A Method for Identification of Face of the Person and Presence of Facial Spot in the Face

Payal Bose[1] and Prof. Samir K Bandhyopadhyay[2]

[1]Research Scholar at Lincoln University College, Malaysia

[2]Distinguish Professor at Lincoln University College, Malaysia

Abstract

Human Face and facial parts are the most significant parts as it reveals a person's true identity. It plays an important role in various biometric applications like crowd analysis, human tracking, photography, cosmetic surgery, etc. There are many techniques are available to detect a facial image. Among them, skin detection is the most popular one. The aim of this paper is to detect first the person's identity from facial image and finally check any spot present the the detected person. The first step is to detect the maximum skin region based on a combination method of RGB and HSV color space model. Next it is to verify the skin areas of human through machine learning approach. The Aggregated Channel Features (ACF) detector is used to identify the different facial parts like eye pairs, nose, and mouth. Bootstrap aggregation decision tree classifier is applied to classify the person's identity based on Histogram Oriented Gradient (HOG) features value. The experimental results show that the proposed method gives the average 97% accuracy.

Keywords: Skin Detection, Color Space Model, Aggregated Channel Features (ACF) Detector, Histogram Oriented Gradient (HOG) Features Detection, Bootstrap Aggregation Decision Tree Classifier, Spot Detection

INTRODUCTION

In computer vision and human-computer interaction, process for face detection and its recognition plays a vital role in many applications like face tracking, tagging, crowd analysis, face authentication and many more. It is very challenging problem because detection of faces depends on lighting conditions, postures, orientations, age, gender and angles. Several methods are available to detect and recognized a face in different conditions.

Skin color detection is one of the most popular approaches to detect a face from a color image since skin color detection does not rely on the poses or orientation or any significant features of a face. In this paper, the first motive is to build a system that identifies the different facial parts from a face. The face is detected by using skin color segmentation techniques and the different facial parts are separated using an aggregated channel features detector. Aa face recognized system is developed next considering with high accuracy. The effectiveness of the face recognition system is to correctly identify the authorized person from a set of facial images. The main key point is to detect all the features from the image correctly through a classifier. In this paper, HOG features extractor used to extract the features from the image. The bootstrap aggregation

decision tree method is an ensemble decision tree classifier and it provides a tree-based classification with high performance measure and a higher classification rate with lower computational cost.

RELATED WORKS

Face is one of the most important external parts of the human being. For an automated face recognition system, face detection is a critical part of every aspect. Face detection aims to decide facial parts from facial image.

Researchers shows an efficient method to detect the human faces based on combinations of three popular color space models RGB, YCbCr, and HSV [1]. Here the thresholding concept was used. Based on the geometrical features and the skin color segmentation this algorithm can able to identify the single or multiple faces from an image with accuracy 96%.

Authors proposed a fuzzy system to detect human skin color[2]. The proposed method accepts RGB, HSV, and YCbCr color space to develop the design and assume each color space as a fuzzy set. The method builds up a three partition fuzzy entropy approach to calculate all the necessary parameters for the system. The average accuracy for the system is 93% with low false-positive and false-negative rate.

Combination of Aggregated Channel Features (ACF) detector and Faster R-CNN are used to improve object detection performance in fetal ultrasound images [3]. Authors proposed a combined method of Aggregated Channel Features (ACF) detector and the Convolution Neural Network (CNN), R-CNN for detection of the cross-sectional area of fetal-limb in an ultrasound image. This method solves the problem where the objects are not properly visible. It is based on a small amount of training set. This method improves the average precision and detection rate with less time for detection of facial image.

Authors studied of the redundancy of image features using HOG feature extraction [4]. This study shows the advantage of inter-channel correlation based on HOG feature extraction from the Bayer pattern image with proper gamma compression. The experimental result can be used in pedestrian detection systems.

Researchers described a face recognized method under illumination variation, facial expression, different poses, face orientation, occlusion, nationality variation, and camera motion conditions [5]. They used a combination of Local Binary Pattern (LBP) and HOG feature extraction for lower computational time. For classification, they used the latest and accurate classification model Random Forest classifier. The recognition rate for the proposed algorithm is 97.6%.

Authors proposed a thyroid disease predicting model based on the ensemble of Bagging with J45 and ensemble of Bagging with SimpleCart to extract useful information and diagnose diseases[6]. The overall accuracy is 99.66%. The data is too small for the proposed method.

An image processing technique to detect and verify the facial mole was proposed by some researchers [7]. The work is divided into three parts. In the first part, they used the Voronoi diagram to partition a mole-face into regions. In the second part, a Laplacian of Gaussian method is used to get the prominent features on the face. Here the aspect ratio and area are two features for mole verification. Finally, they developed an algorithm to detect the user moles to the sample mole-face by Thin face by Thin-plate Spline Analysis for mole recognition.

PROPOSED METHODOLOGY

Face recognition is a process to identify or verify one or more person's appearance in an image frame. This total recognition process categorized into two sections first needs to detect a person's face in an image frame and second, recognizing the person's identity. To identify a person's identity a similar face is searched in a collection of faces and the corresponding label. The basic face recognition process is shown in figure 1.
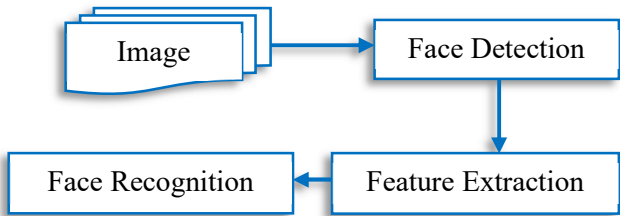

Fig: 1 Basic Process of Face Recognition

The proposed framework of this face recognition system is shown in Fig 2. The whole process divided into three parts. First, read the input image and, detect the face area then identify different facial parts, eyes, nose, or mouth to verify the face. Second, collect the features from the image and, finally identify the person in the image based on the selected features.
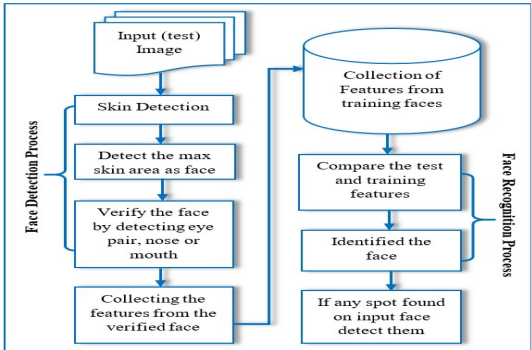

Fig: 2 The Proposed Framework

A. Face Database

In this paper, a face database name Tufts Face Database Computerized Sketches (TD_CS) downloaded from Kaggle (github) is used. In this database more than 100+ images are present. All the images have a very high resolution. For training and test, the images are pre-processed by cropping the extra background spaces [8].

B. Face Detection Process

In this paper, the face detection process divided into two parts. First, detecting the maximum skin region by skin color segmentation and, second verify the skin region as a face by detecting eye pair, nose, or mouth.

1. Skin Detection

Skin region detection is one of the most important parts of the image process. In images, the presence of skin color defines the existence of human or human-like. When a person's image is capturing, it depends on different external factors like light, shadows, etc. In images, the skin color represented by the pixel values with different color ranges. Therefore it is a very much challenging task to identify the suitable color range to detect it properly. To detect the skin color various techniques are available. In this paper, the combination of two color space models RGB and HSV are used to detect the skin region [1,9-10].

1.1. RGB Color Space

The default color space for a digital color image is RGB. It is much easier to transform from RGB to different color spaces. To detect the skin color, it is very much important to choose the correct color space.

In this experiment, each color channel of RGB color space (r, g, b) is segmented with a certain threshold value, and after that, a median filter is used to each channel for reducing the noise. The thresholding range of three color space is described below:

$$Channel_r = \ 0.5 \ < \ r \ < \ 0.6 \ldots\ldots \quad (1)$$
$$Channel_g = 0.25 \ < \ g \ < \ 0.3 \ \ldots\ldots (2)$$
$$Channel_b = \ 0.13 \ < \ b \ < \ 0.2 \ \ldots\ldots (3)$$

After thresholding, combine all the channels to get the final segmented channel.

$$RGB_{seg} = \ combine \ (Channel_r, Channel_g, Channel_b) \ldots (4)$$

1.2. HSV Color Space

HSV stands for Hue, Saturation, and Value where hue defines the dominant color, saturation defines the brightness area and value defines the luminance area of an image. This color space differentiates luminance and chrominance very finely and the transformation from RGB to HSV is unchanged in different lighting conditions. This is the reason this color model is very much useful and popular in skin color detection.

Similarly, like RGB color segmentation for skin detection, first segmented all the color of HSV channels (h, s, v) with a certain color threshold range and after that use a median filter to reduce the noise. The thresholding range of three color space is described below:

$$Channel_h \; = \; 0.25 \; < \; h \; < \; 0.32 \; ..... \,(5)$$

$$Channel_s \; = \; 0.65 \; < \; s \; < \; 0.73 \; ...... \,(6)$$

$$Channel_v \; = \; 0.47 \; < v \; < \; 0.51 \; ..... \,(7)$$

After thresholding, combine all the channels to get the final segmented channel.

$$HSV\_seg \; = \; combine(Channel_h, Channel_s, Channel_v \,) \,.... \,(8)$$

Now, combine equation 4 and 8 to get the required skin detection for this experiment,

$$Proposed\_Skin_{detection} \; = \; combine \,(RGB_{seg}, HSV_{seg}) \; ...... \,(9)$$

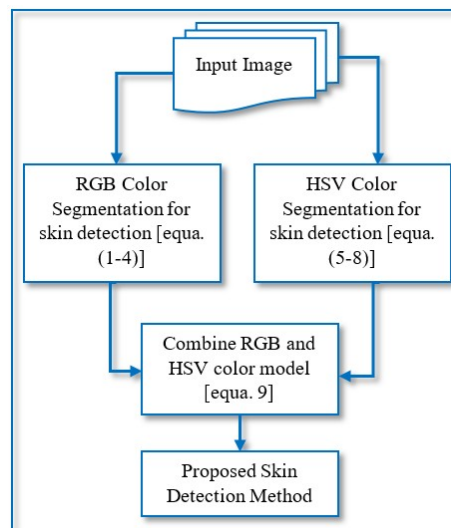The proposed skin detection method for an image is shown in Fig: 3.

Fig: 3 The Structure of proposed Skin detection method

Fig: 4 shows the result of segmentation details of combined skin detection methods.

(a) Original Image

(b) RGB color Segmentation

(c) HSV Color segmentation

(e) Combine Skin detection Method

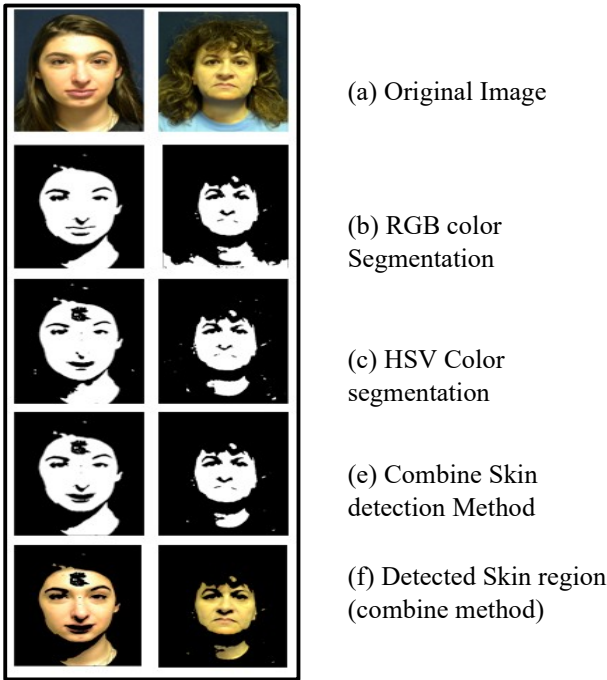(f) Detected Skin region (combine method)

Fig: 4 Result of Skin Detection Method

2. Detect The Face Area

After detecting the skin region, to detect the face area first, calculate all skin areas. Then find the maximum area among them. That maximum area accepted as the face region. These are shown in figure 5 and figure 6.
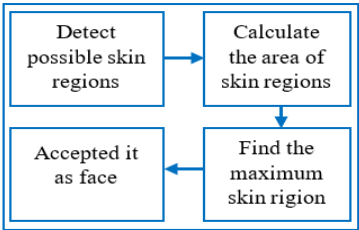

Fig: 5 The steps to detect a face

(a)   (b)   (c)

Fig: 6 a) Original Image, b) Detected maximum skin (as face) area, c) cropped the area

3. Validity Checking of Detected Face

After detecting the maximum face area, it is needed to check the face region belongs to a face or not. In this experiment to check the validity of the detected face, the Aggregated Channel Features (ACF) detector [3] is used. In this paper, for checking the detected face validity try to detect all the facial parts eye pair, nose, and mouth using the ACF detector. If eye pair with nose or mouth any two or all three found, then it accepted as a valid face.

A. Aggregated Channel Features (ACF) detector

An object detector is a computer program that used computer vision, image processing, and artificial intelligence algorithm to detect the feature values of the region of interest in an image. To classify an unknown data these object detectors used computer vision to extract the features values from an image. Then use machine learning or deep learning techniques to classify the features.

The Aggregated Channel Features (ACF) detector [11] is one of the machine learning techniques to classify data. This is a fast and effective sliding window detector. It is an evolution of the Viola & Jones (VJ) detector. ACF is best suited for quasi-rigid object detection, like detecting face, facial parts, cars, etc.

Before using this detector it is required to collect all the information provided by the empirical evidence and observation called the ground truth. In this case, it is known as a set of labeled images. These labeled images contain, the location of the images, their object classes, and the location of regions of interest of those images. The workflow of this detector is shown in figure 7. The detection of eye pairs, nose and mouth using ACF detector is shown in figure 8.
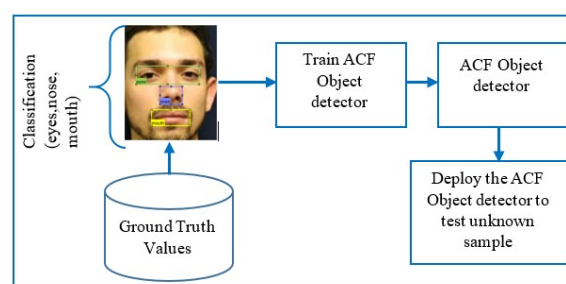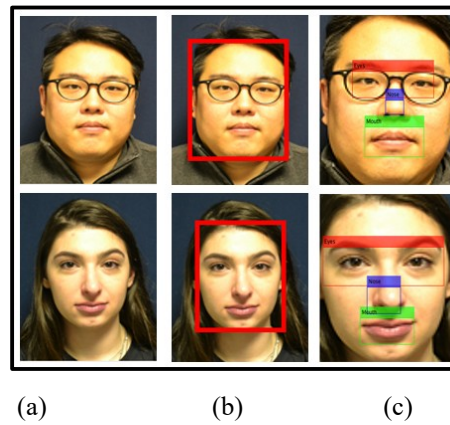

Fig: 7 Workflow of ACF Detector

(a)   (b)   (c)

Fig: 8 a) Original Image, b) Detected face region after skin color segmentation, c) Eye pair, Nose and Mouth detection using ACF detector

C. Face Recognition Process

The next step after successfully detected a human face is recognized the human identity [12]. To identify a face, it is necessary to extract facial features or the characteristics of the image. Feature extraction is a task to reduce high-dimensional training images to a set of features of investigation properties. In this experiment, the HOG used to extract the features.

1. Histogram of Oriented Gradients (HOG):

Histogram of oriented Gradients [4-5,13] method often used in computer vision to detect an object. This feature extraction method is done by extracting the gradient and orientation at the edge of an image. This method broke the complete image into many smaller regions, then calculate the gradients and orientation for each region. Finally, calculate the histogram each of these regions separately.

This HOG feature extraction method follows three main steps, 1) Prepossessing the data, 2) Calculate gradients, 2) then calculate the magnitude and orientation by using the below equations:

$$magnitude = \sqrt{g_x^2 + g_y^2} \ldots \ldots (10)$$

$$Orientation(\emptyset) = arctan \frac{g_x}{g_y} \ldots \ldots (11)$$

Where, $g_x,$ and g_y are the gradient values along x and y direction.

In this experiment, the image size (150×120) pixel is taken for processing the data. Then the above steps are used to calculate the HOG features of all the images. This size is good enough to create more than 8000 feature values for every single image.

2. Bootstrap Aggregation Decision tree Classifier (TreeBagger)

Bootstrapping is a powerful statistical method for estimating a quantity from a dataset. It creates many sub-sample datasets from the original dataset and calculates the mean value of them. Then it takes the mean value of all the mean sub-samples to estimate the final quantity of the given dataset. Bootstrap Aggregation (Bagging) [6][14-15] is a very powerful ensemble technique. An ensemble is a technique which combines all the prediction from different machine learning algorithms and predicts more accurate result than any other individual models. The bootstrap algorithm is the general procedure to reduce the high variances of the algorithm. The decision tree is very sensitive to the trained data. If the data changed then the tree and the prediction value will be changed.

In this experiment, the bagging classifier, TreeBagger algorithm is used to predict the model. The workflow of this model is given in figure 9.
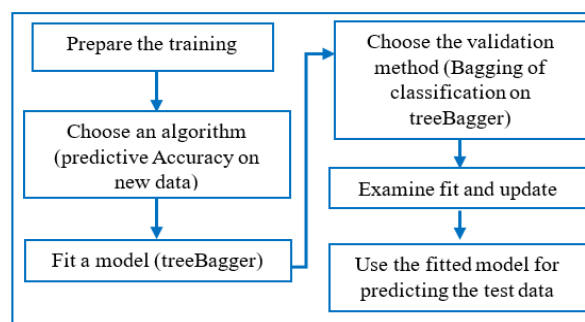

Fig: 9 The Workflow of the TreeBagger Classifier

D. Facial Spot Detection

After face recognition, the next step is to detect whether any facial spot present in the input query image. In this experiment to detect the facial spot first convert the input query image and the matched person image into grayscale, then calculate the difference between two images. Finally, use the dilation operation of morphological operation to identify the spot perfectly and give the boundary to the detected areas [7,16]. Figure 10 shows the works flow of spot detection of the identified face.
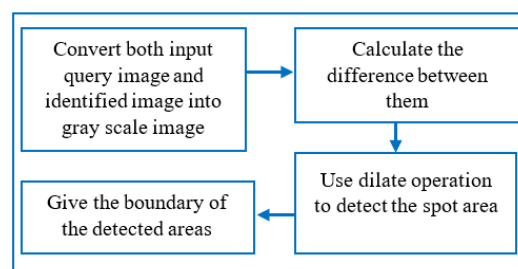

Fig: 10 The workflow of facial spot detection

I. EXPERIMENTAL RESULTS AND DISCUSSIONS

The experiment has three major stages. First stage is to detect the human skin and find the maximum skin area. Second stage is to check if the maximum detected skin area belongs to a human face or not. And third stage is to identify the face identity. The next stage is to check if any spot found on the identified face or not. If found then mark it.

The first experiment is to detect the skin region based on equations 1 to 9 and then determine the maximum skin area. The second experiment is to determine whether the maximum skin area belongs to a human face. For this experiment ACF detector used to identify all the facial parts, eyes pair, nose, and mouth. Eyes are the main part of a face, therefore eye pair detection is one of the most important parts in this experiment. If no eye pair found that implies face validation error. Again, if the eye pair found and nose or mouth any one or both found then either way, the maximum skin region accepts as the valid face. It is shown in figure 11. The detection of valid face is shown in figure 12
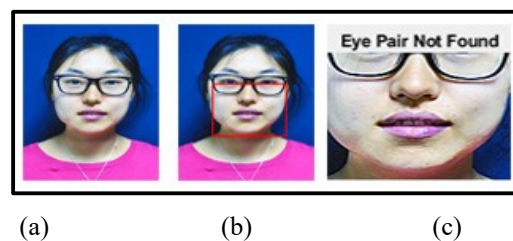

(a)　　　　(b)　　　　(c)

Fig: 11 a) Original image, b) Detected Maximum Skin area, c) Validation Checking (Not a Valid Face)
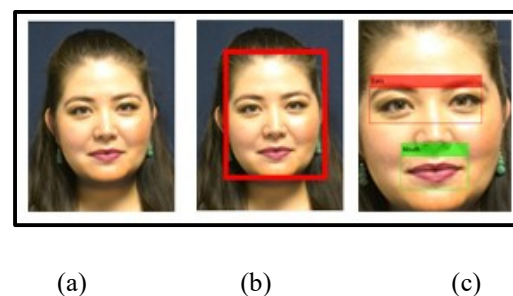

(a)　　　　(b)　　　　(c)

Fig: 12 a) Original image, b) Detected Maximum Skin area, c) Validation Checking (Valid Face)

In this experiment, the test data divided into two categories such as person without glass and person with glass. It is shown in figure 13.
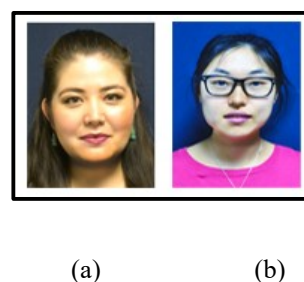

(a)　　　　　(b)

Fig: 13 a) person without glass, b) person with glass

The face detection rate for these categories is shown in table 1.

Table:1 Face Detection Rate

	Person without glass	*Person with glass*
Face Detection rate	98.6%	91.4%

The average precision values for eye pair, nose, and mouth detection for valid faces are given in table 2.

Table 2: Average Precision values for Eyes, Nose, and Mouth Detection

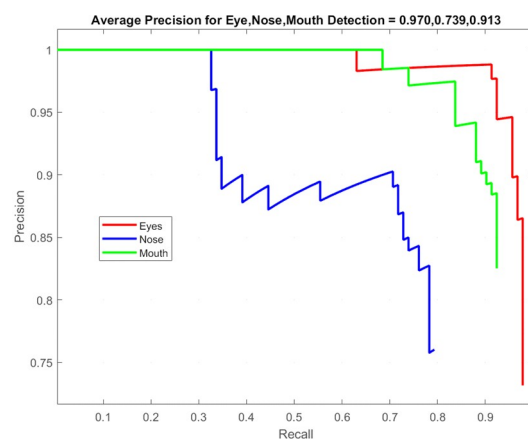|	*Eyes*	*Nose*	*Mouth*
Avg. Precision value	97%	73.9%	91.3%

Fig: 14 Precision vs Recall graph for Eyes, Nose, and Mouth Detection

Precision vs Recall graph for Eyes, Nose, and Mouth Detection is shown in figure 14.

The third experiment is to identify valid faces. In this experiment, a Bootstrap Aggregation classifier is used to identify the test face..The decision making tree classifier treeBagger trained more than 100+ training sample images with 8000+ dataset of each image. The recognition rate for the experimental classifier and the comparison with other classifiers are shown in table 3. The classification result of the proposed method is shown in figure 15.
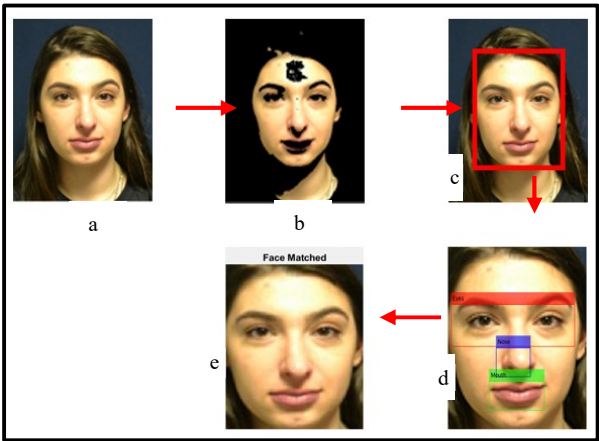
Fig: 15 a) Input Query image, b) Detected Skin area, c) Maximum Skin Area, d) Validation Checking (Valid face) e) Face Classification (Face Matched)

Table: 3 Comparison of Different Classifier

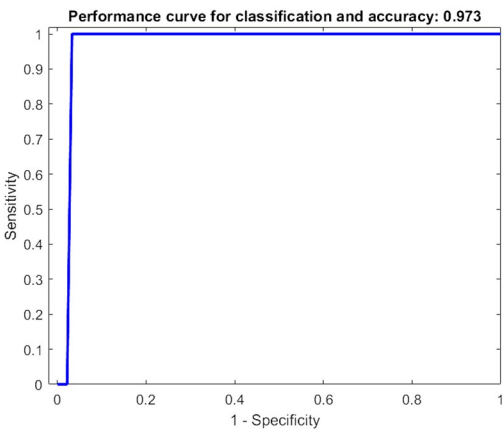| Classifier Name	Accuracy	Time Taken (sec)
Proposed Method	97.3%	80.80
Multi-Class SVM	95.2%	545.67
KNN	95.6%	10.12
Fine Tree	64.5%	616.90

Fig: 16 Average Performance Curve of the Classifier

The next step after identifying the image label of the input query image is to detect whether any spot present on the face. If present then simply detect the place. The average performance curve of the classifier is shown in figure 16. Figure 17(a) shows training images where as figure 17 (b) detects spot in the face.
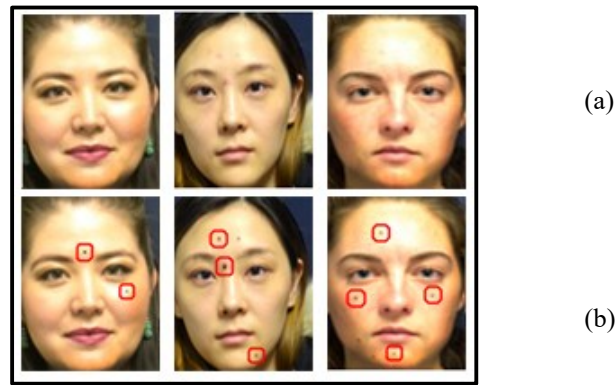
(a)

(b)

Fig: 17 a) Training Images, b) Input Query Images (spot detection)

CONCLUSIONS

Automated face recognition system is very demanding in society. This study shows an improved and effective approach to detect and identify a face from an image accurately. The appearance of a face detected by the combination of two color space model RGB and HSV skin segmentation method. Then the face validity checked by detected the facial parts eye pair, nose, or mouth by ACF detector. HOG feature extraction method is used to measure all features of a valid face and then a Bootstrap Aggregation Decision tree Classifier is proposed to classify the image class. The proposed method is defined, performed, and tested on TD_CS. After recognized the person's identity the system also checks if any kind of facial spot is present in the input face or not. The calculation is based on the difference between the input query image and matched person's image. Finally, a graphical user interface (GUI) is created for better use for all users.

REFERENCES

[1] R. Mohanty and M. V Raghunadh, "Skin Color Segmentation based Face Detection using Multi-Color Space," vol. 5, no. 5, pp. 470–475, 2016, doi: 10.1021/la900547b.

[2] F. A. Pujol, M. Pujol, A. Jimeno-Morenilla, and M. J. Pujol, "Face detection based on skin color segmentation using fuzzy entropy," *Entropy*, vol. 19, no. 1, pp. 1–22, 2017, doi: 10.3390/e19010026.

[3] F. A. Hermawati, H. Tjandrasa, and N. Suciati, "Combination of Aggregated Channel Features (ACF) detector and Faster R-CNN to improve object detection performance in fetal ultrasound images," *Int. J. Intell. Eng. Syst.*, vol. 11, no. 6, pp. 65–74, 2018, doi: 10.22266/ijies2018.1231.07.

[4] W. Zhou, S. Gao, L. Zhang, and X. Lou, "Histogram of Oriented Gradients Feature Extraction from Raw Bayer Pattern Images," *IEEE Trans. Circuits Syst. II Express Briefs*, vol. 67, no. 5, pp. 946–950, 2020, doi: 10.1109/TCSII.2020.2980557.

[5] H. Mady and S. M. S. Hilles, "Face recognition and detection using Random forest and combination of

LBP and HOG features," *2018 Int. Conf. Smart Comput. Electron. Enterp. ICSCEE 2018*, no. July, pp. 1–7, 2018, doi: 10.1109/ICSCEE.2018.8538377.

[6] A. O. J and F. Ogwueleka, "Effective and Accurate Bootstrap Aggregating (Bagging) Ensemble Algorithm Effective and Accurate Bootstrap Aggregating (Bagging) Ensemble Algorithm Model for Prediction and Classification of Hypothyroid Disease," no. July, 2020, doi: 10.5120/ijca2020920542.

[7] C.-C. Hsieh and J.-A. Lai, "Face Mole Detection, Classification and Application," *J. Comput.*, vol. 10, no. 1, pp. 12–23, 2015, doi: 10.17706/jcp.10.1.12-23.

[8] Panetta, Karen, Qianwen Wan, Sos Agaian, Srijith Rajeev, Shreyas Kamath, Rahul Rajendran, Shishir Rao et al. "A comprehensive database for benchmarking imaging systems." IEEE Transactions on Pattern Analysis and Machine Intelligence (2018). Website: http://tdface.ece.tufts.edu/

[9] N. Razmjooy, B. S. Mousavi, and F. Soleymani, "A hybrid neural network Imperialist Competitive Algorithm for skin color segmentation," *Math. Comput. Model.*, vol. 57, no. 3–4, pp. 848–856, 2013, doi: 10.1016/j.mcm.2012.09.013.

[10] M. Z. Osman, M. A. Maarof, and M. F. Rohani, "Improved dynamic threshold method for skin colour detection using multi-colour space," *Am. J. Appl. Sci.*, vol. 13, no. 2, pp. 135–144, 2015, doi: 10.3844/ajassp.2016.135.144.

[11] P. Blondel *et al.*, "Collaborative training of far infrared and visible models for human detection," *Int. J. Simul. Multidiscip. Des. Optim.*, vol. 10, 2019, doi: 10.1051/smdo/2019016.

[12] S. Il Choi, S. S. Lee, S. T. Choi, and W. Y. Shin, "Face Recognition Using Composite Features Based on Discriminant Analysis," *IEEE Access*, vol. 6, pp. 13663–13670, 2018, doi: 10.1109/ACCESS.2018.2812725.

[13] C. Q. Lai and S. S. Teoh, "An efficient method of HOG feature extraction using selective histogram bin and PCA Feature reduction," *Adv. Electr. Comput. Eng.*, vol. 16, no. 4, pp. 101–108, 2016, doi: 10.4316/AECE.2016.04016.

[14] C. Beaulac and J. S. Rosenthal, "Predicting University Students' Academic Success and Major Using Random Forests," *Res. High. Educ.*, vol. 60, no. 7, pp. 1048–1064, 2019, doi: 10.1007/s11162-019-09546-y.

[15] T. H. Lee, A. Ullah, and R. Wang, "Bootstrap Aggregating and Random Forest," *Adv. Stud. Theor. Appl. Econom.*, vol. 52, pp. 389–429, 2020, doi: 10.1007/978-3-030-31150-6_13.

[16] U. R. Gogoi, M. K. Bhowmik, P. Saha, D. Bhattacharjee, and B. K. De, "Facial mole detection: An approach towards face identification," *Procedia Comput. Sci.*, vol. 46, no. Icict 2014, pp. 1546–1553, 2015, doi: 10.1016/j.procs.2015.02.080.