

FACE DETECTION TECHNIQUES USING DATABASE IMAGES

Manisha M. Kasar¹, S.H. Patil², Debnath Bhattacharyya^{3*} and Jason Levy⁴

¹*Department of Information Technology,
Bharati Vidyapeeth Deemed to be University College of Engineering,
Pune-411043, India*

²*Department of Computer Engineering,
Bharati Vidyapeeth Deemed to be University College of Engineering,
Pune-411043, India*

³*Department of Computer Science and Engineering,
Vignan's Institute of Information Technology,
Visakhapatnam-530049, India*

⁴*University of Hawaii, Hawaii, USA*

kasarmanisha@gmail.com, shpatil@bvucoep.edu.in, debnathb@gmail.com,
jlevy@hawaii.edu

Abstract— The paper presents a review of proposed face detection techniques. Face Detection using captured face pictures still a complicated task in the area of image processing, and presently it's an operative area of analysis. Automatic face detection is a critical process in face recognition, facial expression recognition. Using an ineffective algorithm, there will be a negative impact on the performance of face detection and face recognition. The face detection technique is not only a step in face recognition technique; however, it is also a self-sufficient widely-used technique. Developing a sustainable feature for face detection technique is necessary. Face detection has attracted massive attention as a result of its several demands in computer vision, pattern recognition, and Biometrics. Face detection may be a technique to analyze a face from several images that have many features in that image. The face detection process is necessary for facial emotion recognition, movement of the face, and posture assessment. Face detection may be a difficult task as a result of looks don't seem to be clear, and it may modify the length, height, width, appearance, tone, etc. Face detection becomes an additional difficult process whereas inputted image not visible and hidden by some various factors and not at all correct light conditions, etc. The suggested technique should be capable of analyzing the facial region including high accuracy, efficiency, and minimum errors.

Keywords— Face detection, face recognition, facial expression recognition, image processing

1. INTRODUCTION

Face detection has the aim to search out whether the face is present in an available image or not. This is the first process in facial expression recognition area. Face detection from the frontal face and partial frontal face, the face detection process has been widely

Received: April 16, 2019

Reviewed: May 31, 2019

Accepted: June 3, 2019

* Corresponding Author



used. For an operating image, to compact an image, improve the quality of an image or to take out any critical details against the image, an image processing technique will be used. Because of the broad utilization of computer and people interaction now a day's face detection has been rigorously studied. A face detection system composed of three stages: image acquisition, face detection, and image categorization shown in the following figure.

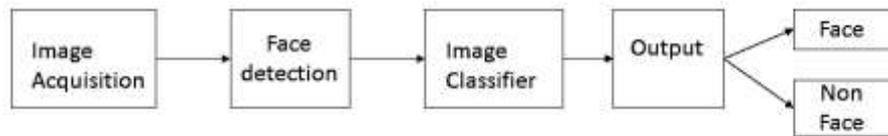


Fig. 1 Architecture of Face Detection [1]

The Fig.1 shows architecture of face detection. Image acquisition phase is used to acquire face image from user input using the face database or model captured by the camera. Images may change in pattern, scale, and resolution and may have frames of videos. Then some image-processing methods will apply to remove unbidden noise, lighting condition to normalize the image. Then classifier will decide either the image is in the face or non-face category depends on knowledge accomplished from the image in training.

Face detection is used to analyze either any face image is apparent in image or not. At present specific method has available to investigate a face from the image. A unique technique is viola Jones technique [1] is used to identify a face, to take away the other information from that original image, a method should thoroughly analyze a complete image. In the face detection process, there are two types of mistakes may be examined first is the false negative rate [1] and another is a false positive rate [1]. Inside a false positive area of an image is observed that image doesn't found the face still a face is available but in a false negative rate, faces are not visible due to minimum detection rate. Every method have some characteristics like detection rate [1], knowledge acquisition period, Implementation period and a proportion among the analysisfrequency false values [1]. The system honestly analyzes the correlation among the bunch of faces examined by a person, and the bunch of faces is called a detection rate [1].Some people work on face detection to detect how the condition affects biometric functions. Several images feature that act on the detection or recognition performance. There is an unacceptable difference among scene-based ultimatums like unclear information of image, lighting conditions, etc., the methods proposed by some researchers are explain below.

The technology recommends by Shengcai Liao, Anil K. Jain, and Stan Z. Li [7]. He suggested that NPD method has achieved good accomplishment in analyzing uncontrolled faces including random posture changes and obstructions with disarrange positions, but there are some issues in this method. Obstructions and indistinct shape these two are the important exceptions for uncontrolled face detection; this is the limitation of NPD method. Some of these issues were solved using Normalizes Pixel Difference (NPD) [2], and haar classifier [2] hybrid method and Deep Convolutional Neural Network suggested by Neethu A, Athi Narayanan S, Kamal Bijlani [2] and Shuo Yang, Ping Luo, Chen Change Loy [13], and Xiaoou Tang [13]. To improve the correctness of face analysis and to minimize the false detections rate, haar-classifier [2] is applicable as another detector. A conjunction of the face mentioned above detector may increase the face detection execution mainly in uncontrolled conditions, but there is some issue in this method. Video face detection is also a major challenge for this hybrid method which has not been addressed by NPD and haar classifier method. Some of these issues are solved by Haar-like features [2], motion and color filtering and hard hat color detection method suggested by Shan Du, Mohamed Shehata and WaelBadawy [15]. This

method has various video processing methods to build architecture for quick and strong hard hat detection [14] in fabrication areas. The suggested method may analyze the face and hard hat in the actual measure. But there are some issues in this method. The forehead part the hair, this feature has not tested. This issue has been addressed by PASCAL VOC Precision-recall protocol for object detection [17] method suggested by Xiangxin Zhu Deva Ramanan [17]. This method gives a combined prototype for face detection, posture analysis and localization of certain key points using a composition of trees with a combine pool of sections. These model are unexpectedly functional in catching universal elastic deformation, Even though it makes it simple to gain but some limitations are in this method like data size used for face detection is less this issued addressed by Viola-Jones algorithm[1] suggested by JatinChatrath, Pankaj Gupta, Puneet Ahuja, Aryan Goel, ShaifaliM.Arora [18]. The researcher suggested a technique for face detection and tracking with minimum computation time and achieved high detection rate and tracking accuracy. But still, there are some issues of extracting features form face for face detection which is addressed by AdaBoost classifier method suggested by Cheng Jiang, Yifeng Zhang [19].After studying this paper, we analyze the accuracy of the proposed method has increased at some level, and it is robust in rotation and scaling transformation. But still, there is an issue of finding face features including background knowledge which is addressed by 2.5 Two-dimensional Gaussian function method suggested by Zhu Youlian, Huang Cheng, Zhai Kun, Pan Lingjiao [10]. This technique has better acceptability for various facial images and the considerable changes in angels of the face. Furthermore, a different facial pattern is used to match with all feasible area of the face image. But still there is the issue of multi-view face detection which is addressed by Seyed Mohammad Hassan Anvar, Wei-Yun Yau, and EamKhwang Teoh, The training procedure needs manually labeled images of all positions. A moving window of various sizes entirely finds out the given image to check the presence of a face. This process will redo for each posture of the face by handling it across the group of classifiers. This kind of techniques are moderate and operate well only on the limited poses will train. Hence, the ability of the proposed technology will just spread with other varying images so long as the interclass dissimilarity is not so vast that enough correlation points cannot found. All these are mention below section.

2. EARLIER FACE DETECTION TECHNIQUES

Now a day's, the advanced techniques have to keep remarkable input to the theory and exercise of automatic face analysis. Presently, various methods available for face analysis, some of the published techniques have concentrated on face detection of the single image. Proposed research on face detection summarized in the below sections. Each method has some advantages and limitations. Developing much systematic face detection technique is the target of several researchers.

2.1 NORMALIZED PIXEL DIFFERENCE (NPD):

Normalized Pixel Difference (NPD) method has suggested by Shengcai Liao, Anil K. Jain, and Stan Z. Li [14] in 2016. The Normalized Pixel Difference (NPD) components among a pair of pixels in an image are determined in the equation (1) [7]

$$f(x, y) = \frac{x-y}{x+y} \quad (1)$$

Where $x, y \geq 0$ are darkness and brightness values of the pair of pixels, and $f(0, 0)$ is determine as 0 when $x = y = 0$. The NPD characteristics calculate the comparative dissimilarity among the pair of pixel values. The sign of $f(x, y)$ exhibits the ordinal correlation among the pair of pixels x and y , and the volume of $f(x, y)$ computes the comparative dissimilarity among x and y . A various variety of characteristics, known as

NPD that is well organized to calculate and has some advantageous characteristics, which includes scaling invariable, enclosed, and allow to the reassembling of the or actual image. A deep quadratic tree [7] beginner is predicted to find out the most appropriate set of NPD characteristics to increase their dissimilarities. Likewise, just a one only soft-cascade AdaBoost [7] classifier is necessary to hold uncontrolled faces including obstructing and random posture, in the absence of posture labeling or grouping in the training step. Secondly, a deep quadratic tree learning [7] technique and build a single soft-cascade AdaBoost classifier [7] to pick up multiple composite face and arbitrary posture and occlusion constraints. A period of time unique NPD characteristics could have "weak" discriminating capacity, researcher work shows that a set of NPD characteristics most appropriate, understand and merge to build a large number of particular characteristics in a wide quadratic tree. Likewise, various kinds of faces could be naturally separated within separate leaves of a tree classifier, along with the compound face manifold in excessive extent area may be dividing up in the gain and understanding of procedures. It is a "divide and conquers" approach to fear uncontrolled face detection in one classifier. The resultant face detector is prosperous to dissimilar in posture, occlusion, and variation in lighting condition, also unclear and lowest image determination rate [7].

2.2 SKIN DETECTION USING TWO COLOR SPACES HSV AND YCGCR:

This method has proposed by Bashir Muhammad, Syed Abd Rahman Abu-Bakar [8] in 2015. A skin determination technique suggests that make a combination of two color spaces Hue, Saturation, and Value (HSV) and luminance, chrominance in green, and chrominance in red (YCgCr) [8]. The HSV is an intuitive color space mainly exhibited as a non-linear integration of the RGB values. The conversion from RGB to HSV is shown in the equation (2) (3) (4) (5) [8].

$$v = \frac{R+G+B}{3} \quad \dots (2)$$

$$s = 1 - \frac{3*\min(R,G,B)}{R+G+B} \quad \dots (3)$$

$$H = \begin{cases} \cos^{-1}(z) & \text{if } (G \geq R) \\ 2\pi - \cos^{-1}(z) & \text{if } (G < R) \end{cases} \quad (4)$$

Where

$$z = \frac{(2B-G-R)}{\sqrt{(B-G)^2 + \frac{(B-R)(G-R)}{2}}} \quad (5)$$

The HSV spaces are beneficial for image processing as they segregate the color details in methods that equivalent to the person visual system. Skin determination outcome presents that suggested technique may answer effectively to various skin color modulation including a minimum sensibility to the skin has background pixels. The luminance (Y) element is estimated as a weighted average of RGB values although the chrominance (Cb and Cr) elements calculated by subtracting the luminance element from B and R values. That is given in the equation (6) (7) (8) [8]:

$$y = 16 + 65.481R + 128.533G + 24.966B \quad (6)$$

$$Cb = 128 - 37.79R - 74.203G + 112B \quad (7)$$

$$Cr = 128 + 112R - 93.78G - 18.214B \quad (8)$$

Where R, G, and B elements varies from 0 to 1.

Using HSV and YCbCr color spaces we can build skin analyzer which depends on the S, Cg and Cr elements. The S element only divides skin pixels finely, and moreover, the

Cg and Cr element achieves effective rate for skin detection. These three elements are applied here for better skin detection. The detector accepts RGB image as input then it is converted into the SCgCr space, and every pixel has arranged in classes as of whether it will be skin or other than that. A skin detection study reports that the S element singly can separate skin pixels exactly and also the Cg and Cr elements accomplish effectively for skin detection. The three items are available for excellent skin detection manipulating the below-expected threshold [8]:

$$0.05 \leq S \leq 0.64$$

$$104 \leq Cg \leq 126$$

$$141 \leq Cr \leq 170$$

The detector gives as an input in the form of RGB image that converted to the SCgCr space, and all pixels are arranged in a group using the accepted threshold as either skin or otherwise. The suggested skin detector established on the S, Cg and Cr elements from the HSV and YCbCr color spaces. Whereas images consist of too many people, a few skin areas may be connected, and usage of edge detection may be useful to divide complained skin areas. The Canny edge detection technique is more useful on grayscale image accepting a high threshold of 0.45 [8]. Edge and divided binary images are well adjusted, and additional morphological procedures like erosion, holes filling and dilation [8] are put in to gain the last skin patches. The outcome indicates that the suggested technique can divide skin patches of separate skin tones including minimum consideration to skin-like surrounding pixels. The extreme face analysis measure of the recommended technology is a symptom of preferable skin regionalization potentiality and therefore may be functional for face detection approaches.

2.3 REAL-TIME MULTI-FACE DETECTION SYSTEM:

Nai-Jian Wang, Sheng-Chieh Chang, Pei-Jung Chou [9] has proposed this method in 2012. The YCbCr space is used to perform skin color detection. The CbCr distribution between different faces and skin colors are distressed and indicate small variations in creating it is not difficult to differentiate skin or non-skin colors. In the area of face detection including the eye knowledge and the differentiation among lip color and position are examined. Here, through skin color detection the face candidate region is estimated which shows in the following figure [9].

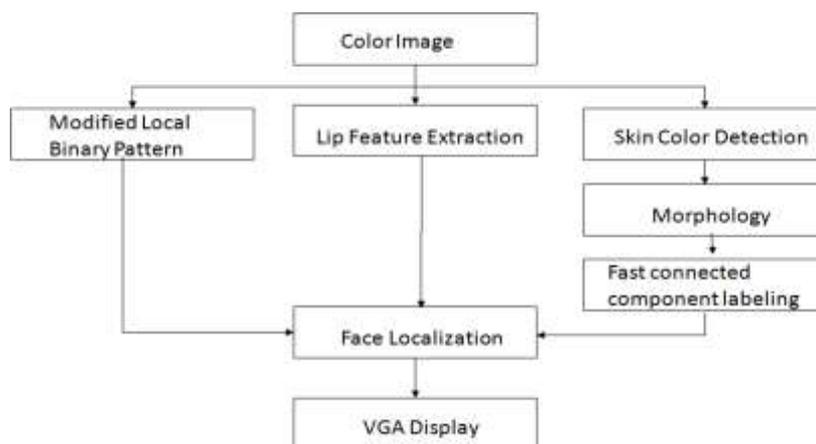


Fig. 2 Process of face detection [9]

The Fig. 2 shows the overall process of face detection. Initially, a color image has given as input then skin color detection procedure is applied to the inputted image using

the YCbCr color space method. The CbCr distribution during different faces and skin color are focused and display some discrepancies, creating it more straightforward to differentiate skin and non-skin colors. YCbCr space separated the information of lightness such as the computation outcomes are not a concern to the fluctuation of surrounding lightness. YCbCr space can be easily used to determine the color range. Through the analysis, the researcher has defined the scope of CbCr as shown in the following equation. Skin color could explain if it falls within the range is given in equation (9) [16].

$$\text{Skin} = \begin{cases} 255, & 80 < Cb < 120 \text{ and } 140 < Cr < 165 \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

After detection of skin color, two primary morphological operation dilation and erosion can be applied, and then a fast connected labeling algorithm is used to speed up the constructed relationship by employing a pre-arranged table. The time needed to analyze whether the two adjacent are considerably minimize due to the reduction in some iteration. As per the analysis for a distention point, we have 16 available conditions by its four adjoin points the left, upper left, upper and upper right. Afterword the finishing of the primary categorization, the fast connected-element marking technique can accomplish easily. For target point the code of the location in left, upper left, a top, and high right sequence can be determined as C1, C2, C3, and C4 correspondingly in the label assign to the technique, there are three categories. First, $m=m+1$ presents a new object to the current code allocated. Second, no working means selecting the system from its next point, and the third used to compare the code of its neighboring points. First, the targeted code is the same as C3's if C3 has a valid code. There are eight circumstances in this step, and the rest can accomplish in the same manner. The 16 cases can check through the label assign algorithm. Here they are using "else if" statement the decision process will stop as long as it meets the requirement. To make the result correctly, two scans the connected labeling algorithm requires [9]. However, meanwhile, the mouth portion is defined as a skin area also. Therefore we have to detect further methods to analyze the lip area. Lip color of the face usually is dipper than neighboring skin color, the importance of lip color in the RGB color space modifies significantly, but a change in the R/G ration doesn't. Color information will use to choose the possible lip candidate points and could be estimated using the equation (10) [9]:

$$\text{Lip} = \begin{cases} 1, & 1.7 < \frac{R}{G} < 2.5 \text{ and } 50 < Y < 220 \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

A converted LBP suggested to overthrow the imperfection of edge detection. Researcher changes the typical LBP commands using the equation (11) as express in figure [9].

$$M_LBP(x) = \sum_{(p=0)}^7 (g_n - g_\lambda) 2^p, s(u) = \begin{cases} 1, u > 10 \\ -1, u < -10 \\ 0, \text{otherwise} \end{cases} \quad (11)$$

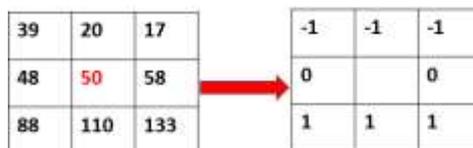


Fig. 3 Result of M_LBP encoding [9]

The Fig. 3 shows result of M_LBP encoding which Considers the calculation in diagram across the equation, we will get an encircling sample digit 01110-1-1-1 and

transform them into fraction value 105. There are 320*240 merits for 320*240 images while doing M_LBP for every pixel in an image that is dissimilar from conventional LBP [9]. The algorithm will be improved when the concept of image processing and hardware design combines into the design, and hardware specifications have taken into consideration. The hardware method may be smoothly put it into effect, and it may reduce the processing time and the hardware resources.

2.4 VIOLA JONES ALGORITHM:

Viola Jones algorithm has proposed by DivyaMeena, Ravi Sharan [1] in 2016. To analyze the face and principal component analysis [1] for face recognition Viola-Jones algorithm will use, and the procedure of face detection for an individual image various steps were used as shown below [1]:

1. The image is captured by a camera or by video sequence and keep in a database. Further images may be useful later that is held earlier in a dataset.
2. The image may read using imread function which is used by the author. This technique removes the shape of the feature for the nose, chin, mouth, eye, eyebrows, and head position.
3. The prime objective of the face detection step is to recognize the face location in an image and identify it. Furthermore, generate a border throughout the face by investigating the surrounding of the image.
4. Feature point will be detected using the equation (12):

$$I(X) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(x, y) \quad (12)$$

5. The equation (13) will find edges of the image:

$$V1 + F1 - E1 = 2 \quad (13)$$

Where, V1= vertices, F1= no. of faces, and E1= no. of edges

6. Images were crop after recognition of a face, and only face value can be useful at coordinating step. Evaluation time and complexity can reduce due to a significant part of the image has been crop and applied cropping functionality.
7. This process builds a “powerful” classifier as a linear conjunction of weighted easy “weak” classifiers.

$$H_j(x) = \begin{cases} -s_j, & \text{if } f_j < \theta_j \\ s_j, & \text{otherwise} \end{cases} \quad (14)$$

As express in equation (14) the threshold value θ_j and the difference s_j , $\epsilon \pm 1$ are resolved in training, also the coefficients. Viola Jones algorithm gives the conclusion with rapid detection and high accuracy. A few disadvantages of this method are the top estimating period through the image dimension is enormous, and resolution is high [1].

2.5 TWO-DIMENSIONAL GAUSSIAN FUNCTION:

Two-dimensional Gaussian function has proposed by Zhu Youlian, Huang Cheng, Zhai Kun, and PanLingjiao [10] in 2015. First, the skin color prototype became constructed in the RGB color space, and skin color areas were preparatory subdivided in proportion to the skin color characteristics. Lastly, facial areas were coordinate with facial patterns to understand face detection. In RGB color space [10], skin color dissimilarities have lessened, and tightness of the skin color pixel is superior. The correlation among the RGB and the rgb color pattern has express here. If [R1, G1, B1] and [R2, G2, B2] is two pixels in the RGB color space, they have a proportionate interrelation express in equation (15) [10]:

$$\frac{R_1}{R_2} = \frac{G_1}{G_2} = \frac{B_1}{B_2} \quad (15)$$

After that, we treated that the two pixels have the same color, but dissimilar lightness. The skin color dissimilarity mostly focuses on the intensity, and the color dissimilarity may reduce by formalizing the light. The normalization equation is express in equation (16) [10]

$$\begin{cases} r = \frac{R}{R+G+B} \\ g = \frac{G}{R+G+B} \\ b = \frac{B}{R+G+B} \end{cases} \quad (16)$$

After the normalization, respective lightness elopements have windrowed from [R, G, B], and the lightness effect of skin tone has weakened. Then he has examined the two pixels have the same color, still dissimilar lightness. The skin tone differences typically analyzed on the brightness, and the color dissimilarities can be minimized across standardize the lightness. After that, the standardization, respective lightness elements have taken away from [R, G, B], and the lightness effect of skin tone has minimized. To sub-divide the skin tone areas major correctly, he chooses the natural and instinctive Gaussian model to calculate the feasibility of every pixel which is homogenous to the skin tone. Overall feasibility efficiency is continual knowledge documentation that represents the skin feasibility chart, and the feasibility value is useful to calculate skin tone areas. The 2D Gaussian function is helpful in this work to express in equation (17) [10]:

$$P(r, g) = exp[(-0.5) * (x - m) * n^{-1} * (x - m)^T] \quad (17)$$

Where, x is a 2D typical pixel matrix [r, g] in the rgb space, *i.e.*, $x = [r, g]$; m is a 2D skin tone typical matrix $[\bar{r}, \bar{g}]$ in the rgb space, *i.e.*, $m = [\bar{r}, \bar{g}]$; n is a covariance matrix of the skin tonelikeness pattern, *i.e.*, $n = E [(x - m) * (x - m)^T]$; we use following two equations (18) and (19) to calculate m and n [10].

$$\bar{r} = \frac{1}{n} \sum_{i=1}^n r_i, \bar{g} = \frac{1}{n} \sum_{i=1}^n g_i \quad (18)$$

$$n = \begin{bmatrix} \sigma_{y,y} & \sigma_{y,g} \\ \sigma_{g,y} & \sigma_{b,b} \end{bmatrix} \quad (19)$$

The merit of m and n are calculated in equation (20) and (21) [10].

$$m = [0.4144, 0.3174] \quad (20)$$

$$n = \begin{bmatrix} 0.0031 & -0.0004 \\ -0.0004 & 0.0003 \end{bmatrix} \quad (21)$$

Thereafter Gaussian skin illustration, skin color areas, and non-skin color areas were introductorily partitions. Similarly, a significant number of arbitrary difference in the image may produce in the designing procedure. The smoothing filter can be applied to compromise the random change in the image. The volume of the pattern is $n \times n$, and M (i, j) is the actual gray merit of the midpoint (i, j), after that the flatten midpoint pixel gray merit may be shown in equation (21) [10]

$$g(x, y) = \frac{1}{n * n} \sum_{i=j-k}^{i+k} \sum_{j=j-k}^{j+k} g(x, y) * M(i, j) \quad (21)$$

Where, n is often an odd, and k is equals to $(n - 1) / 2$. Transcendental information is applied to detach skin-color areas successfully and make better detection reliability. Including facial pattern similarity, the capacity and the diversion point of view of the facial pattern are a balance as specified by the available facial areas capacity and diversion point of view [10].

2.6 QUALHOG FEATURES:

QualHOG features method has proposed by SuriyaGunasekar, Joydeep Ghosh, and Alan C. Bovik [11] in 2014. For sturdy face detection, a modern set of QualHOG characteristics has suggested that expands face-typical HOG characteristics with intuitive standard well informed spatial Natural Scene Statistics (NSS) [11] characteristics. Face analyzers prepared on those common characteristics which supply a statistically notable development in understanding to image deformation higher than the critical guideline. Deformation-reliant and deformation-ignorant variations on the face detectors were estimated and calculated on a vast collection of facial images presenting a broad scope of distortion. A partial change of training method strategy has additionally determined that additionally increase the hardness of the indicated face detectors. To make possible the present analysis, the author designed a latest Deformed Face Database, holds face and non-face areas from images reduced along with a collection of general deformed categories and measures. Firstly constructed that the simply estimable NR image grade count, NIQE is thriving as a representative for real bias measures when calculating the replacement among face detection achievement across image disablement appear from three general deformation categories, AWGN, gaussian blur, and JPEG. The author examines three fundamental varieties of manipulations that frequently happen in numerical appliances and above transmission mediums. A matrix denotes the image I, thus I (i, j) shows the (i, j) the pixel in the image I. AWGN (σ_N^2), Additive White Gaussian Noise That is a regional deformation, in which a zero mean Gaussian noise of variation parameter σ_N^2 is include separately to every pixel that is calculated using equation (22) [11].

$$\tilde{I}(i, j) = I(i, j) + N_{ij}, \text{ such that } N_{ij} \sim N(0, \sigma_N^2) \quad (22)$$

Where $N(\mu, \sigma^2)$ is a Gaussian distribution with mean μ and variance σ^2 . That is a standard pattern for a wideband mechanism or channel noise. Gblur (σ_B), Gaussian Blur is a worldwide deformation in that every point is unclear between convolutions with a Gaussian low pass filter of standard deviation σ_B . For commutating facility, the Gaussian kernel has abbreviated at $6\sigma_B$. The discrete abbreviated Gaussian filter in two extents has expressed in equation (23) [11]:

$$G(x, y) = \frac{1}{2\pi\sigma_B^2} e^{-\frac{x^2+y^2}{2\sigma_B^2}} \quad (23)$$

Where $-[3\sigma_B] \leq x \leq [3\sigma_B]$ and $-[3\sigma_B] \leq y \leq [3\sigma_B]$. An image with Gaussian blur deformation has shown in following equation (24) and (25) [11].

$$\tilde{I} = I * G \quad (24)$$

$$\tilde{I}(i, j) = \sum_{x=-[3\sigma_B]}^{[3\sigma_B]} \sum_{y=-[3\sigma_B]}^{[3\sigma_B]} I(i+x, j+y) G(x, y) \quad (25)$$

The present pattern is standard for lens blur. JPEG (Q), JPEG compression is the best regularly utilized lossy compression technique for a numeric snapshot. The tradeoff among image capacity and image correctness is admired by characteristics $0 \leq Q \leq 100$, where $Q = 100$ equivalent to no condensation while lessening values of Q inspire higher compression and lower image quality. The Q factor controls the degree of quantization. If G is the DCT matrix of the image I, the quantized DCT matrix, \tilde{G} is express in equation (26) [11]:

$$\tilde{G}(i, j) = \text{round} \left(\frac{G(i, j)}{Q(i, j)} \right) \quad (26)$$

Where the quantization matrix Q is of the same size as G, is designed to provide higher resolution in frequencies that are hypothesized to be perceptually more critical. An unassuming development in recognizing image standard carefully which computes face

detection presentation. This area was successfully select once assigning assets in exhausting positions. After that, the author presents that QualHOG characteristics that collaborate face expressive HOG characteristics class–common particular NSS characteristics are better functional at studying a face analyzer which is hardy to usual and principal image deformation. The spatial–NSS characteristics utilized in QualHOG to achieve unsighted IQA and made up of variables representing the natural position stochastic of essential elements. The image area, I , is preprocessed using mean local elimination and schematic normalization using equation (27) [11]:

$$\hat{I}(i, j) = \frac{I(i, j) - \mu(i, j)}{\sigma(i, j) + c} \quad (27)$$

Where (i, j) are a unique indicator, $\mu(I, j)$ and $\sigma(i, j)$ are the mean and variance, accordingly, of surrounding points weighted by an abbreviated proportional 2–D Gaussian, and C is the saturation continual (generally $C = 1$) that balances the splitting. QualHOG depends face analyzer indicates significant development above their HOG depends on analogs when working out on misshaped images. The partially modified QualHOG face analyzer additionally upgrades the toughness of suggested face analyzer. Therefore, the QualHOG Especially formed face detectors are suited to acquire allowable face detection achievement at considerably superior measures of visual disablement than what is currently achievable [11].

2.7 NEURAL NETWORKS WITH SKIN SEGMENTATION:

Neural Networks with Skin Segmentation has proposed by Samantha Denise F. Hilado, Elmer P. Dadios, and Reggie C. Gustilo [12] in 2011. Firstly, skin colors are subdivided all over the image. After that, search window pass across the image up until it identifies an acceptable face. Lastly, merging similar elements reanalyzed the analysis. This procedure is done again up to the complete image has been examined. This technique has enough to analyze other than single face inside the image. However skin partitioning may previously analyze faces inside certain images, further images can have another skin-like framework which unable to remove by differentiating it into a face mask. Depend on ten skin models from arbitrary images, the minimal and maximal Y , Cb and Cr merits have recognized. The skin has discriminated by correlating the pixel merits of the image is prepared to find the range of values determined in this system as skin. It goes as follows [12]:

```

If (mi Y > I pixel > ma H) AND (miCb > I pixel > maCb) AND (mi Cr > I pixel > ma Cr,)
Name the image as 1 (skin)
ELSE
Name the image as 0 (non_skin)
    
```

Although skin separation can previously analyze faces in individual images, further images may hold new skin-like surrounding that cannot eliminate by differentiating it to a face veil. Moreover, concurrent faces unable to recognize. The feed forward back propagation architecture has used in this work. A variable knowledge acquisition ratio is a composition of adaptable knowledge acquisition ratio, and training will apply to train images. At the start, the images have straightforwardly given to the neural network. Further edge detection method usually has operated before detecting regions within the image having immediate modification inside illumination, color or emotion. The dissimilarity over the images was rearranged preliminary for implementing edge detection to generate other determines characteristics. The canny technique is used to identify the edge. Inside the canny method, a threshold has defined for each powerful and a delicate edge. As long as the fine edge is uniquely incorporated when attached to a

commanding edge in the canny technique, already stated statement withdraws recognition of noise expected from unattached delicate edges. Information from edge recognition is useful as taken into the neural network [12]. The neural network has utilized for training an 18x27 image of face edge. Instead of using a raw grayscale image with values differentiating from 0 to 255 as input to the artificial neural network, a binary data of faces has used. Neural networks and skin partitioning [12] is merged to increase recognition precision and outcome indicates that the suggested technique is moreover competing than utilizing neural networks uniquely.

2.8 DEEP CONVOLUTIONAL NEURAL NETWORK:

This method has proposed by Shuo Yang, Ping Luo, Chen Change Loy, and Xiaoou Tang [13] in 2018. A complete image x is given as input to a CNN to produce the part level response map for every face segment. A collection of CNN's, called as attribute-aware networks, are utilize to create the part level response map of various components. The part level response map is acquired by weighted mapping altogether the acknowledgement maps at its upper convolutional layer. The map shows the position of a particular facial element given in the image for example hair, eyes, nose, mouth, and beard represented by h_a , h_e , h_n , h_m , and h_b , respectively [13]. For the demonstration, the author did summation entire these maps into a face class map ' h_f ' that recommends faces' positions considering a collection of candidate windows $\{w\}$ which are produced by current object proposition techniques like a region proposal network (RPN). They classify those windows in accordance with their faces scores Δw that has obtained from the part level response maps for dissimilar facial sections arrangements. Contender window 'A' encloses a particular area of h_a , that is, hair and its faceness count have computed by categorizing the merits at its topper section for the merits at its bottom section for a reason that hair is better probable to here at the upper layer of a face area. The lower section demonstrates the particular arrangements of five facial sections. The facial arrangements may accomplish from the training information. To decrease the amount of the suggested windows, we apply non-maximal suppression (NMS) to well order the scores by fractioning the specific correlations between those windows [13]. We conduct additional loop of NMS to decrease the number of suggested windows further using faceness result The author provides a case under on utilizing a part level response map of hair, h_a , for pattern approach every merit of position (i, j) on the part level response map show the possibility of the presence of the hair element. They choose a collection of M positions $\{(h_i, h_j)\}_{(i=1)^M}$ with a possibility $p(h_i, h_j)$ greater than t . For every chosen position, various pattern schemes have produced, where the amount of maximal feasible schemes for every position fixes as K . The schemes have acquired from decided referral boxes that we name as patterns. For every face region, patterns have centralized at various positions in consideration of the formation of the human face. Also, they connected with a particular range and condition proportion. For example, the patterns of the hair area centralized at $(W/2, H/3)$ and the patterns of eyes have centralized at $(W/2, H/2)$, where W and H present the width and height of presenter [13]. These patterns are transferrable invariable to the network's entire step, and the technique doesn't sustain additional valuation for forwarding scales. A faceless outcome on those windows coming again to the ordered group of top-scoring face schemes. The strategy obtaining the faceness merit from the part level response maps of hair and eyes, let Δw is the faceness outcome of a window w . For instance, stated a part level response map of hair, h_a , Δw has addressed by classifying the summation of merits in first green by the summation of merits in the second green area. Likewise indicates that Δw has acquired by separating the summation of values in a third green area for first + fourth of him. For a couple of the above instances, a huge value of Δw shows an extreme interweaving proportion of w with a face. It is feasible to calculate the faceness outcome using the new formation of handmade characteristics that can successfully catch the face construction across response

maps. Among the produced face approach, the author trained a powerful face detector that exhibited excellent performance on different face detection standard databases.

2.9. HAAR-LIKE FEATURES:

Shan Du has proposed Haar-like features, motion, and color filtering and hard hat color detection, Mohamed Shehata and WaelBadawy [14] in 2011. This proposed system has shown in the following Figure.

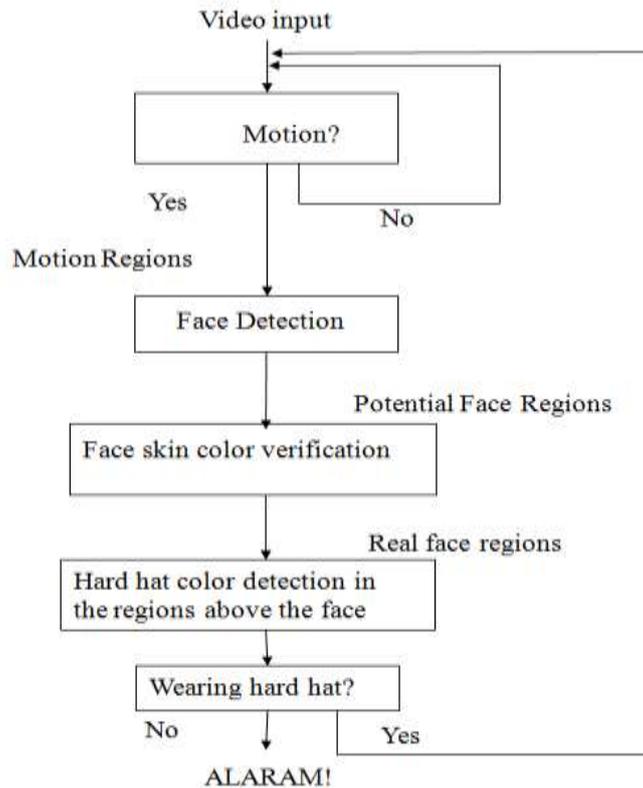


Fig. 4 Process of Face Detection [14]

The Fig. 4 shows the process of face detection. The initial section is the face detection depends on Haar-like features, author has detected the entire feasible face areas. Previous to face detection, they first analyze fluctuations in video series. If there is no movement or small movement in the video, he does not do anything more. If there is movement, they acquire the movement areas and deal with face detection uniquely on those areas to keep away examining the entire image. That may store the calculating period and eliminate false alarms on the surrounding. They use the face skin color technique to strain the non-face blocks. Another step is hard hat color detection in this step, we analyze that he tends to observe all attainable face regions. They use fourteen Haar-like characteristics pattern express in Figure [14].

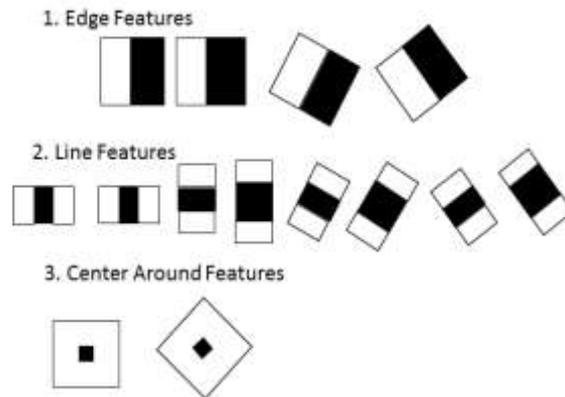


Fig. 5 Haar-like Characteristic Pattern [14]

The Fig. 5 shows haar-like characteristics pattern. It involves four boundary characteristics, eight edge characteristics, and two midpoint-adjacent characteristics. The total of the pixels that lie within the white rectangles has deducted from the submission of pixels in the dark rectangles [14]. These patterns are mounted separately in the vertical and horizontal regulation to create a rich, above complete collection of characteristics. To eliminate the non-face regions, they used to face skin color. Then another is hard hat color analysis method [14]. The hard hat detection has formed on the color knowledge in the areas right over every first face area. To investigate those areas, initially, they subdivide those using particular morphological functions. That is due to those areas composed of hair or hard hat as well as may consist of forehead and surrounding. Then we require ignoring the forehead part to study if there is a hard hat protecting the hair. They examine the areas of right above faces on the image to analyses if there is a hard hat or not [14].

2.10 FASTER R-CNN:

Faster R-CNN method has proposed by Huaizu Jiang, Erik Learned-Miller [15] in 2017. The original benefaction of the researcher is to calculate the modern Faster R-CNN [15] on a massive database of faces. The author illustrates new face detection outcomes by applying the Faster R-CNN [15] on tree-like face analysis standard. The author further differentiates unique creations of segment-based CNN object detection [15] representation and separate into a variety of unlike modern high-operating analyzers. To decrease the computing time of program generation, the Faster R-CNN has suggested. That involves two sections firstly, known as the Regional Proposal Network (RPN) [15], and is a fully convolutional network [15] for producing desire programs.

In the RPN, the convolution layers of an advanced trained network have followed by a 3×3 convolutional layer. That proportionate to measuring a particular huge window or active area in the input image to a small-geographical characteristics vector at a midpoint stalk, for example, 16 for VGG16. Two 1×1 convolutional layers have included for categorization and progression sections of all particular windows. To achieve various scopes and particular proportion of entities, anchors has inaugurated in the RPN. The author uses three ranges of 1282, 2562, and 5122 midpoints and three unique propositions as 1: 1, 1: 2, and 2: 1, dominant to $k = 9$ anchors at every position. Every approach has to describe in terms of parameter comparative to an anchor. They mark that the identical characteristics of every slither posture are utilized to decline $k = 9$ proposals, in place of extracting k positions of components and training a unique image. Training of the RPN may be finished in endwise fashion using stochastic gradient descent (SGD) for each categorization and reversion sections. For the complete structure, we have to keep both the RPN and Fast R-CNN components as long as they divide convolutional layers. The

RPN and Fast R-CNN have trained endwise as they are free. Consider that the input of the Fast R-CNN is reliant on the output of the RPN [15].

2.11. NORMALIZES PIXEL DIFFERENCE (NPD), HAAR CLASSIFIER:

Normalizes Pixel Difference (NPD), haar classifier has proposed by Neethu A, Athi Narayanan S, Kamal Bijlani [2] in 2016. A composite face detection technique for a single calculation by margining three well-organized face detection approaches Normalized Pixel Difference (NPD) [2], haar classifier and haar classifier for profile face [2]. NPD is a scale-invariable technique, and its presentation is immensely better in fluctuating surrounding and modest lightning. Composition of those three face analyzers is utilized to upgrade the correctness and to decrease false detections. Firstly, the NPD face detector [2] is used to analyze faces in the image. The face detection rate of NPD [2] is reasonable in comparison to another face analysis technique. To improve the correctness of face analysis and to decrease the false detection rate, haar-classifier [2] has utilized as a next detector. The third classifier is haar-classifier for profile face [2], and it executes right in profile and moving face images. Composition of those face analyzer may increase the achievement of face analysis generally in uncontrolled conditions, hence initially, second and third detectors are merged, and individual face bounding box gets back. Post processing is completed to prevent false detections and various detections. A unique face in an image is detected several times by the noticed face detectors, and those multiple detections have integrated into a dependable one. False detections are these sub-windows that do not face incorrectly recognized as faces. For avoiding false detections, a threshold value computed by has to verify some images and removed these adjoining boxes that have values below the threshold. The figure expresses the architecture of the suggested composite face detector.

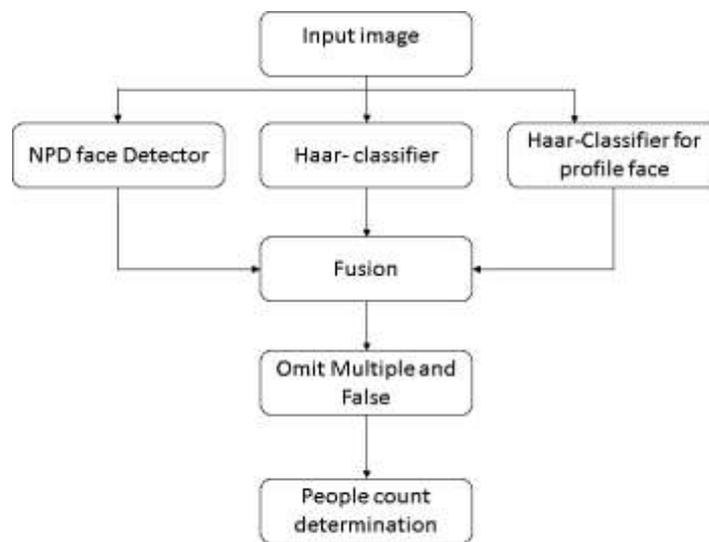


Fig. 6 Block Diagram of Hybrid Face detector [2]

The Fig. 6 shows the block diagram of hybrid face detector which describes the process of face detection including several steps. To prevent those false detections various images are examined, and a lower and upper threshold has computed, and merits not among those thresholds have ignored. The upper and lower limitation has calculated in equation (28) and (29).

$$t_{upper} = \text{median}(\text{areaofboundingbox}) * 1.6 \quad (28)$$

$$t_{lower} = \text{median}(\text{areaofboundingbox}) * 0.4 \quad (29)$$

2.12. 2-D DCT

2-D DCT method has been proposed by M HanumaTeja [16] in 2011. In the suggested technique, face pattern matching has utilized for rapid detection of face-like substances. That is theoretically easy and fast technique for face detection. The fundamental aim here is to observe areas of the image that come out to be equivalent to the mention face pattern. The mentioned pattern has produced by hand cropping and choosing the model that may be the leading estimated familiar face. Preprocessing steps includes following steps [16]:

1. The Extermination of the image by an aspect to decrease estimation price
2. Transformation of the input RGB images to the YCbCr color space by employing the subsequent conversion using equation (30), (31) and (32) [23].

$$Y'=16+ (65.481.R^'+128.553.G^'+24.966.B^') \quad (30)$$

$$C_B=128+ (-37.797.R^'-74.203.G^'+112.0.B^') \quad (31)$$

$$C_R=128+ (112.0.R^'-93.786.G^'-18.214.B^') \quad (32)$$

The inverse transformation is shown in equation (33), (34) and (35) [16]

$$R'_D= (298.082.Y^')/256+ (408.583.C_R)/256-222.921 \quad (33)$$

$$G'_D=(298.082.Y^')/256-(100.291.C_B)/256-(208.120.C_R)/256+135.576 \quad (34)$$

$$B'_D= (298.082.Y^')/256+ (516.412.C_B)/256-276.836 \quad (35)$$

Matching accomplished by formalized 2-D cross-correlation of the input image to the mentioned pattern. That has further followed by using two thresholds on the merits of cross-correlation γ –

1. If $\max(\gamma) > 0.55$, then areas throughout maxima proportionate to a face
2. If $0.37 < \max(\gamma) < 0.55$ then areas throughout maxima proportionate to feasible face entity
3. If $\max(\gamma) < 0.37$ no face is there in the image structure

The recognition of skin has accomplished by employing thresholds on the merits of the pixel that has enclosed to the YCbCr color space before refining. For speedy categorization, only the chrominance Cr element has examined. The area designated as a feasible face object by the pattern matching step now recognized for the proportion of pixels equivalent to the skin. If these merits are larger than 85%, then the areas equivalent to a face entity has rejected. The suggested technique includes the investigation of the image's DCT energy [16] helps to detect wink of eye and pupil fluctuation. To reshape the image to its frequency domain presentation, we did the 2- D Discrete Cosine Transform as express in the equation (36) [16]

$$F(u, v) = \left(\frac{2}{N}\right)^{1/2} \left(\frac{2}{M}\right)^{1/2} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \Lambda(i) \cdot \Lambda(j) \cdot \cos \left[\frac{\pi \cdot \mu}{2 \cdot N} (2i + 1) \right] \cdot \cos \left[\frac{\pi \cdot \mu}{2 \cdot M} (2j + 1) \right] \cdot f(i, j) \dots (36) \text{ Where } \Lambda(\xi) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } \xi = 0 \\ 1 & \text{Otherwise} \end{cases} \quad (37)$$

The suggested technique, face pattern matching has utilized for quick analysis of face-like things. This technique has established on the reality which considering the area of the

photo is lesser than the live face. High-frequency elements of photo images should be minimum than actual face images. The previously analyzed method has certain disadvantages that have resolved by several authors. Still, there are specific problems in face detection as face image gallery does not consider different views of the face image and they had only worked on neutral images, Accuracy of the proposed system is weak, the efficiency of detecting the faces from large databases. Inputted image having a composite background and different lighting situations may be additionally somewhat complicated in pinpointing or tracking. Face emotions analysis turns to decline if the test image has a various illumination background as well as the training images. If lightning is not constant, the facial point may be analyzed incorrectly. To control the issue of an unconstrained environmental problem, the face detection which detects faces from the image/frame using R-CNN (Regional Convolutional Neural Network) will apply. Which reduces the efficiency of detecting the faces from an extensive database because R-CNN improves face detection effectiveness by distributing estimation along a feature pyramid permits for detection at a small number of frames per second [16].

Table I. Comparative Analysis of Face Detection Techniques

Sr. No.	Name of Author and year of publication	Methodology used	Accuracy
1	M HanumaTeja, 2011	2-D DCT	85%
2	Shan Du, Mohamed Shehata and WaelBadawy, 2011	Haar-like features	86%
3	Samantha Denise F. Hilado, Elmer P. Dadios, and Reggie C. Gustilo,2011	Neural Networks with Skin Segmentation	87%
4	Nai-Jian Wang, Sheng-Chieh Chang,2012	Real-time multi-face detection system	94.9%
5	SuriyaGunasekar, Joydeep Ghosh, and Alan C. Bovik, 2014	QualHOG features	90%
6	Bashir Muhammad, Syed Abd Rahman Abu-Baka,2015	skin detection method of combining two colorspace HSV and YCgCr	70%
7	DivyaMeena, Ravi Sharan,2016	Viola Jones algorithm	90%
8	Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik, 2016	Convolutional Neural Network	81%
9	Shengcai Liao, Anil K. Jain, and Stan Z. Li,2016	Normalized Pixel Difference (NPD)	80%
10	Neethu A, Athi Narayanan S, Kamal Bijlani, 2016	Normalizes Pixel Difference (NPD), haar classifier	90.20%
11	Huaizu Jiang, Erik Learned-Miller,2017	Faster R-CNN	90.3%
12	Shuo Yang, Ping Luo, Chen Change Loy, and Xiaoou Tang, 2018	Deep convolutional Neural Network	92.84%

The Table I Gives Comparative analysis about the fair work of various Face detection techniques accuracy along with the use of different databases, here we examine that the few techniques have reached more exceptional recognition accuracy, but there are few obstacles of face detection.

3. CONCLUSION

Based on the above discussion the paper summarized the information evaluation of research equivalent to face recognition technique. Face detection is an inspirational and prime section in facial expression and recognition system, between complete the biometric methods. The evaluation time, correctness and the presentation of the face detection methods under unconstrained lightning with different postures is a significant

concern in the facial expression recognition system. Here, we consider unique designs, viewpoints, techniques, a dataset for training or testing images and presentation calculation of face detection technique utilized in every research. Each author carries owned method for detecting a face from a database or from video several authors has tried to resolve the issues related to the previously suggested techniques still, and there are few profits and constraints in those analyzed methods. There is a necessity to implement a useful technology for face recognition that will decrease estimation time and improve effectiveness.

4. FUTURE SCOPE

Our future work will establish on the real time information of images using R-CNN (Regional Convolutional Neural Network) [20] with many hidden layers to analyze the outcome in the best presentation values of the face detection method.

REFERENCES

- [1] DivyaMeena, Ravi Sharan, "An Approach to Face Detection and Recognition", IEEE International Conference on Recent Advances and Innovations in Engineering. Jaipur, India, December 23-25, 2016, pp. 1-6.
- [2] Neethu A ; Athi Narayanan S ; Kamal Bijlani, "People Count Estimation Using Hybrid Face Detection Method", International Conference on Information Science, Kochi, India, 12-13 Aug. 2016, pp. 144-148.
- [3] S Wiegand, C Igel, U Handmann, "Evolutionary Optimization of Neural Networks for Face Detection", European Symposium on Artificial Neural Networks Bruges, (2004), pp. 139-144.
- [4] H., Boughrara, M Chtourou., C.B., Amar, "MLP neural network based face recognition system using constructive training algorithm" International conference on Multimedia computing and system, Tangier, (2012), pp. 233 – 238.
- [5] H Sahoolizadeh, D Sarikhanimoghadam, and H. Dehghani, "Face recognition system using neural network with Gabor and discrete wavelet transform parameterization", International Conference of Soft Computing and Pattern Recognition, Tunis, (2014), pp. 17 – 24.
- [6] Y Lu , N Zeng , Y Liu , NZhang, "A hybrid Wavelet Neural Network and Switching Particle Swarm Optimization algorithm for face direction recognition", International Journal on Neurocomputing, Volume 155, (2015), pp. 219–224.
- [7] Shengcai Liao, Anil K. Jain, and Stan Z. Li, "A Fast and Accurate Unconstrained Face Detector", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 38, Issue 2, February, 1, 2016, pp. 211 – 223.
- [8] Bashir Muhammad, Syed Abd Rahman Abu-Baka, "A Hybrid Skin Color Detection using HSV and YCgCr Color Space for Face Detection", International Conference on Signal and Image Processing Applications, Kuala Lumpur, Malaysia, 19-21 Oct. 2015, pp. 95-98.
- [9] Nai-Jian Wang, Sheng-Chieh Chang, Pei-Jung Chou, "A Real-time Multi-face Detection System Implemented on FPGA", International Symposium on Intelligent Signal Processing and Communication Systems, November 4-7, 2012, pp. 333-337.
- [10] Zhu Youlian, Huang Cheng, Zhai Kun, Pan Lingjiao, "Face Detection Method Using Template Feature and Skin Color Feature in rgb Color Space", Chinese Control and Decision Conference, Qingdao, China, 23-25 May 2015, pp. 6133- 6137.
- [11] SuriyaGunasekar, Joydeep Ghosh, and Alan C. Bovik, "Face Detection on Distorted Images Augmented by Perceptual Quality-Aware Features", IEEE Transactions on Information Forensics and Security, Volume 9, Issue 12, December 2014, pp. 2119 – 2131.
- [12] Samantha Denise F. Hilado, Elmer P. Dadios, and Reggie C. Gustilo, "Face Detection using Neural Networks with Skin Segmentation", International Conference on Cybernetics and Intelligent Systems, Qingdao, China, 17-19 Sept. 2011, pp. 261-265.
- [13] Shuo Yang, Ping Luo, Chen Change Loy, and Xiaoou Tang, "Faceness-Net: Face Detection through Deep Facial Part Responses", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 40, Issue 8, August, 1, 2018, pp. 1845 – 1859.
- [14] Shan Du, Mohamed Shehata and WaelBadawy, "Hard Hat Detection in Video Sequences based on Face Features, Motion and Color Information", International Conference on Computer Research and Development, Shanghai, China, 11-13 March 2011, pp. 25-29.
- [15] Huaizu Jiang, Erik Learned-Miller, "Face Detection with the Faster R-CNN", International Conference on Automatic Face & Gesture Recognition, Washington, DC, USA, 30 May-3 June 2017, pp. 650-657

- [16] M HanumaTeja, "Real-time Live Face Detection using Face Template Matching and DCT Energy Analysis", International Conference of Soft Computing and Pattern Recognition, Dalian, China, 14-16 Oct. 2011, pp. 342-346.
- [17] Xiangxin Zhu Deva Ramanan, "Face Detection, Pose Estimation, and Landmark Localization in the Wild", IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16-21 June 2012, pp. 2879- 2886.
- [18] JatinChatrath, Pankaj Gupta, Puneet Ahuja, Aryan Goel, ShaifaliM.Arora, "Real Time Human Face Detection and Tracking", International Conference on Signal Processing and Integrated Networks, Noida, India, 20-21 Feb. 2014, pp. 705-710.
- [19] Cheng Jiang, Yifeng Zhang, "The Face Detection Algorithm Based on Local Elastic Potential Energy Feature", International Conference on Wireless Communications and Signal Processing, Nanjing, China, 11-13 October 2017.
- [20] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik, "Region-based Convolutional Networks for Accurate Object Detection and Segmentation", IEEE Transaction on Pattern Analysis and Machine Intelligence, Volume 38, Issue 1, January, 1,2016, pp. 142-158.