RDUP3: Relative Distance based User Profiling from Profile Picture in Multi-Social Networking

Vasanthakumar G. U.
Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore University, Bangalore, India.

P. Deepa Shenoy
Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore University, Bangalore, India.

Venugopal K. R.
Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore University, Bangalore, India.

ABSTRACT

User Profiling in Online Social Network (OSN) requires the frontal photographs of the users as thier Profile Pictures in Multi-Social Networking. The existing algorithms are ineffective in detecting the facial features like eyes, mouth and nose on the face appropriately, making it inefficient. This work proposes a novel approach to efficiently detect the facial features and improve the effectiveness of face detection and recognition by bifurcating the detected face horizontally, vertically and cropping it. The algorithm is effectively run only on the portion of the detected face Bounded Box (BB) and area to generate bounded boxes of other facial objects and later the Euclidian Distance (ED) between those BBs with respect to that of the face is computed to get Logarithm of Determinant of Euclidian Distance Matrix (LDEDM) in Relative-Distance (RD) method and stored in the database. The LDEDM so computed is unique for every user under consideration and is further utilized for identity matching recognizing from the database. The results show that the Equal Error Rate (EER) is considerably low indicating accurate threshold fixation for better performance with the proposed Relative Distance based User Profiling from Profile Picture (RDUP3) algorithm.

Keywords

Data Mining, Face Detection, Online Social Networks, Profile Picture, Relative Distance, User Profiling

1. INTRODUCTION

In online social networks, usage of photos and video sharing has tremendously increased in recent years leading to accumulation of huge data every minute. Online sharing of images are imparting numerous problems like misuse of photos, creating fake profiles, crime, forgery and so on. User name, profile picture, email id, current place, date-of-birth, interests, education and various other information are available for each profile of the user created in any social network. Tremendous data available on social media may be used appropriately for betterment of lives of common people by performing data mining dynamically [1] [2].

Identifying people on the web has become mandatory for various purposes in the present world. As and how the social network usage increases, identification of its users becomes prerequisite for security reasons [3] [4] [5]. There are many identification techniques like Password/Personal Identification Number etc., available, but reliability on such methods are risky since there exists high chances of lapses and forgery. To overcome the said problems, the use of pattern recognition techniques increased in biometric identification systems [6] for people identification using their physiological characteristics. Typical application areas of face detection mechanism include smart cards, entertainment, law enforcement, surveillance and Information security.

Automatic recognition of persons is a challenging task and is an active research topic in recent past. Since there is no specific technique available that can be applied in all situations and applications, depending on the applications, new techniques are being developed with advancements. Application of techniques and its results vary with the type of image as well. Normal images may provide different results when compared to 2D/3D images. Face Recognition Techniques are of various types: i) Geometric characteristics, ii) Template matching, where former provides limited results and later uses Linear Discriminant Analysis (LDA) or Principal Component Analysis (PCA) and deformable templates like Elastic Graph Matching (EGM) and neural network classification, iii) appearance-based, iv) Point features and v) Regional features. Based on type and quality of image, one of these techniques can be applied to detect face(s) in the given image.

Face detection is basic and an important part of face recognition. By localizing the objects in the given image, face is detected and recognized [7]. Some common techniques of face detection are: i) finding face(s) with controlled background, ii) finding face(s) by color (Color Segmentation), iii) finding face(s) by motion (video), iv) finding face(s) by combing color and motion techniques, v) finding face(s) in unconstrained scenes: a) Model based face tracking and b) Weak classifier cascades. Real-time face recognition has few fundamental issues which include size, different positions of face appearance, pose, illumination, age, lighting, background and so on. Some issues like size, feature

location, grey level normalization, increase the challenge of face detection. Aging and 3D pose variations are the two difficult problems expecting more concentration while resulting in better performance during face recognition. Occlusion, complex background, different facial expressions, result in decreasing the performance of face detection algorithm, hence such images need more attention while addressing. The overall performance of the algorithm is calculated based on the face detection in real-time with high accuracy and detection rate with less computational time.

Trends in technology vary day-by-day with new advancements and as the applications increase, the need for automated face detection and recognition increase. General face detection mechanism follows segmentation, detection and followed by Recognition. With temporal segmentation, the full image is scanned to extract the key frame from the given image which contains the face. By extracting the key frame from the image, the non-facial objects are excluded. Later, restricting on the key frame, the regional segmentation is processed to detect the possible facial region. This binary classification method helps in avoiding false detection of objects as face in the given image. Region Identification is part of detection which identifies the region of face upon which the actual process of face recognition starts since the exact region of face is obtained. Feature extraction mechanism is followed to get the facial features and then the normalization is applied. These steps followed in face recognition ultimately help recognize the face in the given image and avoid false detection.

The challenges in forensic face recognition mechanism are facial aging, forensic sketch recognition, and facial marks. Person aging directly implies changes in their facial features, making it more challenging to identify accurately. The database may contain old images compared to the existing face structure of humans creating a need for extra attention during face recognition. Though marks on the face gives some identification to the person, yet in few scenarios, these marks may trigger challenges for face recognition since there exist various options to hide these facial marks in the image using Photoshop or other such freely available softwares. Even though normalization technique is used to avoid false recognition with variable poses, it is still a challenging factor during face recognition. Different lighting conditions and variations in lighting effects on face make it difficult to recognize. More often, these techniques of face detection are restrained with 'Noise' problem. It is practically difficult to get images without noise and due to the existence of noise in the images, detection rate and percentage of accuracy reduces.

Initially the given image is classified using classifiers to localize the face. Then the face is detected after segmentation. Here binary classification helps in avoiding risks of detecting non-face objects since prior information about the image will not be known. Then the featured-based detection along with edge detection or other algorithms are applied to detect face in the image and later that face is matched against the images in the database. This process is followed in many applications and in specific, identification of criminals using face detection being one of the major applications of face recognition techniques, during criminal activity investigation, this procedure helps identify the criminals accurately [8].

Motivation: A generalized face detection / recognition algorithm to meet the needs of all possible scenarios is still a requirement. For any algorithms to execute, there needs to be some pre-assumptions made. Changing needs in the applications of face detection motivates the development of new methodologies to achieve high accuracy. Some sensitive areas like national security data, banks, reserved places etc., need tight security with limited access to only authorized and authentic users with proper permissions especially for online activities. These applications critically depend on face recognition techniques. Here, the need for face detection algorithms with high accuracy and less false detection rate acquires high importance. Hence the requirement for developing such algorithms without any drawbacks or restrictions is an active research topic.

Contribution: Given a database of profile pictures, a novel approach is proposed in this work to improve the effectiveness and efficiency of matching the User Profile Pictures by detecting their face and later bifurcating and cropping the detected face to obtain the facial features. The algorithm runs only on the portion of the detected face BB to generate bounded boxes of other facial objects, and later the Euclidian distance between the BBs with respect to that of the face is computed to get LDEDM in RD method and stored in the database. The LDEDM so computed is unique for the user image under consideration and is used for the purpose of matching the identity of the user images from the database. Our proposed approach detects facial features of both single and multiple faces in the given input image.

The remainder of the paper is organized as follows: We briefly review the Related Work as well as the Background Work in Sections-2 and 3 respectively. The problem is defined in Section-4. The Proposed System Model is illustrated in Section-5 and the Relative Distance based User Profiling from Profile Picture (RDUP3) algorithm is presented in Section-6. Section-7 demonstrates Simulation Results and Performance Analysis of the proposed system. Section-8 is summarized with Conclusions and scope for future work.

2. RELATED WORK

Face detection / recognition is achieved in many ways. A geographical face model proposed by Kang-Seo et al [9] uses images of gradient magnitude including an algorithm for face detection from 33 block rank pattern of images. Experiments conducted proved that the method is robust in face detection and performs better to illumination changes. For the purpose of real time face detection using enhanced Haar-like features, Kailash et al. [10] proposed Graphics Processing Unit (GPU) architecture and proved that the approach is faster.

If security is the purpose for which face detection method is used, then the minute aspects should be keenly studied to achieve high performance with accuracy and also to minimize false detection rate. Here, different facial expressions also play a vital role in detection accuracy. Hence for this purpose, Feature Extraction holds good which extracts and examines each feature in the image and then detects the face accurately. Viola-Jones algorithm [11] is the well-known algorithm available for face detection having many versions with improvisation on its basic framework. Two step face tracking method is proposed by Arundati et al. [12] to detect faces in real time live video. Firstly, the face is detected using Haar based algorithm and then from the detected face, the nose, lips, eyes and eyebrows are detected. These features are extracted using the method proposed Shi et al. [13]. The advantage of this method is that, it tracks human faces even in unexpected occlusions, harsh lighting conditions and horizontal or vertical rotation of the face.

To improve the face detection process and to decrease the time taken, Ijaz et al. [14] proposed an improved and hybrid version of Viola-Jones algorithm combining it with skin color pixel detection.

Wide range of algorithms are available for scaling the image to extract its features. These algorithms combine with available Viola-Jones framework in order to detect the face in the given image or during real time face detection. Himanshu et al. [15] examined the effects of Extended Linear [16], Nearest Neighbor [17], Bicubic [18], Bilinear [19] and Piece-wise Extended Linear [20] algorithms on Viola-Jones framework by integrating with available OpenCV library. Parameters like SSIM and PSNR were considered while evaluating the algorithms and the experiments prove that Bicubic algorithm performs better than all others.

To speed up face detection, Liuliu et al. [21] used skin color detection along with Uniform Color Scales (UCS) space which reduces the time taken by Adaboost to search and detect multi-pose faces by image segmentation. Once the face area is detected, then the Adaboost algorithm is used to detect candidate area. Even for colored images, the process of detection varies. Hence for the purpose of face detection in colored images, Yi-Qing [22] proposed learning code and learning algorithm. With the support of Adaboost and attentional cascade, the proposed face detector works faster and detects accurately.

Adrian et al. [23] presented a framework on latest Graphics Processing Unit (GPU) accelerated with the help of Viola-Jones algorithm upon frontal face based on Adaboost utilizing OpenCV and Compute Unified Device Architecture (CUDA). This process improves the overall system performance and the results of the experiment proved that the proposed method performs upto 18 times better than conventional algorithms. Sometimes, the actual challenge lies in the identification of face in the given image and extracting the exact face from the image becomes difficult because the other objects in the image may have some properties which might be similar to that of the face. Hence to solve the problem in this type of situation, Sureshkumar et al. [24] proposed a high performance image retrieval system, which is the combination of attribute-enhanced sparse coding and attribute embedded inverted indexing.

In all situations, it is not possible to consider the whole image for face detection and still achieve high performance rate. So the researchers started localizing the objects in the image and then detecting. By using a formulation of probabilistic classifier, Seyed et al. [25] proposed a face constellation algorithm to localize multi views and to enable multi-view face detection. This method performed better compared to other well known methods. Raphael et al. [26] proposed an active test framework for faster face detection and localization by using hierarchical model and mutation information gain heuristic novel search techniques which efficiently prune search space. From the results of experiments conducted, it is proved that the proposed method detects the face faster by locating the faces in the image by significantly reducing the number of classifier evaluations and increasing the speed exponentially.

Since localization on the complete image may result in less accuracy and also consume more time, Classification / Segmentation method came into existence. Classifying the image into parts and then either localization or detection algorithm is applied to achieve high detection rate. For security purpose, Hongliang et al. [27]

proposed a new matting algorithm which mainly segments human faces. Based on the process of segmentation by this algorithm, the application of this process extended even for verification during security issues, computer vision and face recognition. To automatically determine temporal segments and for their modality fusion synchronization, Hatice et al. [28] proposed a method, which recognizes their role affectively. In order to achieve the synchronization, individual models are trained separately. Firstly, using only the detected neural and apex frames, monomodal affect recognition is obtained. Later, affective face and body modalities are synchronized based on apex phase which is fused for classification at the feature level. Lastly, in decision level from all single modalities, probabilities are combined.

A new methodology with automatic techniques is proposed by Mauricio et al. [29] for face segmentation. By using only the depth information as input, this method achieved landmark detection in images and proved robust in finding the facial expressions in the images. Global face recognition becomes highly challenging while having a single training sample, since it is difficult to achieve high accuracy and feasibility. As a challenge, a single-sample face recognition algorithm [30] is developed based on Locality Preserving Projection (LPP) feature transfer. It uses generalization ability promotion and image segmentation. Initially, using whitened cosine similarity transfer, sources are screened, thus obtaining selective sample source. Then, the vectors of source and target faces are projected using LPP into feature sub-space and the mapping relationship between faces using feature transfer matrix is calculated. Face recognition is achieved based on the nearest neighbor classifier. Experiments conducted on ORL, Yale and FERET data sets proved that proposed LPP outperforms than the available single-sample algorithms like Block Fisher Linear Discriminant Analysis (FLDA) [31] and (PC)2A [32].

Cascade face detection architecture, based on two field feature extraction scheme is proposed by Chih-Rung et al. [33]. This architecture achieves high face detection rate during experiments conducted. Feature extraction is faster since it reduced the amount of memory required for storing and reduced the average detection time of features. For feature extraction or detection purpose even with different facial expressions, Mandeep et al. [34] presented Singular Value Decomposition (SVD) for classifying emotions. This method works efficiently in both real time and in JAFEE database. Experimental results prove that this method distinguishes different expressions by identifying features effectively. Jian et al. [35] proposed a framework which simultaneously recognizes LR (Low Resolution) faces and hallucination. This framework used SVD for face hallucination and recognition. Results of experiment proved that the method produces plausible HR (How Resolution) images with high frequency details and a holistic structure.

Faster internal image calculation is achieved using the Semi-Local Structure Pattern (SLSP) architecture, which is a novel feature extraction method proposed on the basis of regional local differences by Kyungjoong et al. [36]. Statistical analysis of the proposed method confirms the transformation of noise in the face image into uniform patterns. Compared with the conventional face detector methods, it performs better by obtaining binary code and encoding relative sizes of central region with its neighboring regions. By combining high level attributes and low level features in a face image, Christina et al. [37] proposed Secure Me, an android application providing accurate result while extraction of similar

faces. It also uses authentication techniques to provide security for the data storage. Huang et al. [38] proposed Local Vector Pattern, a feature with weighting mechanism to detect 'M' possible face candidates using local vector pattern scheme. Second stage uses feature-point Bilateral Recognition (BR) scheme to produce final result. BR contains both backward and forward recognition and it uses feature point detection. The experimental results proved that the proposed method performs better than Elastic Bunch Graph Matching (EBGM) [39] and Patterns of Oriented Edge Magnitudes (POEM) [40] and it can be applied in real time.

Juefie-Xu et al. [41] conducted a study on forensic face processing techniques for occlusion and cropping. They used commercial off-the-shelf (COTS) components to build the face recognition system. There exist many methods for face recognition in 2D images, but 3D face recognition approaches with variations in illumination and changes in facial expressions are investigated by Dhara Marvadi et al. [42]. Globally 2D-LDA and 2D-PCA approaches are employed compared to local PCA or LDA. Experiments proved that 2D-LDA based 3D face recognition outperforms 2D-PCA. This provides close to 96 to 99 percent recognition rate with EER of 8.96 percent. For handling pose and expression variations in the given image, a 3D Probabilistic Facial Expression Recognition Generic Elastic Model (3D PFER-GEM) is proposed by Moeini et al. [43]. This method reconstructs a 3D model from frontal human face and then iterative process is continued using Feature Library Matrix (FLM) to extract features of rotated images. The proposed method is experimented on CMU-PIE, LFW and FERET databases, doe pose-invariant and on Bosphorus database for both pose and expression variations. Demonstrated results prove that the method performs better than state-of-the-art approaches.

3. BACKGROUND WORK

Human face detection is carried out using various algorithms. Viola-Jones [11] algorithm being one such, extracts facial lineament of two dimensional human face image and detects the features on the occult as well as distant faces. This framework provides competitive object detection rates in real-time. The algorithm has mainly four stages: (1) Haar Feature Selection, (2) Creating Integral Image, (3) Adaboost Training algorithm and (4) Cascaded Classifiers.

The inbuilt functions of the algorithm are effectively run on the human face to find the prominent attributes. These generate bounded box matrix of size 1x4, having x, y coordinates, i.e. the starting point of detected attribute box, along with the width and the height of the box. The matrix stores the values dynamically with every new image run. Initially, this method creates Integral Image representation which computes the features and passes to detector. Secondly, efficient classier is developed by using a learning algorithm based on AdaBoost which selects less number of critical features. At last, a method for combining all the classifiers is described. Hence this framework detects objects with high accuracy by minimizing computational time.

4. PROBLEM DEFINITION

As the combination of detected objects yield better image processing, there exists Viola-Jones [11] algorithm which is popular and developed by Paul et al. Even though it is widely used, it still fails in achieving high detection rate since it detects non-facial objects also as facial objects and it wrongly identifies facial objects like

nose, mouth and eyes outside the facial area as well. Sometimes it detects more than the available facial objects in the given image which may create confusion during face recognition by matching.

Given the Profile Pictures of Social Network Users, recognizing user faces from different social network images requires front facial view of the users. The facial feature detection is restricted only to the detected frontal face area of the user to match with the profile pictures of them in multi-social networks.

5. PROPOSED SYSTEM MODEL

To correct false detection of facial features outside the facial area and to improve the accuracy of face detection, this work proposes an approach which accurately detects facial features in the given image and improves the overall efficiency of face detection and recognition. Using the proposed approach, both single and multiple faces are detected in the given image.

5.1 Facial Feature Extraction and their Bounded Box values

The proposed method is an improvisation made upon the existing Viola-Jones [11] framework. Initially, the face bounded box is generated by running the Viola-Jones Face Detection Function on the given image. The detected facial area is then bifurcated horizontally to detect the mouth in the lower half by running Viola-Jones Mouth Detection Function to generate mouth bounded box. Then, the upper half is bifurcated vertically to detect both right and left eyes accurately by the Viola-Jones Eye Detection Functions to generate respective bounded boxes. To effectively detect the nose, the facial area is cropped to an extent such that the nose area is obtained and by running the Viola-Jones Nose Detection Function, the nose bounded box is generated accurately. Here, each facial feature along with their respective bounded
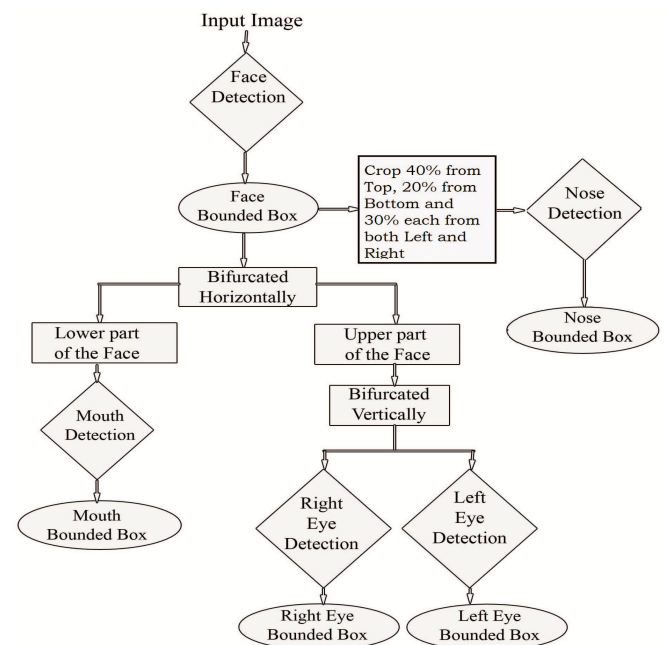

Fig. 1: Facial Feature Extraction Process

(a) Profile Picture (b) Detected Face (c) Horizontally Bifurcated Face to get Mouth Area

(d) Upper Face Vertically Bifurcated to get Left and Right Eye Areas (e) Face Cropped to get Nose Area (f) Facial Features with their BBs
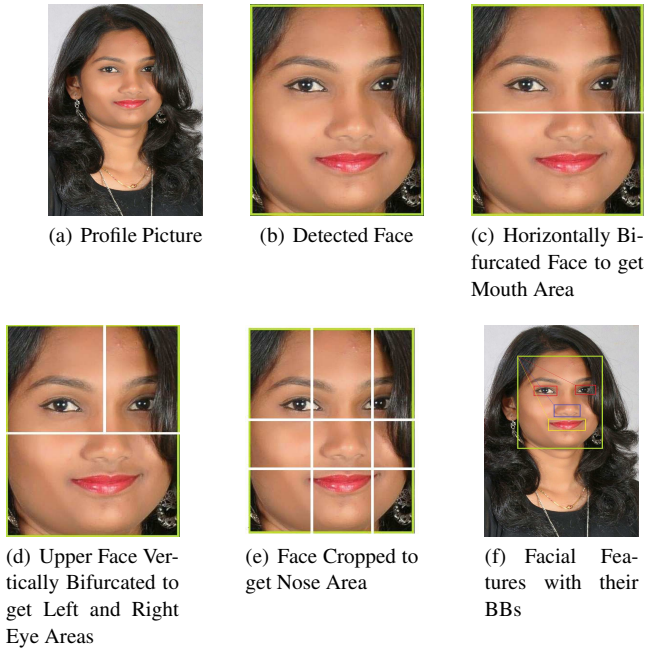
Fig. 2: Facial Feature Extraction Illustration

boxes are obtained individually. By this, the entire face is divided into parts and each object's bounded box value is obtained after respective object detection. The Facial Feature Extraction Process diagram shown in Fig. 1 clearly describes the bifurcation and cropping mechanisms followed during the process.

As already said, nose detection has been improved by first identifying and obtaining the nose area in the face. For this purpose, the considered facial area is cropped from top by 40 percent, whereas 20 percent from the bottom. Like wise, 30 percent respectively from both left and right side, cropping is done to obtain the nose area. This helps in avoiding false detection of nose in areas other than the face. Now the Nose Detection Function of Viola-Jones algorithm is run on the cropped nose area to detect the nose accurately and obtain appropriate nose bounded box.

Fig.2 illustrates the Facial Feature Extraction considering a sample input test image consisting of an user's profile picture as shown in Fig. 2a. The detected face after running the Face Detection Function is as shown in Fig. 2b. The detected face which is bifurcated horizontally for getting the mouth area in the lower portion is shown in Fig. 2c which is fed to Mouth Detection Function whereas the upper portion of face is bifurcated vertically for getting the areas of both the eyes is as shown in Fig. 2d which are then subjected to Left and Right Eye Detection Functions respectively. The detected face which is then cropped to get the nose area and subjected to Nose Detection Function is as shown in Fig. 2e. Finally, Fig. 2f shows the BBs of all the detected facial features in the subjected image.

5.2 Relative-Distance Method

The bounded box matrix generated during the algorithm run time is used in RD method. The x,y coordinate values of the bounded boxes of mouth, nose, left eye and right eye are used to calcu-

late the relative distance of each lineament with respect to that of the face. The ED between the BBs is computed to get LDEDM in RD method. The LDEDM so computed is used for the purpose of matching the identity of the user image from the database for any incoming image.
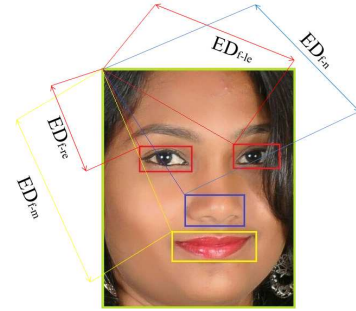

Fig. 3: Relative-Distance Method Illustration

The ED between x,y coordinates of face bounded box with that of the x,y coordinates of other features bounded boxes on the face as shown in Fig. 3 are computed and stored in a 2 X 2 Euclidian Distance Matrix (EDM) as per Equation (1).

$$EDM = \begin{bmatrix} ED_{f-m} & ED_{f-n} \\ ED_{f-le} & ED_{f-re} \end{bmatrix} \tag{1}$$

$$DEDM = |EDM| = [ED_{f-m} * ED_{f-re}] - [ED_{f-le} * ED_{f-n}] \tag{2}$$

$$LDEDM = log|EDM| \tag{3}$$

The Determinant of Euclidian Distance Matrix (DEDM) is computed and the logarithmic value of the determinant i.e. LDEDM of the image which is unique for every user is computed as per Equations (2) and (3) respectively and the LDEDM values of every user so computed are stored in the database.

5.3 System Architecture

An inverted image is subjected to the Viola-Jones algorithm [11] to check if the prominent attributes are being detected correctly. The test revealed an inaccurate detection of features for the inverted image. The issue is resolved in the proposed RDUP3 algorithm by rotating of the image based on the upper body feature detection on the image. The proposed RDUP3 algorithm is used further for the facial detection and recognition and proved to be accurate and efficient in facial attributes detection.

The basic information of the users collected and stored in the database for each individual is displayed on successful result of matching component of the proposed RDUP3 algorithm which checks the computed LDEDM of the current input image with the values stored in the database for the match. The absolute difference value between the generated LDEDM and the stored LDEDM is defined within the range for a successful match as shown in Fig.4.
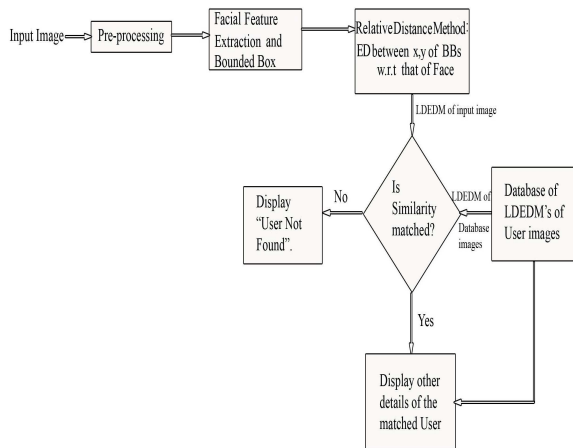
Fig. 4: System Architecture Diagram

6. ALGORITHM

Considering one image at a time, facial features in the image are detected. If the image is inverted, then it is rotated upright to locate the face. The LDEDM of the considered image is stored in database using LDEDM-Function as shown in Algorithm 1.

Algorithm 1 *LDEDM Function*

1: **while** True **do**
2:　　**for** Every Profile Image in the Network **do**
3:　　　　**if** Upper body detected **then**
4:　　　　　　goto Detection
5:　　　　**else**
6:　　　　　　Rotate the Image
7:　　　　**end if**
8:　　　　Detection: Detect the Face and Compute its BB value.
9:　　　　Consider the detected Face BB and Crop it 40 percent -from Top, 20 percent from Bottom and 30 percent each -from Left and Right.
10:　　　　Detect the Nose and compute its BB value in the -cropped nose area.
11:　　　　Consider the detected Face BB and divide it hori -zontally at the middle.
12:　　　　Detect the Mouth in the lower half of the division and -compute its BB values.
13:　　　　Bisect the upper half of the divided image vertically.
14:　　　　Detect the Right Eye on left half and Left Eye on right -half of bisected image and compute their BB values.
15:　　　　**for** Every BB != that of the Face **do**
16:　　　　　　Calculate the Euclidean Distance between x, y -coordinates of the BB and that of the Face and -Record in a 2 X 2 Euclidean Distance Matrix.
17:　　　　**end for**
18:　　　　Compute the Logarithmic of Determinant of Euclidean -Distance Matrix (LDEDM) and store it.
19:　　**end for**
20: **end while**

Algorithm 2 *Relative Distance based User Profiling from Profile Picture (RDUP3) Algorithm*

1: **while** True **do**
2:　　**for** Every Test Image Considered **do**
3:　　　　Compute LDEDM using LDEDM-Function
4:　　　　**for** Every LDEDM of computed data set **do**
5:　　　　　　Find the Absolute Difference between LDEDM of -Test Image and that of data set.
6:　　　　　　**if** Absolute Difference is within defined range **then**
7:　　　　　　　　Display Other details of the Matched User -correspondingly.
8:　　　　　　**else**
9:　　　　　　　　Display "User Not Found"
10:　　　　　　**end if**
11:　　　　**end for**
12:　　**end for**
13: **end while**

Each test image considered is fed to the proposed RDUP3 algorithm shown in Algorithm 2. The LDEDM value of the image under consideration is computed using the LDEDM-Function shown in Algorithm 1. Preprocessing of the image under consideration is done to make sure it is not an inverted one and that it is the frontal face image. The LDEDM of test image is compared with that of the LDEDMs of dataset images. If the absolute difference between those two LDEDMs fall within the threshold range, then the algorithm outputs all other details of the Matched User found in the database, else User Not Found message will be displayed.

7. SIMULATION RESULTS AND PERFORMANCE ANALYSIS

The proposed RDUP3 algorithm is implemented with MATLAB 2013a in a system having Intel Pentium i7 with 4GB RAM configuration, upon Windows-8 platform. Initially, we constructed a database by collecting 500 User's details along with their Profile Pictures from Linked-In social networking site. For each collected Profile Picture, LDEDM is computed and stored in the database. The database created consists of Name of the User in the image, User image and its LDEDM, Highest Education of the User and their University/College name.

7.1 THRESHOLD SETTING

For setting-up of Threshold, 50 User Profile Pictures were considered, out of which 40 Profile Pictures of the same users appearing in Linked-In are collected from the Facebook. Other 10 pictures were collected from outside the social networking sites which do not appear in constructed database as well.

The performance of the proposed RDUP3 algorithm for the test profile pictures is analyzed based on the following measures:

i) False Acceptance Rate (FAR) is the measure that incorrectly/falsely matches an image to a wrong user in the database.

$$FAR = \frac{\text{Number of Falsely Accepted Images}}{\text{Total Number of User Images in Database}}$$

ii) False Rejection Rate (FRR) is the measure that incorrectly/falsely rejects an image of the right user in the database.

$$FRR = \frac{\text{Number of Falsely Rejected Images}}{\text{Total Number of User Images in Database}}$$

iii) Equal Error Rate (EER) is a threshold independent performance measure obtained at the point where both FAR and FRR values are same.

The proposed RDUP3 algorithm is run on each picture under consideration for testing upon varying threshold from 0.0 to 1.0 and the output of the algorithm is recorded. Based on the output of the algorithm, the FAR and FRR values are computed.
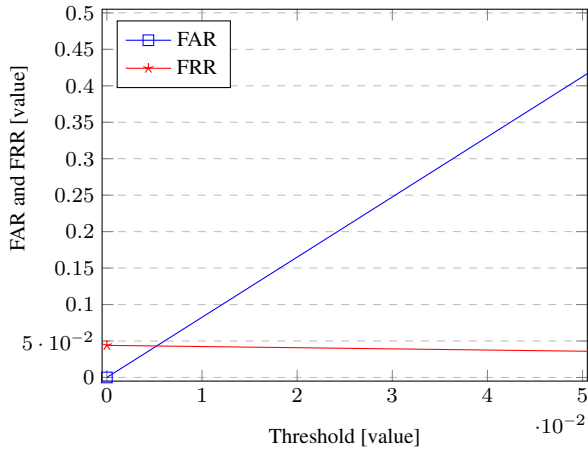
Fig. 5: Equal Error Rate of the System.

Fig.5 Shows the FAR and FRR values against the threshold considered during testing. The EER value 0.043 obtained from the testing shown in Fig.5 depicts better performance of the proposed RDUP3 algorithm as it is very low at 0.05 threshold.

7.2 RESULT ANALYSIS

The important drawback of Viola-Jones [11] algorithm is that it fails in multi-face detection by detecting more than the available faces in given image and by falsely identifying non-facial objects as facial objects both inside and outside the face area as shown in Fig. 6.
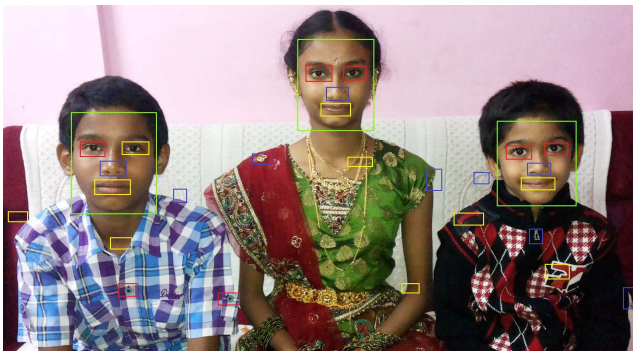

Fig. 6: Facial Features Detected using Viola-Jones [11] Algorithm.

The proposed RDUP3 algorithm is applied on both single-faced and multi-faced images which detected all the facial features accurately at their respective places in all the faces by following iterative loop process as shown in Fig. 7. Thus, proposed RDUP3 algorithm detects all the facial features accurately in both single and multi-faced images and the bounded box values of facial features are stored for each user face individually.

Fig. 7: Facial Features Detected using Proposed RDUP3 Algorithm.

While computing the LDEDM of profile pictures from the constructed data set of 500 images, it is observed that the proposed RDUP3 algorithm performed well with 100 percent accuracy in Face, Eyes and Mouth detection whereas 82 percent accuracy for nose detection.

Table 1. : Average Detection Accuracies of Face, Eyes, Mouth and Nose along with Time Consumed per Image

Average	Viola-Jones [11]	Hybrid Design [14]	Proposed RDUP3
Face Detection Accuracy (in percentage)	100	99	100
Eyes Detection Accuracy (in percentage)	65	99.5	100
Mouth Detection Accuracy (in percentage)	57	98	100
Nose Detection Accuracy (in percentage)	42	60	82
Time Consumed per Image (in seconds)	1.467	0.907	1.053

The comparison of the proposed RDUP3 algorithm with Viola-Jones [11] and Hybrid Design [14] is as shown in Table. 1, which depicts that though the accuracy of Viola-Jones [11] in detecting face is excellent, but when it comes to the detection of eyes, mouth and nose, it loses in achieving accuracy as well as in time consumption. On the other hand, though the Hybrid Design [14] consumes less time, but the detection of face, mouth, eyes and nose is outperformed by the proposed RDUP3 algorithm. Thus the proposed RDUP3 algorithm is proved to be robust and efficient in Face Detection and Recognition at a Threshold of 0.05, achieving 99.957 percent accuracy for User Profiling in Multi-Social Networking.

8. CONCLUSIONS

The paper presents RDUP3 algorithm for Profiling Online Social Network Users. The algorithm is evaluated and a novel approach is proposed which improves the effectiveness and efficiency of face detection by bifurcating and cropping the detected face. The LDEDM computed during the process is unique for the user image under consideration and is used for the purpose of matching and recognizing the identity of the user images from the database. Simulation results show that the EER is considerably low with the proposed RDUP3 algorithm indicating better performance, while overcoming the drawbacks of State-of-the-Art works.

The RDUP3 algorithm when applied on Multi-Social Networking sites, helps in identifying the criminals upon matching with the criminal database. This approach is also suitable in recruiting process to cross validate the candidature of applicants through their Online Social Network Status and Activities. The avenues for future work are in carving the User Profiles with their Multi-Social Networking activities through one single platform by identifying the same user in different social media.

9. REFERENCES

[1] P Deepa Shenoy, Srinivasa K G, Venugopal K R and Lalit M Patnaik, "Evolutionary approach for mining association rules on dynamic databases," *Advances in Knowledge Discovery and Data Mining*, pp. 325–336, April 2003.

[2] P Deepa Shenoy, Srinivasa K G, Venugopal K R and Lalit M Patnaik, "Dynamic association rule mining using genetic algorithms," *Intelligent Data Analysis*, vol. 9, no. 5, pp. 439–453, September 2005.

[3] Vasanthakumar G U, P Deepa Shenoy and Venugopal K R, "PFU: Profiling Forum users in online social networks, a knowledge driven data mining approach," *IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, pp. 57-60, December 2015.

[4] Vasanthakumar G U, P Deepa Shenoy and Venugopal K R, "PTIB: Profiling Top Influential Blogger in Online Social Networks," *International Journal of Information Processing (IJIP-2016), IK International Publishing*, vol. 10, no. 1, pp. 77–91, March 2016.

[5] Vasanthakumar G U, Priyanka R, Vanitha Raj K C, Bhavani S, Asha Rani B R, P Deepa Shenoy and Venugopal K R, "PTMIB: Profiling Top Most Influential Blogger using Content Based Data Mining Approach," *IEEE International Conference on Data Science and Engineering (ICDSE-2016)*, Cochin, India, August 2016.

[6] Ramachandra A C, Pavithra K, Yashasvini K, Raja K B, Venugopal K R and Lalit M Patnaik, "Cross-validation for graph matching based offline signature verification," *IEEE Annual India Conference, INDICON 2008*, vol. 1, pp. 17–22, 2008.

[7] Vasanthakumar G U, Aakriti Kumari Upadhyay, Pradeep F Kalmath, Sthita Dinakar, P Deepa Shenoy and Venugopal K R, "UP3: User profiling from Profile Picture in Multi-Social Networking," *Annual IEEE India Conference (INDICON)*, pp. 1-6, December 2015.

[8] Veena H Bhat, Prashanth G Rao, Abhilash, P Deepa Shenoy, Venugopal K R and L M Patnaik, "A Novel Data Generation Approach for Digital Forensic Application In Data Mining," *IEEE Second International Conference on Machine Learning and Computing*, February 2010.

[9] Kang-Seo Park,Young-Gon Kim and Rae-Hong Park, "Face Detection Using The 33 Block Rank Patterns Of Gradient Magnitude Images," *Signal and Image Processing : An International Journal (SIPIJ)*, vol. 4, no. 5, October 2013.

[10] Kailash Devrari and K.Vinay Kumar, "Fast Face Detection Using Graphics Processor," *(IJCSIT) International Journal of Computer Science and Information Technologies*, vol. 2, no. 3, pp. 1082–1086, 2011.

[11] Paul Viola and Michael Jones, "Robust Real-time Object Detection," *Second International Workshop On Statistical And Computational Theories Of Vision Modeling, Learning, Computing, and Sampling*, July 2001.

[12] Arundhati Das, Mameeta Pukhrambam and Ashim Saha, "Real-Time Robust Face Detection and Tracking using extended Haar functions and improved Boosting Algorithm," *IEEE International Conference on Green Computing and Internet of Things (ICGCIoT)*, pp. 981-985, 2015.

[13] Jianbo Shi and Carlo Tomasi, "Good features to track," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'94)*, pp. 593-600, 1994.

[14] Ijaz Khan, Hadi Abdullah and Mohd Shamian Bin Zainal, "Efficient Eyes and Mouth Detection Algorithm using Combination of Viola Jones and Skin Color Pixel Detection," *International Journal of Engineering and Applied Sciences*, vol. 3, no. 4, June 2013.

[15] Himanshu Sharma, Saurav Sumeet, Sanjay Singh, Anil K. Saini and Ravi Saini, "Analyzing impact of image scaling algorithms on viola-jones face detection framework," *IEEE International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 1715-1718, 2015.

[16] Chung-chi Lin, Ming-hwa Sheu, Huann-keng Chiang, Wenkai Tsai and Zeng-chuan Wu, "Real-time FPGA architecture of extended linear convolution for digital image scaling," *IEEE International Conference on ICECE Technology*, pp. 381-384, 2008.

[17] Parker J. Anthony, Robert V. Kenyon and Donald E. Troxel, "Comparison of interpolating methods for image resampling," *IEEE Transactions on medical imaging*, vol. 2, no. 1, pp. 31-39, 1983.

[18] Hou Hsieh and H. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 26, no. 6, pp. 508-517, 1978.

[19] Lehmann, Thomas Martin, Claudia Gonner and Klaus Spitzer, "Survey: Interpolation methods in medical image processing," *IEEE transactions on medical imaging*, vol. 18, no. 11, pp. 1049-1075, 1999.

[20] Gour, Pranav Narayan, Sujay Narumanchi, Sumeet Saurav and Sanjay Singh, "Hardware accelerator for real-time image resizing," *IEEE 18th International Symposium on VLSI Design and Test*, pp. 1-6, 2014.

[21] Wei Liuliu and Liu Mingyang, "Multi-pose Face Detection Research Based on Adaboost," *IEEE Eighth International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, pp. 409-412, 2016.

[22] Yi-Qing Wang, "An Analysis of the Viola-Jones Face Detection Algorithm," *Image Processing On Line*, pp. 128-148, 2014.

[23] Adrian Wong Yoong Wai, Shahirina Mohd Tahir and Yoong Choon Chang, "GPU acceleration of real time Viola-Jones face

detection," *IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, pp. 183-188, 2015.

[24] R.Sureshkumar and N.Arthi, "Generate Attribute-Enhanced Sparse Codewords To Retrieve Image From Large Image Database," *International Journal of Engineering Science Invention*, vol. 2, no. 1, pp. 2319–6726, October 2013.

[25] Seyed Mohammad Hassan Anvar,Wei-Yun Yau and Eam Khwang Teoh, "Multiview Face Detection and Registration Requiring Minimal Manual Intervention," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 10, October 2013.

[26] Raphael Sznitman and Bruno Jedynak, "Active Testing for Face Detection and Localization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 10, October 2010.

[27] Hongliang Li, King N. Ngan and Qiang Liu, "FaceSeg: Automatic Face Segmentation for Real-Time Video," *IEEE Transactions on Multimedia*, vol. 11, no. 1, pp. 77-88, 2009.

[28] Hatice Gunes and Massimo Piccardi, "Automatic Temporal Segment Detection and Affect Recognition From Face and Body Display," *IEEE Transactions on Systems, Man and CyberneticsPart B: Cybernetics*, vol. 39, no. 1, February 2009.

[29] Mauricio Pamplona Segundo, Luciano Silva, Olga Regina Pereira Bellon and Chaua C. Queirolo, "Automatic Face Segmentation and Facial Landmark Detection in Range Images," *IEEE Transactions on Systems, Man and CyberneticsPart B: Cybernetics*, vol. 40, no. 5, October 2010.

[30] Jie Pan, Xue-Song Wang and Yu-Hu Cheng, "Single-sample Face Recognition Based on LPP Feature Transfer," *IEEE*, 2016.

[31] M. Ko and A. Barkana, "A new solution to one sample problem in face recognition using FLDA," *Applied Mathematics and Computation*, vol. 217, no. 24, pp. 10368-10376, August 2011.

[32] J. Wu and Z. H. Zhou, "Face recognition with one training image per person," *Pattern Recognition Letters*, vol. 23, no. 14, pp. 1711-1719. December 2002.

[33] Chih-Rung Chen, Wei-Su Wong and Ching-Te Chiu, "A 0.64mm RealTime Cascade Face Detection Design Based on Reduced Two-Field Extraction," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 19, no. 11, November 2011.

[34] Mandeep Kaur, Rajeev Vashisht and Nirvair Neeru, "Recognition of Facial Expressions with Principal Component Analysis and Singular Value Decomposition," *International Journal of Computer Applications*, vol. 9, no. 12, pp. 36-40, November 2010.

[35] Muwei Jian and Kin-Man Lam, "Simultaneous Hallucination and Recognition of Low-Resolution Faces Based on Singular Value Decomposition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 11, pp. 1761-1772, 2015.

[36] Kyungjoong Jeong, Jaesik Choi and Gil-Jin Jang, "Semi-Local Structure Patterns for Robust Face Detection," *IEEE Signal Processing Letters*, vol. 22, no. 9, September-2015.

[37] Christina Joy, Roshlin Anie Abraham and Raji, "A Survey on Face Matching and Retrieval of Images," *International Journal of Computer Science and Mobile Computing*, vol. 4, no. 2, pp. 33–37, February-2015.

[38] Yea-Shuan Huang and Suen-Yu Chen, "A geometrical-model-based face recognition," *IEEE International Conference on Image Processing (ICIP)*, pp. 3106-3110, 2015.

[39] Wiskott L, Fellous J M, Kuiger N and von der Malsburg C, "Face Recognition by Elastic Bunch Graph Matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19 , no. 7, pp. 775-779, July 1997.

[40] Ngoc-Son Vu and Alice Caplier, "Face Recognition with Patterns of Oriented Edge Magnitudes," *European conference on computer vision (ECCV), Springer Berlin Heidelberg*, pp. 313326, 2010.

[41] Felix Juefei-Xu, Dipan K. Pal, Karanhaar Singh and Marios Savvides, "A Preliminary Investigation on the Sensitivity of COTS Face Recognition Systems to Forensic Analyst-style Face Processing for Occlusions," *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 25-33, 2015.

[42] Dhara Marvadi, Maulin Joshi, Chirag Paunwala and Aarohi Vora, "Comparative Analysis of 3D Face Recognition Using 2D-PCA and 2D-LDA Approaches," *IEEE 5th Nirma University International Conference on Engineering (NUiCONE)*, pp. 1-5, 2015.

[43] Ali Moeini and Hossein Moeini, "Real-World and Rapid Face Recognition towards Pose and Expression Variations via Feature Library Matrix," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 5, pp. 969-984, 2015.