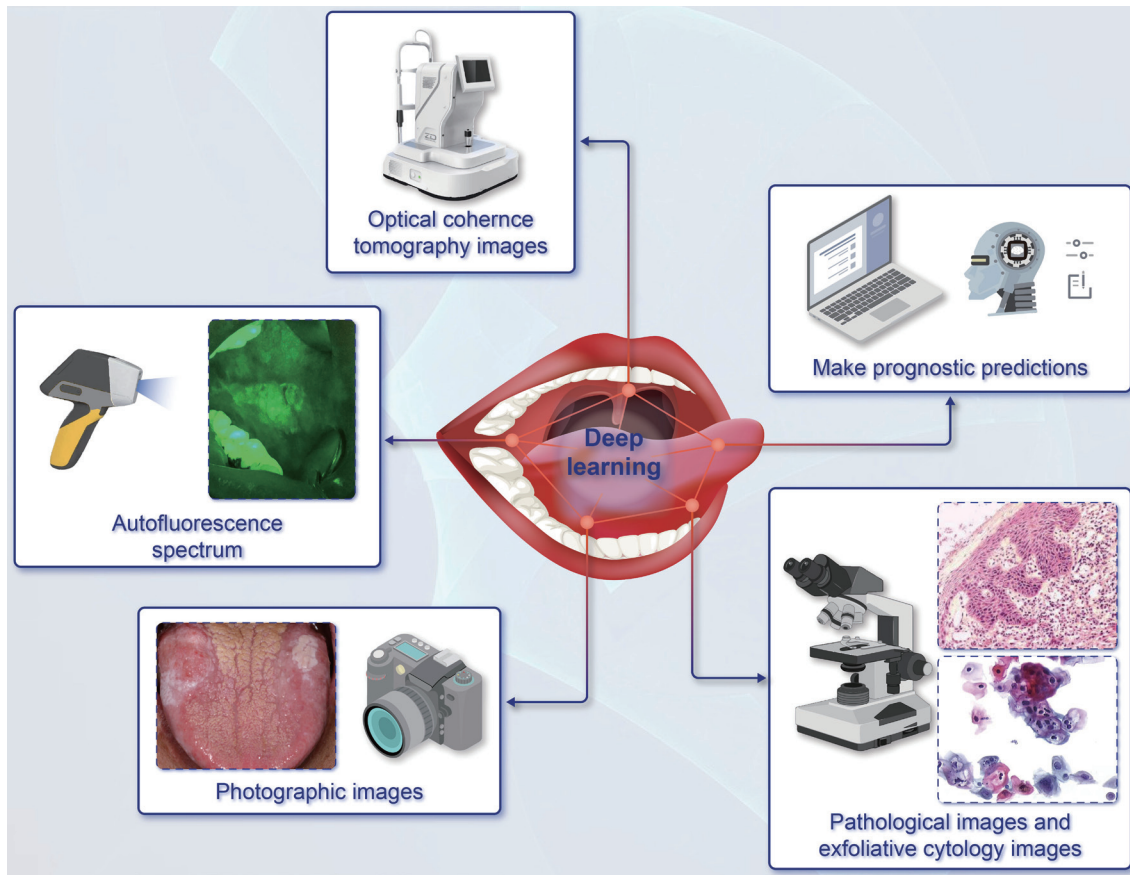
Fig. 1. Deep learning models can be applied to assist in the diagnosis and prognosis of oral potentially malignant disorders.

nologies include deep belief network convolutional neural network (CNN), and recurrent neural network. The advantage of CNN is that it can directly input data such as images, automatically segment and extract features from each part of the data, and integrate convolution layers for data processing, thus avoiding errors caused by manual input and minimizing susceptibility to interference. The number of image features that can be recognized by CNN has increased significantly. As a result, the introduction of CNN has brought substantial improvements to fields such as image processing and natural language processing, and holds great potential for medical applications requiring image recognition, such as medical imaging and histopathology.[12] With the rapid development of AI-related technology, its application in the medical field is expected to become increasingly widespread. This review provides an overview of emerging DL techniques and their applications in the diagnosis and prognosis of OPMDs.

Application of DL in the auxiliary diagnosis and screening of OPMDs

There are various clinical examination methods for OPMDs, such as toluidine blue staining, autofluorescence, exfoliative cytology, chemiluminescence, and pathological examination. The results of these tests are typically reviewed by experienced doctors, which can make the diagnosis subjective. DL technology, represented by CNN algorithm, has unique advantages in image processing, which is helpful for the DL of pathological or imaging image clas-

sification. The internal structure of CNNs is achieved by providing data, mainly model parameters, in a convolutional and repetitive manner. The training process is repeated until incremental improvements in model detection ability allow the input image to be mapped to a specific label.[13] With the assistance of CNN, the computer automatically identifies the best features to match the target image by predesigning the learned feature variables and directly classifies the image without relying on a large amount of data preprocessing or human operation and interference. AI technology has been widely used in clinical practice. Figure 1 demonstrates that it can assist doctors in the diagnosis of oral cancer and precancerous lesions bycompressing, enhancing, reducing, matching, describing, and recognizing various images.[14]

DL in clinical photographic images of OPMDs

At present, visual examination and palpation are the most routine methods for detecting oral cancer and OPMDs. These two examination methods are intuitive and convenient but require a high level of diagnostic ability and experience from the doctor. Some less developed regions may not have specialists available to diagnose OPMDs by conventional methods.

Algorithms such as DL-based CNN and DenseNet have been used to detect lesions in clinical images of the skin and larynx with results comparable to those of experts.[15–17] However, identifying OPMDs from photographic images is much more difficult than classifying skin lesions because mucosal lesions are often hidden or masked in a complex background by overlapping

teeth, tongue, and palate.[18] Despite facing more challenges, some DL algorithms have been successfully applied to clinical photographic image detection and analysis of OPMDs. In general, DL models are divided into segmentation models, image classification models, and object detection models. Segmentation models are used to distinguish the lesion area from normal tissue, image classification models usually identify whether the lesion is cancerous, and object detection models diagnose the lesion as a more specific disease. For instance, Warin *et al*.[19] used DenseNet-169, ResNet-101, SqueezeNet, and Swe-S as classification models and adopted Faster R-CNN, YOLOv5, RetinaNet, and CenterNet2 as detection models to analyze 980 images (365 oral squamous cell carcinoma [OSCC] images, 315 OPMDs images, and 300 benign lesions images). Compared with the results of oral and maxillofacial surgeons and general practitioners, the sensitivity and specificity of DenseNet-169 and ResNet-101 classification models were better than those of specialists and general practitioners. The sensitivity of R-CNN was between that of specialists and general practitioners, indicating that the CNN algorithm-based model has an expertlevel ability to distinguish oral cancer and OPMDs from benign lesions.[19] Keser *et al*.[20] developed a CNN model based on GoogleNet Inception V3 to identify photographic images of OLP. They found that this model was highly effective in distinguishing normal mucosa from OLP lesions. Tanriver *et al*.[21] realized that excellent object detection models and classification models could be concatenated for identifying oral lesions. In their study, they proposed an end-to-end two-stage model combining the classification model EfficientNet-b4 with the detection model YOLOv5l to classify lesions into three categories: benign, OPMDs, and oral cancer. This is a low-cost OPMDs screening method that can automatically detect and classify various types of oral lesions in realtime.[21]

Ferrer-Sánchez *et al*.[22] used a U-Net-based lesion segmentation model and a multi-task CNN classifier model to analyze 261 clinical photographs of oral leukoplakia to predict the risk of epithelial dysplasia and malignant transformation. One of the innovations of this study was the construction of two attention heatmaps of the images to interpret the predictions made by the model. The results showed that for predicting malignant transformation, the model achieved a sensitivity of 1 and specificity of 0.692, while for predicting high-risk dysplasia, the model achieved a specificity of 0.740 and sensitivity of 0.928. These two attention heatmaps, explaining the risk of malignant transformation and epithelial dysplasia, respectively, greatly increased the confidence in the prediction model.[22]

With the popularity of smartphones, mobile phones have become one of the most convenient tools for taking pictures. In a retrospective study by Lin *et al*.,[23] they used the HRNet model based on the CNN algorithm to analyze mucosal lesion images taken using smartphones. After standardized training in photographic methods, Lin *et al*.[23] used the HRNet model to analyze 688 images of mucosal lesions (251 recurrent aphthous ulcers, 231 low-risk OPMDs, 141 high-risk OPMDs, and 65 oral cancer) and 760 images of normal mucosa. The results showed that the sensitivity of the HRNet model was better than other classification models, such as VGG16, ResNet50, DenseNet169, and HRNET-W18, though it still misdiagnosed 1.9% of high-risk OPMDs as recurrent aphthous ulcer.[23] The missed diagnosis of high-risk diseases is more harmful than the misdiagnosis of low-risk diseases in screening, as patients may lose the best opportunity for treatment. Despite the good results of the classification model in Lin's study, the ability of the CNN model to analyze images taken by smartphones compared with specialists could not be determined. Without further optimiza-

tion of the photographing method or classification model, smartphones cannot replace the role of digital single-lens reflex cameras in image acquisition.

Application to autofluorescence spectrum analysis

Autofluorescence is a safe and convenient method for screening OPMDs and early oral cancer. When normal tissues are exposed to blue light, they absorb part of the photon energy and emit lower-energy photons, a phenomenon known as autofluorescence. The fluorescence-producing molecules involved are mainly nicotinamide adenine dinucleotide, flavin adenine dinucleotide, and some elastin, which makes the image of normal tissue appear with green fluorescence. Abnormal tissues, due to changes in porphyrin metabolism and the breakdown of elastin, emit less fluorescence in areas at risk of malignant transformation, showing a lack of fluorescence, which appears black in the image.[24] In a study by Morikawa, the evaluation of autofluorescence images was entirely subjective, showing high sensitivity (98.0%) and low specificity (43.2%) for detecting squamous cell carcinoma. Factors such as inflammation may interfere with the findings. Therefore, Morikawa concluded that a new evaluation method should be developed to obtain an objective evaluation.[25]

van Staveren *et al*.[26] trained a DL model based on artificial neural network algorithms to analyze the autofluorescence spectrum of oral leukoplakia to determine its characteristics and grade of dysplasia. The results of the study showed that DL could effectively distinguish between the autofluorescence spectra of leukoplakia and normal tissue, as well as between homogeneous and heterogeneous leukoplakia.[26] Although the interpretation of autofluorescence spectra by artificial neural network algorithms may not necessarily be the best method for evaluating autofluorescence detection, this study found that the edges of areas with fluorescence loss may not be distinguishable by the naked eye. Real-time intraoperative detection of the spectrum at each pixel can help the physician determine the extent of resection required for oral leukoplakia.

DL technology combined with exfoliative cytology

Exfoliative cytology is considered an effective method for mass screening of high-risk populations, and its role in cervical cancer screening has been well established.[27] Previous research from our group has confirmed that exfoliative cytology hassignificant potential as an accurate and simple diagnostic method for clinically suspected oral precancer and oral cancer.[28] In the traditional exfoliative cytology test, oral mucosal brush specimens are directly applied to glass slides for staining and observation. The analysis of cell morphology is a burdensome task that relies on experienced specialists.[29,30]

Sunny *et al*.[31] used the Inception V3 model to analyze exfoliative cytology images of oral cancer and OPMDs, which was employed to automatically diagnose and risk-grade OPMDs. The test results of this model showed sensitivity and specificity of 73% and 100%, respectively. The study concluded that the tele-cytology platform combined with the CNN model improved accuracy by 30% compared to the traditional manual method.[31]

Exfoliative cytology is a qualitative rather than quantitative method for OPMD detection. Liu *et al*.[32] developed the oral cancer risk index 2 (OCRI2) for the quantitative evaluation of leukoplakia. In this study, the peak random forest model was used to analyze the exfoliative cytology results of 68 patients with oral leukoplakia and calculate the corresponding OCRI2 values. The results showed that an OCRI2 value of 0.5 was the best threshold for predicting

the malignant transformation of leukoplakia (OCRI2 values higher than 0.5 indicated high-risk leukoplakia, while values lower than 0.5 indicated low-risk leukoplakia). Using OCRI2 to quantitatively predict the risk of malignant transformation of leukoplakia based on exfoliative cytology can reduce the cost of patient follow-up.[32] OLP may be severe on one side of the buccal mucosa and mild on the other side, but the mild side may still have a risk of malignant transformation.[3,33,34] Despite this, it is clinically difficult to perform a pathological examination of the minor side mucosa, and noninvasive exfoliative cytology is acceptable. Therefore, the mild side mucosa of patients with OLP can be screened by DL-assisted exfoliative cytology when necessary to determine the subsequent treatment plan.

DL in pathological images of OPMDs

Many studies have shown that DL algorithms can assist pathologists in the diagnosis of oral malignant tumors.[35] For instance, Aubreville *et al*.[36] used the LeNet-5 model to identify confocal laser endomicroscopy images of OSCC, showing that the average accuracy was 88.3%, sensitivity was 86.6%, and specificity was 90%, which was even better than classifiers based on conventional AI algorithms.

Pathological examination is the gold standard for the diagnosis of OPMDs. AI-based analysis of histological images can minimize human interference and reduce the subjective influence of pathologists. Epithelial dysplasia is an important feature of malignant transformation in OPMDs, and the nuclei of these abnormal epithelial cells undergo varying degrees of change. Alshawwa *et al*.[37] used multiple CNN models to analyze changes in nuclear entropy in tissue sections to identify oral leukoplakia and PVL. The Mask R-CNN model was used to segment the nucleus images and extract image features. The average accuracy of the model in segmenting the nuclei of leukoplakia, PVL, and SCC was 92.95%. The use of polynomial classifiers to distinguish leukoplakia and PVL also achieved good results, with average sensitivity of 95.83%, average specificity of 98.29%, and average accuracy of 97.05%, demonstrating that the classifier could distinguish between the two lesions.[37,38] Idrees *et al*.[38] combined whole slide imaging with computer-aided image analysis technology to construct an artificial neural network-multilayer perceptron to analyze and identify OLP. The principle of this study is that the total number of monocytes and granu locytes increases in OLP lesions. The results showed that the model could determine the critical point between OLP and other lichi-like diseases based on the number of inflammation cells, with a sensitivity of 100% and an accuracy of 94.62%.[38] It is well known that AI mainly relies on the segmentation and identification of nuclei to analyze pathological images. Therefore, compared with the interpretation of nuclear atypia, the detection of OLP by inflammatory cell count is less reliable due to the difficulty in differentiating it from other chronic inflammatory diseases, and related studies are rarely reported.

DL for optical coherence tomography (OCT) images

OCT is a non-invasive and radiation-free real-time imaging method. The advantage of OCT is that it provides real-time images with a resolution comparable to that of pathological examination, which can be used to determine the edge of the tumor.[39,40] The principle of OCT is similar to that of ultrasound examination, where electromagnetic waves of a certain wavelength pass through the tissue, and the optical properties of the tissue determine the optical path and depth of the light. Compared with normal tissues, tumors and precancerous lesions have a larger nuclear/plasma ratio, widened

epithelial spikes, and a thickened basement membrane, leading to speckle-like images on OCT.[41] The major drawback of OCT is that its evaluation depends on the operator's professional knowledge. Since the imaging configuration differs from conventional radiological examinations, many specialists find it difficult to evaluate the image reports.

James *et al*.[42] used a DL-based Support Vector Machine (SVM) to automatically diagnose 347 OCT images (151 normal oral mucosa, 121 OPMDs, and 75 malignant lesions). The model achieved a sensitivity of 93% and specificity of 74% for OPMDs, and a sensitivity of 95% and specificity of 76% for malignant tumors. The Inception-ResNet-v2-SVM model showed the highest sensitivity (83%) in differentiating mild dysplasia from moderate to severe dysplasia. These studies indicate that the accuracy of the SVM model in interpreting OCT images is comparable to that of pathological examination.[42] Heidari *et al*.[43] developed a CNN model to classify OCT images of OSCC and dysplasia in order to determine the boundary between normal and abnormal tissues in 3D images. Surprisingly, the sensitivity and specificity of the model were 100% and 70%, respectively, while the sensitivity and specificity of expert pathologists were 85% and 78%, respectively. Although the performance of the CNN model is close to that of pathologists, the model may misjudge uninvolved tissue as abnormal.[43]

Application of DL in predicting and monitoring the prognosis of OPMDs

With the advent of AI in medicine, DL techniques have become a common method for predicting disease development and outcomes based on informative data. Many studies have reported the use of DL algorithms to predict tumor development, including malignant transformation, lymph node metastasis, and prognosis.[44,45] These algorithms learn from health data to provide automatic and exclusive prediction or classification of clinical outcomes without direct programming by the user. Many products based on this technology are used in precision medicine to support clinical decision-making and encourage individualized treatment choices for patients.[46–48]

For OPMDs, it is important to estimate the rate of malignant transformation of lesions when designing treatment. Unfortunately, the malignant transformation rate of many OPMDs, including oral leukoplakia, is highly variable. Clinicians should consider the location, size, color, and other characteristics of the lesion, as well as the grade of epithelial dysplasia, to analyze the prognosis. Most models used to predict the prognosis of oral cancer focus more on tumor metastasis, clinical outcomes, and treatment effects, while they often overlook the malignant transformation of OPMDs. Some models that claim to predict the rate of malignancy in OPMDs simply classify the predictions into high- and low-risk categories, rather than including real-time data as a dynamic variable of transformation probability over time. Obviously, the latter has greater clinical application value but is less explored in research.

Wang *et al*.[49] developed two random forest classification models: the baseline model and the personalized model (model-P), to predict the malignant risk level of OPMDs. This study collected personal information (age, gender, lifestyle, and lesion status), non-invasive oral examinations (toluidine blue staining and autofluorescence), oral tissue biopsies, histopathological analysis, and treatment options. After comparing the performance of the baseline model, model-P, and clinical experts in predicting the risk of malignant transformation, the specificity of all three was comparable (about 90%), while the sensitivity of model-P was better than the other two (over 80%). The study also found that the scope

of autofluorescence loss, toluidine blue staining score, and degree of lesion infiltration were important factors affecting carcinogenesis.[49] Adeoye *et al*.[50] compared three DL models (DeepSurv, time-dependent neural network Cox model, and DeepHit) with two conventional statistical models (random survival forest and Cox proportional hazards) to predict the risk probability of malignant transformation of OPMDs. The predictors entered into these models were gender, age, history of smoking and drinking, abstinence from bad habits, history of cancer, family history, hypertension, hyperlipidemia, diabetes, autoimmune disease, viral hepatitis, and the location and classification of the lesion. The DeepSurv algorithm was found to have the best discrimination performance when simulating the malignant transformation risk of oral leukoplakia and oral lichenoid lesions. However, the conventional random survival forest model outperformed other models in probability calibration.[50]

Difficulties in image analysis and AI algorithms

Automatic classification of medical images is one of the main applications of AI in the medical field. A recent review of AI/ML-based medical devices approved in the United States and Europe between 2015 and 2020 found 126 devices approved or CEmarked for radiological use in Europe and 129 devices in the United States, both representing more than half of the total reviewed.[51] However, the identification and analysis of medical images by AI require the corresponding image database. Unfortunately, large databases of OPMDs and oral cancer have not been established for either clinical photographic images or pathological images, which limits the promotion of current AI-related technologies. Although Warin *et al*. found that DenseNet121, a newly developed deep CNN model, can reduce task overfitting caused by small databases, the ongoing challenges of database limitations still cannot be ignored.[47,52,53] For this reason, Davenport *et al*.[47] claim that humans are unlikely to see substantial changes in healthcare employment due to AI over the next 20 years. To address low-resolution images, Chen *et al*.[54] proposed a joint framework called SRFBN+, which contains a novel transfer learning strategy and a deep super-resolution framework for generating high-resolution slice images from low-resolution slice images. The test results show that SRFBN+ performs well in generating super-resolution pathological examination images.[54] Most of the DL-based models for detecting OPMDs rely on various clinical images, and the quality and standardization of these images present a major challenge for future AI applications.

The CNN algorithm has unique advantages in image processing, which is helpful for DL of pathological or photographic image classification. With the assistance of CNN, the computer automatically identifies the best features to match the target image by pre-designing the learned feature variables and directly classifies the image without relying on a large amount of data preprocessing or human operation.[14] However, CNN has two obvious drawbacks: first, training and testing the model takes a lot of time, and second, details may be lost due to glare when processing images. It is crucial to test how well the model works, and evaluation includes three aspects: statistical validity, clinical utility, and economic utility.[48] Therefore, improving the time-consuming problem remains difficult. The loss of detail in resolution is related to the normalization method and the convolutional layer architecture of the CNN. Both the imaging equipment and methods affect the quality of the images, which in turn impacts the auxiliary diagnosis results of the AI model. This calls for the development of a set of image quality standards similar to those being developed for imaging. The

problem of convolutional layer architecture can only be improved with the help of other DL algorithms, such as Fully Convolutional Networks.[35]

For the algorithm itself, it is better to propose a new algorithm than to optimize the one already in use, as the latter may lead to the conclusion that some subdomain algorithms are not improving.[35] The development of new algorithms requires creating a multidisciplinary team that includes computer and social scientists, operations and research leaders, clinical stakeholders (physicians, caregivers, and patients), and experts in related disciplines. However, algorithm optimization must take various trade-offs into account due to the limitations of data and computing power.

Currently, some studies use manual cropping of the region of interest to preprocess images, while other studies use segmentation models to divide images into patches of a certain size to avoid errors in manual process.[55] However, the limitation of this method is that the network cannot directly analyze the image as a whole and can only focus on a small part, which may also account for the loss of details. Wirtz *et al*.[56] believe that image segmentation is limited by the amount of data and computing power. In the future, algorithms that can analyze the whole image without human intervention in the preprocessing stage should be developed, allowing for more accurate results per unit of computational cost.[56]

Discussion

In an era of increasing digitization of human health data, DL is expected to play a significant role in the development, validation, and implementation of decision support tools to facilitate precision medicine. In this review, we have demonstrated many promising applications of DL in various fields of OPMDs, including digital clinical photographic image analysis, autofluorescence spectroscopy analysis, exfoliated cell examination, pathological section analysis, and OCT image analysis for auxiliary diagnosis. Additionally, we discussed the performance of different DL models in predicting the prognosis of OPMDs (Table 1).[19–23,26,31,32,37,38,42,43,49,50] Currently, most DL models for the auxiliary diagnosis of OPMDs appear to have more potential in screening, and the use of AI technology to interpret exfoliative cytology is one of the most convenient and promising screening methods. The promotion of these technologies could address some of the difficulties associated with the diagnosis of OPMDs in medically underserved areas, potentially reducing the workload of clinicians and pathologists and minimizing the interference of their subjectivity. However, DL models still face challenges such as insufficient image databases, low resolution of images, and limitations in the performance of the algorithms themselves. In the future, there may be technologies that diagnose OPMDs based on salivary markers, with corresponding DL models potentially offering more meaningful screening. Another emerging frontier is histologic inference of genomic features. As research progresses, the future of DL applications in the field of OPMDs will likely focus on multimodal learning to integrate medical images and omics data to identify biologically meaningful biomarkers.[57] AI algorithms for predicting patient prognosis have long been applied in oncology,[58–60] and even some AI models in commercial treatment planning systems, such as RapidPlan and Auto-Planning, can automatically plan and design treatment strategies.[61,62] Unfortunately, few of these frontier studies have been applied to the field of OPMDs. In the future, integrated AI algorithms capable of analyzing mucosal lesions and automatically generating diagnostic reports and evaluating treatment options may be developed. Rapidly evolving DL technologies are expected to continue

Table 1. Summary of DL models for OPMD diagnosis and prognosis

Num-ber	Author (Year, Country)	Purpose	Learning machine	Sample type	Sample number or information	Main outcome	Refer-ences
1	Warin *et al* (2022, Thailand)	To use a novel deep convolutional neural network for early detection of oral cancer and oral lesions.	DenseNet-169, ResNet-101, SqueezeNet, Swe-S, Faster R-CNN, YOLOv5, RetinaNet, CenterNet2	Clinical photographic image	365 images of OSCC, 315 images of OPMDs, 300 images of benign lesions	The model based on CNN algorithm had expert level in distinguishing oral cancer and OPMDs from benign lesions.	19
2	Keser *et al* (2022, Turkey)	To develop a deep learning method to identify oral lichen planus lesions using photographic images.	GoogleNet Inception V3	Clinical photographic image	65 images of healthy oral mucosa, 72 images of the buccal mucosa of OLP	The deep learning model provided classification of all tested images of healthy and diseased mucosa with 100% success rate.	20
3	Tanriver *et al* (2021, Turkey)	The potential application of computer vision techniques in the field of oral cancer within the range of photographic images, and the promise of an automated system for the detection of OPMD was investigated.	EfficientNet-b4, YOLOv5l	Clinical photographic image	162 images of OSCC, 248 images of OPMDs, 274 images of benign lesions	Preliminary results demonstrate the feasibility of the deep learning-based method for real-time automatic detection and classification of oral lesions.	21
4	Ferrer-Sanchez *et al* (2022, Spain)	To estimate the probability of malignancy in oral leukoplakia using deep learning.	U-Net based lesion segmentation model, multi-task CNN classifier model	Clinical photographic image	261 images of OLK	The sensitivity and specificity of the model in predicting malignant lesions were 1 and 0.692, respectively. The specificity and sensitivity of predicting high-risk dysplasia were 0.740 and 0.928, respectively.	22
5	Lin *et al* (2021, China)	A DLD based smartphone image diagnosis method for automatic detection of oral diseases	HRNet-W18	Clinical photographic image	251 images of ulcer, 231 images of low-risk OPMDs, 141 images of high-risk OPMDs, 65 images of oral cancer, 60 images of normal mucosa	The sensitivity of HRNet model was better than that of VGG16, ResNet50, DenseNet169 and HRNET-W18 classification models.	23

(continued)

Table 1. *(continued)*

Num-ber	Author (Year, Country)	Purpose	Learning machine	Sample type	Sample number or information	Main outcome	Refer-ences
6	van Staveren *et al* (2000, Netherlands)	A deep learning model based on artificial neural network algorithm was used to analyze the autofluorescence spectra of oral leukoplakia to determine the characteristics and grading of dysplasia.	Trained ANN model	Autofluorescence spectrum	22 images of OLK, 6 images of normal mucosa	Abnormal tissues can be distinguished from normal tissues by neural networks with a sensitivity of 86% and a specificity of 100%. In addition, the classification of homogeneous or heterogeneous tissue performed reasonably well.	26
7	Sunny *et al* (2019, India)	The risk stratification model based on ANN was used for the early detection of OPMDs.	Inception V3 model	Exfoliative cytology image	981 images	The sensitivity and specificity of the model were 73% and 100%, respectively. The remote cytology platform combined with CNN model improved the accuracy by 30% compared with the traditional manual method.	31
8	Liu *et al* (2017, China)	An oral cancer risk index using DNA index values was developed to quantitatively assess cancer risk in patients with oral leukoplakia.	Peak random forest model	Exfoliative cytology image	18 normal, 28 OLK, 41 OSCC subjects	OCRI2 can distinguish between low-risk and high-risk OLK and may improve the cost-effectiveness of patients with OLK during clinical follow-up.	32
9	Alshawwa *et al* (2022, Saudi Arabia)	Tissue sections were analyzed using multiple CNN models to differentiate oral leukoplakia from proliferative verrucous leukoplakia.	Mask R-CNN, polynomial classifier	Pathological images	568 OLK images, 45 PVL images, 58 SCC images	The average accuracy of Mask R-CNN model in the nuclear segmentation of leukoplakia, PVL and SCC was 92.95%. The average sensitivity of the polynomial classifier was 95.83%, the average specificity was 98.29%, and the average accuracy was 97.05%, indicating that the classifier could distinguish between the two lesions.	37
10	Idrees *et al* (2021, Australia)	A machine-learning artificial neural network was created to detect OLP by identifying and quantifying monocytes and granulocytes in inflammatory infiltrates in digitized hematoxylin and eosin microscope slides.	ANN-MLP	Pathological images	1,606 images of OLP	The machine learning method was able to reliably detect the critical point between OLP and other lichi-like diseases based on the number of inflammatory cells and monocytes, with a sensitivity of 100% and an accuracy of 94.62%. Artificial intelligence has the potential to improve the accuracy of pathologists in the diagnosis of OLP.	38

(continued)

Table 1. *(continued)*

Number	Author (Year, Country)	Purpose	Learning machine	Sample type	Sample number or information	Main outcome	References
11	James *et al* (2021, India)	A point-of-care optical coherence tomography device was used to detect potentially malignant and malignant lesions in the oral cavity.	Support Vector Machine (SVM), Inception-ResNet-v2-SVM model	OCT images	75 images of malignant lesions, 121 images of OPMDs, 152 images of benign lesions	The sensitivity and specificity of the SVM model for OPMDs were 93% and 74%, and the sensitivity and specificity for malignant tumors were 95% and 76%. The Inception-ResNet-v2-SVM model had the highest sensitivity (83%) in differentiating mild dysplasia from moderate and severe dysplasia. This study shows that the accuracy of the SVM model in OCT image interpretation is close to that of pathological examination.	42
12	Heidari *et al* (2020, USA)	A CNN model was developed to classify OCT images of OSCC and dysplasia to determine the boundaries between normal and abnormal tissues in 3D images.	Trained CNN model	3D OCT images	1,000 images from 6 cases of oral cancer and 1 case of dysplasia	The sensitivity and specificity of the model were 100% and 70%, respectively. This approach has the potential to be used as a real-time analysis tool to assess surgical margins.	43
13	Wang *et al* (2020, China)	To predict the malignant risk level of OPMDs.	Baseline model (model-B), personalized model (model-P)	266 patients with OPMDs	Personal information, non-invasive oral examination, oral tissue biopsy and histopathological analysis, treatment and follow-up	The specificity of the two models was comparable to that of experts (about 90%), while the sensitivity of Model-P was better than the other two models.	49
14	Adeoye *et al* (2021, Hong Kong)	To compare the ability of three deep learning models and two traditional statistical models in predicting the risk probability of malignant transformation of OPMDs.	DeepSurv, time-dependent neural network Cox model, DeepHit, random survival forest, Cox proportional hazards	716 patients with a clinical diagnosis of oral leukoplakia, oral lichen planus, or oral lichenoid lesions	Gender, age, history of smoking, drinking, abstinence from bad living habits, history of cancer, family history, history of hypertension, history of hyperlipidemia, history of diabetes, history of autoimmune disease, history of viral hepatitis, lesion location and classification	DeepSurv and RSF are two excellent performing models based on the discrimination and calibration after internal validation, and DeepSurv is more stable than RSF when cross-validated. Overall, the time-to-event model successfully predicted malignant transformation of oral leukoplakia and oral lichenoid lesions,	50

ANN, artificial neural network; CNN, convolutional neural network; DL, deep learning; DLD, dynamic low-resolution distillation; MLP, multi-layer perceptron; OCT, optical coherence tomography; OLK, oral leukoplakia; OLP, oral lichen planus; OPMDs, oral potentially malignant disorders; OSCC, oral squamous cell carcinoma; PVL, proliferative verrucous leukoplakia; RSF, random survival forest; SCC, squamous cell carcinoma.

having a significant impact on the field of OPMDs in the near future. Both researchers and physicians need to be prepared for this revolutionary era.

Conclusions

This review summarizes five image analysis methods based on deep learning neural networks for the diagnosis of OPMDs and the prediction of malignant risk. Furthermore, the current limitations and future development prospects of deep learning in OPMDs are evaluated. The combination of these emerging technologies and diagnostic methods will change the clinical diagnosis and treatment of OPMDs.

Acknowledgments

Funding

Conflict of interest

The authors declare that they have no conflict of interest.

Author contributions

Writing - original draft (XLL), writing - review & editing, supervision, funding acquisition (GZ). Both authors have approved the final version and publication of the manuscript.

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(3):209–249. doi:10.3322/caac.21660, PMID:33538338.

[2] Baykul T, Yilmaz HH, Aydin U, Aydin MA, Aksoy M, Yildirim D. Early diagnosis of oral cancer. J Int Med Res 2010;38(3):737–749. doi:10.1177/147323001003800302, PMID:20819411.

[3] Warnakulasuriya S, Kujan O, Aguirre-Urizar JM, Bagan JV, González-Moles MÁ, Kerr AR, *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 Dis 2021;27(8):1862–1880. doi:10.1111/odi.13704, PMID:33128420.

[4] Morikawa T, Shibahara T, Nomura T, Katakura A, Takano M. Non-Invasive Early Detection of Oral Cancers Using Fluorescence Visualization with Optical Instruments. Cancers (Basel) 2020;12(10):2771. doi:10.3390/cancers12102771, PMID:32992486.

[5] Simonato LE, Tomo S, Scarparo Navarro R, Balbin Villaverde AGJ. Fluorescence visualization improves the detection of oral, potentially malignant, disorders in population screening. Photodiagnosis Photodyn Ther 2019;27:74–78. doi:10.1016/j.pdpdt.2019.05.017, PMID:31116999.

[6] Tomo S, Miyahara GI, Simonato LE. History and future perspectives for the use of fluorescence visualization to detect oral squamous cell carcinoma and oral potentially malignant disorders. Photodiagnosis Photodyn Ther 2019;28:308–317. doi:10.1016/j.pdpdt.2019.10.005, PMID:31600576.

[7] Farah CS, McIntosh L, Georgiou A, McCullough MJ. Efficacy of tis-

[8] Mehrotra R, Singh M, Thomas S, Nair P, Pandya S, Nigam NS, *et al*. A cross-sectional study evaluating chemiluminescence and autofluorescence in the detection of clinically innocuous precancerous and cancerous oral lesions. J Am Dent Assoc 2010;141(2):151–156. doi:10.14219/jada.archive.2010.0132, PMID:20123872.

[9] Trogdon JG, Falchook AD, Basak R, Carpenter WR, Chen RC. Total Medicare Costs Associated With Diagnosis and Treatment of Prostate Cancer in Elderly Men. JAMA Oncol 2019;5(1):60–66. doi:10.1001/jamaoncol.2018.3701, PMID:30242397.

[10] Niel O, Bastard P. Artificial Intelligence in Nephrology: Core Concepts, Clinical Applications, and Perspectives. Am J Kidney Dis 2019;74(6):803–810. doi:10.1053/j.ajkd.2019.05.020, PMID:3145 1330.

[11] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521 (7553):436–444. doi:10.1038/nature14539, PMID:26017442.

[12] Rabual JR, Dorado J. Artificial Neural Networks in Real-Life Applications. IGI Global; 2005. doi:10.4018/978-1-59140-902-1.

[13] Schwendicke F, Golla T, Dreher M, Krois J. Convolutional neural networks for dental image diagnostics: A scoping review. J Dent 2019;91:103226. doi:10.1016/j.jdent.2019.103226, PMID:31704386.

[14] Reher R, Kim HW, Zhang C, Mao HH, Wang M, Nothias LF, *et al*. A Convolutional Neural Network-Based Approach for the Rapid Annotation of Molecularly Diverse Natural Products. J Am Chem Soc 2020;142(9):4114–4120. doi:10.1021/jacs.9b13786, PMID:32045230.

[15] Xiong H, Lin P, Yu JG, Ye J, Xiao L, Tao Y, *et al*. Computer-aided diagnosis of laryngeal cancer via deep learning based on laryngoscopic images. EBioMedicine 2019;48:92–99. doi:10.1016/j.ebiom.2019.08.075, PMID:31594753.

[16] Phillips M, Marsden H, Jaffe W, Matin RN, Wali GN, Greenhalgh J, *et al*. Assessment of Accuracy of an Artificial Intelligence Algorithm to Detect Melanoma in Images of Skin Lesions. JAMA Netw Open 2019;2(10):e1913436. doi:10.1001/jamanetworkopen.2019.13436, PMID:31617929.

[17] Mohammed MA, Ghani MKA, Arunkumar N, Hamed RI, Abdullah MK, Burhanuddin MA. A real time computer aided object detection of nasopharyngeal carcinoma using genetic algorithm and artificial neural network based on Haar feature fear. Future Gener Comput Syst 2018;89:539–547. doi:10.1016/j.future.2018.07.02.2.

[18] Fu Q, Chen Y, Li Z, Jing Q, Hu C, Liu H, *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, PMID:33150326.

[19] Warin K, Limprasert W, Suebnukarn S, Jinaporntham S, Jantana P, Vicharueang S. AI-based analysis of oral lesions using novel deep convolutional neural networks for early detection of oral cancer. PLoS One 2022;17(8):e0273508. doi:10.1371/journal.pone.0273 508, PMID:36001628.

[20] Keser G, Bayrakdar İŞ, Pekiner FN, Çelik Ö, Orhan K. A deep learning algorithm for classification of oral lichen planus lesions from photographic images: A retrospective study. J Stomatol Oral Maxillofac Surg 2023;124(1):101264. doi:10.1016/j.jormas.2022.08.007, PMID:35964938.

[21] Tanriver G, Soluk Tekkesin M, 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, PMID:34199471.

[22] Ferrer-Sánchez A, Bagan J, Vila-Francés J, Magdalena-Benedito R, Bagan-Debon L. Prediction of the risk of cancer and the grade of dysplasia in leukoplakia lesions using deep learning. Oral Oncol 2022;132:105967. doi:10.1016/j.oraloncology.2022.105967, PMID:35763911.

[23] 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(8):086007. doi:10.1117/1.JBO.26.8.086007, PMID:34453419.

[24] Luo X, Xu H, He M, Han Q, Wang H, Sun C, *et al*. Accuracy of autofluorescence in diagnosing oral squamous cell carcinoma and oral potentially malignant disorders: a comparative study with aerodigestive lesions. Sci Rep 2016;6:29943. doi:10.1038/srep29943,

PMID:27416981.

[25] Morikawa T, Kozakai A, Kosugi A, Bessho H, Shibahara T. Image processing analysis of oral cancer, oral potentially malignant disorders, and other oral diseases using optical instruments. Int J Oral Maxillofac Surg 2020;49(4):515–521. doi:10.1016/j.ijom.2019.08.016, PMID:31500953.

[26] van Staveren HJ, van Veen RL, Speelman OC, Witjes MJ, Star WM, Roodenburg JL. Classification of clinical autofluorescence spectra of oral leukoplakia using an artificial neural network: a pilot study. Oral Oncol 2000;36(3):286–293. doi:10.1016/s1368-8375(00)00004-x, PMID:10793332.

[27] Lee ES, Kim IS, Choi JS, Yeom BW, Kim HK, Han JH, *et al*. Accuracy and reproducibility of telecytology diagnosis of cervical smears. A tool for quality assurance programs. Am J Clin Pathol 2003;119(3):356–360. doi:10.1309/7ytvag4xnr48t75h, PMID:12645336.

[28] Ye X, Zhang J, Tan Y, Chen G, Zhou G. Meta-analysis of two computer-assisted screening methods for diagnosing oral precancer and cancer. Oral Oncol 2015;51(11):966–975. doi:10.1016/j.oraloncology.2015.09.002, PMID:26384539.

[29] Sekine J, Nakatani E, Hideshima K, Iwahashi T, Sasaki H. Diagnostic accuracy of oral cancer cytology in a pilot study. Diagn Pathol 2017;12(1):27. doi:10.1186/s13000-017-0618-3, PMID:28298213.

[30] Sukegawa S, Ono S, Nakano K, Takabatake K, Kawai H, Nagatsuka H, *et al*. Clinical study on primary screening of oral cancer and precancerous lesions by oral cytology. Diagn Pathol 2020;15(1):107. doi:10.1186/s13000-020-01027-6, PMID:32912249.

[31] Sunny S, Baby A, James BL, Balaji D, N V A, Rana MH, *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, PMID:31730638.

[32] Liu Y, Li Y, Fu Y, Liu T, Liu X, Zhang X, *et al*. Quantitative prediction of oral cancer risk in patients with oral leukoplakia. Oncotarget 2017;8(28):46057–46064. doi:10.18632/oncotarget.17550, PMID:28545021.

[33] Cheng YS, Gould A, Kurago Z, Fantasia J, Muller S. Diagnosis of oral lichen planus: a position paper of the American Academy of Oral and Maxillofacial Pathology. Oral Surg Oral Med Oral Pathol Oral Radiol 2016;122(3):332–354. doi:10.1016/j.oooo.2016.05.004, PMID:27401683.

[34] van der Meij EH, van der Waal I. Lack of clinicopathologic correlation in the diagnosis of oral lichen planus based on the presently available diagnostic criteria and suggestions for modifications. J Oral Pathol Med 2003;32(9):507–512. doi:10.1034/j.1600-0714.2003.00125.x, PMID:12969224.

[35] Wang X, Li BB. Deep Learning in Head and Neck Tumor Multiomics Diagnosis and Analysis: Review of the Literature. Front Genet 2021;12:624820. doi:10.3389/fgene.2021.624820, PMID:33643386.

[36] Aubreville M, Knipfer C, Oetter N, Jaremenko C, Rodner E, Denzler J, *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, PMID:28931888.

[37] Alshawwa SZ, Saleh A, Hasan M, Shah MA. Segmentation of Oral Leukoplakia (OL) and Proliferative Verrucous Leukoplakia (PVL) Using Artificial Intelligence Techniques. Biomed Res Int 2022;2022:2363410. doi:10.1155/2022/2363410, PMID:35909480.

[38] Idrees M, Farah CS, Shearston K, Kujan O. A machine-learning algorithm for the reliable identification of oral lichen planus. J Oral Pathol Med 2021;50(9):946–953. doi:10.1111/jop.13226, PMID:34358361.

[39] Hamdoon Z, Jerjes W, McKenzie G, Jay A, Hopper C. Optical coherence tomography in the assessment of oral squamous cell carcinoma resection margins. Photodiagnosis Photodyn Ther 2016;13:211–217. doi:10.1016/j.pdpdt.2015.07.170, PMID:26210067.

[40] Wang J, Xu Y, Boppart SA. Review of optical coherence tomography in oncology. J Biomed Opt 2017;22(12):1–23. doi:10.1117/1.JBO.22.12.121711, PMID:29274145.

[41] Ramezani K, Tofangchiha M. Oral Cancer Screening by Artificial Intelligence-Oriented Interpretation of Optical Coherence Tomography Images. Radiol Res Pract 2022;2022:1614838. doi:10.1155/2022/1614838, PMID:35502299.

[42] James BL, Sunny SP, Heidari AE, Ramanjinappa RD, Lam T, Tran AV, *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, PMID:34298796.

[43] Heidari AE, Pham TT, Ifegwu I, Burwell R, Armstrong WB, Tjoson T, *et al*. The use of optical coherence tomography and convolutional neural networks to distinguish normal and abnormal oral mucosa. J Biophotonics 2020;13(3):e201900221. doi:10.1002/jbio.201900221, PMID:31710775.

[44] Cruz JA, Wishart DS. Applications of machine learning in cancer prediction and prognosis. Cancer Inform 2007;2:59–77. PMID:19458758.

[45] Adeoye J, Tan JY, Choi SW, Thomson P. Prediction models applying machine learning to oral cavity cancer outcomes: A systematic review. Int J Med Inform 2021;154:104557. doi:10.1016/j.ijmedinf.2021.104557, PMID:34455119.

[46] Alabi RO, Youssef O, Pirinen M, Elmusrati M, Mäkitie AA, Leivo I, *et al*. Machine learning in oral squamous cell carcinoma: Current status, clinical concerns and prospects for future-A systematic review. Artif Intell Med 2021;115:102060. doi:10.1016/j.artmed.2021.102060, PMID:34001326.

[47] Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J 2019;6(2):94–98. doi:10.7861/futurehosp.6-2-94, PMID:31363513.

[48] Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc J 2021;8(2):e188–e194. doi:10.7861/fhj.2021-0095, PMID:34286183.

[49] Wang X, Yang J, Wei C, Zhou G, Wu L, Gao Q, *et al*. A personalized computational model predicts cancer risk level of oral potentially malignant disorders and its web application for promotion of non-invasive screening. J Oral Pathol Med 2020;49(5):417–426. doi:10.1111/jop.12983, PMID:31823403.

[50] Adeoye J, Koohi-Moghadam M, Lo AWI, Tsang RK, Chow VLY, Zheng LW, *et al*. Deep Learning Predicts the Malignant-Transformation-Free Survival of Oral Potentially Malignant Disorders. Cancers (Basel) 2021;13(23):6054. doi:10.3390/cancers13236054, PMID:34885164.

[51] Muehlematter UJ, Daniore P, Vokinger KN. Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015-20): a comparative analysis. Lancet Digit Health 2021;3(3):e195–e203. doi:10.1016/S2589-7500(20)30292-2, PMID:33478929.

[52] 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, PMID:34358372.

[53] Warin K, Limprasert W, Suebnukarn S, Jinaporntham S, Jantana P. Performance of deep convolutional neural network for classification and detection of oral potentially malignant disorders in photographic images. Int J Oral Maxillofac Surg 2022;51(5):699–704. doi:10.1016/j.ijom.2021.09.001, PMID:34548194.

[54] Chen J, Ying H, Liu X, Gu J, Feng R, Chen T, *et al*. A Transfer Learning Based Super-Resolution Microscopy for Biopsy Slice Images: The Joint Methods Perspective. IEEE/ACM Trans Comput Biol Bioinform 2021;18(1):103–113. doi:10.1109/TCBB.2020.2991173, PMID:32356757.

[55] Hwang JJ, Jung YH, Cho BH, Heo MS. An overview of deep learning in the field of dentistry. Imaging Sci Dent 2019;49(1):1–7. doi:10.5624/isd.2019.49.1.1, PMID:30941282.

[56] Wirtz A, Mirashi SG, Wesarg S. Automatic Teeth Segmentation in Panoramic X-Ray Images Using a Coupled Shape Model in Combination with a Neural Network. International Conference on Medical Image Computing and Computer-Assisted Intervention; 2018 Sep 16-20; Granada, Spain. Cham: Springer; 2018doi:10.1007/978-3-030-00937-3_81.

[57] Tran KA, Kondrashova O, Bradley A, Williams ED, Pearson JV, Waddell N. Deep learning in cancer diagnosis, prognosis and treatment selection. Genome Med 2021;13(1):152. doi:10.1186/s13073-021-00968-x, PMID:34579788.

[58] Buchner A, Kendlbacher M, Nuhn P, Tüllmann C, Haseke N, Stief CG, *et al*. Outcome assessment of patients with metastatic renal

cell carcinoma under systemic therapy using artificial neural networks. Clin Genitourin Cancer 2012;10(1):37–42. doi:10.1016/j.clgc.2011.10.001, PMID:22056212.

[59] Tseng WT, Chiang WF, Liu SY, Roan J, Lin CN. The application of data mining techniques to oral cancer prognosis. J Med Syst 2015;39(5):59. doi:10.1007/s10916-015-0241-3, PMID:25796587.

[60] Jang BS, Jeon SH, Kim IH, Kim IA. Prediction of Pseudoprogression versus Progression using Machine Learning Algorithm in Glioblastoma. Sci Rep 2018;8(1):12516. doi:10.1038/s41598-018-31007-2, PMID:30131513.

[61] Nawa K, Haga A, Nomoto A, Sarmiento RA, Shiraishi K, Yamashita H, *et al*. Evaluation of a commercial automatic treatment planning system for prostate cancers. Med Dosim 2017;42(3):203–209. doi:10.1016/j.meddos.2017.03.004, PMID:28549556.

[62] Krayenbuehl J, Norton I, Studer G, Guckenberger M. Evaluation of an automated knowledge based treatment planning system for head and neck. Radiat Oncol 2015;10:226. doi:10.1186/s13014-015-0533-2, PMID:26555303.