Contemporary Orofacial Sciences (2023) 1(1):13-16
DOI: 10.30486/cofs.2023.1982110.1002

ORIGINAL RESEARCH

Evaluation of the Use of Artificial Intelligence (AI) for Low & High-Grade OSCC Diagnosis from Normal Mucosa

Mehran Arshadi Fard [1], Nader Kalbasi [2]

Received: 2023-03-11 / Accepted: 2023-05-10 / First publication date: 2023-05-10
© The Author(s) 2023

Abstract

Background: Artificial Intelligence (AI) is a term that implies using a computer to model intelligent behavior with minimal human intervention. Their potential to exploit the meaningful relationship with a dataset can be used in diagnosis, treatment, and outcome prediction in many clinical scenarios. The aim was to evaluate the use of artificial intelligence algorithms to successfully differentiate histopathologic images of grades of OSCC and healthy oral mucosa.

Material & Methods: In this cross-sectional study, Inception-ResNet-V2, a recently created artificial intelligence system, was used to analyze 844 pictures captured from the histopathological view of the connective tissues from three groups, low-grade OSCC, high-grade OSCC and normal mucosa.

Result: The results obtained from this research and comparable articles emphasize that deep learning-based systems have a high ability to analyze histopathological images and can be very useful and effective in cancer diagnosis and grading. According to the results of the ROC analysis from this research, Inception-ResNet-V2 has shown robust results in successfully differentiating Low-Grade OSCC, High-Grade OSCC and normal mucosa with over 95% accuracy.

Conclusion: According to the results of the present and previous studies, it can be concluded that CNN, and particularly Inception-ResNet-V2 have immense potential in analyzing histopathology pictures and could be very helpful for pathologists in cancer diagnosis.

Key Words: Artificial Intelligence, Diagnosis, Oral Squamous Cell Carcinoma, Inception-Resnet-V2

Introduction

Artificial intelligence (AI) could be well-defined as the result of combining engineering and science, which includes an understanding of intelligent behavior and creating systems that display these behaviors. In other words, a set of programs that enable computers to illustrate intelligent behavior. Alan Turing, a British mathematician, is believed to be one of the founders of modern computer science and artificial intelligence. He described artificial intelligence as; the ability to achieve human-like performance in conscious activities (1). In medical grounds, Artificial intelligence was first investigated in 1976 by Gunn for diagnosing acute abdominal pain. In the last two decades, we have seen a dramatic increase of interest in using artificial intelligence in medicine in diagnosis and prognosis, drug prescription, examination and analysis technologies and even surgical interventions. (2, 3).

Oral squamous cell carcinoma is the most common oral cancer, and although localized in a region that is accessible and has the potential to be detected early, this usually does not occur. The golden standard procedure for the diagnosis of oral cancer relies on histopathological examination by an oral pathologist, however, this process does have a subjective component of the examination that could directly impact the diagnosis. For this reason, artificial intelligence (AI) algorithms are widely used today as an aid in the diagnosis for the classification of tumours, to balance the subjective variability (4).

Rahman et al. used traditional machine learning techniques to achieve a 99.78% accuracy by applying a decision tree in classifying OSCC using morphological and textural features of 452 hand-cropped cell nuclei (5). These traditional machine learning techniques rely on predefined features that describe regular patterns in the data extracted from regions of interest (ROIs) with the parameters defined based on expert knowledge (6).

Corresponding Author: Nader Kalbasi
Email: dr.naderkalbasi@yahoo.com

[1] School of Dentistry, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran.
[2] Department of Oral and Maxillofacial Pathology, School of Dentistry, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran.

In 2006, Hinton et al. published an article in the journal "Science" that presented an artificial neural network (ANN) with multiple hidden layers and excellent feature learning ability, leading to further studies related to the topic of deep learning (7). ANN is the main algorithm for guiding deep learning, and CNN is the most common type of ANN for deep learning. In fact, in the early eighties and nineties, CNN performed excellent results in various areas of pattern recognition, especially handwriting recognition. However, these systems were only suitable for small image recognition. Since the developed CNN achieved the best classification performance in the ImageNet Large-Scale Image Recognition Challenge (LSVRC) in 2012, more researchers have paid attention to it, and these systems achieved advances daily (8).

As a subset of artificial intelligence, deep learning relies on a neural network structure inspired by the human brain. In terms of feature selection and extraction, deep learning algorithms do not need predefined features and they don't necessarily need to place complex shaped ROIs on the images. They can directly learn features by exploring the data space and performing image classification and processing tasks. This data-driven mode makes it more functional and even faster. Today, convolutional neural networks (CNN) are the most popular type of deep learning system on the bases of medical image analysis (9).

In January 2020, Tschandl et al. (10) conducted a study titled "Human-Computer Collaboration for Skin Cancer Recognition", A resnet34 was used in differentiating seven diagnostic categories, which concluded that a high-quality artificial intelligence-based support alongside clinical decision-making improved diagnostic accuracy compared to AI or pathologists alone, also less experienced pathologists benefit the most from this AI-based supported system.

In another research conducted in 2020 by navarun et al. titled automatic classification of epithelial tissue cells in oral squamous cell carcinoma using artificial intelligence pre-trained systems such as AlexNet, Resnet, and VGG 16&19 were used, which the highest accuracy reported being the Resnet system, which recorded an accuracy of 92.15% (4). In 2021, Musilin et al. (11) conducted research based on artificial intelligence to diagnose oral squamous cell carcinoma. They reached diagnostic accuracy of 96.3% using the artificial intelligence system deepLabv3+ and Xception. One of the characteristics of our research is to examine the different grades of oral squamous cell carcinoma, using 322 histopathologic images as a dataset. Images were captured in a 10x magnification zoom.

Materials & Methods

In this cross-sectional study (ethic code: IR.IAU.KHUISF.REC.1401.058), paraffin blocks (n=16) of high-grade and low-grade OSCC were gathered from the archives of the pathology department of Al-Zahra Hospital and Khorasgan University. For approval of previous diagnoses, the pathologist re-evaluated the slides of each block. Two high-grade and three low-grade OSCCs were excluded from the study. The samples of normal oral mucosa were obtained from crown lengthening surgeries at Khuisf University. The collected tissues were immediately placed in 10% formalin and delivered to the laboratory, where each sample was placed in a separate paraffin block. From paraffin blocks, a 3–4-micron section was prepared and placed overnight on a special glass slide in Poly-L-lysine. The data was acquired from the photographs taken of the histopathologic feature of the lesions, using the Nikon digital sight ds-l2 digital photography device and Nikon eclipse e200 microscope in two magnifications of 10x and 40x without overlap. Photographs without appropriate accuracy were excluded.

Finally, 844 images comprising 437 and 407 images with magnifications of 10x and 40x, respectively, were obtained. These images were divided into three classes Normal oral tissue, Low-grade and High-grade OSCC, as shown in table 1.

Table 1. Details of the whole dataset

	Normal Tissue	Low-Grade	High-Grade	Total
10X Zoom	88	174	175	437
40X Zoom	126	141	140	407
Number of patients	7	8	3	18
Total	214	315	315	844

600 images were chosen randomly from the entire images to form a study group, and the remaining 244 images were the test group. The details of both groups

Table 2. Details of the training set

	Normal Tissue	Low-Grade	High-Grade	Total
10X Zoom	67	127	126	320
40X Zoom	87	105	88	280
Total	154	232	214	600

Table 3. Details of the test dataset

	Normal Tissue	Low-Grade	High-Grade	Total
10X Zoom	21	47	49	117
40X Zoom	39	36	52	127
Total	60	83	101	244

are shown in Tables 2 and 3.

The backbone of the model was the Inception-ResNet-V2 (12) trained on the ImageNet dataset and created in 2017 by Szegedy et al. by combining two previous efficient artificial intelligence systems, Inception (GoogLeNet) and Reset. Which reportedly provided shrewd results during the ImageNet Large-Scale Image Recognition Challenge (LSVRC) in 2014 and 2015, respectively. This model has shown robust results on image recognition tasks (13). The general

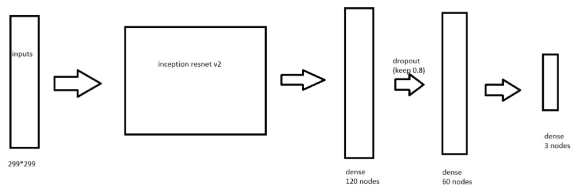

Figure1. General architecture of the Inception-ResNet-v2 based model.

architecture of our model is shown in Figure 1.

Considering the size of our dataset, it was important to avoid overfitting the model. That is when a model “memorizes” the images instead of learning the pattern to “recognize” them. Different techniques were used to avoid this phenomenon. At first, an excellent choice of a pre-trained base model improves the pattern recognition results. The second step was positioning dropout layers in the architecture. Finally, a method of randomized augmentation was implemented to increase the dataset size virtually to a large number. This method is discussed in the training process section.

Training process: The images were labelled into three groups of 0, 1 and 2 for Normal Tissue, Low-grade and High-grade respectively. For the augmentation step, every image was randomly cropped into five 920 *920 sections from the original 1280 * 960 pixels. A random rotation chosen from 0, 90, 180 and 270 degrees was applied to four of these new images. Then all of the images are resized to 299 * 299 pixels and their pixel values were normalized to be between -1 and +1 for the model input. This randomized step was done from scratch for every image on every epoch. That mathematically produces (1280-920) width * (960-920) height * 4 rotations * 5 augments = 288,000 different images from every image and reduces the chance of overfitting significantly as the model may not see the exact same image twice. In the testing process, the final predicted label, for every image is chosen as the most-voted-for label from the 5 augmented images. If any two classes had the same number of votes, for example, 2 votes for class 0 and two votes for class 1, then the class with the lower number was selected which was 0. The prevalence of this case was exceedingly small in our results. In the training process, the layers of the pre-trained section in the model were frozen and the rest of the layers were trained with Adam optimizer with a learning rate of 0.0001and the loss was calculated with categorical cross-entropy function. Training was stopped after 110 epochs. All of the training processes were done on a pc with a Nvidia 1660 Super with 8GB of graphical memory, 11600H Intel CPU, and 16 GB of RAM on Ubuntu 20.04 operation system.

Results

Run one:

combined: 233 true, 11 false, acc = 95.49, false positive = 1, false negative = 1

zoom 10: 111 true, 6 false, acc = 94.87, false positive = 0, false negative = 1

zoom 40: 122 true, 5 false, acc = 96.06, false positive = 1, false negative = 0

Table 4. Normal vs OSCC stats crosstabulation

			act		Total
			normal mucosa	oscc	
pred	normal mucosa	Count	59	1	60
		% within act	98.3%	0.5%	24.6%
	oscc	Count	1	183	184
		% within act	1.7%	99.5%	75.4%
Total		Count	60	184	244
		% within act	100.0%	100.0%	100.0%

Table 5. Normal vs OSCC Area Under the Curve

Test Result Variable(s): pred				
Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
0.989	0.010	0.000	0.969	1.000

The test result variable(s): pred has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

a. Under the nonparametric assumption
b. Null hypothesis: true area = 0.5

Table 6. Detailed Stats Crosstabulation

			act			Total
			Normal mucosa	Low-Grade	High-Grade	
pred	Normal mucosa	Count	59	1	0	60
		% within act	98.3%	1.2%	0.0%	24.6%
	Low-Grade	Count	1	77	4	82
		% within act	1.7%	92.8%	4.0%	33.6%
	High-Grade	Count	0	5	97	102
		% within act	0.0%	6.0%	96.0%	41.8%
Total		Count	60	83	101	244
		% within act	100.0%	100.0%	100.0%	100.0%

Table 7. Detailed Stats Area Under the Curve

Test Result Variable(s): pred				
Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
.948	.018	.000	.912	.984

The test result variable(s): pred has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

a. Under the nonparametric assumption
b. Null hypothesis: true area = 0.5

normal vs cancer: 242 true, 2 false, acc = 99.18

Discussion

The aim was to evaluate the use of Artificial Intelligence (AI) for Low & High-Grade OSCC Diagnosis from Normal Mucosa. An inception-ResNet-V2-based AI was used to divide 244 histopathology images of Low-Grade OSCC, High-grade OSCC and Normal Mucosa. Our findings were quite similar to other Recent research.

Navarun et al .used pre-trained systems to grade classes of OSCC, which reported an accuracy of 92.15% related to the ResNet system (8).

In 2021, Musilin et al. (10) used artificial intelligence and 322 histopathology images (10X zoom) of oral squamous cell carcinoma as a dataset. These researchers reached a diagnostic accuracy of 96.3% using the artificial intelligence system deepLabv3+ and Xception. Rahman et al. (11) researched traditional artificial intelligence systems to identify 450 photos of cell nuclei from forty different biopsy slides. With the morphological features of the cells, they reached 99.78% accuracy from combined techniques with five various factors. Cropping the nuclei of the relevant cells was done manually. Also, the biopsied slides were divided, simply into two healthy and malignant groups. The present study used an inception-ResNet-V2, a recently produced CNN system, to diagnose 244 images with a dataset with more than six hundred histopathology images captured with the X10 and X40 zoom. An overall accuracy of 95.49% was observed in the present study.

The main limitation of this study was the small number of samples of the high-grade OSCC. We suggest further studies involving intermediate-grade OSCC and other types of oral cancers.

Conclusion

In this research, the Obtained results revealed that CNN and in particular, Inception-ResNet-V2 have great potential in analyzing histopathology pictures and could be very helpful for pathologists in cancer diagnosis.

Conflicts of interest: none

References

1. Turing AM. Computing machinery and intelligence. Mind 1950; 59: 433–60.
2. Ramesh AN, Kambhampati C, Monson JR, Drew PJ. Artificial intelligence in medicine. Ann R Coll Surg Engl. 2004;86:334-8.
3. Bohr A, Memarzadeh K. The rise of artificial intelligence in healthcare applications. Artificial Intelligence in Healthcare. 2020:25-60.
4. Das N, Hussain E, Mahanta LB. Automated classification of cells into multiple classes in epithelial tissue of oral squamous cell carcinoma using transfer learning and convolutional neural network. Neural Netw. 2020; 128:47-60.
5. Rahman TY, Mahanta LB, Choudhury H, Das AK, Sarma JD. Study of morphological and textural features for classification of oral squamous cell carcinoma by traditional machine learning techniques. Cancer Rep (Hoboken). 2020:3: e1293
6. Zhou LQ, Wang JY, Yu SY, Wu GG, Wei Q, Deng YB, et al. Artificial intelligence in medical imaging of the liver. World J Gastroenterol. 2019; 25:672-682.
7. Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science. 2006; 313:504-7.
8. Russakovsky O, Deng J, Su H , Krause J , Satheesh S , Ma S , et al. ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision 2015; 115: 211-252
9. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017; 42:60-88.
10. Tschandl P, Rinner C, Apalla Z, Argenziano G, Codella N, Halpern A, et al. Human–computer collaboration for skin cancer recognition. Nat Med. 2020;26:1229-1234
11. Musulin J, Štifanić D, Zulijani A, Ćabov T, Dekanić A, Car Z. An Enhanced Histopathology Analysis: An AI-Based System for Multiclass Grading of Oral Squamous Cell Carcinoma and Segmenting of Epithelial and Stromal Tissue. Cancers (Basel). 2021;13:1784
12. Szegedy CH, Sergey I, Vanhoucke V, Alemi AA. Inception-v4, Inception-ResNet and the impact of residual connections on learning. Proceedings of the AAAI Conference on Artificial Intelligence 2016;31:
13. Zhang H, Liu C, Zhang Z, Xing Y, Liu X, Dong R, et al. Recurrence Plot-Based Approach for Cardiac Arrhythmia Classification Using Inception-ResNetv2. Front Physiol. 2021; 12:648950.