An Efficient Segmentation of U-area and T-area on Facial Images by Using Matlab with Hough Transform and Viola-Jones Algorithm Base

Indriyani[1,2*], Ida Ayu Dwi Giriantari[2], Made Sudarma[2], I Made Oka Widyantara[2]
{indry.joice@gmail.com[1], dayu.giriantari@unud.ac.id[2], msudarma@unud.ac.id[2]}

Institute of Technology and Business (ITB) STIKOM Bali, Indonesia[1],
Udayana University, Indonesia[2]

Abstract. The left and right cheeks and chin (U-area) and the forehead and nose (T-area) are useful for examining skin types. These areas provide crucial information to determine facial characteristics. The Matlab application function, particularly the Viola-Jones Algorithm base, is helpful in detecting the U and T areas. This application accurately detects normal images for face, including the eyes, nose, and mouth position. The T-area was determined by referring to the position of the eyes and nose, while the U-area was identified based on the eyes, nose, and mouth position. The accuracy in determining the two areas using this method was close to 100%.

Keywords: U-area; T-area; Viola-Jones Algorithm

1 Introduction

Some parts of the face, such as the forehead and nose (T-area), chin and cheeks (U-area), are useful in examining skin types. The two parts have distinct features, with the T-area being oily while the U-area being dry [1][2]. These observations are supported by the results [3], which stated that most Chinese women with oily skin have shiny noses, foreheads, and cheeks. More research has been conducted regarding facial skin characteristics, including the level of oil production and facial contours, based on the features extracted from the T-area and U-area [4, 5].

The location of facial organs such as the eyebrows, eyes, and mouth help determine the T-area and U-area. A lot of research concerning automatic eyebrow detection for facial recognition has been conducted. Previous research shows that the Hough transformation application can be used to obtain the approximate position of the eyebrows based on the inclination level of the face [6, 7]. For instance, the Prewitt method locates the eyebrow line through edge detection by calculating the average value of the green channel gradient on the image of the upper part of the face according to the X-axis [8]. This method is proposed based on the consideration that the eyebrows are at the top of the face and have a darker color than the skin in surrounding area. Another method is performed by extracting the features of the eyes, mouth, and facial edges using measurements obtained from the image color space [9]. The eye usually contains dark and bright pixels on the luma component, making its detection easier using morphological surgery [10].

MIC 2022, November 12, Semarang, Indonesia

DOI 10.4108/eai.12-11-2022.2327390

The neural network method helped to detect the mouth and lips while searching for objects with the longest major axis value was conducted using Principal Component Analysis. Furthermore, opening morphology is used to detect eyebrows in grayscale images of the face, and oral detection is performed using the 4-connected context method [11, 12]. However, mouth detection becomes difficult when the mouth is open to the extent of exposing the teeth. This is because opening the mouth makes it hard to enhance the contrast and perform hole filling.

Previous research shows that the detection of brow contours might use K-means clustering [13]. Another research showed that eyebrow detection could be conducted by initially detecting the position of the eyes and then searching the dark area above [14]. Furthermore, the Gaussian model was used in the HSV color space to detect non-skin areas and the Laplacian operator to perform edge detection. The mouth detection was performed from the left end part of the mouth using a subclass-based classifier, then used to generate a Gaussian model and Laplacian operator. An efficient way to yield a 96% success level for detecting eyebrows involves a computational method with a gradient and signed edge map [15]. In general, eyebrow detection methods are inaccurate when the brow line is not visible because it is covered by another object (glasses or hair), a thin eyebrow line, or uneven lighting.

The T-area detection can be conducted by initially using canny edge detection and Hough transform [16]. The success of determining the correct position of eyebrows affects the location of the T-area. Furthermore, to detect the U-area [17] requires several stages, including detecting the fuzzy edge to produce edge images. The Hough transformation application is applied to detect facial shapes, and finally, the Viola-Jones algorithm detects the location of the eyes and mouth. However, this method requires time and extensive memory since the two areas use different methods and require less efficient algorithms and computations.

This research applied a simpler, faster, and more efficient method by utilizing one of the Matlab functions, which is also based on the Viola-Jones algorithm to obtain facial, eye, nose, and mouth position features. There were 25 facial images from 4 different races samples that were analyzed, and the results of U-area and T-area position detection obtained 100% accuracy.

2 Methodology

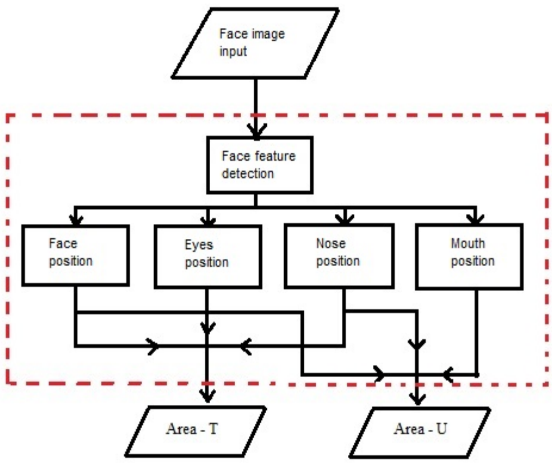

Fig. 1. U-area and T-area detection methods.

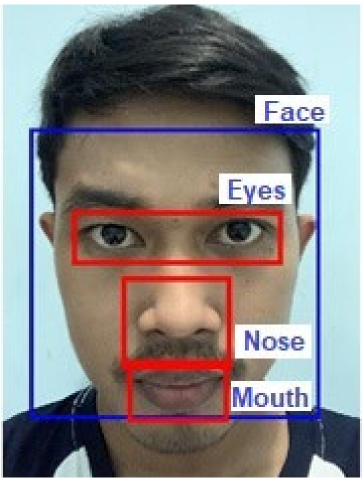

Fig. 2. Position of face, eyes, nose, and mouth.

The information provided by Matlab vision.cascadeObjectDetector function was used to detect and segment the U and T areas. This function is based on the Viola-Jones algorithm, which accurately determines the normal face, eyes, nose, and mouth position. The T-area was determined based on the position of the face, eyes, and nose, while the U-area was determined with the help of information on the position of the face, nose, and mouth. The image below shows the results of the face, eyes, nose, and mouth position detection.

2.1 U-area

The U-Area involves the area around the cheeks and chin, which forms the letter U and is helpful in classifying facial skin types. This area requires extra attention and treatment because it has more oil glands than other skin parts. The example of the U-area on the face is shown in Figure 3.

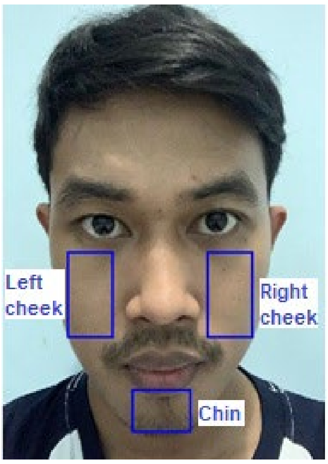

Fig. 3. U-area on the facial images.

2.2 T-area

The T-area includes the forehead and nose, which form the letter T on the face, and it is deemed the oiliest part of the skin because of its numerous oil glands. Figure 4 below shows this part of the face.

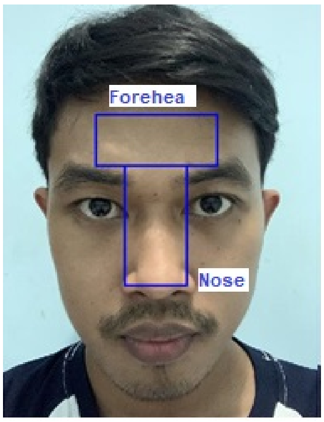

Fig. 4. T-area on the facial images.

3 Result and Discussion

3.1 Image dataset and groundtruth

The detection system experiments of U-area and T-area on facial images used a sample data of 24. The data for the front facial images were drawn from four Indonesian races, including mixed, Mongoloid Malay, Papuan Melanezoid, and Weddoid. The experts selected the U-area and T-area groundtruth best images, which were used for comparison.

3.2 Position detection of face, eye, nose, and mouth

Before detecting the U and T areas, the first step involved detecting the facial border, position of eyes, nose, and mouth. With the help of the vision.cascadeObjectDetector Matlab function, information was detected in the form of coordinate points of the upper left corner of xy, x = data 1 and y = data 2, horizontal distance (data 3), and vertical distance (data 4). When detecting a face, the following data will be obtained:

$$BBf=[\ f1\ f2\ f3\ f4] \quad (1)$$

Description :
f1 = x-coordinate of the top left corner of the face box
f2 = y-coordinate of the top left corner of the face box
f3 = distance between the x-coordinate and the right border of the face box
f1 = distance between the y-coordinate and the lower bound of the face box

This data was used to determine the size of the box in which the face was detected.

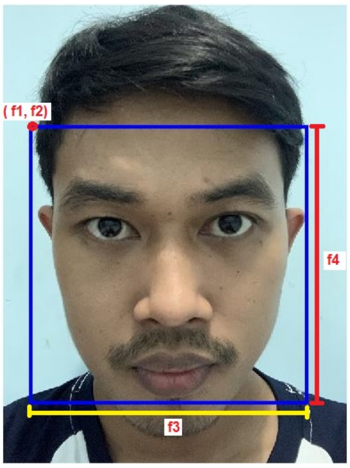

Fig. 5. Information of boundary and position.

The upper horizontal line information was used as the upper border for the forehead area (horizontal T-area). Also, the lower horizontal line was used as the lower border of the chin (lower U-area).

Additionally, the eye position detection results were extracted using the same function. The information obtained from the eye position is given below.

$$BBe=[e1\ e2\ e3\ e4] \quad (2)$$

Description :
e1 = x-coordinate of the upper left corner of the eye box
e2 = y-coordinate of the upper left corner of the eye box
e3 = horizontal distance of eye box
e1 = vertical distance of eye box

The relationship between BBe and the eye position is shown in Figure 6.

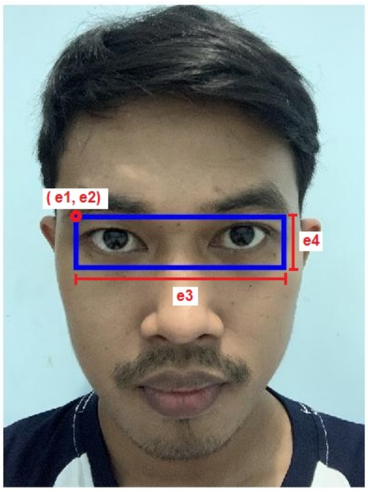

Fig. 6. Eye position.

The information from the center point was used as a reference for the vertical T-area. Also, two-thirds of the eye box length was used as a reference to obtain the length of the forehead box (forehead / horizontal T-area). The width of the eye box was used as a reference to get the distance between the eye box and the forehead box.

The results of nose position detection were extracted with the same function, and the information obtained is given in the equation below.

$$BBn=[n1\ n2\ n3\ n4] \quad (3)$$

Description :
n1 = x-coordinate of the top left corner of the nose box
n2 = y-coordinate of the top left corner of the nose box
n3 = horizontal distance of nose box
n1 = vertical distance of nose box

The relationship between body weight and nose position is presented in Figure 7.

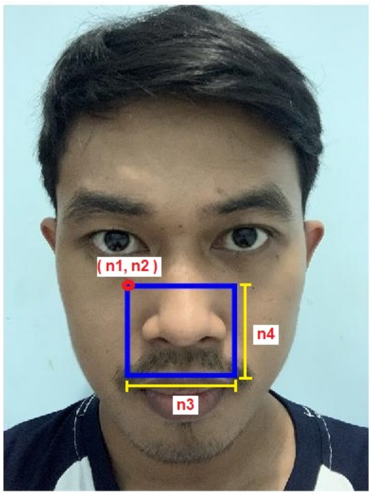

Fig. 7. Nose position.

Information regarding the nose box midpoint was used as a reference for the lower end of the vertical T-area. Therefore, the vertical T-area distance was obtained easily by drawing the distance between the lower bound of the horizontal T-area / forehead to the midpoint of the nose box. Furthermore, the midpoint of the nose was also used as the center of the vertical T-area. The width of the vertical T-area was obtained by an applied deviation to the left and the right. Right and left borders together with the upper border of the mouth and the distance between the lower border of the eye are used to get the U-area of the left-cheek and right-cheek.

On the same note, the detection of the mouth position was extracted, and the obtained information was recorded as follows.

$$BBm=[m1\ m2\ m3\ m4] \quad (4)$$

Description :
m1 = x-coordinate of the upper left corner of the mouth box
m2 = y-coordinate of the upper left corner of the mouth box
m3 = horizontal distance of mouth box
m1 = vertical distance of mouth box

The relationship between BBm and mouth position is shown in Figure 8 below.

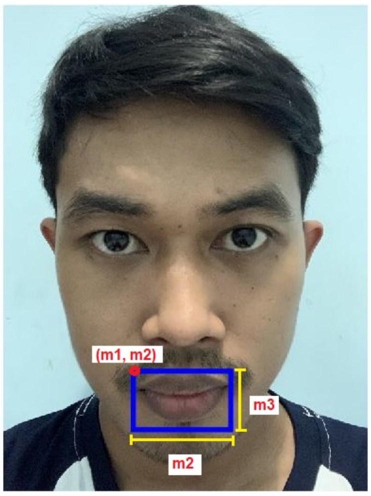

Fig. 8. Mouth position.

The information from the lower border of the mouth was used as a reference for the upper limit of the lower U-area (chin). Also, the lower limit of the chin used the lower border of the face area, while the width of the mouth was used as a reference to get the length of the box in the chin area.

3.3 T-area detection

The T-area detection comprised the forehead and nose, whereby the forehead area detection used the information from the face area, eye equation-2, and nose, given in equations-1 to equation-3. Also, the T-area includes the forehead (horizontal T-area) and the nose (vertical T-area). The following equation was used to get the forehead area:

The coordinates of the right corner (x) were obtained from the left border of the eye box plus 1/6 of the length of the eye box, because most oily facial features in the forehead area are located at 2/3 of the forehead width in the middle position, and the upper limit (y) used the upper border of the face area:

$$Th1=e1+e3/6 \quad (5)$$

Description :
Th1 = x-coordinate of the top left corner of the forehead box
e1 = x-coordinate of the upper left corner of the eye box
e3 = horizontal distance of eye box

$$Th2=f2 \quad (6)$$

Description :
Th2 = y-coordinate of the top left corner of the forehead box
f2 = y-coordinate of the top left corner of the face box

The width of the forehead measured 2/3 of the length of the eye area, while the lower border of the forehead area is above the upper border of the eye area as far as 1/3 of the eye area:

$$Th3=e3*2/3 \quad (7)$$

Description :
Th3 = horizontal distance of forehead box
e1 = x-coordinate of the upper left corner of the eye box

$$Th4=(e2-e4/2)-f2 \quad (8)$$

Description :
Th4 = vertical distance of forehead box
e2 = y-coordinate of the upper left corner of the eye box
e4 = vertical distance of eye box
f2 = y-coordinate of the top left corner of the face box

The results of the T-area detection on the forehead (horizontal) are as shown in Figure 9.

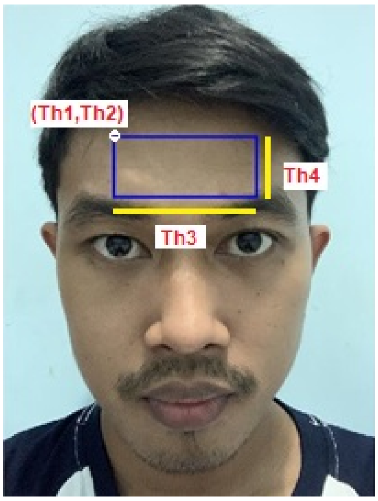

Fig. 9. T-area of the forehead (horizontal).

The information regarding the nose and forehead areas was fundamental in calculating the T-area of the nose. Therefore, the x coordinate is the midpoint of the nose area minus ½ the width of the vertical T-area. Because the width of the vertical T- area is, half on the left of the center line and the other half on the right. However, the width of the vertical T-area is 2/3 the length of the nose area, because the width of the vertical T-area, located between the eyes, and the width is about 2/3 the width of the nose.

$$Tv3=n3*2/3 \text{ (the width of the T-area)} \quad (9)$$

Description :

Tv3 = the horizontal distance of the nose box (the width of the T-area of the nose)
e2 = y-coordinate of the upper left corner of the eye box
n3 = horizontal distance of nose box

$$Tv1=(n1+n3/2)-Tv3/2 \quad (10)$$

Description :
Tv1 = x-coordinate of the top left corner of the nose box (vertical T-area)
n1 = x-coordinate of the top left corner of the nose box
n3 = horizontal distance of nose box
Tv3 = the horizontal distance of the nose box (the width of the T-area of the nose)

The upper limit (y) was obtained from the lower boundary of the forehead given by:

$$Tv2=Th2+Th4 \quad (11)$$

Description :
Tv2 = y-coordinate of the top left corner of the nose box (vertical T-area)
Th2 = y-coordinate of the top left corner of the forehead box (horizontal T-area)
Th4 = vertical distance of forehead box (horizontal T-area)

The height of the nose was obtained by determining the distance between the midpoint and the upper border of the nose area:

$$Tv4=(n2+n4/2)-(Th2+Th4) \quad (12)$$

Description :
Tv4 = nose box vertical distance (vertical T-area)
n2 = y-coordinate of the upper left corner of the nose
n4 = vertical distance of nose box
Th2 = y-coordinate of the top left corner of the forehead box (horizontal T-area)
Th4 = vertical distance of forehead box (horizontal T-area)

The detection results of the nose around the T-area (vertical) are as shown in Figure 10.

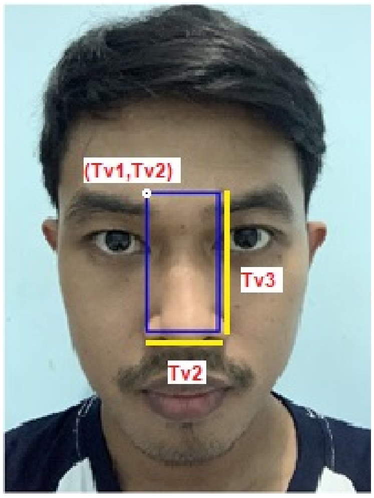

Fig. 10. T-area of the nose.

The total T-area was obtained when Figures 8 and 9 were combined, as shown in Figure 11.

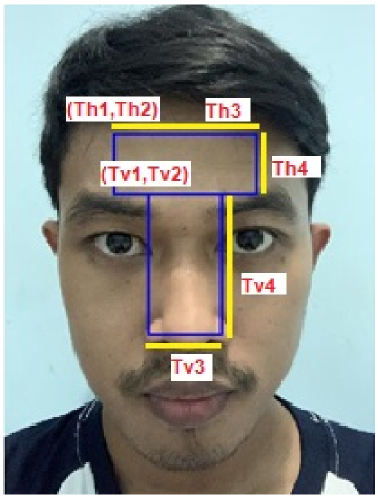

Fig. 11. T-area.

3.4 U-area detection

The U-area detection is also divided into two parts, including the cheek area (right-left) and chin. The cheek area used information for the eye and nose, where the first step involved searching the right cheek using the xy coordinate. These coordinates were obtained by utilizing information from the eyes area.

$$Uka1=e1 \quad (13)$$

Description :
Uka1 = x-coordinate of the upper left corner of the right cheek (U-area)
e1 = x-coordinate of the upper left corner of the eye box

$$Uka2=e2+e4 \quad (14)$$

Description :
Uka2 = y-coordinate of the upper left corner of the right cheek (U-area)
e2 = y-coordinate of the upper left corner of the eye box
e4 = vertical distance of eye box

The width and height of the right cheek were obtained by utilizing information from eyes and nose areas, with the following equation:

$$Uka3=n1-Uka1 \quad (15)$$

Description :
Uka3 = horizontal distance of the top left cheek right cheek (U-area)
n2 = x-coordinate of the top left corner of the nose box
Uka1 = x-coordinate of the top left corner of the right cheek box

$$Uka4=n4 \quad (16)$$

Description :
Uka4 = vertical distance of the upper left box of the right cheek (U-area)
n4 = nose box vertical distance

Detecting the xy coordinates of the left cheek used information obtained from the nose and eyes areas, with the following equation:

$$Uki1=n1+n3 \quad (17)$$

Description :
Uki1 = x-coordinate of the upper left corner of the left cheek (U-area)
n1 = x-coordinate of the top left corner of the nose box
n3 = horizontal distance of nose box

$$Uki2=e2+e4 \quad (18)$$

Description :
Uka2 = y-coordinate of the upper left corner of the left cheek (U-area)
e2 = y-coordinate of the upper left corner of the eye box
e4 = vertical distance of eye box

The width and height of the left cheek were obtained by utilizing the information from the nose and eye areas, with the following equation:

$$Uki3=(e1+e3)-(n1+n3) \tag{19}$$

Description :
Uki3 = left cheek horizontal distance (U-area)
e1 = x-coordinate of the upper left corner of the eye box
e3 = horizontal distance of eye box
n1 = x-coordinate of the top left corner of the nose box
n3 = horizontal distance of nose box

$$Uki2=n4 \tag{20}$$

Description :
Uki4 = vertical distance of left cheek (U-area)
n4 = nose box vertical distance
ok top left eye box
e4 = vertical distance of eye box

The U-area detection results for the right and left cheeks are shown in Figure 12.

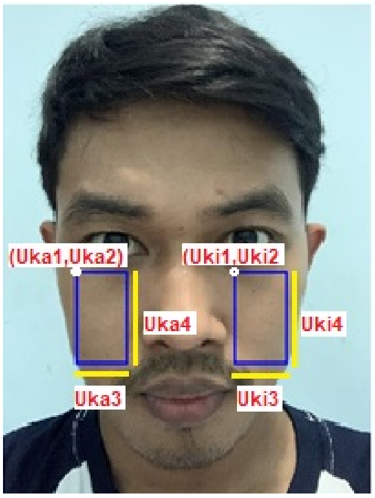

Fig. 12. U-area of the right cheek and left cheek.

Furthermore, the information from the mouth and nose was used to detect the area of the chin. The left border of the chin area is equal to the left border of the mouth area and slightly shifted by 1/6 width of the mouth to the center, Because usually the chin is narrower than the mouth, and the width is about 2/3 the width of the mouth. Additionally, the upper limit of the chin area used the information on the lower boundary of the mouth area (ascend ¼ part), this is because the mouth box extends down and slightly takes up the chin area, which led to the following equation:

$$Ub1=m1+m3/6 \tag{21}$$

Description :
Ub1 = x-coordinate of the top left corner of the chin box (U-area)
m1 = x-coordinate of the upper left corner of the mouth box
m3 = horizontal distance of mouth box

$$Ub2=m2+m4*3/4 \quad (22)$$

Description :
Ub2 = y-coordinate of the top left corner of the chin box (U-area)
m2 = y-coordinate of the upper left corner of the mouth box
m4 = vertical distance of mouth box

The width of the chin is equal to 2/3 of the mouth width, considering that the lower part of the chin is contracting. Also, the height of the chin includes the distance between the lower border of the face slightly down (1/4 the height of the mouth) with the upper chin limit as obtained by the following equation:

$$Ub3=m3*2/3 \quad (23)$$

Description :
Ub3 = chin box horizontal distance (U-area)
m3 = horizontal distance of mouth box

$$Ub4=(f2+f4+m4/4)-Ub2 \quad (24)$$

Description :
Ub4 = vertical distance of chin box (U-area)
f2 = y-coordinate of the top left corner of the face box
f4 = vertical distance of face box
m4 = vertical distance of mouth box
Ub2 = y-coordinate of the top left corner of the chin box (U-area)

If the lower limit of the face area is smaller than the lower limit of the mouth. Since most chins are slightly longer than their mouths by about 25%, the length of the chin is the length of the mouth times 1.25, then Ub4 will be obtained using the following equation:

$$Ub4= m4*1.25 \quad (24)$$

Ub4 = vertical distance of chin box (U-area)
m4 = vertical distance of mouth box

Based on the above equation, the U-area of the chin (bottom) was obtained, as shown in Figure 13.

Fig. 13. Chin area.

When Figures 12 and 13 were combined, the total U-area obtained was as shown in Figure 14.

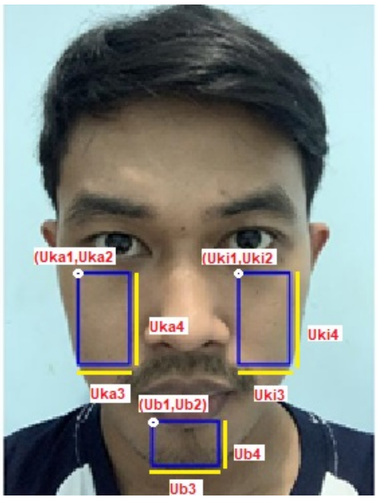

Fig. 14. Total U-area.

When the T and U areas were combined in the same image, the results were as shown in Figure 15.

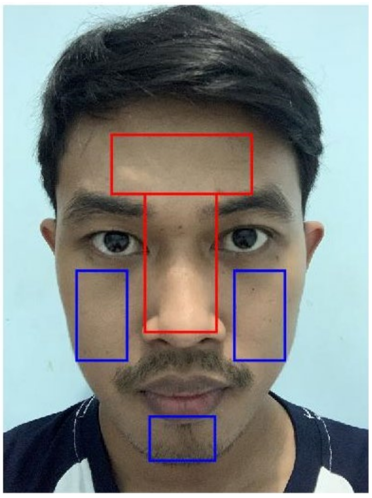

Fig. 15. T-area and U-area.

3.5 Accuracy

To calculate the accuracy of the U and T areas, the segmentation results of the face, applied the precision of the forehead, nose, right-cheek, left-cheek, and chin positions. The detection could be declared correct when all the areas mentioned above are detected. However, if one area is incorrect, the accuracy for the image drops to 80%. If all areas are not detected, the detection accuracy for the U-area and T-area in the image is stated as 0%. Based on the results, all the image samples detected using this method obtained 100% accuracy.

$$Acc = \sum_{i=1}^{c} A_i \frac{1}{C} \quad (7)$$

If n tests are carried out, and the correct one is denoted by 1 while the wrong one is denoted by 0. Then the accuracy is the sum of all the values of 1 and 0 divided by the number of examiners. For example, 10 tests were carried out and 7 were correct, then accuracy = (1+1+0+1+1+0+1+1+0+1)/10=7/10=70%. A denotes the accuracy of detection per image with 5 segmentations comprising the U and T areas in percentage, while C is the number of images detected. The accuracy of 25 trials on different images totals 100%.

4 Conclusion

The vision.cascadeObjectDetector function was instrumental in calculating the U-area, which includes the right and left cheeks along with the chin and the T-area, comprising the forehead and nose. Furthermore, the function accurately detected the face area, including the position of the eyes, nose, and mouth. The success level in determining the two areas adds to 100%, as long as the boundaries of the face, eye, nose, and mouth are given correctly, which highly depends on the frontal position of the image before the camera and the straight neck.

References

[1] Youn SW. Sebum secretion, skin type, and pH. In: Zouboulis C, Katsambas A, Kligman A (eds) *Pathogenesis and Treatment of Acne and Rosacea*. Berlin: Springer, 2014, pp. 299–303.

[2] Ovhal AG, Jerajani HR, Dhurat RS. A comparative study of sebum secretion rates and skin types in individuals with and without acne. *MedPulse Int J Med* 2018; 7: 88–92.

[3] Nouveau-Richard S, Zhu W, Li YH, et al. Oily skin: Specific features in chinese women. *Ski Res Technol* 2007; 13: 43–48.

[4] Park SG, Kim YD, Kim JJ, et al. Two possible classification of facial skin type by two parameters in korean women excretion rate (SER) and skin surface relief (SSR). *Ski Res Technol* 1999; 5: 189–194.

[5] Youn SW, Kim SJ, Hwang IA, et al. Evaluation of facial skin type by sebum secretion: Discrepancies between subjective descriptions and sebum secretion. *Ski Res Technol* 2002; 8: 168–172.

[6] Athanasios N, Pitas I. Facial feature extraction and pose determination. *Pattern Recognit* 2000; 33: 1783–1791.

[7] Vezhnevets V. Face and facial feature tracking for natural human-computer interface. *Proc Graph* 2002; 86–90.

[8] Hadid A, Pietikainen M, Martinkauppi B. Color-based face detection using skin locus model and hierarchical filtering. In: *2002 International Conference on Pattern Recognition*. 2002, pp. 196–200.

[9] Hsu RL, Abdel-Mottaleb M, Jain AK. Face detection in color images. *IEEE Trans Pattern Anal Mach Intell* 2002; 24: 696–706.

[10] Ioannou S, Caridakis G, Karpouzis K, et al. Robust feature detection for facial expression recognition. *EURASIP J Image Video Process*; 2007.

[11] Sohail ASM, Bhattacharya P. Detection of facial feature points using anthropometric face model. In: Damiani E, Yétongnon K, Schelkens P, et al. (eds) *Signal Processing for Image Enhancement and Multimedia Processing*. Boston: Springer, 2008, pp. 189–200.

[12] Wilson JN, Ritter GX. *Handbook of computer vision algorithms in image algebra*. 2nd ed. CRC Press, 2000.

[13] Chen Q, Cham WK, Lee KK. Extracting eyebrow contour and chin contour for face recognition. *Pattern Recognit* 2007; 40: 2292–2300.

[14] Ding L, Martinez AM. Precise detailed detection of faces and facial features. *IEEE Conf CVPR 2008 Comput Vis Pattern Recognit* 2008; 1–7.

[15] Sathyanarayana S, Satzoda RK, Sathyanarayana S, et al. A compute-efficient algorithm for robust eyebrow detection. *Proc IEEE Conf Comput Vis Pattern Recognit Work* 2014; 650–655.

[16] Indriyani, Suciati N, Navastara DA. Automatic determination of t-zone on face area using canny edge detection and hough transformation for skin type classification. *Eur J Sci Res*; 143.

[17] Indriyani, Sudarma M. Automatic segmentation of u-zone area on fafial images using fuzzy edge detection, hough transformation, and viola-jones algorithm. *Int J Adv Sci Technol* 2019; 133: 19–30.