Title
Automatic classification of dual-modality, smartphone-based oral dysplasia and malignancy images using deep learning.

Permalink
https://escholarship.org/uc/item/2mr098vc

Journal
Biomedical optics express, 9(11)

ISSN
2156-7085

Authors
Song, Bofan
Sunny, Sumsum
Uthoff, Ross D
et al.

Publication Date
2018-11-01

DOI
10.1364/boe.9.005318

Peer reviewed
Automatic classification of dual-modality, smartphone-based oral dysplasia and malignancy images using deep learning

Bofan Song,1,8 Sumsum Sunny,2,3 Ross D. Uthoff,1 Sanjana Patrick,4 Amritha Suresh,2,3 Trupti Kolur,2 G. Keerthi,5 Afarin Anbarani,6 Petra Wilder-Smith,6 Moni Abraham Kuriakose,2,3 Praveen Birur,4,5 Jeffrey J. Rodriguez,7 and Rongguang Liang1,9

1 College of Optical Sciences, The University of Arizona, Tucson, AZ, USA
2 Mazumdar Shaw Medical Centre, Bangalore, India
3 Mazumdar Shaw Medical Foundation, Bangalore, India
4 Biocon Foundation, Bangalore, India
5 KLE Society Institute of Dental Sciences, Bangalore, India
6 Beckman Laser Institute, University of California, Irvine, Irvine, CA, USA
7 Dept. of Electrical and Computer Engineering, The University of Arizona, Tucson, AZ, USA
8 songb@email.arizona.edu
9 rliang@optics.arizona.edu

Abstract: With the goal to screen high-risk populations for oral cancer in low- and middle-income countries (LMICs), we have developed a low-cost, portable, easy to use smartphone-based intraoral dual-modality imaging platform. In this paper we present an image classification approach based on autofluorescence and white light images using deep learning methods. The information from the autofluorescence and white light image pair is extracted, calculated, and fused to feed the deep learning neural networks. We have investigated and compared the performance of different convolutional neural networks, transfer learning, and several regularization techniques for oral cancer classification. Our experimental results demonstrate the effectiveness of deep learning methods in classifying dual-modal images for oral cancer detection.

© 2018 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

1. Introduction

Oral cancer ranks the sixth most common malignant tumor globally with high risk in low- and middle-income countries (LMICs). Globally, there are an estimated 529,000 new cases of oral cancer with more than 300,000 deaths each year [1]. The most common risk factors include tobacco chewing, smoking, alcohol consumption, human papillomavirus (HPV) infection and local irritation [2]. Overall, the five-year survival rate for oral and oropharyngeal cancers is about 65% in the United States [3] and 37% in India [4]. Early diagnosis increases survival rates. In India, the five-year survival rate of patients diagnosed at an early stage is 82% whereas the rate is 27% when diagnosed at an advanced stage [5]. With half of all oral cancers worldwide diagnosed at an advanced stage, tools to enable early diagnosis are greatly needed to increase survival rates [6].

Biopsy, the gold standard for diagnosis of oral cancer, is often not ideal as a screening tool, due to invasive nature and the availability of limited expertise at point-of-care [7]. The clinical standard for oral lesion detection is conventional oral examination (COE), which involves visual inspection by a specialist. However, most high risk individuals don’t have access to doctors and dentists in LMICs due to inadequate trained specialists and limited health services [8]. Therefore, a point-of-care oral screening tool is urgently needed. In our previous study using a mobile phone, it was proved that remote specialist consultation
through images transmitted by front line health workers improved early detection of oral cancer [9].

Autofluorescence visualization (AFV) has gained interest for early detection of oral malignant and oral potentially malignant lesions (OPML), and preliminary research indicates AFV can help detect some oral malignant and OPMLs, which are not visible by COE [10] and assist in differentiating dysplasia from normal mucosa by comparing fluorescence loss [11]. However, not all dysplastic lesions display loss of autofluorescence [12–14], and some research concludes the combination of AFV and COE has a better accuracy than either AFV or COE alone [15,16].

Using autofluorescence imaging techniques, several tools have recently been developed as screening adjuncts [13,15–18], but each of them has some major limitations like portability or convenient data transmission capabilities, preventing them from becoming an effective solution for oral cancer screening, particularly in low resource settings. We have developed a dual-mode intraoral oral cancer screening platform which social workers can use to screen a high-risk population, communicate with remote specialists, and access cloud-based image processing and classification as shown in Fig. 1 [19].

![Low-cost, dual-modality smartphone-based oral cancer screening platform.](image)

*Fig. 1. Low-cost, dual-modality smartphone-based oral cancer screening platform. The platform consists of a commercial LG G4 Android smartphone, an external peripheral with intraoral imaging attachment, LEDs, LED driver and batteries, a cloud-based image process and storage server. A custom Android application provides a user interface, controls the phone along with external peripherals, and enables communication with the cloud server and remote specialist.*

While a number of oral cavity detection devices have been developed [10,13,20–23], one challenge is how images are analyzed and presented to the medical users as most of them are not familiar with imaging techniques. A number of automatic image classification methods have been developed for oral cavity images [17,24–26]. Darren et al. developed a simple classification algorithm based on the ratio of red-to-green fluorescence and linear discriminant analysis [26]. Rahman et al. used a similar algorithm to evaluate a simple, low-cost optical imaging system, and identified cancerous and suspicious neoplasia tissue [17]. Recently, convolutional neural networks (CNNs) have proven to be successful in solving image classification problems [27,28], and many medical image analysis tasks [29–32], including detection of diabetic retinopathy in retinal fundus photographs [31], classification of skin cancer images [32], confocal laser endomicroscopy imaging of oral squamous cell
carcinoma [33], hyperspectral imaging of head and neck cancer [34], and orofacial image analysis [35]. In this paper, we propose a dual-modal CNN classification method for oral cancer for point-of-care mobile imaging applications.

2. Materials and methods

2.1 Dual-modal imaging platform

Both autofluorescence and white light imaging have been used for oral cancer detection independently, but each has its own limitation. We propose dual-modal imaging to improve the system sensitivity and specificity. The platform [19] consists of a commercial LG G4 Android smartphone, a whole mouth imaging attachment (not shown), an intraoral imaging attachment, an LED driver, a mobile app, and a cloud-based image processing and storage server as shown in Fig. 1. The system utilizes six 405 nm Luxeon UV U1 LEDs (Lumileds, Amsterdam, Netherlands) to enable the autofluorescence imaging (AFI), and four 4000 K Luxeon Z ES LEDs for white light imaging (WLI). The illumination LEDs are driven by a switching boost voltage regulator (LT3478, Linear Technology, Milpitas, CA). Autofluorescence excitation wavelengths of 400 nm to 410 nm have been shown to be effective in previous studies [26,36,37]. The optical power of UV LEDs delivered to the sample is 2.01 mW/cm², measured by a Newport optical power meter (1916-R, Newport, Irvine, CA), which is below the maximum permissible exposure (MPE) limit [38]. A custom Android application provides a user interface, controls the phone along with external peripherals, and enables communication with the cloud server and remote specialist.

The workflow of the proposed mobile imaging platform is shown in Fig. 2. Autofluorescence and white light image pairs are captured on-site, and then uploaded to a cloud server. The automatic classification algorithm and a reminder email to remote specialists are both triggered when upload is completed. A deep learning approach classifies the image pairs for oral dysplasia and malignancy. After receiving the reminder email, the remote specialist can log into the platform and provide diagnosis from anywhere via web browser. Results from both the deep learning algorithm and remote specialist are saved on the cloud server, and a PDF (portable document format) report is automatically generated when both results are available. The patient will be referred to the specialist if the result turns out suspicious.

![Workflow Diagram](image)

Fig. 2. Workflow of the proposed mobile imaging platform. Autofluorescence and white light images acquired from smartphone are uploaded to cloud server and classified for oral dysplasia and malignancy based on deep learning. Remote diagnosis could be provided by remote specialists anywhere with internet connections. Results can be viewed on-site through the customized mobile app.
2.2 Study population

The study was carried out among patients attending the outpatient clinics of Department of Oral Medicine and Radiology, KLE Society Institute of Dental Sciences (KLE), India. Head and Neck Oncology Department of Mazumdar Shaw Medical Center (MSMC), India, and Christian Institute of Health Sciences and Research (CIHSR), Dimapur, India, during the period of October 2017 to May 2018. Approval of the institutional ethical committee was obtained prior to initiation of the study and written informed consent was obtained from all subjects enrolled in the study. The subjects who were clinically diagnosed with OPML and malignant lesions were recruited in the study. Images of normal mucosa were also collected. Patients who did not consent or had undergone biopsy previously were excluded from the study.

2.3 Data acquisition

Dual-modal image pairs from 190 patients were labelled and separated into two categories by oral oncology specialists from MSMC, KLE, and CIHSR. Normal images were categorized as ‘normal’ and OPML and malignant lesions as ‘suspicious’. Images with low quality due to saliva, defocus, motion, low light, and over-exposure were labeled as ‘Not Diagnostic’ and were not used in the classification study.

Figure 3 shows examples of dual-modal image pairs. Figures 3(a) (AFI) and 3b (WLI) are from the palate of a healthy patient, Figs. 3(c) (AFI) and 3(d) (WLI) are from the buccal mucosa of a patient with an oral potentially malignant lesion, and Figs. 3(e) and 3(f) are image pairs of a malignant lesion from the lower vestibule. The autofluorescence images look yellowish, partially due to the cross-talk between the green and red channels in the CMOS sensor.

We observe that for healthy regions, the autofluorescence signal is relatively uniform across the whole image, and the white light images look smooth and the color is uniform. In contrast, for OPML and malignant regions, there is loss of autofluorescence signal and the color in white light images is not uniform. These features are used to classify the images in the CNN.

![Figure 3. Examples of dual-modal image pairs captured from the dual-modal mobile imaging device. (a) and (b) are autofluorescence image and white light image from the palate of a healthy patient, (c) and (d) are from the buccal mucosa of a patient with oral potentially malignant lesion, and (e) and (f) are image pairs of a malignant lesion from the lower vestibule.](image)

2.4 Image enhancement

The most common artifacts in the data set are low brightness and contrast. An adaptive histogram equalization method is used to improve the quality for the images with low brightness and contrast. The method increases contrast by equalizing the distribution of pixel intensities in the image. Since the ratio between red and green color channels can be an important feature to discriminate OPML and malignant lesion from normal ones [26,39], equalizing the fluorescence images for each channel ensures they have the same intensity scale to maintain a consistent interpretation of ratio between color channels. The image is also sharpened by subtracting a Gaussian blurred version of the image.
2.5 Data preparation for dual-modal image classification

For dual-modal CNN-based classification, a new, three-channel data set is created from the autofluorescence and white light images. Since a long pass filter is added in front of the CMOS sensor to block the excitation wavelengths, the blue channel of the WLI has low signal and high noise (Fig. 4) and is excluded. In autofluorescence images, the ratio of the red and green channels contains information related to the loss of fluorescence signal and is correlated with OPML and malignant lesions [26,39]. Instead of using both three-channel images to train the CNN, we create a single three-channel image using the green and red channels from the white light image and the normalized ratio of red and green channels from the autofluorescence image as the third channel. The fused data feeds the neural network.

![Fig. 4. Overview of the data preparation for dual-modal image classification. A new, three-channel data set is created from the autofluorescence and white light image pairs. The blue channel of the white light image which has low signal and high noise is excluded. The new three-channel image uses the green and red channels from the white light image and the normalized ratio of red and green channels from autofluorescence image as the third channel.](Image)

2.6 Dual-modal deep learning methods

There are several deep learning frameworks available [40–42]. MatConNet [42], an open source CNN toolbox designed by Visual Geometry Group (Oxford University), is used in our application [43]. Since we have a small number of image data to train the CNN classifier, the complexity and the number of free parameters of the deep neural network may cause an overfitting problem, resulting in low accuracy or training difficulty. We investigate transfer learning, data augmentation, dropout, and weight decay to improve the performance and avoid overfitting.

**Transfer learning**: One common method to solve the problem of limited training data is transfer learning, where a model trained for one task is reused as the starting point for another task [44], leveraging knowledge learned from a different task to improve learning in the target task. The key benefits of transfer learning include reduced time to fully learn the new image data set and overall better performance. Transfer learning has been effective in the use of deep CNNs in medical imaging applications [45,46]. We compare the performance of transfer learning with networks VGG-CNN-M and VGG-CNN-S from Chatfield et al. [47] and VGG-16 from Simonyan et al. [48]. Our study uses CNN-M pre-trained using ImageNet [49], the last two layers (the fully connected layer and softmax layer) are replaced by a new dense layer and softmax layer. The transfer learning process is completed by training the new network with our fused dual-modal image data. The same data augmentation strategy is
applied during each training process, and each network is trained for 100 epochs. The results and network comparisons are discussed in Section 4.

Data augmentation: Data augmentation has been widely used to improve the classification performance of deep learning networks [32,33]. Here we enrich the data provided to the classifier by flipping and rotating original images since our probe images have no natural orientation. Seven additional images are generated from each image by rotating and flipping the original as shown in Fig. 5. The same data augmentation strategy has been applied to each training process.

![Data augmentation](image)

Fig. 5. Data augmentation. Example of the data-augmented images. A total of 8 images are obtained from a single image by rotating and flipping the original image.

Regularization techniques: A network may over-fit the data when the network is complex and the training data set is small. Regularization is a technique to avoid overfitting. Proposed by Srivastava et al. in 2014 [50], dropout is an effective regularization technique [28,50]. This method drops neurons along with their connections in hidden layers during the training process with certain probability. By randomly dropping neurons in a certain layer, some dependencies with specific neurons which only help to classify training images but do not work with test images are removed. The remaining dependencies work well with both training and test data. Because our data set is very small, overfitting is likely to occur. Therefore, a dropout layer with a drop probability of 0.5 is added between the second fully connected layer and the last one.

Weight decay is another regularization technique used to address overfitting problems [51]. It limits the growth of the weights by introducing a penalty to large weights in the cost function:

$$E(w) = E_0(w) + \frac{\lambda}{2} \sum_i w_i^2$$

(1)

where $E_0$ is the original cost function, each $w_i$ is the weight for free parameter of the network, and $\lambda$ governs how strongly large weights are penalized.

3. Results

After excluding low quality images due to saliva, defocus, motion, low light, and over-exposure, 170 image pairs are used in the study and are assigned to categories ‘normal’ or ‘suspicious’, where ‘suspicious’ includes images labeled with OPML and malignant lesions. There are 84 image pairs in the ‘suspicious’ category and 86 images pairs in the ‘normal’ category; 66 normal images and 64 suspicious images are used to train the CNN, and 20 images in each category are used for validation (Table 1).
Table 1. Image data set used for developing deep learning image classification method.

|       | Training set | Validation set |
|-------|--------------|----------------|
| Normal| 66           | 20             |
| Suspicious | 64           | 20             |

In the study, we have compared the performance of different networks, applied data augmentation and regularization techniques, and compared our fused dual-modal method with single modal WLI or AFI. For each comparison, all settings except the one under investigation are fixed.

3.1 Architecture

While increasing the depth of network is typically beneficial for the classification accuracy [48], sometimes the accuracy becomes worse when training a small image data set [52]. To achieve the best performance with our small data set, we have evaluated a number of neural network architectures. Figure 6(a) shows the performances of three different architectures: VGG-CNN-M, VGG-CNN-S, and VGG-16. VGG-16 contains 16 layers, of which 13 are convolutional layers. VGG-16 has the largest number of convolutional layers of the three networks, but the worst performance. Both VGG-CNN-M and VGG-CNN-S have five convolutional layers. In VGG-CNN-M the stride in layer conv2 is two and the max-pooling window in layers conv1 and conv5 is 2x2. In VGG-CNN-S the stride in layer conv2 is one, and the max-pooling window in layers conv1 and conv5 is 3x3. VGG-CNN-M has the best performance, and we use this architecture in our study.

3.2 Data augmentation

Data augmentation is implemented to increase the data set size. Figure 6(b) shows the increased accuracy from using data augmentation. As expected, the accuracy increases with a larger data set.

3.3 Regularization

Various weight decay parameters have been explored and $\lambda = 0.001$ has the best performance as shown in Fig. 6(c). Additional studies could help to find the optimal weight decay value. As shown in Fig. 6(d), we can further improve the performance by applying the dropout technique.
Fig. 6. (a) The validation errors of three different neural network architectures: VGG-CNN-M, VGG-CNN-S, and VGG-16. VGG-CNN-M performs best among these three networks. (b) The comparison between the neural networks trained with and without augmented data. (c) The performances for different weight decays. (d) The performance with and without Dropout.

3.4 Dual-modal image vs. single-modal image

To compare classification performance, we train the VGG-CNN-M independently with AFI, WLI, and dual-modal images (Fig. 7). Clearly, a neural network trained with dual-modal images has the best performance while the one with only white light images has the worst performance. This result is expected as more features can be extracted from dual-modal images to classify images. This study also partially validates that autofluorescence imaging has better performance in detecting oral cancer than white light imaging. Finally, the best result uses transfer learning of the VGG-CNN-M module along with the data augmentation of the dual-modal images, resulting in a 4-fold cross-validation average accuracy of 86.9% at a sensitivity of 85.0% and specificity of 88.7%, as shown in Fig. 7.

Fig. 7. The performance of neural networks trained with dual-modal images, AFI, and WLI.
4. Discussion

The images in this study were captured from buccal mucosa, gingiva, soft palate, vestibule, floor of mouth and tongue. One challenge associated with the current platform is that different tissue structural and biochemical compositions at different anatomical regions have different autofluorescence characteristics, so the performance of the trained neural network will be potentially impacted. This problem could be addressed in the future when we have sufficient image data at every anatomical region to train neural networks and classify each anatomical region separately.

In this study, our goal is to develop a platform for community screening in LMICs, we currently focus on distinguishing dysplasia and malignancy tissue from normal ones. While, in clinical practice, it would be of interest for clinicians to distinguish different stages of cancerous tissue, this work provides the platform for further studies towards a multi-stage classification in the future.

One remaining question on is whether the images marked as clinically normal or suspicious actually reflect the true condition of the patients because the diagnosis results from the specialists are used as the ground truth, not the histopathological results. The device will act as a screening device at rural communities, and patients will be referred to hospitals and clinics for further test and treatment if a suspicious result appears.

5. Conclusion

Smartphones integrate state-of-art technologies, such as fast CPUs, high-resolution digital cameras, and user-friendly interfaces. An estimated six billion phone subscriptions exist worldwide with over 70% of the subscribers living in LMICs [53]. With many smartphone-based healthcare devices being developed for various applications, perhaps the biggest challenge of mobile health concerns is the volume of information these devices will produce. Because adequately training specialists to fully understand and diagnose the images is challenging, automatic image analysis methods, such as machine learning, are urgently needed.

In this paper, we present a deep learning image classification method based on dual-modal images captured from our low-cost, dual-mode, smartphone-based intraoral screening device. To improve the accuracy, we fuse AFI and WLI information into one three-channel image. The performance with fused data is better than using either white light or autofluorescence image alone.

We have compared convolutional networks having different complexities and depths. The smaller network with five convolutional layers works better than a very deep network containing 13 convolutional layers, most likely due to the increase of the network’s complexity causing overfitting when training a small data set. We have also investigated several methods to improve the learning performance and address the overfitting problem. Transfer learning can greatly reduce the number of parameters and number of images needed to train. Data augmentation used to increase the training data set size and regularization techniques such as weight decay and dropout are effective in avoiding overfitting.

In the future, we will capture more images and increase the size of our data set. With a bigger image data set, the overfitting problem will reduce and more complex deep learning models can be applied for more accurate and robust performance. Furthermore, due to the variation in tissue types in the oral cavity, performance is potentially degraded when classifying them together. With additional data, we will be able to perform more detailed multi-class (different anatomical regions) and multi-stage classification using deep learning.
Funding

National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health (UH2EB022623); National Institutes of Health (NIH) (S10OD018061); National Institutes of Health Biomedical Imaging and Spectroscopy Training (T32EB000809).

Acknowledgments

We are grateful to TIDI Products for manufacturing custom shields for our device.

Disclosures

The authors declare that there are no conflicts of interest related to this article.

References

1. World Health Organization, “Oral cancer,” http://www.who.int/cancer/prevention/diagnosis-screening/oral-cancer/en/.
2. S. Petti and C. Scully, “Oral cancer knowledge and awareness: primary and secondary effects of an information leaflet,” Oral Oncol. 43(4), 408–415 (2007).
3. American Society of Clinical Oncology, “Oral and Oropharyngeal Cancer: Statistics,” https://www.cancer.net/cancer-types/oral-and-oropharyngeal-cancer/statistics.
4. M. K. Mallath, D. G. Taylor, R. A. Badwe, G. K. Rath, V. Shanta, C. S. Pramesh, R. Digumarti, P. Sebastian, B. B. Borthakur, A. Kalwar, S. Kapoor, S. Kumar, J. L. Gill, M. A. Kuriakose, H. Malhotra, S. C. Sharma, S. Shukla, L. Viswanath, R. T. Chucko, J. L. Pautu, K. S. Reddy, K. S. Sharma, A. D. Purushotham, and R. Sullivan, “The growing burden of cancer in India: epidemiology and social context,” Lancet Oncol. 15(6), e205–e212 (2014).
5. National Institute of Cancer Prevention and Research, “Statistics,” http://cancerindia.org.in/statistics/.
6. A. C. O. V. Le Campion, C. M. B. Ribeiro, R. R. Luiz, F. F. da Silva Junior, H. C. S. Barros, K. C. B. dos Santos, S. J. Ferreira, L. S. Goncalves, and S. M. S. Ferreira, “Low survival rates of oral and oropharyngeal squamous cell carcinomas,” Int. J. Dent. 2017, 1–7 (2017).
7. K. D. Shield, J. Ferlay, A. Jemal, R. Sankaranarayanan, A. K. Chaturvedi, F. Bray, and I. Soerjomataram, “The global incidence of lip, oral cavity, and pharyngeal cancers by subsite in 2012,” CA Cancer J. Clin. 67(1), 51–64 (2017).
8. K. R. Coelho, “Challenges of the oral cancer burden in India,” J. Cancer Epidemiol. 2012, 701932 (2012).
9. P. N. Birur, S. P. Sunny, S. Jena, U. Kandasarma, S. Raghavan, B. Ramaswamy, S. P. Shanmugam, S. Patrick, R. Kuriakose, J. Malliaiah, A. Suresh, R. Chigurupati, R. Desai, and M. A. Kuriakose, “Mobile health application for remote oral cancer surveillance,” J. Am. Dent. Assoc. 146(12), 886–894 (2015).
10. P. M. Lane, T. Gilhuly, P. D. Whitehead, H. Zeng, C. Poh, S. Ng, M. Williams, L. Zhang, M. Rosin, and C. E. MacAulay, “Simple device for the direct visualization of oral-cavity tissue fluorescence,” J. Biomed. Opt. 11, 24006–24007 (2006).
11. I. Pavlova, M. Williams, A. El-Naggar, R. Richards-Kortum, and A. Gillenwater, “Understanding the Biological Basis of Autofluorescence Imaging for Oral Cancer Detection: High-Resolution Fluorescence Microscopy in Viable Tissue,” Clin. Cancer Res. 14(8), 2396–2404 (2008).
12. R. Mehrrota, M. Singh, S. Thomas, P. Nair, S. Pandya, N. S. Nigam, and P. Shukla, “A cross-sectional study evaluating chemiluminescence and autofluorescence in the detection of clinically innocuous precancerous and cancerous oral lesions,” J. Am. Dent. Assoc. 141(2), 151–156 (2010).
13. K. H. Awan, P. R. Morgan, and S. Warnakulasuriya, “Evaluation of an autofluorescence based imaging system (VELscope™) in the detection of oral potentially malignant disorders and benign keratoses,” Oral Oncol. 47(4), 274–277 (2011).
14. M. Rana, A. Zapf, M. Kuehle, N.-C. Gellrich, and A. M. Eckardt, “Clinical evaluation of an autofluorescence diagnostic device for oral cancer detection: a prospective randomized diagnostic study,” Eur. J. Cancer Prev. 21(5), 460–466 (2012).
15. C. S. Farah, L. McIntosh, A. Georgiou, and M. J. McCullough, “Efficacy of tissue autofluorescence imaging (velscope) in the visualization of oral mucosal lesions,” Head Neck 34(6), 856–862 (2012).
16. V. Jayaprakash, M. Sullivan, M. Merzianu, N. R. Riguil, T. R. Loree, S. R. Popat, K. B. Moysich, S. Ramamanta, T. Johnson, J. R. Marshall, A. D. Hutson, T. S. Mang, B. C. Wilson, S. R. Gill, J. Frustino, A. Bogaards, and M. E. Reid, “Autofluorescence-guided surveillance for oral cancer,” Cancer Prev. Res. 2(11), 966–974 (2009).
17. M. S. Rahman, N. Ingole, D. Roblyer, V. Stepancek, R. Richards-Kortum, A. Gillenwater, S. Shastri, and P. Chaturvedi, “Evaluation of a low-cost, portable imaging system for early detection of oral cancer,” Head Neck Oncol. 2(1), 10 (2010).
18. C. F. Poh, C. E. MacAulay, L. Zhang, and M. P. Rosin, “Tracing the “At-Risk” Oral Mucosa Field with Autofluorescence: Steps Toward Clinical Impact,” Cancer Prev. Res. 2(5), 401–404 (2009).
24. M. M. R. Krishnan, V. Venkatraghavan, U. R. Acharya, M. Pal, R. R. Paul, L. C. Min, A. K. Ray, J. Chatterjee, J. L. N. Roodenburg, “Classification of clinical autofluorescence spectra of oral leukoplakia using an artificial neural network: a pilot study,” Oral Oncol. 44(11), 1059–1066 (2008).

25. M. Rahman, P. Chaturuvedi, A. Gillenwater, and R. R. Richards-Kortum, “Low-cost, multimodal, portable screening system for early detection of oral cancer,” J. Biomed. Opt. 13, 30502–30503 (2008).

26. D. Roblyer, C. Kurachi, V. Stepanski, M. D. Williams, A. K. El-Naggar, M. R. Descour, and R. R. Richards-Kortum, “In vivo imaging of oral neoplasia using a miniaturized fiber optic confocal reflectance microscope,” Oral Oncol. 36(3), 286–293 (2000).

27. D. Robbyr, C. Kurachi, V. Stepanski, M. D. Williams, A. K. El-Naggar, J. J. Lee, A. M. Gillenwater, and R. Richards-Kortum, “Objective detection and delineation of oral neoplasia using autofluorescence imaging,” Cancer Prev. Res. (Phila.) 2(5), 423–431 (2009).

28. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” in Advances in Neural Information Processing Systems 25, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, eds. (Curran Associates, Inc., 2012), pp. 1097–1105.

29. Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature 521(7553), 436–444 (2015).

30. F. Bastien, P. Lamblin, R. Pascanu, J. Bergstra, I. Goodfellow, A. Bergeron, N. Bouchard, D. Warde-Farley, and Y. Bengio, “Theano: new features and speed improvements,” arXiv Prepr. arXiv1211.5590 (2012).

31. V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widener, T. Madams, J. Cuadros, R. Kim, R. Raman, P. C. Nelson, J. L. Mega, and D. R. Webster, “Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs,” JAMA 316(22), 2402–2410 (2016).

32. A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, “Dermatologist-level classification of skin cancer with deep neural networks,” Nature 542(7639), 115–118 (2017).

33. M. Aubreville, C. Knipfer, N. Oetter, C. Jaremenko, E. Rodner, J. Denzler, C. Bohr, H. Neumann, F. Stelzle, and A. Maier, “Automatic Classification of Cancerous Tissue in Laserendomicroscopy Images of the Oral Cavity using Deep Learning,” Sci. Rep. 7(1), 11979 (2017).

34. M. Halicek, G. Lu, J. V Little, X. Wang, M. Patel, C. C. Griffith, M. W. El-Deiry, A. Y. Chen, and B. Fei, “Deep convolutional neural networks for classifying head and neck cancer using hyperspectral images,” J. Biomed. Opt. 22, 60503–60504 (2017).

35. R. Anantharaman, V. Anantharaman, and Y. Lee, “Oro Vision: Deep Learning for Classifying Orofacial Diseases,” in 2017 IEEE International Conference on Healthcare Informatics (ICHI) (2017), pp. 39–45.

36. D. L. Heintzelman, U. Utzinger, H. Fuchs, A. Zuluaga, K. Gossage, A. M. Gillenwater, R. Jacob, B. Kemp, and R. R. Richards-Kortum, “Optimal excitation wavelengths for in vivo detection of oral neoplasia using fluorescence spectroscopy,” Photochem. Photobiol. 72(1), 103–113 (2000).

37. E. Sivistun, R. Alizadeh-Naderi, A. El-Naggar, R. Jacob, A. Gillenwater, and R. Richards-Kortum, “Vision enhancement system for detection of oral cavity neoplasia based on autofluorescence,” Head Neck 26(3), 205–215 (2004).

38. Laser Institute of America, American National Standard for the Safe Use of Lasers, ANSI Z136.1–2007 (2007).

39. M. C. Pierce, R. A. Schwarz, V. S. Bhattar, S. Mondrik, M. D. Williams, J. J. Lee, R. Richards-Kortum, and A. M. Gillenwater, “Accuracy of in vivo multimodal optical imaging for detection of oral neoplasia,” Cancer Prev. Res. (Phila.) 5(6), 801–809 (2012).

40. F. Bastien, P. Lamblin, R. Pascanu, J. Bergstra, I. Goodfellow, A. Bergeron, N. Bouchard, D. Warde-Farley, and Y. Bengio, “Theano: new features and speed improvements,” arXiv Prepr. arXiv1211.5590 (2012).

41. M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng, “TensorFlow: A System for Large-scale Machine Learning,” in Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation, OSDI’16 (USENIX Association, 2016), pp. 265–283.
43. “MatConvNet,” http://www.vlfeat.org/matconvnet/.
44. S. J. Pan and Q. Yang, “A Survey on Transfer Learning,” IEEE Trans. Knowl. Data Eng. 22(10), 1345–1359 (2010).
45. H. C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, “Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning,” IEEE Trans. Med. Imaging 35(5), 1285–1298 (2016).
46. N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and Jianming Liang, “Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?” IEEE Trans. Med. Imaging 35(5), 1299–1312 (2016).
47. K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman, “Return of the Devil in the Details: Delving Deep into Convolutional Nets,” CoRR abs/1405.3, (2014).
48. K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” CoRR abs/1409.1, (2014).
49. J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, “ImageNet: A large-scale hierarchical image database,” in 2009 IEEE Conference on Computer Vision and Pattern Recognition (2009), pp. 248–255.
50. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” J. Mach. Learn. Res. 15, 1929–1958 (2014).
51. A. Krogh and J. A. Hertz, “A Simple Weight Decay Can Improve Generalization,” in Proceedings of the 4th International Conference on Neural Information Processing Systems, NIPS ’91 (Morgan Kaufmann Publishers Inc., 1991), pp. 950–957.
52. K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” CoRR abs/1512.0, (2015).
53. International Telecommunication Union, “Statistics,” https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx.