

TONGUE IMAGE CLASSIFICATION USING RANDOM FOREST CLASSIFIER

R. Charanya¹, J. Josphin Mary², G. Sridevi³, V. Shanthi⁴

^{1,2,3}Assistant Professor, Faculties of Humanities Science, Meenakshi Academy of Higher Education Research Chennai, TamilNadu, India

⁴Professor, Faculties of Humanities Science, Meenakshi Academy of Higher Education Research Chennai, TamilNadu, India

E-mail: charanyar@maherfhs.ac.in¹, josphinmaryj@maherfhs.ac.in², sridevig@maherfhs.ac.in³, principal@maherfhs.ac.in⁴

Abstract:

Diabetes people who take antibiotics regularly to treat multiple infections are more likely to develop a fungal mouth and language infection. The fungus thrives in people with uncontrolled diabetes at high levels of glucose in the saliva. A yeast infection called oral thrush is common among people with diabetes. It looks like a white layer coating your language and your cheeks' insides. The Yeast grows in a higher amount of sugar found in your saliva. The early diagnosis is required for tongue image classification. In this study, the automatic classification of tongue image classification for diabetes detection system is discussed. Initially, the input tongue images are given to median filter for pre-processing. Then the Gray Level Co-occurrence Matrix (GLCM) and Haralick features are extracted. Finally, Random Forest (RF) classifier is used for Prediction. The performance of proposed system produces the classification accuracy of 95% using RF classifier.

Keywords: *Tongue image classification, Gray level co-occurrence matrix, haralick features, Random Forest classifier, Diabetes detection*

INTRODUCTION

Creation of a language learning program for mobile phones [1]. While a range of methods have over the years been developed for the collection, retrieval and analysis of photographs, most have limitations in reliability, precision, cost and usability. Analysis of a visual language image using the ColorChecker language in medical applications [2]. This ColorChecker acts as a color guide and can be used to standardize the captured digital tongue images in color calibration algorithms. At first, a statistical tongue color gamut based on a broad tongue image dataset is created. Different amounts of color patches are calculated by experimenting on the ColorChecker tongue.

An assessment tool of the consistency of tongue images in Chinese traditional medicine [3]. In the form of a 17-dimensional feature vector for the evaluation of language content, geometric features, texture features or spectral entropy features, are extracted respectively based

on the spatial spectral entropy-dependent quality index for tongue images. Tongue statistical analysis function extraction and diagnostics in [4]. Based on a wide database conveying over 9000 language images from a colorimetric digital camera specially built non-contact imaging system, a single class SVM algorithm is used to describe the color gamut in the colour scheme.

Color decomposition and thresholding of language image segmentation [5]. In order to define the region of gap between the tongue body root and the upper lip, a picture threshold on the RGB color model is then carried out. Finally, the initial tongue region is optimized with the aid of morphological operations in the above-mentioned field by eliminating fake object regions like the upper lip to achieve the final result. Segmentation of the language picture by clusters and thresholds [6-7]. Eventually, by extracting fake tongue body regions, including the high lip, with the aid of the gap region, the tongues are removed from the initial object region

Tongue image classification is presented in this study. The rest of the paper is organized as follows: The methods and materials used for proposed system is discussed in section 2. The experimental results and discussion are explained in section 3. The last section concludes the tongue image classification system.

Methods and Materials

Initially, the input tongue images are given to median filter for pre-processing. Then haralick and GLCM features are used for feature extraction. Finally, RF classifier is used for prediction. Figure 1 shows the processing of tongue image classification system.

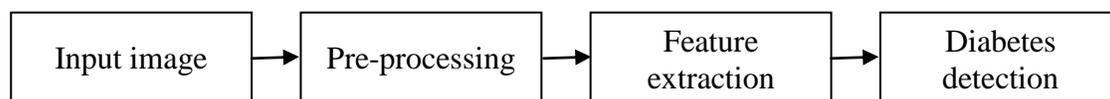


Figure 1 proposed tongue image classification system processing

Median filter-preprocessing:

Median filtering is important for the processing of digital images because it preserves edges and noise removal under certain conditions, and also has signal processing applications. The median filter is a non-optical filter system sometimes used to eliminate image or signal noise [8-9]. This reduction of noise is a conventional pre-processing method aimed at increasing subsequent processing.

Tongue image classification based GLCM feature extraction:

The GLCM is a square matrix having the size of the gray level N in the area of interest. The GLCM functions define the texture of a picture by calculating the number of pixels in a picture, creating GLCM, and then removing statistically determined measures from that matrix in a given spatial relationship [10-11]. The texture features of Haralick are calculated using a GLCM, a matrix that shows the co-occurrence of adjacent gray levels in the image.

RF based Prediction:

A random forest is a meta-estimator that matches a number of decision-tab classificatory for different data samples and uses averages to boost prediction precision and overfit power. The

algorithm of the random forest comprises of different decision-making trees. When each tree is built, it uses bagging and randomness to try to create a non-relating tree forest whose prediction by comitology is more accurate.

Results and Discussion

The input tongue images are given to median filter for pre-processing, then the GLCM and haralick features are used for feature extraction. The sample tongue images and normal images are shown in figure 2.

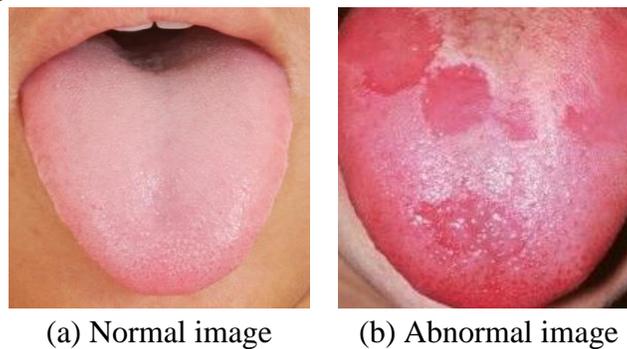


Figure 2 Sample tongue images

The prediction is made by RF classifier. The classification accuracy, sensitivity and specificity are shown in figure 3.

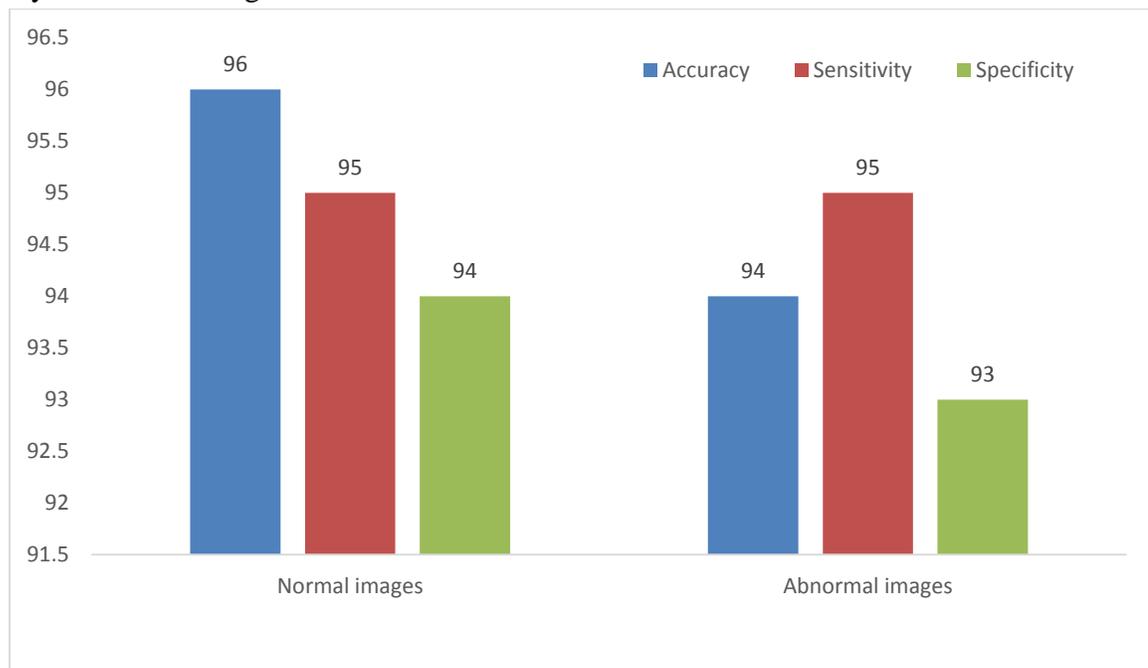


Figure 3 Graphical representation of RF prediction

From the above figure, it is observed that the overall classification accuracy is 96%. The classification accuracy of normal images is 95% and its sensitivity and specificity are 94% and 95%. The classification accuracy of abnormal images is 94% and its sensitivity and specificity are 95% and 93% by using GLCM based haralick features and RF classifier.

Conclusion:

An automatic classification of tongue image for diabetes detection using random forest classifier is described in this study. Initially the input tongue images are preprocessed by median filter. Then extracted by using GLCM and haralick features. The RF classifier is used for final output prediction. The overall classification accuracy is 96 % by using statistical features and RF classifier. It's sensitivity and specificity are 95 % and 94%. The accuracy for abnormal tongue images are 92 % and also its sensitivity and specificity are 95% and 94% by using RF classifier.

REFERENCE:

- [1] Zhao, Q., Zhang, D., & Zhang, B. (2016, October). Analysis of digital language images with ColorChecker in medical applications. In 2016 2nd IEEE International Conference on Computer and Communications (ICCC) (pp. 803-807). IEEE.
- [2] Kim, K. H., Do, J. H., Ryu, H., & Kim, J. Y. (2008, November). Method of tongue diagnosis for successful area exclusion and language classification. In 2008 First Workshops on Image Processing Theory, Tools and Applications (pp. 1-7). IEEE.
- [3] Zhang, B., Kumar, B. V., & Zhang, D. (2013). Detection of diabetes with the tone, shape and geometry of the language diabetic retinopathy;. IEEE transactions on biomedical engineering, 61(2), 491-501.
- [4] Wang, X., Zhang, B., Yang, Z., Wang, H., & Zhang, D. (2013). Tongue images mathematical analysis for the extraction and diagnosis of features. IEEE Transactions on Image Processing, 22(12), 5336-5347.
- [5] Liang, R., Wang, Z. P., Yang, X. Y., Ren, Y. J., Zhang, Y., & Yao, X. Y. (2012, August). Applied colorimetric studies in language diagnostic education. In 2012 International Symposium on Information Technologies in Medicine and Education (Vol. 1, pp. 426-429). IEEE.
- [6] Akgul, Y. S., Kambhamettu, C., & Stone, M. (1999). Automatic tongue contour extraction and recording. IEEE Transactions on Medical Imaging, 18(10), 1035-1045.
- [7] Ranganathan, S. (2019). Images with filter bank for rain removal. International Journal of MC Square Scientific Research, 11(1), 9-14.
- [8] Manahoran, N., & Srinath, M. V. (2017). Marine Picture Segmentation K-Means Clustering. International Journal of MC Square Scientific Research, 9(3), 26-29.
- [9] Santosh Kumar Srivastava. (2020). S-Transformation and Gaussian model for the classification of acoustic scene. International Journal Of Advances In Signal And Image Sciences, 6(1), 29-37.
- [10] Prasath, R., Kumanan, T., Department of CSE, Meenakshi Academic of Higher Education and Research, Chennai, India, imaging Science Journal, Volume 67, Issue 2, Pages 76-89.
- [11] J. Aswini, N. Malarvizhi, T. Kumanan , Department of CSE, Meenakshi Academic of Higher Education and Research, Chennai, India, International Journal of Innovative Technology and Exploring Engineering ,Volume 8, Issue 5, 2019, Pages 91-96.