ISSN 2278-3091
Volume 14, No.1, January - February 2025
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse011412025.pdf
https://doi.org/10.30534/ijatcse/2025/011412025

A Robust Framework for Detection of Human Faces in Clutter Images

Mokhtar H. Mohamed
Assistant Professor, Department of Natural and Applied Sciences, Qassim Applied College, Qassim University, Saudi Arabia, h.hassien@qu.edu.sa
Assistant Professor Department of Computer and Information Systems, Sadat Academy, Assiut, Egypt

Received Date: December 11, 2024 Accepted Date: January 14, 2025 Published Date: February 06, 2025

ABSTRACT

Automatic finding of human face from a complex background image is a challenging task in computer vision. This task is the first step in a large number of applications such as security control, human computer interface, face recognition, and image database management. This paper proposes a complete framework for face detection in clutter color images in presence of varying lighting conditions as well as presence of complex background. The proposed framework relies in using a group of neural network classifiers. These classifiers are applied sequentially to detect faces in the image. Initial skin regions over the entire image are detected using one of these classifiers by labeling the image pixels as skin/ non-skin pixels. Skin regions are classified into face/non-face candidates using the correlation between skin region and a prototype face-outline shape image. The facial features mouth and eye are detected within that face candidate using another two classifiers, one for each feature. Finally, the face candidate is verified by applying some face geometrical constrains. Experimental results demonstrate that the proposed framework can be used as a face detection tool over wide range of facial variations in color, position, scale, orientation and expressions.

Key words: Face detection, Luminance compensation, Skin tone, Facial feature, and Geometrical rules.

1. INTRODUCTION

Finding the human faces in an image is a problem that has gained importance in the last decade. It is one of the visual tasks, which humans can do effortlessly. However, in computer vision this task is challenging. The goal of face finding process is to determine whether or not there are any human faces in the image, and if present return their location and spatial extend. The solution of the problem involves segmentation, extraction, and verification of faces and possibly facial features from an uncontrolled background.

There are many closely related problems to face detection, namely facial feature detection, face tracking, teleconferencing [1], video surveillance and security control [2], human computer interface [3], face recognitions [4] and image database management [5]. All of these problems involve face detection as the first processing step. The face detection problem can be viewed as a two-class (face versus non-face) classification problem. There are many factors that challenge it, such as:

- **Variations in position:** The faces in the image vary due to rotation, and translation.
- **Facial Expression:** The appearance of a face is largely affected by the expression on the face. Also the presence or absence of additional features like glasses adds to this variability.
- **Complex Background:** The background which defines the profile of the face cannot be ignored.
- **Occlusion:** Other objects may occlude the face or a group of faces.
- **Imaging conditions:** When the image is formed, factors such as lighting and camera characteristics affect the appearance of a face.

Numerous methods have been proposed over the recent years to detect faces in single intensity (grayscale) or color images. These methods can be classified to the following categories.

a. **Knowledge based methods** [6]: These methods are rule based methods that describe a face based on rules.

b. **Feature invariant methods** [7-10]: These methods extract features (e.g. eyes, nose, and mouth) and then use them to detect a face.

c. **Template matching methods** [11–12]: Several standard patterns of a face are stored to describe the face as a whole or the facial features separately.

d. **Appearance based methods** [13-17]: In these methods, statistical and machine learning techniques are used to learn the characteristics of face and non-face images from examples.

The proposed framework starts with an algorithm which is implemented as a group of neural networks. These networks are used to classify the human face features such as skin, mouth and eye. The classification is based on a pixel color. The algorithm detects the skin over the entire image to generate a number of face candidates. Then mouth and eyes are detected inside the candidates. Geometrical feature constrains and pointing system between face features are used to verify the face candidate. The proposed framework is able to handle a

wide range of variations in color images. The present paper is organized as follows: Section 2 describes the face detection algorithm. Section 3 presents the results of the proposed

2. FACE DETECTION ALGORITHM

An overview of the face detection algorithm is depicted in Figure 1, which contains three major modules: 1) Finding face candidate, 2) Finding facial features, and 3) Verifying the face candidate.

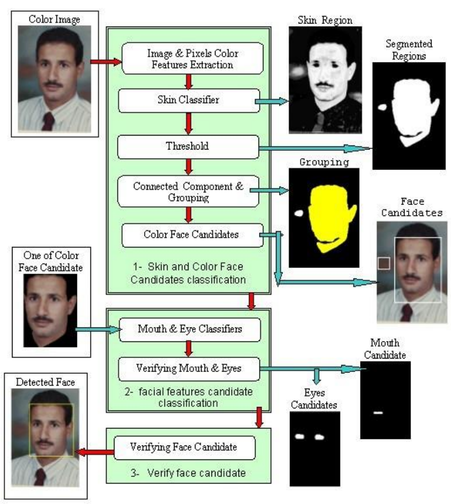


Figure 1: Flow chart for face detection algorithm.

In this algorithm the neural network is used as a pixel color classifier because of its power in classification field. In this algorithm, 14 parameters are used to classify the skin color pixel and other facial features. Each of these parameters has direct effects on the appearance of the image. The detected skin pixels are grouped in regions. Each region is correlated with a prototype face outline shape model to eliminate the non-face shape regions. The mouth and eyes are detected within these regions. Geometrical rules and pointing system are used to verify faces inside the image.

2.1 Classifiers Design

Various classifiers are designed using feed forward neural networks type with back . The procedure for achieving this preparation is passed through two main steps 1) Collecting the training sets, and 2) Training the Neural Network. Over 300 images are used from different sources such as, family celebrations, WEB news, video stream and so on.

2.1.1 Skin Classifier (SC) Design

While face detection based on color seems like an obvious choice to human eyes, identifying the color unique or dominant to the face region is not that obvious for machine based techniques. Different people have skin of different colors ranging from fair to dark skin tones. As a result, it is very difficult to find a region or regions in RGB space that can be defined as a skin color. In this paper the two major color spaces RGB and HSV are used in classifying the face skin-tone and facial features (mouth and eyes). This choice is due to the following reasons:

a. **RGB** (Red, Green, Blue) color space is the color space used in cameras. but not enough for detecting of a wide range of skin-tone color. For that reason, HSV color space is used with RGB.

b. **HSV** (Hue, Saturation, Value) color space imitates the medical fact that variation in human skin color is due to the melanin pigment (H) and its saturation (S) in the skin. The (value) component represents by the intensity value in the pixel. The skin colors form a tight cluster in the H and S space and are distributed in intensity plane.

Selecting the skin samples from each image is carried out by extracting the color features components listed in Table 1 for the entire image. Then, color feature components listed in Table 2 are extracted from each sample.

The training skin set contains over (30,000) of such vectors. SC is created by training the previous feed forward neural network.

Table 1: Image Color Features Components

No.	Feature	Description
1	Rav.	average of Red in the image.
2	Gav.	average of Green in the image.
3	Bav.	average of Blue in the image.
4	Hav.	average of Hue in the image.
5	Sav.	average of Saturation in the image.
6	Vav.	average of Intensity in the image.
7	HImax.	Average Hue at pixels that has over 95 % of maximum intensity.
8	SImax.	average saturation at pixels that has over 95 % of maximum intensity.

2.1.2 Mouth Classifier (MC) Design

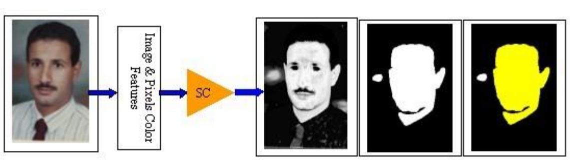
Mouth is an important feature in the face because of its size and its color contrast. The mouth samples are collected by the same manner as in skin samples set. But the main difference is that the mouth samples are taken from images that have faces only and black background. Experimentally, it is found that the most significant color features are:
(**R, G, R/B, H, S, Rav, Gav, Bav, Hav, Sav, HImax, SImax**). **HImax** and **SImax,** are extracted from the main image not from the face image. The mouth training set contains over (12,000) patterns. MC is created by training a feedforward neural network with previous training mouth set.

Table 2: Pixel Color Features components

No.	Feature	Description
1	R	red color component in the pixel.
2	G	green color component in the pixel.
3	B	blue color component in the pixel.
4	H	hue color component in the pixel.
5	S	Saturation color component in the pixel.
6	V	Intensity value color component in the pixel.

2.1.3 Eye Classifier (EC) Design

Eye is a popular classification facial feature in face detection methods. Human eyes have many colors as black, blue, green and honey-color. Selecting eye sample technique is the same as mouth selecting in previous section. The color features components are the same in the MC and consequently the training pattern length is the same as in MC. Creating EC is also done by training a feedforward neural network with the same construction using eye training set.



(a) (b) (c) (d) (e) (f)

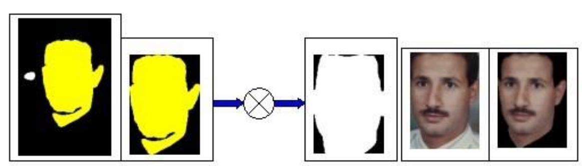
Figure 2: Human skin detection.
(a) color image.
(b) Color Feature vectors.
(c) Skin Classifier.
(d) Detected skin regions.
(e) Threshold image.
(f) Labeled image produced by connected component process.

2.2 Face Candidate (FC) Segmentation

Finding and locating the human skin is the main step in many of face detection methods. Also many of the Human Computer Interface (HCI) applications are based on the assumption that the areas of human skin are already detected and located. The algorithm here achieves skin detection task in the following steps (see Figure 2):

a. Reading the image and resize it to (360 X 480). The goal of this action is to reduce the time and memory size consumed in skin detection process.
b. Extraction Image and pixels color features as listed in Tables-1 and 2.
c. Classification of skin pixel is done by feeding the features from the last step to the input of the SC.
d. The gray-scale image in the last step is then converted to a binary image using an appropriate threshold (experimentally). Regions with 1 values represent skin region and those with 0 values represent the background. A connecting component technique is used to grouping the skin region as shown in Figure (2).
e. Extracting the color face candidate (FC) is carried out by cropping the rectangle that contains only one region and treating it as a separate image and referring to it by Binary Skin Image (BSI).

The face model Figure (6) is extracted from a real human face. According to the correlation coefficient value with the BSI will be accepted if the value exceeds a suitable threshold or reject if not. The desired FC image results from multiplying CSI with BSI leaving the FC with black background as shown in Figure 3.



(a) (b) (c) (d) (e) (f)

Figure 3: Color face candidate detection.
(a) Labeled image.
(b) One of the labeled region.
(c) Correlation process.
(d) Prototype face model.
(e) Cropped face candidate from original image.
(f) face candidate without background.

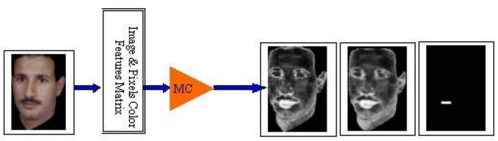
2.3 Facial Feature Detection

The skin region and its shape is important feature but it is not enough to say that the region is a face. For that reason, verification process needs some additional face features like mouth, eyes, brows and nose.

2.3.1 Mouth Detection

The extracted FC in section 2.2 is treated as a separate image as mentioned before. The color features components for that image and individual pixel are extracted by the same manner in FC detection section. The following color features:

(**R, G, R./B, H, S, Rav, Gav, Bav, Hav, Sav, HImax, SImax**) are arranged in a vectors with 12 elements. The vectors are applied on the input of MC. Morphological dilation process is used on the output image to enhance mouth tone region. The output grayscale image is converted to a binary image using a suitable threshold see Figure 4.



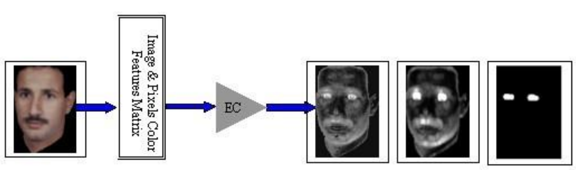
(a) (b) (c) (d) (e) (f)

Figure 4: Mouth detection.
(a) face candidate with black background.
(b)Transformed matrix. (c) Mouth classifier.
(d)Detected mouth regions.
(e)Mouth region morphological process.
(f) threshold mouth candidate.

2.3.2 Eye Detection

The same steps used in mouth detection process are carried out in the eye detection process. The following color features components:

(R, G, B/ R, H, S, Rav, Gav, Bav, Hav, Sav, HImax, SImax) are arranged in vectors 12 elements These vectors are applied on the input of EC. As in mouth detection, morphological dilation is used on the output image from EC to enhance eye-tone region. Converting the gray-scale image to a binary image using suitable threshold. The region with the 1 value in the binary image represents the eye candidates as shown in Figure 5.

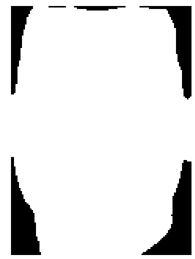


(a) (b) (c) (d) (e) (f)

Figure 5: Eyes detection.
(a) face candidate without background.
(b) Transformed matrix.
(c) Eye classifier.
(d) Detected eye regions.
(e) Eye regions after morphological process.
(f) Threshold eye candidates.

2.4 Verification

Human face is a regular object and its features (mouth, eyes) are arranged in a certain spatial manner. The diagram in the Figure 6 shows the binary face and Figure 7 represents the layout of the mouth and eyes and also labels the distances between them. Some rules for these distances are extracted experimentally from real faces pictures. Because of the human face pose specially the rotation, the following rules are valid only for face candidate that has rotation around 60 degrees right or left. So for simplicity the algorithm assumes only the face candidate that has the rotation range from 0 to 60 degree.



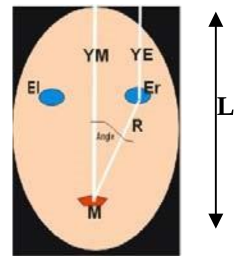


Figure 6: Binary face. **Figure 7:** Facial features layout

2.4.1 Verify Mouth and Eye candidates

The distances and angels between mouth and eye and also between them and the boundary of the face are determined as follows: (see Figures 6, 7).

a. The height of the face candidate (L)
b. The distance between the mouth candidate center and the upper border of the face candidate (YM).
c. The distance between the eye candidate center and the upper boarder of the face candidate (YE).
d. The diagonal distance between the mouth candidate center and the eye candidate center (R).
e. The angle between YM and R (Angle). Naturally there is no fixed distance between mouth and eye.

Naturally there is no fixed distance between mouth and eye because of the variance in human face tallness and breadth. The male and female faces are different in some characteristics. So, the proposed algorithm uses the percentages between these distances instead of the distances itself. Experimentally these percentages are found in a narrow range and are used as verification rules. The proposed algorithm verifies the mouth and eyes candidates according to their spatial arrangement inside the human face using the verification rules listed in Table 3.

Table 3: Mouth and Eye Verification rules.

No.	Rule	Notes
1	YM>0.5L	Mouth at lower half
2	YE<0.5L	Eye at upper half
3	$0.35 \le \frac{YM - YE}{YM} \le 0.75$	Experimental
4	$0 \le Angle \le 60$	Experimental
5	$0.45 \le \frac{R}{YM} \le 0.65$	Experimental

2.4.2 Verifying Face Candidate

A pointing system is used to verify the overall face candidate. The pointing system in brief is similar to a voting process. In the present work, each face candidate feature is given weighting points. The face candidate that has a score greater than a specific threshold is classified as a face in the image, otherwise it is rejected. In future work the pointing system is enhanced using another technique for giving the weighting points.

3. EXPERIMENTAL RESULTS

The proposed framework has been tested using an image database that contains over 300 images. These images are collected from different sources as, web, families' celebration, and video streams. These images contain multiple faces with variations in color, position, scale, orientation and facial expressions as shown in the following figures. The proposed framework is able to find the human faces with black skin-tone and faces with bright skin-tone as shown in figures (8). Human faces with facial expressions and also in presence of eyeglasses can be detected as shown in Figure (9). Frontal and no frontal faces as long as the eyes and mouth are visible in half-profile views are also detected as demonstrated in Figures (10). Human faces with a head cover and facial hair are detected as shown in Figure (11). The algorithm can also find more than one face in the image as appears in Figures (11, 12). A summary of the results is listed in Table 4.

Table 4: Detection Results

Image Database	Our Image Data Base	HHI Image Data Base
No. Images	300	206
No. Faces	786	206
No. Detected Faces	724	182
No. False Detected Faces	118	25
Detection Ratio (%)	92.11	88.34
False Detection Ratio(%)	15.01	12.14

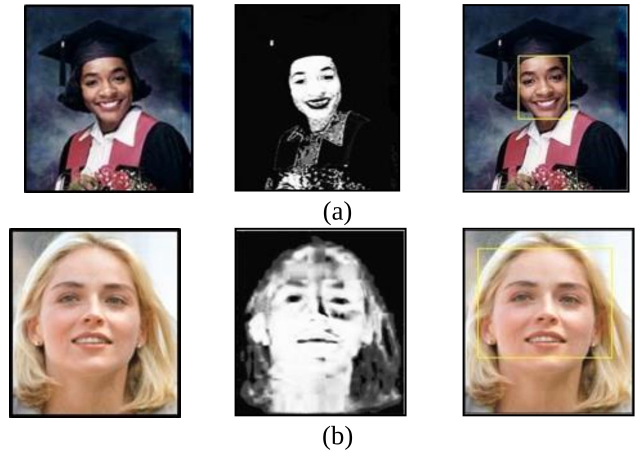


Figure 8: Face detection examples contains. (a) dark skin-tone face image. (b)western skin-tone face image. Each example has the original image, skin-tone regions and the original overlaid with detected face.

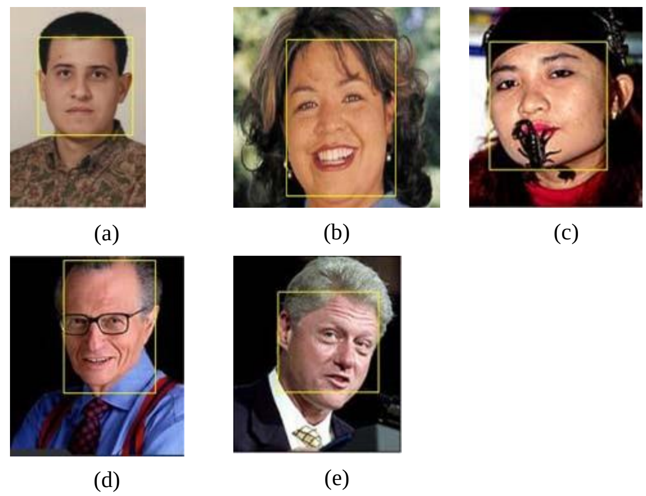
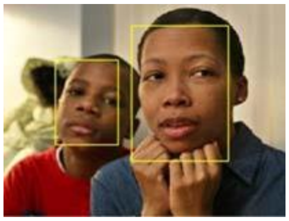


Figure 9: Face detection examples in presence of facial variations and eye glasses

Size: 250x179 FLOPs: 2180140
Size:220x168 FLOPs :8645823
Size: 451x299 FLOPs: 56854112
Size: 120x150 FLOPs: 2533823
Size: 300x225 FLOPs: 32802541
Size:150x190 FLOPs: 3522468

Figure 10: Face detection results on non-frontal view faces with a sign on the facial features.


(a)


(b)


(b)


(c)

Figure 11: Face detection results on four photos. Each image contains more than one face
(a) Size:450x360 : FLOPs:76877112
(b) Size: 649x669 : FLOPs:60456663
(c) Size: 217x297 : FLOPs: 25895098
(d) Size: 568x365 : FLOPs: 36366252


(a)


(b)

Figure 12: Face detection results on a family photo. Each image contains multiple faces. False negatives are due to extreme lighting conditions and background is similar to human skin.

(a) Size: 568x366: FLOPs:27589375
(b) Size: 564x365: FLOPs:22853775

4. CONCLUSION

A color pixel-based face finding framework in clutter images is presented in this paper. The proposed framework relies on using pre-trained neural network classifiers. The training sets are based on the pixel color not on the window that contains face which is the common case in previously proposed algorithms. The framework first finds the skin regions over the entire image and then generates face candidates based on the spatial arrangement of these skin patches. Mouth and eyes are detected inside the face candidate and are verified using the geometrical rules. The overall face candidate is verified using a point system. The proposed framework overcomes the lighting conditions and color bias by appending the effective color features components in the training sets. The proposed framework combines color spaces (RGB and HSV) and also combines color features with geometrical features which enhance the detection results. Results on several image collections have been presented.

REFERENCES

1. M. S. Yehea Al Atrash, **Detecting and Counting People's Faces in Images Using Convolutional Neural Networks**, IEEE, no. Palestinian International Conference on Information and Communication Technology (PICICT), p. 116, 2021.
2. T. Nguyen-Trang, **A new efficient approach to detect skin in color image using Bayesian classifier and connected component algorithm**, Math. Problems Eng., vol. 2018, Aug. 2018.
3. M. A. Chandra and S. Bedi, **A twin support vector machine based approach to classifying human skin**, Proc. 4th Int. Conf. Comput. Commun. Automat. (ICCCA), pp. 1-5, 2018.
4. E. Thabet, F. Khalid, P. S. Sulaiman and R. Yaakob, **Low cost skin segmentation scheme in videos using two alternative methods for dynamic hand gesture detection method**, Adv. Multimedia, vol. 2017, Mar. 2017.
5. M. Kawulok, J. Kawulok and J. Nalepa, **Spatial-based skin detection using discriminative skin-presence features**, Pattern Recognit. Lett., vol. 41, pp. 3-13, May 2014.
6. F. Awad and R. Shamroukh, **Human detection by robotic urban search and rescue using image processing and neural networks**, Int. J. Intell. Sci., vol. 4, no. 2, pp. 39, 2014.
7. V. Bhandiwad and B. Tekwani, **Face recognition and detection using neural networks**, 2017 International Conference on Trends in Electronics and Informatics (ICEI), Tirunelveli, India, 2017, pp. 879-882,
8. N. Brancati et al., **Human skin detection through correlation rules between the YCb and YCr subspaces based on dynamic color clustering**, Comput. Vis. Image Understand., vol. 155, pp. 33-42, Feb. 2017.
9. Li, C.; Li, X.; Chen, M.; Sun, X. **Deep Learning and Image Recognition**. In Proceedings of the 2023 IEEE 6th International Conference on Electronic Information and Communication Technology (ICEICT), Qingdao, China, 21–24 July 2023; pp. 557–562.
10. Sadeghi, H.; Raie, A.A. HistNet: **Histogram-based convolutional neural network with Chi-squared deep**

metric learning for facial expression recognition. Inf. Sci. 2022, 608, 472–488.

11. Yu G, Katz LN, Quaia C, Messinger A, and Krauzlis RJ. **Short-latency preference for faces in primate superior colliculus depends on visual cortex**, Neuron. Epub Jul 2, 2024.
12. A.F. D. S. MD. Tahmid Hasan Fuad, **Recent Advances in Deep Learning Techniques for Face Recognition**, IEEE Access,p. 23, 2021.
13. Yadav, S.P.; Jindal, M.; Rani, P.; de Albuquerque, V.H.C.; dos Santos Nascimento, C.; Kumar, M**. An improved deep learning-based optimal object detection system from images.** Multimed. Tools Appl. 2024, 83, 30045–30072.
14. Y. S. a. H. C. Jing Huang, **Improved Viola-Jones face detection algorithm based on HoloLens** EURASIP Journal on Image and Video Processing, 2019.
15. L. W.-y. Ming, **Face Detection Based on Viola-Jones Algorithm Applying Composite Features** International Conference on Robots & Intelligent System (ICRIS), p. 45, 2019.
16. U. H., X. H. a. L.-P. C. Yi Wang, **A Self-Training Approach for Point-Supervised Object Detection and Counting in Crowds** IEEE TRANSACTIONS ON IMAGE PROCESSING, vol. 30, no. 2876, p. 2876, 2021.
17. K. B. Shaik et al., **Comparative study of skin color detection and segmentation in HSV and YCbCr color space** Procedia Comput. Sci., vol. 57, pp. 41-48, Jan. 2015.