Person Identification Using Latent Face Images

Sayali S. Avhad[1], Dr. Ameya K. Naik[2], Dr. Kiran Ajetrao[3]

[1] Department of Electronics and telecommunication, K.J. Somaiya college Of Engineering, Mumbai, Maharashtra, India

[2] Department of Electronics and telecommunication, K.J. Somaiya college Of Engineering, Mumbai, Maharashtra, India

[3] Department of Electronics and telecommunication, K.J. Somaiya college Of Engineering, Mumbai, Maharashtra, India

Abstract

Face recognition is one of the most important areas in computer vision like identification, authentication, security, surveillance system, human-computer interaction, psychology and so on. Eyes, nose, and mouth are the key features for face recognition. Eyes and mouth bear some sort of facial expressions and eyes are the most crucial facial feature due to its inter-ocular distance, which is constant among people and unaffected by moustache or beard. Therefore, face recognition is distinctly influenced by these features. This paper presents to detect the face parts in an image. The detection of the facial parts like eyes, nose, mouth, and face are an important task in this process. This technique is used to recognize and detect the parts of the human facial factors in an image. The study involves the algorithm of Viola-Jones Cascade Object Detector which provides various combination of filters and methods to detect these facial expressions.

Keywords: Face detection; Viola-Jones algorithm; nose detection; mouth detection; eyes detection.

1.Introduction

The technique of identifying one or more human faces in image or videos is known as face detection. Many biometric, security, and surveillance systems, as well as images and video indexing systems, rely on it. Face detection is a computer technique that recognizes human faces in digital images and is used in a range of applications. It is the psychological mechanism by which humans locate each other. The detection of frontal human faces is the focus of face-detection algorithms. It's similar to image detection, where a person's image is matched bit by bit. Some facial algorithms identify by extracting facial features or assessing the relative position, size, and shape of eyes, cheekbones, and other facial features. Any facial feature changes in the database will invalidate the matching process.
The major goal is to create a system that can detect and recognize the textures of human body parts in images and videos. The numerous parameters of facial features are linked with the estimated parameters of human body components. Face detection with a computer is a difficult operation since it necessitates recognizing and identifying faces of various sizes, shapes, textures, and intensities of colors on it. This may be extended to real-world applications such as facial recognition in online tests, gender/age identification, and much more. Face detection using computers works by detecting and distinguishing between facial and non-facial features, then returning the facial elements that are present in the human body.
The high detection rate was achieved using a machine learning methodology for visual object detection. A face detection technique in which a single image can be used to recognize several human faces. A face detector must determine whether or not the image contains a face. There have been several approaches devised to detect a face from an image, one of which being the Viola Jones Algorithm which is discussed in paper. To detect a face, the system must properly analyze the entire image and remove the background image. First, an integral picture is proposed in this paper, which allows detectors to be computed relatively quickly. The second approach is based on the Adaboost method, which quickly selects a few of the most prominent traits from a vast range of features to produce an incredibly efficient classifier. The third way involves merging more complex classifiers in a cascade to easily eliminate the image's background.

2. Related Work

The problem related to facial expression recognition may vary between several factors such as illumination, pose invariant and rotation, etc. There are many researches based on the detection of face and people tracking and counting the number of peoples in either an image or video such as [1] and [2]. But detecting facial parts in several images is a challenging task as the accuracy won't be good for every images. The company named 'Omran' [7] is a sensing company has released a smile measurement software which detects the smile of a several persons at the same time with the percentage of 0-100. It uses 3D face mapping technology, and its detection rate will be greater than 90%. Also, the facial expressions given by the human beings have been recorded and analyzed using [4]. The algorithm such as PCA, LDA has been used for the recognition of face which gives a good result [5] and [6]. The FGNET face database have been used [3] and the results have a detection rate of 88.5% and false alarm rate of 12.04% whereas Sony T300 performs the detection rate of 72.7% and false alarm rate of 0.5%. The emotional recognition on the face has been achieved [8] with a best result of 94% on a Raspberry Pi II and the faces that are invariant to the orientation and posehave been detected [9] with the Gabor wavelet methods.
This paper presents a method to detect face, eyes, nose and mouth detection by viola jones algorithm. The rest of report is organized as follow: detail information about proposed work, algorithm structure, experimental results of GUI and conclusion.

3.Methodology

3.1 Viola Jones Algorithm:

The Viola Jones algorithm is named after Paul Viola and Michael Jones, two computer vision researchers who proposed the method in 2001. The method Vision was used to implement this algorithm in the software 'Matlab.' CascadeObjectDetector. Despite being an older framework, Viola-Jones is highly effective, and its application in real-time face identification has proven to be extremely important. The flowchart of Viola Jones algorithm is given below:

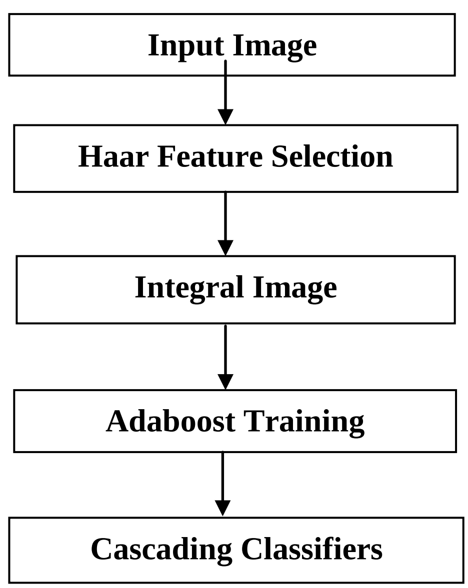


Fig 1. Flowchart of Viola jones algorithm

The Viola Jones algorithm has four main steps, which we shall discuss in the sections to follow:

1. Selecting Haar-like features
2. Creating an integral image
3. Running AdaBoost training
4. Creating classifier cascades

1. Selecting Haar-like features:

The idea of implementing Haar wavelets was adopted by Voila and Jones, who created the so-called Haar-like features [10]. Digital image features called Haar-like features are employed in object recognition. Some universal characteristics of the human face may be found in all human faces, such as the eye region being darker than its neighbour pixels and the nose region being brighter than the eye region. Summarizing and comparing the pixel values of both regions is a simple way to determine which area is lighter or darker. The sum of pixel values in the darker area will be less than the sum of pixel values in the lighter area. If one side is lighter than the other, it might be an eyebrow or the center area of the box being shinier than the surrounding boxes, which can be interpreted as a nose. This can be accomplished using Haar-like features and with the help of them, we can interpret the different parts of a face.

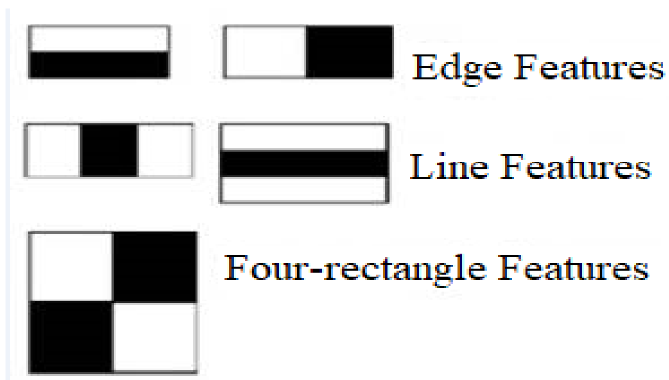


Fig 2. Haar-like features

There are 3 types of Haar-like features as shown above in fig 2.that Viola and Jones identified in their research:

1. Edge features
2. Line-features
3. Four-sided features

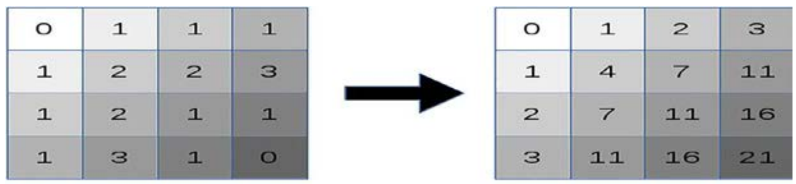
Edge and line features help to recognize edges and lines, respectively. Finding diagonal features is done with the four-sided features. The sum of pixel values in the black area minus the sum of pixel values in the white area is used to determine the feature's value. The value is zero for a plain surface in which all the pixels have the same value, and thus, provide no useful information. Since facial features are complicated shapes with darker and brighter areas, a Haar-like feature gives you a large number when the areas in the black and white rectangles vary significantly. We can extract a bit of useful information from the image by using this value. A Haar-like feature should give you a large number in order to be useful, suggesting that the areas in the black and white rectangles must be considerably distinct.

2. Creating an Integral image:

ISSN (Online) 2348 – 7968 | Impact Factor (2020) – 6.72

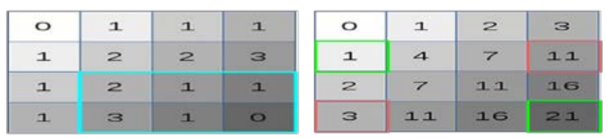
In the previous section, we have seen that to calculate a value for each feature, we need to perform computations on all the pixels inside that feature. These calculations can be very intensive since the number of pixels would be much greater when we are dealing with a large feature. The integral image contributes to our ability to quickly run these extensive calculations in order to determine whether a feature of multiple features fits the criteria. An integral image (also known as a summed-area table) is the name of a data structure as well as the technique used to build it. It is used to calculate the sum of pixel values in an image in a quick and efficient manner.

In an integral image, the value of each point is the sum of all pixels above and to the left, including the target pixel:



Original Image Integral Image

It performs calculations on four edges of the rectangle. See the example below to understand.



When we add the pixels in the blue box, we get 8 as the sum of all pixels and here we had six elements involved in your calculation. Now to calculate the sum of these same pixels using the integral image, you just need to find the corners of the rectangle and then add the vertices which are green and subtract the vertices in the red boxes. Now doing that here

$$21+1-11-3=8$$

We get the same answer and only four numbers are involved in calculations. No matter how many pixels are in the rectangle box, we will just need to compute on these 4 vertices. Now to calculate the value of any Haar-like feature, you have a simple way to calculate the difference between the sums of pixel values of two rectangles.

3.Running AdaBoost training:

Each Haar-like feature in the Viola-Jones method represents a weak learner. AdaBoost evaluates the performance of all classifiers you submit to it to determine the kind and size of a feature that will be included in the final classifier [10]. To calculate the performance of a classifier, you evaluate it on all subregions of all the images used for training. Positives indicate that the classifier believes the image has a human face. Subregions that do not generate a strong response, in the view of the classifier, do not contain a human face. Negatives will be given to them. The importance or weight of the classifiers that performed well is increased. As a result, a strong classifier, also known as a boosted classifier, is created, which includes the best performing weak classifiers. So, when we train the AdaBoost to identify key features, we feed it information in the form of training data and then train it to learn from that data to predict.

4.**Creating classifier cascades:**

The cascade's job is to quickly eliminate non-faces and save time and computations. As a result, the speed required for real-time face detection is achieved. We developed a cascaded method in which the process of detecting the face is divided into several steps. We have a classifier in the first stage that is made up of our best features; in other words, the subregion goes through the best characteristics in the first stage, such as the feature that identifies the nose bridge or the feature that identifies the eyes. In the following steps, we will complete all of the remaining tasks. The first step evaluates an image subregion when it enters the cascade. The stage's output is maybe if it evaluates the subregion as positive, indicating that it considers it is a face. When a subregion receives a maybe, it moves on to the next stage of the cascade, and so on until we reach the final stage.

3.2 Block Diagram:

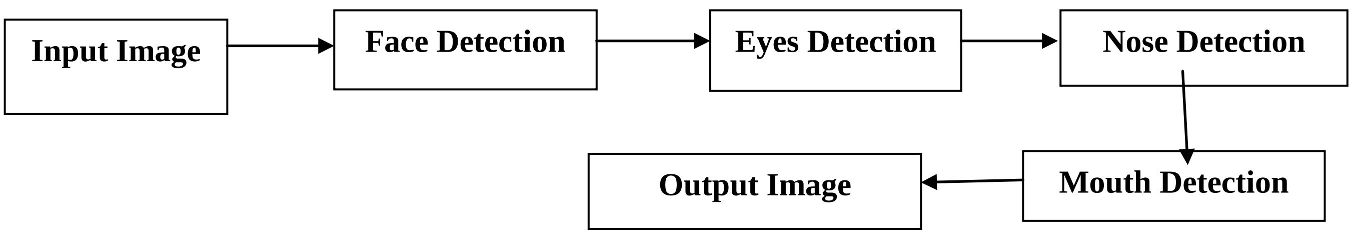
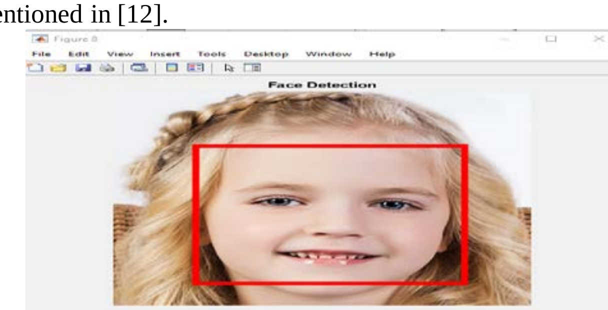


Fig 3. Block Diagram

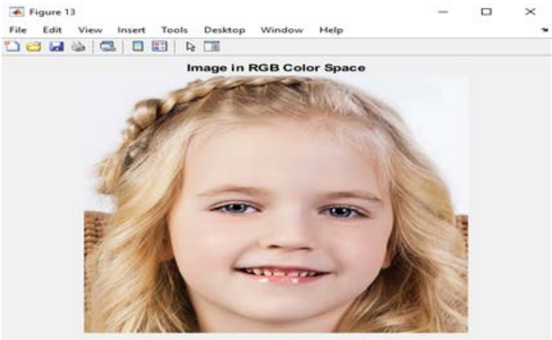
Face Detection:

The face detection is done by Viola Jones [11]. Face detection's major goal is to reduce the number of false positives in facial expression recognition. This part is important for correctly locating the eyes. Skin segmentation [12] is performed by converting the image to YCbCr domain once the face has been detected, as illustrated in Fig 4. b. The most significant benefit of transforming the image to the YCbCr domain is that it eliminates the influence of luminance by focusing exclusively on the chromatic components. Each component of the image, red, green, and blue, has a particular brightness in the RGB domain, as illustrated below in Fig 4. cHowever, in the YCbCr domain, as illustrated in Fig.4.d., the Y- component provides all brightness information, as the Cb (blue) and Cr (red) components are fully independent of luminosity. The domain conversions are used to segment the RGB image into Y, Cb, and Cr components, as illustrated in Fig. c. into Y, Cb, Cr components. Even though skin colors differ from person to person, and race to race, it was found that [11] the color remains distributed over a very tiny region in the chromiance plane. This method detects skin regions over the entire face image and rejects most of the non-face image. Fig.3.3 shows the detected face and the corresponding skin regions using the YCbCr segmentation mentioned in [12]. Despite the fact that skin colours vary from person to person and race to race, [11] it was discovered that the colour is dispersed over a very small region in the chromiance plane. This method detects skin regions over the whole face image and rejects the majority of non-facial images. Fig.4. shows the detected face and the corresponding skin regions using the YCbCr segmentation mentioned in [12].

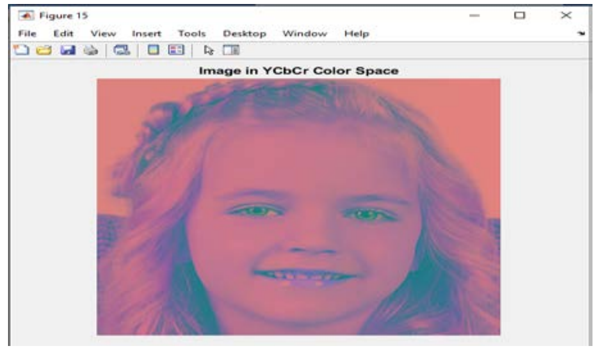
a. Original Image



b. Face Detection



c.Image in RGB Color Space

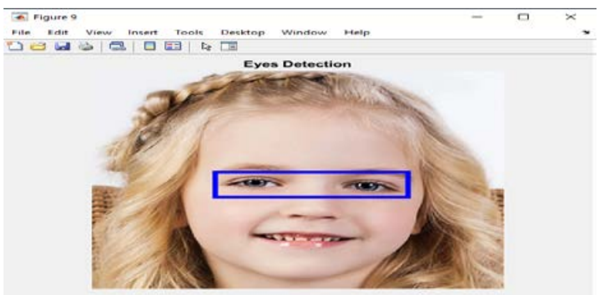


d. Image in YCbCr Color Space

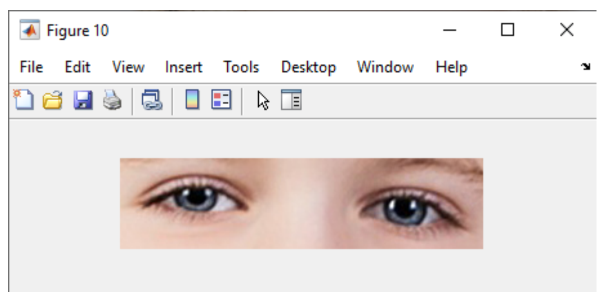
Fig 4. Face Detection

Eyes Detection:

The region of the eye is darker related to other parts of the face, so finding the regions of the eye detection as shown below in fig 5. is based on segmenting a small region of the image which is specified as a darker region [13]. Because the centre of the eye region is darker than the rest, the eyebrow region has been removed in this model. After that, the histogram analysis is used to determine the region of the selected eye region, as the eye region has two peaks whereas the eyebrow region has only one. The final constraint here is the alignment of the two major axes, which ensures that the two eye regions are aligned correspond to the same line.



a. Eyes detected.



b. Eyes cropped

Fig 5. Eyes Detection

Nose Detection:

The nose has different properties on it to detect easily as shown below in fig 6

a) Dark White Dark Pixels: The nostril region will be identified when an image is taken and convolved with these Dark White Dark Pixels [13]. The dark pixels are represented by the two regions of holes on the nose, while the white pixels are described by the middle region of the nose.

b) Similarity of region on both the sides: Both the left and right sides of the nose have the same region of black areas in the nostrils. On both sides of the region, these features have been classified as a similarity.

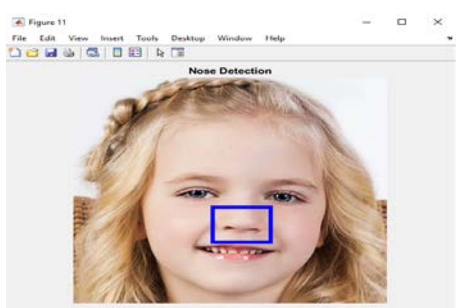


Fig 6. Nose Detection

IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 8 Issue 6, June 2021
ISSN (Online) 2348 – 7968 | Impact Factor (2020) – 6.72
www.ijiset.com

Mouth Detection:

The mouth detection can be classified as weak classifiers, as shown in fig 7., where the detection and extraction of features from the mouth region is based on a typical decision stump that uses Haar features to encode the details of the mouth [13].

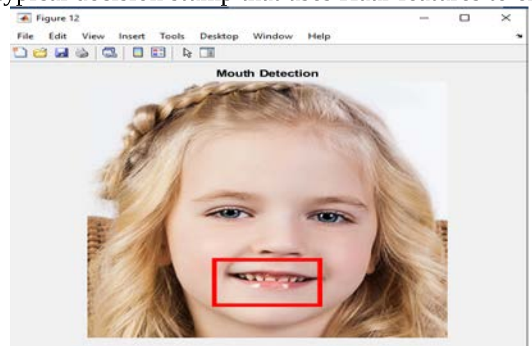
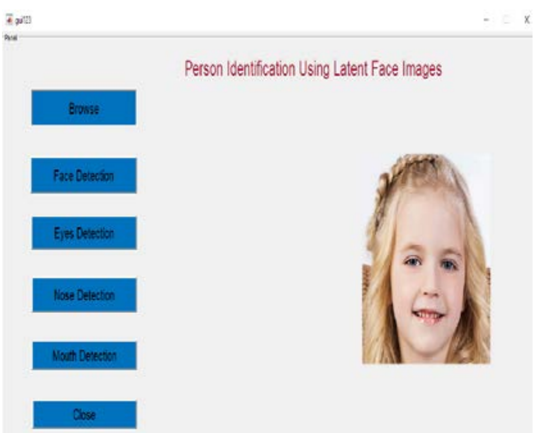


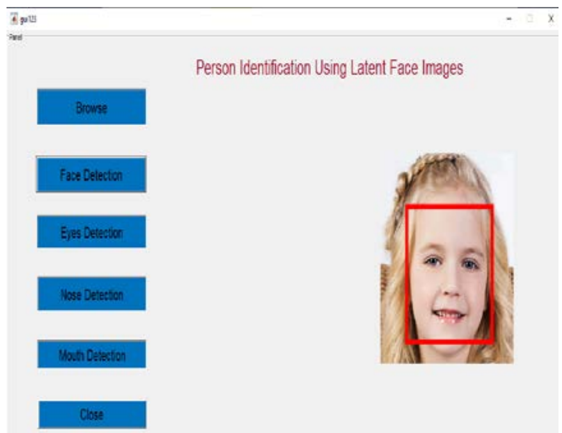
Fig 7. Mouth Detection

4.Result

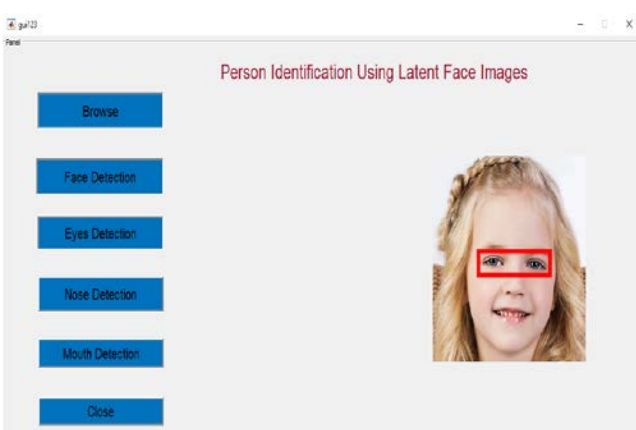
4.1 GUI of Face detection parts:



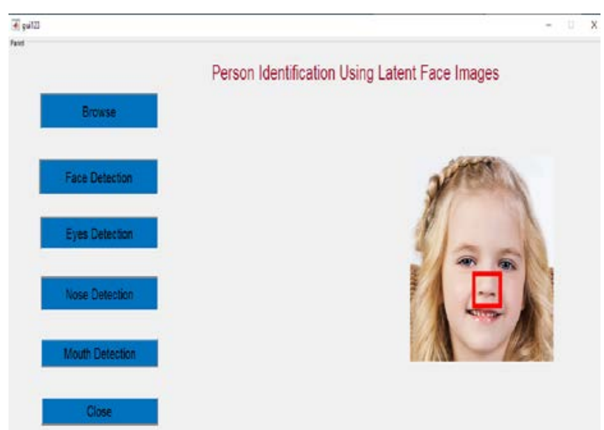
a. Browse image



b. Face detection



c..Eyes detection



d. Nose detection

e. Mouth Detection

4.2 Performance Evaluation:

Here performance evaluation is done considering parameter like accuracy. There are several terms that are commonly used along with the description accuracy. They are true positive (TP), true negative (TN), false negative (FN), and false positive (FP). If the sample is positive and it is classified as positive, i.e., correctly classified positive sample, it is counted as a true positive (TP); if it is classified as negative, it is considered as a false negative (FN). If the sample is negative and it is classified as negative it is considered as true negative (TN); if it is classified as positive, it is counted as false positive (FP). Table 4.2 as shown below shows the obtained performance evaluation parameter.

1. Accuracy:

Accuracy is one of the most commonly used measures for the classification performance, and it is defined as a ratio between the correctly classified samples to the total number ofsamples

$$Accuracy = (TP+TN)/(TP+FP+FN+TN)$$

Table 1. Performance evaluation

Input Images	**Eyes**	**Nose**	**Mouth**
	Accuracy		
Image1	0.7068	0.6163	0.6265
Image2	0.5265	0.4582	0.4965
Image3	0.6926	0.6045	0.6488
Image4	0.5109	0.4791	0.6340
Image5	0.5908	0.5166	0.5363
Image6	0.6092	0.5549	0.6761
Image7	0.4708	0.4378	0.5220
Image8	0.6309	0.5900	0.6347
Image9	0.5967	0.5099	0.6347
Image10	0.6214	0.6114	0.5754

4.Conclusion

This paper describes an efficient method for Face parts detection by Viola Jones Algorithm. These three phases are facial features detection using Viola Jones, the face detection, eyes detection, nose detection and mouth detection. Once the face is detected, the system is made illumination invariant by segmenting the skin part alone and considering only the chromatic components to reject most of the non-face image backgrounds based on skin color. The accuracy for eyes detection is 60%. On comparing accuracy for eyes detection is more than the nose and mouth detection. So, with the help of eyes detection person will be identified easily. Face parts detection can be used for masked face i.e., in COVID pandemic which will help to identify person easily.

References

[1] Xi Zhao, Emmanuel Dellandrea and Liming Chen, “A People-Counting System based on Face-Detection and Tracking in a Video”, Advanced Video and Signal BasedSurveillance, (2009).

[2] Tsong-Yi Chen, Chao-Ho Chen, Da-Jinn Wang, and Yi-Li Kuo, “A People-Counting System Based on Face-Detection”, Fourth International Conference on Genetic and Evolutionary Computing, (2010).

[3] J. L. Crowley, T. Cootes, “FGNET, Face and Gesture Recognition Working Group”, http://wwwprima. inrialpes.fr/FGnet/html/home.html, (2009).

[4] B. Fasel and J. Luettin, “Automatic Facial Expression Analysis: A Survey,” Pattern Recognition, Vol. 36, pp. 259-275, (2003).

[5] Abhishek Bansal, Kapil Mehta, Sahil Arora, “Face Recognition using PCA & LDA algorithm”, Second International Conference on Advanced Computing and Communication Technologies, (2012).

[6] Arulkumar, C. V., and P. Vivekanandan. "Multi-feature based automatic face identification on kernel eigen spaces (KES) under unstable lighting conditions." Advanced Computing and Communication Systems, 2015 International Conference on. IEEE, 2015.

[7] Tahia Fahrin Karim, Molla Shahadat Hossain Lipu, Md. Lushanur Rahman, Faria Sultana, “Face Recognition Using PCA Based Method”, (2010).

[8] Omron, “OKAO Vision,” http://www.omron.com/r_d/coretech/vision/okao.html, (2009).

[9] R. Karthika, Parameswaran Latha, B.K., P., and L.P., S., “Study of Gabor wavelet for face recognition invariant to pose and orientation”, Proceedings of the International Conference on Soft Computing Systems, Advances in Intelligent Systems and Computing, vol. 397, pp.501-509, 2016.

[10] Mrs. Monali Nitin Chaudhari, Ms. Gayatri Ramrakhiani, Ms. Rakshita Parvatikar, Ms. Mrinal Deshmukh,” Face Detection using Viola Jones Algorithm and Neural Networks”, Fourth International Conference on Computing Communication Control and Automation (ICCUBEA) 2018.

[11] Basilio, Jorge Alberto Marcial, et al. "Explicit image detection using YCbCr space color model as skin detection." Proceedings of the 2011 American conference on applied mathematics and the 5th WSEAS international conference on Computer engineering and applications, IEEE ICCE. 2011.

[12] Mahmood, Arif, and Sohaib Khan. "Correlation coefficient-based fast template matching through partial elimination." Image Processing, IEEE Transactions on 21.4 (2012): 2099- 2108.

[13] Vikram K, Dr.S.Padmavathi, "FACIAL PARTS DETECTION USING VIOLA JONES ALGORITHM",2017 International Conference on Advanced Computing and Communication Systems (*ICACCS* -2015), Jan. 06 – 07, 2017, Coimbatore, INDIA.