

Tongue fissure visualization by using deep learning – an example of the application of artificial intelligence in traditional medicine

Wen-Hsien Chang

China Medical University

Han-Kuei Wu

China Medical University

Lun-chien Lo

China Medical University

William W. L. Hsiao

British Columbia Public Health Microbiology & Reference Laboratory

Hsueh-Ting Chu (✉ prof.htchu@gmail.com)

Asia Univerity <https://orcid.org/0000-0003-0618-5598>

Hen-Hong Chang

China Medical University

Research article

Keywords: Chinese medicine, tongue diagnosis, AI, deep learning, class activation mapping

Posted Date: April 7th, 2020

DOI: <https://doi.org/10.21203/rs.2.19210/v3>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background Traditional Chinese medicine (TCM) describes physiological and pathological changes inside and outside the human body by the application of four methods of diagnosis. One of the four methods, tongue diagnosis, is widely used by TCM physicians, since it allows direct observations that prevent discrepancies in the patient's history and, as such, provides clinically important, objective evidence. The clinical significance of tongue features has been explored in both TCM and modern medicine. However, TCM physicians may have different interpretations of the features displayed by the same tongue, and therefore intra- and inter-observer agreements are relatively low. If an automated interpretation system could be developed, more consistent results could be obtained, and learning could also be more efficient. This study will apply a recently developed deep learning method to the classification of tongue features, and indicate the regions where the features are located.

Methods A large number of tongue photographs with labeled fissures were used. Transfer learning was conducted using the ImageNet-pretrained ResNet50 model to determine whether tongue fissures were identified on a tongue photograph. Often, the neural network model lacks interpretability, and users cannot understand how the model determines the presence of tongue fissures. Therefore, Gradient-weighted Class Activation Mapping (Grad-CAM) was also applied to directly mark the tongue features on the tongue image.

Results Only 6 epochs were trained in this study and no graphics processing units (GPUs) were used. It took less than 4 minutes for each epoch to be trained. The correct rate for the test set was approximately 70%. After the model training was completed, Grad-CAM was applied to localize tongue fissures in each image. The neural network model not only determined whether tongue fissures existed, but also allowed users to learn about the tongue fissure regions.

Conclusions This study demonstrated how to apply transfer learning using the ImageNet-pretrained ResNet50 model for the identification and localization of tongue fissures and regions. The neural network model built in this study provided interpretability and intuitiveness, (often lacking in general neural network models), and improved the feasibility for clinical application.

Background

TCM (Traditional Chinese Medicine) physicians learn about the status of internal and external organs, meridians, and blood-Qi circulation in the human body, infer physiological and pathological changes, and choose appropriate treatments through the application of four methods of diagnosis: inspection, listening and smelling examination, inquiry, and palpation. Tongue examination is part of the inspection diagnosis, since health status and disease courses are often highly correlated with the condition of the tongue. TCM physicians directly observe the tongue to corroborate with patient's self-reported medical history. It is evident that tongue diagnosis provides clinically important objective evidence, and therefore is widely used by TCM physicians.

Many tongue features are clinically examined in TCM, including tongue fissures, tooth marks, thin and thick furs, etc., as shown in Fig. 1. The clinical significance of these features can be interpreted from the perspective of both TCM and modern medicine. From the perspective of TCM, tongue fissures indicate

excessive “heat” or inadequate body fluid in the human body, and therefore, many studies in modern medicine have focused on tongue fissures. Ching *et al.* found that patients with burning mouth syndrome were more prone to tongue fissures than the average person [1]. Feil *et al.* found that the occurrence of tongue fissures is directly related to age, gender (more men have tongue fissures than women), and burning mouth syndrome [2]. Dudko *et al.* found that of the 104 patients with tongue fissures or geographic tongue, 70% had mold detected in the mouth, 35% exhibited idiopathic pain and burning, and 10% had dry mouth [3]. This is consistent with the concept of heat in TCM theory as previously discussed. Sjögren’s syndrome is a disease which affects moisture-producing glands in the human body, and usually results in dry mouth. Soto-Rojas *et al.* found that 70% of patients with Sjögren’s syndrome had red tongue and fissures [4], which is consistent with the aforementioned concept of “inadequate body fluid” in TCM. Sudarshan *et al.* designed a detailed classification method based on the direction, location, and number of tongue fissures, and whether burning mouth syndrome is present. This method may be used as a reference for disease assessment in the future [5]. Based on the appearance of psoriasis patient’s tongues, Daneshpazhooh *et al.* found that 66% of those patients had fissures [6], while Zargari found that only 8.2% had fissures [7], and Qahtani *et al.* found that only 4% had fissures [8]. The large differences found in these reports suggest that the determination of tongue fissures may be subjectively, thus different observers report different results.

TCM physicians may have different interpretations for the features on the same tongue, and, in addition, the same TCM physician may have different interpretations of the same tongue at different times. Therefore, inter-observer agreement on tongue features is relatively low [9, 10]. It is likely that if an automated interpretation system could be developed, more consistent results would be obtained, and human error would be reduced. In addition, junior physicians and students could also learn about tongue diagnosis more efficiently. Hence, the objective of this study is to develop a computerized interpretation system to supplement the human diagnostic process.

Methods

Classification models

This study describes the application of Grad-CAM to ResNet in order to localize tongue fissures. The system framework is shown in Fig. 2. ResNet is a neural network model that won first prize in the image classification competition during the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2015. ResNet’s error rate is only 3.57%, which is lower than the reported error rate of 5.1% made by a human expert [14]. Grad-CAM provides a visual interpretation to a neural network model. [16]

Training dataset and training process

This study screened 489 de-identified tongue images from the Department of Chinese Medicine of a medical center in Taiwan, all of which were interpreted by a TCM physician with over 30 years of experience. After interpretation, all the images were divided into two groups: Group F comprising 312 images with tongue fissures, and Group N comprising 177 images without tongue fissures. Horizontal

flipping was applied to all the images to achieve data augmentation. Then, all the images were randomly divided into a training group (80%) and a test group (20%). The neural network did not have any information about the location of tongue fissures. In this study, transfer learning was performed using the ImageNet-pretrained ResNet50 model. Only the last layer of the model was replaced by a binary classifier, and was retrained so that the other layers remained intact without any modification. After training was completed, Grad-CAM was added to the model to localize tongue fissures.

The deep neural network model is easily misled by adversarial images. However, when the model is unable to determine the ground truth class of an adversarial image, Grad-CAM is still able to localize and correctly classify the relevant regions. Based on these observations, Grad-CAM was applied in this study to localize tongue fissures in all images, regardless of whether the model recognized the presence of tongue fissures in the images or not.

Results

Only 6 epochs were trained in this study and no GPU (graphics processing unit) was used. It took less than 4 minutes for each epoch to be trained. The correct rate for the test set was approximately 70%. After the model training was completed, Grad-CAM was applied in order to localize tongue fissures in each image.

Additional experiments assured our proposed method was robust and repeatable. We ran 100-200 epochs for several optimizers such as stochastic gradient descent (SGD), Adam and Root mean square propagation (RMSprop). Batch size of 128, 256, 512 and 1024 and learning rates of 0.05, 0.005, 0.001 and 0.0001 have been tried. After observing several tests we chose a batch size of 128 to train for 150 epochs. With different shuffling of the data, each optimizer was used to train for 20 times. Better results are as the following:

With SGD, loss of 0.48 and accuracy of 0.74 were achieved in the 45th epoch. With RMSprop, loss of 0.30 and accuracy of 0.88 were achieved in the 34th epoch. With Adam, loss of 0.51 and accuracy of 0.78 were achieved in the 7th epoch. Grad-CAM was applied to these three models to localize tongue fissure.

Tongue fissures localized by our network model

As previously discussed, Grad-CAM was still able to localize tongue fissure regions in images in Group F, namely those regarded by the neural network as having no tongue fissure. For instance, Fig. 3 shows the results of localization of tongue fissures that are closer to their actual positions.

Other fissures localized by our neural network

In addition to tongue fissures, fissures on the face may also be localized by Grad-CAM. For example, some nasolabial folds, prejowl sulci and the philtrum would also be marked as shown in Fig. 4. Although these fissures were not the target of this study (*i.e.* tongue fissures), it can be assumed from these results that this neural model has accurately learned the pattern of a fissure.

Not all tongue fissures were covered in the localized regions in some images. In addition, some other features, such as lip wrinkles and grooves between the upper lip and the tongue, similar to fissures but not actually fissures, were also marked as shown in Fig. 5.

Discussion

In the past few decades, automated interpretation of the tongue was performed through conventional feature extraction algorithms and statistical methods. L.C. Lo et al. and Hsu, Y.C. et. al. used conventional image processing techniques to detect tongue features, and located corresponding regions. However, these studies did not provide assessment methods and results in detail [10, 11]. In recent years, artificial intelligence (AI) has been actively applied to medical technology, and significant progress has been made with deep learning in image processing, thereby eliminating the need for image processing experts to extract image features manually [12]. Furthermore, transfer learning, a deep learning model that is pretrained using big data sets, can often be easily applied to different big data sets to interpret image categories. In this study, a pretrained model was applied to the classification of tongue features, and tongue features were directly marked on tongue images.

Some studies have applied deep learning to the analysis of tongue images, but deep learning is yet to be applied to the clinical interpretation of tongue diagnoses in TCM. For example, Meng et al. designed the CHDNet model, which combined deep learning and support vector machine classifiers to extract and classify tongue features [13]. However, the digital features extracted by this model did not visualize the tongue features mentioned in TCM. As a consequence, these digital features could not be applied to clinical inspection diagnosis. Additionally, the classification results showed either “gastritis” or “no gastritis”, which were not related to either the name of a disease or diagnosis in TCM. Hou et al. analyzed tongue color using deep learning, outperforming the conventional methods [14]. The present study applied deep learning visualization techniques to tongue diagnosis, and used specific tongue features as reported by TCM physicians as an example to determine whether those features existed, and to locate the region where they were distributed. We used a well-known deep learning model named ResNet50 to discover and localize tongue fissures. [17] To our knowledge, this is the only study that has applied deep learning visualization techniques to tongue diagnosis in TCM.

Deep learning has been widely used in image classification. Using tongue fissures as an example, a deep neural network model—after the training is completed—can determine whether any tongue image contains tongue fissures, but it is unable to provide a clear basis for users to understand its logic. Essentially, the model functions as a black box that does not allow intuitive interpretation. For example, if the model identifies a certain tongue image as having a tongue fissure, but the user visually interprets it as having no tongue fissure, then the user cannot understand why the model arrived at that interpretation. An explanation might be that perhaps the tongue fissures identified by the model are vague, and cannot be easily spotted through manual inspection. On the other hand, if an incorrect judgment is made by the model, it would be very difficult for the user and the engineer to correct the model. Therefore, many studies have attempted to increase the interpretability of deep neural networks through various types of

algorithms, such as visualization. For instance, Zhou et al. proposed Class Activation Mapping (CAM), which is able to locate class-specific regions in images [15]. However, CAM requires a neural network to satisfy specific requirements, i.e., a fully-convolutional neural network followed by a global average pooling layer and then a linear prediction layer. Hence, the network model often requires modifications. In another study, Selvaraju et al. proposed Gradient-weighted Class Activation Mapping (Grad-CAM), which is a generalization of CAM that may be easily applied to any existing neural network model without having to modify and train it [16]. Grad-CAM uses the gradient information of the last convolutional layer to differentiate the importance of each neuron and heat maps to show the degree of correlation between each region and class-specific regions. For example, red represents high specificity, while blue represents low specificity.

There were some limitations in this study. If tongue segmentation can be performed on tongue images first, where only the tongue in the image is retained, fissures outside the tongue will not interfere with the learning process of the neural network, and the localization of tongue fissures should be more accurate. However, the quality of localization cannot be accurately assessed because large numbers of tongue fissure images which have been recognized by the academic community in TCM, or on which consensus has been reached and marked, are currently not available as ground truths. As a result, inter-observer agreement is not high; hence, it is not easy to obtain consistent ground truths [9]. But automatic tongue fissure localization still can be used as a screen tool in a medical environment without experienced TCM.

Conclusions

This study demonstrated how to quickly complete transfer learning by application of the ImageNet-pretrained ResNet50 model to identify tongue fissures, and how to locate tongue fissure regions using Grad-CAM based on this network model. The results of this study show that our approach is feasible. In future, other deep neural networks may be applied and fine-tuned to obtain better results. We hope that there will be more AI (artificial intelligence) applications developed that are related to tongue diagnosis in TCM, so that this traditional technique, which is so rich with expert experience, can be passed down and more widely used.

Abbreviations

TCM: Traditional Chinese medicine

AI: artificial intelligence

CAM: class activation mapping

Grad-CAM: gradient-weighted class activation mapping

GPU: graphics processing unit

Declarations

Ethics approval and consent to participate

This study was approved by the Institutional Review Board and Ethics committee of China Medical University Hospital. All the study participants gave informed consent and signed a partnership agreement.

Consent for publication

Not applicable.

Availability of data and materials

The results and data sets used in this study are available at:

<https://github.com/htchu/DeepTongueVis>.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was financially supported by Asia University (ASIA-108-CMUH-22), China Medical University Hospital (DMR-107-166, DMR-108-178), the “Chinese Medicine Research Center, China Medical University” from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education(MOE) in Taiwan (CMRC-CMA-6), Taiwan Ministry of Health and Welfare Clinical Trial and Research Center of Excellence (MOHW105-TDU-B-212-133019), and National Research Program for Biopharmaceuticals from National Science Council, Taiwan (DOH96-TD-I-111-006). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' Contributions

WHC did the analysis on the tongue data. HHC interpreted the tongue images and discussed with LCL and HKW. WHC, HKW, LCL, WWLH, HTC and HHC discussed the project and jointly wrote the manuscript. HTC and HHC leded the project. All authors read and approved the final manuscript.

Corresponding Author

Correspondence to Hsueh-Ting Chu or Hen-Hong Chang

Acknowledgments

Not applicable.

References

1. Ching, V., M. Grushka, M. Darling, and N. Su, *Increased prevalence of geographic tongue in burning mouth complaints: a retrospective study*. *Oral Surg Oral Med Oral Pathol Oral Radiol*, 2012. **114**(4): p. 444-8.
2. Feil, N.D. and A. Filippi, *Frequency of fissured tongue (lingua plicata) as a function of age*. *Swiss Dent J.*, 2016. **126**(10): p. 886-7.
3. Dudko, A., A.J. Kurnatowska, and P. Kurnatowski, *Prevalence of fungi in cases of geographical and fissured tongue*. *Ann Parasitol*, 2013. **59**(3): p. 113-7.
4. Soto-Rojas, A.E., J. Villa Ar Fau - Sifuentes-Osornio, D. Sifuentes-Osornio J Fau - Alarcon-Segovia, A. Alarcon-Segovia D Fau - Kraus, and A. Kraus, *Oral manifestations in patients with Sjogren's syndrome*. *J Rheumatol*, 1998 May. **25**(5): p. 906-10.
5. Sudarshan, R., G. Sree Vijayabala, Y. Samata, and A. Ravikiran, *Newer Classification System for Fissured Tongue: An Epidemiological Approach*. *J Trop Med*, 2015. **2015**: p. 262079.
6. Daneshpazhooh, M., H. Moslehi, M. Akhyani, and M. Etesami, *Tongue lesions in psoriasis: a controlled study*. *BMC Dermatol*, 2004. **4**(1): p. 16.
7. Zargari, O., *The prevalence and significance of fissured tongue and geographical tongue in psoriatic patients*. *Clin Exp Dermatol*, 2006. **31**(2): p. 192-5.
8. Al Qahtani, N.A., A. Deepthi, N.M. Alhussain, B.A.M. Al Shahrani, H. Alshehri, A. Alhefzi, and B. Joseph, *Association of geographic tongue and fissured tongue with ABO blood group among adult psoriasis patients: a novel study from a tertiary care hospital in Saudi Arabia*. *Oral Surg Oral Med Oral Pathol Oral Radiol*, 2019. **127**(6): p. 490-497.
9. Lo, L.C., Y.F. Chen, W.J. Chen, T.L. Cheng, and J.Y. Chiang, *The Study on the Agreement between Automatic Tongue Diagnosis System and Traditional Chinese Medicine Practitioners*. *Evid Based Complement Alternat Med*, 2012. **2012**: p. 505063.
10. Hsu, Y., Y. Chen, L. Lo, and J.Y. Chiang. *Automatic tongue feature extraction*. in *2010 International Computer Symposium (ICS2010)*. 2010.
11. Schmidhuber, J., *Deep learning in neural networks: an overview*. *Neural Netw*, 2015. **61**: p. 85-117.
12. Meng, D., G. Cao, Y. Duan, M. Zhu, L. Tu, D. Xu, and J. Xu, *Tongue Images Classification Based on Constrained High Dispersal Network*. *Evid Based Complement Alternat Med*, 2017. **2017**: p. 7452427.
13. Hou, J., H. Su, B. Yan, H. Zheng, Z. Sun, and X. Cai. *Classification of tongue color based on CNN*. in *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)*(2017).
14. He, K., X. Zhang, S. Ren, and J. Sun. *Deep Residual Learning for Image Recognition*. in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016.
15. Zhou, B., A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba. *Learning Deep Features for Discriminative Localization*. in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016.
16. Selvaraju, R.R., M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, *Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization*. 2017 IEEE International

Figures

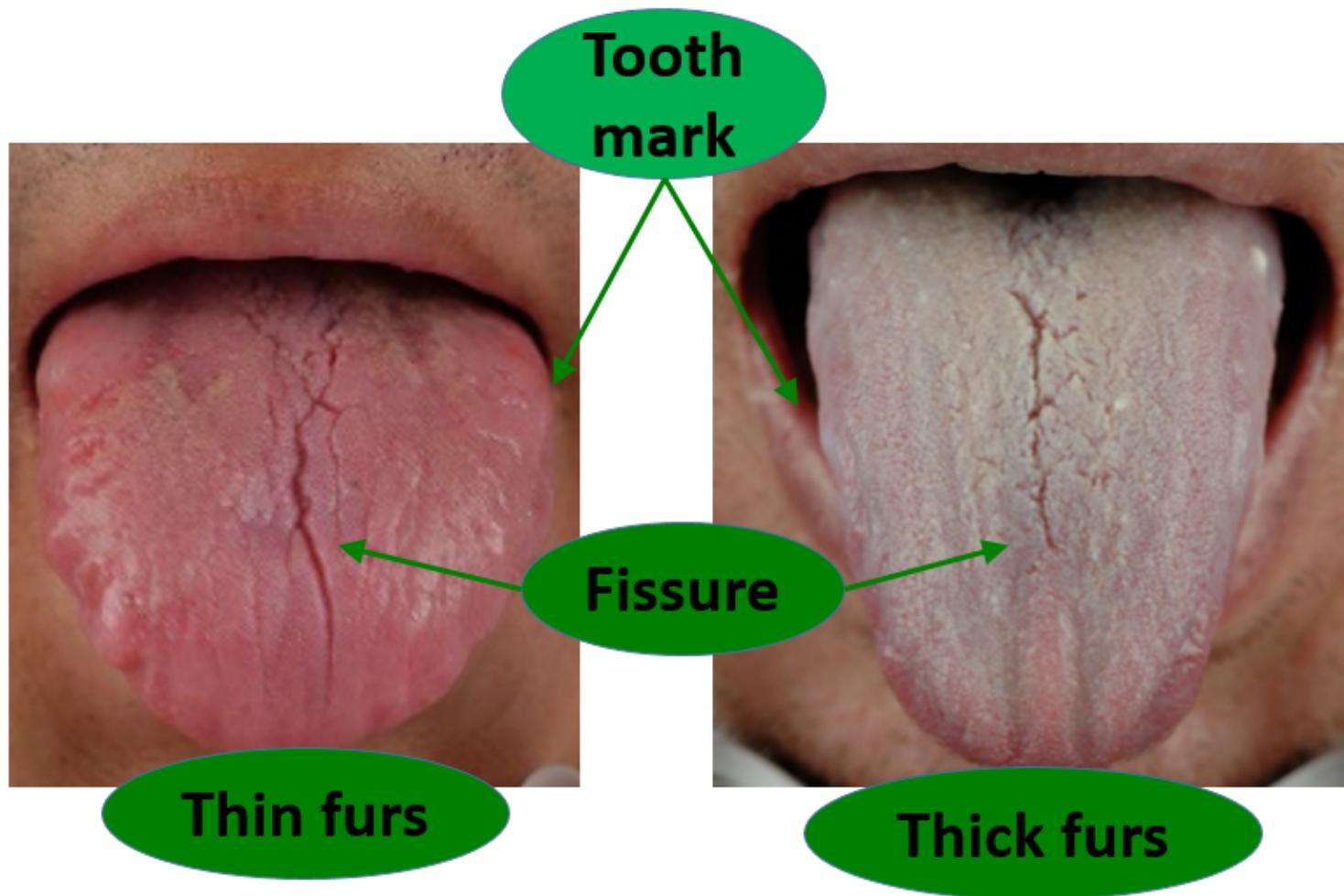


Figure 1

Examples of tongue features. Both tongues have tooth marks on the edges and fissures in the middle. The tongue on the left has thin furs, while the one on the right has thick furs.

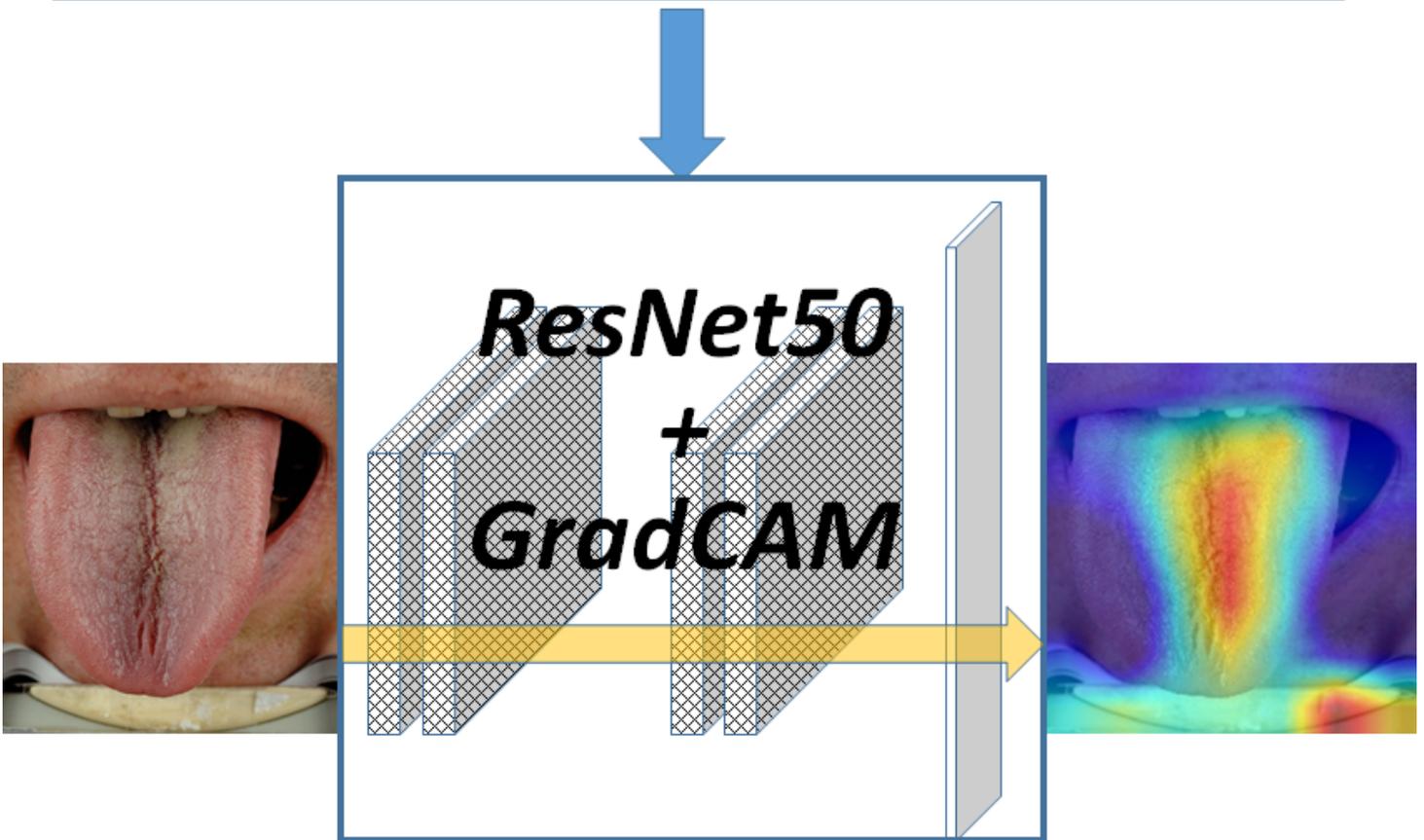
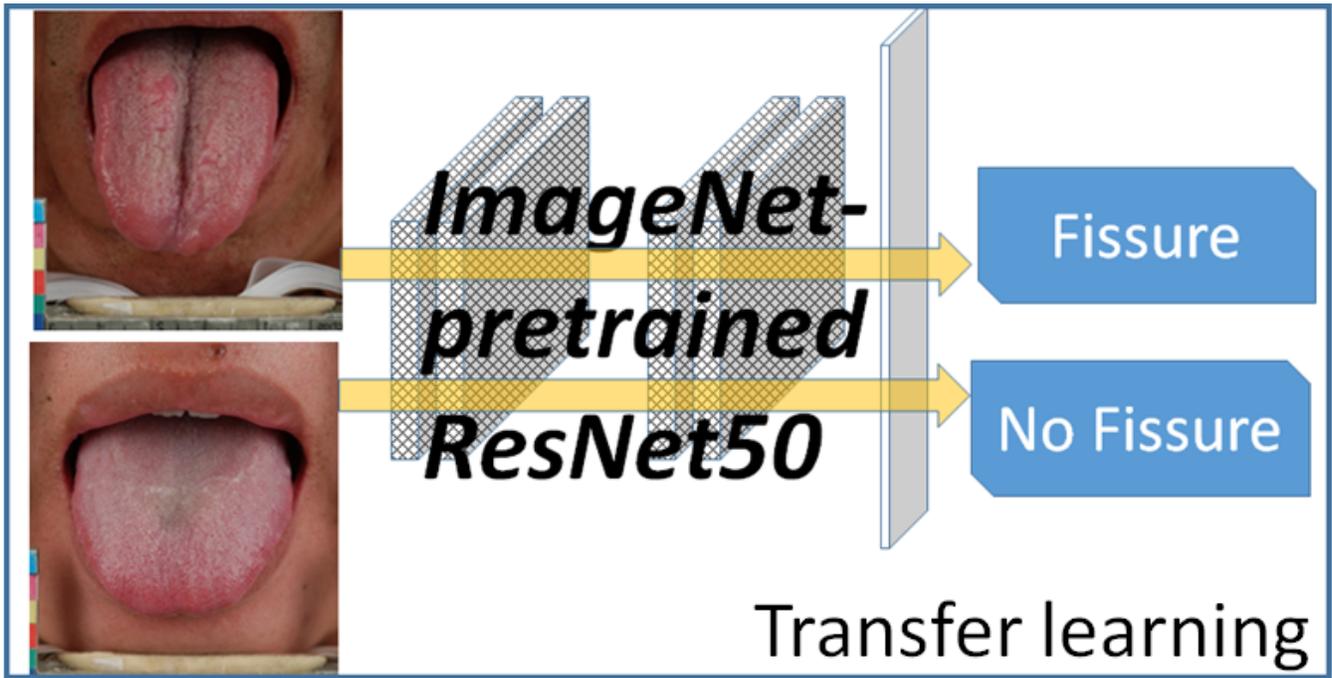


Figure 2

System architecture for tongue fissure visualization

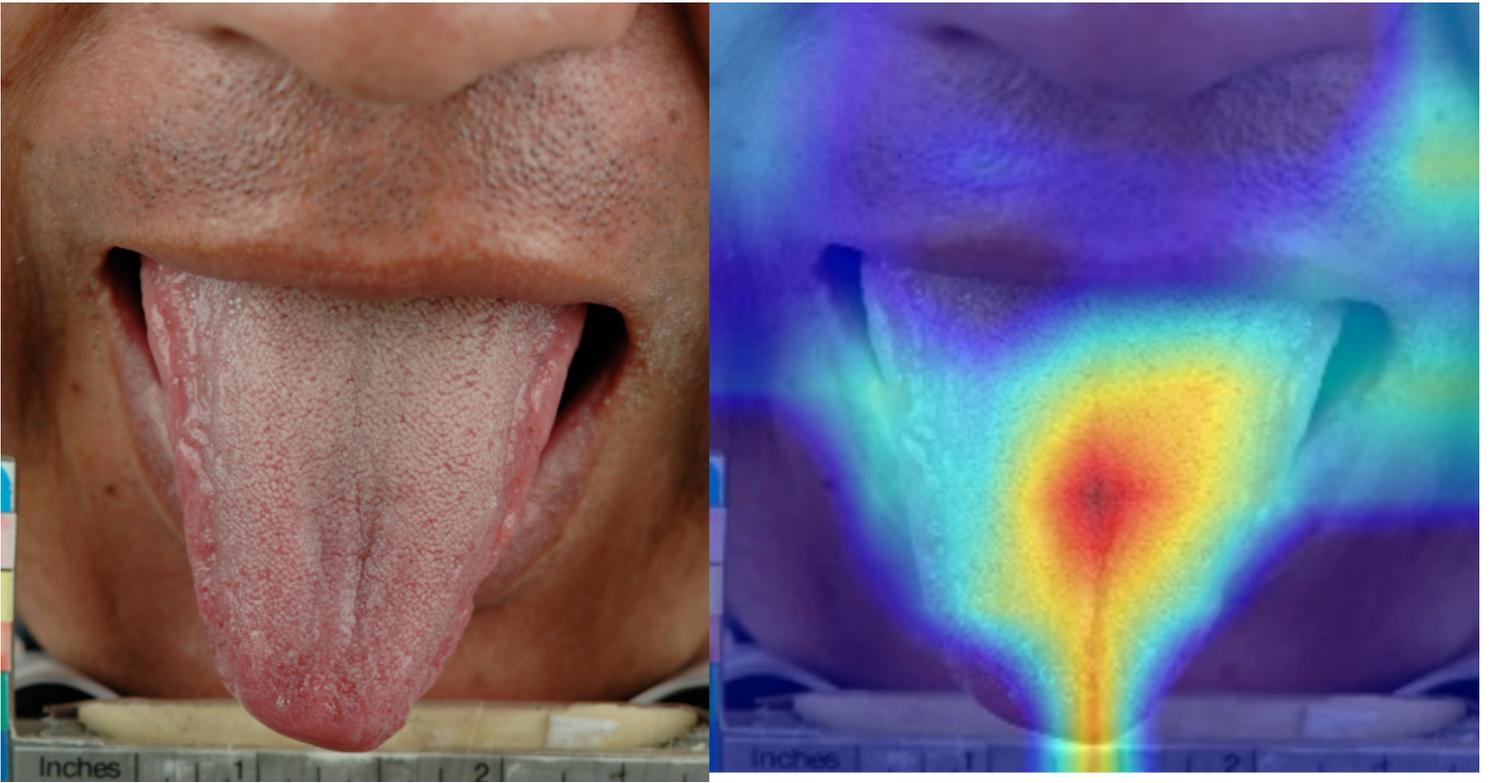


Figure 3

Tongue fissures localized by our neural network

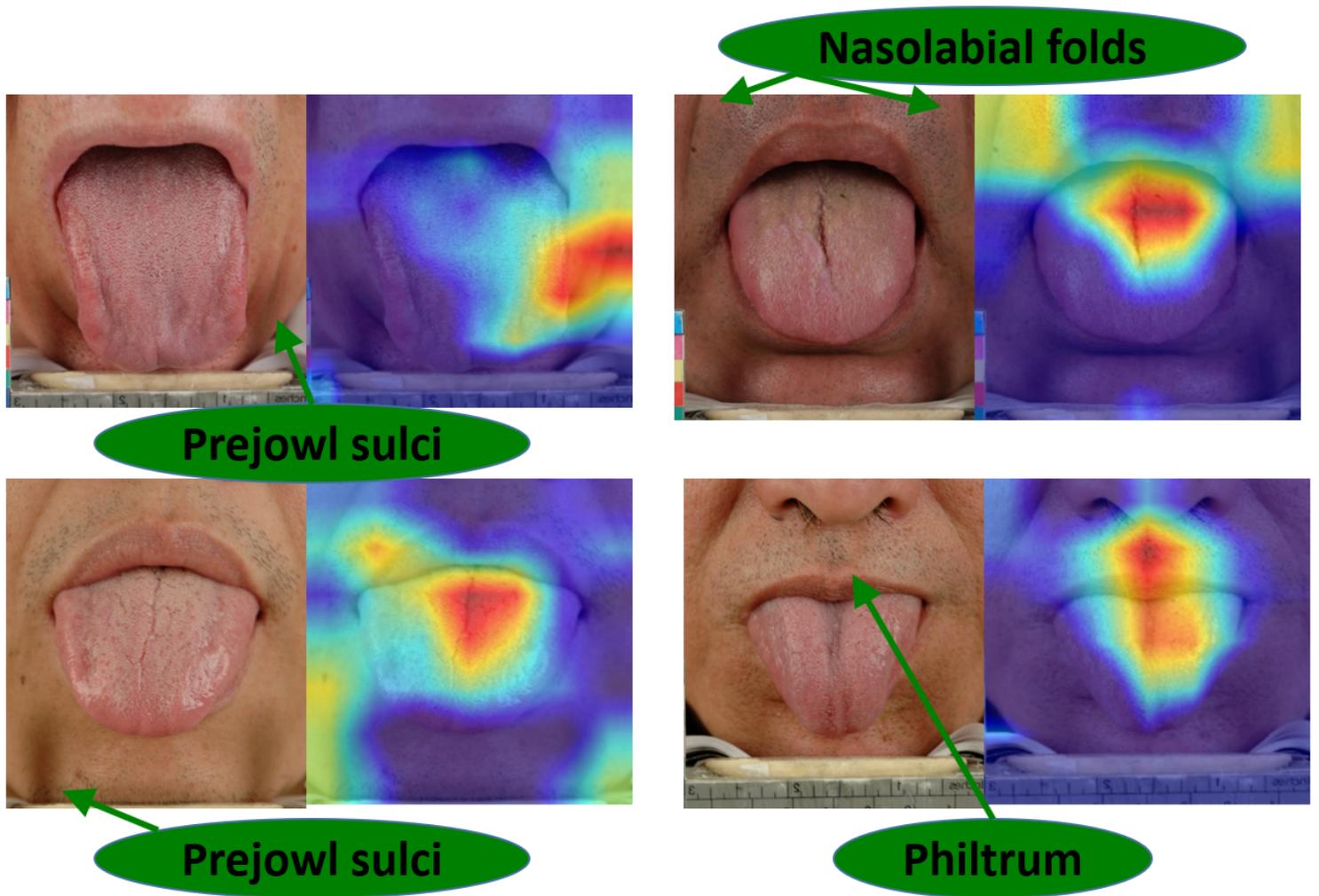
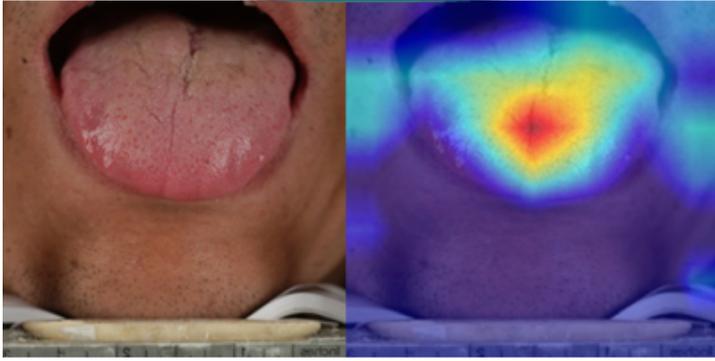


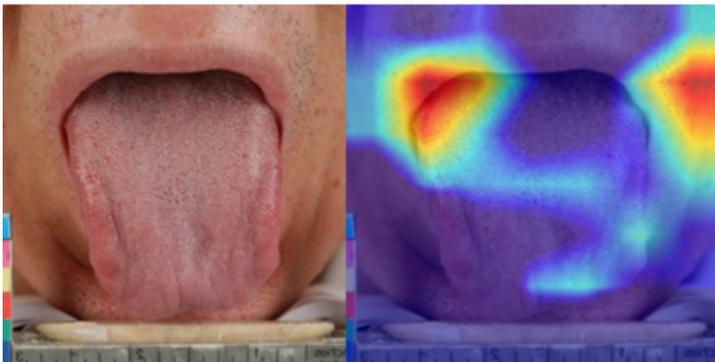
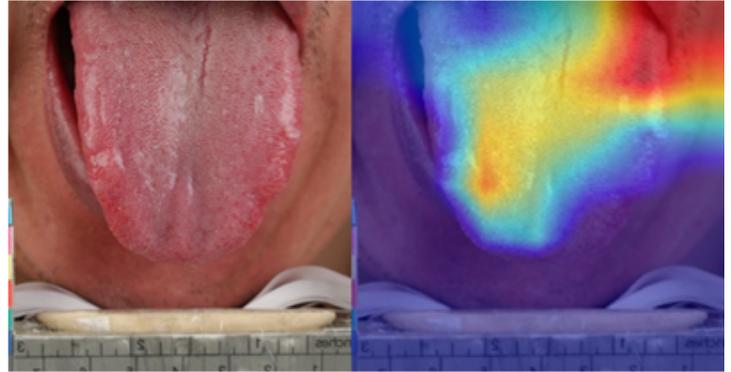
Figure 4

Other identified fissures. Nasolabial folds, prejowl sulci and the philtrum are also localized by our model.

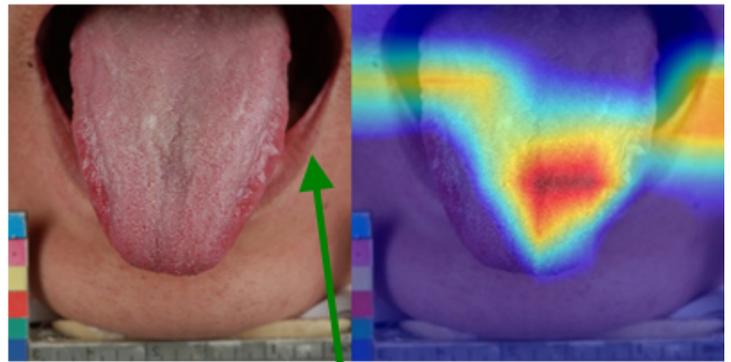
Worse localization



Gap



Gap



Lip fissure

Figure 5

Incorrect tongue fissures. Grooves between the upper lip and the tongue, as well as lip wrinkles were mistaken by our model for tongue fissures.