Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

Pose and Head Orientation Invariant Face Detection Based on Optimised Aggregate Channel Feature

P. Naveen[1*], P. Sivakumar[2]
[1]Assistant Professor, Department of ECE, Kalasalingam Academy of Research and Education
[2] Professor, Department of ECE, Kalasalingam Academy of Research and Education
*naveenamp88@gmail.com, sivapothi@gmail.com

ABSTRACT

Face detection plays a critical part in the identification of people in biometrics. Face detection with the pose or head orientation image is really a challenging problem with surveillance applications since most techniques used in face detection only show better results when the front-facing images exist. The size of the image dataset is more and takes more time for training and detecting. The goal of our proposed work is to establish an effective method with minimum training time for detecting human faces, which can deal with both front and pose or head orientation faces. To accomplish this goal, (ACF) Aggregate Channel feature face detector with optimized parameters was proposed. Our proposed ACF face detector is a learning-based algorithm implemented with Pointing'04 and validated with FEI datasets and wild dataset. We implemented our proposed ACF face detector in MATLAB 2019a and compared their performance with other existing algorithms. The performance parameters of our proposed method such as average precision and log average miss rate are calculated, and the values are 0.90 and 0.18, respectively.

Keywords
Face detection, ACF (Aggregate Channel feature), pose and illumination invariant face.

Introduction

Face detection has gained widespread attention for its wide range of applications in computer vision and machine learning science, including human robot interaction, biometric verification, and surveillance.

The face, which incorporates significant human biological information, is a very fascinating object in images and videos. Obviously, face detection at the very beginning, is considered a crucial part of any automated face treatment system. It is also a difficult task, as the difficulties of designing a reliable face detector arise not only from the variation in human faces (e.g., changeability in size, location, pose, direction, and expression) as well as from changes in environmental conditions (e.g., illumination, lighting, occlusion, etc.)[1].

Face recognition helps to recognise individuals from still images or videos. Side vision face recognition has a variety of uses because when portraits of the front view are not available, the side view face can be used as a compatible authentication / verification feature to boost accuracy. Home safety applications are one potential deployment area for face recognition techniques. Because of the parents and siblings' busy schedule, underestimated hazards or external threats, numerous people experience injuries and fatalities that arise in the environment at home. Such incidents can be avoided by situational awareness, and face recognition is one of the tools which can be used to this end.

Hence face recognition and retrieval of different face features should be very accurate without much error using such methods. But when dealing with facial recognition using side view or

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

multi-angled image it is a difficult task as stated by authors [2, 3]. In this paper, we aim to create a fast and accurate pose or head orientation invariant face detector that can detect faces with head orientation varying from -90 to + 90 degrees.

In their Robust Real-Time Face Detection, Paul Viola and Michael J. Jones showed a new algorithm with considerable accuracy in a stable, real-time system.[4] and this technique has already been implemented in OpenCV in the operation of cvHaarDetectObjects() which is basically used during face detection.[5]. Even though it correctly detects the front face, the most important flaw or fault of the algorithm is that the algorithm is not reliable if the faces were tilted above (in plane) and (out of plane).

In this research work, we proposed optimised ACF based face detector to overcome difficulties in detecting faces due to pose and head orientation angle. Our proposed algorithm is a learning-based algorithm that detects the face exits in the image based on the Region of Interest rectangle trained by the ACF face detector trainer. This algorithm was implemented with Pointing'04 and validated with FEI recognition datasets.

The average precision and log average miss rate were considered as evaluation parameters. The optimum value for negative sample factor and number of stages were tuned based on these parameters along with quality measure. The face detection performance of our proposed ACF face detector was compared with Viola Jones algorithm implemented in MATLAB 2019a and the results showed that our proposed ACF face detector detects better.

The remaining sections are planned as follows; Section 2 briefly consolidates the related work which is necessary for our proposed work. Section 3 discusses our proposed algorithm in detail. Section 4 discusses the dataset, evaluation parameter and examines the performance. Section 5 concludes our proposed work and states our future plan.

Related Work

This section discusses various methods carried out on face detection. Face detection has attracted much publicity since computer vision began in the early days.

An algorithm proposed by Viola Jones consists of three aspects such as quick computation of features through integral image representation, classifier learning using Adaboost, and structure of the attention cascade [4] and, they stated that the algorithm was not reliable if the faces were tilted above (in plane) and (out of plane).

Q. Zhu et al.[6] implemented a fast human face detection system by integrating a cascade of rejecters approach with Hog features. This method improved the speed but still the drawback was not rectified. Liu et al [7] has developed KL-Boosting to consolidate hair like features to a more discriminating KL function. Wang et al [8] introduced RNDA methodology as a hair-like function for the construction of complex optimal feature sets without geometric constraints. All of these approaches typically provide limited feature sets for the actual classifier with strong generalization capability, however struggle by more computational strain due to lots of roaming procedures engaged for each feature computation.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

Simpler and more open interpretation of the basic features such as Pair-wise point [9] and control points [10]. Each specified their characteristics clearly on the principle of pixels that was the most basic component of an image, whereas the normalization of average and standard deviation was rejected to prevent so much computing load. The move created a double-edged blade, i.e., the extraction of the feature wad extremely quick and the point-based feature appears to be ineffective for complex tasks.

Turk et al [11], created set of Eigen faces of some training images by applying principal component analysis. The distance from face space was calculated during reconstruction and the minimum distance was represented as face area. For each image location the distance from face was calculated. This method was not suitable for identifying face in all the angles.

Mikael Nilsso et al [12], suggested a method based on the local Successive Mean Quantization Transform features (SMQT) and the split up sparse Network of Winnows (SNOW). Even though the speed of face detection was improved, this method was able to detect faces of front pose, up and down vertical angles only.

Kirti Dang et al [13], conducted experiment on face detection with various face detection methods such as Viola-Jones, SMQT features & SNOW Classifier, Neural Network-Based Face Detection and Support Vector Machine-Based face detection and compared their performance based on precision and recall rate. According to their conclusion Viola Jones face detector has better precision and recall than SMQT features and SNOW classifier.

B. Wu et al[14], implemented Adaboost algorithm for rotation invariant face detection and they tested with CMU+MIT frontal face image test set and their algorithm detected 89.8% of the faces from the test face set of CMU profile with 221 false alarms. The image dataset did not cover all 360-degree angles.

Y. Li et al [15], suggested a support vector machine to train as a pose estimator for the view pattern and cascaded detectors which were individually learned. The complicated multi view face detection problem was little bit reduced by the pose estimator. But this method resulted in unstable pose estimation which weakened the stability of the system.

R. Brunelli et al [16] suggested template matching methods for detection face. Initial step of this method is defining a template i.e., a standard face pattern. Correlation values were calculated by comparing the template with different parts of the input image. On the other hand, an uncomplicated template matching approach is not stable because it cannot manage with faces of different expressions, poses, scales and lighting conditions.

The human face class was modelled as a set of characteristic components organized in spatial structure vectors and applied this model to face cluttered detection pictures by M.C. Burl et al [17]. Morphological operations were used by Han et al.[18] to identify Eye-analogy pixels in the input images. Since eyes furthermore, eyebrows are the most notable and stable highlights of the human face, a naming procedure was utilized to create the eye-simple sections that guided the quest for potential regions of the face.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

Xiaowei Zhao [19], proposed discontinuous Haar like feature for characterising facial context. In their method powerful representation of targeted tasks was formed by combining traditional Haar-like features (characterizing local texture information) and discontinuous Haar-like features (characterizing context constraints in global sense). Additionally, this method is a method of learning and the classifier required using huge instances of statistical learning.

Stan Z. Li et al [20], suggested a system which can learn to identify multi-view faces by the application of FloatBoost. The system used coarse to fine and simple to complex construction represented as detector pyramid. They used a CMU dataset for validation.

Deepak et al [21], proposed a face detection system based on skin colour segmentation. From the input image facial features were extracted and faces were detected by some set of rules. Wu-Chih et al [22], proposed feature based three stage algorithms based on skin colour. Light compensation was done at first stage using white as reference and followed by resizing the image using discrete wavelet transform, the YCbCr skin colour model and morphological processing. In the second stage, face candidates were obtained from face template and suitable face box measure. At the last stage, the face was detected by measuring facial features.

Various methods were proposed to modify Viola Jones method by Masek et al [24], Pandey et al [25], Zhu [26] et al, Egorov[27] et al, and Mutneja [28]. They utilized Haar features, skin colour extraction with AdaBoost algorithm to detect faces in various illuminating conditions, pose and occlusion.

A novel Direction-Sensitivity Features Ensemble Network for rotation-invariant face detection (DFE-Net) was proposed by to learn an end-to-end convolutional model for RIFD from coarse to fine [29]. A Direction-Sensitivity Features Ensemble Module (DFEM) was adopted to focus on the awareness of different face angles, which can learn and accurately extract features of rotated regions and locate rotated faces precisely [30].

H. Wu et al [31] put forward a multi-task convolutional neural network cascade framework for simultaneous face detection and pose estimation, in real-time. Face detection task resulted in gain boosted performance from pose estimation task using multi-task learning. Through cascade structure, this method achieved real-time performance for resolving these two tasks.

A simple but effective architecture, named Angle-Sensitivity Cascaded Networks (ASCN), for rotation-invariant face detection and alignment consists of three consecutive cascaded networks, first stage, the rotation angle is predicted and candidate bounding boxes are implemented [32]. ASCN refines the candidates and their orientations in the second stage. In the last stage, the accuracy of bounding boxes and alignment is improved. A pose-equitable loss to balance the faces with large poses is also employed.

To summarise, as stated at the start of this section, pose assessment depends on the different perspectives whereas face detection necessitates finding likeness of various views on removing non faces as fast as possible. Most used dataset was CMU+MIT. The dataset was bulkier and training time was also more.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

Proposed Algorithm

The workflow of our proposed algorithm is displayed in figure 1. Initially pre-processing was carried out. In this pre-processing process, RGB image was converted into gray image and gamma correction was done for balancing uneven illumination. After that on the face of the training images Region of Interest (ROI) was marked with bounding boxes which should be detected by the ACF based face detector. The ACFs were then extracted from the ROI. ACF face detector has a classifier, ROI with bounding boxes as positive samples and the rest of the images as negative samples. Positive and negative samples were then grouped as gtruth. With this gtruth the classifier was trained to detect the face. Then the classifier was used as an ACF based face detector to detect the face of the test image.

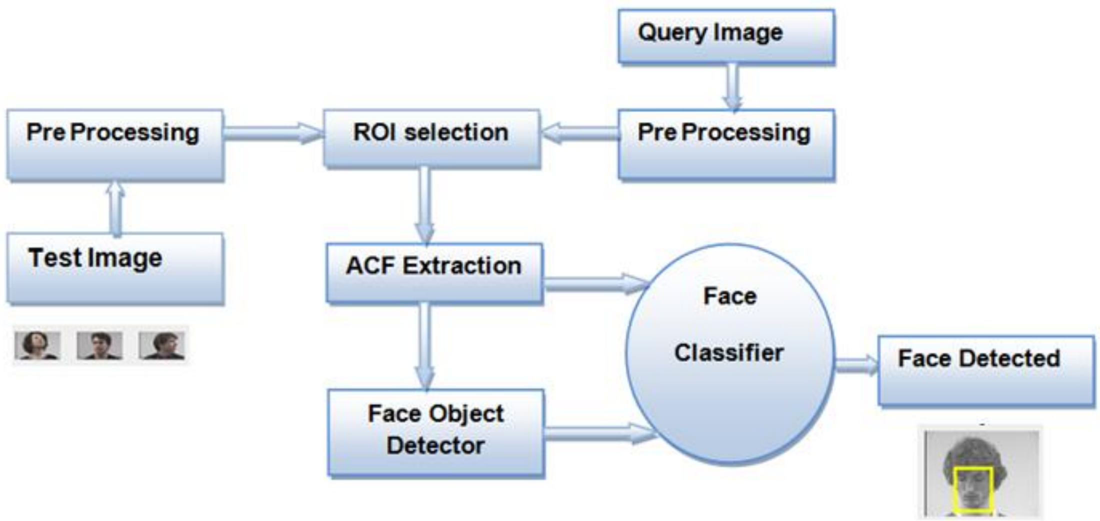

Figure 1. Proposed Algorithms' workflow

Selection of ROI

For detecting the face, it is necessary to minimize the region of the image. ROI was prepared in terms of gtruth using 60% of images. These images were selected randomly for ROI labelling and training. Images labelled with faces were defined as a table with two columns. The first column will include the gray scale image paths and file names. The second column includes M-by-4 matrices, which include the bounding box positions linked to the corresponding image. The second column represents a positive sample i.e., face. During the training process, negative instances were obtained automatically from the images. After choosing the ROI for the face detection phase, the ACFs were extracted from the face's detected region in the ROI and trained to detect faces.

Aggregated Channel Features Extraction

The ACFs have the aspect to easily extract inimitable features from the image. The input image was first created as a multi-resolution pyramid of images at a high speed to quickly extract the

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

ACFs. The feature information was extracted for the images included in the pyramid including the same and special features extracted from the high-resolution image to the low-resolution image. The ACFs have the characteristic of retaining the (k-1) extracted feature information, even if it is the extracted feature information of kth down sampled.

The input gray image (I) was computed as several ten feature channels (C = Ω(I)). These are uniform gradient magnitude, histogram of orientation gradient (six channels), and LUV colour channels (three channels).

Gradient magnitude of the image is represented by equation (1) and orientation of the image is represented by equation (2). Convolution operation was done between a triangular filter [1 2 1]/4 and the gradient image. As a result of this convolution, a smooth image was obtained. For getting the details of gradient scale the smoothen image was normalised by equation (3).

$$M(i,j) = \sqrt{\left(\frac{\partial I(i,j)}{\partial x}\right)^2 + \left(\frac{\partial I(i,j)}{\partial y}\right)^2} \quad (1)$$

$$O(i,j) = \left(\frac{\frac{\partial I(i,j)}{\partial y}}{\frac{\partial I(i,j)}{\partial x}}\right) \quad (2)$$

$$\widetilde{M}(i,j) = \frac{M(i,j)}{S(i,j)+c} \quad (3)$$

where,

I(x, y) – m X n RGB image

$\frac{\partial I(i,j)}{\partial x}$ - Derivative of image I at (i, j) coordinates in x direction

$\frac{\partial I(i,j)}{\partial y}$ - Derivative of image I at (i, j) coordinates in y direction

$S(i,j)$ - Smoothen image at (i, j) coordinates

c- Constant

The following steps explain the ACF extraction.

1. Images were converted into multiple low-resolution images using LUV transformation which form the first 3 channels with the addition of pixels in every 4× 4 block non-overlapping.
2. First derivatives of image in X – direction and Y-direction were calculated, and the next channel was formed with the normalized gradient magnitude with the addition of pixels in every 4× 4 block non-overlapping.
3 For six orientation bins such as 0°, 30°, 60°, 90°, 120°, 150° were calculated as Histogram of gradients for a cell size of four which results in six low-resolution channels.
4. These 10 channels were combined as a column matrix to represent the final feature vector.

Thereafter, a classification procedure was conducted to decide if the extracted feature parameters correspond to the target of interest. Among the ACF sources, the LUV colour model struck by the

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

reality that people are more sensitive to brightness than to colour. The channel L is the norm for light, the channel U for red and green and the channel V for blue and purple [23].

Face detection

Training was carried out with AdaBoost learning algorithm after extracting ACFs from the ROI. AdaBoost is an ensemble learning algorithm which combines several weak classifiers to build a strong classifier. At the end of the training process a powerful classifier was built as a face detector based on ACF. The face detector was tested with the test images. During the testing process, ACFs were extracted and compared with the ACFs which were trained already. Based on the comparison it detected the face with high precision.

Results and Discussion

In this section results and the image dataset used for our research work are discussed. The Viola Jones and Proposed ACF based face detection algorithm were implemented in MATLAB 2019a software and the results were compared.

Dataset

Our main aim is to detect face from pose or head orientation images with all possible angles from 0 to 360 degrees. So, we selected two public dataset Pointing'04[29] and FEI datasets [30]. Pointing'04 dataset has 2 sets of images of 15 persons and the dimension is 288 X 384 pixels. Every set includes 2 sequences of 93 photos of the same individual in various poses. The collection contains 15 people, having or not glasses and having different skin colours. The position, or head angle, is defined by 2 angles (h, v), varying from -90 to + 90 degrees. Figure 2 shows the sample images from the Pointing'04 dataset.

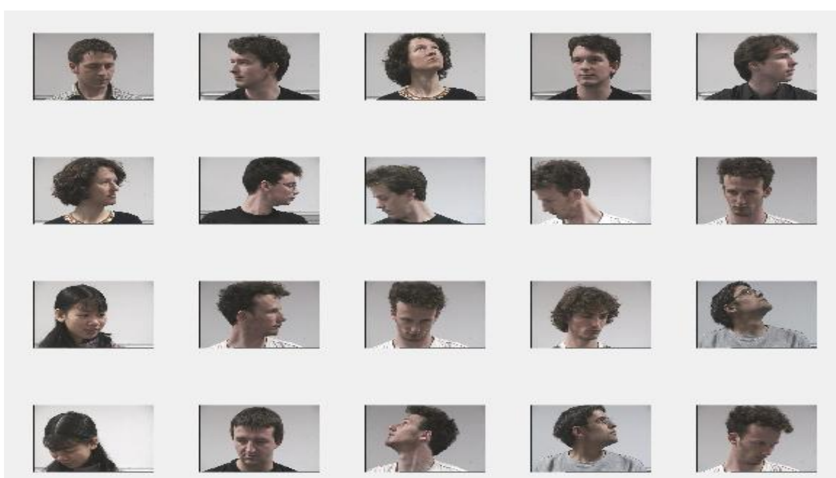

Figure 2. Sample images from the Pointing'04 dataset

FEI dataset has 14 images of 200 people. Both photographs are vibrant and taken in an upright frontal position with a profile rotation of up to around 180 degrees against a white homogenous backdrop. The scale will differ by about 10%, and the actual dimension of each photograph is 640x480 pixels. Figure 3 shows the sample images from the FEI dataset.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

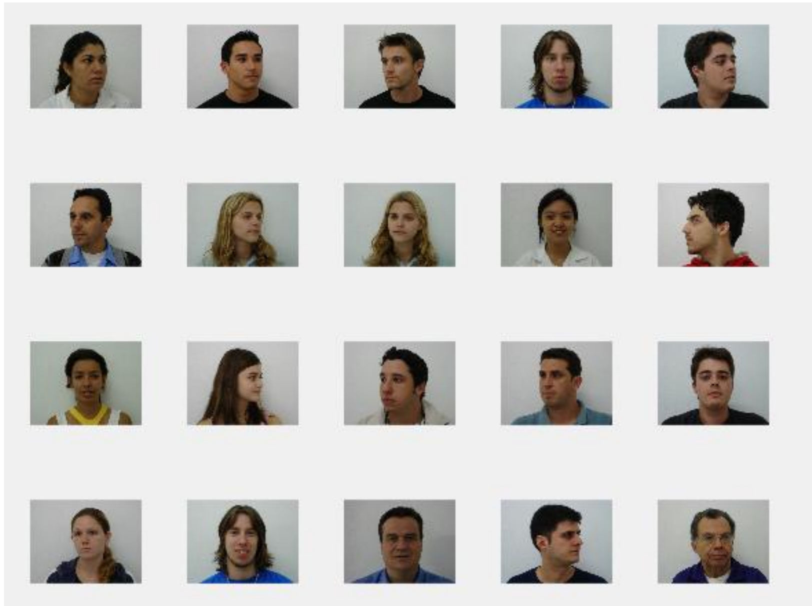

Figure 3. Sample images from the FEI dataset

Experimental Steps

For our research work we took 70% of Pointing'04 dataset as training image set and 30% as testing image set. Then the trained ACF face detector was validated with FEI dataset. The proposed algorithm was implemented by the following steps.

1. 60% of Pointing'04 dataset was selected randomly as a training image set and 40% images were used for testing.
2. Pre-processing such as RGB to gray image conversion and gamma correction were done.
3. ROI was labelled as a bounding box and saved along with the images as gtruth.
4. Then the ROI with images were trained by extracting ACF features specified in ROI.
5. The trained ACF face detector was tested with the testing image set.
6. The ACF parameters were evaluated and optimised based on average precision and log-average miss rate to result in better face detection.
7. Face detecting capability of our proposed algorithm was compared with Viola Jones algorithm.
8. The proposed algorithm was also validated with FEI dataset.

Evaluation Parameters

Several criteria were implemented to determine the efficiency of an ACF based face detector. They could be divided primarily into right metrics based on detection, and error metrics based on detection. Right detection related metrics include precision, recall and average precision.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

Precision is the amount of true positive over positive number. Positive numbers are the sum of true positive and false positive ones.

$$Precision = \frac{True\ positive}{True\ positive + False\ positive} \quad (4)$$

Recall is the amount of true positive over sum of true positive and false negative ones.

$$Recall = \frac{True\ positive}{True\ positive + False\ negative} \quad (5)$$

Average precision describes the proportional increase in accuracy with each recall shift for the thresholds in the precision-recall curve given by

$$Average\ precision = \sum_{i=1}^{n} (R_i - R_{i-1})\, P_i \quad (6)$$

where,
P_i- Precision at the ith threshold

R_i- Recall at the ith threshold

n − Number of thresholds

Metrics based on miss detection include miss rate (MR), false positive per image (FPPI), and log-average miss rate. Miss rate is the number of false negatives over the sum of the number of true positives and false negatives.

$$Miss\ rate = \frac{False\ negative}{True\ positive + False\ negative} \quad (7)$$

Log-average fault rate is calculated by an average fault rate of nine FPPI values uniformly distributed in log-space between 10^{-2} and 10^{0} [16].

Performance Analysis

There are two parameters needed to tune the proposed ACF face detector algorithm such as Number of stages and Negative Samples Factor. The chosen optimal parameters for our research are shown in Table 1.

While increasing the number of stages the training time increases and reduces the training error at each stage. Training error is given by the equation (8) and negative sample factor is given by equation (9)

$$Training\ error = \frac{1}{n} \sum_{i=1}^{n} (y_i - f(y)) \quad (8)$$

where,

n- Number of images in training set

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

y_i- response for i[th] image

$f(y)$ − Function which maps the set of training images

$$\text{Negative Samples Factor} = \frac{\text{Number of negative samples at each stage}}{\text{Number of positive samples used at each stage}} \quad (9)$$

Average Precision, Log average Miss Rate and Training time were considered to select the optimum parameters of the ACF face detector. Along with this quantitative measure, a quality measure was also carried out. Quality measure was done by selecting 20 random images from the testing dataset and the number of correctly detected faces was counted for selecting the optimum parameter. The optimum values were selected based on trial-and-error methods.

Table 1. Optimum parameter tuning

No. of Stages	Negative Samples Factor	Quantitative Measure			Quality Measure
		Average Precision	Log average Miss rate	Training time (Sec)	Correctly detected out of 20
4	*5*	*0.85*	*0.26*	*85.318*	*17*
	6	*0.86*	*0.25*	*89.21*	*18*
	7	*0.90*	*0.18*	*99.13*	*19*
5	*5*	*0.85*	*0.26*	*107.85*	*16*
	6	*0.86*	*0.25*	*109.88*	*19*
	7	*0.90*	*0.18*	*122.65*	*19*
6	*5*	*0.85*	*0.27*	*128.4*	*16*
	6	*0.86*	*0.25*	*134.7*	*17*
	7	*0.90*	*0.18*	*156.4*	*20*

Table 1 lists all the quantitative and qualitative measures based on the number of stages and negative sample factor. Highest average precision and lowest log average miss rate were obtained when negative sample factor was 7. Even though the number of stages was changed there is no variation in the precision and log average miss rate, but training time increased as well as change in the quality measure. Increasing the number of stages results in better quality measure. Hence the optimum parameters for our proposed algorithm were selected based on the quality measure. The optimum value for the number of stages was 6 and negative sample factor was 7. More specifically, with these parameters, the proposed ACF face detector achieved the highest performance in terms of average precision and log-average error rate. Figures 4 and 5 display the respective log-average error incidence curve and Precision-Recall curve.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

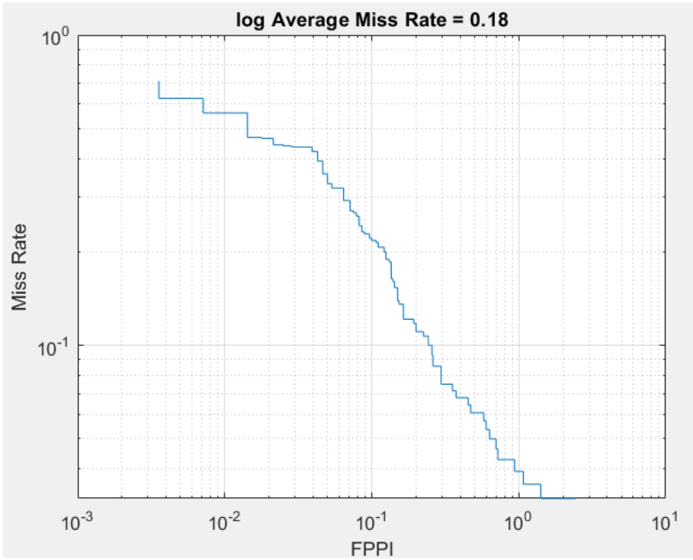

Figure 4. Average error incidence curve of ACF face detector

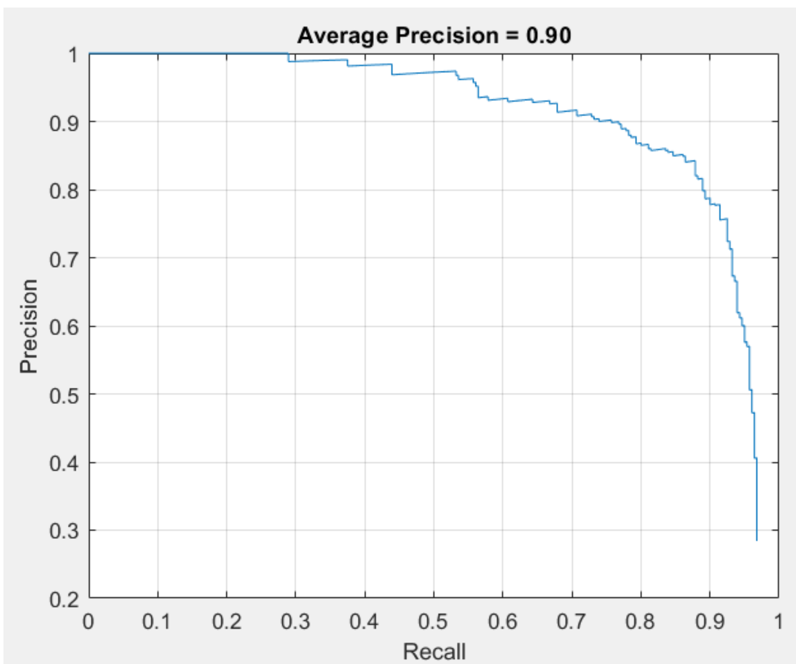

Figure 5. Precision-Recall curve of ACF face detector

With this optimum parameter the ACF based face detector was trained and the trained ACF detector was tested with the testing data. Figure 6 shows the result of our proposed face detector for the randomly selected 20 images from the testing dataset and it is inferred that the face was detected by our proposed detector irrespective of pose and head orientation of the face.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

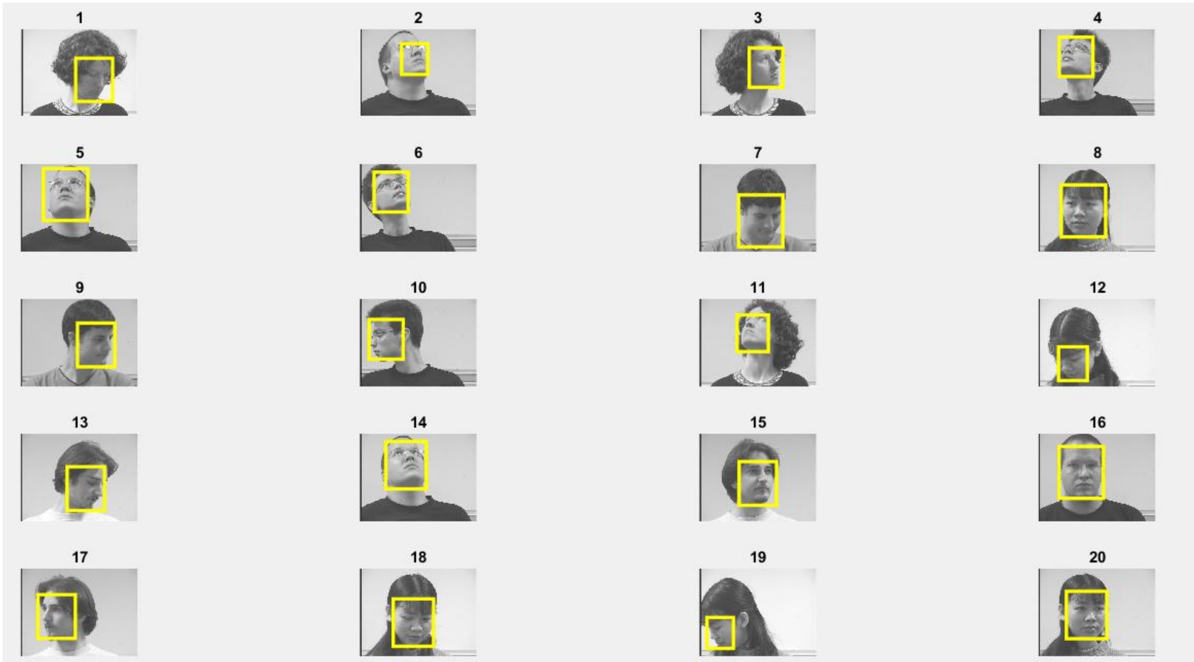

Figure 6. Proposed algorithm, testing with testing dataset

Figure 7 shows the result of our proposed method for randomly selected 20 images from the FEI dataset. All the faces of 20 images were detected by our proposed ACF detector irrespective of pose and illumination.

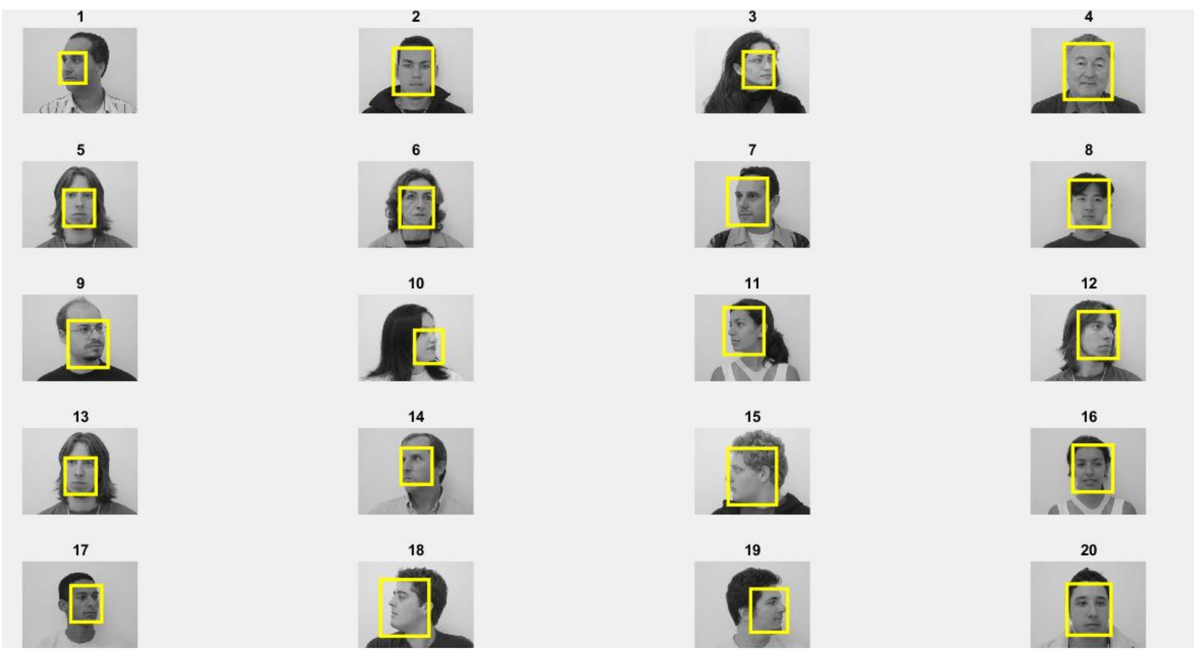

Figure 7. Proposed algorithm, validating with FEI dataset

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

Comparative Analysis

Table 2 displays all the possible vertical and horizontal positions of the head and the comparison of face detection capability of Viola Jones and our proposed algorithm using Piontingʼ04 dataset. The Head Orientation column lists all the 93 angles of the head of each person both vertical and horizontal. 93 images of one person were tested with both algorithms and the results were listed in Table 2.

Face detection from the images of single person from Pointing'04 dataset

With MATLAB 2019a, Viola Jones algorithm was implemented and tested with 93 images of one person as listed in Table 2. As per table 2 nearly 52 faces were not detected by Viola Jones algorithm. Figure 8 shows the face detection results of Viola Jones algorithm.

The same 93 images were tested by our proposed ACF face detector and all the faces were detected correctly irrespective of pose and head orientation. Figure 9 shows the results of our proposed ACF face detector with 20 images which were not detected by Viola Jones algorithm.

Table 2. Comparison among existing Algorithm with Proposed one with images of single person

S. No.	Head Orientation (vertical / horizontal)	Viola Jones Algorithm	Mutneja et.al.	SVM based method	CNN based Algorithm	Proposed Algorithm
1.	-90 / +0	Not Detected	Detected	Not Detected	Not Detected	Detected
2.	-60 / -90	Not Detected	Detected	Not Detected	Detected	Detected
3.	-60 / -75	Not Detected	Detected	Detected	Detected	Detected
4.	-60 / -60	Not Detected	Detected	Detected	Detected	Detected
5.	-60 / -45	Not Detected	Detected	Detected	Detected	Detected
6.	-60 / -30	Not Detected	Not Detected	Detected	Detected	Detected
7.	-60 / -15	Not Detected	Not Detected	Detected	Not Detected	Detected
8.	-60 / +0	Not Detected	Not Detected	Not Detected	Not Detected	Detected
9.	-60 / +15	Not Detected	Not Detected	Not Detected	Not Detected	Detected
10.	-60 / +30	Not Detected	Not Detected	Not Detected	Detected	Detected
11.	-60 / +45	Not Detected	Not Detected	Not Detected	Detected	Detected
12.	-60 / +60	Not Detected	Not Detected	Detected	Not Detected	Detected
13.	-60 / +75	Not Detected	Detected	Detected	Not Detected	Detected
14.	-60 / +90	Not Detected	Detected	Detected	Not Detected	Detected
15.	-30 / -90	Not Detected	Detected	Not Detected	Detected	Detected
16.	-30 / -75	Not Detected	Detected	Not Detected	Detected	Detected
17.	-30 / -60	Not Detected	Detected	Not Detected	Detected	Detected
18.	-30 / -45	Not Detected	Not Detected	Detected	Detected	Detected

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

S. No.	Head Orientation (vertical / horizontal)	Viola Jones Algorithm	Mutneja et.al.	SVM based method	CNN based Algorithm	Proposed Algorithm
19.	-30 / -30	Not Detected	Not Detected	Detected	Detected	Detected
20.	-30 / -15	Detected	Detected	Detected	Detected	Detected
21.	-30 / +0	Detected	Detected	Detected	Detected	Detected
22.	-30 / +15	Not Detected	Not Detected	Not Detected	Not Detected	Detected
23.	-30 / +30	Not Detected	Not Detected	Not Detected	Detected	Detected
24.	-30 / +45	Not Detected	Detected	Detected	Detected	Detected
25.	-30 / +60	Not Detected	Detected	Detected	Detected	Detected
26.	-30 / +75	Not Detected	Detected	Detected	Detected	Detected
27.	-30 / +90	Not Detected	Not Detected	Detected	Detected	Detected
28.	-15 / -90	Not Detected	Not Detected	Detected	Not Detected	Detected
29.	-15 / -75	Not Detected	Not Detected	Detected	Not Detected	Detected
30.	-15 / -60	Not Detected	Detected	Detected	Not Detected	Detected
31.	-15 / -45	Detected	Detected	Detected	Detected	Detected
32.	-15 / -30	Detected	Detected	Detected	Detected	Detected
33.	-15 / -15	Detected	Detected	Detected	Detected	Detected
34.	-15 / +0	Detected	Detected	Detected	Detected	Detected
35.	-15 / +15	Detected	Detected	Detected	Detected	Detected
36.	-15 / +30	Detected	Detected	Detected	Detected	Detected
37.	-15 / +45	Detected	Detected	Detected	Detected	Detected
38.	-15 / +60	Detected	Detected	Detected	Detected	Detected
39.	-15 / +75	Not Detected	Not Detected	Not Detected	Not Detected	Detected
40.	-15 / +90	Not Detected	Not Detected	Detected	Detected	Detected
41.	+0 / -90	Not Detected	Not Detected	Detected	Detected	Detected
42.	+0 / -75	Not Detected	Not Detected	Not Detected	Not Detected	Detected
43.	+0 / -60	Detected	Detected	Detected	Detected	Detected
44.	+0 / -45	Detected	Detected	Detected	Detected	Detected
45.	+0 / -30	Detected	Detected	Detected	Detected	Detected
46.	+0 / -15	Detected	Detected	Detected	Detected	Detected
47.	+0 / +0	Detected	Detected	Detected	Detected	Detected
48.	+0 / +15	Detected	Detected	Detected	Detected	Detected
49.	+0 / +30	Detected	Detected	Detected	Detected	Detected
50.	+0 / +45	Detected	Detected	Detected	Detected	Detected
51.	+0 / +60	Detected	Detected	Detected	Detected	Detected
52.	+0 / +75	Not Detected	Not Detected	Not Detected	Not Detected	Detected
53.	+0 / +90	Not Detected	Not	Detected	Detected	Detected

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

S. No.	Head Orientation (vertical / horizontal)	Viola Jones Algorithm	Mutneja et.al.	SVM based method	CNN based Algorithm	Proposed Algorithm
			Detected			
54.	+15 / -90	Not Detected	Not Detected	Not Detected	Detected	Detected
55.	+15 / -75	Not Detected	Detected	Detected	Not Detected	Detected
56.	+15 / -60	Not Detected	Not Detected	Not Detected	Not Detected	Detected
57.	+15 / -45	Detected	Detected	Detected	Detected	Detected
58.	+15 / -30	Detected	Detected	Detected	Detected	Detected
59.	+15 / -15	Detected	Detected	Detected	Detected	Detected
60.	+15 / +0	Detected	Detected	Detected	Detected	Detected
61.	+15 / +15	Detected	Detected	Detected	Detected	Detected
62.	+15 / +30	Detected	Detected	Detected	Detected	Detected
63.	+15 / +45	Detected	Detected	Detected	Detected	Detected
64.	+15 / +60	Detected	Detected	Detected	Detected	Detected
65.	+15 / +75	Detected	Detected	Detected	Detected	Detected
66.	+15 / +90	Not Detected	Not Detected	Detected	Detected	Detected
67.	+30 / -90	Not Detected	Detected	Detected	Detected	Detected
68.	+30 / -75	Not Detected	Detected	Detected	Detected	Detected
69.	+30 / -60	Not Detected	Not Detected	Not Detected	Detected	Detected
70.	+30 / -45	Detected	Detected	Detected	Detected	Detected
71.	+30 / -30	Detected	Detected	Detected	Detected	Detected
72.	+30 / -15	Detected	Detected	Detected	Detected	Detected
73.	+30 / +0	Detected	Detected	Detected	Detected	Detected
74.	+30 / +15	Detected	Detected	Detected	Detected	Detected
75.	+30 / +30	Detected	Detected	Detected	Detected	Detected
76.	+30 / +45	Detected	Detected	Detected	Detected	Detected
77.	+30 / +60	Detected	Detected	Detected	Detected	Detected
78.	+30 / +75	Not Detected	Not Detected	Detected	Detected	Detected
79.	+30 / +90	Not Detected	Not Detected	Detected	Detected	Detected
80.	+60 / -90	Not Detected	Detected	Not Detected	Detected	Detected
81.	+60 / -75	Not Detected	Detected	Detected	Detected	Detected
82.	+60 / -60	Not Detected	Detected	Detected	Detected	Detected
83.	+60 / -45	Not Detected	Not Detected	Detected	Not Detected	Detected
84.	+60 / -30	Not Detected	Not Detected	Detected	Not Detected	Detected
85.	+60 / -15	Detected	Detected	Detected	Detected	Detected
86.	+60 / +0	Detected	Detected	Detected	Detected	Detected
87.	+60 / +15	Detected	Detected	Detected	Detected	Detected
88.	+60 / +30	Not Detected	Not Detected	Not Detected	Not Detected	Detected
89.	+60 / +45	Not Detected	Detected	Not Detected	Detected	Detected

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

S. No.	Head Orientation (vertical / horizontal)	Viola Jones Algorithm	Mutneja et.al.	SVM based method	CNN based Algorithm	Proposed Algorithm
90.	+60 / +60	Not Detected	Not Detected	Not Detected	Detected	Detected
91.	+60 / +75	Not Detected	Not Detected	Not Detected	Detected	Detected
92.	+60 / +90	Detected	Detected	Detected	Detected	Detected
93.	+90 / +0	Detected	Detected	Detected	Detected	Detected

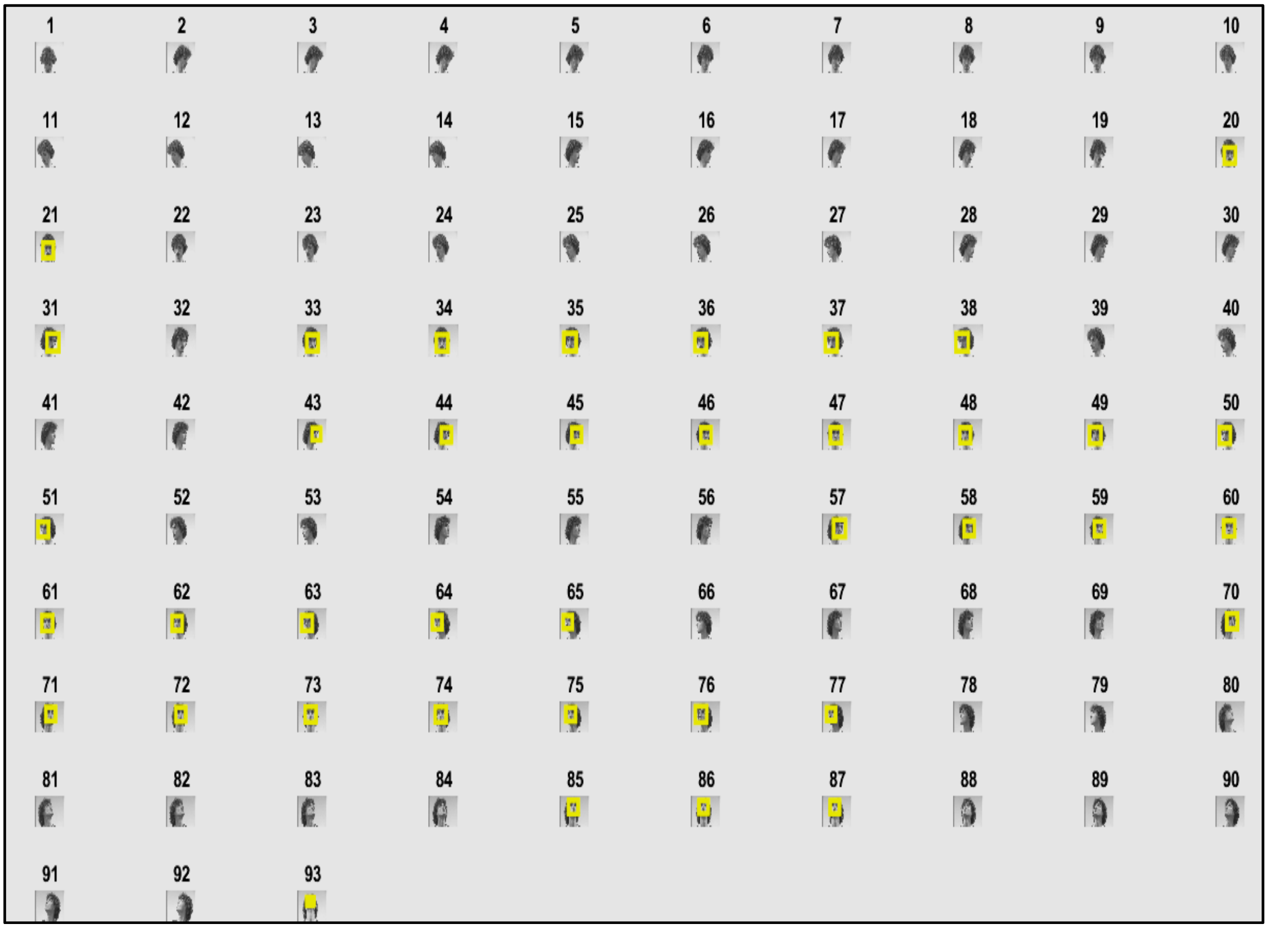

Figure 8. Viola Jones Algorithm tested with single Person image from Pointing’04 dataset

Face detection from the randomly selected images

Randomly 20 images from Pionting’04 dataset were selected and tested with Viola Jones algorithm and Figure 10 shows the results. According to the figure 10Viola Jones algorithm can be able to detect faces of only 12 images out of 20 images.

The same randomly 20 images from Pionting’04 dataset were selected and tested with our proposed ACF face detector and Figure 11 shows the results. According to figure 11 our ACF face detector can be able to detect faces of 20 images out of 20 images.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

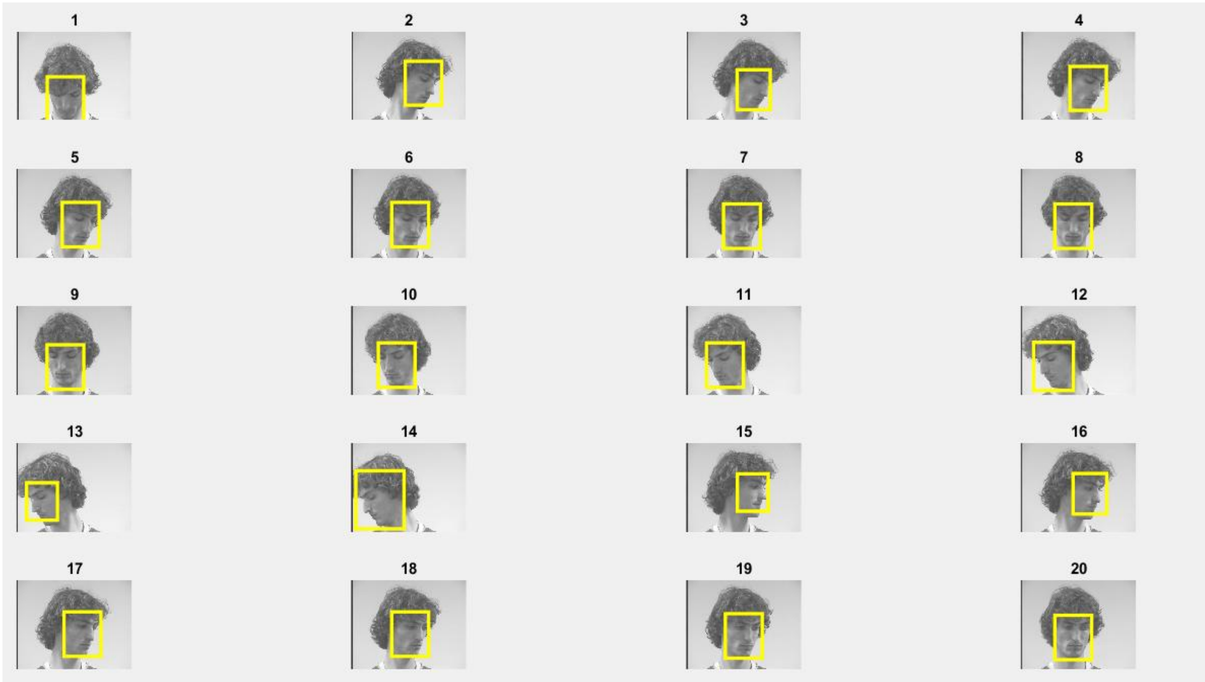

Figure 9. Proposed ACF detector tested with Single Person images from Pointing'04 dataset

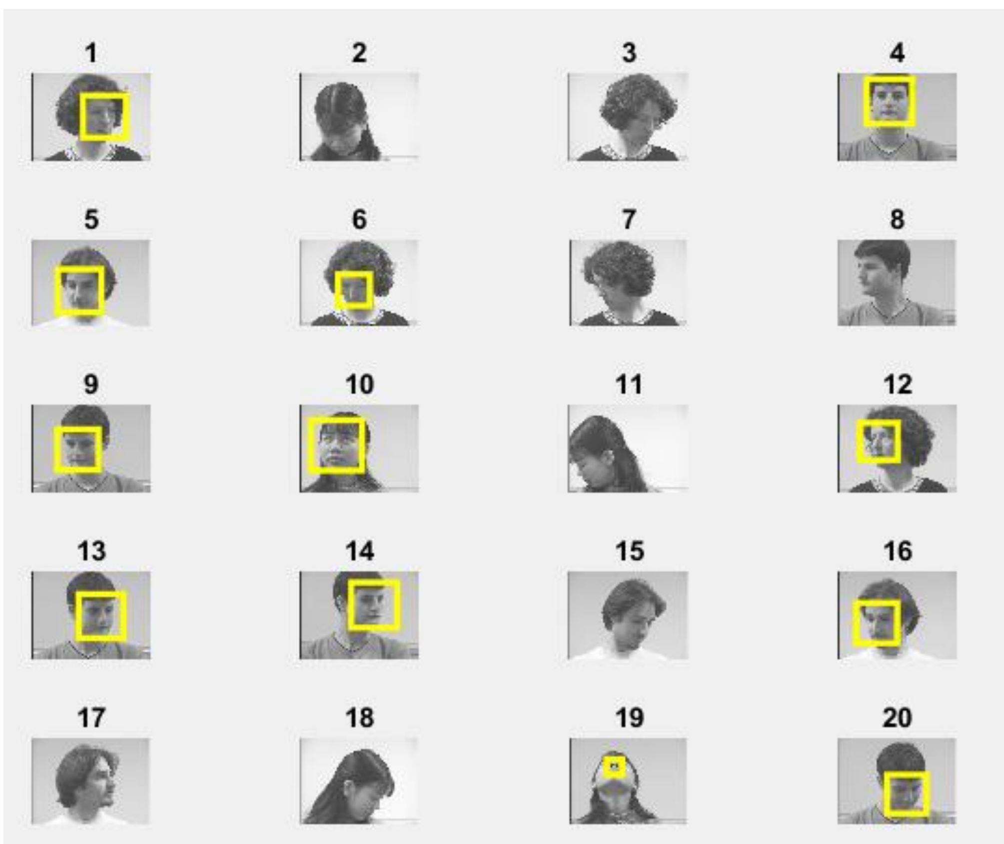

Figure 10. Viola Jones Algorithm tested with randomly selected images from Pointing'04 dataset

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

Figure 11. Proposed face detector tested with randomly selected images from Pointing'04 dataset

Face detection from the randomly selected imagesfrom the wild dataset

The proposed method was tested with the faces in a wild dataset which is available in [31]. The dataset has 30,281 faces from news photographs. Figures 12 and 13 show the result. Analysing the detection results, detection of the faces from the front pose image or head orientation images without the emotion was good. To get better performance from the low-resolution images our training dataset quality should be improved in our upcoming work.

Comparison with the earlier work

The training time and the detection time of our proposed work is compared with other works proposed by Viola Jones et.al., Vikram Mutneja el.al., SVM based method, CNN based Algorithm and our proposed method. From Table 3 our proposed method provides less training time and detection time.

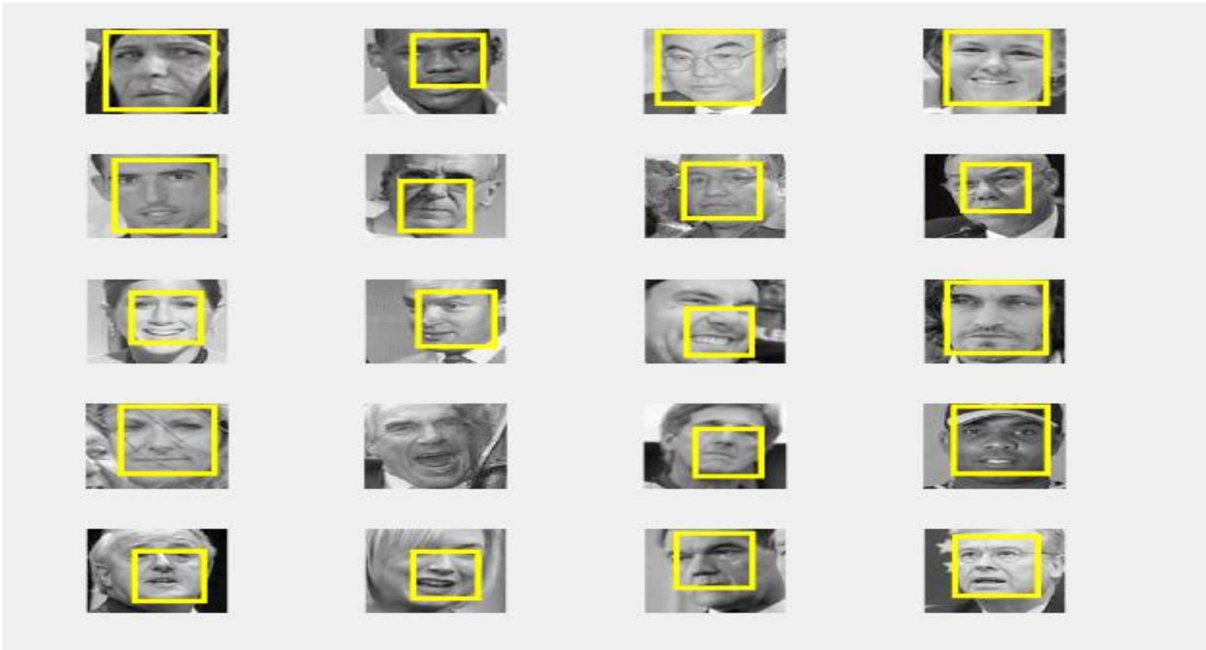

Figure 12. Performance of proposed method tested with randomly selected images from wild dataset

Figure 13. Performance of proposed method tested with randomly selected images from wild dataset.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

Table 3. Comparison with other work

Algorithms	No. of images	Training time	Detection Time(ms)
Viola Jones et.al.	9500	24 hours	67
Mutneja et.al.	6976	1.42 hours	39.69
SVM based method	3700	58 minutes	24.8
CNN based Algorithm	2500	2160 seconds	30.92
Our proposed method	840	447.30seconds	17.67

Conclusion

Our aim is to reduce the training time and achieve better face detection technique irrespective of the pose and head variation. We proposed an approach for detecting faces using the Aggregate Channel feature irrespective of pose and head orientation while achieving better average precision and log average miss rate. First contribution is extracting ACF and along with negative sample factor, number of stages and training time, quality measure was also included to tune the proposed ACF face detector. Second contribution of our work is a simple and efficient AdaBoost as a classifier. Cascading AdaBoost with ACF improved precision and log average miss rate. The proposed ACF face detector was implemented with Piointing’04 dataset and validated with FEI dataset. The face detection performance also compared with Viola Jones and other improved algorithms in MATLAB 2019a. Though our proposed method outperforms well for the Piointing’04 and FEI dataset, for faces with emotion and low-resolution images in a wild dataset is a challenge. To enhance the performance of our proposed method to handle low resolution and hard images, the quality of training images will be improved. Since face recognition is most important in surveillance and home safety applications. For recognising faces, it is very important to detect the face from the video or image irrespective of pose and head orientation angle. Our proposed ACF face detector can detect faces from -90 degree to +90 degree. Our future work will include face recognition and improving the quality of the training.

References

[1] Ming-Hsuan Yang, D. J. Kriegman and N. Ahuja. (Jan. 2002). Detecting faces in images: a survey.*IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(1), 34-58.doi: 10.1109/34.982883

[2] A.F. Abate, M. Nappi, D. Riccio, G. Sabatino. (2007). 2D and 3D face recognition: a survey. *Pattern Recognit. Lett.* 28(14), 1885–1906.

[3] Kurmi, Uma & Agrawal, Dheeraj &Baghel, Rishav. (2014). Study of Different Face Recognition Algorithms and Challenges. *International Journal of Engineering Research*, 3, 112-115. doi: 10.17950/ijer/v3s2/216

[4] Viola, P., Jones, M.J. Robust. (2004). Real-Time Face Detection. *International Journal of Computer Vision*, 57, 137–154. https://doi.org/10.1023/B:VISI.0000013087.49260.fb

[5] Mohammad Ashraful Islam, Md. Anin Naeem, Md. Nazmul Hasan. (May 2017). Comparison Between Viola-Jones andKlt Algorithms and Error Correction of Viola-Jones Algorithm.*International Journal of Computer Engineering and Applications*, 11(5).

[6] Zhu, Q., Avidan, S., Yeh, M. C., & Cheng, K. T. (2006). Fast human detection using a cascade of histograms of oriented gradients. *In Proceedings - IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1491-1498. https://doi.org/10.1109/CVPR.2006.119

[7] Ce Liu and Hueng-Yeung Shum. (2003). Kullback-leibler boosting. *In Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, USA*, 587–594.

[8] Peng Wang and Qiang Ji. (2005). Learning discriminant features for multi-view face and eye detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA*, 1, 373-379.doi: 10.1109/CVPR.2005.200

[9] S. Baluja, M. Sahami and H. A. Rowley. (2004) Efficient face orientation discrimination.*International Conference on Image Processing, Singapore*, 1, 589-592. doi:10.1109/ICIP.2004.1418823

[10] Yotam Abramson, Bruno Steux, and Hicham Ghorayeb. (June 2007). Yet Even Faster (YEF) real-time object detection. *Int. J. Intell. Syst. Technol. Appl.* 2, 2/3, 102–112. doi:https://doi.org/10.1504/IJISTA.2007.012476

[11] Matthew Turk and Alex Pentland. (Winter 1991). Eigenfaces for recognition. *J. Cognitive Neuroscience*, 3 (1), 71–86. doi:https://doi.org/10.1162/jocn.1991.3.1.71

[12] M. Nilsson, J. Nordberg and I. Claesson. (2007). Face Detection using Local SMQT Features and Split up Snow Classifier.*IEEE International Conference on Acoustics, Speech and Signal Processing - Honolulu*, II-589-II-592, doi: 10.1109/ICASSP.2007.366304

[13] K. Dang and S. Sharma. (2017). Review and comparison of face detection algorithms. *International Conference on Cloud Computing, Data Science & Engineering - Confluence, Noida*, 629-633.doi: 10.1109/CONFLUENCE.2017.7943228

[14] C. Huang, H. Ai, Y. Li and S. Lao. (April 2007). High-Performance Rotation Invariant Multiview Face Detection.*IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4), 671-686.doi: 10.1109/TPAMI.2007.1011

[15] Yongmin Li, Shaogang Gong, Jamie Sherrah, Heather Liddell. (2004). Support vector machine based multi-view face detection and recognition.*Image and Vision Computing*,22 (5), 413-427. https://doi.org/10.1016/j.imavis.2003.12.005

[16] R. Brunelli and T. Poggio. (1993). Face recognition: features versus templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*,15(10), 1042-1052.doi: 10.1109/34.254061

[17] Burl M.C., Weber M., Perona P. (1998). A probabilistic approach to object recognition using local photometry and global geometry. *In: Burkhardt H., Neumann B. (eds) Computer Vision — ECCV'98. Lecture Notes in Computer Science, 1407. Springer, Berlin, Heidelberg*. https://doi.org/10.1007/BFb0054769

[18] Chin-Chuan Han, Hong-Yuan Mark Liao, Gwo-Jong Yu, Liang-Hua Chen. (2000). Fast face detection via morphology-based pre-processing.*Pattern Recognition*, 33(10),1701-1712.https://doi.org/10.1016/S0031-3203(99)00141-7

[19] X. Zhao, X. Chai, Z. Niu, C. Heng and S. Shan. (2011). Context constrained facial landmark localization based on discontinuous Haar-like feature.*IEEE International Conference on Automatic Face & Gesture Recognition (FG), Santa Barbara, CA*, 673-678, doi: 10.1109/FG.2011.5771329

[20] Li S.Z., Zhu L., Zhang Z., Blake A., Zhang H., Shum H. (2002). Statistical Learning of Multi-view Face Detection. *In: Heyden A., Sparr G., Nielsen M., Johansen P. (eds) Computer Vision. Lecture Notes in Computer Science, 2353. Springer, Berlin, Heidelberg*.https://doi.org/10.1007/3-540-47979-1_5

[21] Ghimire, Deepak & Lee, Joonwhoan. (2013). A Robust Face Detection Method Based on Skin Color and Edges. *Journal of Information Processing Systems*, 9, 141-156. 10.3745/JIPS.2013.9.1.141

[22] Hu, Wu-Chih& Yang, Ching-Yu & Huang, Deng-Yuan & Huang, Chun-Hsiang. (2011). Feature-based Face Detection Against Skin-color Like Backgrounds with Varying Illumination. *Journal of Information Hiding and Multimedia Signal Processing, 2(2), 123-132.*

[23] Pritam, D., & Dewan, J. (2017). Detection of fire using image processing techniques with LUV color space. *2nd International Conference for Convergence in Technology (I2CT),* 1158-1162.

[24] J. Masek, R. Burget, V. Uher and S. Güney. (2013). Speeding up Viola-Jones algorithm using multi-Core GPU implementation.*36th International Conference on Telecommunications and Signal Processing (TSP), Rome*,808-812, doi: 10.1109/TSP.2013.6614050

[25] Shaily Pandey, Sandeep Sharma. (2015). An Optimistic Approach for Implementing Viola Jones Face Detection Algorithm in Database System and in Real Time.*International Journal of Engineering Research & Technology*, 04 (07). doi: http://dx.doi.org/10.17577/IJERTV4IS070758

[26] J. Zhu and Z. Chen. (2015). Real Time Face Detection System Using Adaboost and Haar-like Features. *2nd International Conference on Information Science and Control Engineering, Shanghai*, 404-407.doi: 10.1109/ICISCE.2015.95

[27] Egorov, A.D., Shtanko, A.N. &Minin, P.E. (2015). Selection of Viola–Jones algorithm parameters for specific conditions. *Bull. Lebedev Phys. Inst.,* 42, 244–248. https://doi.org/10.3103/S1068335615080060

[28] Mutneja, V., Singh, S. (2019). Modified Viola–Jones algorithm with GPU accelerated training and parallelized skin color filtering-based face detection. *J Real-Time Image Proc* 16, 1573–1593. https://doi.org/10.1007/s11554-017-0667-6

[29] Li-Fang Zhou,Yu Gu, Patrick S. P. Wang, Fa-Yuan Liu, Jie Liu, Tian-Yu Xu. (2020). Rotation-Invariant Face Detection with Multi-task Progressive Calibration Networks. *Pattern Recognition and Artificial Intelligence*, 513-524.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4368 - 4390
Received 25 April 2021; Accepted 08 May 2021.

[30] Li-Fang Zhou,Yu Gu, Shan Liang,Bang-Jun Lei,Jie Li. (2020). Direction-Sensitivity Features Ensemble Network for Rotation-Invariant Face Detection.*Pattern Recognition and Computer Vision*, 581-590.

[31] H. Wu, K. Zhang and G. Tian. (2018). Simultaneous Face Detection and Pose Estimation Using Convolutional Neural Network Cascade.*IEEE Access*, 6, 49563-49575.

[32] Bowen Yang, Chun Yang, Qi Liu, Xucheng Yin, and Xu-Cheng Yin. (2019). Joint Rotation-Invariance Face Detection and Alignment with Angle-Sensitivity Cascaded Networks. *In Proceedings of the 27th ACM International Conference on Multimedia (MM '19). Association for Computing Machinery, New York, NY, USA, 1473–1480*. doi:https://doi.org/10.1145/3343031.3350877

[33] https://www-prima.inrialpes.fr/Pointing04/data-face.html(dataset)

[34] https://fei.edu.br/~cet/facedatabase.html(dataset)

[35] http://tamaraberg.com/faceDataset/index.html(dataset)