# Face Descry System in Exams Committees


## Eman A.E. Ahmed

*Mathematics Department, College of Sciences, Imam Abdulrahman Bin Faisal University, Dammam, KSA.*



**ABSTRACT:** *The main objective of the proposed face descry system is to use the face in the exams committees as a password to authenticate the students during exams periods. In such places, more security is needed than only the existing of the person. Different applications for this work can be applied for example, authentication for the persons using their faces in front of ATM instead of using password and Person verification for self-serviced clearance using E-passport is another important typical application. All these application can be considered as cooperative user scenarios in which the user is willing to be cooperative by presenting his/her face in a proper way (for example, in a frontal pose with neutral expression and eyes open) in order to be granted the access or privilege. Face recognition has several advantages over other biometric modalities such as fingerprint and iris: besides being natural and nonintrusive, the most important advantage of face is that it can be captured at a distance and in a covert manner. In this paper, face recognition is used to face verification (or authentication) which involves a one-to-one match that compares a query face image against an enrollment face image whose identity is being claimed. The proposed system consists of four phases. The first phase is the pre-processing phase. The second phase is the face detection based on Viola-Jones method. The third phase is the proposed Moving Average Window conversion. The last phase is the face verification phase, in this phase the similarity between the query face image and the enrollment face image whose identity is being claimed is calculated and the decision is taken depending on two classification parameters (Structural Similarity Measure (SSIM) and Color Measurement Committee (CMC) distance. The experimental results were viewed and compared.*



**Keywords:** *Face Recognition, Moving Average Window, Structural Similarity Measure (SSIM), Viola-Jones method.*


## I. INTRODUCTION

Facial recognition (or face recognition) is a biometric method of identifying an individual by comparing live capture or digital image data with the stored record for that person. Facial recognition systems are commonly used for security purposes but are increasingly being used in a variety of other applications. In the registration phase in the collage an image must be taken for the student and stored in the faculty database to be used in the recognition phase of the descry system. The Processing Workflow of the proposed system is shown in **Fig. 1.**

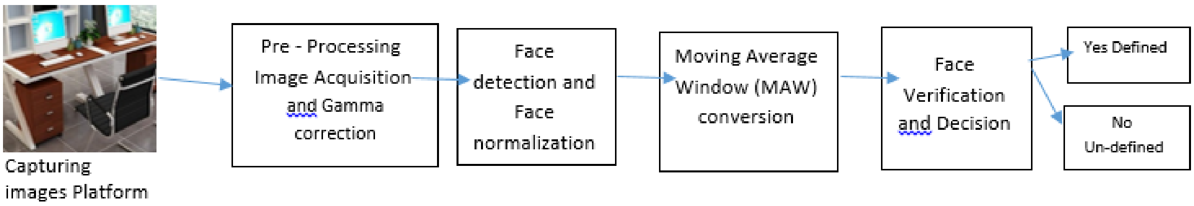



**Fig. 1:** A block diagram of face Descry system.

## II. BACKGROUND

Face recognition system is a computer vision and efficient software to automatically identify or verify a person from a digital image or a video frame from a video source. Face recognition system must include four main phases: preprocessing - face detection – features extraction – face classification. Each phase of this system has many techniques. Pre-processing is implemented to reduce the noise from the image and in this phase, the color images are acquired. Face detection phase is a necessary step in face recognition systems, with the purpose of localizing and extracting the face region from the background. **Fig. 2** gives different approaches associated with face detection technology [1].

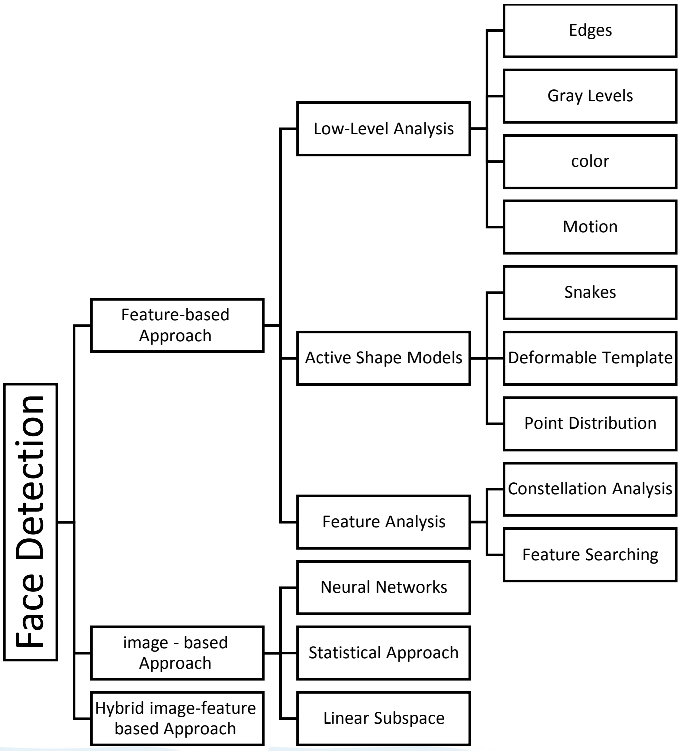



**Fig. 2:** Different approaches of face detection techniques.

**Feature extraction:** Three types of feature extraction methods can be distinguished: (1) generic methods based on edges, lines, and curves (global feature extraction); (2) feature-template-based methods that are used to detect facial features such as eyes (local feature extraction); (3) structural matching methods that take into consideration geometrical constraints on the features (local feature extraction). Early approaches focused on individual features. **Fig.3** shows Different approaches of face features extraction.

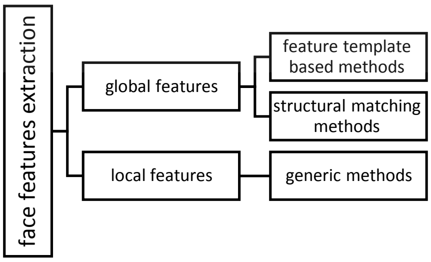



**Fig. 3:** Different approaches of face features extraction.

In global feature extraction methods, feature vectors from the whole face region are extracted. Principal component analysis (PCA) is the most widespread method for global feature extraction and is used for feature extraction from 2D face images and also from range images. Other popular global feature extraction methods, such as linear discriminant analysis (LDA) and independent component analysis (ICA) are also used on range images. The use of global features is prevalent in face recognition systems based on images acquired in a controlled environment. In Local feature extraction methods extract a set of feature vectors from a face, where each vector holds the characteristics of a particular facial region. The local features extraction methods have advantages over the global features in uncontrolled environments, where the variations in facial illumination, rotation, expressions and scale are present. The process of local feature extraction can be divided into two parts. In the first part, the interest points on the face region are detected. In the second part, the points of interest are used as locations at which the local feature vectors are calculated.

**Face classification (recognition):** All existing face recognition techniques can be classified into five types based on the way they identify the face. **Fig (4-a)** shows different approaches associated with face classification methods. Another classification based on the nature of capturing the face image can be shown in **Fig (4-b).**

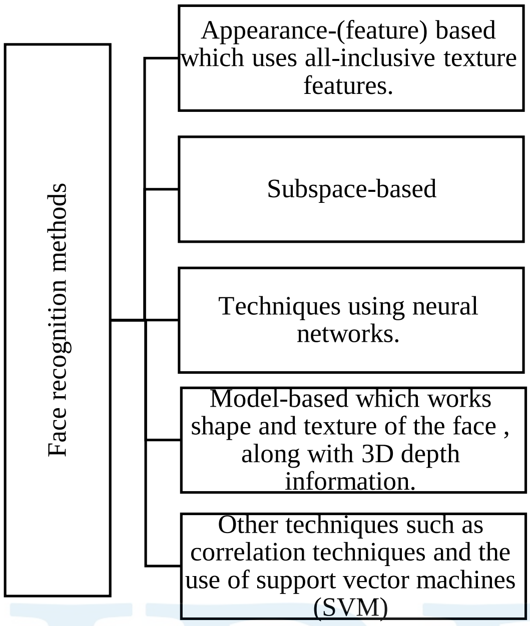



**Fig (4-a):** shows different approaches associated with face classification methods.

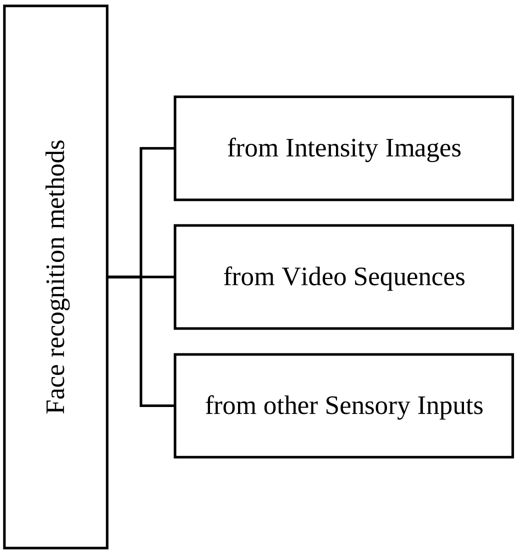



**Fig (4-b):** shows face classification methods based on the nature of capturing the face image.

## III. RELATED WORK

There are two families of face detection algorithms as mentioned in [1]: The first family based on rigid-templates and included:

– The Viola-Jones face detection algorithm and its variations [2].

– Algorithms that are based on Convolutional Neural Networks (CNNs) and Deep CNNs (DCNNs) such as [3]; recently, a DCNN showed exceptional performance in multi-object class detection [4], thus currently its learning architectures have been investigated for face detection [5].

– Methods such as [6], which apply strategies inspired by image-retrieval and Generalized Hough Transform [7].

The second family of algorithms that learn and apply a Deformable Parts-based Model (DPM) [8, 9] to model a potential deformation between facial parts. Other notable parts-based methods include [10, 11].

For face Features extraction phase, The introduction of robust feature extraction methodologies, such as Scale Invariant Feature Transform (SIFT) features [12], Histograms of oriented Gradients (HoGs) [13], Local Binary Patterns (LBPs) and their variations[14,15], their fast counterparts such as Speeded Up Robust Features (SURF) [16] and DAISY [17], as well as transformations that combine the above features with integral images, such as Integral Channel Features (ICF) [18]. These features are densely or sparsely sampled and are used to describe face appearance.

For face recognition phase the techniques can be broadly divided into three categories : methods that operate on intensity images, those that deal with video sequences, and those that require other sensory data such as 3D information or infra-red imagery[19]). Face recognition methods for intensity images fall into two main categories: feature-based and holistic [20].

For video-based face recognition system Howell and Buxton [21] employed a two-layer RBF network [22] for learning/training and used Difference of Gaussian (DoG) filtering and Gabor wavelet analysis for the feature representation, while the scheme from [23] was utilized for face detection and tracking. A detailed survey of recent schemes for face recognition from video sequences is provided in [24]. For Face Recognition from Other Sensory Inputs : for the 3D Model-based applications the Comprehensive recent surveys of literature on 3D face recognition can be found in [25] and for Infra-red application A comprehensive review of recent advances in face recognition from infra-red imagery may be found in [26].

## IV. PROPOSED METHOD

The proposed face descry system consists of four separate parts:

***a)*** **Pre processing phase** includes two steps the first is image acquisition: in image acquisition all images have been taken using Specific Platform as described in **Fig. 1**. If the student sit in wrong angles so, the system MUST ask him/her to sit in right pose with 0 vertical 0 horizontal. The second step in this phase is applying gamma correction on the capturing frame to increase the dynamic range and this means, increase the contrast in the image because both low lighting and high lighting conditions tend to conceal features of the face. Gamma correction is optimal to increase the contrast in the image.

***b)*** **Face detection using the Viola-Jones Haar Cascade Method.**

In this paper face detection using Viola-Jones Haar Cascade Method was applied to detect the face.Viola-Jones Face Detection is used using Open CV - Python environment. The Viola-Jones face detector contains three main ideas (the integral image, classifier learning with AdaBoost, and the attentional cascade structure) that make it possible to build a successful face detector that can run in real-time. At a high level, the method scans an image with a window looking for features of a human face. If enough of these features are found, then this particular window of the image is said to be a face [27-29].

After face detection is applied the second step in this phase is Face normalization this step is necessary because state-of-the-art recognition methods are expected to recognize face images with varying pose and illumination. The geometrical normalization process transforms the face into a standard frame by face cropping. The photometric normalization process normalizes the face based on such gray scale.

***c)*** **Moving Average Window (MAW) conversion:**

Construct the moving average image by applying Moving Average Window (MAW) of size 3*3 for both the template and the training images as described in Equation (1).

$$Z(i,j) = \frac{1}{9}M(i,j) \qquad Equation(1)$$

Where M(i,j) is the pixels of the image and Z(i,j) is the corresponding moving average image.

***d)*** **Face Verification and decision:**

I used two classification parameters in this work:

The first classification parameter is Structural Similarity Measure (SSIM) which is calculated between the moving average of the template and the moving average of the training images according to Equation (2). SSIM attempts to separate the task of similarity measurement of two images into Luminance, contrast and structure [30]. SSIM is designed to improve on traditional methods like peak signal-to-noise ratio (PSNR) and mean squared error (MSE), which have proven to be inconsistent with human eye perception.

$$SSIM(P1,P2) = \frac{(2\mu_{p1}\mu_{p2} + C_1)(\sigma_{p1p2} + C_2)}{(\mu_{p1}^2 + \mu_{p2}^2 C_1)(\sigma_{p1}^2 + \sigma_{p2}^2 + C_2)} \qquad \text{Equation( 2 )}$$

Where P1 and P2 are two M $\times$ N images being compared (one of them is the template image), and $\sigma$ are mean and standard deviation of the corresponding images as shown in Equations (3-5) and C1, C2 and C3 are constants used for stability of equations when $\mu$ and $\sigma$ are extremely small. SSIM defines, $\sigma$p1, $\sigma$p2 and $\sigma$p1p2 as the following:

$$\mu_{P1} = \frac{1}{M \times N}\sum_{y=0}^{M-1}\sum_{x=0}^{N-1} P_1(x,y) \qquad Equation\ (3)$$

$$\sigma_{p1} = \sqrt{\frac{1}{(M \times N - 1)}\sum_{y=0}^{M-1}\sum_{x=0}^{N-1}\left(P_1(x,y) - \mu_{p1}\right)^2} \qquad Equation\ (4)$$

$$\sigma_{p1}\sigma_{p2} = \frac{1}{M \times N - 1}\sum_{y=0}^{M-1}\sum_{x=0}^{N-1}\left(P_1(x,y) - \mu_{p1}\right)\left(P_2(x,y) - \mu_{p2}\right) \qquad Equation(5)$$

Mean SSIM can be calculated by dividing SSIM by number of pixels in image. I Construct the moving average images for both the training and the template images then calculate the Mean SSIM between each template / train images to get the Mean SSIM THRESHOLD. The Mean SSIM THRESHOLD value= the minimum Mean SSIM value obtained between the template and the training images.

The second classification parameter is the Color Measurement Committee (CMC) distance, which is used to calculate the color difference between pixels as this color metric gives good results with CIELab color space [31]. CMC is one of most accurate metrics according to human perception. The mathematical equation for calculating CMC can be found in [32].
The Mean SSIM THRESHOLD value= the maximum Mean SSIM value obtained between the template and the training images. To decide if the student is Positive or Negative calculate SSIM and CMC between each template / test images and compare it with the MEAN SSIM THRESHOLD to give the decision.

If the calculated Mean SSIM > = Mean SSIM THRESHOLD so, POSITIVE PERSON
Else If the calculated Mean SSIM <= Mean SSIM THRESHOLD so, NEGATIVE PERSON
If the calculated CMC <= CMC THRESHOLD so, POSITIVE PERSON
Else If the calculated CMC > CMC THRESHOLD so, NEGATIVE PERSON

## V. EXPERIMENTAL RESULTS

Database: In this work, I used two sets of databases. The first one is The Specs On Face ( SoF ) dataset [33] which was assembled to support testing and evaluation of face detection, recognition, and classification algorithms using standardized tests and procedures and I chose 21 images for 12 subject with the same background from this database. I captured the second set of database that included 21 images for 12 subject with the same pose and different backgrounds but for faces without glasses. This means my complete database include 21 * 24 images.
In this paper, I divided the images for each person into three categories for each person:

a) The first group is the stored template image for the person in database. The template image for each student is the image, which is received and saved in the university database for this student in the registration process.
b) The second group is the training group (10 images) were chosen randomly from the database. This group is used to get the Mean SSIM THRESHOLD and the CMC THRESHOLD values.
c) The third group is the test group (the rest 10 images for the same person + 10*23 images for the other persons in the selected database). This group is used by the proposed system to test the performance as described in the classification phase.

Step (V-b) and step (V-c) are executed after and before applying moving average window on the template and the test images. The performance in this work is calculated and compared based on some important statistics:

- **True Positive (TP)** means accept the right person.
- **False positive (FP)** means reject the right person.
- **True Negative (TN)** means reject the wrong person.
- **False Negative (FN)** means accept the wrong person.

**Sensitivity:** The sensitivity tells us how likely the test is come back positive in someone who has the characteristic. This is calculated as TP/(TP+FN).

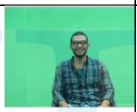
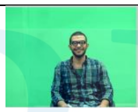
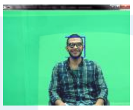
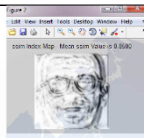
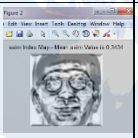
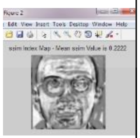
**Specificity:** The specificity tells us how likely the test is to come back negative in someone who does not have the characteristic. This is calculated as TN/(TN+FP).

**Accuracy:** (TP+TN)/(TP+FP+TN+FN)
An ideal system would have very high accuracy.
Experiments shows that the proposed MAW +SSIM approach and the proposed MAW+ CMC techniques give better performance than SSIM and CMC without using the proposed MAW method. **TABLE 1** reports the best performance of different methods. **TABLES 1 and 2** show an example of the experimental results and display the performance of the proposed method.

**TABLE 1:** Performance of the proposed method.

| | The Accuracy without using MAW | The Accuracy without using MAW | The Accuracy for the proposed MAW approach | The Accuracy for the proposed MAW approach |
|---|---|---|---|---|
| Threshold Value | Mean SSIM =0.75 | CMC=8.2 | Mean SSIM=0.8 | CMC=4.17 |
| TP | 8 | 9 | 10 | 9 |
| TN | 210 | 212 | 228 | 227 |
| FP | 2 | 1 | 0 | 1 |
| FN | 20 | 18 | 2 | 3 |
| Sensitivity | 8/28 | 9/27 | 10/12 | 9/12 |
| Specificity | 210/212 | 212/213 | 100% | 227/228 |
| Accuracy | 218/240 | 221/240 | 238/240 | 236/240 |

TABLE 2: An example on the proposed MAW method.

| | Capturing image | Gamma Correction | Face Detection | MAW image | Mean SSIM | CMC |
|---|---|---|---|---|---|---|
| Template Image |  |  |  |  | | |
| Test Image#1 |  |  |  |  |  0.8500 | 3.87 |
| Test Image#2 |  |  |  |  |  0.3434 | 10.54 |
| Test Image#3 |  |  |  |  |  0.2222 | 32.92 |

## VI. CONCLUSION AND FUTURE WORK:

In this paper, a simple new approach was introduced for the Face Descry System in Exams Committees. SSIM and CMC values were compared after and before applying the proposed MAW method. The results showed that the proposed system is excellent especially when the tested image was for a different person. The drawback was that this system require a special environment for capturing the image and this will be a challenge for a future work.

## REFERENCES:

[1] S. Zafeiriou, C. Zhang, Z. Zhang, A Survey on Face Detection in the wild: past, present and future, Computer Vision and Image Understanding (2015)

[2] P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, in: Proc. of CVPR, 2001.

[3] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, 2012, pp. 1097–1105.

[4] R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, arXiv preprint arXiv:1311.2524.

[5] C. Zhang, Z. Zhang, Improving multiview face detection with multi-task deep convolutional neural networks, in: Applications of ComputerVision (WACV), 2014 IEEE Winter Conference on, IEEE, 2014, pp. 1036–1041.

[6] H. Li, Z. Lin, J. Brandt, X. Shen, G. Hua, Efficient boosted exemplar-based face detection, in: Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, IEEE, 2014.

[7] B. Leibe, A. Leonardis, B. Schiele, Robust object detection with interleaved categorization and segmentation, International journal ofcomputer vision 77 (1-3) (2008) 259–289.

[8] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, D. Ramanan, Object detection with discriminatively trained part-based models, PatternAnalysis and Machine Intelligence, IEEE Transactions on 32 (9) (2010) 1627–1645.
[9] P. F. Felzenszwalb, D. P. Huttenlocher, Pictorial structures for object recognition, International Journal of Computer Vision 61 (1) (2005) 55–79.
[14] M. A. Fischler, R. A. Elschlager, The representation and matching of pictorial structures, IEEE Transactions on Computers 22 (1) (1973) 67–92.
[10] H. Schneiderman, T. Kanade, Object detection using the statistics of parts, International Journal of Computer Vision 56 (3) (2004) 151–177.
[11] K. Mikolajczyk, C. Schmid, A. Zisserman, Human detection based on a probabilistic assembly of robust part detectors, in: Proc. of ECCV, 2004.
[12] D. G. Lowe, Distinctive image features from scale-invariant keypoints, International journal of computer vision 60 (2) (2004) 91–110.
[13] N. Dalal, B. Triggs, Histogram of oriented gradients for human detection, in: Proc. of CVPR, 2005.
[14] T. Ojala, M. Pietik¨ainen, T.M¨aenp¨a¨a, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Trans. on PAMI 24 (2002) 971–987.
[15] T. Ahonen, A. Hadid, M. Pietik¨ainen, Face recognition with local binary patterns, in: Proc. of ECCV, 2004.
[16] H. Bay, A. Ess, T. Tuytelaars, L. Van Gool, Speeded-up robust features (surf), Computer vision and image understanding 110 (3) (2008) 346–359.
[17] E. Tola, V. Lepetit, P. Fua, Daisy: An efficient dense descriptor applied to wide-baseline stereo, Pattern Analysis and Machine Intelligence,
IEEE Transactions on 32 (5) (2010) 815–830.
[18] P. Dollar, Z. Tu, P. Perona, S. Belongie, Integral channel features., in: BMVC, Vol. 2, 2009, p. 5.
[19] W. Zhao, R. Chellappa, P. Phillips, and A. Rosenfeld, "Face Recognition: A Literature Survey," ACM Computing Surveys , Vol.35, pp.399-458, 2003.
[20] R. Brunelli and T. Poggio, "Face recognition: features versus templates," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.15,pp.1042-1052, 1993.
[21] A. Howell and H. Buxton, "Towards unconstrained face recognition from image sequences," in Proceedings of the Second IEEE International Conference on Automatic Face and Gesture Recognition, 1996, pp.224-229.
[22] J. Moody and C. Darken, "Learning with localized receptive fields," in Proceedings of the 1988 Connectionist Models Summer School, D. Touretzky, G. Hinton, and T. Sejnowski, Eds.: Morgan Kaufmann, 1988, pp.133-143.
[23] S. McKenna and S. Gong, "Combined motion and model-based face tracking," in Proceedings of British Machine Vision Conference. Edinburgh, UK, 1996, pp.755-765.
[24] S. Zhou and R. Chellappa, "Beyond a single still image: Face recognition from multiple still images and videos," in Face Processing: Advanced Modeling and Methods: Academic Press, 2005.
[25] A. Ruifrok, A. Scheenstra, and R. C. Veltkamp, "A Survey of 3D Face Recognition Methods," in Audio- and Video-based Biometric Person Authentication, Vol.3546, Lecture Notes in Computer Science: Springer Berlin / Heidelberg, 2005, pp.891-899.
[26] S. G. Kong, J. Heo, B. R. Abidi, J. Paik, and M. A. Abidi, "Recent advances in visual and infrared face recognition - a review," Computer Vision And Image Understanding, Vol.97, pp.103-135, 2005.
[27] Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In:
Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Kauai, Hawaii (2001)
[28] Viola, P.A., Jones, M.J.: Fast and robust classification using asymmetric AdaBoost and a detector cascade. In: Advances in Neural Information Processing Systems, vol. 14, pp. 1311–1318 (2001)
[29] Viola, P., Jones, M.J.: Robust real-time face detection. Int. J. Comput. Vis. 57(2), 137–154 (2004)
[30] Wang, Z., Bovik, A., Sheik, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Process 4, 13(4), 1–14 (2004).
[31] "Color Image Segmentation Using CIELab Color Space Using Ant Colony Optimization "Seema Bansal et al IJCSET |August 2011 | Vol 1, Issue 7, 415-420
[32] https://en.wikipedia.org/wiki/Color_difference
[33] https://sites.google.com/view/sof-dataset