Received: 23 October 2020 | Revised: 30 January 2021 | Accepted: 6 February 2021

DOI: 10.1111/odi.13825

ORIGINAL ARTICLE

ORAL DISEASES Leading in Oral, Maxillofacial, Head & Neck Medicine

WILEY

A novel lightweight deep convolutional neural network for early detection of oral cancer

Fahed Jubair[1] | Omar Al-karadsheh[2] | Dimitrios Malamos[3] | Samara Al Mahdi[2] | Yusser Saad[2] | Yazan Hassona[2]

[1]Computer Engineering Department, School of Engineering, The University of Jordan, Amman, Jordan

[2]Department of Oral and Maxillofacial Surgery, Oral Medicine, and Periodontics, School of Dentistry, The University of Jordan, Amman, Jordan

[3]Oral Medicine Clinic, 1st Regional Health District of Attica, National Organization for the Provision of Health Services, Athens, Greece

Correspondence
Yazan Hassona, Oral Medicine and Special Care Dentistry, The University of Jordan, Queen Rania Street, Amman 11942, Jordan.
Email: yazan@ju.edu.jo

Abstract

Objectives: To develop a lightweight deep convolutional neural network (CNN) for binary classification of oral lesions into benign and malignant or potentially malignant using standard real-time clinical images.

Methods: A small deep CNN, that uses a pretrained EfficientNet-B0 as a lightweight transfer learning model, was proposed. A data set of 716 clinical images was used to train and test the proposed model. Accuracy, specificity, sensitivity, receiver operating characteristics (ROC) and area under curve (AUC) were used to evaluate performance. Bootstrapping with 120 repetitions was used to calculate arithmetic means and 95% confidence intervals (CIs).

Results: The proposed CNN model achieved an accuracy of 85.0% (95% CI: 81.0%–90.0%), a specificity of 84.5% (95% CI: 78.9%–91.5%), a sensitivity of 86.7% (95% CI: 80.4%–93.3%) and an AUC of 0.928 (95% CI: 0.88–0.96).

Conclusions: Deep CNNs can be an effective method to build low-budget embedded vision devices with limited computation power and memory capacity for diagnosis of oral cancer. Artificial intelligence (AI) can improve the quality and reach of oral cancer screening and early detection.

KEYWORDS

artificial intelligence, computer-aided diagnosis, convolutional neural network, deep learning, early detection, oral cancer, tongue cancer

1 | INTRODUCTION

Oral cancer is a global health issue (Siegel et al., 2019). It has been estimated that oral cancer is responsible for more than half a million incident cases and more than 200,000 deaths in 2012 (Chaturvedi et al., 2013). The incidence of oral cancer is expected to rise to 856,000 cases by the year 2035 (Shield et al., 2017).

The term oral cancer is rather a general term that includes cancers of the tongue, gingiva (maxilla and mandible), palate, buccal mucosa, floor of the mouth and lips. The majority of cancers arising from these sub-sites are squamous cell carcinoma (SCC) and are preceded by noticeable mucosal lesions known as oral potentially malignant disorders (OPMDs; Ojeda et al., 2020; Rivera, 2015; Warnakulasuriya et al., 2007).

Early detection of oral SCC can reduce cancer-specific mortality and morbidity (Schutte et al., 2020; Thomas et al., 2020). However, despite significant improvement in the understanding of molecular mechanisms underlying the pathogenesis of oral SCC and the malignant transformation of OPMD, the majority of oral cancers remain diagnosed at late stage (Califano et al., 1996; Karunakaran & Muniyan, 2020; Thomas et al., 2020). Unfortunately, it is unclear why oral cancer is often diagnosed late, but possible reasons include the lack of public and professional awareness, unfamiliarity of healthcare professionals with the variable and non-specific clinical presentation of oral SCC, and the lack of reliable and practical adjuvant diagnostic techniques (Hassona et al., ,2016, 2018; Madhura et al., 2020; Zhang et al., 2019).

It has been recommended that opportunistic screening, especially for high risk groups, can help diagnose oral SCC at earlier stages (AAOM, 2016; Rajaraman et al., 2015; Speight et al., 2017). However, the conventional visual and tactile oral examination is limited in its ability to differentiate between benign lesions and oral SCC or OPMDs. Adjuvant diagnostic techniques for early detection of oral SCC and OPMDs have been suggested and include the use of optical aids, vital staining, cytopathologic platforms, omics technology, synthetic biology, lab-on-chip, microfluidics, nanodiagnostics and salivary diagnostics (Huber, 2018; Madhura et al., 2020). Most of these techniques have shown proof for concept, but the available evidence to support their use is limited (Lingen et al., 2017).

Artificial intelligence (AI) is defined as the use of machines (i.e., computers) to mimic cognitive functions that humans associate with the human mind such as learning and problem solving. The use of AI is gaining an increasing attention in various field of medicine. Recently, the use of AI for cancer diagnosis and classification has witnessed significant improvement, and the evolution of deep convolutional neural networks (CNNs) led to the emergence of intelligent computer-aided diagnosis systems (Muthukrishnan et al., 2020). Esteva et al. (2017) demonstrated, using nearly 130,000 clinical images, that CNNs are capable of classifying skin cancers with a level of competence comparable to expert dermatologists (Esteva et al., 2017). Promising results regarding cancer diagnosis were also reported in other types of cancer including lung, breast, brain and colon (Attardo et al., 2020; Cho et al., 2020; Niu et al., 2020; Sathyakumar et al., 2020).

The use of AI methods for image analysis can potentially provide a practical, low-cost and universal approach for early detection of oral cancer (Ilhan et al., 2020; Kar et al., 2020; Thankappan et al., 2020; Waal, 2018). Only few studies applied AI methods for early detection of oral cancer using clinical images. Most of the previous studies were based on small date sets and applied traditional machine learning models that relied on image preprocessing (Chodorowski et al., 2008; Chodorowski et al., 1999; Thomas et al., 2013). The purpose of the present study is to develop a lightweight deep CNN for discrimination between benign and malignant or potentially malignant oral lesions using a data set of verified clinical images, and the use of EfficientNet-B0 transfer model.

2 | MATERIALS AND METHODS

2.1 | Data set collection and preparation

A total of 716 clinical images for various tongue lesions were used for the purpose of the present study. The data set was divided into two categories: (1) “suspicious” lesions which was made of 236 (33%) clinical images of either SCC or oral epithelial dysplasia; and (2) benign lesions which was made of 480 (67%) clinical images of various benign tongue conditions (Figure 1). Definitive diagnoses of the included images were confirmed by reviewing patients' record and pathology reports. The clinical images were collected by authors from a total of 543 patients over a period of 4 years using various types of digital cameras and smartphones. As a result, the images have variable sizes, zoom, angle, light conditions and picture orientations. All images were manually cropped, resized to 224 × 224 × 3 and converted into JPG file format.

The data set was randomly split, using random sampling in Python, into three categories: (1) training set (79%; n = 566 image); (2) validation set (7%; n = 50 image); and (3) test set (14%; n = 100 image). To ensure that there is no overlap between clinical images between the three categories (i.e. data leakage), all images were independently assessed for redundancy by two authors (Y.H and F.J) (Table 1).

2.2 | Convolutional neural network development

EfficientNet-B0 transfer model was used for the purpose of the present study. EfficientNet-B0 is a recently developed lightweight

FIGURE 1 Example of clinical images used in the study

TABLE 1 Data set distribution into train, validation and test sets

Set	Benign lesions	Malignant or potentially malignant lesions	Total
Train	405 images from 317 patients	161 images from 76 patients	566 images from 393 patients
Validation	25 images from 25 patients	25 images from 25 patients	50 images from 50 patients
Test	50 images from 50 patients	50 images from 50 patients	100 images from 100 patients

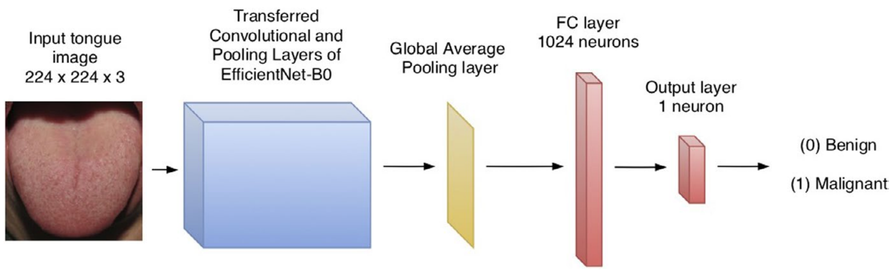

FIGURE 2 The proposed convolutional neural network for binary classification between benign and malignant/potentially malignant oral lesions

TABLE 2 List of transfer learning models investigated in this study

Transfer learning model	Non-trainable parameters	Trainable parameters	Total parameters
EfficientNet-B0	2,706,803	2,655,537	5,362,340
VGG19	20,024,384	526,337	20,550,721
ResNet101	42,658,176	2,099,201	44,757,377

transfer learning model that was shown to reach state-of-the-art accuracy on common large transfer learning data sets, while being significantly smaller and faster than other CNNs (Tan & Le, 2019). Transfer learning is a widely used technique for countering overfitting in medical applications, especially when data set is small and susceptible to learning noise (Litjens et al., 2017; Shen et al., 2017).

EfficientNet-B0 was used as a transfer learning model to develop a CNN for binary classification of clinical images of the tongue into either benign or "suspicious" (i.e., cancer or OPMD). The CNN model was built using five layers: (1) an input layer with an expected image size of 224 × 224 × 3; (2) transferred convolutional and pooling layers of EfficientNet-B0; (3) a global average pooling layer; (4) a fully connected layer that has 1,024 neurons with rectified linear unit (ReLU) activation functions and a dropout rate of 0.5; and (5) an output layer with a sigmoid activation function (Figure 2).

In order to determine whether there is any performance advantage of using heavyweight transfer learning models compared to our lightweight model, the performance of the proposed CNN was examined using bigger-size transfer learning models, namely VGG19 and ResNet101 (He et al., 2016; Simonyan & Zisserman, 2015; Table 2). Implementation and training configurations of VGG19 and ResNet101 were the same as EfficientNet-B0, except that best performance was achieved for both models (VGG19 and ResNet101) when using Adam optimizer with a learning rate of 1×10^{-3}.

2.3 | Convolutional neural network training and testing

The proposed CNN model was implemented using the Keras 2.4.0 library (keras.io) and was trained using Adam optimizer with a learning rate of 1×10^{-4}, and a mini-batch size of 32 (Kingma & Ba, 2015). All parameters of the convolutional and pooling layers of EfficientNet-B0 were initialized using the pretrained weights on ImageNet, a widely popular transfer learning data set with 1.2 million images and 1,000 classes (Russakovsky et al., 2015). Other parameters in the fully connected layers were initialized using the Glorot uniform initializer (Glorot & Bengio, 2010). For further calibration of the CNN parameters to our classification task, we unfreeze the top 20 hidden layers of EfficientNet-B0, allowing their parameters to be trainable.

The number of clinical images for benign tongue conditions was roughly double (2.5×) the number of clinical images for cancer and OPMD in the training set, an issue called "class imbalance." This can introduce bias in CNNs training and lead to poor predictive performance in testing (Buda et al., 2018). To minimize the impact of "class imbalance," we used the "weighted cross entropy loss" function, which accounts for the number of samples in each class (i.e., the contribution of each image of cancer or OPMD to the training error is 2.5× the contribution of each image of benign tongue conditions).

Two callbacks were used during the training process. The first callback was "early stopping," a mechanism used to continuously monitor the validation loss and stop the training if no further improvement is observed (patience = 5 epoch). The second callback was "model checkpoint," a mechanism made available by the Keras library that allows saving the model with the minimum validation loss during training.

Bootstrapping with 120 repetitions is used to calculate arithmetic means and 95% confidence intervals (CIs) for accuracy, sensitivity and specificity of the proposed CNN model on the test set. Furthermore, the receiver operating characteristic (ROC) curve and area under curve (AUC) were determined (Table 3). All experiments were executed using Google Colab (colab.research.google.com).

3 | RESULTS

The performance parameters of the proposed CNN in binary classification of clinical images of the tongue into either benign or "suspicious" (i.e., cancer or OPMD) are summarized in Table 3. When using EfficientNet-B0 as a transfer learning model, the mean accuracy was 85.0% (95% CI: 81.0%–90.0%), the mean specificity was 84.5% (95% CI: 78.9%–91.5%), and the mean sensitivity was 86.7% (95% CI: 80.4%–93.3%) (Table 4).

The receiver operating characteristic (ROC) curve indicated that the class separability of the proposed model at various classification boundaries is high and close to the ideal curve (i.e., AUC = 1.0; Figure 3). The AUC for EfficientNet-B0 model was 0.928 (95% CI: 0.88–0.96), while it was 0.911 (95% CI: 0.87–0.95) for VGG19 and 0.915 (95% CI: 0.88–0.95) for ResNet101.

4 | DISCUSSION

Early diagnosis of oral cancer is essential to minimize cancer related morbidity and mortality (Schutte et al., 2020; Thomas et al., 2020). Unfortunately, oral cancer is often diagnosed at late stages despite that the oral cavity is readily accessible for visual inspection, and the majority of oral cancers are preceded by noticeable mucosal changes (Ojeda et al., 2020). The use of artificial intelligence to identify suspicious oral mucosal changes can provide a practical, low-cost and universal approach for early detection of oral SCC and OPMDs (Kar et al., 2020; Thankappan et al., 2020; van der Waal, 2018).

There are only few studies that applied AI methods in the context of oral cancer. In a recent study, Kouznetsova et al. (2020) applied machine learning to develop a model for distinction between oral cancer and periodontitis using metabolites from patients' saliva (Kouznetsova et al., 2020). Shaban et al. (2019) used deep learning model to quantify tumour-infiltrating lymphocytes and predict prognosis in oral SCC, Chen et al. (2019) developed one-class learning algorithms for identification of novel genes related to oral SCC, and Chu et al. (2020) described the use of machine learning algorithms for prediction of treatment outcome in oral cancer (Chen et al., 2019; Chu et al., 2020; Shaban et al., 2019). Deep CNNs were demonstrated to have high sensitivity and specificity in discrimination between tongue SCC and normal tissues examined by fibre-optic Raman spectroscopy following surgical excision (Jeng et al., 2019). A recent systematic review analysed the use of CNNs for histopathological diagnosis of oral precancerous and cancerous lesions, reported that the quality of evidence is low, with most studies showing a high risk of bias (Mahmood et al., 2020).

The present study proposed a deep learning CNN for identification of suspicious oral lesions using a data set of verified clinical images. Earlier studies about oral cavity coloured image analysis applied traditional machine learning models that relied on image preprocessing to extract specific features such as colour variations, shapes and textures of oral lesions (Chodorowski et al., 2008; Chodorowski et al., 1999; Thomas et al., 2013). In contrast to traditional machine learning models, deep CNN models can automatically extract discriminating features by examining pixels of a colour image, without the need for handcrafted feature extraction methods. The powerful learning nature of deep CNNs allowed them to overtake traditional machine learning approaches and rapidly become the state-of-the-art technique in computer-aided cancer detection (Choi et al., 2020).

TABLE 3 Metrics used to evaluate the performance of the proposed CNN model on the test set ($n = 100$)

Metric	Definition
Accuracy	The percentage of correctly classified lesions, calculated using the equation: $(TP + TN)/n$
Specificity	The percentage of correctly classified benign lesions, calculated using the equation: $TN/(TN + FP)$
Sensitivity	The percentage of correctly classified malignant lesions. calculated using the equation: $TP/(TP + FN)$
Receiver operating characteristic (ROC) curve	A plot that shows sensitivity against 1 − specificity at various classification boundaries (the higher the curve, the better the model at distinguishing between benign and malignant lesions)
Area under curve (AUC)	Represents the area under the ROC curve (ideal AUC = 1.0 indicates that the model has perfect class separability between benign and malignant lesions)

Note: True positive (TP) represents the count of correctly classified malignant/potentially malignant lesions. True negative (TN) represents the count of correctly classified benign lesions. False positive (FP) represents the count of incorrectly classified malignant lesions. False negative (FN) represents the count of incorrectly classified benign lesions.

Deep learning CNNs were used in the context of oral cancer for image analysis utilizing various types of specialized imaging techniques such as hyperspectral images, confocal laser endomicroscopy imaging, auto-fluorescence and white light images, and conventional microscopy images (Aubreville et al., 2017; Folmsbee et al., 2018; Jeyaraj & Samuel Nadar, 2019; Morikawa et al., 2020; Song et al., 2018). Our study utilized verified clinical images of benign and malignant or potentially malignant lesions to develop a deep learning CNN for early detection of oral cancer. Clinical images included in the present study were obtained using various types of digital cameras and smartphones. As a result, the images have variable sizes, zoom, angle, light conditions and picture orientations reflecting real-time scenarios. Findings of our study indicated that the proposed CNN can be used for detection of cancerous or potentially malignant tongue lesions based on standard colour images with high levels of accuracy (85%), sensitivity (84.5%) and specificity (86.7%) that are comparable to human performance. It has been estimated that the diagnostic accuracy of the specialist practitioner for detection of suspicious oral lesions is estimated at 90%, while the diagnostic accuracy of general dental practitioner is estimated at 75%, when compared to a specialist (Downer et al., 1995). Studies applying deep learning CNN for image analysis in other types of cancer, such as skin cancer, reported similar results with an estimated sensitivity of 83.7% to 97.1% and specificity of 43.6% to 82.7% (Ferrante di Ruffano et al., 2018; Huang et al., 2020; Maron et al., 2021).

The present study adopted a novel approach for CNN utilizing EfficientNet-B0 transfer model. EfficientNet-B0 is a recently developed lightweight transfer learning model that was shown to reach state-of-the-art accuracy on common large transfer learning data sets, while being significantly smaller and faster than other conventional CNNs (Tan & Le., 2019). The transfer learning technique is widely popular in medical applications. This technique is especially useful in situations where the number of input images is small in order to avoid overfitting if a CNN is trained from scratch using a small sample (Litjens et al., 2017; Munir et al., 2019). Findings of the present study indicated that the overall performance of CNN developed using EfficientNet-B0 transfer model was comparable

TABLE 4 Test results for the proposed CNN model with each transfer learning model

Transfer learning model	Accuracy Mean (95% CI)	Specificity Mean (95% CI)	Sensitivity Mean (95% CI)
EfficientNet-B0	85.0 (81.0–90.0)	84.5 (78.9–91.5)	86.7 (80.4–93.3)
VGG19	83.0 (79.0–88.0)	81.5 (75.0–88.1)	86.4 (77.6–92.7)
ResNet101	84.0 (80.0–89.0)	84.4 (77.0–94.6)	83.9 (78.2–92.7)

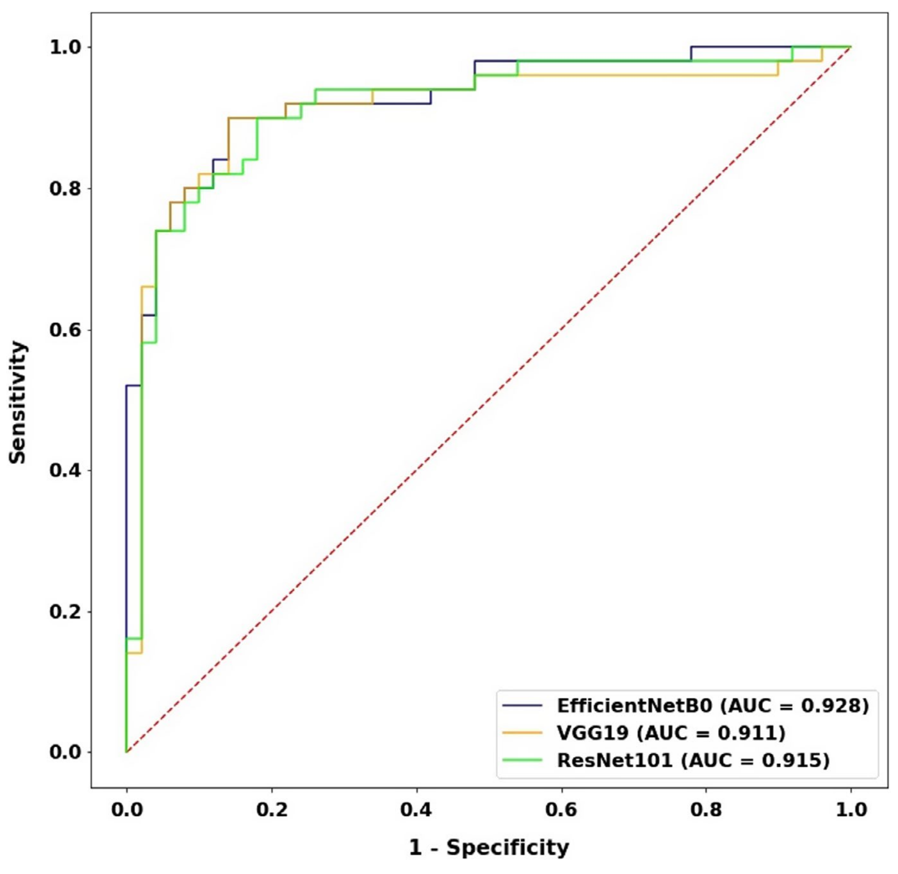

FIGURE 3 The receiver operating characteristic (ROC) curves

to CNN developed using conventional models such as VGG19 and ResNet101 (Shamim et al., 2019; Welikala et al., 2020). These findings demonstrate the feasibility of using a deep lightweight CNN to build a low-budget embedded vision CAD system with limited computation power and memory capacity for automatic tongue cancer detection. Our findings suggest that the proposed CNN can be used as a mobile size application; this might be a potentially practical, low-cost and universal access approach for oral cancer early detection given that 6.3 billion smartphone subscriptions will exist by the year 2021(Cerwall & Report, 2016).

Although our study reported a novel CNN for oral cancer detection, findings of the present study possess some limitations. The data set used in our study was made of clinical images for tongue lesions only and did not include lesions from other mucosal sites. In the main, the tongue is the most common site of oral cancer; however, other mucosal sites can be affected (Warnakulasuriya et al., 2007). Further studies are therefore needed to examine the performance of the proposed CNN for oral cancer detection in other anatomic sub-sites. The present study utilized a data set made of 716 clinical images. The number of images in our data set is relatively large when compared with other studies applying deep learning for image analysis in the context of oral cancer (Chodorowski et al., 2008; Chodorowski et al., 1999; Thomas et al., 2013). However, the size of our data set is small when compared to the size of data sets in other types of cancer such as skin cancer (Esteva et al., 2017). International collaboration is needed to create a large data set of oral cancer and OPMD clinical images. This will facilitate the process of developing and optimising reliable CNNs for oral cancer detection. Furthermore, the number of clinical images for benign conditions was significantly more than the number of clinical images for cancer and OPMDs in the training set (i.e., class imbalance). This can introduce bias in CNNs training and lead to poor predictive performance in testing. To overcome this potential source of bias, we used the "weighted cross entropy loss" function, which accounts for the number of samples in each class. The present study aimed to demonstrate the use of artificial intelligence for identification of suspicious oral lesions using binary classification of tongue lesions into either "benign" or "suspicious," but did not attempt to include images of normal tongue or classify benign tongue lesions by diseases type (i.e., multiclass classification). Further studies are required to examine the performance of artificial intelligence for multiclass classification of oral mucosal lesions.

5 | CONCLUSION

Artificial intelligence is a potentially novel and practical approach for oral cancer early detection. Deep CNN using EfficientNet-B0 transfer model can be used for detection of cancerous or potentially malignant oral lesions with high levels of accuracy, sensitivity and specificity with the additional advantages of being small in size and need small computation power and memory capacity.

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

Fahed Jubair: Conceptualization; Investigation; Methodology; Writing-review & editing. **Omar Al-karadsheh:** Conceptualization; Formal analysis; Methodology; Writing-review & editing. **Dimitrios Malamos:** Conceptualization; Data curation; Resources; Writing-review & editing. **Samara Al Mahdi:** Data curation; Resources; Writing-review & editing. **Yusser Saad:** Resources; Writing-review & editing. **Yazan Hassona:** Conceptualization; Data curation; Investigation; Methodology; Project administration; Resources; Supervision; Writing-original draft; Writing-review & editing.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/odi.13825.

DATA AVAILABILITY STATEMENT

Data are provided in the Sections 2 and 3.

ORCID

Dimitrios Malamos https://orcid.org/0000-0001-7824-0862
Yazan Hassona https://orcid.org/0000-0003-1480-8134

REFERENCES

AAOM Clinical Practice Statement (2016). Subject: Oral cancer examination and screening. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology*, *122*(2), 174–175.

Attardo, S., Chandrasekar, V. T., Spadaccini, M., Maselli, R., Patel, H. K., Desai, M., Capogreco, A., Badalamenti, M., Galtieri, P. A., Pellegatta, G., Fugazza, A., Carrara, S., Anderloni, A., Occhipinti, P., Hassan, C., Sharma, P., & Repici, A. (2020). Artificial intelligence technologies for the detection of colorectal lesions: The future is now. *World Journal of Gastroenterology*, *26*(37), 5606–5616. https://doi.org/10.3748/wjg.v26.i37.5606

Aubreville, M., Knipfer, C., Oetter, N., Jaremenko, C., Rodner, E., Denzler, J., Bohr, C., Neumann, H., Stelzle, F., & Maier, A. (2017). Automatic classification of cancerous tissue in laserendomicroscopy images of the oral cavity using deep learning. *Scientific Reports*, *7*, 11979.

Buda, M., Maki, A., & Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks*, *106*, 249–259.

Califano, J., Van der Riet, P., Westra, W., Hawroz, H., Clayman, G., Piantadosi, S., Corio, R., Lee, D., Greenberg, B., Koch, W., & Sidransky, D. (1996). Genetic progression model for head and neck cancer: Implications for head and neck cancer: Implications for field cancerization. *Cancer Research*, *56*, 2488–2492.

Cerwall, P., & Report, E. M. (2016). *Ericssons mobility report*. https://www.ericsson.com/res/docs/2016/ericsson-mobility-report-2016.pdf

Chaturvedi, A. K., Anderson, W. F., Lortet-Tieulent, J., Curado, M. P., Ferlay, J., Franceschi, S., Rosenberg, P. S., Bray, F., & Gillison, M. L. (2013). Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers. *Journal of Clinical Oncology*, *31*, 4550–4559.

Chen, L., Zhang, Y. H., Huang, G., Pan, X., Huang, T., & Cai, Y. D. (2019). Inferring novel genes related to oral cancer with a network embedding method and one-class learning algorithms. *Gene Therapy*, *26*(12), 465–478. https://doi.org/10.1038/s41434-019-0099-y

Cho, S. J., Sunwoo, L., Baik, S. H., Bae, Y. J., Choi, B. S., & Kim, J. H. (2020). Brain metastasis detection using machine learning: A systematic review and meta-analysis. *Neuro Oncology*, *23*(2), 214–225. https://doi.org/10.1093/neuonc/noaa232

Chodorowski, A., Choudhury, C. R., & Gustavsson, T. A. (2008). Image analysis and CADx system for mucosal lesions. *8th IEEE International Conference on BioInformatics and BioEngineering*.

Chodorowski, A., Mattssonb, U., & Gustavssona, T. (1999). *Oral lesion classification using true-color images*. Proc. SPIE 3661, Medical Imaging 1999: Image Processing.

Choi, K., Choi, S. J., & Kim, E. S. (2020). *Computer-aided diagnosis for colorectal cancer using deep learning with visual explanations*. 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) (pp. 1156–1159). https://doi.org/10.1109/EMBC44109.2020.9176653

Chu, C. S., Lee, N. P., Adeoye, J., Thomson, P., & Choi, S. W. (2020). Machine learning and treatment outcome prediction for oral cancer. *Journal of Oral Pathology and Medicine*, *49*(10), 977–985. https://doi.org/10.1111/jop.13089

Downer, M. C., Evans, A. W., Hughes Hallet, C. M., Jullien, J. A., Speight, P. M., & Zakrzewska, J. M. (1995). Evaluation of screening for oral cancer and precancer in a company headquarters. *Community Dentistry and Oral Epidemiology*, *23*(2), 84–88. https://doi.org/10.1111/j.1600-0528.1995.tb00206.x

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115–118. https://doi.org/10.1038/nature21056

Ferrante di Ruffano, L., Takwoingi, Y., Dinnes, J., Chuchu, N., Bayliss, S. E., Davenport, C., Matin, R. N., Godfrey, K., O'Sullivan, C., Gulati, A., Chan, S. A., Durack, A., O'Connell, S., Gardiner, M. D., Bamber, J., Deeks, J. J., Williams, H. C., Cochrane Skin Cancer Diagnostic Test Accuracy Group (2018). Computer-assisted diagnosis techniques (dermoscopy and spectroscopy-based) for diagnosing skin cancer in adults. *Cochrane Database of Systematic Reviews*, *12*(12), CD013186. https://doi.org/10.1002/14651858.CD013186

Folmsbee, J., Liu, X., Brandwein-Weber, M., & Doyle, S. (2018). Active deep learning: Improved training efficiency of convolutional neural networks for tissue classification in oral cavity cancer. In *2018 IEEE 15th International Symposium on Biomedical Imaging* (pp. 770–773).

Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, JMLR Workshop and Conference Proceedings*, *9*, 249–256.

Hassona, Y., Sawair, F., Matarweh, D., Abdalhamid, A., Thweib, D., & Scully, C. (2018). Oral cancer early detection: What do patients need to know? *Journal of Cancer Education*, *33*(4), 865–869. https://doi.org/10.1007/s13187-017-1191-x

Hassona, Y., Scully, C., Shahin, A., Maayta, W., & Sawair, F. (2016). Factors influencing early detection of oral cancer by primary health-care professionals. *Journal of Cancer Education*, *31*(2), 285–291. https://doi.org/10.1007/s13187-015-0823-2

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778).

Huang, H. W., Hsu, B. W., Lee, C. H., & Tseng, V. S. (2020). Development of a light-weight deep learning model for cloud applications and remote diagnosis of skin cancers. *The Journal of Dermatology*. https://doi.org/10.1111/1346-8138.15683

Huber, M. A. (2018). Adjunctive diagnostic techniques for oral and oropharyngeal cancer discovery. *Dental Clinics of North America*, *62*(1), 59–75.

Ilhan, B., Lin, K., Guneri, P., & Wilder-Smith, P. (2020). Improving oral cancer outcomes with imaging and artificial intelligence. *Journal of Dental Research*, *99*, 241–248.

Jeng, M. J., Sharma, M., Sharma, L., Chao, T. Y., Huang, S. F., Chang, L. B., Wu, S. L., & Chow, L. (2019). Raman spectroscopy analysis for optical diagnosis of oral cancer detection. *Journal of Clinical Medicine*, *8*(9), 1313. https://doi.org/10.3390/jcm8091313

Jeyaraj, P. R., & Samuel Nadar, E. R. (2019). Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm. *Journal of Cancer Research and Clinical Oncology*, *145*(4), 829–837. https://doi.org/10.1007/s00432-018-02834-7

Kar, A., Wreesmann, V. B., Shwetha, V., Thakur, S., Rao, V. U. S., Arakeri, G., & Brennan, P. A. (2020). Improvement of oral cancer screening quality and reach: The promise of artificial intelligence. *Journal of Oral Pathology & Medicine*, *49*(8), 727–730. https://doi.org/10.1111/jop.13013

Karunakaran, K., & Muniyan, R. (2020). Genetic alterations and clinical dimensions of oral cancer: A review. *Molecular Biology Reports*, *47*(11), 9135–9148. https://doi.org/10.1007/s11033-020-05927-0

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations (ICLR)*.

Kouznetsova, V. L., Li, J., Romm, E., & Tsigelny, I. F. (2020). Finding distinctions between oral cancer and periodontitis using saliva metabolites and machine learning. *Oral Diseases*. https://doi.org/10.1111/odi.13591

Lingen, M. W., Abt, E., Agrawal, N., Chaturvedi, A. K., Cohen, E., D'Souza, G., Gurenlian, J., Kalmar, J. R., Kerr, A. R., Lambert, P. M., Patton, L. L., Sollecito, T. P., Truelove, E., Tampi, M. P., Urquhart, O., Banfield, L., & Carrasco-Labra, A. (2017). Evidence-based clinical practice guideline for the evaluation of potentially malignant disorders in the oral cavity: A report of the American dental association. *Journal of the American Dental Association*, *148*(10), 712–727.e10.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, *42*, 60–88. https://doi.org/10.1016/j.media.2017.07.005

Madhura, M. G., Rao, R. S., Patil, S., Fageeh, H. N., Alhazmi, A., & Awan, K. H. (2020). Advanced diagnostic aids for oral cancer. *Disease-a-Month*, *66*(12), 101034. https://doi.org/10.1016/j.disamonth.2020.101034

Mahmood, H., Shaban, M., Indave, B. I., Santos-Silva, A. R., Rajpoot, N., & Khurram, S. A. (2020). Use of artificial intelligence in diagnosis of head and neck precancerous and cancerous lesions: A systematic review. *Oral Oncology*, *110*, 104885. https://doi.org/10.1016/j.oraloncology.2020.104885

Maron, R. C., Haggenmüller, S., von Kalle, C., Utikal, J. S., Meier, F., Gellrich, F. F., Hauschild, A., French, L. E., Schlaak, M., Ghoreschi, K., Kutzner, H., Heppt, M. V., Haferkamp, S., Sondermann, W., Schadendorf, D., Schilling, B., Hekler, A., Krieghoff-Henning, E., Kather, J. N., … Brinker, T. J. (2021). Robustness of convolutional neural networks in recognition of pigmented skin lesions. *European Journal of Cancer*, *145*, 81–91. https://doi.org/10.1016/j.ejca.2020.11.020

Morikawa, T., Shibahara, T., Nomura, T., Katakura, A., & Takano, M. (2020). Non-invasive early detection of oral cancers using fluorescence visualization with optical instruments. *Cancers*, *12*(10), E2771. https://doi.org/10.3390/cancers12102771

Munir, K., Elahi, H., Ayub, A., Frezza, F., & Rizzi, A. (2019). Cancer diagnosis using deep learning: A bibliographic review. *Cancers*, *11*(9), 1235. https://doi.org/10.3390/cancers11091235

Muthukrishnan, N., Maleki, F., Ovens, K., Reinhold, C., Forghani, B., & Forghani, R. (2020). Brief history of artificial intelligence. *Neuroimaging Clinics of North America*, *30*(4), 393–399. https://doi.org/10.1016/j.nic.2020.07.004

Niu, S., Huang, J., Li, J., Liu, X., Wang, D., Zhang, R., Wang, Y., Shen, H., Qi, M., Xiao, Y. I., Guan, M., Liu, H., Li, D., Liu, F., Wang, X., Xiong, Y. U., Gao, S., Wang, X., & Zhu, J. (2020). Application of ultrasound artificial intelligence in the differential diagnosis between benign and malignant breast lesions of BI-RADS 4A. *BMC Cancer*, *20*(1), 959. https://doi.org/10.1186/s12885-020-07413-z

Ojeda, D., Huber, M. A., & Kerr, A. R. (2020). Oral potentially malignant disorders and oral cavity cancer. *Dermatologic Clinics*, *38*(4), 507–521. https://doi.org/10.1016/j.det.2020.05.011

Rajaraman, P., Anderson, B. O., Basu, P., Belinson, J. L., D'Cruz, A., Dhillon, P. K., Gupta, P., Jawahar, T. S., Joshi, N., Kailash, U., Kapambwe, S., Katoch, V. M., Krishnan, S., Panda, D., Sankaranarayanan, R., Selvam, J. M., Shah, K. V., Shastri, S., Shridhar, K., & Mehrotra, R. (2015). Recommendations for screening and early detection of common cancers in India. *The Lancet Oncology*, *16*(7), e352–e361.

Rivera, C. (2015). Essentials of oral cancer. *International Journal of Clinical and Experimental Pathology*, *8*, 11884–11894.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, *115*, 211–252.

Sathyakumar, K., Munoz, M., Singh, J., Hussain, N., & Babu, B. A. (2020). Automated lung cancer detection using artificial intelligence (AI) deep convolutional neural networks: A narrative literature review. *Cureus*, *12*(8), e10017. https://doi.org/10.7759/cureus.10017

Schutte, H. W., van den Broek, G. B., Steens, S. C. A., Hermens, R. P. M. G., Honings, J., Marres, H. A. M., Merkx, M. A. W., Weijs, W. L. J., Arens, A. I. J., van Engen-van Grunsven, A. C. H., van Herpen, C. M. L., Kaanders, J. H. A. M., van den Hoogen, F. J. A., & Takes, R. P. (2020). Impact of optimizing diagnostic workup and reducing the time to treatment in head and neck cancer. *Cancer*, *126*(17), 3982-3990. https://doi.org/10.1002/cncr.33037

Shaban, M., Khurram, S. A., Fraz, M. M., Alsubaie, N., Masood, I., Mushtaq, S., Hassan, M., Loya, A., & Rajpoot, N. M. (2019). A novel digital score for abundance of tumour infiltrating lymphocytes predicts disease free survival in oral squamous cell carcinoma. *Scientific Reports*, *9*(1), 13341. https://doi.org/10.1038/s41598-019-49710-z

Shamim, S., Syed, M., Shiblee, M. U., & Ali, S. (2019). *Automated detection of oral pre-Scancerous tongue lesions using deep learning for early diagnosis of oral cavity cancer*. arXiv e-print: arXiv:1909.08987.

Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, *19*, 221–248.

Shield, K. D., Ferlay, J., Jemal, A., Sankaranarayanan, R., Chaturvedi, A. K., Bray, F., & Soerjomataram, I. (2017). The global incidence of lip, oral cavity, and pharyngeal cancers by subsite in 2012. *CA: A Cancer Journal for Clinicians*, *67*, 51–64.

Siegel, R. L., Miller, K. D., & Jemal, A. (2019). Cancer statistics. *CA: A Cancer Journal for Clinicians*, *69*(1), 7–34.

Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. In *3rd International Conference on Learning Representations (ICLR)*.

Song, B., Sunny, S., Uthoff, R. D., Patrick, S., Suresh, A., Kolur, T., Keerthi, G., Anbarani, A., Wilder-Smith, P., Kuriakose, M. A., Birur, P., Rodriguez, J. J., & Liang, R. (2018). Automatic classification of dual-modalilty, smartphone-based oral dysplasia and malignancy images using deep learning. *Biomedical Optics Express*, *9*(11), 5318–5329. https://doi.org/10.1364/BOE.9.005318

Speight, P. M., Epstein, J., Kujan, O., Lingen, M. W., Nagao, T., Ranganathan, K., & Vargas, P. (2017). Screening for oral cancer – A perspective from the global oral cancer forum. *Oral Surgery*, *Oral Medicine*, *Oral Pathology, and Oral Radiology*, *123*(6), 680–687.

Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In *International Conference on Machine Learning* (vol. *97*, pp. 6105–6114).

Thankappan, K., Birur, P., Raj, M., Djalalov, S., Subramanian, S., Iyer, S., & Kuriakose, A. M. (2020). Oral cancer screening using mobile phone-based(mHealth) approach versus conventional oral examination approach, protocol of a cluster randomized study with cost-effectiveness analysis. *International Journal of Surgery Protocols*, *23*, 1–5. https://doi.org/10.1016/j.isjp.2020.07.001

Thomas, B., Kumar, V., & Saini, S. (2013). Texture analysis based segmentation and classification of oral cancer lesions in color images using ANN. In *2013 IEEE International Conference on Signal Processing, Computing and Control* (pp. 1–5).

Thomas, A.., Manchella, S.., Koo, K.., Tiong, A.., Nastri, A.., & Wiesenfeld, D.. (2020). The impact of delayed diagnosis on the outcomes of oral cancer patients: A retrospective cohort study. *International Journal of Oral and Maxillofacial Surgery*. https://doi.org/10.1016/j.ijom.2020.08.010

van der Waal, I. (2018). Skin cancer diagnosed using artificial intelligence on clinical images. *Oral Diseases*, *24*(6), 873–874. https://doi.org/10.1111/odi.12668

Warnakulasuriya, S., Johnson, N. W., & van der Waal, I. (2007). Nomenclature and classification of potentially malignant disorders of the oral mucosa. *Journal of Oral Pathology and Medicine*, *36*(10), 575–580.

Welikala, R. A., Remagnino, P., Lim, J. H., Chan, C. S., Rajendran, S., Kallarakkal, T. G., Zain, R. B., Jayasinghe, R. D., Rimal, J., Kerr, A. R., Amtha, R., Patil, K., Tilakaratne, W. M., Gibson, J., Cheong, S. C., & Barman, S. A. (2020). Automated detection and classification of oral lesions using deep learning for early detection of oral cancer. *IEEE Access*, *8*, 132677–132693.

Zhang, X., Liu, D., Dong, H., Li, Y., Zhang, Y., Wang, X., Zhang, B., & Bian, L. (2019). Factors associated with delay in presentation among patients for oral cancer. *Journal of Comparative Effectiveness Research*, *8*(12), 1003–1071. https://doi.org/10.2217/cer-2019-0067

How to cite this article: Jubair F, Al-karadsheh O, Malamos D, Al Mahdi S, Saad Y, Hassona Y. A novel lightweight deep convolutional neural network for early detection of oral cancer. *Oral Dis*. 2021;00:1–8. https://doi.org/10.1111/odi.13825