



# A novel technique for automated concealed face detection in surveillance videos

Hanan A. Hosni Mahmoud<sup>1,2</sup> · Hanan Abdullah Mengash<sup>3</sup>

Received: 24 March 2020 / Accepted: 18 May 2020 / Published online: 12 June 2020  
© Springer-Verlag London Ltd., part of Springer Nature 2020

## Abstract

Face detection perceives great importance in surveillance paradigm and security paradigm areas. Face recognition is the technique to identify a person identity after face detection. Extensive research has been done on these topics. Another important research problem is to detect concealed faces, especially in high-security places like airports or crowded places like concerts and shopping centres, for they may prevail security threat. Also, in order to help effectively in preventing the spread of Coronavirus, people should wear masks during the pandemic especially in the entrance to hospitals and medical facilities. Surveillance systems in medical facilities should issue warnings against unmasked people. This paper presents a novel technique for concealed face detection based on complexion detection to challenge a concealed face assumption. The proposed algorithm first determine of the existence of a human being in the surveillance scene. Head and shoulder contour will be detected. The face will be clustered to cluster patches. Then determination of presence or absent of human skin will be determined. We proposed a hybrid approach that combines normalized RGB (rgb) and the YCbCr space color. This technique is tested on two datasets; the first one contains 650 images of skin patches. The second dataset contains 800 face images. The algorithm achieves an average detection rate of 97.51% for concealed faces. Also, it achieved a run time comparable with existing state-of-the-art concealed face detection systems that run in real time.

**Keywords** Face detection · Security · Human skin detection · YCbCr space color

## 1 Introduction

In this research, we are concerned by two concepts; the first is face detection in surveillance videos and the second is concealed face identification. Face detection is concerned by discovery of faces in a video or video frames and if discovered, then the image location should be noted. The challenges of face detection techniques are numerous. Face pose in a

video frame can be non-conform because of the locus of camera-face dependency. Camera locus can be frontal, inclined, or profile. Also, faces can be concealed partially or totally due to innocent actions like presence of beards or glasses, or due to threatening actions like presence of mask. Another problem that face detection systems may face is the problem that faces may be the occlusion where faces can be partially concealed by other objects in the video frame. Also, lighting state and camera characteristics can distress the appearance of a face. To handle these complications, researchers have proposed different techniques. Robust face recognition from multi-view videos is proposed by Du, Sankaranarayanan, and Chellappa [10]. Advanced face detection techniques can handle adversative conditions such as lightning setting and profile angles. Nowadays, techniques utilize neural networks and skin color identification. Skin detection using color processing mechanism was proposed by Wu et al. [43]. Skin color detection utilizing neural network was proposed by Kim, Hwang, and Cho [24]. This algorithm achieved high-performance face detection time. Shearlet neural network for face detection using one sample per person was achieved by Borgi et al. [3]. Ejbali, Zaied, and Ben

---

✉ Hanan Abdullah Mengash  
hamengash@pnu.edu.sa

Hanan A. Hosni Mahmoud  
HAhosni@pnu.edu.sa

<sup>1</sup> Department of Computer Science, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia

<sup>2</sup> Department of Computer and Systems Engineering, Faculty of Engineering, University of Alexandria, Alexandria, Egypt

<sup>3</sup> Department of Information Systems, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia

Amar [11] implemented face recognition model based on elastic graph equivalence, skin segmentation, and consequent. A literature survey of face recognition is presented by Zhao, Chellappa, Phillips, and Rosenfeld [49]. Filali et al. [12] introduced texture classification of melanoma skin cancer utilizing an efficient convolutional neural network. Chai, Shan, Chen, and Gao [5] utilized locally linear regression for pose-invariant face recognition. Khan and Khan [23] produced pioneered algorithms with high reliability for face localization in multifaceted images.

Identification of skin area is one of the well-known techniques to firstly identify faces in images or video frame. Distinguishing the skin area reduces the time complexity of the face detection algorithms. On the contrary, we want to detect covered faces, so we are proposing an exclusion algorithm, where we have to identify faces in frames by applying the head and shoulder identification; second, if we can detect movement in the video, then we can locate the face; we then exclude the concealed face assumption if skin complexion is detected.

Face detection techniques are classified in four areas: template matching techniques, feature invariant techniques, Knowledge-based techniques, and appearance training-based techniques.

1. *Template matching techniques* In this technique, it saves various patterns that describe faces as a whole face or as facial features, as described by Wang and He [41]. It locates faces utilizing correlation function in respect with a standard face feature. The technique experiences some problems such as the following: it is difficult to characterize typical templates fit for various poses, angles, facial manifestation, and illumination setting.
2. *Feature invariant techniques* It utilizes different facial features such as eyes, nose, and skin color for face detection. This technique, as explained by Song et al. [37], is robust under different lightning settings. At the first step, the technique extracts discriminative local features. In the second step, the technique employs a spatial pyramid to construct a local-holistic face image. It then utilizes a support vector machine for classification.
3. *Knowledge-based techniques* These techniques, as described by Devadethan et al. [8] and [18], utilizes rules extraction for face detection. The rules describe the associations among facial features. These techniques are utilized for face localization. Also, these techniques suffer from low accuracy as the accuracy is greatly affected by rules considered for face detection. Restrictive constraints yield low detection rate while loose rules yield false detection rate. [18] explained the assessment of face detection in angular positions. They assess matched features utilizing minimum distance measures within clusters of the facial area. On the other hand, they utilized maximum distance measures among classes of non-facial areas.

4. *Appearance training-based techniques* These techniques, as described by Molder and Oancea [31], are trained utilizing set of facial images. No predefined templates are required. Convolution neural network face detection utilizes such training methodology. Khan M. Z. et al. [22] utilized scale-invariant transforms for feature extraction. They proposed a training technique that utilizes labeled faces in the wild dataset presented in Huang G. B. et al. [19].

Our research presents a technique for face detection which is based on skin color segmentation using hybrid color model combining the normalized RGB and the YCbCr models. This paper combines the skin segmentation and facial features to detect faces in images. It has been found that by using skin segmentation accuracy, the algorithm is improved and gives the better result than other approaches.

This paper is organized in the following manner: Section 2 presents background around face detection in general and in constrained environment. Section 3 describes the proposed technique. Section 4 describes the proposed hybrid approach for skin detection. Section 5 gives experimental results, while section 6 summarizes the conclusions.

## 2 Background

A face shape prototype with local adaptive morphology handling is presented by Liu, Guo, Liu, Lee, and Yao [29] as an alignment standard to overcome geometric distortion artifacts due to various poses. Du, Hu, Qiao, and Pitas [9] proposed a very well-connected face recognition system utilizing low-rank sparse illustration. Zaman, Shafie, and Mustafah [46] presented a facial recognition system that is robust against different expressions and camera occlusions. Jin, McCann, Froustey, and Unser [21] proposed a deep convolutional neural network to resolve ill-posed inverse image problems.

Zhu, Mai, and Shao [50] utilized a color attenuation methodology for haze removal from hazy images. They introduced linear modelling paradigm for scene depth utilizing a supervised learning method, where the depth information can be modelled. Yu, Bampis, Gupta, and Bovik [45] introduced bi-step image quality estimate methodology which is very important in any image prediction system. According to Fu, Xu, Li, Liu, Ye, and Zhu [13], crowd density estimation is carried utilizing convolutional neural networks, which is very helpful in surveillance system to detect and estimate people in crowds. Wang, Wang, Wang, Zhang, and Qiao [42] proposed scene recognition utilizing local patches. Wu, Lin, Dong, Yan, Bian, and Yang [44] utilized one example methodology for person re-identification through progressive learning.

Chen, Papandreou, Kokkinos, Murphy, and Yuille [6] presented semantic image segmentation utilizing deep

convolutional nets; the proposed methodology gave a very good performance in image segmentation that utilizes semantic information. Zhang K., Zhang Z., Li Z., and Qiao Y. [48] proposed a framework that influenced a cascaded architecture of deep convolutional networks. They indicated high performance in forecasting face location in a coarse-to-fine fashion. Also, Badrinarayanan, Kendall, and Cipolla [2] proposed image segmentation technique using neural network. Li et al. [26] as well, utilized semantic information in maps to solve salient object identification problem; they modelled the semantic attributes of salient objects and utilized convolutional neural network with raw images as input and saliency maps as output. Romera, Álvarez, Bergasa, and Arroyo [36] also focused on extracting semantic information to achieve semantic segmentation in real time. Although Gao, Li, Woo, and Tian [14] emphasized on image segmentation of thermography imaging using genetic algorithms, still their proposal can be extended to normal images. While Liu, Xiao, and Yang [28] focused on edge detection utilizing coastline detection algorithm, they mixed the region and edge active contour methods.

Real time for surveillance systems is very crucial in their success. Awais M. et al. [1] proposed video surveillance system with enhanced accuracy and less computational complexity. Their system comprises face localization and recognition, and it takes real-time videos of faces. The system then extracts key frames and compares it with stored facial images. It uses histogram of oriented gradients (HOG) features. Their simulation results show almost 92% success rates which are comparable with the deep learning approaches, but deep learning techniques have higher computational cost. Ullah H. et al. [40] presented a real-time and new face recognition technique with occlusion. The system utilized 68 points to detect the face in the input image. Linear discriminant techniques are then utilized to extract face features. At the last stage, a classifier with nearest center measures is used. The system is proven to act in real time through experimentation results. Haq M. et al. [17] proposed a novel technique to boost the performance of low-resolution face recognition. Many other articles have studied high-performance face recognition such as the authors Zhang J. et al. [47]. They designed a high-performance face recognition system that utilizes edge computing.

As we surveyed many face detection and recognition techniques, still the most important issue for our research are the systems that recognize masked faces or faces under occlusion. Qezavati H., Majidi B., and Manzuri M. T. [34] introduced a methodology for the detection of partially covered face. It is utilized in surveillance videos containing partially concealed faces including headscarves and eyeglasses. The methodology combines Haar and binary histogram for face classification.

Rajeshwari, Karibasappa, and GopalKrishna [35] surveyed face detection based on skin detection. Liao, Jain, and Li [27] addressed problems in face detection with no prior constraints. They utilized normalized pixel difference image features. These features are extracted by experimental psychology. Bu W. et al. [4] presented a novel cascaded CNN (convolutional neural network) framework to detect masked faces. They also constructed a dataset for masked faces. Ge S. et al. [15] also proposed a LLE-CNNs for occluded face detection. Pre-trained CNNs are utilized to exclude facial regions from the image expressing them with descriptors. These descriptors are converted into similarity-based measures. They tested the system on a large pool of synthesized faces, occluded faces, and also on non-faces. Ghiasi G. and Fowlkes C.C. [16] presented a hierarchical deformable face detection model. They presented occlusions in a structured model. They also enhanced training data with synthetically occluded face images. Nair A. and Potgantwar A. [32] proposed an automated masked person detection in less time.

In a recent study, Ud Din N. et al. [39] proposed mask object removal in face images. They faced challenges because facial masks usually cover a large part of the face, and they also faced the problem of the lack of training datasets for face image with and without mask. They introduced a solution for mask detection. They also utilized a generative adversarial network (GAN) of two discriminators.

### 3 The proposed technique for human detection in surveillance scene

The proposed technique is utilized in surveillance systems. It aims in detecting concealed faces in surveillance images or videos. It comprises many algorithms starting with an algorithm to take images for the surveillance scene under different conditions. The scene under surveillance will detect new objects that enter the scene. The object will be identified as a human being using height and width measurement extraction. If a human being is detected, face and shoulder areas will be extracted utilizing pattern learning from a training phase. Clustering of patches is performed and determines if it is skin patch or a concealed face patch. Human complexion detector using a hybrid technique combining normalized RGB and the YCbCr is proposed. An overview of the proposed technique is depicted in Fig. 1. Head and shoulder detection algorithm is depicted in Fig. 2. The proposed technique comprises many algorithms. Algorithm 1 depicts the training phase for head and shoulder detection. Algorithm 2 is utilized to determine head and shoulder for an unknown image. Algorithm 3 depicts the face detection algorithm.

Algorithm 1: Training phase for head and shoulder detection  
(Output {pattern})

**Start**

```

For i=1 to n (n is a large number of images of head and shoulder images)
{
Read (image i)
    Detect contour of head and shoulder;
    Pattern learning;
    Save the set of patterns {pattern} in the database;
}

```

Algorithm 2: Determine Head-and Shoulder (Input: Clus (i), R  
(Clus (i)))

**Start**

```

{Detect contour (Clus (i));
Find match from {pattern};
If found (match)
Then
    {Output the contour (Clus (i));
    Output rectangular block that contains the contour R (Clus (i));
    Return contour Clus (i), R (Clus (i));}
Else Return null
}
End

```

Algorithm 3: Face Detection Algorithm

**Start**

1. Enter scene id
2. Determine lighting, L :Day or Night or Artificial
3. Read a video frame  $V_{id}(L)$ .
4. Compute histogram of  $V_{id}(L)$ : Histogram ( $V_{id}(L)$ )
5. Take a video of the actual scene now:  $V_{now}$ .
6. For each frame  $i$  of  $V_{now}$ 
  - a. Read a video frame  $V_i$ .
  - b. Compare  $V_i$  with  $V_{id}(L)$  using subtraction algorithm  
Residue= Histogram ( $V_i - V_{id}(L)$ )
  - c. If (Residue – Histogram ( $V_{id}(L)$ ) >  $\epsilon$ ) //  $\epsilon$  is a threshold  
Then  
    {Call **Procedure: Perform Face Detection** ( $V_i$ );  
    Continue;}  
Else  
    {i= i+1;  
    Continue;}

**End**

**Procedure: Perform Face Detection** ( $V_i$ )

**Start**

1. Call procedure: Image Fusion ( $V_i$ ,  $V_{id}(L)$ ,  $IMAGE_{Futilized}$ );
  2. Get histogram of  $IMAGE_{Futilized}$ : Histogram ( $IMAGE_{Futilized}$ )
  3. Compare clusters of  $V_{id}(L)$  and  $IMAGE_{Futilized}$
  4. Determine all new n clusters from step 3 that are found in  $IMAGE_{Futilized}$  and not in  $V_{id}(L)$   
: Clus (i), i= 1 to n
  5. for i = 1 to n  
{Extract Height (Clus (i));  
Extract Width (Clus (i));  
If (Height (Clus (i)) – Average Height (Person) <  $\epsilon_{human}$   
And Width (Clus (i)) - Average Height (Person) <  $\epsilon_{human}$ )  
Then  
    Call Procedure Determine Head-and Shoulder (Clus (i))}
- End**

## 4 The proposed hybrid approach skin detector

The RGB color space is mainly utilized for digital image as explained by Cheng, Liu, and Haifeng [7]. There is a high correlation between the RGB components, which creates sensitivity. To solve the sensitivity issue, each component of the RGB should undergo normalization process: normalized RGB color space by Loesdau et al. [30]. The normalization in Eqs. 1, 2, and 3 helps to reduce the dependency between the RGB components.

$$r = \frac{R}{R + G + B} \quad (1)$$

$$g = \frac{G}{R + G + B} \quad (2)$$

$$b = \frac{B}{R + G + B} \quad (3)$$

HSV model (hue saturation value) is described by Jang and Ra [20] as a discernment model that discriminates between luminance and chrominance components. H, S, and V components are depicted in Eqs. 4, 5, and 6.

```

Start
1. Take images for the surveillance scene under different illumination settings
2. Detect new objects that enter the scene( every few seconds)
3. If the object is not human being
   Then Goto Step 2
   Else Extract Head and Shoulder area
4. Perform Clustering of patches into n patches (patch(i))
5. Concealed-patch =0
6. For i = 1 to n
   If patch (i) is not a skin patch
   Then Concealed- patch = Concealed- patch+1
   End For
7. If Concealed- patch > n/2+1
   Then {
   Concealed-face = True;
   Generate Alarm
   }
   Else Concealed-face =False
End
    
```

Fig. 1 An overview of the proposed technique

$$H = \cos^{-1} \frac{\frac{1}{2} ((R-G) + (R-B))}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \tag{4}$$

$$S = 1-3 \left( \frac{\min(R, G, B)}{R + G + B} \right) \tag{5}$$

$$V = \frac{1}{2} (R + G + B) \tag{6}$$

Another model is *YCbCr* to encode RGB images as introduced by Lei et al. [25]. This model utilizes linear transform to separate illumination and chrominance components. This

model is very effective the detection of human skin. *YCbCr* is calculated as follows in Eqs. 7, 8, and 9.

$$Y = 0.299 R - 0.587 G - 0.114 B \tag{7}$$

$$C_b = R - Y \tag{8}$$

$$C_r = B - Y \tag{9}$$

We simulated the three models using 100 labelled images of skin and non-skin patches; 70 patches were skin and 30 patches were non-skin patches. The computed confusion matrix is depicted in Tables 1, 2, and 3. The results are not convincing enough, as the best one was the *YCbCr*, which gives only 71.4% true-positive detection and 53.3% true-negative detection. We decided to utilize a hybrid model of the normalized RGB and the *YCbCr* as follows in Eqs. 10, 11, and 12.

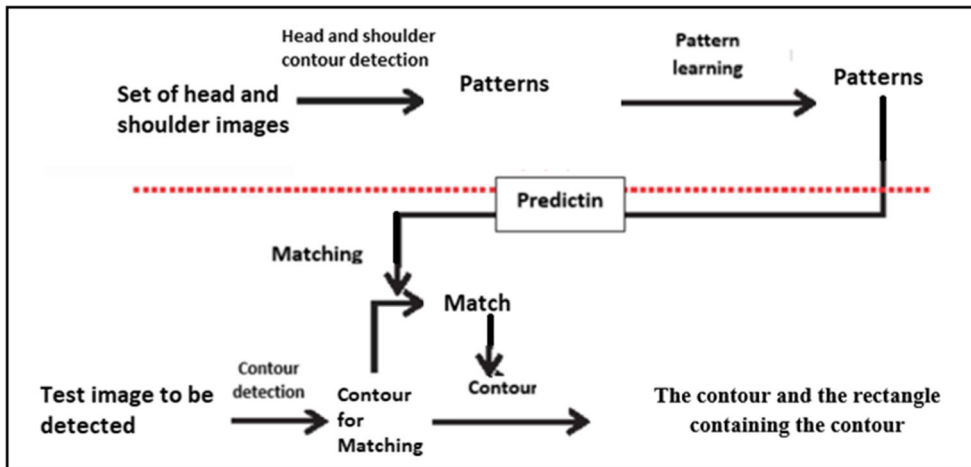
$$Y = 0.299 r - 0.587 g - 0.114 b \tag{10}$$

$$C_b = r - Y \tag{11}$$

$$C_r = y - Y \tag{12}$$

The hybrid model takes advantage of the normalization process of the RGB model to reduce the dependency between the RGB components, and takes advantage of the *YCbCr* model to encode RGB images and to separate illumination and chrominance components. The confusion matrices of the normalized RGB, HSV model, *YCbCr* model, and our proposed hybrid model are shown in Tables 1, 2, 3, and 4. In the experiments, we used 2000 patches; 1000 are human skin patches and 1000 are masked human skin patches. The hybrid model gives 97.3% true-positive detection and 98.2% true-negative detection, outperforming the other three models. The true-

Fig. 2 Head and shoulder training and detection





**Table 1** The normalized RGB

		Predicted cases	
		Positive skin detection	Negative skin detection
Actual cases	Human skin	560	440
	Non-human skin	320	680

**Table 2** HSV model

		Predicted cases	
		Positive skin detection	Negative skin detection
Actual cases	Human skin	680	320
	Non-human skin	240	760

negative percentage is very important for our proposed system, because if we are sure that the area is face and it did not give us skin detection, therefore we can assume the face is covered. The proposed algorithm is depicted in Algorithm 4. Identification of the covered face is depicted

Algorithm 4. Training for presence of human complexion (Clus (i))

Training phase of a human complexion detector is as follows:

**Start**

1. Collect skin patches from different images under different illumination: day; night; artificial illumination
2. Perform the hybrid model on the collected skin patch
3. Extract features using F-measure.
4. Classify the patches according to the F-Measure.

**End**

Algorithm 5. Identification of the covered face (Clus (i))

The Identification of covered face in a given cluster (Clus (i)) is as follows:

**Start**

1. Convert the cluster into the same hybrid model that was utilized in the training phase.
2. Classify cluster using the skin classifier as either a skin or non-skin patch.
3. If Patch is non skin  
Then the face is covered  
Else  
The face is not covered

**End**

## 5 Experimental result

We tested the hybrid model utilizing a set of training images that are selected from two databases, the first one (TAN) is presented by Tan et al. [38], and the second dataset (FvNF) is presented by Nanni and Lumini [33]. The first image database includes 650 skin and non-skin patches. Color images are obtained from the various sources and under different illumination settings. The second dataset (FvNF) is (face vs. nonface), which is composed by 800 face images. This dataset has been collected and used by Nanni and Lumini

in Algorithm 5. The whole system is depicted in Fig 3 a and b.

[33] to evaluate the capability of a skin detector method to detect the presence of a face, based on the number of pixels classified as skin. For our experiments, we created a synthetic dataset by occluding parts of the skin in 450 images in FvNF and labelled them with concealed human set while the rest of the images were left not occluded.

For classification of the skin patches, we applied the FMeasure as stated in Eq. 13. It is shown from Table 5 that the proposed hybrid technique has high detection rate compared with existing models. Metrics such as FMeasure and specificity will be utilized for comparison

**Table 3** YCbCr model

		Predicted cases	
		Positive skin detection	Negative skin detection
Actual cases	Human skin	799	201
	Non-human skin	120	880

between the proposed algorithm (Algorithms 4 and 5) and other algorithms in the literature. Specificity describes the true-negative rate that measures the ratio of actual negatives that are correctly detected as such. FMeasure is a metric of the experiment accuracy. It comprises the precision and the recall. Also, the metrics recall, false-positive ratio (FPR), and false-negative ratio (FNR) are utilized in the experimentation, and the comparison result is depicted in Tables 5 and 6, and Fig. 6. All those metrics are defined in Eqs. 13–18.

$$FMeasure = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)} \tag{13}$$

$$Precision = \frac{TP}{(TP + FP)} \tag{14}$$

$$Recall = \frac{TP}{(TP + FN)} \tag{15}$$

$$Specificity = \frac{TN}{(TN + FP)} \tag{16}$$

$$FPR = \frac{FN}{(FN + TP)} \tag{17}$$

**Table 4** Our proposed hybrid model

		Predicted cases	
		Positive skin detection	Negative skin detection
Actual cases	Human skin	973	27
	Non-human skin	18	982

**Training phase**

**Start**

1. Load image from database
2. Implement color transformation from RGB to the Hybrid model
3. Perform skin detection using F-Feature
4. If (F-Feature < threshold)
  - Then label as Non-Skin
  - Else
    - {
    - Perform skin extraction process;
    - Perform classification;
    - Output classification Rules;
    - }

**End**

**a**

**Covered Face Detection**

**Start**

1. Load image from database
2. Perform color transformation from RGB to Hybrid model
3. Perform skin detection using F-Feature
4. If (F-Feature < threshold)
  - Then
    - {
    - Label as Non-Skin;
    - Output "The face is covered"
    - }
  - Else
    - {
    - Label as Skin;
    - Output "The face is uncovered"
    - }

**End**

**b**

**Fig. 3** **a** Classification of skin and non-skin patches. **b** Detection of concealed faces

$$FNR = \frac{FP}{(FP + TN)} \tag{18}$$

Some of the used skin patches, clustering, and skin patched after concealment (used in actual experiments) are depicted in Fig. 4 (dataset TAN) and Fig. 5 (dataset FvNF). The proposed hybrid approach achieved high-concealed face detection performance for frontal faces. Performance is degraded when images contain non-frontal faces and for dark skin images if concealed with brown cover. The hybrid approach achieves average specificity of 96.8%, which is an enhancement in specificity

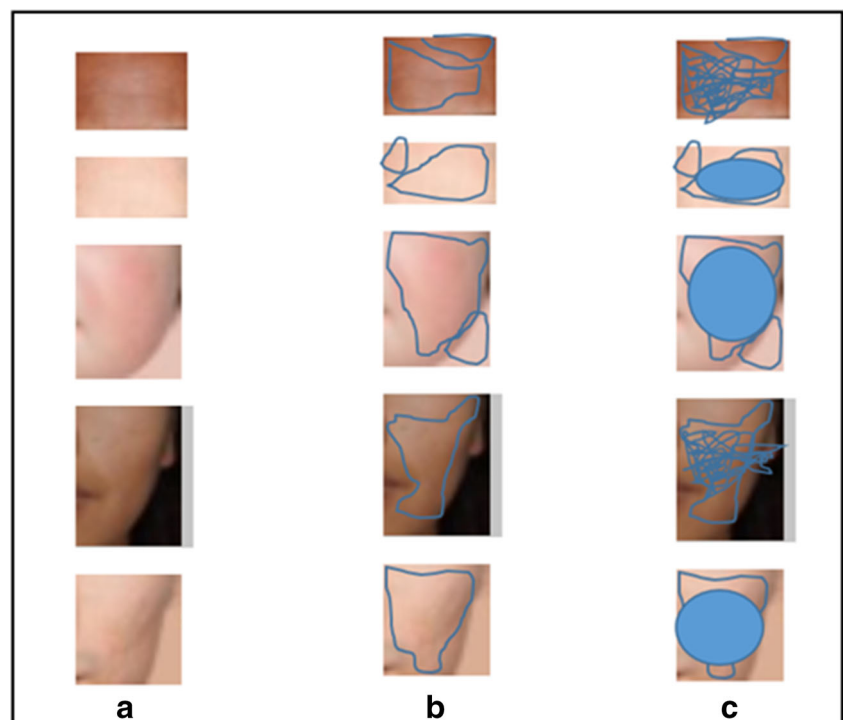
**Table 5** The performance measures of the proposed hybrid approach against other approaches using the TAN dataset of 600 face images

Technique	Average FNR	Average FPR	Average FMeasure	Average specificity	Average detection rate
The normalized RGB	0.4244	0.4814	0.5844	0.5844	0.5449
HSV model	0.4194	0.2544	0.6804	0.6804	0.6449
YCbCr model	0.1288	0.0999	0.7288	0.7561	0.6859
Proposed hybrid technique	0.0512	0.0434	0.9498	0.9688	0.9751

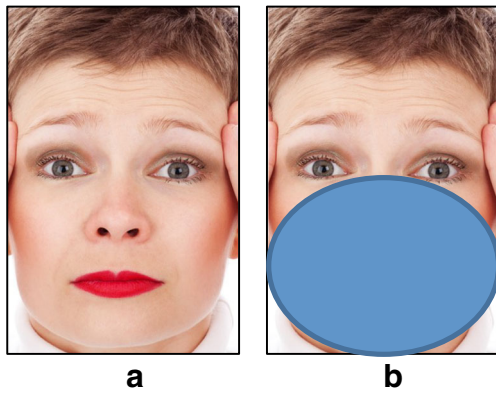
**Table 6** The performance measures of the proposed hybrid approach against other approaches using the FvNF dataset of 450 images

Technique	Average FNR	Average FPR	Average FMeasure	Average specificity	Average detection rate
The normalized RGB	0.431	0.481	0.591	0.582	0.545
HSV model	0.419	0.255	0.681	0.681	0.645
YCbCr model	0.127	0.099	0.727	0.757	0.686
Proposed hybrid technique	0.06	0.051	0.945	0.971	0.989

**Fig. 4** **a** Patches of skin. **b** Clustering. **c** Concealing (dataset TAN)







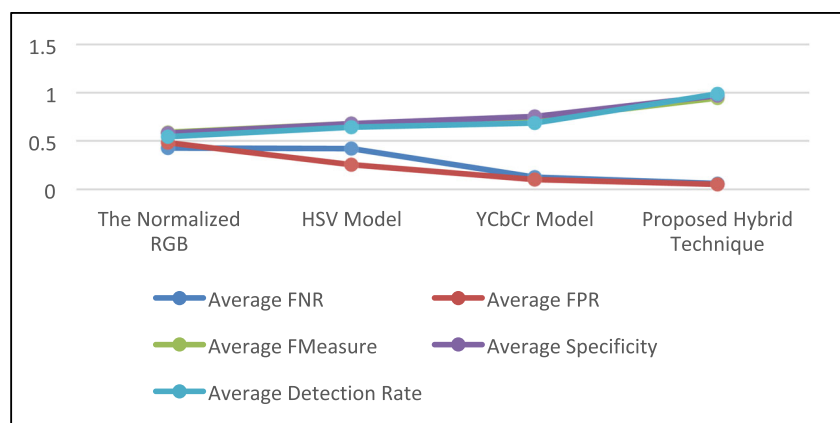
**Fig. 5** a Example of an image from FvNF. b Synthetic masked image

by 28%, and average detection rate of 97.5%, which is an enhancement by 42.12% than the second best algorithm *YCbCr* model (Fig. 6). While FMeasure is a measure that combines precision and recall, the proposed hybrid approach achieves enhancement by 30%.

In Table 7, we compared some of existing systems for concealed face detection. Different features are sought and compared with our proposed system. The systems that we compared are as follows:

1. GAN: A novel generative network for unmasking of masked face. It was introduced by Ud Din N. et al. [39].
2. Head scarf : Qezavati H., Majidi B., and Manzuri M. T. [34] detected partially covered face with headscarf.
3. LLE: Ge S. et al. [15] succeeded in detecting masked faces with LLE-CNNs.
4. Occlusion coherence: Ghiasi G. and Fowlkes C.C. [16] introduced a local occluded faces utilizing a hierarchical deformable model.

**Fig. 6** The performance measures of the proposed hybrid approach against other approaches using the FvNF dataset of 450 images



5. Viola: Nair A. and Potgantwar A. [32] proposed a masked face detection using the viola algorithm.

We carried experiments to determine the runtime of our proposed approach. It is very crucial to carry our algorithm in real time. We compared our implementation with Viola (Nair A. and Potgantwar A. [32]) and GAN (Ud Din N. et al. [39]). The Viola system detects people’s face and determines if it is masked or not in video based setting. The Viola system comprises several phases such as distance from camera, identifies the eye line and the face, and finally detects if the face is masked. While with GAN, we compared part of our proposed system (masked face detection) with the first phase of GAN which includes detection of mask in images. We added the detection of face and shoulder to GAN. Both systems are known for their real-time occluded face detection.

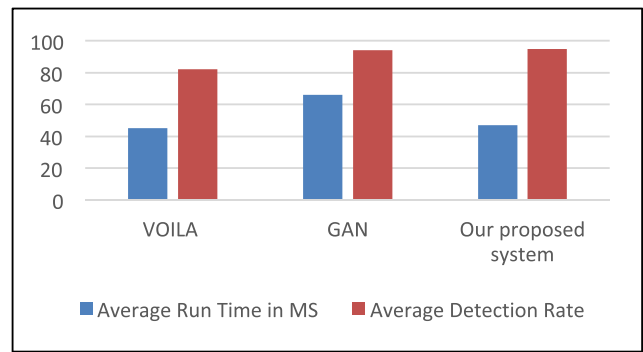
In Fig. 7, we show average runtime comparison combined with average detection rate. We used 450 images from the FvFN database with synthetic mask.

As shown in Fig. 7, Voila has a slight less average runtime than our proposed system but with lower detection rate. Our system has 12% more detection rate while having average runtime of 8% which is still considered real time. Also, GAN has more average runtime than our proposed system but with comparable detection rate. Our system is 30% faster than GAN.

Figure 8 presents the average runtime for VOILA, GAN versus, and our proposed system for 11 different masked face images.

**Table 7** Comparison of existing systems for concealed face detection

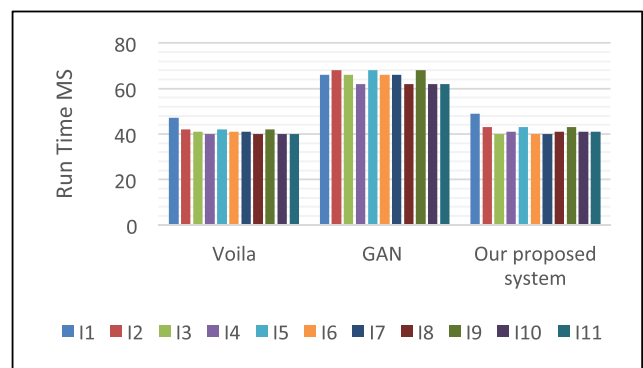
System	Goal	Input	Methodology	Output	Data set
GAN Ud Din N. et al. [39]	Mask detection and removal from facial images	Masked facial image	Feeding the input image and the mask's binary map into a GAN network	An image without the mask	Synthetic Dataset
Head scarf Qezavati et al. [34]	Partially covered face detection	Face with headscarf	Feature extraction using Haar cascade	Face detection in videos	A large dataset of a crowded office environment in Middle East
LLE Ge S. et al. [15]	Extraction and illustration of face candidates	Masked faces	Cascading two CNNs for feature extraction	A 4096 descriptor	MIFA
Occlusion coherence Ghiasi G. and Fowlkes C.C. (2014)	Face detection with occlusions	Occluded faces	Augmentation training data with large instances of synthetically occluded faces	Sets featuring significant occlusion	Synthetically occluded instances
Viola Nair A. and Potgautwar A. [32]	To detect masked person in less time	Videos with masked faces	A progressive approach of masked face detection: for less time consumption	Region of masked faces	Real time inputs
Our proposed method for concealed face	Detection of presence or absent of human skin	Videos	A hybrid approach that combines normalized RGB (rgb) and the YCbCr space color	Detection of masked people	Synthetically occluded instances



**Fig. 7** Average runtime and average detection rate for VOILA, GAN, and our proposed system

## 6 Conclusions

This paper proposes a robust approach for concealed face detection under different camera angle and illumination settings. Human skin patches are identified by a hybrid non-linear transform model that combines the RGB (rgb) color space model and the YCbCr color model. The concealed face is detected by negating the presence of skin patches in a true identified face. The novel technique for concealed face detection based on complexion detection to challenge a concealed face assumption. The proposed algorithm first determine of the existence of a human being in the surveillance scene. Head and shoulder contour is detected. The face will be clustered to cluster patches. Then determination of presence or absent of human skin will be determined. We proposed a hybrid approach that combines normalized RGB and the YCbCr space color. This technique is tested on two datasets; the first one contains 650 skin patches, and the second one contains 800 face images. Masks are synthesized on 60% of the images in the data sets. The algorithm achieves an average masked face detection rate of 97.51% for concealed faces in real time.



**Fig. 8** Average runtime in MS for VOILA, GAN versus, and our proposed system for 11 different masked face images

**Funding information** This research project was funded by the Deanship of Scientific Research, Princess Nourah bint Abdulrahman University, through the Program of Research Project Funding After Publication, grant NO (40 - PRFA - P - 2).

## References

- Awais M et al (2019) Real-time surveillance through face recognition using HOG and feedforward neural networks. *IEEE Access* 7: 121236–121244
- Badrinarayanan V, Kendall A, Cipolla R (2017) SegNet: A deep convolutional encoder–decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell* 39(12):2481–2495
- Borgi MA et al (2014) Regularized shearlet network for face recognition using single sample per person. *ICASSP IEEE International Conference on Acoustics Speech and Signal Processing*, no 6853649:514–518
- Bu W et al (2017) A cascade framework for masked face detection. In: 2017 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM), Ningbo, pp 458–462
- Chai X, Shan S, Chen X, Gao W (2007) Locally linear regression for pose-invariant face recognition. *IEEE Trans Image Processing* 16(7):1716–1725
- Chen L-C, Papandreou G, Kokkinos I, Murphy K, Yuille AL (2017) DeepLab: semantic image segmentation with deep convolutional nets atrous convolution and fully connected CRFs. *IEEE Trans Pattern Anal Mach Intell* 40(4):834–848
- Cheng Y, Xu L, Li H (2007) Uniform color space for color storage. In: 2007 Asia Optical Fiber Communication and Optoelectronics Conference, Shanghai, pp 555–557
- Devadethan S et al (2014) Face detection and facial feature extraction based on a fusion of knowledge based method and morphological image processing. In: 2014 Annual International Conference on Emerging Research Areas: Magnetics, Machines and Drives, Kottayam, pp 1–5
- Du HS, Hu QP, Qiao DF, Pitas I (2015) Robust face recognition via low-rank sparse representation-based classification. *Int J Autom Comput* 12(6):579–587
- Du M, Sankaranarayanan AC, Chellappa R (2001) Robust face recognition from multi-view videos. *IEEE Trans Image Process* 23(3):1105–1117
- Ejbalı R, Zaided M, Ben AC (2014) Face recognition based on beta 2D elastic bunch graph matching. In: 13th International Conference on Hybrid Intelligent Systems art, no. 6920461, pp 88–92
- Filali Y., & et al. (2019) Texture Classification of skin lesion using convolutional neural network. 2019 International Conference on Wireless Technologies, Embedded and Intelligent Systems, Morocco, pp. 1-5.
- Fu M, Xu P, Li X, Liu Q, Ye M, Zhu C (2015) Fast crowd density estimation with convolutional neural networks. *Eng Appl Artif Intell* 43:81–88
- Gao B, Li X, Woo WL, Tian GY (2018) Physics-based image segmentation using first order statistical properties and genetic algorithm for inductive thermography imaging. *IEEE Trans Image Process* 27(5):2160–2175
- Ge S et al (2017) Detecting masked faces in the wild with LLE-CNNs. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, pp 426–434
- Ghiasi G. & Fowlkes C.C. (2014) Occlusion coherence: localizing occluded faces with a hierarchical deformable part model. *IEEE CVPR*, pp. 1899-1906.
- Haq M et al (2019) Boosting the face recognition performance of ensemble based LDA for pose non-uniform illuminations and low-resolution images *KSII Transactions on Internet and Information Systems*, p 2019
- Hashemi SM, Reza et al (2019) An overview of face detection methods in angular positions. In: 2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI), Tehran, Iran, pp 484–491
- Huang GB et al (2008) Labeled faces in the wild: a database for studying face recognition in unconstrained environments. In: Proc. Workshop Faces ‘Real-Life’ Images Detection Alignment Recognition, pp 1–11
- Jang JH, Ra JB (2008) Pseudo-color image fusion based on intensity-hue-saturation color space. In: 2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, Seoul, pp 366–371
- Jin KH, McCann TM, Froustey E, Unser M (2017) Deep convolutional neural network for inverse problems in imaging. *IEEE Trans Image Process* 26(9):4509–4522
- Khan MZ et al (2019) Deep unified model for face recognition based on convolution neural network and edge computing. *IEEE Access* 7:72622–72633
- Khan NA, Khan MH (2004) Towards reliable face localization in complex images. In: Proceedings of INMIC 2004, Lahore, pp 3–8
- Kim Y, Hwang I, Cho NI (2017) Convolutional neural networks and training strategies for skin detection. In: 2017 IEEE International Conference on Image Processing, Beijing, pp 3919–3923
- Lei L et al (2013) Image retrieval based on YCbCr color histogram. In: 2013 IEEE 12th International Conference on Cognitive Informatics and Cognitive Computing, New York, NY, pp 483–488
- Li X et al (2016) DeepSaliency: Multi-task deep neural network model for salient object detection. *IEEE Trans Image Process* 25(8):3919–3930
- Liao S, Jain AK, Li SZ (2016) A fast and accurate unconstrained face detector. *IEEE Trans Pattern Anal Mach Intell* 38:211–223
- Liu C, Xiao Y, Yang J (2017) A coastline detection method in polarimetric SAR images mixing the region-based and edge-based active contour models. *IEEE Trans Geosci Remote Sens* 55(7): 3735–3747
- Liu Y, Guo J, Liu P, Lee J, Yao C (2018) Panoramic face recognition. *IEEE Transactions on Circuits and Systems for Video Technology* 28(8):1864–1874
- Loesdau M et al (2017) Chromatic indices in the normalized rgb color space. In: 2017 International Conference on Digital Image Computing: Techniques and Applications, Sydney, NSW, pp 1–8
- Molder C, Oancea R (2012) Appearance-based facial detection for recognition. In: 2012 9th International Conference on Communications, Bucharest, pp 119–122
- Nair A, Potgantarwar A (2018) Masked face detection using the Viola algorithm: a progressive approach for less time consumption. *IJES* 6(4)
- Nanni, Lumini (2017) FvNF, Skin detection for reducing false positive in face detection, V.M. Petrova (Ed.), *Adv. Eng. Res.*, Vol.16, Nov. Publ. 2017, Nova Publisher. <https://www.novapublishers.com/>
- Qezavati H, Majidi B, Manzuri MT (2019) Partially covered face detection in presence of headscarf for surveillance applications. In: 2019 4th International Conference on Pattern Recognition and Image Analysis (IPRIA), Tehran, Iran, pp 195–199
- Rajeshwari J, Karibasappa K, GopalKrishna M (2014) Survey on skin based face detection on different illumination poses and occlusion. In: 2014 International Conference on Contemporary Computing and Informatics (IC3I), Mysore, pp 728–733
- Romera E, Álvarez JM, Bergasa LM, Arroyo R (2018) ERFNet: efficient residual factorized ConvNet for real-time semantic segmentation. *IEEE Trans Intell Transp Syst* 19(1):263–272

37. Song T et al (2014) Face Recognition based on scale invariant feature transform and spatial pyramid representation. In: 2014 IEEE International Conference on Systems, Man, and Cybernetics, San Diego, CA, pp 4113–4118
38. Tan, et al. (2012) Image and skin Database. IEEE T-II. [http://cs-chan.com/downloads\\_skin\\_dataset.html](http://cs-chan.com/downloads_skin_dataset.html)
39. Ud Din N et al (2020) A Novel GAN-Based Network for Unmasking of Masked Face. *IEEE Access* 8:44276–44287
40. Ullah H et al (2019) A robust face recognition method for occluded and low-resolution images. 2019 International Conference on Applied and Engineering Mathematics (ICAEM). Taxila, Pakistan 2019:86–91
41. Wang R, He X (2019) Face detection based on template matching and neural network. In: 2019 International Conference on Communications, Information System and Computer Engineering (CISCE), Haikou, China, pp 547–550
42. Wang Z, Wang L, Wang Y, Zhang B, Qiao Y (2017) Weakly supervised patchnets: describing and aggregating local patches for scene recognition. *IEEE Trans Image Process* 26(4):2028–2041
43. Wu QX et al (2012) Skin detection using color processing mechanism inspired by the visual system. In: IET Conference on Image Processing, London, pp 1–5
44. Wu Y, Lin Y, Dong X, Yan Y, Bian W, Yang Y (2019) Progressive learning for person re-identification with one example. *IEEE Trans Image Process* 28(6):2872–2881
45. Yu X, Bampis CG, Gupta P, Bovik AC (2019) Predicting the quality of images compressed after distortion in two steps. *IEEE Trans Image Process* 28(12):5757–5770
46. Zaman FK, Shafie AA, Mustafah YM (2016) Robust face recognition against expressions and partial occlusions. *Int J Autom Comput* 13(4):319–337
47. Zhang J et al (2019) Design and implementation of a face recognition system based on edge computing. 2019 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS). Shenyang, China 2019:363–366
48. Zhang K, Zhang Z, Li Z, Qiao Y (2016) Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters* 23(10):1499–1503
49. Zhao W, Chellappa R, Phillips PJ, Rosenfeld A (2003) Face recognition: a literature survey. *J ACM Comput Survey* 35(4):399–458
50. Zhu Q, Mai J, Shao L (2015) A fast single image haze removal algorithm using color attenuation prior. *IEEE Trans Image Process* 24(11):3522–3533

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.