Multimedia Tools and Applications (2021) 80:14565–14590
https://doi.org/10.1007/s11042-020-10457-9

Face detection in still images under occlusion and non-uniform illumination

Ashu Kumar[1] · Munish Kumar[2] · Amandeep Kaur[3]

Received: 12 June 2020 / Revised: 20 September 2020 / Accepted: 22 December 2020 /
Published online: 26 January 2021

Abstract
Face detection is important part of face recognition system. In face recognition, face detection is taken not so seriously. Face detection is taken for granted; primarily focus is on face recognition. Also, many challenges associated with face detection, increases the value of TN (True Negative). A lot of work has been done in field of face recognition. But in field of face detection, especially with problems of face occlusion and non-uniform illumination, not so much work has been done. It directly affects the efficiency of applications linked with face detection, example face recognition, surveillance, etc. So, these reasons motivate us to do research in field of face detection, especially with problems of face occlusion and non-uniform illumination. The main objective of this article is to detect face in still image. Experimental work has been conducted on images having problem of face occlusion and non-uniform illumination. Experimental images have been taken from public dataset AR face dataset and Color FERET dataset. One manual dataset has also been created for experimental purpose. The images in this manual dataset have been taken from the internet. This involves making the machine intelligent enough to acquire the human perception and knowledge to detect, localize and recognize the face in an arbitrary image with the same ease as humans do it. This article proposes an efficient technique for face detection from still images under occlusion and non-uniform illumination. The authors have presented a face detection technique using a combination of YCbCr, HSV and $L \times a \times b$ color model. The proposed technique improved results in terms of Accuracy, Detection Rate, False Detection Rate and Precision. This technique can be useful in the surveillance and security related applications.

Keywords Face detection · Occlusion · Human computer interaction · Illumination

✉ Munish Kumar
munishcse@gmail.com

Extended author information available on the last page of the article

1 Introduction

In computer vision, face detection is an important area of research, because it is the first step for any face processing system, such as face recognition, faces tracking, surveillance systems. In many problems, it is assumed that the facial region has been easily identified and perfectly localized. Detecting and recognizing faces in a crowd is an important task in intelligent surveillance application. For developing a fully automated system such as surveillance system which looks for facial information in images containing faces, face detection step must produce an accurate result. Therefore, there is a need for robust face detection algorithm. Locating the face and extracting the facial feature are the other problems related to face detection. The early work on face detection dates to early 70s, where simple heuristics and knowledge about the facial structure were used to convert them into simple rules. Face detection is important for the face recognition problem because it is the preprocessing step for a face recognition system. This technology has been integrated into many products and services including digital billboards, online social networks, and mobile apps. Face detection also has very useful applications in forensic applications, crowd surveillance, secure access control, financial transaction, pedestrian detection, driver alertness monitoring systems, video conferencing etc. In Human-Computer Intelligent Interaction, areas of body analysis, face analysis and complete systems which integrate multiple cues or modalities to give intuitive interaction capabilities to computers, are the dominant clusters of research currently. Tracking the motion and location of the face or body and modeling the face or the human body is considered at face and human body analysis. These are considered fundamental enabling technologies towards Human-Computer Interaction (HCI). In the surveillance system, to track all the people who are present in a place in given time, face detection methodologies are useful. Firstly, surveillance camera is used to capture and later these images are used to detect the regions containing human faces. Regions extracted as facial regions will be compared with a relevant face database available beforehand from sources such as passport, identity card, digital record etc., to accurately identify people.

Face detection is a challengeable task due to large variations in shape, color, size and textural differences between faces. Differences in facial appearance of different people also contribute to the complexity of the problem. Additional features such as spectacles or a moustache, occlusion, lighting conditions and a cluttered image background can make face detection a challengeable task. In the past several years, many techniques for face detection have been proposed throughout the literature that addresses most of the above problems. Pose variations, hardware characteristics and setting the conditions of the camera as well as imaging environment add more constraints to feature space used for face detection. However, there is still no single face detection technique available that fully addresses these problems. Face detection is a complex pattern recognition problem, because of the large level of change in facial appearance. Faces are non-rigid dynamic objects with a large diversity in shape, color, and texture. Main factors like facial expressions, head pose, occlusions (spectacles, muffler, hand), lighting conditions (shadows, contrast) and other facial features (tresses, beard, and moustache) make face detection a complex task. Humans can effortlessly detect and recognize faces with all these variations. So, this work motivates the authors to provide an efficient technique for face detection of still images in the presence of occlusion. Main research objective is to improve the accuracy of face recognition system. Face detection is part of face recognition system. In face recognition, face detection is taken for granted. After literature review it has been found that face detection in presence of face occlusion and non-uniform illumination is not done. So, the authors have motivated to present this research work.

1.1 Challenges

Developing the efficient machine for detecting and recognizing faces with all variations of faces is a big challenge as discussed in Section 1 of this article. In general, a good face detection system should handle faces of different shapes, sizes (face size and appearance vary from person to person) as well as different possible poses in which human faces occur in different image clippings. Humans can very easily recognize and distinguish actual faces from cartoons and drawings. Machines should be trained to effectively distinguish the actual faces from the cartoons/drawings and precisely locating the faces irrespective of the image background and illumination condition. Human beings possess an inherent capability to identify and recognize human faces in any position and even if it is partially visible with any type of occlusion. Good face detection system should search only the regions which contain faces and avoid searching regions which do not contain faces and thus reduce the computational costs and increase the running time efficiency of the algorithm, as algorithm's running time and space are inversely related. Having a robust feature space with reduced dimension demands a good feature extraction method, and in turn reducing the computational cost is one of the key aspects of an efficient face detection system. Some of the research issues involved in face detection are extraction of strong and robust features representing the face in any type of imaging conditions, locating faces of different sizes, finding the precise locations of faces in images, speeding up the process of detecting and locating faces in camouflaged images. There is a need for robust feature set which consumes less time for training. Another challenge for face detection is occlusion, i.e. hiding of face by any object (Fig. 1). It may be glasses, scarf, hand, hairs, hats and any other object etc. It also reduces the accuracy rate of face detection.

The authors observed few challenges during detecting and locating faces in skin-tone regions. Skin regions may not be properly identified as skin pixel during skin segmentation due to variation in illumination. Locating faces in these circumstances are more difficult as compared to find faces with uniform, non-skin-tone background. Detecting and localizing human faces in color images containing multiple faces with skin-tone background is difficult as depicted in Fig. 2a and the entire skin-tone background is also segmented as skin region as

Fig. 1 Example of occluded face images

depicted in Fig. 2b. The entire image of the person is detected as facial region while segmenting the faces of people wearing skin-tone dresses using skin pixel segmentation preprocessing technique, as shown in Fig. 3a and b.

Localizing multiple faces in images with cluttered and complex background is also more difficult as the segmentation of these regions is not straight forward as shown in Fig. 4a and b. Apart from these issues, overlapping face regions are also a constraint while finding the faces.

2 Related work

Hjelmas and Low [5] has organized face detection techniques into two categories, namely, feature-based approaches and image-based approaches as depicted in Fig. 5. Feature based approaches are further classified as low-level analysis, feature analysis and active shape model based. Low level analysis deals with edges, height to width ratio, gray information, color, and motion.

Feature analysis deals with feature searching, constellation searching whereas active shape models are categorized into snakes, deformable templates, and point distribution models. In feature based techniques, the probable face region is checked for the presence of various facial features such as eyes, nose and mouth. Image based approaches are further classified as subspace based, neural network based, and statistical approach based. These images based approaches are holistic in nature and use a scanning window which scans across the whole image at different scales and resolutions to locate the faces. Face detection methods can be broadly classified into four categories; knowledge-based approaches, feature invariant approaches, template matching approaches and appearance-based approaches. Tao et al. [16] proposed a decoupled probabilistic algorithm, which is named Bayesian tensor analysis (BTA). Khandait et al. [8] have used combination of colour space and edges for facial feature detection in colour and grayscale images. In the combined skin segmented regions, it looks for the presence of facial features such as eye, nose, and mouth through image projections. It works well for frontal faces without any unwanted background, but this approach fails for persons with spectacles. Zhipeng et al. [21] have located eye analogue segments at a given

(a) (b)

Fig. 2 **a** Color images containing multiple faces with skin-tone background, **b** Segmented image after various morphological operations

(a) (b)

Fig. 3 **a** Color images containing multiple faces with skin-tone dress and variations in illumination, **b** Segmented image after various morphological operations

scale by finding regions which are roughly as large as eyes and darker than their neighborhoods. A pair of eye analogue segments is hypothesized to be eyes in a face and combined into a face candidate if their placement is consistent with the anthropological characteristic of human eyes. This method is robust as it can deal with illumination change and works well for frontal faces with eyes wide open.

Hu et al. [6] have used a feature-based three-stage scheme to achieve real-time and reliable face detection. In the first stage, after performing the lighting compensation using the reference white, image resizing is done using the discrete wavelet transform. Later segmentation using YCbCr skin-colour model, followed by morphological processing, and fast 4-connected component labeling are used to detect skin regions. Finally, facial features (luminance variation and histogram bin-based skin distribution) are measured to detect face. Viola and Jones [17] described a face detection framework that is capable of processing images extremely rapidly while achieving high detection rates. Instead of directly using pixel information, the proposed framework used a set of simple Haar like features. A feature is computed by summing the pixels in the white region and subtracting those in the dark region. Haar-like features can be computed efficiently with the integrated image representation. Bonnen et al. [2] proposed framework for component-based face alignment and representation that

(a) (b)

Fig. 4 **a** Color images containing multiple faces with cluttered & complex Background, **b** Segmented image after various morphological operations

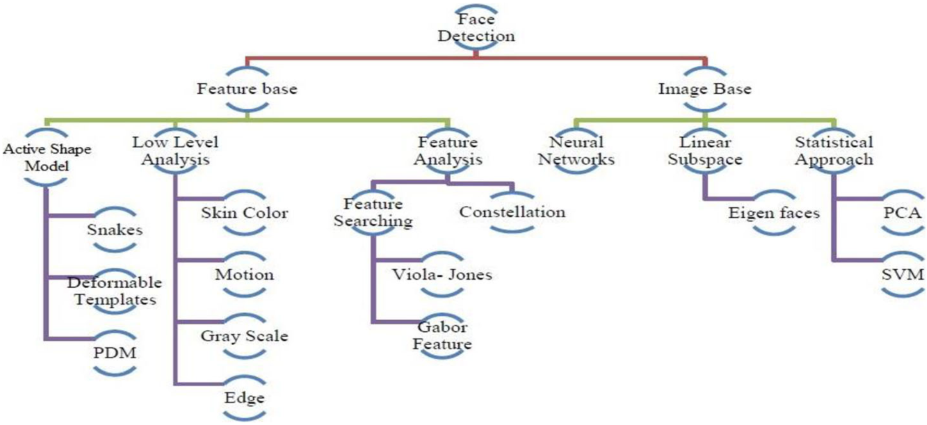


Fig. 5 Different techniques for face detection

demonstrates improvements in matching performance over the more common holistic approach to face alignment and representation. Cheney et al. [3] presented a large scale study of the accuracy and efficiency of face detection algorithms on unconstrained face imagery. Nine different face detection algorithms are studied, which are acquired either by government rights, open source, or commercial licensing. The most notable finding is that top performing detectors still fail to detect many faces with extreme pose, partial occlusion, and/or poor illumination. The speed of the detectors was generally correlated with accuracy: faster detectors were less accurate than their slower counterparts. Ranjan et al. [11] presented a face detection algorithm based on Deformable Part Models (DPM) and deep pyramidal features. The proposed method called DP2MFD can detect faces of various sizes and poses in unconstrained conditions. By adding a normalization layer to the deep convolutional neural network (CNN), it reduces the gap in training and testing of DPM on deep features. Experiments on four publicly available unconstrained face detection datasets showed that the method can capture the meaningful structure of faces and performs significantly better than many competitive face detection algorithms. Bellil et al. [1] proposed a new 3D face identification and recognition method based on Gappy Wavelet Neural Network (GWNN) that is able to provide better accuracy in the presence of facial occlusions. Jiang and Miller [7] discussed that while deep learning based methods for generic object detection have improved rapidly, but most approaches to face detection are still based on the R-CNN framework, leading to limited accuracy and processing speed. Authors applied the Faster R-CNN on face detection. By training a Faster R-CNN model on the large scale WIDER face dataset, the authors reported state-of-the-art results on the WIDER test set as well as two other widely used face detection benchmarks, FDDB and the recently released IJB-A. Wang et al. [18] proposed a novel face detector called Face Attention Network (FAN) in which new anchor-level attention highlighted the features from the face region. Rowley et al. [12] have presented a neural network-based face detection system. Their system arbitrates between multiple networks to improve performance over a single network. They used a bootstrap algorithm for training the networks, which adds false detections into the training set as training progresses. Soundararajan and Biswas [14] have assessed machine vision quality for robust face detection. They defined machine vision quality for face detection as a weighted combination of precision and recall. And they used machine vision quality for image enhancement to improve face detection performance

Guo et al. [4] presented a face detection technique with large area occlusion based on rules of facial physiological characteristics. Adaboost cascade classifier based on the Haar-like feature is underlying detection algorithm. They used opencv cascade classifier to detect the human eye and the mouth. Then, according to the relationship between the human eye and the human face and the relationship between the face and the mouth, the face detection of the occlusion is realized. According to them, this method achieves a relatively higher accuracy compared with other conventional face detection methods. Sharma et al. [13] proposed a novel real-time face detection system which detects the faces that are tilted, occluded or with different illuminations, any difficult pose. The problem with cascaded classifier is that it does not detect the tilted or occluded faces with different illuminations. So, to overcome this problem, they proposed a system using Modified Affine Transformation with Viola Jones. They tested the proposed system with three different datasets namely FDDB, YALE and "Google top 25 'tilted face'" image datasets. The results were 6% better than Viola–Jones method. Sun et al. [15] presented a new face detection scheme using deep learning. They performed the experiments with FDDB face detection and achieved the state-of-the-art detection performance. They improved the state-of-the-art Faster RCNN framework by combining a few strategies, including feature concatenation, hard negative mining, multi-scale training, model pre-training, and proper calibration of key parameters. According to them, proposed scheme obtained the state-of-the-art face detection performance and was ranked as one of the best models in terms of ROC curves of the published methods on the FDDB benchmark. Zhang et al. [20] proposed a fast and robust face occlusion detection algorithm for ATM surveillance, to be effective and efficient to handle arbitrarily occluded faces. They first construct a novel energy function for elliptical head contour detection and then develop a fast and robust head tracking algorithm, which utilizes the gradient and shape cues in a Bayesian framework. Lastly, to verify whether a detected head is occluded or not, they discussed method to fuse information from both skin color and facial structure using the AdaBoost algorithm. Experimental results on real world data show that the proposed algorithm can achieve 98.64% accuracy on face detection and 98.56% accuracy in face occlusion detection, even though there are severe occlusions in faces, at a speed of up to 12 frames per second. Nasir et al. [10] presented an algorithm, which consists of three main processes, namely, Haar feature, AdaBoost and cascading. These processes require scanning the patterns of a human face to obtain all pixels in the image. Hybridization of skin colour segmentation prior to using Viola–Jones algorithm is discussed. Different types of environments (light conditions and face orientations) and various ethnicities (Malay, Chinese and Indian) images are tested using the improved algorithm and the original Viola–Jones algorithm. They concluded that combining YCbCr color and Viola–Jones algorithm is the optimal model for human face detection under various conditions. Other color models, such as hue saturation value (HSV) and standard algorithms have slightly low detection rates. They also determined that the YCbCr and HSV colour models demonstrate certain limitations when handling dark faces and lighting conditions because the skin colour becomes slightly out of range from the normal skin color normalization. Already existing face detection or face recognition systems detect only front faces, but they are not able to detect/recognize occluded face or non-uniform illuminated face. This is the major drawback of existing systems. This step has motivated us to propose an efficient technique to detect occluded faces from non-uniform illuminated images. So, in this article, the authors have proposed a technique to detect occluded faces from non-uniform illuminated images, with acceptable accuracy. This technique can be useful in the surveillance and security related applications. Kumar et al. [9] have presented a review on face detection techniques. They have also discussed various challenges for face detection. Wang et al. [19] proposed re-ranking approach, which achieved significantly higher precision in the top ten

Table 1 Comparison of feature based and image based approach

	Feature based approach	Image based approach
Technique	Find invariant features of faces for detection. The underlying assumption is based on the observation that humans can effortlessly detect faces and objects in different poses and lighting conditions, so there must exist properties or features which are invariant over these variabilities. Facial features such as skin-color, eyebrows, eyes, nose, mouth and hair-line are commonly extracted using edge detectors. Based on the extracted features, a statistical model is built to describe their relationships and to verify the existence of a face.	Templates are learned from examples in images. In general, appearance-based methods rely on techniques from statistical analysis and machine learning to find the relevant characteristics of face and non-face images. The learned characteristics are in the form of distribution models or discriminant functions that is consequently used for face detection
Examples	Skin color, motion, edge, Viola–Jones, snakes etc.	Neural-networks, HMM, SVM, AdaBoost learning etc.
Advantages	Easy to implement.	Difficult to implement
Disadvantages	Image features can be severely corrupted due to illumination, noise, and occlusion. Feature boundaries can be weakened for faces, while shadows can cause numerous strong edges which together render perceptual grouping algorithms useless.	Dimensionality reduction is usually carried out for the sake of computation efficiency and detection efficacy

different SBIR methods and datasets. Table 1 depicts the comparative study of feature based and image based approaches.

2.1 Standard database for face detection

Face image databases are collection of different type of faces, which may be used as test set for face detection system. Some standard face image databases are available, which are following (Table 2).

3 Dataset

In this proposed system, we used two standard image datasets and one manual collected dataset. First one is the AR face database, second one is a Color FERET database and the third one is manual collected images from internet. The AR face database contains 4000 color images of 126 people's faces (70 men and 56 women). In Fig. 6, the authors have shown a few samples of AR face database. The Color FERET database contains 1564 sets of images. A total of 14,126 images that include 1199 individuals and 365 duplicate set of images. The few specimens of Color FERET database appeared in Fig. 7. It contains 1005 images having problem of face occlusion and non-uniform illumination. In Fig. 8, the authors have shown a few samples of manual dataset. Brief information about these databases is presented in Table 3.

A few samples of the AR face database, Color FERET database and manual dataset are shown in Figs. 6, 7 and 8, respectively.

Table 2 Standard face datasets

Database	Website	Description
MIT dataset	http://cbcl.mit.edu/softwaredatasets/FaceData2.html	19×19 Gray-scale PGM format images Training set: 2429 faces, 4548 non-faces Test set: 472 faces, 23,573 non-faces
PIE database, CMU	http://www.ri.cmu.edu	A database of 41,368 images of 68 people, each person under 13 different poses, 43 different illumination conditions, and with 4 different expressions
FERET database	http://www.itl.nist.gov/iad/humanid/fer et/feret_master.html	It consists of 14,051 eight-bit gray-scale images of human heads with views ranging from frontal to left and right profiles
The Yale face database	http://www.face-rec.org/databases/	Contains 165 Gy-scale images in GIF format of 15 individuals. There are 11 images per subject, one per different facial expression or configuration: center-light, w/glasses, happy, left-light, w/no glasses, normal, right-light, sad, sleepy, surprised, and wink
Indian face database	http://www.pics.stir.ac.uk/Other_face_databases.htm	11 images of each of 39 men, 22 women from Indian Institute of Technology Kanpur
AR database	http://www.2.ece.ohio-state.edu/~aleix/	It contains over 4000 color images corresponding to 126 people's faces (70 men and 56 women). Features based on frontal view faces with different facial expressions, illumination conditions, and occlusions (sun glasses and scarf)
SCface—surveillance cameras face database	http://www.scface.org	Images were taken in uncontrolled indoor environment using five video surveillance cameras of various qualities. Database contains 4160 static images (in visible and infrared spectrum) of 130 subjects

4 Proposed methodology

4.1 Key technologies

HSV, YCbCr and $L \times a \times b$ color models are used in this purposed methodology. Fusion of output of all three models is done. It works efficiently. The output of this fusion helps to segment the face with the help of skin color. All three models are explained as below:

- HSV [10, 21]

Fig. 6 Occluded faces with different illumination (AR Face Database)

HSV is a cylindrical color model that remaps the RGB primary colors into dimensions that are easier for humans to understand. Like the Munsell Color System, these dimensions are hue, saturation, and value.

- Hue specifies the angle of the color on the RGB color circle. A 0° hue results in red, 120° results in green, and 240° results in blue.
- Saturation controls the amount of color used. A color with 100% saturation will be the purest color possible, while 0% saturation yields grayscale.

Fig. 7 A few samples of Color FERET database

Fig. 8 A few samples taken from manual dataset

- Value controls the brightness of the color. A color with 0% brightness is pure black while a color with 100% brightness has no black mixed into the color. Because this dimension is often referred to as brightness, the HSV color model is sometimes called HSB.
- YCbCr (Luminance, Chrominance) Color Model [10, 21]

YCbCr is a commonly used color space in the digital video domain. Because the representation makes it easy to get rid of some redundant color information, it is used in image and video compression standards like JPEG, MPEG1, MPEG2 and MPEG4. Transformation simplicity and explicit separation of luminance and chrominance components makes YCbCr color space. In this format, luminance information is stored as a single component (Y), and chrominance information is stored as two color-difference components (Cb and Cr). Cb represents the difference between the blue component and a reference value. Cr represents the difference

Table 3 Brief information about datasets

Database	Size of image	Total number of images
AR face database	768 × 576	4000
Color FERET database	512 × 768 or 256 × 484 or 128 × 192	14,126
Manual dataset (collected from internet)	786 × 576	1005

between the red component and a reference value. YCbCr values can be obtained from RGB color space. YCbCr color space may be used for skin detection.

- L × a × b [10, 21]

In this color model, L is for lightness of color. The 'a' component stands between green and red, wherein the 'b' component stands between blue and yellow. L × a × b color is designed to approximate human vision. It aspires to perceptual uniformity, and its L component closely matches human perception of lightness.

Algorithm for skin segmentation

Image (I) is original image, n is total number of pixel in image.

Step 1: Convert original image (I) into YCbCr image
Step 2: Then Convert YCbCr image into binary image with some thresholding values of Y, Cb and Cr components.

for $i := 0$ **to** n **step** 1 pixel **do**
if Pixel value between thresholding values **then** in binary image set pixel value 1, otherwise set 0
End if
End for

Step 3: Convert original image (I) into HSV image.
Step 4: Then Convert HSV image into binary image with some thresholding values of H, S and V components.

for $i := 0$ to n **step** 1 pixel **do**
if Pixel value between thresholding values **then** in binary image set pixel value 1, otherwise set 0
End if
End for

Step 5: Convert original image (I) into L×a×b image.
Step 6: Then Convert L×a×b image into binary image with some thresholding values of L, a and b components

for $i :=0$ to n **step** 1 pixel **do**
if Pixel value between thresholding values **then** in binary image set pixel value 1, otherwise set 0
End if
End for

Step 7: Fused the outputs of Step 2, 4 and 6. We got one new merged binary image, having skin segmented.

4.2 Block diagram of proposed system

Block diagram of the proposed methodology is depicted in Fig. 9. Proposed methodology consists of four phases. First of all, read the .jpeg image. First phase is pre-processing. In a pre-processing phase, resizing of images and normalization of illumination has been done. For normalized the illumination, method Histogram Equalization is used. Second phase is color model conversion. In second phase, image is converted into YCbCr image, HSV image and

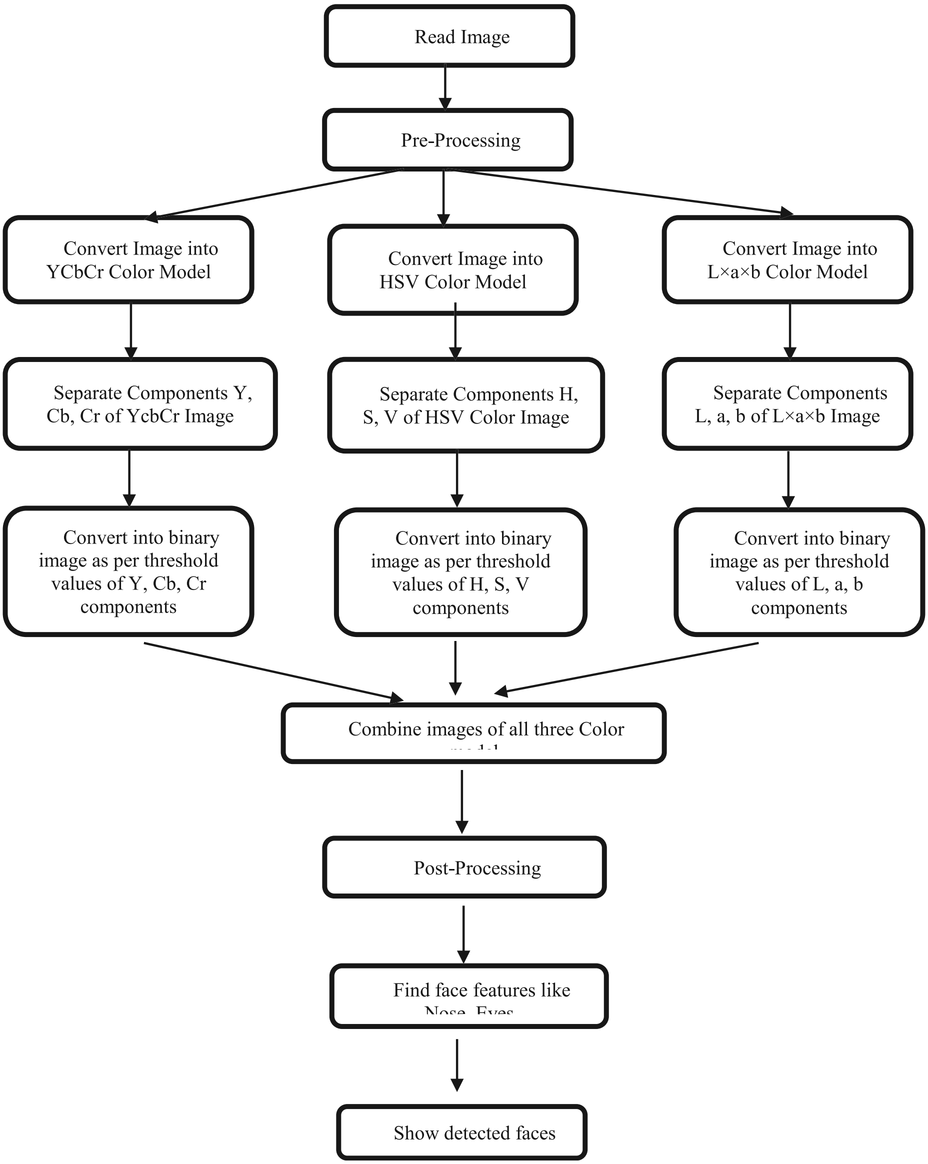


Fig. 9 Block diagram of proposed methodology for face detection

L × a × b image. Respective components of all three color models have been found. After that, according to threshold value, each image is converted into binary image.

For YCbCr image, we used threshold values

$$72 <= Cb <= 122 \text{ and } 128 <= Cr <= 171$$

For HSV image, we used threshold values

$$V >= 38$$

$0.2 <= S < 0.5$
$0^0 < H < 25^0$ or $335^0 < H < 360^0$

After that, binary images of YCbCr, HSV and L × a × b are fused into single image. This helps to skin segmentation in image. Third phase is post-processing. In this phase, morphological operations have been performed on binary image, gotten from second phase. Fourth phase is face detection. In this phase, authors try to find the face features like the eyes and nose in the image. With the help of CascadeClassifier, facial features like nose, eyes have been detected in image. These facial features find the positions of face in the image. On the basis of these face features, face has been detected using method detectMultiScale().

5 Experimental results and discussion

Initially, standard datasets, namely, AR face dataset and Color FERET dataset have been acquired. AR face dataset consists of approximately 8000 images of various persons. These images were in .raw format. Authors have converted all the .raw images into. JPEG format. Color FERET dataset consists of 1564 sets of images that include 1199 individuals and 365 duplicate sets of images. These images were in .ppm format. Authors have converted all the .raw images into. Jpeg format. They also created a manual dataset of 1005 images, collected from the internet, having problem of non-uniform illumination and occluded faces. After that, they read the. Jpeg image. In a preprocessing step, resizing of images and normalization of image has been done. After that, the authors try to find the face features like the eyes and nose in the image. These facial features find the positions of face in the image.

To measure the performance of the proposed system, the authors have performed experiments using the AR face dataset, color FERET dataset and manual dataset. Images having problem of occluded faces and non-uniform illumination have been considered for experimental work. Performance is measured in terms of Precision, Detection Rate and False Detection Rate. Precision, Detection Rate and False Detection Rate has been calculated with the help of TP, FP, TN and FN (Ture Positive, False Positive, True Negative and False Negative respectively). In Table 4 different test images taken from AR face dataset, color FERET dataset and manual dataset have been shown. In Table 4 input image and its respective output image with detected face has been shown. In remarks column remarks has been given about the problem of image i.e. occluded face or non-uniform illumination.

5.1 Face detection experimental results

Experimental comparison of the various face detection algorithms has been presented. Accuracy, Precision, detection rate and false detection rate are the various performance metrics for the analysis of the results. Accuracy shows the proportion of true results, both true positives and true negatives. The accuracy is obtained in all the four algorithms stated using the Equation.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

Where.

True positive (TP) is the number of faces that are detected.

True negative (TN) is the number of faces that are not detected.

Table 4 Experimental results for face detection using proposed methodology

Input Image	Output Image	Remarks
		Problem of non-uniform illumination
		Problem of face occlusion
		Problem of non-uniform illumination
		Problem of non-uniform illumination
		Problem of non-uniform illumination and face occlusion
		Problem of non-uniform illumination
		Problem of face occlusion

Table 4 (continued

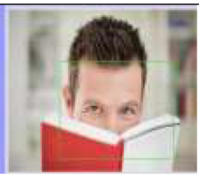
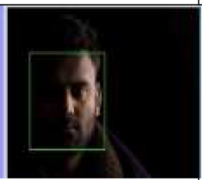
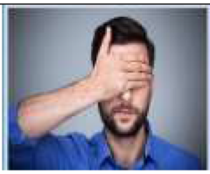
		Problem of face occlusion
		Problem of face occlusion
		Problem of non-uniform illumination
		Problem of non-uniform illumination
		Problem of non-uniform illumination
		Problem of face occlusion
		Problem of non-uniform illumination
		Problem of non-uniform illumination and face occlusion

Table 4 (continued

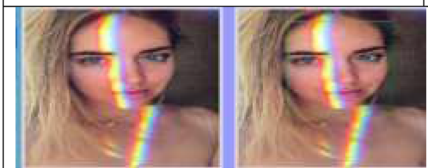
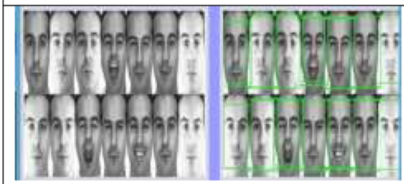
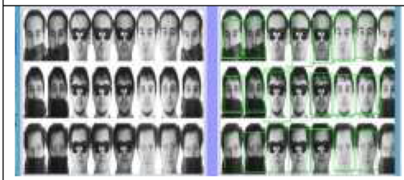
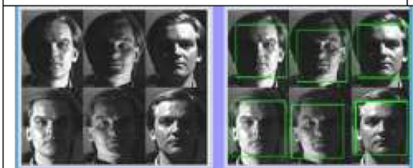
	Problem of face occlusion
	Problem of face occlusion
	Problem of non-uniform illumination
	Problem of face occlusion
	Problem of face occlusion
	Problem of non-uniform illumination
	Problem of face occlusion
	Problem of non-uniform illumination

Table 4 (continued)

	Problem of non-uniform illumination
	Problem of non-uniform illumination
	Problem of non-uniform illumination
	Problem of face occlusion
	Problem of non-uniform illumination
	Problem of non-uniform illumination and face occlusion
	Problem of non-uniform illumination
	Problem of non-uniform illumination

False positive (FP) is the number of non-faces falsely detected as faces.

False negative (FN) is the number of non-faces rejected from the classifier.

Precision or positive prediction value is defined as the proportion of the true positives against all positive results.

$$Precision = \frac{TP}{TP + FP}$$

Detection rate gives us a percentage of faces correctly detected.

$$Detection\ Rate = \frac{TP}{Total\ Faces}$$

False detection rate shows the percentage of non-faces falsely detected as face.

$$False\ Detection\ Rate = \frac{FP}{Total\ Non\text{–}Faces}$$

5.2 Results of skin color based face detection in YCbCr color space

Experiments have been performed on the YCbCr color space. The accuracy is found to be 87.14%. Table 5 shows that the false detection rate is high and dismissal rate is low, thus causing a bit low accuracy in detecting the face. It is found that skin color region is more effectively extracted in YCbCr color space. This is because Cb and Cr have some distinct color range for skin region. Thus the accuracy of this algorithm is quite good.

5.3 Results of skin color based face detection in HSI color space

Same experiments have been performed on the HSI color space as on YCbCr. Experimental images taken from AR Face dataset, FERET Color dataset and Manual created dataset. Images having

Table 5 Face detection results using YCbCr Color Space

Dataset	Criterion	No. of Images	TP	FP	TN	FN	Accuracy (%)	Detection Rate (%)	Precision (%)	False Detection Rate (%)
AR Face Dataset	Occluded Faces	130	61	7	9	53	87.69	87.14	89.70	11.66
	Non-Uniform Illumination	130	59	8	11	52	85.38	84.28	88.05	13.33
FERET Color Dataset	Occluded Faces	130	62	7	8	53	88.46	88.57	89.85	11.66
	Non-Uniform Illumination	130	59	7	11	53	86.15	84.28	89.39	11.66
Manual Created Dataset	Occluded Faces	130	60	8	10	13	86.15	85.71	88.23	13.33
	Non-Uniform Illumination	130	58	7	12	13	85.38	82.85	89.23	11.66

Table 6 Face detection results using HSI color space

Dataset	Criterion	No. of Images	TP	FP	TN	FN	Accuracy (%)	Detection Rate (%)	Precision (%)	False Detection Rate (%)
AR Face Dataset	Occluded Faces	130	58	9	12	51	83.84	82.85	86.56	15.00
	Non-Uniform Illumination	130	57	8	13	52	83.84	81.42	87.69	13.33
FERET Color Dataset	Occluded Faces	130	57	8	13	52	83.84	81.42	87.69	13.33
	Non-Uniform Illumination	130	55	8	15	52	82.30	78.57	84.61	13.33
Manual Created Dataset	Occluded Faces	130	57	8	13	52	83.84	81.42	87.69	13.33
	Non-Uniform Illumination	130	56	7	14	53	83.84	80.00	86.15	11.66

problem of occluded faces and non-uniform illumination have been considered for experiments work. The accuracy is found to be 82.85% which is better, but little less compared to YCbCr color space. Table 6 shows that the false detection rate is high and dismissal rate is low, thus causing a bit low accuracy in detecting the face. It is concluded skin color region is effectively extracted in HSI color space. Thus the accuracy of this algorithm is quite good against all three dataset.

5.4 Results of skin color based face detection in L × a × b color space

Experiments have been performed on the L × a × b color space. Experimental images taken from AR Face dataset, FERET Color dataset and Manual created dataset. Images having problem of occluded faces and non-uniform illumination have been considered for experimental work. The accuracy of this

Table 7 Face detection results using L × a × b color space

Dataset	Criterion	No. of Images	TP	FP	TN	FN	Accuracy (%)	Detection Rate (%)	Precision (%)	False Detection Rate (%)
AR Face Dataset	Occluded Faces	130	42	18	28	42	64.61	60.00	70.00	30.00
	Non-Uniform Illumination	130	43	16	27	44	66.92	61.42	72.88	26.66
FERET Color Dataset	Occluded Faces	130	43	18	27	42	65.38	61.42	70.49	30.00
	Non-Uniform Illumination	130	42	17	28	43	65.38	60.00	71.11	28.33
Manual Created Dataset	Occluded Faces	130	41	17	29	43	64.61	58.57	70.68	28.33
	Non-Uniform Illumination	130	40	18	30	42	63.07	57.14	68.96	30.00

Table 8 Face detection results using proposed model (YCbCr + HSI + L × a × b)

Dataset	Criterion	No. of Images	TP	FP	TN	FN	Accuracy (%)	Detection Rate (%)	Precision (%)	False Detection Rate (%)
AR Face Dataset	Occluded Faces	130	69	3	1	57	96.92	98.57	95.83	5.00
	Non-Uniform Illumination	130	68	3	2	57	96.15	97.14	95.77	5.00
FERET Color Dataset	Occluded Faces	130	68	3	2	57	96.15	97.14	95.77	5.00
	Non-Uniform Illumination	130	68	3	2	57	96.15	97.14	95.77	5.00
Manual Created Dataset	Occluded Faces	130	67	4	3	56	94.61	95.71	94.36	6.00
	Non-Uniform Illumination	130	66	4	4	56	93.84	94.28	94.28	6.00

experiment is found to be 60.0%. Table 7 shows that the false detection rate and dismissal rate are very high, thus causing very low accuracy in detecting the face. In L × a × b color space it is found that it represents some non-skin region also as the skin region and hence false detection rate is quite high.

5.5 Results of the proposed algorithm

The experiments show very good results for proposed algorithm (YCbCr + HCI + L × a × b) as in Figure. The True positive, true negative, false positive and false negative are shown in Table 8. The low false detection rate shows that the algorithm is able to distinguish between actual skin and background color with skin color appearance. The detection rate is found to be 98.57%. Sample results from proposed algorithm are graphically depicted in Fig. 10.

6 Comparative analysis of proposed method vs. existing methods

In this section, experimental results based on the proposed method are compared with the existing methods for face detection. The face detection results for 70 face images and 60 non-

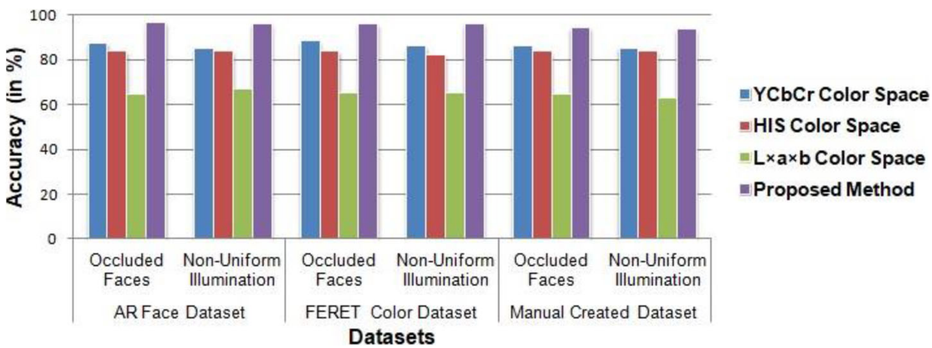


Fig. 10 Comparison of various face detection algorithms (with Accuracy rate)

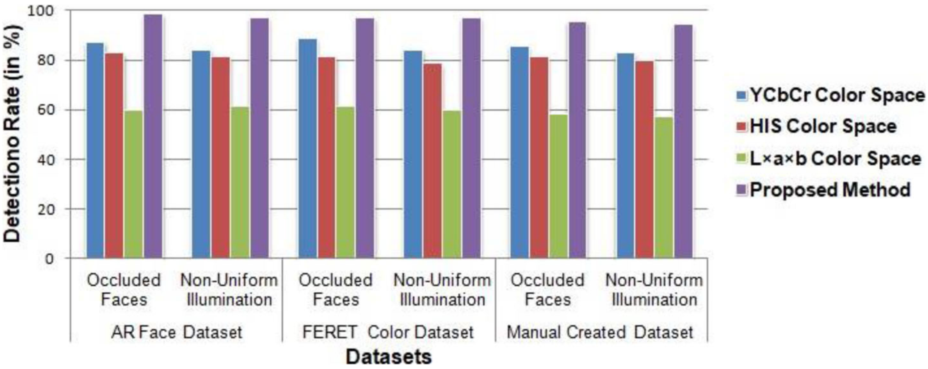


Fig. 11 Comparison of various face detection algorithms (with detection rate)

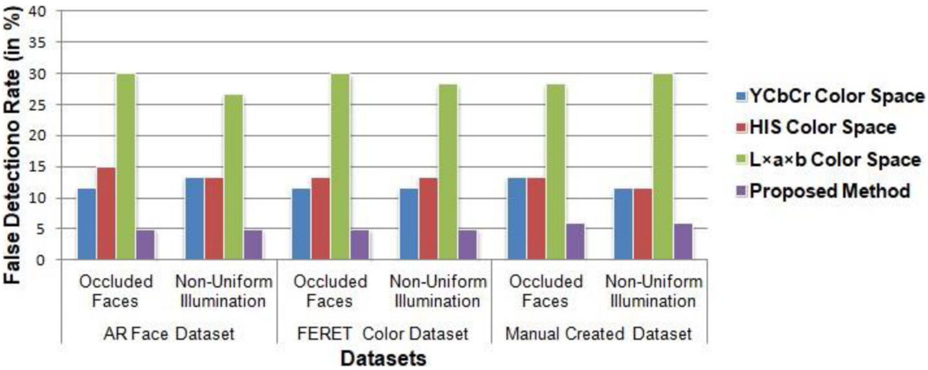


Fig. 12 Comparison of various face detection algorithms (with false detection rate)

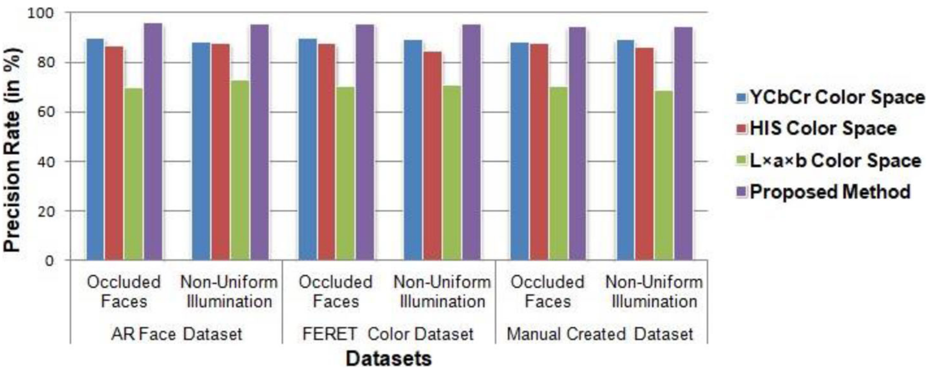


Fig. 13 Comparison of various face detection algorithms (with Precision Rate)

Table 9 Comparative analysis based on the Accuracy

Dataset	Criterion	YCbCr color space	HSI color space	L × a × b Color Space	Proposed Method (YCbCr+ HSI+L × a × b)
AR Face Dataset	Occluded Faces	87.69	83.84	64.61	**96.92**
	Non-Uniform Illumination	85.38	83.84	66.92	**96.15**
FERET Color Dataset	Occluded Faces	88.46	83.84	65.38	**96.15**
	Non-Uniform Illumination	86.15	82.30	65.38	**96.15**
Manual Created Dataset	Occluded Faces	86.15	83.84	64.61	**94.61**
	Non-Uniform Illumination	85.38	83.84	63.07	**93.84**

face images using YCbCr, HSI, L × a × b color spaces and proposed method (YCbCr + HSI + L × a × b) are represented in Tables 5, 6, 7, and 8, respectively. Total 130 × 2 (Occlusion and Non – uniform Illumination) × 3 (Three datasets) = 980 Images has been taken for experiment work. Experimental results of proposed method are compared with the existing methods based on accuracy, detection rate, false detection rate and precision rate. Results of comparative study based on accuracy, detection rate, false detection rate and precision rate are presented in Tables 9, 10, 11 and 12 respectively. It is calculated that proposed method (Fusion of YCbCr, HSI and L × a × b) has better accuracy, detection rate, false detection rate and precision rate than existing method of YCbCr, HIS and L × a × b.

Table 10 Comparative analysis based on the detection rate

Dataset	Criterion	YCbCr color space	HSI color space	L × a × b Color Space	Proposed Method (YCbCr+ HSI+L × a × b)
AR Face Dataset	Occluded Faces	87.14	82.85	60.00	**98.57**
	Non-Uniform Illumination	84.28	81.42	61.42	**97.14**
FERET Color Dataset	Occluded Faces	88.57	81.42	61.42	**97.14**
	Non-Uniform Illumination	84.28	78.57	60.00	**97.14**
Manual Created Dataset	Occluded Faces	85.71	81.42	58.57	**95.71**
	Non-Uniform Illumination	82.85	80.00	57.14	**94.28**

Table 11 Comparative analysis based on the false detection rate

Dataset	Criterion	YCbCr color space	HSI color space	L × a × b Color Space	Proposed Method (YCbCr+ HSI+L × a × b)
AR Face Dataset	Occluded Faces	11.66	15.00	30.00	**5.00**
	Non-Uniform Illumination	13.33	13.33	26.66	**5.00**
FERET Color Dataset	Occluded Faces	11.66	13.33	30.00	**5.00**
	Non-Uniform Illumination	11.66	13.33	28.33	**5.00**
Manual Created Dataset	Occluded Faces	13.33	13.33	28.33	**6.00**
	Non-Uniform Illumination	11.66	11.66	30.00	**6.00**

Table 12 Comparative analysis based on the precision rate

Dataset	Criterion	YCbCr color space	HSI color space	L × a × b Color Space	Proposed Method (YCbCr+ HSI+L × a × b)
AR Face Dataset	Occluded Faces	89.70	86.56	70.00	**95.83**
	Non-Uniform Illumination	88.05	87.69	72.88	**95.77**
FERET Color Dataset	Occluded Faces	89.85	87.69	70.49	**95.77**
	Non-Uniform Illumination	89.39	84.61	71.11	**95.77**
Manual Created Dataset	Occluded Faces	88.23	87.69	70.68	**94.36**
	Non-Uniform Illumination	89.23	86.15	68.96	**94.28**

7 Conclusion and future scope

This article proposes an efficient technique for face detection from still images under occlusion and non-uniform illumination. The authors have presented a face detection technique using a combination of YCbCr, HSV and L × a × b color model. Proposed method contains four phases. First phase is pre-processing. In a pre-processing phase, resizing of images and normalization of illumination has been done. For normalized the illumination, method Histogram Equalization is used. Second phase is color model conversion. Third phase is post-processing. In this phase, morphological operations have been performed on binary image, gotten from second phase. Fourth phase is face detection. In this phase, authors try to find the face features like the eyes and nose in the image. Experimental results for existing methods and the proposed methods are graphically depicted in Figs. 11, 12, and 13. Figure 11 presents results based on detection rate, Fig. 12 depicts the false detection rate and Fig. 13 depicts the precision rate of the proposed system and existing systems for face detection.

After comparison of proposed algorithm with other three methods with respect of dataset and problems, it is concluded that proposed method provides better results. Experimental work has been tested on two standard dataset AR Face dataset and FERET Color dataset. One manual dataset also has been created and tested for real time images. Comparison has been done existing and proposed method on basis of accuracy, detection rate, false detection rate and precision rate. Proposed technique of face detection for still images under the presence of problems of occlusion and non-uniform illumination, will helps to increase the accuracy of face detection system and as well of face recognition system. There is always scope for further research in any work. So here is also scope of face detection in complex images. Complex images mean images having small multiple faces, and many other objects in image.

Compliance with ethical standards

Conflict of interest All the authors declare that they have no conflict of interest in this work.

References

1. Bellil W, Brahim H, Amar CB (2016) Gappy wavelet neural network for 3D occluded faces: detection and recognition. Multimed Tools Appl 75:365–380

2. Bonnen K, Klare BF, Jain AK (2013) Component-based representation in automated face recognition. IEEE Trans Inf Forensics Secur 8(1):239–253
3. Cheney J, Klein B, Jain AK, Klare BF (2015) Unconstrained face detection: state of the art baseline and challenges. Proceedings of the International Conference on Biometrics (ICB), 229-236
4. Guo ZH, Zhou W, Xiao L, Hu X, Zehao Z and Zhou H (2018) Occlusion face detection technology based on facial physiology. Proceedings of the IEEE International Conference on Computational Intelligence and Security (CIS), 106-109.
5. Hjelmas E, Low BK (2001) Face detection: A Survey. Comput Vis Image Underst 83:236–274
6. Hu WC, Yang CY, Huang DY, Huang CH (2011) Feature-based face detection against skin-color like backgrounds with varying illumination. J Inf Hiding Multimedia Signal Process 2(2):123–132
7. Jiang H, Miller EL (2017) Face detection with the faster R-CNN. Proceedings of the 12th IEEE International Conference on Automatic Face & Gesture Recognition, 650-657
8. Khandait S, Khandait PD, Thool R (2009) An efficient approach to facial feature detection for expression recognition. Int J Recent Trends Eng 2(1):179–183
9. Kumar A, Kaur A, Kumar M (2019) Face detection techniques: a review. Artif Intell Rev 52(2):927–948
10. Nasir AFA, Ghani ASA, Zakaria MA, Majeed APPA, Ibrahim AN (2019) Automated face detection using skin colour segmentation and Viola–Jones algorithm. J Intell Manuf Mechatron 1(1):58–63
11. Ranjan R, Patel VM, Chellappa R (2015) A deep pyramid deformable part model for face detection. Proceedings of the IEEE 7th International Conference on Biometrics Theory, Applications and Systems 1-8
12. Rowley HA, Baluja S, Kanad T (1996) Neural network-based face detection. Comput Vis Pattern Recognit 20:203–208
13. Sharma R, Ashwin TS, Guddeti RMR (2018) A novel real-time face detection system using modified affine transformation and Haar cascades. Proceedings of the Recent Findings in Intelligent Computing Techniques 193-204
14. Soundararajan R, Biswas S (2019) Machine vision quality assessment for robust face detection. Signal Process Image Commun 72:92–104
15. Sun X, Wua P, Hoia SCH (2018) Face detection using deep learning: an improved faster RCNN approach. Neurocomputing 299:42–50
16. Tao D, Song M, Li X, Shen J, Sun J (2008) Bayesian tensor approach for 3-D face modeling. IEEE Trans Circ Syst Video Technol 18(10):1397–1410
17. Viola P, Jones MJ (2004) Robust real-time face detection. Int J Comput Vis 57(2):137–154
18. Wang J, Yuan Y, Yu G (2017) Face attention network: an effective face detector for the occluded faces. CoRR
19. Wang L, Qian X, Zhang Y, Shen J, Cao X (2020) Enhancing sketch-based image retrieval by CNN semantic re-ranking. IEEE Trans Cybern 50(7):3330–3342
20. Zhang T, Li J, Jia W, Sun J, Yang H (2018) Fast and robust occluded face detection in ATM surveillance. Pattern Recogn Lett 107:33–40
21. Zhipeng C, Junda H, Wenbin Z (2010) Face detection system based on skin color model. Proc Int Conf Netw Digit Soc 2:664–667

Affiliations

Ashu Kumar[1] **· Munish Kumar**[2] **· Amandeep Kaur**[3]

Ashu Kumar
ashu_sa@pbi.ac.in

Amandeep Kaur
aman_k2007@hotmail.com

1 Research Scholar, Department of Computer Science, Punjabi University, Patiala, Punjab, India

2 Department of Computational Sciences, Maharaja Ranjit Singh Punjab Technical University, Bathinda, Punjab, India

3 Department of Computer Science & Technology, Central University of Punjab, Bathinda, Punjab, India