Received: 31 March 2020 | Revised: 29 April 2020 | Accepted: 9 May 2020

DOI: 10.1111/jop.13042

REVIEW

Journal of Oral Pathology & Medicine WILEY

# The use of artificial intelligence, machine learning and deep learning in oncologic histopathology

Ahmed S. Sultan[1] | Mohamed A. Elgharib[2] | Tiffany Tavares[3] | Maryam Jessri[4] | John R. Basile[1,5]

[1]School of Dentistry, University of Maryland, Baltimore, MD, USA

[2]Max Planck Institute for Informatics, Saarbrucken, Germany

[3]School of Dentistry, University of Missouri-Kansas City, Kansas City, MO, USA

[4]Oral Health Centre of Western Australia, Perth, WA, Australia

[5]University of Maryland Greenebaum Cancer Center, Baltimore, MD, USA

**Correspondence**
Ahmed S. Sultan, School of Dentistry, University of Maryland, Baltimore, MD 21201, USA.
Email: asultan@umaryland.edu

## Abstract

**Background:** Recently, there has been a momentous drive to apply advanced artificial intelligence (AI) technologies to diagnostic medicine. The introduction of AI has provided vast new opportunities to improve health care and has introduced a new wave of heightened precision in oncologic pathology. The impact of AI on oncologic pathology has now become apparent, and its use with respect to oral oncology is still in the nascent stage.

**Discussion:** A foundational overview of AI classification systems used in medicine and a review of common terminology used in machine learning and computational pathology will be presented. This paper provides a focused review on the recent advances in AI and deep learning in oncologic histopathology and oral oncology. In addition, specific emphasis on recent studies that have applied these technologies to oral cancer prognostication will also be discussed.

**Conclusion:** Machine and deep learning methods designed to enhance prognostication of oral cancer have been proposed with much of the work focused on prediction models on patient survival and locoregional recurrences in patients with oral squamous cell carcinomas (OSCC). Few studies have explored machine learning methods on OSCC digital histopathologic images. It is evident that further research at the whole slide image level is needed and future collaborations with computer scientists may progress the field of oral oncology.

KEYWORDS

artificial intelligence, computational pathology, deep learning, digital pathology, head and neck cancer, machine learning, oncologic histopathology

## 1 | BACKGROUND

### 1.1 | Artificial intelligence

The integration of artificial intelligence (AI), machine learning, deep learning, robotics, and Big Data into our physical daily lives has defined the Fourth Industrial Age. Clinicians are now provided with vast new opportunities to improve health care and optimize their approach to patient care. AI can be conceptualized into two main fields, namely artificial general intelligence (AGI) and artificial narrow intelligence (ANI). AGI is generally concerned with the design of machines that can mimic human intelligence with the goal of creating machines that can perform functions similar to humans or even functions that surpass what a human can do (superintelligence).

The peer review history for this article is available at https://publons.com/publon/10.1111/jop.13042

However, most of the progress in AI has been occurring in ANI and a widely used application of ANI is its integration in self-driving cars. The foundational cores of AI are knowledge representation and reasoning (KR$^2$), Bayesian inference, and self-awareness.

## 1.2 | Machine learning

AI broadly encompasses both machine learning and deep learning, the latter of which was most recently introduced. Machine learning is a core discipline of AI that utilizes algorithms that detect patterns within existing data and then trains itself to make predictions on new data.[1,2] Machine learning algorithms are designed to learn different predictive models via a training phase whereby parameters of an artificial neural network model are estimated. The predictive accuracy of the learned model is then tested on blinded data *via* a testing/validation phase to determine the model with the lowest generalization error.[2] The testing/validation phase determines the network's ability to handle and classify unseen data. Modern-day applications of machine learning algorithms include email spam filtering, online advertising, online recommendation systems used for movies or books, traffic prediction for GPS navigation systems, and storm prediction models for the weather.

One of the main forms of machine learning is reinforcement learning (Figure 1). Reinforcement learning is a continuous cycle of learning within a computational environment that occurs through positive and negative reinforcement. The system learns as a consequence of its past actions and is rewarded when it achieves successful outcomes. The system operates on maximizing cumulative

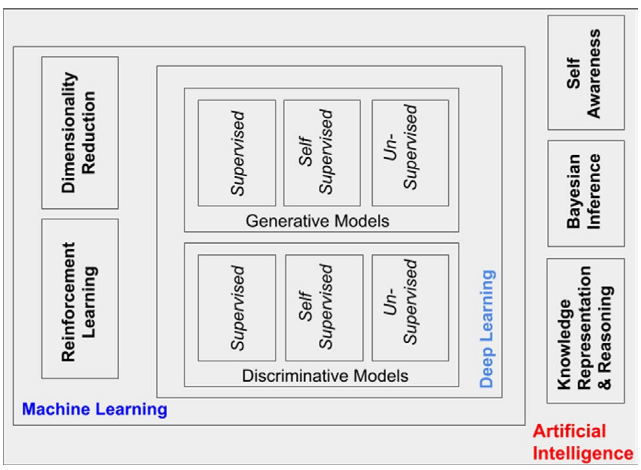


**FIGURE 1** Classification of Artificial Intelligence (AI) Systems. Overview of the main components of each discipline of AI. Machine learning is a sub-branch of AI with capabilities of robust prediction modeling that have many important applications in medicine and in the commercial sector. Deep learning is the newest generation of machine learning that uses artificial neural networks and its performance is determined by large datasets. Deep learning is particularly effective when the input data are unstructured data (eg, an image) and have important capabilities of image classification and in visual editing applications

rewards and makes new choices based on past positive and negative experiences. Dimensionality reduction is another form of machine learning and is the process of reducing the dimension of a large feature-based dataset. Dimensionality reduction is very good at correcting for over-fitting. Over-fitting occurs when a machine learning model becomes heavily reliant on the training dataset and demonstrates poor performance with new test datasets. Dimensionality reduction aims to improve accuracy and speed of models by reducing the amount of misleading and redundant data. A common approach to achieve dimensionality reduction is by utilizing a method referred to as feature selection. Feature selection selects for only the most relevant features in a large dataset and narrows the model to focus on this data.

## 1.3 | Deep learning

Deep learning is considered the most recent evolution of machine learning and is appropriately designated as a subdiscipline of machine learning. Deep learning can accomplish all the aforementioned applications of machine learning but in addition, has more complex functionality, can provide decision-making capabilities, and can handle extremely large datasets. For example, where machine learning can provide image detection for self-driving cars, deep learning can make decisions on which direction the car should go to avoid objects. Deep learning has also shown to be effective in language translation, speech recognition, image classification, and face recognition. Furthermore, a recent exponential rise in the development of deep learning algorithms has revolutionized the performance in several visual computing fundamental problems, for example, image classification and image segmentation (discriminative models). Deep learning has also shown heightened performance in visual editing applications, for example, synthesis and rendering (generative models).

Machine and deep learning techniques are data-driven, where they need supervision to learn from data. Such supervision constraints the learned model to our human knowledge. The degree of training of learning models varies from fully supervised learning to completely un-supervised learning (Figure 1). Self-supervised learning is where the data are autonomously labeled by the machine and no manual human labeling is required (no manual annotations). Weakly supervised learning exists on a spectrum between supervised and self-supervised learning where a limited amount of manual labeling of data is needed (some manual annotations). The ultimate goal of deep learning is a fully automated un-supervised system that can generate accurate predictions.

In medicine, the most prominent application of deep learning has been its use in the detection of abnormalities or disease in radiology and pathology.[3,4] Deep learning-based pattern-recognition applications have been effective in analyzing large medical images and can accomplish many tasks including object detection (identifying the location of the lesion of interest) and image classification (generating differential diagnoses).[3]

## 1.4 | Artificial neural networks

Artificial neural networks (ANNs) are computational models designed to mimic biological neural networks in a similar manner to the operation of a human brain. The most basic form of ANNs is traditional neural networks (ie, multilayer perceptron) that feature three main components or layers, namely an input layer, a hidden layer, and an output layer. The input is analogous to human dendrites, and the output is analogous to human axons. The input layer receives the initial data source (eg, image), and the output layer produces the final result. Layers are connected together, and each layer contains a set of artificial neurons that are programmed to execute a desired task. The number of neurons in each layer is determined by the size and type of training data. Hidden layers exist between input and output layers and are introduced into neural networks to solve more complex computational tasks. The network size and performance of ANNs is directly proportional to the volume of data the network has been trained on. Very large ANNs that have been trained on large quantities of data perform significantly better than networks that have encountered less data. However, it is important to note that training large ANNs with limited training data will learn a model that is over-fitted to the training data and will not generalize to the test data (ie, data not seen during training).

## 1.5 | Convolutional neural networks

Deep learning techniques are commonly driven by deep convolutional neural networks (CNNs).[5,6] CNNs are comprised of a large set of neurons, their connections, and parameters. During the training phase, the network observes positive and negative samples and the neurons learn to co-ordinate their operations in order to achieve the desired task. The optimum method neurons co-ordinate is very complicated as it involves millions of neurons and their parameters (connectivity). Thus, large computational power must be exploited through the use of Graphics Processing Units (GPUs) in order for deep learning to work more efficiently with extremely large datasets and to be significantly rapid in its operation. This level of computational power is usually not required for machine learning.

Convolutional neural networks differ from traditional neural networks in that CNNs incorporate many different types of layers. These layers include convolutional layers and pooling layers.[7] In essence, CNNs use convolutional layers to act as feature extractors and to identify important features from their input signals. Pooling layers function to minimize the network size and maintain the system steady at a computational threshold. CNNs are much better adapted at classifying large images than traditional neural networks. This is so because CNNs demonstrate heightened performance in terms of spatial features.

In summation, deep learning differs from machine learning in that deep learning can handle much larger datasets and has a neural-inspired architectural design.

## 1.6 | Recurrent neural networks

Recurrent neural networks (RNNs), another type of ANNs, function by taking in a standard input but also take inputs they have previously perceived in the past. RNNs work by containing hidden layers that have recurrent feedback neurons that store previous information and then merge that stored information with the current incoming inputs. Therefore, RNNs are advantageous models that exploit previously stored data to make informed decisions with new data. Whereas CNNs are effective at image classification, RNNs are particularly useful when working on sequences such as text, audio, or multiple sets of images.

## 1.7 | Transfer learning

A major advantage of machine and deep learning is that they can facilitate transfer learning. Transfer learning works by transferring the knowledge of an existing successful machine learning system or CNN to a newly created one. The process of using a pre-built and pre-trained classifier previously trained on a large dataset and transferring this knowledge to a new model with a smaller dataset significantly reduces training time and reduces the computational power requirement.[7] The most effective method to achieve transfer learning is fine-tuning. Fine-tuning is the process of freezing the early convolutional layers of a network and tasking them to extract generic low-level features that are found across all images, the last few layers can then be modified to focus on specific features that are relevant to the new dataset.

## 1.8 | Generative adversarial networks

Convolutional neural networks achieved revolutionizing success in fundamental computer vision and image processing problems, including image classification[5] and segmentation.[8]

The latest class of deep neural networks are known as Generative Adversarial Networks (GANs).[9] GANs are a special class of CNNs that consist of two main components: a generator and a discriminator. While the generator encodes and decodes images into and from low-dimensional latent space representation, the discriminator is trained to discern “real” versus “fake” images. Millions of “real” and “fake” data points are required to adequately train the discriminator, and hence, GANs heavily rely on Big Data. Big Data can be defined as extremely large volumes (in the order of petabytes) of different types of data, available from multiple various sources, which are generated at very high rates of acquisition.[10]

Generative Adversarial Networks are trained using the principle of game theory,[9,11] in which the generator learns to efficiently generate new data samples which can only be classified as real images by the discriminator. The use of GANs has paved the way for advanced solutions for a number of computer graphics and computer vision problems and is largely responsible for the recent success in revolutionizing the field of visual computing.[11,12] For instance, GANs have proven to generate pictures of extremely photo-realistic looking human faces.[12] Using modern technological advances and enhanced neural networks, this technology is rapid in providing its output, to the point that it can perform an accurate analysis in a fraction of a second.[6,13]

### 1.9 | Oncologic pathology and computer vision

By combining the two niche disciplines of oncologic pathology and computer vision, recent studies have demonstrated a unique approach to enhancing diagnostic accuracy through utilizing robust AI-driven solutions. The intent is to significantly reduce misdiagnosis and human error secondary to work-related fatigue and burnout. This has become even more apparent in this day and age of increasing work demands as a consequence of fewer specialists and an increase in administrative duties. For instance, automated approaches that facilitate the completion of elaborate staging checklists not only reduce the burden of these time-intensive tasks for the pathologist but also limit operator variability.

Initial studies which have incorporated AI and oncologic pathology are highly laudable as correct diagnosis in oncology are critical. The main advantages of integrating machine learning with automated microscopic image analysis are the ability to significantly reduce intra- and inter-observer variability and to improve objectivity and reproducibility.[1] By achieving such reproducibility and standardization among pathologists, this will likely improve research reporting and ultimately enhance our understanding of the natural history of diseases.

## 2 | DISCUSSION

### 2.1 | Big data and whole slide images

The success of AI in oncologic pathology depends on Big Data, and thousands of data points are required to generate reliable predictive models. This is so because the training process requires an innumerably large number of training examples.[14] Currently, the establishment of extremely large digital databases of pathological entities is the most significant challenge in the field. This is especially the case in oncologic pathology as several cancers are relatively rare, and therefore, databases of such large magnitudes are difficult to obtain without multicenter collaboration. However, with the advent of whole slide images (WSIs), in which billions of pixels are contained in a single digitally scanned specimen, there has been great opportunity to exploit such large volumes of data points and to design scalable algorithms to facilitate microscopic image analysis.[1] WSIs can contain gigapixel resolution, and therefore, in order to facilitate deep learning on such large image sizes, patch-based CNNs have been introduced which have been shown to effectively decompartmentalize WSIs.[15]

### 2.2 | Traditional computational pathology

Much of the initial studies have focused on breast, lung, and prostate cancer.[4,16] Computational pathology platforms have enabled rapid and sophisticated image analysis of digital microscopy, and many of these initial studies have relied on traditional computational pathology methods for mitosis counting.[17] This approach is important as the current process of manual mitosis counting by pathologists is time consuming, highly subjective, and does not allow for standardized reporting of mitotic scores across pathology laboratories. Dong et al[18] in 2014 designed a computational pathology method to identify and quantify nuclear features from diagnostic tumor regions of interest (ROIs) of intraductal proliferative lesions of the breast. This computational method was designed to extract 392 features of the mean and standard deviation in nuclear size and shape, intensity, and texture. This study demonstrated that computational pathology methods could achieve high accuracy for distinguishing between benign breast ductal hyperplasia and malignant ductal carcinoma in situ (DCIS).[18] Furthermore, this study showed that a high level of accuracy could be achieved in distinguishing between low-grade and high-grade DCIS (an area under the receiver operating characteristic curve [AUC] = 0.98).[18] A significant limitation of many of these studies that incorporated automated mitosis counting is that mitosis detection was done on the patch-level or on predetermined ROIs as the input, whereas to more appropriately mimic the real-world scenario, WSIs should be used as the input for automated mitosis detection.[17]

### 2.3 | CNNs in histopathology

The Tumor Proliferation Assessment Challenge 2016 (*TUPAC16*) was the first challenge to predict breast tumor proliferation scores based on automated mitosis detection from WSIs.[17] Bejnordi et al[19] 2017 utilized CNNs to diagnose breast cancer in 646 breast tissue samples. Their system achieved an AUC of 0.92 at the WSI level for distinguishing breast cancer from benign breast tissue.[19] A subsequent study by Bejnordi et al[20] utilized context-aware stacked CNNs to classify breast pathology from WSIs into three categories. Their system achieved an AUC of 0.96 for binary classification (nonmalignant and malignant) and a three-class accuracy of 81% for classification of WSIs into normal/benign, DCIS, or invasive ductal carcinoma.[20]

### 2.4 | RNNs in histopathology

Most recently, RNNs have shown excellent results in the classification of histopathological images. Yao et al[21] demonstrated superior results with a parallel system of CNN and RNN in the image classification of four classes of breast biopsy histopathology images (normal breast tissue, benign, carcinoma in situ, and invasive carcinoma) compared to CNN alone. Iizuka et al[22] used a combined CNN and RNN on WSIs of stomach and colon biopsies. Image classification

**TABLE 1** Overview of artificial intelligence methods in oral oncologic histopathology

| Study | Year | Method | Lesion type | Study outcome | Input | Results |
|---|---|---|---|---|---|---|
| Chang et al[33] | 2013 | 1. SVM<br>2. ANFIS<br>3. LR<br>4. ANN | OSCC | Comparison of different ML oral cancer prognostication models | 31 OSCC cases from an Oral Cancer Database<br>Data inputted:<br>Clinicopathologic data<br>Genomic data | Best ML accuracy achieved with ANFIS<br>ANFIS AUC: 0.90 |
| Lu et al[35] | 2017 | Computer-assisted histomorphometric ML classifier that was based on nuclear architecture (shape and texture) | OSCC | Stratification between high- and low-risk patients in terms of disease-specific survival | Tissue microarray of 115 OSCC | ML AUC: 0.72 |
| Folmsbee et al[41] | 2018 | CNN | OSCC | Comparison of active learning and random learning in the identification of seven tissue classes (stroma, lymphocytes, tumor, mucosa, keratin pearls, blood, and adipose) | 143 OSCC WSIs<br>Data inputted:<br>WSIs | Active learning performed 3% better than random learning |
| Karadaghy et al[26] | 2019 | 2-class decision forest architecture | OSCC | ML prediction of 5-year overall survival compared to TNM staging system model | 33 065 patients from the NCDB<br>Data inputted:<br>clinicopathologic data | ML AUC: 0.80<br>TNM staging system AUC: 0.68 |
| Bur et al[27] | 2019 | 1. LR<br>2. Decision forest<br>3. Kernel support vector machines<br>4. Gradient boosting ML architectures | OSCC | ML prediction model of pathologic nodal metastasis compared to methods based on tumor DOI | 654 patients from the NCDB and 71 patients treated at a single academic institution<br>Data inputted:<br>clinicopathologic data | ML AUC: 0.84<br>DOI model AUC: 0.66 |
| Alabi et al[28] | 2019 | MLP | OSCC | ML prediction of locoregional recurrences in early OSCC compared to LR model | 311 patients from multiple centers<br>Data inputted:<br>clinicopathologic data | ML accuracy: 93%<br>LR model accuracy: 87% |
| Shaban et al[36] | 2019 | CNN | OSCC | Quantification of tumor-infiltrating lymphocytes (TILs) and prediction of disease-free survival in OSCC patients | 60 OSCC cases from a single center<br>Data inputted:<br>clinicopathologic data | CNN accuracy: 96%<br>TILs are a strong prognostic indicator of disease-free survival |
| Halicek et al[42] | 2019 | CNN | HNSCC | Ability of a CNN to localize and detect SCC from WSIs | 381 WSIs from 156 patients with HNSCC<br>Data inputted:<br>WSIs | CNN AUC: 0.92 |
| Lalithmani & Punitha[43] | 2019 | 1. SVM<br>2. Bayes Net<br>3. Decision Stump<br>4. AdaBoostM1<br>5. KNN<br>6. DNAFS | OSCC | Comparison of different classifier systems in the diagnosis of OSCC | Data inputted:<br>Clinicopathologic data | Best accuracy achieved with DNAFS<br>DNAFS accuracy: 96% |

Abbreviations: ANFIS, Adaptive Neuro-Fuzzy Inference System; ANN, Artificial Neural Network; AUC, Area Under the receiver operating characteristic Curve; CNN, Convolutional Neural Network; DNAFS, Deep Neural Based Adaptive Fuzzy System; DOI, Depth of Invasion; HNSCC, Head & Neck Squamous Cell Carcinoma; KNN, K-Nearest Neighbors; LR, Logistic Regression; ML, Machine Learning; MLP, Multilayer Perceptron; NCDB, National Cancer Database; OSCC, Oral Squamous Cell Carcinoma; SVM, Support Vector Machine; TILs, Tumor-infiltrating Lymphocytes; TNM, Tumor, Node, Metastasis; WSIs, Whole-Slide Images.

into non-neoplastic, adenoma, and adenocarcinoma was achieved with an AUC of up to 0.99.

## 2.5 | Man vs machine

A common approach to validate deep learning algorithms in oncologic histopathology is to assess the performance of these algorithms as compared to an expert pathologist. This was highlighted in the Cancer Metastases in Lymph Nodes Challenge 2016 (*CAMELYON16*), which compared the diagnostic accuracy of an expert pathologist to algorithms designed to identify breast cancer from WSIs.[23] In the *CAMELYON16* challenge, 11 pathologists under time constraints were tasked with detecting breast cancer lymph node metastases by evaluating 129 WSIs (49 with breast cancer metastases and 80 without metastases) and were compared with deep learning algorithms which evaluated the same test set.[23] This validation approach has demonstrated that some deep learning algorithms provide superior (and more rapid) diagnostic performance capabilities than pathologists in the detection of breast cancer lymph node metastases.[23] More recently, researchers have focused on using deep CNNs from WSIs to assess the tumor-associated stroma.[24] In one such study, 928 WSIs from 330 patients were used in the testing phase and the developed algorithm accurately distinguished between benign breast proliferations and invasive breast cancer (AUC = 0.96).[24] Most recently, Wang et al[25] 2019 developed a weakly supervised deep learning system for the detection of lung cancer from WSIs. Their system demonstrated 97.3% accuracy.[25]

## 2.6 | AI in oral oncology

While deep learning has made significant advances and progressed the field of oncologic pathology, its use with respect to oral oncology is still in the nascent stage (Table 1). Machine and deep learning methods designed to enhance prognostication of oral cancer have been proposed with much of the work focused on prediction models on patient survival and locoregional recurrences in patients with oral squamous cell carcinomas (OSCC).[26-29] These models utilized AI-based methods to analyze large data registries of OSCC patients; however, the extent of data analyzed was limited to demographic, clinicopathologic, or genomic data.[26,30-32] Chang et al[33] 2013 developed an oral cancer prognostication model that incorporated clinicopathologic data (demographic information, details regarding clinical and pathological stage) and genomic data (*p53* and *p63* scores from immunohistochemistry (IHC) slides). Through machine learning methods, their model delivered an AUC of 0.90 for oral cancer prognosis based on invasion, *p63*, and alcohol consumption status. The use of optical detection systems to aid in the classification of OSCC has also been employed. A notable example is a recent study by Jeyaraj & Nadar that featured a CNN-driven automated computer-aided hyperspectral image detecting system to classify OSCC.[34] This system achieved an accuracy of 91% for image classification of OSCC vs. benign tissue and achieved an accuracy of 95% for classification of OSCC vs. normal tissue.[34]

Few studies have explored machine learning methods on OSCC digital histopathologic images. Lu et al[35] in 2017 devised a computer-assisted histomorphometric classifier that was based on nuclear architecture (shape and texture). Utilizing a digitized tissue microarray of 115 OSCC, a machine learning classifier analyzed 2 mm areas of OSCC and this classifier yielded an AUC of 0.72 in distinguishing between high- and low-risk patients in terms of disease-specific survival. Although a low specificity of 71% and sensitivity of 62% were achieved with this classifier, this exploratory study highlighted the potential of analyzing small areas of tissue. Shaban et al[36] 2019 trained a novel CNN system to quantify tumor-infiltrating lymphocytes (TILs) from WSIs of OSCC and achieved an accuracy of 96%. Moreover, they demonstrated that TILs are a strong prognostic indicator of disease-free survival.[36]

Although initial efforts of applying AI-based approaches have focused on OSCC, the use of these approaches on oral potentially malignant disorders (OPMDs) has significant potential to improve early detection of OSCC. Shamim et al[37] 2019 used an annotated dataset of clinical images of benign and potentially malignant tongue lesions to train deep CNNs for the purposes of image classification. In this study, their model achieved a mean classification accuracy of 0.98 in distinguishing between benign and potentially malignant tongue lesions.[37] Studies on WSIs of OPMDs are lacking but importantly, parallels may be drawn with AI-based approaches on clinicopathologic parameters of premalignant disorders in different sites such as the esophagus (Barrett's esophagus) and the cervix (cervical intraepithelial neoplasia).[38,39] Thus, future studies focused on OPMDs would greatly benefit the field especially in tackling the large intra- and inter-observer variability that occurs in oral dysplasia grading.

# 3 | CONCLUSION

The recent momentous drive to apply advanced AI technologies to diagnostic medicine can potentially revolutionize care as we know it. AI heralds exciting opportunities to streamline health care and signifies virtually endless possibilities for personalizing cancer care.

Artificial intelligence has achieved excellent and sometimes superior results as compared to human pathologists across many different cancer types, and it has become apparent that initial studies performed on OSCC digital histopathologic images show promise, especially when clinical and genomic data are included in the predictive models. Moreover, the combined power of an expert pathologist and AI system has demonstrated reduced diagnostic errors and superior results compared to either pathologist or machine alone.[40] It has been predicted that researchers will endeavor to use AI to more accurately quantitatively grade immunohistochemistry stains and to reduce the time pathologists spend screening (eg, in cytopathology) by identifying ROIs.[14] This would free up pathologists' time and allow them to focus on higher-level diagnostic tasks such as integrating molecular data which can aid clinicians in treatment decision-making.[14] Further, the use of AI will become important in assisting pathologists in the diagnosis of poorly differentiated tumors and in the diagnosis of exceptionally rare cases.

Given the early promising results in the field of oral oncology, we now have been provided with new opportunities to incorporate deep learning methods to progress the field. Thus, it would be strategic to collaborate with computer scientists with the goal of addressing some of the major challenges in our field such as significantly reducing intra- and inter-observer variability in oral dysplasia grading and improving diagnostic criteria for elusive conditions such as proliferative *verrucous* leukoplakia (PVL).

In conclusion, the overarching goal of AI is not to replace pathologists. By working synergistically with virtual AI pathology assistants, it is hoped that AI will provide pathologists with recommendations and rapid automated assistance that will introduce a new wave of heightened precision in oncologic pathology and oral oncology.

## CONFLICT OF INTEREST

None declared.

## AUTHOR CONTRIBUTION

**Ahmed Sultan:** Conceptualization; Data curation; Methodology; Writing-original draft; Writing-review & editing. **Mohamed Elgharib:** Conceptualization; Data curation; Methodology; Software; Visualization; Writing-original draft; Writing-review & editing. **Tiffany Tavares:** Conceptualization; Methodology; Writing-original draft; Writing-review & editing. **Maryam Jessri:** Conceptualization; Methodology; Writing-original draft; Writing-review & editing. **John Basile:** Conceptualization; Methodology; Writing-original draft; Writing-review & editing.

## ORCID

*Ahmed S. Sultan* https://orcid.org/0000-0001-5286-4562

## REFERENCES

1. Xing F, Yang L. Machine learning and its application in microscopic image analysis. In: Wu G, Shen D, & Sabuncu MR ed. *Machine Learning and Medical Imaging*. Amsterdam, Netherlands: Elsevier Inc.; 2016:97-127.
2. Badrick T, Banfi G, Bietenbeck A, Cervinski MA, Loh TP, Sikaris K. Machine learning for clinical chemists. *Clin Chem*. 2019;65(11):1350-1356.
3. Kim M, Yun J, Cho Y, et al. Deep learning in medical imaging. *Neurospine*. 2019;16(4):657-668.
4. Chang HY, Jung CK, Woo JI, et al. Artificial intelligence in pathology. *J Pathol Transl Med*. 2019;53(1):1-12.
5. Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks. In: Pereira F, Burges CJC, Bottou L, & Weinberger KQ, ed. *Advances in Neural Processing Information Systems (NIPS)*. New York, NY: Curran Associates, Inc.; 2012:1097-1105.
6. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: International Conference on Learning Representations, 2015.
7. Rawat W, Wang Z. Deep convolutional neural networks for image classification: a comprehensive review. *Neural Comput*. 2017;29(9):2352-2449.
8. Lin T, Maire M, Belongie S, et al. Microsoft coco: common objects in context. In: *Computer Vision-ECCV*. Cham, Switzerland: Springer International Publishing; 2014:740-755.
9. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. In: *Advances in Neural Information Processing Systems (NIPS)*. New York, NY: Curran Associates, Inc.; 2014;2672-2680.
10. Favaretto M, De Clercq E, Schneble CO, Elger BS. What is your definition of Big Data? Researchers' understanding of the phenomenon of the decade. *PLoS One*. 2020;15(2):e0228987.
11. Isola P, Zhu J, Zhou T, Efros A. Image-to-image translation with conditional adversarial networks. In Computer Vision and Pattern Recognition (CVPR). 2017:5967-5976.
12. Karras T, Aila T, Laine S, Lehtinen J. Progressive growing of GANs for improved quality, stability, and variation. In: International Conference on Learning Representations. 2018.
13. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In Computer Vision and Pattern Recognition (CVPR). 2016:770-778.
14. Granter SR, Beck AH, Papke DJ. AlphaGo, deep learning, and the future of the human microscopist. *Arch Pathol Lab Med*. 2017;141(5):619-621.
15. Hou L, Samaras D, Kurc TM, Gao Y, Davis JE, Saltz JH. Patch-based convolutional neural network for whole slide tissue image classification. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2016;2016:2424-2433.
16. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60-88.
17. Veta M, Heng YJ, Stathonikos N, et al. Predicting breast tumor proliferation from whole-slide images: the TUPAC16 challenge. *Med Image Anal*. 2019;54:111-121.
18. Dong F, Irshad H, Oh EY, et al. Computational pathology to discriminate benign from malignant intraductal proliferations of the breast. *PLoS One*. 2014;9(12):e114885.
19. Bejnordi BE, Lin J, Glass B, et al. Learning-based assessment of tumor-associated stroma for diagnosing breast cancer in histopathology images. *Proc IEEE Int Symp Biomed Imaging*. 2017;2017:929-932.
20. Bejnordi BE, Zuidhof G, Balkenhol M, et al. Context-aware stacked convolutional neural networks for classification of breast carcinomas in whole-slide histopathology images. *J Med Imaging (Bellingham)*. 2017;4(4):44504.
21. Yao H, Zhang X, Zhou X, Liu S. Parallel structure deep neural network using CNN and RNN with an attention mechanism for breast cancer histology image classification. *Cancers (Basel)*. 2019;11(12):1901.
22. Iizuka O, Kanavati F, Kato K, et al. Deep learning models for histopathological classification of gastric and colonic epithelial tumours. *Sci Rep*. 2020;10:1504.
23. Ehteshami Bejnordi B, Veta M, Johannes van Diest P, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA*. 2017;318(22):2199-2210.
24. Ehteshami Bejnordi B, Mullooly M, Pfeiffer RM, et al. Using deep convolutional neural networks to identify and classify tumor-associated stroma in diagnostic breast biopsies. *Mod Pathol*. 2018;31(10):1502-1512.

25. Wang X, Chen H, Gan C, et al. Weakly supervised deep learning for whole slide lung cancer image analysis. *IEEE Trans Cybern*. 2019;1-13. https://doi.org/10.1109/TCYB.2019.2935141
26. Karadaghy OA, Shew M, New J, Bur AM. Development and assessment of a machine learning model to help predict survival among patients with oral squamous cell carcinoma. *JAMA Otolaryngol Head Neck Surg*. 2019;145(12):1115.
27. Bur AM, Holcomb A, Goodwin S, et al. Machine learning to predict occult nodal metastasis in early oral squamous cell carcinoma. *Oral Oncol*. 2019;92:20-25.
28. Alabi RO, Elmusrati M, Sawazaki-Calone I, et al. Machine learning application for prediction of locoregional recurrences in early oral tongue cancer: a Web-based prognostic tool. *Virchows Arch*. 2019;475(4):489-497.
29. Arora A, Husain N, Bansal A, et al. Development of a new outcome prediction model in early-stage squamous cell carcinoma of the oral cavity based on histopathologic parameters with multivariate analysis: The Aditi-Nuzhat Lymph-node Prediction Score (ANLPS) System. *Am J Surg Pathol*. 2017;41(7):950-960.
30. Patil S, Habib Awan K, Arakeri G, et al. Machine learning and its potential applications to the genomic study of head and neck cancer-A systematic review. *J Oral Pathol Med*. 2019;48(9):773-779.
31. Li S, Chen X, Liu X, et al. Complex integrated analysis of lncRNAs-miRNAs-mRNAs in oral squamous cell carcinoma. *Oral Oncol*. 2017;73:1-9.
32. Schmidt S, Linge A, Zwanenburg A, et al. Development and validation of a gene signature for patients with head and neck carcinomas treated by postoperative radio(chemo)therapy. *Clin Cancer Res*. 2018;24(6):1364-1374.
33. Chang SW, Abdul-Kareem S, Merican AF, Zain RB. Oral cancer prognosis based on clinicopathologic and genomic markers using a hybrid of feature selection and machine learning methods. *BMC Bioinformatics*. 2013;14:170.
34. 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.
35. Lu C, Lewis JS, Dupont WD, Plummer WD, Janowczyk A, Madabhushi A. An oral cavity squamous cell carcinoma quantitative histomorphometric-based image classifier of nuclear morphology can risk stratify patients for disease-specific survival. *Mod Pathol*. 2017;30(12):1655-1665.
36. Shaban M, Khurram SA, Fraz MM, et al. A novel digital score for abundance of tumour infiltrating lymphocytes predicts disease free survival in oral squamous cell carcinoma. *Sci Rep*. 2019;9(1):13341.
37. Shamim M, Syed S, Shiblee M, Usman M, Ali S. Automated detection of oral pre-cancerous tongue lesions using deep learning for early diagnosis of oral cavity cancer. arXiv:190908987 [eessIV]. 2019.
38. Ghatwary N, Ahmed A, Grisan E, Jalab H, Bidaut L, Ye X. Barrett's esophagus digital pathology stage classification through feature enhancement of confocal laser endomicroscopy. *J Med Imaging (Bellingham)*. 2019;6(1):14502.
39. Xu T, Zhang H, Huang X, et al. Multimodal Deep learning for cervical dysplasia diagnosis. *Med Image Comput Comput Assist Interv*. 2016;9901:79-87.
40. Turner OC, Aeffner F, Bangari DS, et al. Society of toxicologic pathology digital pathology and image analysis special interest group article*: opinion on the application of artificial intelligence and machine learning to digital toxicologic pathology. *Toxicol Pathol*. 2020;48(2):277-294.
41. Folmsbee J, Liu X, Brandwein-Weber M, Doyle S. Active deep learning: Improved training efficiency of convolutional neural networks for tissue classification in oral cavity cancer. *IEEE 15th International Symposium on Biomedical Imaging*. 2018;770-773. https://doi.org/10.1109/ISBI.2018.8363686
42. Halicek M, Shahedi M, Little JV, et al. Head and neck cancer detection in digitized whole-slide histology using convolutional neural networks. *Sci Rep*. 2019;9(1):14043.
43. Lalithmani K, Punitha A. Detection of oral cancer using deep neural based adaptive fuzzy system in data mining techniques. *Int J Rec Tech Eng*. 2019;397-404.

**How to cite this article:** Sultan AS, Elgharib MA, Tavares T, Jessri M, Basile JR. The use of artificial intelligence, machine learning and deep learning in oncologic histopathology. *J Oral Pathol Med*. 2020;49:849–856. https://doi.org/10.1111/jop.13042