AN EFFICIENT SEGMENTATION METHOD FOR AUTOMATED TONGUE EXTRACTION USING HSV COLOR MODEL

Saba A. Tuama[1] and Jamila H Saud[2]

[1]*Informatics Institute of Postgraduate Studies, University of Information Technology and Communication, Baghdad, Iraq*
[2]*Mustansiriyah University/ College of Science/ Computer Science Dept., Baghdad, Iraq*
[1]sabaayad91@gmail.com, [2]dr.jameelahharbi@gmail.com

Abstract— **The segmentation step of the tongue is one of the most prerequisite steps in various applications such as diseases diagnosis, and human identification. This paper presents a new method for extraction tongue region in colored images using HSV color model for the identification system. The proposed method consists of many steps. At first, the image color is transformed from RGB color space to HSV color space. Then, thresholding is done on the hue component of the HSV color model. Finally, post preprocessing operation utilized to remove the noise and the fake object regions. Through testing of the proposed method on a variety of tongue images, the method has proved to be effective and satisfactory for tongue image segmentation. The suggested method achieved efficient results in the detection rate for the tongue region.**

Keywords— **Image segmentation; HSV color model; Tongue Identification; Tongue extraction**

1. INTRODUCTION

Human tongue biometric delivers a level of uniqueness for using in identification applications that other biometrics cannot match in the context of that it is difficult to forge and it is well protected in the mouth [1], [2]. The first step of such a system is to segment the tongue region from the collection of images completely and accurately. Automated tongue segmentation is complicated due to the variance of tongue shape, interference of the lips and the complication of the pathological tongue [3]. It is difficult to segment a tongue region from the image because it not only contained tongue print but also lips, mouth itself, teeth and others object are the objects of the mouth. So, the problem of this segmentation process is how to extract the tongue print among other objects in that mouth region.

Tongue segmentation is a challenging task and has been in the focuses of researches all over the world in recent years. So, different methods for automated tongue segmentation have been proposed Zuo *et al.* [4] have presented a method for tongue segmentation by combining active contour model and polar edge detector. Firstly, a polar edge detector was proposed to extract the edge of the tongue body. Secondly, the edge filtering process was utilized to filter out the edge that was useless for tongue segmentation. After that, and local adaptive edge binarization method was applied to binarize edge image. Finally, an initialization and active contour model was used for segmentation of tongue body. The

presented method was tested on 50 images and achieved TPR was 97.2% and FPR was 4.7 %.

Fu Zhicheng *et al.*, [5], have proposed a method for automatic tongue segmentation. Initially, look-up-table based real AdaBoost algorithm is introduced for getting the cascade tongue detector that could locate the tongue in human face images faster. After that, color cancellation was utilized for removing the lip. The presented method was combined radial edge detection with an active contour model to obtain the accurate contour of the tongue. Their proposed method gave experimental results a precise rate of up to 94.91%.

Ning Jifeng *et al.*, [6], proposed a method for automatic tongue extraction based on region merging and gradient vector flow. Initially, the gradient vector was modified in order to diffuse the tongue image whilst preserving the edge of the tongue. After that, segmentation the diffused tongue image into small regions by utilized the watershed algorithm. Then, the maximal similarity-based region merging was used to segment the tongue body region. Finally, the snake algorithm was utilized to refine the result of region merging. The presented method was tested on 200 images by traditional Chinese medicine practitioners and tested on 50 tongue images by using ROC analysis.

Shi *et al.*, [7], presented a method for tongue image segmentation by using the double geo-vector flow. At first, a Tongue image was refinement through a saliency window. Then, they initialized an area of the tongue as the upper binary part and lower level set matrix. After that, they proposed a double geo-vector flow (DGF) in order to detect the tongue edge and extract the tongue area in the image. Their proposed method (DGF) was evaluated by using 100 images provided by the Laboratory of Traditional Medical Syndromes, Shanghai University. The DGF results gave true-positive volume fraction reaching 98.5%, its false-negative volume fraction is 1.42%, and false-positive volume fraction is 1.51%.

Saparudin *et al.*, [8], presented an approach for tongue image segmentation from opened mouth image by using Active Contour method and Canny Edge Detection. Their proposed method had four steps: Smoothing Gaussian Filter, Finding Gradients, Non-maximum Suppression, and Hysteresis Thresholding. Then, Canny Edge Detection was utilized for finding the edges of tongue and followed by Active Contour Model and cropping process to produce tongue image. The proposed method was tested on 40 tongue images and achieved an accuracy rate of 75%.

Priya Mohana *et al.*, [9], proposed a method of tongue segmentation for Detection of Diabetes by using Gradient Vector Flow snake technique. For extraction of the tongue region, they have done many steps. Firstly, the input image was cropped for obtained the exact tongue region and following by rescaling of the tongue region image. Secondly, applied Gray conversion and Gaussian filter. Thirdly, performed a segmentation process by using GVF snake technique which includes three steps: Gradient image, Smoothing image, and edge detection. The proposed system was tested on a collection of tongue image which was captured by using a smartphone.

The organizing of this paper is in four sections. The tongue segmentation method is presented in Section 2. Experience and performance evaluation are presented in Section 3. Finally, conclusions are given in Section 4.

2. PROPOSED METHODOLOGY

We present an efficient extraction method based on color segmentation using HSV color space, binarization, and post preprocessing for segmentation the tongue-print from the color images. The stages of the proposed method are shown in Figure 1.
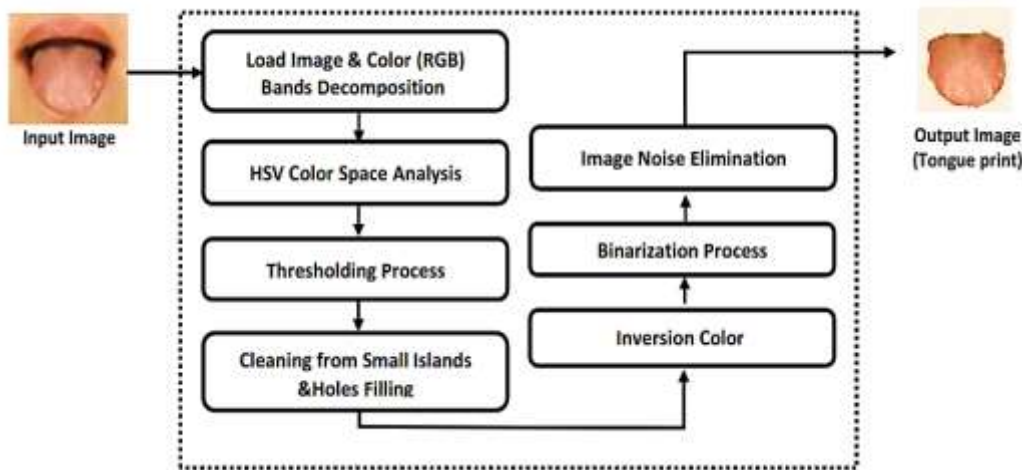
Fig. 1   Steps of the Proposed Method

2.1 HSV COLOR SPACE ANALYSIS

HSV color models are most beneficial in applications wherever segmentation relies heavily on specific features of the colors rather than overall perceptual variance [10]. The HSV color model is represented via a single cone. Fig. 2 that illustration the three components of the cone (hue, saturation, and value) [11]. Color segmentation by using HSV color space is elaborately described in [12], [13], [14].

The tongue image is transformed from RGB into HSV color space by using the equation (1), (2), (3) [15].Thus the tongue image is split into three various components which are hue, saturation and value components based on color and intensity information, as seen in Fig. 3.
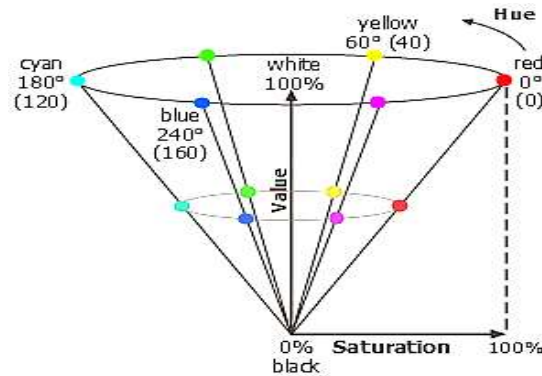

Fig. 2   HSV color model

$$H = \arccos \frac{\frac{1}{2}(2R - G - B)}{\sqrt{(R - G)^2 - (R - B)(G - B)}} \tag{1}$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \tag{2}$$
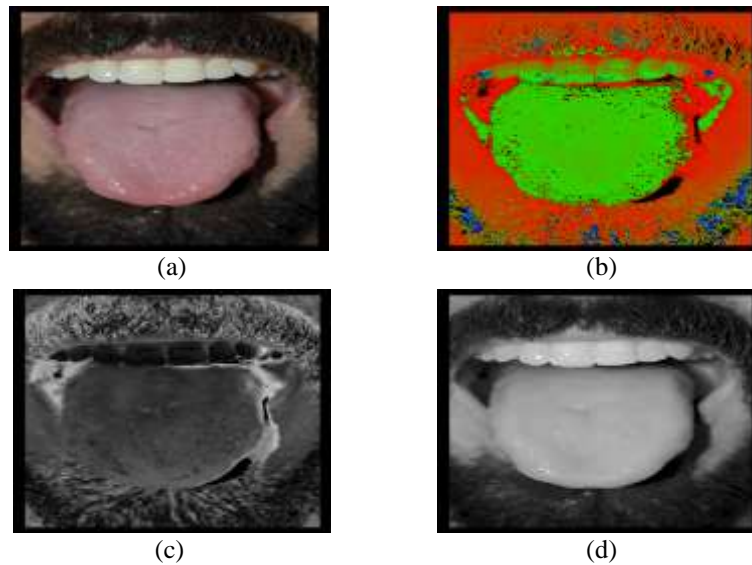
$$V = \max(R, G, B) \tag{3}$$

Fig. 3 Three component of HSV color space (a) input image,
(b) Hue, (c) Saturation, and (d) Value.

2.2 THRESHOLDING PROCESS

After applied hue conversion process; thresholding process is applied in order to isolate the tongue region from other regions. This process is accomplished by using equation (4) [16]. Initially, scanning all pixels of the Hue image and testing the hue values; if the hue value between T1 and T2 then the pixel marked as foreground and take white color (*i.e.*, 1's), else the pixels marked as background and take black color (*i.e.*, 0's). The result after applied thresholding process is shown in Fig. 4.

$$\text{Thr\_img}(x, y) = \begin{cases} 1 & \text{if } T1 \leq h(x, y) \leq T2 \\ 0 & \text{Otherwise} \end{cases} \qquad (4)$$

Where Thr_img() is the produced binary image, T1 & T2 are the thresholding values, and h() is the hue image.
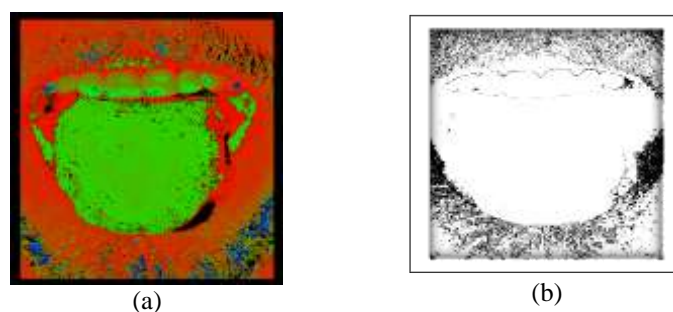

Fig. 4   (a) Before thresholding, (b) After thresholding.

2.3 CLEANING FORM SMALL ISLANDS AND HOLES FILLING

The producing image from the previous step contains some of the small islands which consider as noise patches and appear as tongue regions rather than parts of the background due to miss thresholding, as seen in Fig. (5a). So, this step is major in order to eliminate these islands. Seed filling technique is used to cleaning from the regions of the small islands.

At first, scan every pixel in Thr_Img() image, if its value is 1, then this first pixel will be considered as a first seed pixel, and its location is added in a temporary buffer, and convert its value to 3 (it means visited). after that, test the four direct neighbors of the detected seed point (*i.e.*, left, right, top & bottom), if any one of them has value 1; then its location is registered to the temporary buffer and convert its value to (3). Repeat the check of pixel value sequentially for all pixels listed in a temporary buffer, till reaching the end of the buffer. after that, it would be checking from size of the collected segment; if the number of collected segment (*i.e.*, size of filled buffer) that contains white pixels is less than a predefined threshold value Thr_v1, then it discarded and considered as part of the background (0' s); otherwise, it would be considered as part of tongue (1's). Fig. (5b) shows the effect of removing the small islands.
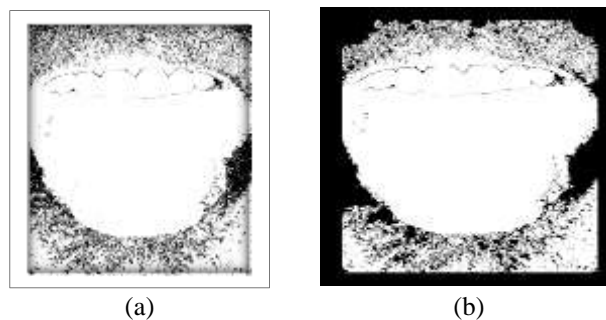

(a)   (b)

Fig. 5   The effect of removing small islands: (a) Binary image, (b) after small islands removal

Some small black area (*i.e.*, gaps) may appear inside tongue area after thresholding process; thus in order to handle this problem, the tiling process was applied for removing all small black areas inside the tongue area in the binary image. This process starts by scanning all pixels of the binary image; in case the pixel value is found zero (i.e., black pixel inside the tongue area); then open a structure element around it (black pixel), and compute the number of all neighbour white pixels (i.e, 1) inside the structure element. In case the count value of white pixels is found more than the threshold value (thr_val), the tested black pixel is considered as gap point and converted to white pixel (1); otherwise, it is kept black (0) without any change. Fig. (6b) shows the effect of gaps filling process by tiling process.
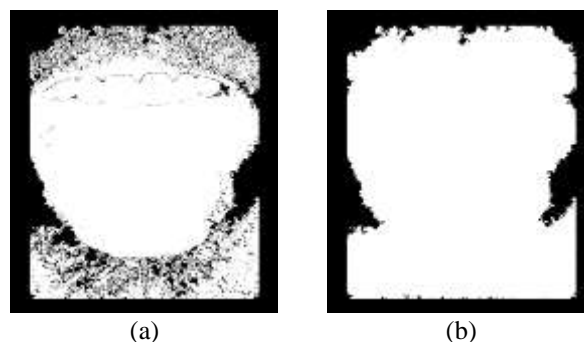

(a)   (b)

Fig. 6   The effect of gaps filling: (a) The binary image, (b) after gaps filling.

2.4 INVERSION COLOR OF IMAGE AND BINARIZATION PROCESS

At first, scanning all pixel in the binary image that produced from the previous holes filling process, and checking if the pixel value is white then it would be converted to the pixel value in the input image; otherwise, it is kept as it without any change (i.e., 0's

value). After that, the producing image contains some unwanted areas (i.e., hair) as seen in Fig. (7a); Thus, color inversion [17] and followed by the binarization process is applied for removing these unwanted areas.

After applied color inversion the binarization process is utilized by checking all pixel in the resulting image from the inversion process; if the pixel value is less than a thresholding value (Thr), then it would be marked as a pixel belonging to the tongue area (i., 1's value); otherwise the pixel value will be changed to zero value. The resulting image after the color inversion and binarization processes is shown in Fig. (7b) (7c), respectively.
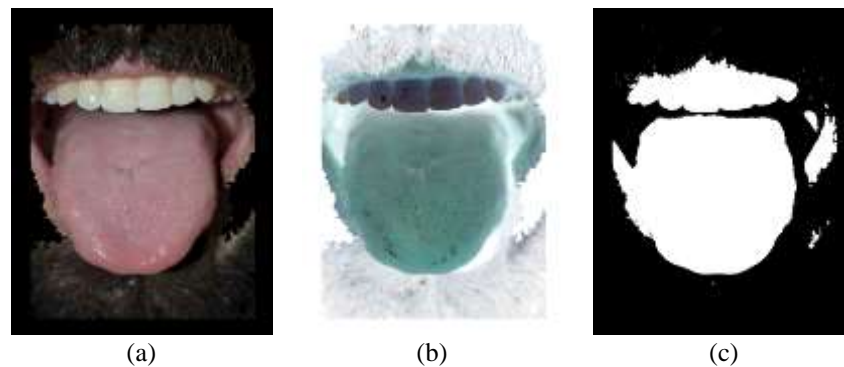

(a)　　　　　　　　(b)　　　　　　　　(c)

Fig. 7　The effect of color inversion and binarization process: (a) The color image, (b) after color inversion, (c) after binarization process

2.5 IMAGE NOISE ELIMINATION

The resulting image from the previous step (*i. e.*, binarization process)　contains some islands that are detected falsely as tongue region. So, to get good segmentation result it would be applied region growing algorithm [18],[19] for removing these small white pieces of pixels; Fig. (8b) shows the resulting image after the applied process of the removal of the small island.
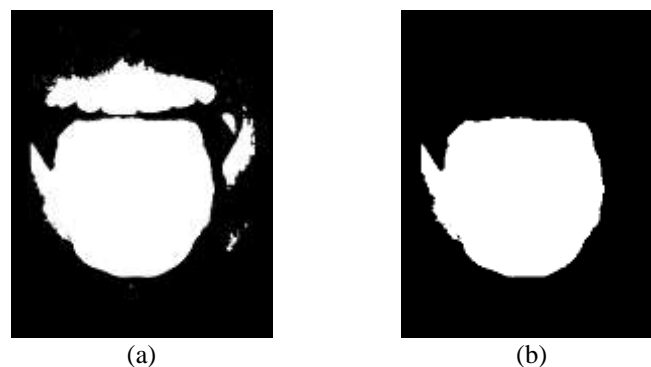

(a)　　　　　　　　(b)

Fig. 8　(a)before noise elimination, (b) after noise elimination.

2.6 ALLOCATING OF TONGUE AREA

At first, scanning all pixel in resulting image from the previous step (i.e., noise elimination process), and checking if the pixel value equals one (i.e., white) then it would be converted to the pixel value in the original image; otherwise, it is kept as it without any change (i.e., 0's value). The resulting image after allocating tongue region is shown in Fig. (9b).
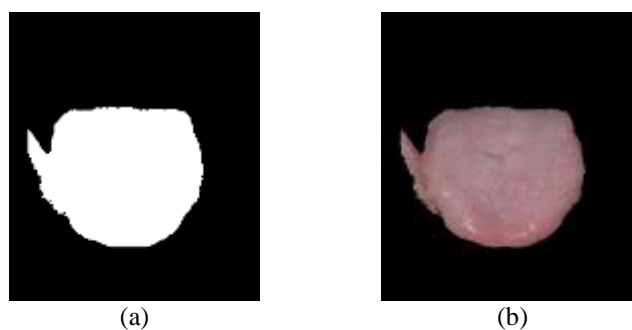
(a)   (b)

Fig. 9   (a) before allocating tongue, (b) after allocating tongue.

3. RESULTS AND DISCUSSION

The proposed segmentation method is tested on our own database of tongue images which is consists of 116 tongue images taken from 116 subjects. The images were acquired over a period of 3 months, and all the subjects in the database are in the age group of (7–47) years. All the images were taken using three cameras (i.e., Nikon D3200, Nikon D5100, and Nikon D7200), under different lighting conditions with the subject placed at a distance of about 0.5 meters from the camera, and we used a 55mm focal length for the tongue images. The resolution of images from camera1, camera 2, and camera 3 are (690×1012), (572×838), and (690×1010) pixels respectively, in jpeg format. Figure 10 shows examples of tongues images captured in different illumination and environment.

Fig. 10. Samples of tongue images from different subjects.

The performance of the suggested method is evaluated by using three segmentation performance metrics: (i) Accuracy, (ii) True Positive Rate (TPR), and (iii) False Positive Rate (FPR); they are defined as follows [20]:

$$TPR = \frac{TP}{TP + FN} \tag{5}$$

$$FPR = \frac{FP}{FP + TN} \tag{6}$$

$$Acuuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{7}$$

where TP refers to the number of pixels that are detected and identified as tongue points correctly, TN refers to the number of pixels that are detected and identified as a background points correctly, FP refers to the number of pixels that are identified incorrectly as tongue points while in truth it is background points, and FN refer to the number of pixels that are identified incorrectly as background points while in truth it is tongue points. Table I shows the resulting of our proposed approach by using the three segmentation performance measures.

Table I. Evaluation results by three segmentation performance measures.

	TPR	FPR	Accuracy
Proposed method	0.99	0.009735923	0.9906%

The suggested approach shows the value of FPR is less than other methods. For instance, Zuo et al. [4] method showed good segmentation tongue; but their approach added many false pixels (FP), and they achieved higher FPR than our method; while achieved accuracy of segmentation 97.2% less than our achieved segmentation accuracy. As well as Shi et al. [7] method achieved FPR 1.51% which is higher than FPR achieved from our method.

Also, our proposed method evaluation by using a performance metric (i.e., detection rate) in order for measuring the success of our suggested method. Detection rate is defined as the number of correctly detected ears over the whole number of tongues in the database, and it's computed by using equation 8. Table II summarizes the results of tongue detection accuracy in the proposed method.

$$\text{Detection Rate} = \frac{\text{No. of correctly detected Tongue}}{\text{Total No. of Tongue Images}} \times 100\% \qquad (8)$$

Table II. Accuracy of the suggested tongue segmentation method on the different datasets: respective numbers (#) and percentages (%) of correct detections.

Tongue Images captured from 3 cameras (#images)	Correct detection (#)	Correct detection (%)
Captured from Nikon D3200 (18)	17	93.2%
Captured from Nikon D5100 (82)	81	98.78%
Captured from Nikon D7200 (16)	16	100%

The experimental results show that the proposed method achieved a correct detection rate of 98.27% for segmentation of all tongue images captured under variant environment and the suggested method is more accurate and uses fewer complexity steps in extracting tongue region in comparison to the other published methods in the literature. For example, Saparudin *et al.*, [8] method hasn't detected the tongue region properly in all test images that used; though they used small samples include 50 images, and they achieved accuracy and detection rate less than the accuracy and detection rate of our suggested method although we used a lot of test images which include 116 images. Also, Fu Zhicheng *et al.*, [5] method achieved an accuracy of 94.91% that is less than our achieved accuracy. Table

III displays some of the results of the suggested method where tongues of various size and shape are properly detected from different tongue images.

Table III. Shows segmentation results of tongue with different shape and size.

Input image	Resulting image

4. CONCLUSIONS

The process of tongue segmentation is complicated due to the variance of tongue shape, a complication of the pathological tongue and interference of the lips. In this paper, an accurate segmentation method is presented for automated tongue extraction by using HSV Color Model, thresholding, and post preprocessing operation. The proposed method is tested on many images that have different qualities. Experimental results explain that the proposed method can segment the tongue precisely. The proposed method was tested on 116 tongue images and the experimental result shows that the proposed method achieved a detection rate of 98.27% for the segmentation of tongues.

REFERENCES

[1] Naaz, Rohaila, Sanjana Yadav, and Manoj Diwakar. "Tongue image extraction technique from face and its application in public use system (banking)." 2012 International Conference on Communication Systems and Network Technologies. IEEE, Rajkot, India, 2012, pp. 203-206.

[2] Paulose, Akhil, et al. "Biometric Applications of Tongue Images." International Journal of Engineering Trends and Technology (IJETT),2016, Vol. 31, No. 2, pp. 69-72.

[3] Bob, Zhang, et al. "Tongue Image Analysis.", first ed. Book published in Springer, Computers, 2017, 335 pages.

[4] Zuo, Wangmeng, et al. "Combination of polar edge detection and active contour model for automated tongue segmentation." Third International Conference on Image and Graphics (ICIG'04). IEEE, 2004, pp. 270-273.

[5] Fu, Zhicheng, et al. "Automatic tongue location and segmentation."2008 International Conference on Audio, Language and Image Processing. IEEE, 2008, pp. 1050-1055. doi:10.1109/ICALIP.2 008.459 0130.

[6] Ning, Jifeng, et al. "Automatic tongue image segmentation based on gradient vector flow and region merging." Neural Computing and Applications 2012, Vol. 21, No. 8, pp. 1819-1826.

[7] Shi, Miao-Jing, et al. "Computerized tongue image segmentation via the double geo-vector flow." Chinese medicine 9.1 (2014), pp. 1-7.

[8] Saparudin, Erwin Erwin, and Muhammad Fachrurrozi. "Tongue Segmentation Using Active Contour Model." Proceeding of the Electrical Engineering Computer Science and Informatics 3.1 (2016): 012041. doi:10.1088/1757899X/190 /1/012041.

[9] Priya, N. Mohana, V. Gayathri, and Mrs Beulah Sherin Ponmalar VJ. "Performance Analysis And Detection Of Diabetes Using Classifiers In Tongue Image." International Journal of Pure and Applied Mathematics 119.15 (2018), pp. 819-825.

[10] Mavrinac, Aaron, et al. "Competitive learning techniques for color image segmentation." 2008 Congress on Image and Signal Processing, IEEE, 2008, Vol. 3, pp. 644-649.

[11] Ibrahim, Abdul-Wahab Sami, and Hind Jumaa Sartep. "Grayscale i mage coloring by using YCbCr and HSV color spaces." International Journal of Modern Trends in Engineering and Research, 2017, Vol. 4, No. 4.

[12] Zarit, Benjamin D., Boaz J. Super, and Francis KH Quek. "Comparison of five color models in skin pixel classification." Proceedings International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems. In Conjunction with ICCV'99 (Cat. No. PR00378), IEEE, 1999, pp. 58-63.

[13] Sigal, Leonid, Stan Sclaroff, and Vassilis Athitsos. "Estimation and prediction of evolving color distributions for skin segmentation under varying illumination." Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662), IEEE, 2000, Vol. 2, pp. 152-159.

[14] Sural, Shamik, Gang Qian, and Sakti Pramanik. "Segmentation and histogram generation using the HSV color space for image retrieval." Proceedings. International Conference on Image Processing, IEEE, 2002, Vol. 2.

[15] Shaik, Khamar Basha, et al. "Comparative study of skin color detection and segmentation in HSV and YCbCr color space." Procedia Computer Science 57, 2015, pp. 41-48.

[16] Gonzalez, Rafael, et al."Digital image processing (Book)." Reading, Mass., Addison-Wesley Publishing Co., Inc.(Applied Mathematics and Computation, Third Edition, 2008.

[17] Agu, E. "Digital Image Processing (CS/ECE 545). Lecture 2: Histograms and Point Operations (Part 1)." Computer Science Dept. Worcester Polytechnic Institute (WPI) http://web. cs. wpi. edu/~ emmanuel/courses/cs545S14, 2013. https://web.cs.wpi.edu/~emmanuel/course s/cs545/S14/slides/lectur e09.pdf.

[18] Gómez, Octavio, Jesús A. González, and Eduardo F. Morales. "Image segmentation using automatic seeded region growing and instance-based learning." Iberoamerican Congress on Pattern Recognition. Springer, Berlin, Heidelberg, 2007, pp. 192-201.

[19] Hojjatoleslami, S. A., and Josef Kittler. "Region growing: a new approach." IEEE Transactions on Image processing, 1998, Vol. 7, No. 7, pp. 1079-1084.

[20] Zhang, Yu Jin. "A review of recent evaluation methods for image segmentation." Proceedings of the Sixth International Symposium on Signal Processing and its Applications (Cat. No. 01EX467), IEEE, 2001. Vol. 1, pp. 148-151.