Real-time classification on oral ulcer images with residual network and image enhancement

Jianbin Guo1 | Haolin Wang2 | Xingsi Xue3 | Mengting Li4 | Zhongxiong Ma1

1 Fujian Key Laboratory of Oral Diseases & Fujian Provincial Engineering Research Center of Oral Biomaterial & Stomatological Key lab of Fujian College and University, School and Hospital of Stomatology, Fujian Medical University, Fuzhou, China
2 Intelligent Information Processing Research Center, Fujian University of Technology, Fuzhou, Fujian, China
3 Fujian Provincial Key Laboratory of Big Data Mining and Applications, Fujian University of Technology, Fuzhou, Fujian, China
4 School of Computer Science and Mathematics, Fujian University of Technology, Fuzhou, Fujian, China

Correspondence
Xingsi Xue, Fujian Provincial Key Laboratory of Big Data Mining and Applications, Fujian University of Technology, Fuzhou, Fujian 350118, China.
Email: jack8375@gmail.com

Funding information
Young Teacher Education and Research Fund of Fujian Provincial Education Department, Grant/Award Number: JT180208, Fujian Province Undergraduate Universities Teaching Reform Research Project, Grant/Award Number: FBJG20190156, 2018 Program for Outstanding Young Scientific Researcher in Fujian Province University, Intelligent Computing and Applied Scientific Research Innovation Team of Concord University College Fujian Normal University in 2020, Grant/Award Number: 2020-TD-001, Research Project on Lifelong Education in Fujian Province in 2020, Natural Science Foundation of Fujian Province, Grant/Award Number: 2020J01646

Abstract
With the advances of deep learning research in the past few years, healthcare and smart medicines have been significantly developed. Inspired by the wide application of deep learning in medical image classification and disease diagnosis, this paper further proposes a variant of the Residual Network framework to classify the oral ulcer images in real-time. In particular, image pre-processing and enhancement techniques are used to enrich the datasets and reduce model overfitting. Besides, the transfer learning is further introduced into the residual blocks to improve the classification accuracy, with the later layers trained from the labeled datasets. To validate the performance of authors’ proposal, it is compared with other classic deep learning models with respect to the classification sensitivity, specificity, and accuracy. The experimental results show that authors’ approach outperforms those classic classification networks when the oral ulcers are classified and diagnosed in real-time.

1 | INTRODUCTION

Deep learning, a subfield of machine learning (ML), has seen a tremendous resurgence over the past few years [1–5], which is mainly due to the increased computational power and the availability of large numbers of new datasets. As medical data and records continue to grow, healthcare and smart medicines can significantly benefit deep learning [6, 7]. Some of the greatest successes of deep learning have been in medical real-time image processing, electronic health records (EHRs), surgical robotics, and DNA sequence marker detection [7–9]. Currently, deep learning has become an essential component for the development of healthcare and medicines.

In recent years, many scholars have applied deep learning to medical image classification and disease diagnosis [10], which has achieved remarkable results. Convolutional neural network (CNN) is the most widely used deep learning model in the smart medicines [11–13], including VGG-16, AlexNet, and GoogleNet. Through medical image segmentation, recognition and classification, the automatic diagnosis of diseases is realized [14–16]. For example, Hazlett et al. [17] applied deep learning models to predict the level of risk of autism in children. However, few studies were found wherein deep learning was used for the oral ulcer diagnosis. Inspired by the success of the deep learning in the disease diagnosis, this work proposes a novel deep learning model and uses it to diagnose the diseases on oral ulcers in real-time. Since the Residual Network (ResNet) has the superiority to other common convolutional neural networks [11], we further propose a modification of the ResNet framework for the detection of oral ulcers. In particular, we use...
image pre-processing and enhancement techniques to improve the generalizability of the fitted model. After that, the transfer learning is applied to the residual blocks, and the later layers are trained using the datasets labeled by dental specialists.

The rest of this paper is as follows: in Section 2, we review the application of convolutional neural networks in medical images; Section 3 presents the proposed model in detail; in Section 4, experimental results are described; finally, a conclusion is provided in Section 5.

2 | RELATED WORK

Recent years have witnessed the wide application of CNN in medical imaging, which has dramatically contributed to the development of smart medicines. Through processing and recognition of images, CNN can assist doctors in diagnosing diseases.

Jeremy et al. [18] used CNN to identify skin cancer images, partitioned the input images into multiple different resolutions, and extracted different features of the images. They trained on the Dermofit skin cancer datasets and obtained a diagnostic accuracy of 78.1%. Hosseini et al. [19] identified brain MRI images to diagnose Alzheimer’s, using 3D convolutional neural networks (3D-CNN). In the proposed model, the parameters of the fully connected upper convolutional layer were fine-tuned and trained by the ADNI datasets. In particular, the diagnostic accuracy of the presented method achieved 89.1%. Besides, Bakkouri et al. [20] provided a novel Computer-Aided Diagnosis (CAD) system based on 3D multi-scale feature blocks for patients screening of MRI imaging modality standard. The experimental results demonstrate that the proposed method obtains state-of-the-art performance compared with the existing conventional methods. Xiaojie Huang et al. [21] presented a 3D-CNN to detect lung nodules types. In the presented model, the lung nodule features were extracted by a local geometric model filter, and the classification accuracy obtained was higher than 90%. Similarly, Shen et al. [22] proposed a hierarchical learning framework—multiscale convolutional neural network (MCNN)—that captures the heterogeneity of lung nodules by extracting features from stacked layers. Dorj et al. [23] presented AlexNet network to extract the features of skin cancer images and classify them with the SVM classifier. Collecting 3753 skin cancer images for training, the recognition accuracy was 95.1%. Khan et al. [24] built a deep learning framework for breast cancer diagnosis which learns features from different convolutional neural networks. The model has higher classification accuracy than other deep learning models. Gour et al. [25] proposed a ResNet-based model for identifying breast cancer on histopathological images. The proposed framework is able to distinguish between benign and malignant classes on histopathological images. The diagnostic accuracy of the model is 92.25% and the F1-score is 93.45%, which is significantly better than other existing convolutional neural networks. Recently, deep learning is widely used in oral disease diagnosis. These works [26–29] all proposed CNN-based approaches to classify and diagnose histopathological images of oral cancer. The presented methods have good performance in terms of accuracy of diagnosis and classification. Different from those existing works, we aim at proposing a novel deep learning model with residual network and image enhancement and apply it to the real-time diagnosis of oral ulcers.

3 | METHODOLOGY

The outstanding performance of CNN in image recognition motivated us to investigate a model for oral ulcer image classification. In this work, we suggest an improved residual network for the classification of oral ulcer images. The proposed framework is illustrated by Figure 1 and described in detail in the subsequent subsections.

3.1 | Image enhancement and pre-processing

The performance of CNN models depends heavily on the quality and number of available datasets [30]. However, there is no publicly labeled oral ulcer datasets available online. It is difficult to get large datasets to train the proposed network. Therefore, image pre-processing and enhancement techniques play an important role [31] in processing the oral images that we already have. Pre-processing is used to remove different types of noise in oral images. Image enhancement is applied to expand the datasets, achieve greater accuracy, and reduce model overfitting.

In this method, we used single-sample image enhancement techniques to create multiple versions of the oral image by color processing, transformations (translation, scaling, and rotation), and flipping, which help prevent the model overfitting. Then, prior to training the network, stain normalization is used to normalize the oral images and eliminate the variations in color and intensity of the datasets, which improves the prediction accuracy. Figure 2 presents the augmented images generated by the proposed method.

3.2 | Transfer learning

Data dependence is one of the most serious problems in deep learning [32]. The most important difference between traditional machine learning and deep learning is the performance when the scale of dataset is scaled up. When the scale of training dataset is very small, the deep learning method’s performance is barely satisfactory. This results in the poor generalization of the model. However, this is not the case in practice. Obtaining a matching dataset is a complex task in most practical applications. Transfer learning is an important tool to solve the short-age of training data, and it can transfer the parameters from the already trained model to the new one. There is no need to train the model from scratch, which can dramatically reduce training data and training time.

The parameters in the trained model are transferred into the new model to learn basic low-level features, which can speed up the problem-solving process of the new model. In
FIGURE 1  Diagram of the proposed framework. Conv means convolutional layer, followed by the parameters of the layer; Avg Pooling means average pooling layer; FC means fully connected layer; the red font means the classification result is ulcerated oral cavity.

this approach, firstly, the ResNet architecture is trained from the ImageNet datasets. Secondly, the weights of the residual blocks in the network are frozen and the network is retrained with the labeled datasets, fine-tuning the parameters of the later layers. Thirdly, in order to adapt to the classification of oral ulcer images, we transform the model into a binary classification model with softmax layer and binary cross-entropy.

3.3 Residual learning

Generally, the number of layers in the network plays a key role in deep learning models. A deep neural network is the neural network with many hidden layers, which allows it to model complex nonlinear functions more effectively than the one with single-layer. However, as the deeper networks begin to converge, a degradation problem is exposed: as the depth of the network increases, the accuracy tends to saturate and degrades rapidly [14]. To solve the degradation problem, ResNet was proposed.

The most important components of the residual network are the residual blocks. The learned feature of stacked nonlinear layers (composed of several layers) is denoted by $H(x)$.
where $x$ is the input, and we use $H(x)$ to learn the residuals $F(x) = H(x) - x$. The original learned features are reconverted to $F(x) + x$. The reason for this is that residual learning is easier than direct learning to learn the original features. The formula $F(x) + x$ can be implemented under a feedforward neural network with “shortcut connections”. Shortcut connections have skip connections to jump over one or more layers, shown in Figure 3.

**4 | EXPERIMENT**

**4.1 | Experimental configuration**

Since there is no publicly labeled dataset available online on oral ulcer, we obtained a dataset labeled by dental specialists from Fujian Stomatological Hospital. The dataset includes 360 images of oral mucosa from 78 patients. Among them, 107 images belong to normal oral mucosa and the rest 253 ones ulcerative oral mucosa. Figure 4 shows the sample images of normal oral mucosa and ulcerative oral mucosa in the dataset. Table 1 presents detailed information on the distribution of normal and ulcerative oral cavity images.

The dataset augmented is divided into three groups according to 1:1:8 as test set, validation set, and training set. Train the model using the training set. Tune the hyperparameters utilizing the validation set to reduce the model overfitting. After freezing the model and hyperparameters, we evaluated the model with the test set.

The proposed model is compared with two classical CNN frameworks: AlexNet [33], GoogleNet [34]. For the controlled model to be well adapted to the application of binary classification, we transform the AlexNet and GoogleNet model for binary classification model with softmax and binary cross entropy. In order to be consistent with the proposed approach, these models also require ImageNet datasets for pre-training, and then the network parameters are configured according to transfer learning. All the three models are trained for 30 epochs.

**4.2 | Evaluation metrics**

Sensitivity, specificity and accuracy were used as evaluation metrics to assess the performance of the method. The accuracy, sensitivity and specificity are defined in this study by following equations, respectively.

$\text{Accuracy} = \frac{TN + TP}{FP + FN + TP + TN}$

$\text{Sensitivity} = \frac{TP}{FN + TP}$

$\text{Specificity} = \frac{TN}{FP + TN}$

where $TP$ means true positive, that is, the correctly classified oral ulcer images, $FP$ false positive, that is, the misclassified oral ulcer images, $FN$ false negative, that is, the misclassified normal oral images, and $TN$ true negative, that is, the correctly classified normal oral images.
4.3 Results and analysis

As is demonstrated in Table 2, the classification accuracy of AlexNet and GoogleNet models is 89.16% and 95.38%, respectively, while the classification accuracy of the proposed method is 98.79%. The results depict that the proposed method has higher accuracy in the diagnosis and classification of oral ulcers compared with the other two models. Moreover, the sensitivity and specificity of our method are both higher than two competitors.

5 CONCLUSION

In this work, we propose a novel deep learning framework for real-time diagnosis and classification of oral ulcers. In particular, pre-trained model parameters are transferred to the proposed model to learn basic low-level features, which speed up the problem-solving process and improve the classification accuracy. In addition, we use the image enhancement and preprocessing techniques to increase the dataset and reduce overfitting of the model. In the experiment, the proposed method is compared with the classic CNN models, and the experimental results show that the classification accuracy of our proposal is 98.79%, the specificity is 99.27%, and the sensitivity is 98.24%. Our approach outperforms the existing CNN image classification methods and has a good classification accuracy.

Although the proposed method is effective for diagnosing the oral ulcers, it requires a larger dataset for training and validation before clinical application. In the future, we are interested in investigating the diagnostic and classification methods for ulcers of different degrees in oral cavity.

ACKNOWLEDGEMENTS

This work is supported by the Fujian Province Undergraduate Universities Teaching Reform Research Project (No. FBJG20190156), the Intelligent Computing and Applied Scientific Research Innovation Team of Concord University College Fujian Normal University in 2020 (No. 2020-TD-001), the Fujian Provincial Natural Science Foundation (No. 2020J01646), and Young Teacher Education and Research Fund of Fujian Provincial Education Department (No. JT180208).

REFERENCES

1. Goodfellow, I., et al.: Deep Learning, vol. 1. MIT Press, Cambridge (2016)
2. Xue, X.: A Compact Firefly Algorithm for Matching Biomedical Ontologies. 62, pp. 2855–2871 (2020)
3. Xue, X., Chen, J.: Matching biomedical ontologies through compact differential evolution algorithm with compact adaption schemes on control parameters. 1–27 (2020)
4. Xue, X., Chen, J.: Optimizing sensor ontology alignment through compact co-firefly algorithm. Sensors 20(7), 2056 (2020)
5. Lin, J.C.W., et al.: ASRNN: A recurrent neural network with an attention model for sequence labeling. Knowl.-Based Syst. 212, 106548 (2020)
6. Razzak, M.I., et al.: Classification in BioApps, pp. 323–350. Springer, Cham (2018)
7. Xue, X., Chen, J.: Using Compact Evolutionary Tabu Search algorithm for matching sensor ontologies. Swarm Evol. Comput. 48, 25 (2019)
8. Esteva, A.: A guide to deep learning in healthcare. Nat. Med. 25(1), 24–29 (2019)
9. Lin, J.C.W., et al.: Privacy preserving multi-objective sanitization model in 6G IoT environments. IEEE Internet Things J. 1–10 (2020)
10. Li, Q., et al.: Medical image classification with convolutional neural network. In: 2014 13th International Conference on Control Automation Robotics & Vision (ICARCV), pp. 844–848. IEEE, Piscataway (2014)
11. Zhang, Q., et al.: A GPU-based residual network for medical image classification in smart medicine. Inf. Sci. 91–100 (2020)
12. Xue, X., Wang, Y.: Using memetic algorithm for instance correspondence resolution. IEEE Trans. Knowl. Data Eng. 28(2), 580–591 (2015)
13. Xue, X., Liu, J.: Collaborative ontology matching based on compact interactive evolutionary algorithm. Knowl.-Based Syst. 137, 94–103 (2017)
14. He, K., et al.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778. IEEE, Piscataway (2016)
15. Lijrens, G., et al.: A survey on deep learning in medical image analysis. Med. Image Anal. 42, 60–88 (2017)
16. Xue, X., Yao, X.: Interactive ontology matching based on partial reference alignment. Appl. Soft Comp. 72, 355–370 (2018)
17. Hazlett, H.C., et al.: Early brain development in infants at high risk for autism spectrum disorder. Nature 542(7641), 348–351 (2017)
18. Kawahara, J., Hamarneh, G.: Multi-resolution-Tract CNN with Hybrid Pretrained and Skin-Lesion Trained Layers. In: International Workshop on Machine Learning in Medical Imaging, pp. 164–171. Springer, Cham (2016)
19. Hosseini-Asl, E., et al.: Alzheimer’s disease diagnostics by adaptation of 3D convolutional network. In: 2016 IEEE International Conference on Image Processing (ICIP), pp. 126–130. IEEE, Piscataway (2016)
20. Bakkouri, I., et al.: Recognition of Alzheimer's disease on sMRI based on 3D multi-scale CNN features and a gated recurrent fusion unit. In: 2019 International Conference on Content-Based Multimedia Indexing (CBMI), pp. 1–6. IEEE, Piscataway (2019)
21. Huang, X., et al.: Lung nodule detection in CT using 3D convolutional neural networks. In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), pp. 379–383. IEEE, Piscataway (2017)
22. Shen, W., et al.: Multi-scale convolutional neural networks for lung nodule classification. In: International Conference on Information Processing in Medical Imaging, pp. 588–599. Springer, Cham (2015)
23. Dorj, U.O., et al.: The skin cancer classification using deep convolutional neural network. Multimedia Tools Appl. 77(8), 9909 (2018)
24. Khan, S., et al.: A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. Pattern Recognit. Lett. 125, 1 (2019)
25. Gour, M., et al.: Residual learning based CNN for breast cancer histopathological image classification. Int. J. Imaging Syst. Technol. 621–635 (2020)

TABLE 2 Evaluation result comparison by AlexNet, GoogleNet and Proposed Method

| Classifier | Accuracy | Sensitivity | Specificity |
|------------|----------|-------------|-------------|
| AlexNet    | 89.16%   | 91.22%      | 84.61%      |
| GoogleNet  | 95.38%   | 94.91%      | 96.22%      |
| Proposed   | 98.79%   | 98.24%      | 99.27%      |

ORCID

Haolin Wang https://orcid.org/0000-0002-0477-4973
Xingsi Xue https://orcid.org/0000-0002-3008-8782
26. Aubreville, M., et al.: Automatic classification of cancerous tissue in laserendomicroscopy images of the oral cavity using deep learning. Sci. Rep. 7(1), 11979 (2017)
27. Shamim, M.Z.M., et al.: arXiv preprint, arXiv:1909.08987, 2019
28. Folmsbee, J., et al.: 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 (ISBI 2018), pp. 770–773. IEEE, Piscataway (2018)
29. Fu, Q., et al.: A deep learning algorithm for detection of oral cavity squamous cell carcinoma from photographic images: A retrospective study. E. Clinical Medicine 27, 100558 (2020)
30. Wong, S.C., et al.: Understanding data augmentation for classification: When to warp? In: 2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA), pp. 1–6. IEEE, Piscataway (2016)
31. Lin, J.C.W., et al.: Incrementally updating the high average-utility patterns with pre-large concept. Appl. Intell. 50(11), 3788 (2020)
32. Tan, C., et al.: A survey on deep transfer learning. In: International Conference on Artificial Neural Networks, pp. 270–279. Springer, Cham (2018)
33. Krizhevsky, A., et al.: Imagenet classification with deep convolutional neural networks. Adv. Neural Inform. Process. Syst. 25, 1097 (2012)
34. Szegedy, C., et al.: Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9. IEEE, Piscataway (2015)

How to cite this article: Guo J, Wang H, Xue X, Li M, Ma Z. Real-time classification on oral ulcer images with residual network and image enhancement. IET Image Process. 2021;1–6. https://doi.org/10.1049/ipr.12144