

ROBUST DETECTION AND RECOGNITION SYSTEM BASED ON FACIAL EXTRACTION AND DECISION TREE

Ansam H. Rashed¹

*Muthana H. Hamd¹

1) Department of Computer Engineering, College of Engineering, Mustansiriyah University, Baghdad, Iraq.

Received 11/12/2020

Accepted in revised form 17/1/2021

Published 1/7/2021

Abstract: Automatic face recognition system is suggested in this work on the basis of appearance based features focusing on the whole image as well as local based features focusing on critical face points like eyes, mouth, and nose for generating further details. Face detection is the major phase in face recognition systems, certain method for face detection (Viola-Jones) has the ability to process images efficiently and achieve high rates of detection in real time systems. Dimension reduction and feature extraction approaches are going to be utilized on the cropped image caused by detection. One of the simple, yet effective ways for extracting image features is the Local Binary Pattern Histogram (LBPH), while the technique of Principal Component Analysis (PCA) was majorly utilized in pattern recognition. Also, the technique of Linear Discriminant Analysis (LDA) utilized for overcoming PCA limitations was efficiently used in face recognition. Furthermore, classification is going to be utilized following the feature extraction. The utilized machine learning algorithms are PART and J48. The suggested system is showing high accuracy for detection with Viola-Jones 98.75, whereas the features which are extracted by means of LDA with J48 provided the best results of (F-measure, Recall, and Precision).

Keywords: *Face Detection, Face Recognition, Principal Component Analysis, Linear Discriminant Analysis, Local Binary Pattern Histogram, Viola-Jones, Feature Extraction, Classification.*

1. Introduction

One of the major developed areas in pattern recognition and computer vision is face detection, also it has been the first significant stage for the approaches of facial analysis and one of the major significant topics in computer vision like face authentication, facial expression, face recognition and people tracking. Also, with the wide utilization of inexpensive digital cameras and internet, along with powerful software for editing images like Photoshop, the typical user has further access to digital processing tools compared to in the past. The main aim of face detection is determining whether there were any faces in images, after that returning the bounding box and location regarding each one of the faces in an image irrespective of expression, facial pose, occlusions, illuminations, and orientation. Automatic human tracking and detection is a vital research field and offering various application areas [1]. Also, tracking can be considered as one of the significant tracking system steps, that is localizing and associating features over frames' series. In addition, face

*Corresponding Author: dr.muthana.hamd@gmail.com

recognition is of high importance due to its considerable possibilities in many applications (civilian applications, smart-card applications, security, surveillance, human computer interaction, criminal justice systems, image data-base investigations, multimedia environment with adaptive human computer interface and video indexing) [2].

The increase in using computer vision to replace humans, surveillance, initiated the studies in the face detection area. Former researches are biased towards human recognition instead of tracking. The tracking requirements were raised via tracking human movement. In addition, movement tracking is vital to identify people and know their attentions [1]. The efficiency regarding many face-based applications, from a standard face verification and recognition to recent face clustering, tagging, and retrieval, is based on accurate and effective face detection. One of the major parts regarding face recognition systems is face detection since it has the capability of focusing the computational resources on significant image parts which contain faces.

Face recognition concerns with recognizing individuals with their intrinsic facial characteristic. In comparison to other biometrics, face recognition is more natural, non-intrusive and can be used without the cooperation of the individual. In general, face recognition systems can be used in two manners: verification and identification. The face verification system (one-to-one matching) involves confirming or denying the identity claimed by a specific individual. Face identification system (one-to-many matching) tries to find the identity of a specific individual against all image templates in face individual database [3].

The techniques of face recognition might be categorized into local and holistic approaches, latter is legally trying to identify faces with the

use of global representation on the basis of the whole image instead of local facial features. The holistic approach might be categorized as nonlinear or linear. Linear dimension reduction is performed via linear appearance based approaches [4], while the projection coefficient is utilized as feature representation with regard to each one of the face images via the projection related to the face image vector onto a basis vector. Linear approaches were LDA and PCA. Local approaches are treating just certain facial features. They have been more sensitive to facial expressions, pose, and occlusions. The major aim of such techniques is discovering distinctive features like LBPH.

Several types of research have been applied in PCA, LDA and LBPH such as

Nawaf Hazim Barnouti [3] in the year 2016 proposed a system of automatic face recognition on the basis of exterior based approaches, while the method of Viola-Jones was utilized for detecting and gathering faces in each one of the databases. In addition, square euclidean distance was applied for determining distances between 2 images, which result in finding the similarity of an image.

S. Haji and A. Varol [5] in 2016, propose a real-time application, based on the Windows operating system, for face recognition. The system has been based on Eigen and Local Binary Patterns for measuring the similarity between the face images, to authenticate users, to reduce the effect of different illumination conditions. Despite the high recognition accuracy of 90%, the study shows that the accuracy is dramatically affected by any changes to the conditions of the environment, such as distance between the individual and the camera or ambient lighting, or when the images are collected using different cameras.

Shan et al., [17] investigated the LBP method for texture encoding in facial expression description. Two methods of feature extraction were proposed, in the first one, features were extracted from a fixed set of patches. In the second method, the features are extracted from most probable patches found by boosting.

Malhotra et al. [18], proposed an illumination invariant face recognition algorithm based on the combination of gradient-based illumination normalization and fusion of two illumination invariant descriptors. The ratio of gradient amplitude and original image intensity allows an invariant visual representation of the illumination. The feature sets obtained from LBP and LTP methods were consolidated into a single feature set by using feature normalization and feature selection. Where an artificial neural network was used in the classification stage.

L.Novamizanti, A. Luhur and B. Satria [19] in 2020, proposed a system for classifying infant crying sounds using Linear Discriminant Analysis (LDA) with Discrete Wavelet Transform (DWT) and Mel-frequency Cepstral Coefficient (MFCC) as a feature extraction method. The system can identify the sound of crying babies grouped into 5 (five) classes, namely discomfort, hunger, colds, burp, and drowsiness. The system achieves an accuracy of 94% and an average computing time of 1.5506 seconds. The performance of Linear Discriminant Analysis (LDA) outperformed Principal Component Analysis (PCA) in the identification of crying babies

2. Proposed Methodology

In real-world applications, face recognition is one of the complex image processing problems. The presented study is providing details for the training process and the approach of suggested face recognition and detection system.

Technologies features and characteristics making face recognition vital and excellent performers on the basis of application.

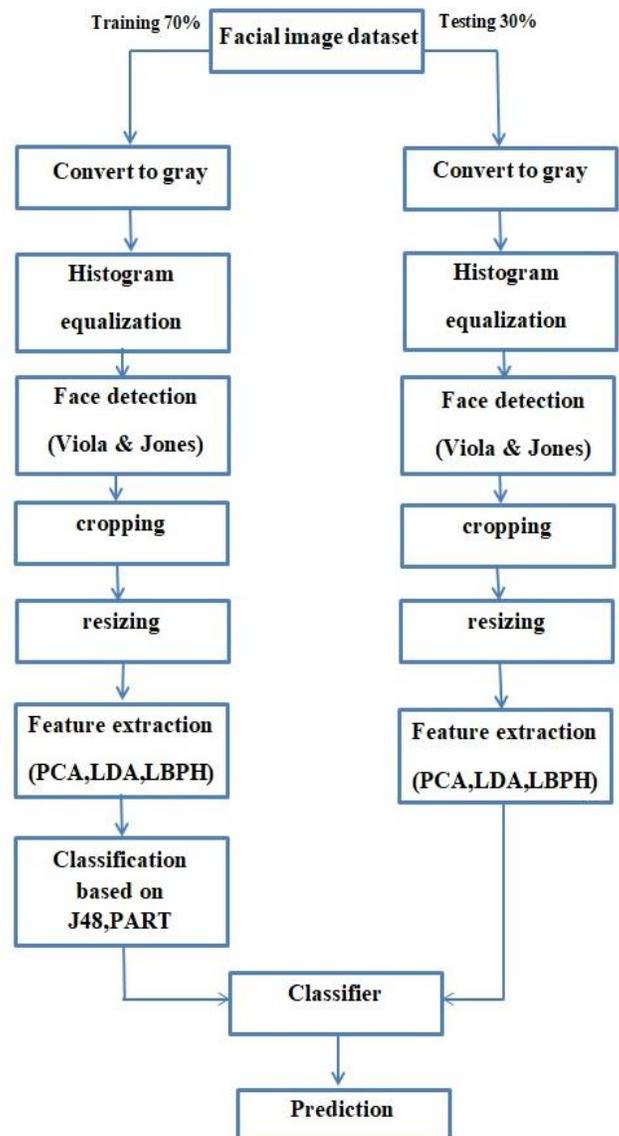


Figure1. Face Recognition Proposed Methodology Process

Basically, face recognition is categorized into 3 steps beginning with face detection continuing with feature extraction as well as ending up with a classification process. Certain database Milborrow University Of Cape Town (MUCT) facial images composed of 240 colored facial images for both men and women that obtained from [76] which are divided into 20 class and

each class consist of 12 images is utilized for testing the performance of the system, database was created to provide more diversity of lighting, age, and ethnicity [6]. In addition, Viola-Jones approach is applied for detecting and cropping the face region in each one of the database images. The approaches of feature extraction LBPH, LDA, and PCA were utilized for dimension reduction and feature extraction. Lastly, PART and J48 methods of classification are utilized. Fig1 is showing the suggested method

2.1 Face Detection

Generally, face detection is specified as one of the cases related to object class detection, also it is one of the significant topics in biometric research. In addition, face detection is specified as a major phase in face recognition systems, the identified faces utilized to classification methods and feature extraction [3], whereas face detection isn't simple, since it is carrying many appearance variations in images, like illumination condition, image orientation, facial expression, occlusion and pose variation. Viola-Jones approach is used in this study.

The method of Viola-jones object detection has been proposed in 2001 via P. Viola as well as M. Jones, the approach was vital in the 2000s and specified as the first framework for object detection for providing relevant object detection running in real-time applications. In addition, Viola-Jones is requiring full view frontal upright faces. At such levels, the approach is reading each one of the input images with a window that looks for features in human faces. In the case when collecting enough features, such image's window type is indicated to be a face [7]. For the purpose of obtaining faces of different sizes, then the window should be scaled, while the process will be repeated a lot of times. With regard to each one of the window scales involve

via the approach in a separate way from the other scales, the approach turned out to be time-consuming due to the estimation of various image sizes. For decreasing the number of features must be checked via each one of the windows, all windows must be passing through levels. Therefore, early levels are including less checked features and simpler to pass, yet later levels have further features and they are demanding. At every one of those levels, the features' evaluation for levels have been gathered and in the case where the collected value is not passing the value of the threshold, then the level fails and such window won't be identified as face. Furthermore, the approach of Viola-Jones has been divided to 3 major parts (which are: Integral image, classifier learning with adaBoost as well as attention cascade structure) which make it likely to develop effective face detection which might be utilized on real-time applications.

A. Creating an Integral Image

Integral images, also referred to as summed-area tables, were evaluated as a step of pre-processing. The initial phase of Viola-Jones is converting an input facial image to an integral one, this might be achieved via making every one of the pixels equal to the whole sum regarding all above pixels and to the left of that pixel [6].

In which I represent the integral image, while O represents the original image.

$$I(u, w) = \sum_{\hat{u} \leq u, \hat{w} \leq w} O(\hat{u}, \hat{w}),$$

Where

I is the integral image and

O is the original image.

$$A(u, w) = A(u, w - 1) + O(u, w)$$

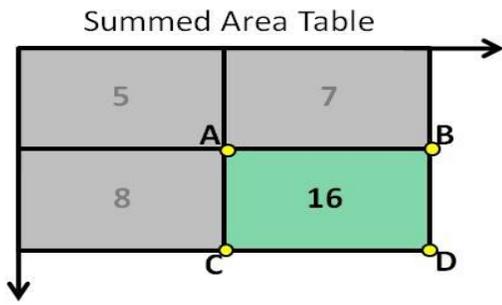
$$I(u, w) = I(u - 1, w) + A(u, w)$$

Where

A (u, w) is the accumulative row summation

$$A(u, -1) = 0, \text{ and } I(-1, w) = 0.$$

The integral image might have been calculated in one override over the initial picture.



Thus, each one of the Haar-like rectangular summations might be done with precision in only 4 reference arrays, whereas the evaluation of various factors within rectangular summations of Haar-like features may be done in 6 references

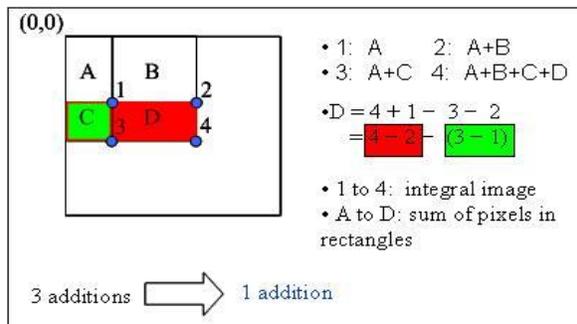


Figure 2. Calculation of integral image

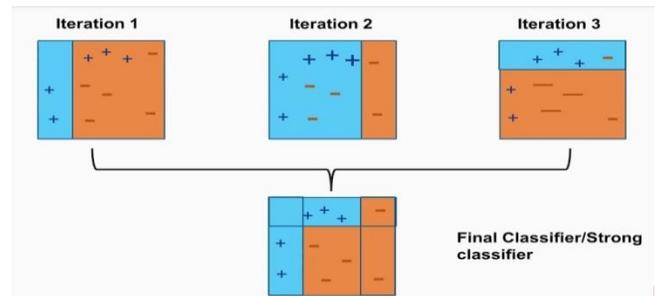
B. Haar-like features

Scaling with regard to Haar-like features was utilized instead of scaling image sub windows for getting a system of high speed detection, while the filter process related to Haar-like features was one the basis of picture’s categorization through Viola-Jones’ system. In addition, Haar-like features were majorly efficient in the processes of face detection. After the identification of integral image, there will be a simultaneous identification of every Haar-like property at every scale or region in image. The different types of rectangle Haar-like features) is shown in figure (3)



Figure3. Haar-like features collection

Adaboost can be defined as one of the machine learning boosting methods with the ability to find a very precise classifier through combining various weak classifiers, each of them with average accuracy. AdaBoost training can be defined as a machine learning approach, where a little is selected in place of the delicate classifiers, each of which is specified accurately



with one Hear- like feature, which is then combined for forming a powerful classifier

Figure 4. AdaBoost training

D. Cascaded Classifiers

Viola-Jones created a cascaded classifiers architecture from a series of strong classifiers for rapidly discarding insignificant image regions not containing faces. In addition, strong classifiers in a cascade were grouped in ascending order of complexity. Thus, various regions that were not likely to include faces have been eliminated through initial classifiers with not many efforts, whereas more computation is done on candidate regions via the later, further advanced classifiers, such approach has resulted in the sufficient increase of the performance of detection and significantly reducing time for computation. The AdaBoost-based detector of Viola-Jones was precise and rapid. Its ease of implementation contributes to its efficiency [16].

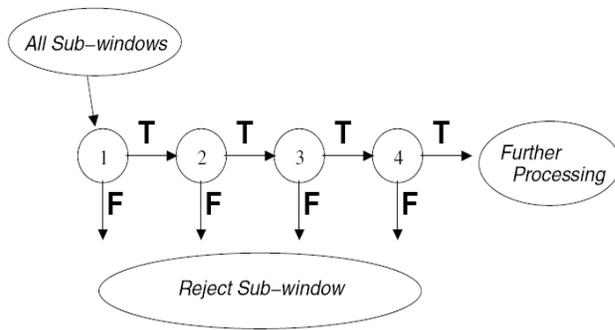


Figure 5. Classifier Cascade Process

2.2 Feature Extraction

Feature extraction includes decreasing the number of the resources that are needed for describing large data amounts. Also, feature extraction from certain data was one of the critical problems for effective applications related to machine learning. In the presented study, LBPH, LDA, and PCA were utilized as dimension reduction and feature extraction approaches from original face images. LBPH, LDA, and PCA are producing feature vectors in the reduced dimensions.

2.2.1 Principal Component Analysis (PCA)

PCA has been specified as a major approach utilized in pattern recognition as well as compression, also it is specified as a dimensions reduction and feature extraction approach [9]. PCA is a major statistical approach that uses a holistic method for finding patterns in the high-dimensional data. The aim of PCA has been taken from the approach of information theory that is breaking down face images to small groups of the characteristic feature images that have been referred to as Eigen-faces utilized for representing new and current faces [10]. In PCA, 2D facial image matrices should be converted to a 1D vector. The 1D vector might be a column or a row. Thus, an image representation results in high-dimensional space [11]. The steps of PCA are:

- 1- Train a set of M images utilized for computing the Average Mean as follows:

$$\text{Average} = \frac{1}{M} \sum_{n=1}^M \text{Training image}(n)$$

- 2- Subtracting the original image from Average Mean as follows:

$$\text{SUB} = \text{Training images} - \text{Average}$$

- 3- Covariance Matrix is computed as follows

$$\text{covariance} = \sum_{n=1}^M \text{Sub}(n)\text{Sub}(n)^T$$

- 4- Computing the eigenvectors and eigenvalues related to covariance matrix.
- 5- The best eigenvalues are sorted and chosen. The highest eigenvalues belonging to group of eigenvectors were selected, such M eigenvectors are describing eigenfaces. Assume encountering new faces, eigenfaces might be recalculated or updated.
- 6- Projecting the training samples on eigenfaces

2.2.2 Linear Discriminant Analysis (LDA)

LDA has been referred to as the fisher-face approach also, which is utilized for overcoming the limitations of PCA in terms of its application that has been kept in a small image data-base. It was accomplished via the projection of an image onto eigenfaces space through PCA, then implement pure LDA over it for classifying eigen face space projected data [10]. In addition, LDA is searching for vectors in underlying space which are best discriminating between the classes. Furthermore, the images of the LDA group that are related to the same class and separating distinctive class images. Mathematically, there have been 2 measures specified (between-class scatter matrix and within-class scatter matrix) [12]. For every class sample the between-class scatter matrix SB and

in-class scatter matrix S_W were specified as follows:

- 1- Computing d-dimensional mean vectors for various classes from the dataset.
- 2- Computing scatter matrices (within class scatter matrix and in-between-class).
- 3- Compute the eigenvectors (e_1, e_2, \dots, e_d) and corresponding eigenvalues ($\lambda_1, \lambda_2, \dots, \lambda_d$) for the scatter matrices.
- 4- Sort the eigenvectors by decreasing eigenvalues and choose k eigenvectors with the largest eigenvalues to form a $d \times k$ dimensional matrix W (where every column represents an eigenvector).
- 5- Using such $d \times k$ eigenvector matrix for transforming the samples into new sub-space, this might be specified via matrix multiplication:

$Y = X \times W$ (in which X is a $n \times d$ - dimensional matrix represents n samples, and y was the transformed $n \times k$ -dimensional samples in new sub-space).

2.2.3 Local Binary Pattern Histogram (LBPH)

This algorithm combines Histograms of Oriented Gradients (HOG) descriptor with Local Binary Patterns (LBP). The latter is a simple, yet significant approach for extracting and labeling image pixels. With the use of LBPH, one might simply specify face images with only a straightforward vector. LBP was initially provided and developed for being a texture analysis for gray scale images. LBPH is specified to be adequate for feature extraction since it is describing the structure and texture of images. The main benefit of the LBP method is that they are combining statistical and structural approaches result in increasing texture analysis performance.

An operator is labeling the image pixels via thresholding 3×3 -neighborhood regarding each one of the pixels with center value and assuming the result as a binary number. The steps of LBPH are [13]

- 1- Assuming an image with dimensions of $N \times M$, it will be divided into regions of the same width and height leading to $(m \times m)$ dimension for each one of the regions.
- 2- A local binary operator was utilized in each one of the regions, LBP operator was specified as a window of 3×3 .

$$LBP(x_c, y_c) = \sum_{p=0}^7 S(i_p - i_c) 2^p$$

In which

i_c : central pixel values

i_p : neighbor pixel values.

- 3- With the use of a median pixel value as a threshold, it is comparing the pixel to its 8 closest pixels with the use of such, in the case when the neighbor value is equal or more than the central value, then it will be 1 or else 0.

$$S(t) = \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases}$$

For neighbors equal to 8 for each one of the labels, there were 256 likely combinations ($2^8 = 256$).

2.3 Classification phase

Classification is the process of building a model of classes from a set of records that contain class labels. Decision Tree Algorithm is to find out the way the attributes-vector behaves for a number of instances. Extracted features have been fed to the classifier. PART and J48 are used for classification.

2.3.1 J48 Algorithm:

It is a type of the greatest famous algorithms of machine learning decision tree. Generally, the constructed tree was top down recursive divide-and-conquer style, in the beginning, all training samples were at the root, attributes have been definite (in the case when values were continuous, then they are separated in advance), the samples are separated recursively depending on the chosen attributes the test attributes were selected based on heuristic or statistical measure (for instance, information gain)[14].

2.3.2 PART Algorithm

PART (Projective Adaptive Resonance Theory) is a partial decision tree algorithm. In particular, PART generates a set of rules according to the divide-and-conquer strategy, removes all instances from the training collection that are covered by this rule and proceeds recursively until no instance remains. To generate a single rule, PART builds a partial C4.5 decision tree for the current set of instances and selects the leaf with the largest coverage as the new rule. Afterwards, the partial decision tree along with the instances covered by the new rule are removed from the training data, in order to avoid early generalization. The process is repeated until all instances are covered by extracted rules [15].

3. Performance Measures

1. Precision: represents the amount of the true positives that are separated by the number of the true positive cases as well as the number of the false positives.

$$\text{Precision} = \frac{TP}{TP+FP}$$

2. Recall: The capability of finding every relevant examples in a data-set, precision represents the ratio of the data points this model states was relevant were in fact relevant

$$\text{Recall} = \frac{TP}{TP + FN}$$

3. F-measure: The harmonic medium value of accuracy and recall, with F1 being the optimal value at one and worse at zero.

$$F_1 = \frac{1}{\frac{1}{\text{recall}} + \frac{1}{\text{precision}}} = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

4. Result and Discussion

The method of Viola-Jones was applied for face detection on each one of the databases, such an approach reached a high detection rate 98.75%, also all the images were detected as well as cropped in the databases as shown in figure (9).



Figure 9. Detection and cropping on MUCT database

After performing the preprocessing, face detection, and resizing, the second part of this work is feature extraction using three methods PCA, LDA, and LBPH, the input of this part is cropped facial image and the output is image features.

Table 1. The experimental results of implemented PCA, LBPH, and LDA on the cropped facial images		
Methods	Features Vector Length	Time of Extraction
LDA	19	5.10 sec
PCA	237	6.32 sec
LBPH	16381	5.59 sec

The table show that LDA give a minimum amount of features with minimum extraction time and this is important point in recognition and identification systems.

In the third part of the work, a description is obtained about the results of applying both PART, and J48 machine learning algorithms for classifying the features obtained from the previous described feature extraction algorithms which include LDA, PCA, and LBPH.

The experimental results of implemented classification algorithms PART and J48 with PCA, LBPH, and LDA are shown in tables (1) and (2).

Table 2. The experimental results of implemented the PART with PCA, LBPH, and LDA.

Methods	Precision	Recall	F-measure
PART+PCA	73.301	72.995	72.221
PART+ LBPH	50.146	51.476	50.232
PART+LDA	91.215	90.717	90.763

Table 3. The experimental results of implemented the J48 with PCA, LBPH, and LDA.

Methods	Precision %	Recall %	F-measure %
J48+PCA	79.815	78.902	78.804
J48+ LBPH	53.011	51.898	51.345
J48+LDA	94.337	94.092	94.64

The proposed system procedures are accomplished on a hp laptop with CORE i7 , 8 GB RAM size , the screen card is nVIDIA QUADRO, 500 GB HDD and the operating system is windows 10 64 bit. The proposed methodology was performed using Python version 3.6 software package.

4. Conclusion

The aim of this study is to implement a system for automatic face recognition. In addition, face

detection with the use of the Viola-Jones approach was applied for detecting and cropping faces in a certain database (MUCT). The approach of Viola-Jones shows a high detection rate of 98.75%. PCA, LDA, LBPH is utilized for dimension reduction and feature extraction. J48 and PART classification algorithms used with features extracted from PCA, LDA, LBPH which leads to finding LDA gives the best results of (Precision, Recall, F-measure) with J48 classification algorithm. The recognition time has been accepted and it took only a small number of the seconds. Results showed an increase in the rates of recognition in the case where the number of the training images has been increased.

Acknowledgements

My sincere thanks and gratitude to Al Mustansiriyah university for the support and encouragement. <https://uomustansiriyah.edu.iq/>

Conflict of interest

The author declare that there are no conflicts of interest regarding the publication of this manuscript.

5. References

1. Hatem, H., Beiji, Z., Majeed, R., Lutf, M. and Waleed, J., 2015. Face Detection and Pose Estimation Based on Evaluating Facial Feature Selection. International Journal of Hybrid Information Technology, 8(2), pp.109-120.
2. Bakshi, U. and Singhal, R., 2014. A survey on face detection methods and feature extraction techniques of face recognition. International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), 3(3), pp.233-237.

3. Barnouti, Nawaf Hazim. 2016. Improve Face Recognition Rate Using Different Image Pre-Processing Techniques, American Journal of Engineering Research (AJER), 5(4), pp.46-53.
4. Rath, S.K. and Rautaray, S.S., 2014. A Survey on Face Detection and Recognition Techniques in Different Application Domain. International Journal of Modern Education and Computer Science, 6(8), pp.34-44.
5. Haji, Suad, and Asaf Varol. "Real time face recognition system (RTFRS)." In 2016 4th International Symposium on Digital Forensic and Security (ISDFS), pp. 107-111. IEEE, 2016.
6. Milborrow, S., Morkel, J. and Nicolls, F., 2010. The MUCT landmarked face database. Pattern Recognition Association of South Africa, 201(0).
7. Cha, Z. and Zhengyou, Z., 2010. A survey of recent advances in face detection. Microsoft Research, Microsoft Corporation.
8. Hefenbrock, D., Oberg, J., Thanh, N.T.N., Kastner, R. and Baden, S.B., 2010, May. Accelerating viola-jones face detection to fpga-level using gpus. In 2010 18th IEEE Annual International Symposium on Field-Programmable Custom Computing Machines (pp. 11-18). IEEE.
9. Kaushik, Sangeeta, Dubey, R. B., and Abhimanyu Madan. 2014. Study of Face Recognition Techniques, International Journal of Advanced Computer Research, 4(17), pp.939-949.
10. Singh, A., Singh, S.K. and Tiwari, S., 2012. Comparison of face Recognition Algorithms on Dummy Faces. The International Journal of Multimedia & Its Applications, 4(4), pp.121-135.
11. Barnouti, Nawaf Hazim. 2016. Face Recognition Using Eigen-Face Implemented On DSP Processor, International Journal of Engineering Research and General Science, 4(2), pp.107-113.
12. Bhattacharyya, S.K. and Rahul, K., 2013. Face Recognition by Linear Discriminant Analysis. International Journal of Communication Network Security, ISSN, 2(2), pp.2231-1882.
13. Tariq, Umair Ulah, Waqar Ahmad, M. Daud Abdullah Asif, and Mubashir Hussain. "Gender Perception From Faces Using Boosted LBPH (Local Binary Patten Histograms)." Carpathian Journal of Electronic & Computer Engineering 6, no. 1 (2013).
14. Chaurasia, Vikas, and Saurabh Pal. "Data mining approach to detect heart diseases." International Journal of Advanced Computer Science and Information Technology (IJACSIT) Vol 2 (2014): 56-66.
15. Eibe F.; Witten I.H. Generating Accurate Rule Sets Without Global Optimization. Fifteenth International Conference on Machine Learning 1998, 144–151.
16. Nawaf Hazim Barnouti ,Sinan Sameer Mahmood Al-Dabbagh ,Wael Esam Matti ,Mustafa Abdul Sahib Naser, —Face Detection and Recognition Using Viola-Jones with PCA-LDA and Square Euclidean Distance International Journal of Advanced Computer Science and Applications, Vol. 7, No. 5, 2016.
17. C. Shan, S. Gong, P. W. J. I. McOwan, and v. Computing, "Facial expression recognition based on local binary patterns: A comprehensive study," Image and vision Computing. vol. 27, no. 6, pp. 803-816, 2009.

18. D. Huang, C. Shan, M. Ardabilian, Y. Wang and L. Chen, "*Local Binary Patterns and Its Application to Facial Image Analysis: A Survey*," in IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 41, no. 6, pp. 765-781, Nov. 2011.
19. Ledy Novamizanti, Anggunmeka Luhur Prasasti , Bangun Satria Utama "Study of Linear Discriminant Analysis to Identify Baby Cry Based on DWT and MFCC", IOP Conf. Series: Materials Science and Engineering 982 (2020).