Face recognition and detection using Random forest and combination of LBP and HOG features

Conference Paper · July 2018

DOI: 10.1109/ICSCEE.2018.8538377

2 authors:

Huda Mady

10 PUBLICATIONS **145** CITATIONS

SEE PROFILE

Shadi M S Hilles

Istanbul Okan University

39 PUBLICATIONS **372** CITATIONS

SEE PROFILE

Face recognition and detection using Random forest and combination of LBP and HOG features

Huda Mady
Dept. Computer Science
Almadinah International University
Kuala lumpur, Malaysia
Huda.lib22@gmail.com

Shadi M. S. Hilles
Dept. Computer Science
Almadinah International University
Kuala lumpur Malaysia
shadihilless@gmail.com

***Abstract*— the effective facial recognition method should perform well in unregulated environments based on video broadcast to satisfy the demands of applications in real-world However, this still remains a big challenge for most current face recognition algorithms that will affect the accuracy of the system. This study was conducted to develop face recognition method based on video broadcast under illumination variation, facial expressions, different pose, orientation, occlusion, nationality variation and motion. Viola-Jones algorithm was applied to improve face detection which is these method have proven to detect the faces in an uncontrolled environment in the real world simply and high accuracy. A combination of Histograms of Oriented Gradients (HOG) and Local Binary Pattern (LBP) descriptors was conducted for faces features extraction purpose. These descriptors have proven to be lower computational time. The latest and accurate technique was applied for face classification based on Random Forest classifier (RF). To evaluate the efficiency of the Random Forest classifier, compared it with Support Vector Machine classifiers (SVM) is done with different existing feature extraction methods. Four experiments were implemented on Mediu staff database and excellent results have reported the efficiency of proposed algorithm average recognition accuracy 97.6% The Computer Vision and Image Processing MAT LAB 2016b Toolboxes was used for coding the desired system, dataset based on videos.**

Keywords—Mediu Staff, Face Recognition, Viola &Jones, Random Forest classifier (RF), Local Binary Pattern (LBP), and Histograms of Oriented Gradients (HOG).

I. INTRODUCTION

Face recognition technique has received considerable attention in recent year, due to the increasing numerous of commercial and law enforcement applications requesting authoritative personal recognition (e.g. attendance management system, access control, surveillance of people in any public places) Face recognition is one of the most vital biometric identification techniques Compared with other biometric identification techniques due to it is easier in acquiring data, natural and non-intrusive[1].

The effectiveness of a face recognition method is measured by its accuracy for identifying authorized individuals and rejecting unauthorized ones. Key performance indicators include the feature extraction and classification stage. In fact, several methods have been proposed to develop of automatic face recognition in uncontrolled environment such as variations in pose, illumination, occlusion, motion and facial expression, including the Principal Component Analysis, Discriminant Analysis, Local Binary Patterns (LBP), neural networks, Elastic Template Matching, and Gabor Filtering .[2].

As one of the most successful local method, Local Binary Pattern (LBP) has been used within the face recognition task over the last years. LBP computes a histogram by taking into account each pixel in the image and considering the values of its neighborhoods [3]. it is Robust against rotation, scale, expressions, illumination variations, Pose and occlusion variations [4]&[1] &[5]& [6]. Another wide and powerful local face descriptor which are Gabor kernels. Despite the success of this technique the great number of kernels to be applied to the image leads to high feature dimensions and makes it thus high computational and unusable in real-time applications [4]. Proven the histogram oriented gradients descriptor (HOG) was used in different areas of image processing as human detection and hand gesture recognition.

Very few publications apply his descriptors to face recognition while it achieves comparable results with other powerful descriptors even better than some of them. Furthermore, HOG has demonstrated to be much lower-complexity in term of computing time when compared to its competitors, which allowed it to be used in many areas of image processing as real-time applications.[7]& [8] & [2] Has been confirmed that the combined performance of POEM descriptor and MBC descriptor gives better results. But, this technique has been evaluated only on the image database which is a FERET datasets and a Wild (LFW) dataset.

Reported combination between [3]LBP and MHOG to extract the weighted from faces of images blocks and WSRC classification are applied to minimize pose and illumination variation. To solve challenging of face recognition under different pose and illumination for face images, the proposed method was evaluated by on the ORL, Yale, and CMU-PIE databases. The experimental results proved accuracy effectively in illumination and pose variable face recognition.

The second key develop of face recognition method is an accuracy and performance of the classifier. Several machine learning techniques have been developed to classify facial in historic based on statistic and artificial intelligence, such as NN [9], support vector machine (SVM) , k-nearest neighbor [10], rule-based classifiers and random forest (RF) classifier [11]. RF is a technique widely used for face classification in recent years. Random Forest (RF) is an ensemble of decision trees, where each tree provides a classification, and the final result is

the majority vote given by all of the trees[12]. It has high discrimination performance, training classifying and higher recognition rate with lower computational cost [13]. RF demonstrated to be better results compared with those of the most well-known classifiers concerning accuracy in several tasks, as well as face recognition tasks.

Although different approaches have achieved great results under constrained face images but under uncontrolled environment based of video broadcast are facing a great number of challenges such as variation of illumination, pose, expression and occlusion or orientation changes, etc. these challenges major drawback poor performance of face recognition rate [1] & [14].

This paper focuses on constructing a new framework approach to improve face recognition and capable to solve challenges of face recognition under uncontrolled environment based of video broadcast. The proposed method reported combination between Histograms of Oriented Gradients and Local Binary Pattern for extracted features and the Random Forest for classification using real-life collected video database which called Mediu staff [15].The proposed approach can be implemented in a real-time face recognition method under uncontrolled environment.

The main contributions of this paper are summarized as follows new approach framework for face recognition under uncontrolled conditions is constructed by using:

- Fast and real-time face detection is reported by the Viola-Jones algorithm based of four classifier.
- Feature extraction technique is developed by Combining robust local binary patterns and Histograms of Oriented Gradients
- Latest and accurate face classification is applied by Random Forest to face recognition.
- The Comparison of two most popular classifier which are random forest and support vector machine with different feature extraction methods on Mediu staff database are utilized for evaluate the suggested method.
- A new formula is introduced into the final process of matching result to achieve the accuracy of face recognition which is a Voting stage.

The following section 2 of the paper Proposed method which explains the Mediu staff database, Face detection HOG and LBP descriptors in details and describe them, random forest classifier and new proposed formula which is Voting stag are explained. Experimental results are presented in section 3, in section 4 discussion of result, and conclusions are given in the final section.

II. PROPOSED FRAMEWORK

Face recognition is the process of identifying or verifying one or more persons appearing in a scene by using Samples from Mediu-S-DB in uncontrolled environments.as showed in figure1 Proposed approach framework is shown in figure 2 and 3, the steps of face recognition in video are as follows A to E stage.

A. Mediu staff database (Mediu-S-DB)

Reported collected Mediu-S-DB of video sequences for studying [15] the problem of unconstrained face classification based on videos broadcast and evaluating the proposed face recognition method. This database contains 22 broadcast videos for ten males from different races which is used to extract and store the templates of enrolled users. This module is also responsible for enrolling users in the evaluate face recognition method. In figure4,

To create the training database, 100 images per subject with varied illumination, pose, and background, expressions, and occlusions are extracted for the training videos. The training database is created by cropping off the labeled faces in Mediu-S-DB after performing the face detection, keeping only the face and most of the background is eliminated. These faces are then preprocessing and converted to grayscale, resized and normalized. Face detection process is accomplished based on the Viola Jones approach based on four classifiers.

Testing data of videos contain 12 videos, divided to three groups, first group contain 10 videos to test the system performance in case there is one person in the video, the second group to test the developed algorithm in case there is more than one person in the video and the third group in case the person is not included in the training database. The algorithm selects from each video 30 frames randomly to present all the duration of the video broadcast.

B. Face detection and Preprocessing image

Face detection is techniques for locating faces in images and video sequences. This study proposed viola&Jons algorithm based on four classifiers to improve face detection process from images and video sequences under unsteadied environment. The viola&Jons process follows by an acquired face frame is scanned firstly to detect a presence of faces and locating their exact location and size. The output of face detection is an image window containing only the face area. Irrelevant information, such as background, hair, neck, and shoulders, ears, etc. After face detect, cropped windows of face, the face of image is normalized by histogram equalization method. The histogram equalization methods are defined to be as a known method to make extraction of data from detect face correctly. Consequently, it is used for the gray-scale transform and as the illumination preprocessing. The main advantage of this method is that it works directly and fairly and use the invertible operator. Therefore, the histogram works on Equalization Function is very well known and it is easy can be managed to recover and the calculation effect is not computationally intensive in the process function[16]

C. Features extraction approach

Combination two local features approach used by the LBP and HOG method to extract features from face classifier to identify or verify the identity of an unknown face.

LBP Local Binary Pattern (LBP) is a local feature widely utilized for texture description in pattern recognition. The original LBP assign a label to each pixel. Each pixel is used as a threshold and compared with its neighbor. If the neighbor is higher it takes the value 1 otherwise it takes the value 0. Finally the thresholder neighbor pixel values are concatenated and considered as a binary number that becomes the label for the central pixel. A graphical representation is showed in Fig 3 (a) and equation 1. Where g_c is the luminance value of the central pixel, g_p is the luminance value of its neighbors, P is the total number of involved neighbors, and R is the radius of the neighborhood in proposed method case R=1. To account for the spatial information the image is then divided in sub-regions. LBP is applied to each sub-region and a histogram of L bins is generated from the pixel labels. Then the histograms different regions are concatenated in a single higher dimensional histogram as represented in Fig 3(b).

$$LBP = \sum_{p=0}^{p=15} s(g_c - g_p)\ 2^p \quad (1)$$

- Histogram of Oriented Gradient (HOG) was suggested [17] to detect objects. The HOG descriptor is robust and is not sensitive to both illumination and geometric variations, and the computational complexity of a HOG feature is much less than that of the original data [18]. The main steps of HOG feature extraction algorithm are as follows:
- Initially, the [144* 144]-pixel window is divided into small regions called cells, by [8*8] pixel cells, as shown in Fig. 6. The luminance gradient is calculated for each pixel (x, y) in equation 3
- A block of [17*17] pixels is created of [2*2] cells with one cell overlapping. Pixel's orientation in the histogram cells is computed by equation 2,
- Hence, features vector is constructed by all orientation histogram by equation 3.

$$M(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \quad (2)$$

$$\theta(x,y) = \tan^{-1}\left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}\right) \quad (3)$$

Where: The luminance gradient is a vector with magnitude M and orientation θ illustrated by the variation in the luminance, and L: is the Luminance value of the pixel. Finally, the normalization process is performed using equation 4 to create HOG feature vector.

$$F(n) = \frac{F(n)}{\sqrt{\left(\sum_{k=1}^{2\times 2\times 9} F(k)^2\right) + 1}} \quad (4)$$

Where, F is a magnitude of each orientation, (2× 2) represents the number of cells in each image and No. 9 represent the number of orientation bins in each cell. Figure illustrates HOG for the original image; figure 4 present the construction of HOG feature vector.

D. RF Classification module

Random Forest RF is a technology that represents a set of an ensemble learning methods for random classification that functioning by making decisions using a multitude of trees votes and predicting the features of data as follow: classification results from each tree are collected for the input image, after that, the majority voting is gathered to give the resulting class label. Random forest (RF) is artificial intelligence technique and the strong modern method to a classification of data and modeling. RF has been applied to compare extracted template (vectors features) from both training and testing stage to match the corresponding person. Suitable features vector that able to characterize, as much as possible is highly recommended for the classification process. In which the extracted vectors features during step 2 from testing data is compared against the stored vectors in training data to generate matching results

E. Voting stag

Voting stage is new formula which is proposed for increase the accurate of verification of face, the matching result from step four is compared to a predefined threshold and based on the result of this comparison. In case of identifying the face, and the face is accepted, a set of matching results between the extracted vectors features and the features of enrolled users is calculated. If the features of user S1 provided the best result, then the unknown face is more similar to S1, than any other face of the person in the dataset. To ensure that the unknown face is actually S1 and not an impostor, the matching result is compared to a predefined threshold. Voting stag method procedures:

- The algorithm is captured 30 frames from each testing video randomly, detect the face from each frame then create the features vector.
- The trained random forest model classifies the extracted features to identify the corresponding person from the database and the algorithm stores the result for each classified person in specific counter.
- Step 2 is repeated for all detected 30 frames.
- The counters of all person in the database are divided by the number of the detected face in step 1 then multiplied by 100 to get the percentage of each person.
- The person who gets the percentage >= 60% represents the person in the testing video.
- To calculate the average accuracy; from step 5, the summation of percentages are calculated and divided by the number of testing videos.

Fig. 1. *Samples from Mediu-S-DB under uncontrolled environments*

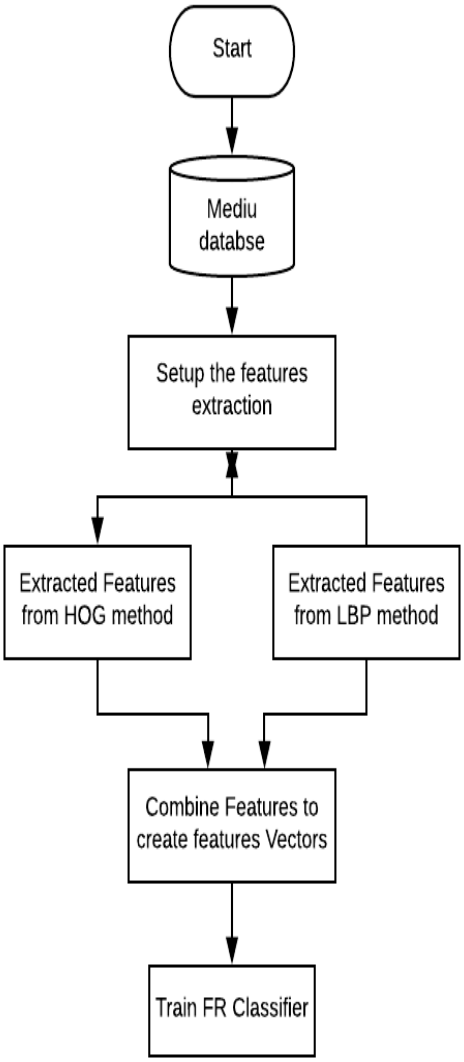


Fig. 2. *Training random forest classifier*

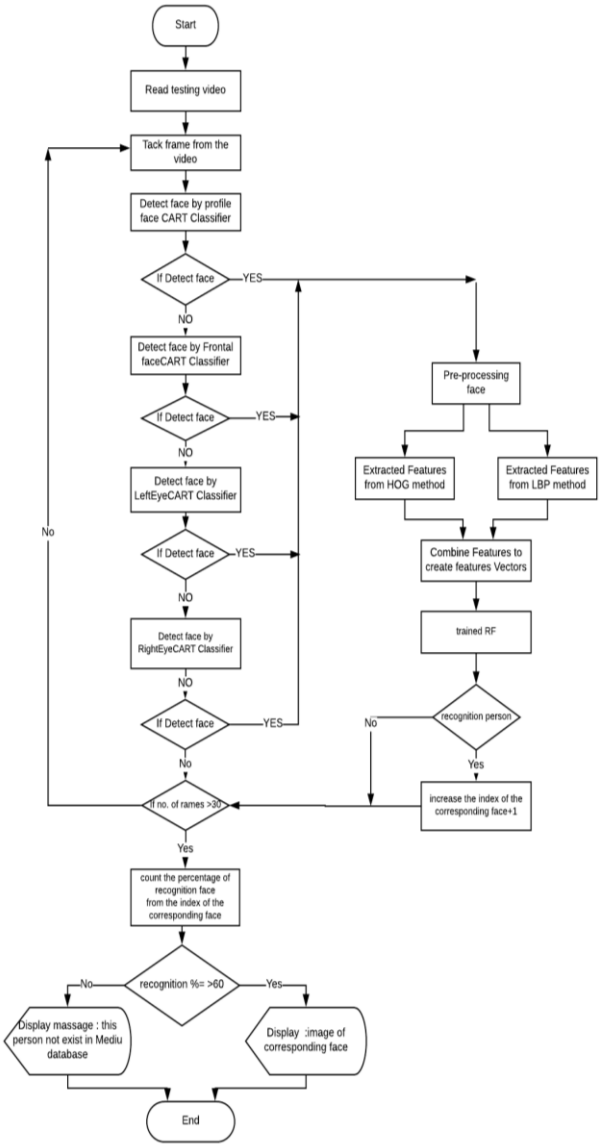


Fig. 3 *Evaluate the proposed method*

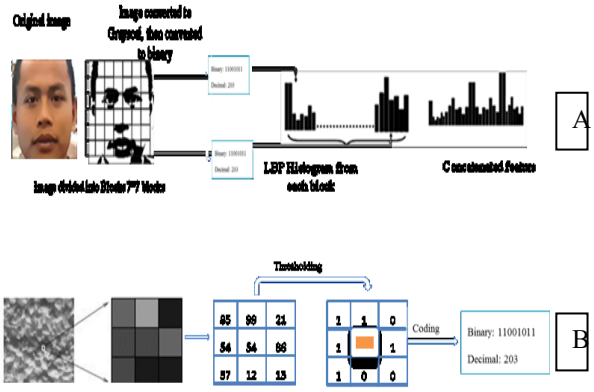


Fig. 4. *LBP labeling procedure (a) and spatial LBP histogram Concatenation (b)*

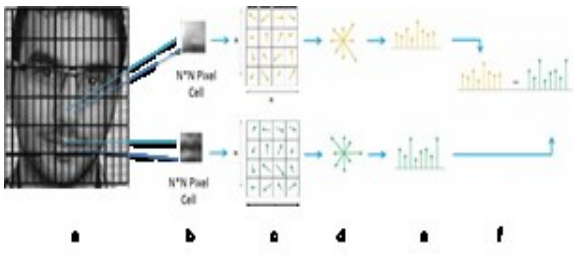


Fig. 5: *Overview of construction HOG feature vector*

III. III. IMPLEMENTATION

In this chapter, results have been withdrawn by addressing the challenges of real-time video. To evaluate the performance of the designed face recognition system, it was implemented based on a database of chosen video samples taken from real life conditions in MEDUE staff considering the recognition of rate's duration and the following six experiments had been done.

A. Experiment 1: Training RF to optimize the number of trees:

To recognition, the rate and accuracy of overall system the RF Classifier were trained with a different number of trees. By using ten training videos for training RF Classifier and ten videos for testing the trained classifier. From the results it is clear that the number of trees deeply effects on the accuracy and performance of the overall system. As shown in Table 1.

TABLE 1: ACCURACY OF THE SYSTEM USING DIFFERENT NUMBER OF RANDOM FOREST'S TREES.

No. of Trees	100	200	**300**
Recognition accuracy	93.8713 %	96.4238 %	**97.8238 %**

B. Experiment 2: Testing of Proposed System Accuracy and Performance:

Table 2 below shows samples and average of accuracy and performance of the proposed method using RF as classifier and combination of LBP and HOG as feature extractors.

C. Experiment 3: Evaluation the Efficiency Of The Proposed System:

To evaluate the efficiency of the Random Forest classifier with LBP, HOG combination and LBP, HOG separately as a feature extractor, we compared it with Support Vector Machine classifiers (SVM) with the same feature extractors. Using ten videos for training and ten videos for testing. The recognition rate as matched person and result is shown in Table 3.

TABLE 3: EVALUATION THE EFFICIENCY OF THE PROPOSED SYSTEM

Percentage of Face Recognition Accuracy		
Method	**RF**	**SVM**
HOG	95.8659 %	92.1101 %
LBP	96.2935 %	96.9975 %
LBP+HOG	97.82 %	95.3396 %

D. Experiment 4: Developing face Recognition method to recognize more than one face:

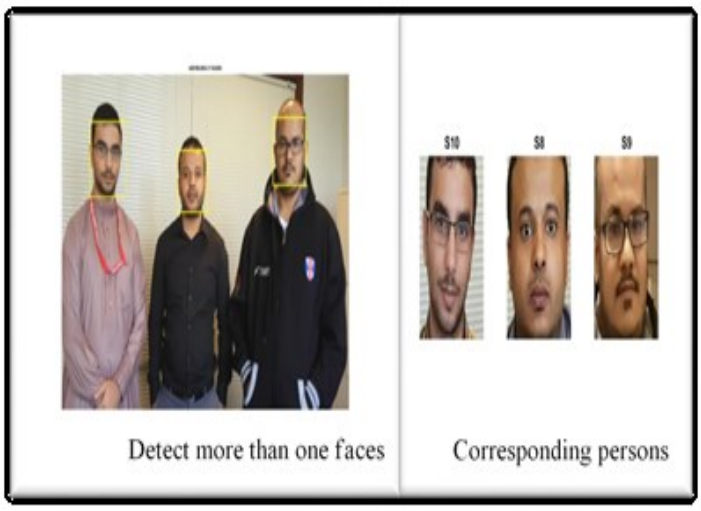
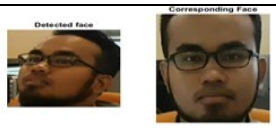
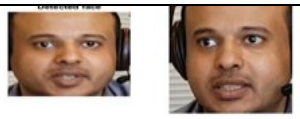


Fig. 6 *Detect the faces and recognize three faces*

The aim of this experiment is to recognize more than one face in the videos at the same time. RF Classifier and HOG and LBP feature extractor were applied to achieve one of the main aims using three staff members in one video, then this video inserts into the system to detect three faces and recognize them at the same time. The results show that the proposed method can detect more than one faces and match the corresponding person as shown in figures 6.

TABLE 2: AVERAGE ACCURACY AND PERFORMANCE OF PROPOSED SYSTEM

Samples of Mediu-S-DB	Percentage of Accuracy for 30 Frame	Percentage of Performance
 S1	100%	6.7016e-05 sec
 S2	100%	6.0375e-06 sec
 S3	100%	8.4524e-06 sec
 S4	88%	2.4754e-05 sec
 S5	95%	1.8716e-05 sec
 S6	100%	1.0264e-05 sec
 S7	100%	1.7509e-05 sec
 S8	95.2381%	:5.4941e-05 sec
 S9	100%	5.4337e-06 sec
 S10	100%	4.83e-06 sec
The Average Consumed Time: 2.1795e-05 sec The Average Accuracy: 97.8238 %		

This experiment is reported to evaluate the proposed method based on the created Mediu Staff database according to the following proposed voting stage method procedures.

IV. DISCUSSION OF RESULT

The aims of this study were achieved by conducting three experiments. Face trained Random Forest classifier to overcome recognition challenges with a different variation of faces by 100 frames for every ten videos from Mediu -S-DB. Training and test videos evaluated the proposed methods and showed the excellent results in experiment 2.

The most important challenge comes when the face recognition system is concentrated on the balance between accuracy and performance with large videos under uncontrolled environment. To solve these problems, face detectors are reported to fast face detection under challenging and variation of faces with low false detection, the combination between LBP and HOG to extract important features from detected faces and RF classifier to obtain best recognition accuracy rate and voting stage is used to improve face matching result. 30 frames from each testing video were taken to reduce the processing time, and the results showed less recognition rate and high accuracy of the overall system. Furthermore, this system carries the characteristics of being simple regarding time computing due to the tagging of both face descriptors (LBP), (HOG) and the RF classifier.

In experiment 3 the results demonstrate the robustness of the system to recognize the face in high accuracy and short time compared with SVM. It was noticed that the combination between LBP and HOG are more robust than using them separately. Despite, increasing number of trees of a random forest leads to raising the accuracy of the system, as well as it is time consuming compared to the fewer trees random forest as shows as experiment 1.

V. CONCLUSION

There is a significant demand for automatic face recognition in society. The study has offered to improve recognition methods by the Viola Jones algorithm to improve the performance of detected face. The appearance of the face was analyzed and extracted using a combination of LBP and HOG descriptors. RF classifier is trained by training data, and it is applied to recognize the face correctly from the testing database. However, it should be noted that, although increasing the random forest tree's number makes the forest more accurate, it makes the system time-consuming. Considering the combination of LBP & HOG, using both methods can improve the overall system accuracy with RF classification and voting stage. The suggested methods proved very efficient by giving excellent results regarding performance and accuracy when compared with existing methods. The proposed methods were defined, performed, tested, and performance evaluated uncontrolled environments on collected Mediu-S-DB. The proposed algorithms proved accurate the identity and verify one person or more people from video broadcast under a controlled environment. In fact, this developed face recognition technique can be employed to improve real-world applications such as attendance management system instead of other biometric applications in any institution educational.

REFERENCES

[1] B. Surekha, Kanchan Jayant Nazare, S. V. Raju, and N. Dey, "Attendance Recording System Using Partial Face Recognition Algorithm," in Intelligent Techniques in Signal Processing for Multimedia Security, vol. 660, 2017, pp. 47–64.

[2] L. Nanni, A. Lumini, and S. Brahnam, "Ensemble of texture descriptors for face recognition obtained by varying feature transforms and preprocessing approaches," Applied Soft Computing Journal, vol. 61, pp. 8–16, 2017.

[3] L. Liu, P. Fieguth, Y. Guo, X. Wang, and M. Pietikäinen, "Local binary features for texture classification: Taxonomy and experimental study," Pattern Recognition, vol. 62, pp. 135–160, 2017.

[4] V. Pali, S. Goswami, and L. Bhaiya, "An Extensive Survey on Feature Extraction Techniques for Facial Image Processing," International Conference on Computational Intelligence and Communication Networks (CICN), pp. 142–148, 2014.

[5] D. Hutchison, Biometric Recognition. 2015.

[6] A. K. Agrawal and Y. N. Singh, "An efficient approach for face recognition in uncontrolled environment," Multimedia Tools and Applications, vol. 76, no. 3, pp. 3751–3760, 2017.

[7] A. Salhi, M. Kardouchi, and N. Belacel, "Fast and efficient face recognition system using Random Forest and Histograms of Oriented Gradients," Biometrics Special Interest Group (BIOSIG), 2012 BIOSIG. pp. 293–303, 2012.

[8] P. Carcagnì, M. Del Coco, M. Leo, and C. Distante, "Facial expression recognition and histograms of oriented gradients: a comprehensive study," SpringerPlus, vol. 4, no. 1, p. 645, 2015.

[9] S. Dabbaghchian, M. P. Ghaemmaghami, and A. Aghagolzadeh, "Feature extraction using discrete cosine transform and discrimination power analysis with a face recognition technology," Pattern Recognition, vol. 43, no. 4, pp. 1431–1440, 2010.

[10] A. Nabatchian, E. Abdel-Raheem, and M. Ahmadi, "Human Face Recognition Using Different Moment Invariants: A Comparative Study," 2008 Congress on Image and Signal Processing, pp. 661–666, 2008.

[11] P. R. Misal, "Face Reconstruction and Authentication Based on Wavelet Transform and Random Forest : A Review," 2016.

[12] X. Pu, K. Fan, X. Chen, L. Ji, and Z. Zhou, "Facial expression recognition from image sequences using twofold random forest classifier," Neurocomputing, vol. 168, pp. 1173–1180, 2015.

[13] Y. Mishina, R. Murata, Y. Yamauchi, T. Yamashita, and H. Fujiyoshi, "Boosted random forest," IEICE Transactions on Information and Systems, vol. E98D, no. 9, pp. 1630–1636, 2015.

[14] N. D. V Santhi, Studies in Computational Intelligence 660 Intelligent Techniques in Signal Processing for Multimedia Security. .

[15] H. H. Mady and S. M. S. Hilles, "Efficient Real Time Attendance System Based on Face Detection Case Study ' MEDIU Staff ,'" vol. 1, no. 2, 2017.

[16] G. Hemalatha and C. P. Sumathi, "Preprocessing techniques of facial image with Median and Gabor filters," 2016 International Conference on Information Communication and Embedded Systems, ICICES 2016, pp. 1–6, 2016.

[17] A. Albiol, D. Monzo, A. Martin, J. Sastre, and A. Albiol, "Face recognition using HOG-EBGM," Pattern Recognition Letters, vol. 29, no. 10, pp. 1537–1543, 2008.

[18] J. L. B, M. Castrill, and E. Ram, "Video Analytics for Audience Measurement," vol. 8811, no. November, pp. 53–65, 2014.

[19] Hilles, S., & Maidanuk, V. P. (2014). Self-organization feature map based on VQ components to solve image coding problem. ARPN Journal of Engineering and Applied Sciences. Vol. 9,№ 9: 1469-1475.

[20] Hilles, S. M., & Hossain, M. A. (2018). Classification on Image Compression Methods. International Journal of Data Science Research, 1(1), 1-7.