An improved human skin detection and localization by using machine learning techniques in RGB and YCbCr color spaces

Mohammad Mortazavi T. [1] , **Omid Mahdi Ebadati E.** Corresp. [2]

[1] Knowledge Engineering and Decision Science, Kharazmi University, Tehran, Tehran, IRAN, ISLAMIC REPUBLIC OF
[2] Mathematics & Computer Science, Kharazmi University, Tehran, Tehran, IRAN, ISLAMIC REPUBLIC OF

Corresponding Author: Omid Mahdi Ebadati E.
Email address: ebadati@khu.ac.ir

Human Skin Detection is one of the most applicable methods in human detection, face detection and so many other detections. These processes can be used in a wide spectrum like industry, medicine, security, etc. The objective of this work is to provide an accurate and efficient method to detect human skin in images. This method can detect and classify skin pixels and reduce the size of images. With the use of RGB and YCbCr color spaces, proposed approach can localize a Region Of Interest (ROI) that contains skin pixels. This method consists of three steps. In the first stage, pre-processing an image like normalization, detecting skin range from the dataset, etc. is done. In the second stage, the proposed method detects candidate's pixels that are in the range of skin color. In the third stage, with the use of a classifier, it decreases unwanted pixels and areas to decrease the accuracy of the region. The results show 97% sensitivity and 85% specificity for support vector machine classifier.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27488v1 | CC BY 4.0 Open Access | rec: 15 Jan 2019, publ: 15 Jan 2019

An Improved Human Skin Detection and Localization by using Machine Learning Techniques in RGB and YCbCr Color Spaces

Mohammad Mortazavi T.[1], Omid Mahdi Ebadati E.[2]

[1] Department of Knowledge Engineering and Decision Science, Kharazmi University, #242, Somayeh Street, Between Qarani & Vila, Tehran, Iran
[2] Department of Mathematics and Computer Science, Kharazmi University, #242, Somayeh Street, Between Qarani & Vila, Tehran, Iran

Corresponding Author:
Omid Mahdi Ebadati E.[1]
Email address: ebadati@khu.ac.ir

Abstract
Human Skin Detection is one of the most applicable methods in human detection, face detection and so many other detections. These processes can be used in a wide spectrum like industry, medicine, security, etc. The objective of this work is to provide an accurate and efficient method to detect human skin in images. This method can detect and classify skin pixels and reduce the size of images. With the use of RGB and YCbCr color spaces, proposed approach can localize a Region Of Interest (ROI) that contains skin pixels. This method consists of three steps. In the first stage, pre-processing an image like normalization, detecting skin range from the dataset, etc. is done. In the second stage, the proposed method detects candidate's pixels that are in the range of skin color. In the third stage, with the use of a classifier, it decreases unwanted pixels and areas to decrease the accuracy of the region. The results show 97% sensitivity and 85% specificity for support vector machine classifier.

Introduction
Machine Vision (MV) is one of the most powerful and applicable approach in machine learning and artificial intelligence. This approach can analyze images and extract the different features from them. Machine vision applications include variety of areas like: Robot guidance, Food packed or food packet checks, final inspection cells and many more.
Textural and spatial features are often exploited for skin modeling. In an approach proposed from [1], they extracted the textural features from the skin probability maps rather than from the luminance channel. The skin detection outcome was achieved by exploiting the discriminative features extracted from the skin probability maps and using them in a spatial analysis scheme.

There is another research that proposes a model for skin color detection without color transformation. This model is based on the Red Green Blue (RGB) color space system. It is especially suitable to implement on hardware for image processing applications.
In this model for the simplification of computational complexity, statistical method is popularly applied [2]. There is an approach that proposes a dynamic model for skin detection in images. This model consists of two stages: training stage and detection stage. In the first stage, the model is used on skin and non-skin images dataset to calculate the histogram for them. In the second stage, it uses the histogram for segmentation of the image to skin and non-skin pixels [3].
One of the most popular applications of skin detection is the face detection. There is a research that proposed a model for face detection based on skin color likelihood. This algorithm computes the similarity between a color region and the skin color. Both Haar-like feature and local binary pattern (LBP) feature are used to build an accurate cascaded classifier. A boosting algorithm is implemented based on skin color emphasis to localize the face region from a color image [4].
A hybrid model for human skin detection was proposed in another research [5], and the authors used k-mean and multilayer perceptron (MLP) for their model. K-mean is an unsupervised method for clustering [6-8].
There is a research about color space which is a comparative study between HSV and YCbCr color space for human skin detection. This research shows which RGB color space is not preferred for color-based detection and color analysis because of mixing of color (chrominance) and intensity (luminance) information. In YCbCr color space, we can separate chrominance and luminance easily. The result shows the efficiency of YCbCr color space for the segmentation and detection of skin color in color images [9].
In another paper, the research is conducted on a multi-skin color clustering model, by using the HSV color space. In this model, skin regions are extracted using four skin color clustering models, and a skin color correction at the shadow-skin layer is used to improve the detection rate and accuracy [10, 11].
Another research [12], proposed an adaptive skin color filter for skin color detection. This method can find candidate regions for faces or hands in color images. This method consists of two stages. Firstly, a thresholding box in HSV color space, which was t updated adaptively by using a color histogram under the assumption of the area of skin color. Secondly, color vectors inside the thresholding box are classified into two groups: skin and background color vectors. This method was tested with variety of sample images from the Internet.
Two Eigen-based fuzzy c-means (FCM) clustering algorithms to segment the color images, is another research, which operates based on the fuzzy model. This method is created by combining the PCT and FCM methods together to extract the desired color images, and it can find and segment the points, which have the same color with the preferred points [13].
This work was structured to describe the proposed method for human skin detection based on supervised learning. The proposed method can detect a region of interest (ROI) based on skin color in images after detection. The calculations and operators are described in Section 2. Experimental results are presented in Section 3, and Section 4 concludes the paper.

Materials & Methods

Skin detection is one the most important approach in image processing. Detecting and finding skin regions can help us in human detection, face detection, localization, etc. Our aim is to detect regions, which consist of skin pixels and cutting this area for classification steps. In the classification step, support vector machine (SVM) is used to classify the new area pixels as skin and non-skin. This method can detect and localize the human skin region in images. All steps of proposed method are shown in Fig. 1.

Color constancy is the ability to perceive colors of objects, invariant to the color of the light source. An object in a different light source can have different views. In Fig. 2, the same flower is depicted four times, each rendered under a different light source.

Therefore, we need an image which is invariant to variant color illuminations. To deal with this issue, color normalization method is used in the proposed model. In this method, the normalization is used in grey level in each channel. The grey world normalization makes the assumption that changes in the lighting spectrum can be modeled by three constant factors applied to the Red, Green and Blue (RGB) channels of color.

$$\left(\frac{R}{\frac{\kappa}{n}\sum_i R},\frac{G}{\frac{\kappa}{n}\sum_i G},\frac{B}{\frac{\kappa}{n}\sum_i B}\right) \qquad (1)$$

Where k is max $\left(\frac{1}{\frac{\kappa}{n}\sum_i R},\frac{1}{\frac{\kappa}{n}\sum_i G},\frac{1}{\frac{\kappa}{n}\sum_i B}\right)$. This method can help us to detect true colors in different situations [14].

Then, the RGB images must convert to YCbCr color space. Before the start of the proposed model, these conversions to color spaces describe shortly that RGB is a color space originated from cathode-ray tube (CRT) displays. It describes the color as a combination of three colored rays (red, green and blue). Some complexities like high correlation between channels, mixing of chrominance and luminance data make RGB not a very favorable choice for color analysis and color-based recognition algorithms. YCbCr is an encoded nonlinear RGB signal that is used by European television studios. YCbCr is used to separate a luma signal (Y) and two chroma components (C_b and C_r) [15]. In Formula 2, the transformation of RGB to YCbCr is formulated.

$$Y = 0.299R + 0.587G + 0.114B \qquad (2)$$
$$C_r = R - Y$$
$$C_b = B - Y$$

C_b and C_r mostly represent the skin–color reference map that is [16]:

$$77 \leq C_b \leq 127 \; and \; 133 \leq C_r \leq 173$$

In this research, SFA dataset [17] and some other human skin images from the Internet are used. There are 3380 numbers of human skin images with a dimension of 35×35 pixels. In this level, conversion of all images to YCbCr color space, and thereafter calculation of the maximum and

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27488v1 | CC BY 4.0 Open Access | rec: 15 Jan 2019, publ: 15 Jan 2019

minimum of luma and chroma has been done. The results show a little change in chroma range, and luma range by using the proposed method. The results are presented in Formula 3.

$$50 \le lu \le 203 \text{ and } 87 \le C_b \le 127 \text{ and } 127 \le C_r \le 175 \tag{3}$$

This range will be used for nominated skin color in continuation. All pixels that are in this range may be selected and classified as pixels of image for skin and non-skin.

$$mask(x,y) = \begin{cases} 1, & pixel(x,y) \text{ is in range} \\ 0, & else \end{cases}$$

With this $mask$, there will be an almost big region that contains the skin pixels. Now, we can calculate the Region Of Interest (ROI). The obtained ROI contains a region that has at least one skin pixel.

In the next stage, we create a dataset of features that exist in skin and non-skin images in RGB color space. This dataset contains 3369 positive and 500 negative images. We extract features like: mean, standard deviation and skewness of images for each channel.

The three-color moments can then be defined as (Formulas 4-6):

$$E(i) = \frac{1}{n}\sum_{j=1}^{n} p_{ij} \tag{4}$$

$$Std(i) = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(p_{ij} - E(i)\right)^2} \tag{5}$$

$$Skw(i) = \sqrt[3]{\frac{1}{n}\sum_{j=1}^{n}\left(p_{ij} - E(i)\right)^3} \tag{6}$$

The i^{th} color channel is defined at the j^{th} image pixel as p_{ij}. A dataset of features creates this calculation.

$$[E(R), E(G), E(B), Std(R), Std(G), Std(B), Skw(R), Skw(G), Skw(B)]$$

Each feature vector has nine columns for each image.

In the next stage, a linear support vector machine (SVM) trains on positive and negative sample images for classification. It includes a set of input examples $\{(x_i,y_i)\}_{i=1,\ldots,m}$ where $x_i \in \mathbb{R}^n$ all are the inputs and $y_i \in \{-1,+1\}$ are the corresponding outputs. The regularization term is $\frac{1}{2}\|w\|^2$ and error variables are $\xi = (\xi_1,\ldots,\xi_m)^t$. Optimization problem is shown in Formula 7.

$$\min_{w,e} \frac{1}{2}w^t w + Ce^t\xi \tag{7}$$

$$s.t.\ y_i\left(w^t\phi(x_i) + b\right) \ge 1 - \xi_i$$

$$\xi_i \ge 0\ \forall i = 1,\ldots,m$$

Where $C > 0$ is the regularization parameter, for this primal problem, SVM solves its Lagrangian dual problem (Formula 8).

$$\min_{\alpha} \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\alpha_j y_i y_j K(x_i,x_j) - \sum_{i=1}^{m}\alpha_i \qquad (8)$$

$$s.t. \sum_{i=1}^{m} y_i\alpha_i > 0$$

$$0 < \alpha_i < C$$

Where $K(.,.)$ is the linear kernel function, which means linear function $K(x,x_i) = x^T x_i$ [18, 19]. Now, we can apply this SVM on each pixel on the ROI to improve the accuracy. For an accurate detection, implementation of this work is used in 1×1 windows. The proposed model results are presented in the next section.

Analysis and Design of the Proposed Model

In this research, SFA dataset and many other images from the Internet are used for training data. Figure 3 is shown as a short view of SFA dataset. Furthermore, we need a non-skin dataset for training the designed SVM. Figure 4 is shown as a short view of the non-skin dataset, which is collected from the internet. All of the samples' resolution is 35×35.

If the image in RGB color space normalizes, then we can find the area of color space that consists of the human skin color. For this purpose, Formula 9 is used to do normalization of the RGB color space.

$$r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B}, b = \frac{B}{R+G+B} \qquad (9)$$

$$r + g + b = 1 \Longrightarrow b = 1 - r - g$$

Now, we can plot the RGB skin and non-skin color space with r and g axes. Figure 5 is illustrated in this area.

In the next stage, we use an image that contains the skin region for grey world normalization. We use a random image from the Internet. In Figure 6, we can see the original image and the normalization of the image.

After that, the grey world normalization of image is converted into the YCbCr color space, and we can take a mask of it from the skin and non-skin human color . Now, with the help of morphological operators we can fill the small holes (Fig. 7).

Filled images should be filtered with median filters in this stage (Fig. 8) and the next morphological closing of the opening of the image is calculated. These works can reduce the white small holes.

Now, the region contains skin, as well as small unwanted regions and noises. For clearing this noise, the regions that have the great area with 900 pixels are selected. To continue, the selected area which contains the skin is shown as (Fig. 9).

For the next step, this method reduces the image to an area with skin pixels, but it contains many other objects. For solving this problem and showing the exact skin region, we use SVM to classify it as skin and non-skin. The result of SVM classifier has high accuracy, and it only contains the skin pixels. The result is shown in Fig. 10.

In continuation, this method is applied on several images and all of them have good results with high accuracy in skin detection. We can create a contingency table that describes all the different combinations of correct and incorrect classifications (details shown in Table 1).

With this data, the proposed method has 97% ($\frac{3267}{3369}$) sensitivity and 85% ($\frac{426}{500}$) specificity, which is a high accuracy for the skin detection, and can classify the human skin pixels with the best performance. The proposed method is implemented on the other images, and the result is presented in figure 11.

Conclusions

With special attention to the image-processing application, machine vision, image processing and machine learning techniques have more important roles in day to day functioning. Skin detection technique can be widely used for face detection, human detection, etc. In this paper, an approach for human skin detection is proposed, by implementing morphological operator and SVM methods. At first, morphological operator used to detect a region that contains human skin pixels, and it was cut to reduce the size of image. The reduced image is a suitable input for a classification. In the next stage, SVM is used to increase accuracy and decrease noise and unwanted pixels. This model with 97% sensitivity and 85% specificity is one of the most accurate models in skin detection, which is presented in the analysis section. For the future of this work, the proposed method can combine with other algorithms like ANN, genetic algorithm and the heuristic algorithm to speed up the method.

References

[1] M. Kawulok, J. Kawulok, and J. Nalepa, "Spatial-based skin detection using discriminative skin-presence features," *Pattern Recognition Letters,* vol. 41, pp. 3-13, 2014.

[2] Y.-H. Chen, K.-T. Hu, and S.-J. Ruan, "Statistical skin color detection method without color transformation for real-time surveillance systems," *Engineering Applications of Artificial Intelligence,* vol. 25, pp. 1331-1337, 2012.

[3] H.-M. Sun, "Skin detection for single images using dynamic skin color modeling," *Pattern recognition,* vol. 43, pp. 1413-1420, 2010.

[4] Y. Ban, S.-K. Kim, S. Kim, K.-A. Toh, and S. Lee, "Face detection based on skin color likelihood," *Pattern Recognition,* vol. 47, pp. 1573-1585, 2014.

[5] H. K. Al-Mohair, J. M. Saleh, and S. A. Suandi, "Hybrid Human Skin Detection Using Neural Network and K-Means Clustering Technique," *Applied Soft Computing,* vol. 33, pp. 337-347, 2015.

[6] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image Segmentation Using K-means Clustering Algorithm and Subtractive Clustering Algorithm," *Procedia Computer Science,* vol. 54, pp. 764-771, 2015.

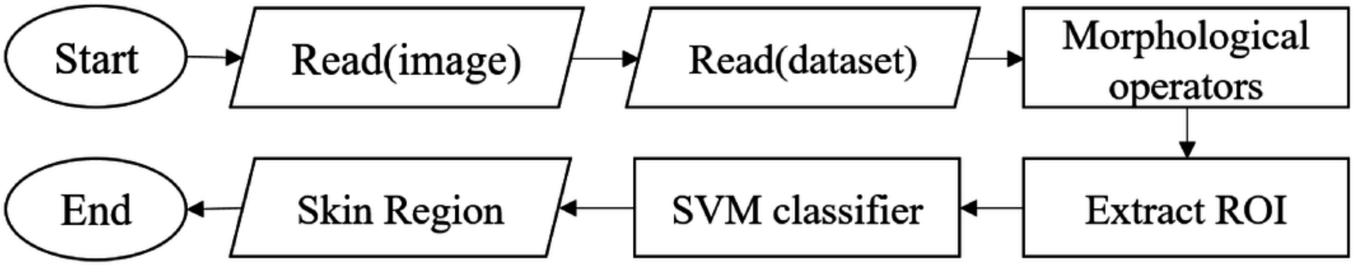
[7] P. Arora and S. Varshney, "Analysis of K-Means and K-Medoids Algorithm For Big Data," *Procedia Computer Science,* vol. 78, pp. 507-512, 2016.

[8] S. Shahrivari and S. Jalili, "Single-pass and linear-time k-means clustering based on MapReduce," *Information Systems,* vol. 60, pp. 1-12, 2016.

[9] K. B. Shaik, P. Ganesan, V. Kalist, B. Sathish, and J. M. M. Jenitha, "Comparative Study of Skin Color Detection and Segmentation in HSV and YCbCr Color Space," *Procedia Computer Science,* vol. 57, pp. 41-48, 2015.
[10] S. A. Naji, R. Zainuddin, and H. A. Jalab, "Skin segmentation based on multi pixel color clustering models," *Digital Signal Processing,* vol. 22, pp. 933-940, 2012.
[11] R. Khan, A. Hanbury, J. Stöttinger, and A. Bais, "Color based skin classification," *Pattern Recognition Letters,* vol. 33, pp. 157-163, 2012.
[12] K.-M. Cho, J.-H. Jang, and K.-S. Hong, "Adaptive skin-color filter," *Pattern Recognition,* vol. 34, pp. 1067-1073, 2001.
[13] J.-F. Yang, S.-S. Hao, and P.-C. Chung, "Color image segmentation using fuzzy C-means and eigenspace projections," *Signal Processing,* vol. 82, pp. 461-472, 2002.
[14] J. M. Buenaposada and B. Luis, "Variations of Grey World for face tracking," *Image Processing & Communications,* vol. 7, pp. 51-61, 2001.
[15] V. Vezhnevets, V. Sazonov, and A. Andreeva, "A survey on pixel-based skin color detection techniques," in *Proc. Graphicon*, 2003, pp. 85-92.
[16] D. Chai and K. N. Ngan, "Face segmentation using skin-color map in videophone applications," *Circuits and Systems for Video Technology, IEEE Transactions on,* vol. 9, pp. 551-564, 1999.
[17] J. P. B. Casati, D. R. Moraes, and E. L. L. Rodrigues, "SFA: A human skin image database based on FERET and AR facial images," in *IX workshop de Visao Computational, Rio de Janeiro*, 2013.
[18] M. Tanveer, "Robust and sparse linear programming twin support vector machines," *Cognitive Computation,* vol. 7, pp. 137-149, 2015.
[19] W. Fenghua, X. Jihong, H. Zhifang, and G. Xu, "Stock Price Prediction based on SSA and SVM," *Procedia Computer Science,* vol. 31, pp. 625-631, 2014.

Figure 1

The view of workflow



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27488v1 | CC BY 4.0 Open Access | rec: 15 Jan 2019, publ: 15 Jan 2019

Figure 2

The same flower under different light source

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27488v1 | CC BY 4.0 Open Access | rec: 15 Jan 2019, publ: 15 Jan 2019

Figure 3

A short view of SFA dataset

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27488v1 | CC BY 4.0 Open Access | rec: 15 Jan 2019, publ: 15 Jan 2019

Figure 4

A short view of non-skin dataset

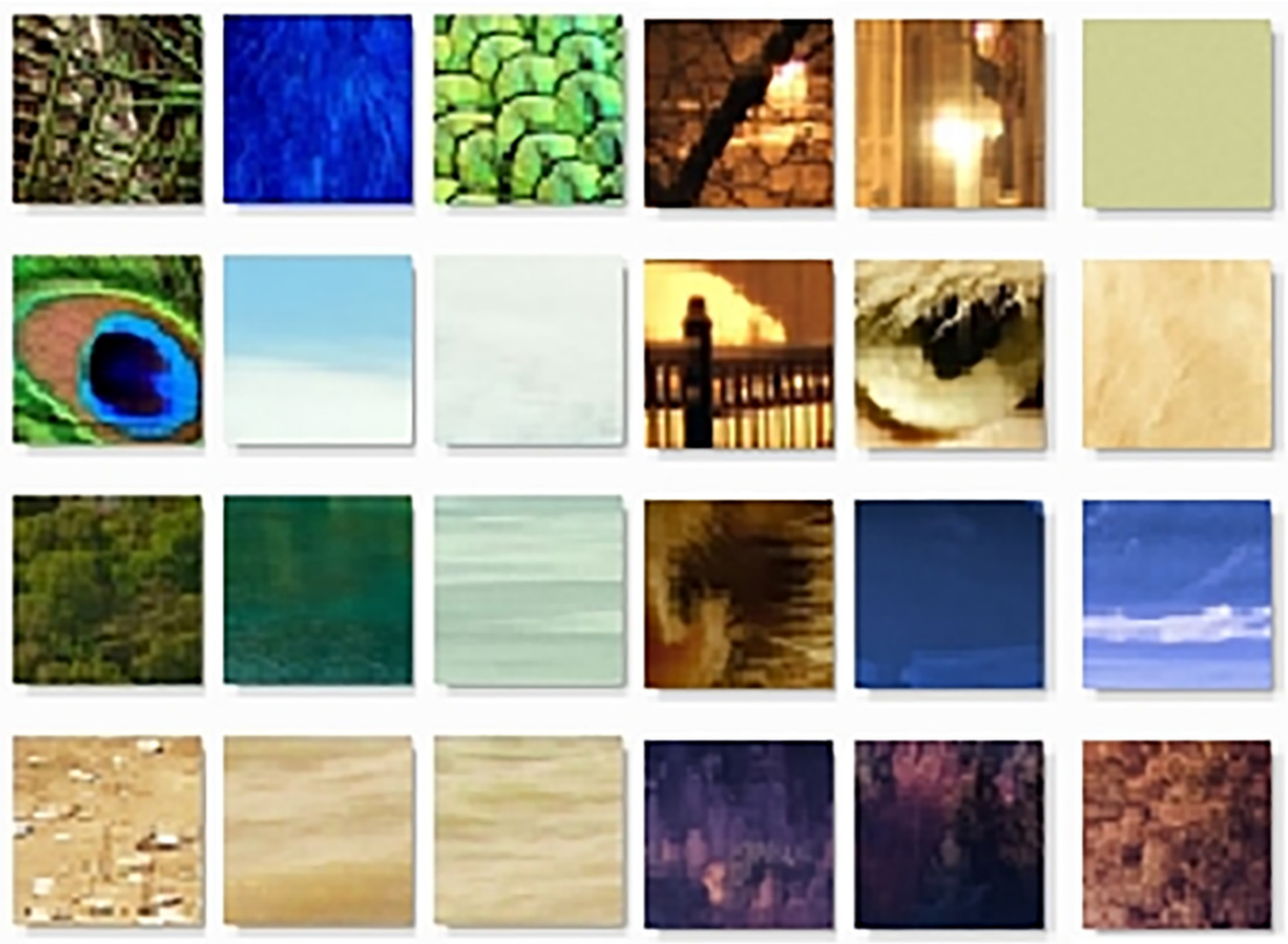


Figure 5

Skin and non-skin area in 2D space (green and red)

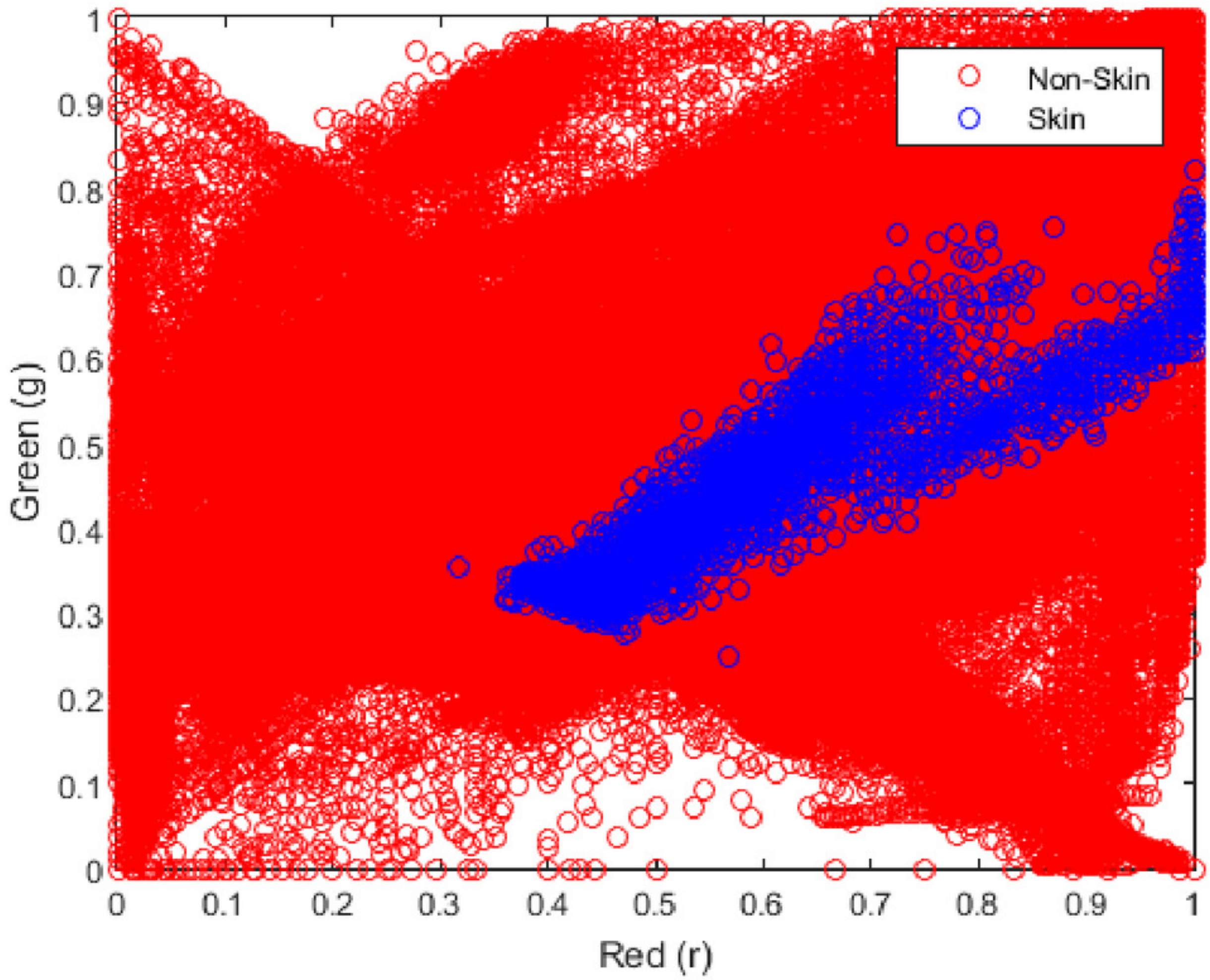


Figure 6

(A) is the original image; (B) grey is the normalization of it

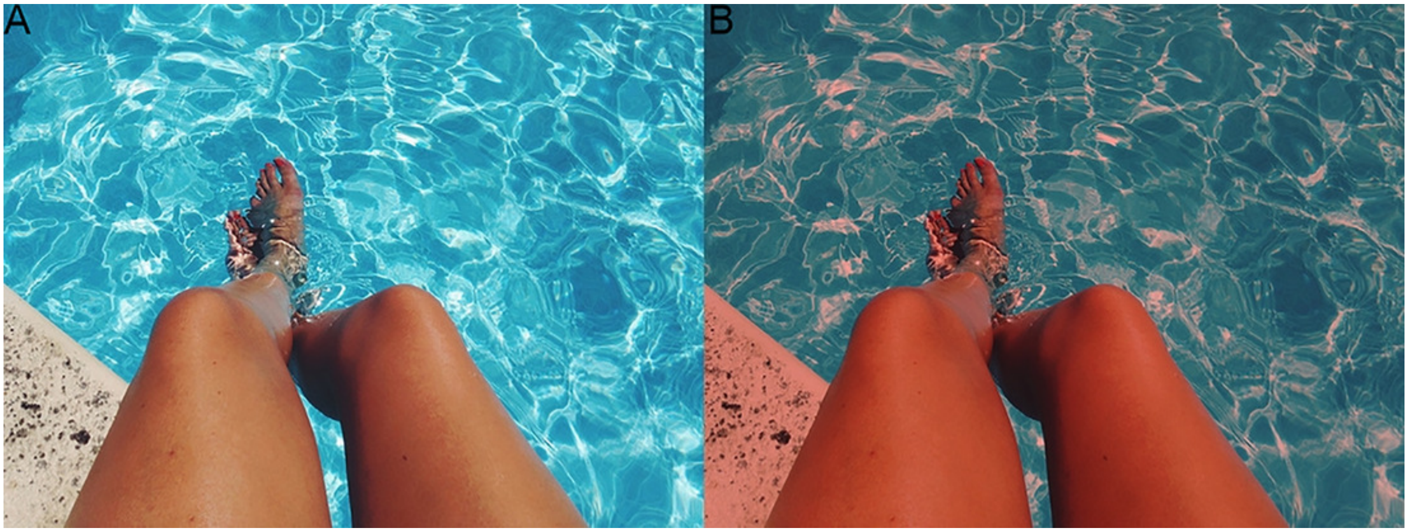


Figure 7

(A) mask image, (B) filled small holes image in the white area

**Note: Auto Gamma Correction was used for the image. This only affects the reviewing manuscript. See original source image if needed for review.*

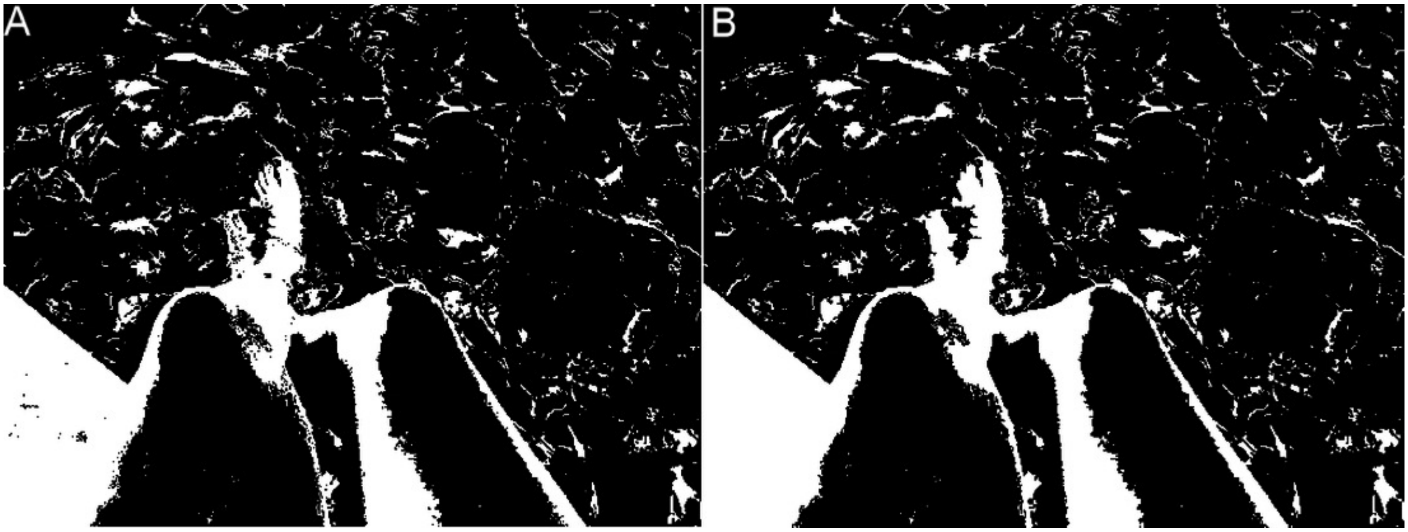


Figure 8

(A) median filter applied to image, (B) morphological closing of the opening of the image is calculated

**Note: Auto Gamma Correction was used for the image. This only affects the reviewing manuscript. See original source image if needed for review.*

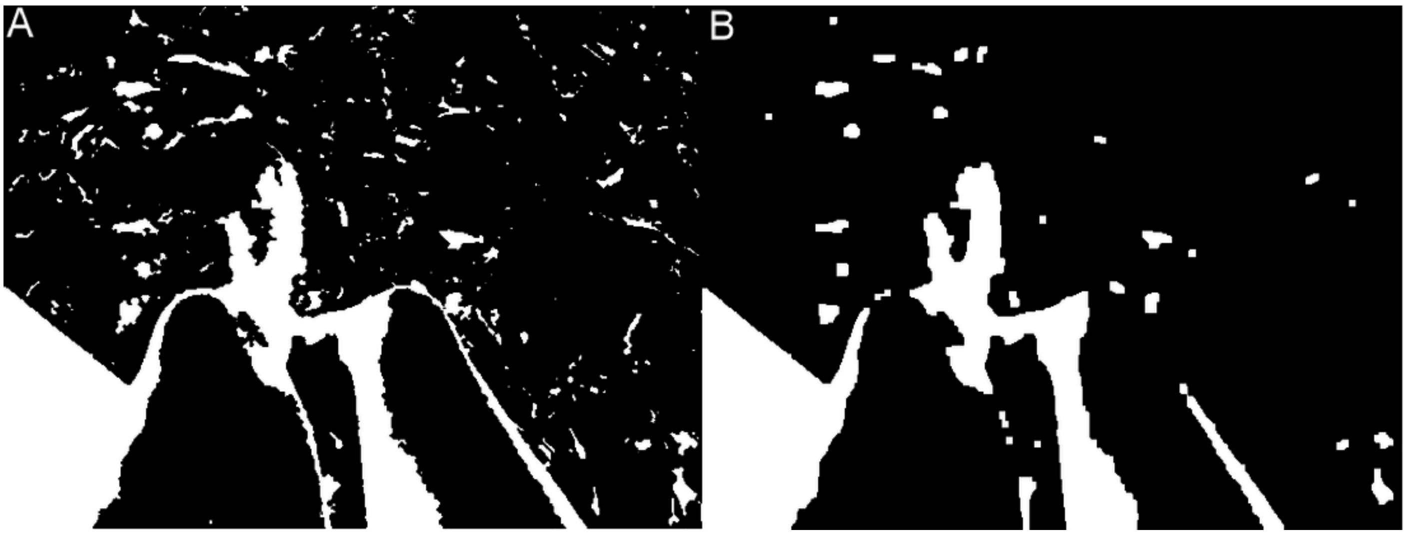


Figure 9

(A) a cleared image from unwanted noise, (B) the region that contains the skin pixels

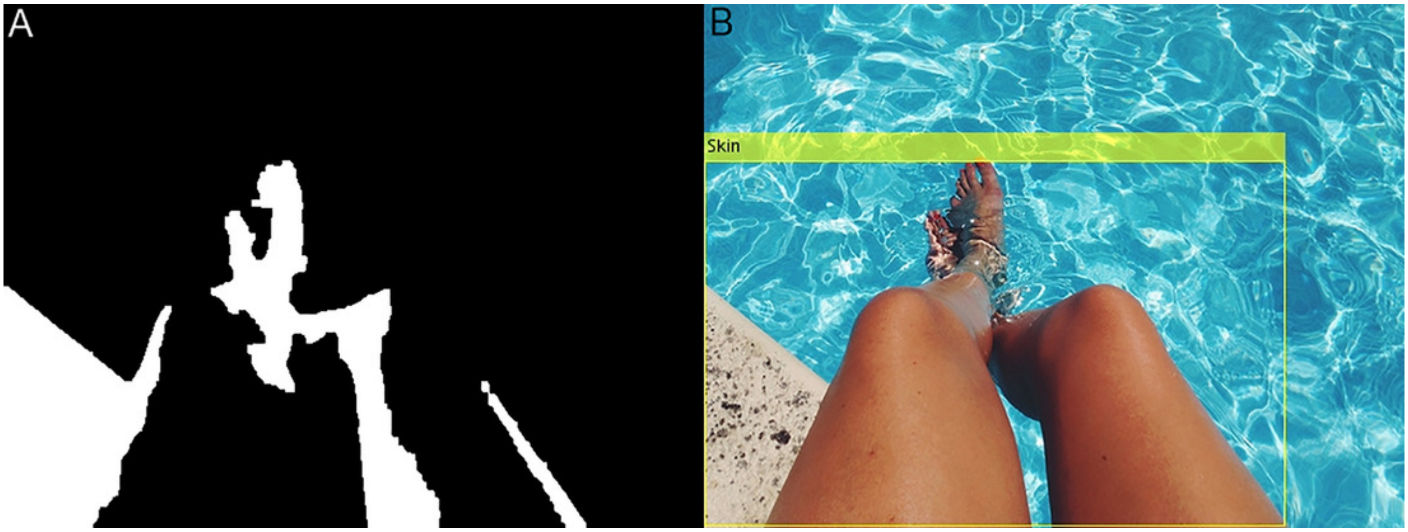
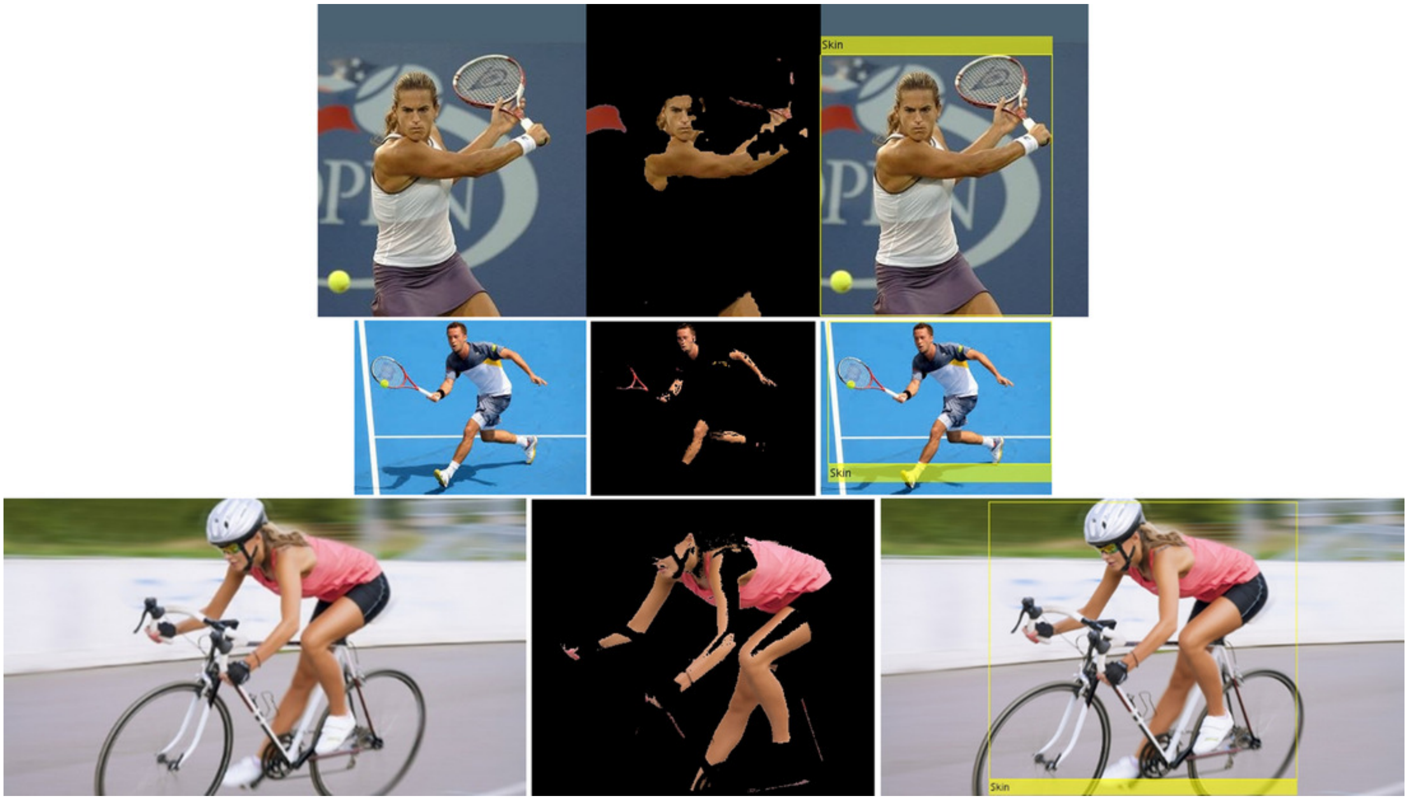


Figure 10

The final result of the proposed model

Figure 11

Several examples of proposed method



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27488v1 | CC BY 4.0 Open Access | rec: 15 Jan 2019, publ: 15 Jan 2019

Table 1 (on next page)

Contingency table

		Gold Standard	
Classification Outcome		Skin (Positive)	Non-Skin (Negative)
	Skin (Positive)	TP=3267	FP=74
	Non-Skin (Negative)	FN=102	TN=426

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27488v1 | CC BY 4.0 Open Access | rec: 15 Jan 2019, publ: 15 Jan 2019