Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 167 (2020) 562–570

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Computational Intelligence and Data Science (ICCIDS 2019)

Feature-based Sketch-Photo Matching for Face Recognition

Sahil Dalal[a,*], Virendra P. Vishwakarma[a] and Sanchit Kumar[b]

[a]University School of Information, Communication & Technology
Guru Gobind Singh Indraprastha University, Sector 16-C, Dwarka, New Delhi, India
[b]Department of Electronics & Communication Engineering
Delhi Technological University, Sector-17, Rohini, Delhi, India

Abstract

Matching of sketch and photo is very important in police verification and intelligence as it is used to track the criminals or some person. Therefore, it should be accurate so that mismatching of the sketch with photo can be avoided. For this, feature based matching method is introduced through this paper. Feature based matching includes a feature vector which contains features of the face image (either it is a sketch or a photo) in terms of histogram of oriented gradients (HoG) features and gray level co-occurrence matrix (GLCM) features. Computing the features, first, increases the chances of correct matches. It can be depicted from the results as well that, comparing with other state-of-art approaches, proposed method is more accurate in matching the sketches with photos of the same person.

Keywords: Sketch; Photo; HoG; GLCM; Matching.

1. Introduction

A sketch is a design that is made quickly without a great deal of information. Artists frequently use sketches as a preparation for more detailed painting or drawing. Sketching techniques are used in law enforcement and are regularly utilized to distinguish accused from an observer recall. These traditional methods of recognizing are

* Corresponding author. Tel.: +91-880-017-8730.
E-mail address: dalalsahil22@yahoo.co.in

10.1016/j.procs.2020.03.318

mostly hand-worked, meticulous and may not lead to capture the correct guilty party. When the sketch is ready, it is matched with the photos available with the department to recognize the suspect. Photo is available only if that person is convicted at least once. In earlier times, this recognition was done manually which is a time-consuming process, then automatic recognition techniques came into existence. Many automatic recognition techniques have been introduced so far and it has been tried to improve these techniques through the work mentioned.
Work on automatic face sketch recognition systems were started by Uhl et al. in 1994 who proposed a method based on Principle Component Analysis (PCA) [1]. After that most of the work in this field was done by Tang and Wang. An automatic method for recognizing photos from a database using a sketch had been introduced in which a sketch was not matched with a photo directly as there was a great difference between a sketch and real photo in terms of shape and texture. So the database was first converted from the photos to sketches using Eigenface method which itself used Karhunen-Loeve Transform. It also created feature vector for the sketches in the database and the testing image. Then, the feature vectors of sketch and face are matched [2]. Han et al. gave an automatic face- sketch recognition system using a component based representation (CBR) [3]. Klum et al. gave a system in which the forensic sketches (made by sketch artist) and composite sketches (made using FACES software) both are used for facial recognition with original photos taken using two methods: (a) Holistic method (b) Component method. Problem with the usage of composite sketches is that sketches design depends totally on the person who is using the software to design them [4]. Liu introduced an automatic face-sketch recognition system in which photos from two different modalities (i.e. Photos and Sketches) were matched directly using Joint dictionary learning. Both the photos and sketches were jointly trained and then the coefficients of the tested data were calculated separately and compared using Cosine Distance. The proposed method was also compared with the previously proposed methods [5]. Roy and Bhattacharjee introduced the automatic face sketch photo matching system utilizing the Local Gradient Fuzzy Pattern (LGFP). Multiscale LGFP method was also tested in the paper and the two proposed methods i.e. single scaled LGFP and Multi scaled LGFP were compared with many previously introduced methods. Method was time consuming and was having high computational cost [6]. Wang et al. gave a synthesized method for sketch to photo matching based on Anchored Neighborhood Index (ANI) [7]. Galea and Farrugia gave an automatic intermodility (i.e. photos of two different modularities like original photo and sketch) face recognition system based on Deep Learning Based Architecture [8]. Setumin and Suandi proposed an automatic face sketch to photo matching method using Difference of Gaussian Oriented Gradient Histogram (DoGOGH) for Face Sketch to Photo Matching. The effectiveness of DoGOGH was evaluated and compared with several popular local descriptors (i.e., MLBP, SIFT, SURF, and HOG) [9]. Chugh et al. introduced a composite face sketch and photo matching method using Transfer Learning based Evolutionary Algorithm. The features were extracted using two feature descriptors methods: one is Histogram of Oriented Gradient (HOG) and other was proposed in the paper itself i.e. Histogram of Image Moments (HIM) descriptor for both the source and target domain. Then these feature vectors were given to the target domain and features of both the domains were matched utilizing genetic learning [10]. Patil and Shubhangi made a matching system using Geometrical Face Model. Face region was identified using AdaBoost algorithm and then facial main components like eyes, nose, mouth etc. were marked using geometrical structure of face. Then from each facial component a texture feature was extracted using Weber Local Descriptor (WLD). Artificial neural network (ANN) was used for classification [11]. Li et al. introduced the system based on generating photorealistic faces guided by Descriptive Attributes. The photorealistic face was generated by the generation algorithm i.e. multi-modal conditional generative adversarial network (MMC-GAN). Matching process was conducted by computing Cosine distance of the representation for face feature matching [12]. Rajput et al. proposed an automated face photo-sketch using HOG descriptors and matched using K-Nearest Neighbours (K-NN) algorithm [13]. Zhang et al. also gave an automatic and effective photo to sketch synthesizing method in 2019 based on Dual-transfer i.e. inter domain transfer and intra domain transfer. This method is very complex and has high computational cost [14].

In this paper, an effective feature-based face recognition system has been introduced using hybrid of two descriptors i.e. histogram of oriented gradient (HOG) descriptors and gray level co-occurrence matrix (GLCM) descriptors so that limitation of one can be overcome by other. The proposed paper is organised in the following way: Section 2 tells about the techniques used for features extraction and matching of the sketched with their respective photos. Section 3 gives a discussion on the experimental results followed by conclusion in section 4.

2. Feature Extraction and Matching

The steps involved in the methods are: (1) Preprocessing of both sketches and photos. (2) Feature extraction of both the images using hybrid of HOG descriptors [15] and GLCM descriptors [16]. (3) Matching of sketch to photo using Euclidean distance. The block diagram of the proposed method has been given in Fig. 1.

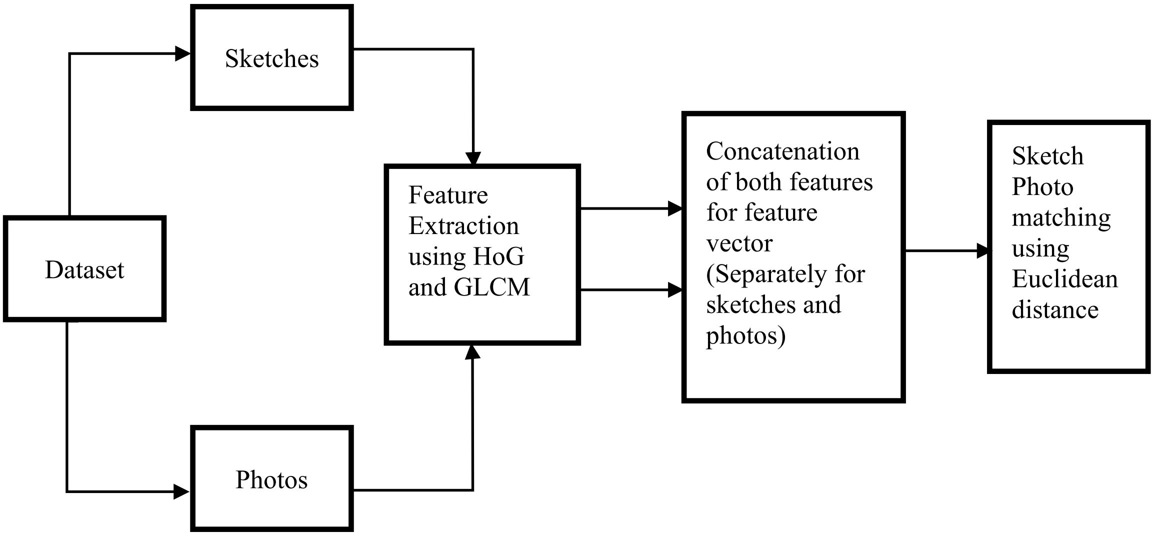

Fig. 1. Block diagram of the proposed method

2.1. *Preprocessing of face images*

Preprocessing is one of the pivotal advances in image processing. It not only modifies the pictures into the alluring areas yet additionally expels pointless data like noise from the pictures to give better outcomes. Here, preprocessing comprises of conversion of images from RGB (coloured) to gray (Black and White) and removal of noise from images using median filter. The dataset consists of photos and sketches having more than 2 gray levels i.e. white and black. So, both photos and sketches are converted to gray images to make their gray levels into the same domain. Then, median filtering is utilized to eliminate noise from the images while preserving the edges. So, after transformation to gray images both sketch and photograph are processed with 2-D median filter which expels noise from the images while protecting their edges. It helps in giving great outcomes as the edges are saved and edges are extremely useful in matching sketch and photographs.

2.2. *Feature extraction*

It is the way towards characterizing a set of feature which will represent the data most productively that is essential for investigation and classification. Features are computed using two different techniques i.e. HOG descriptors and GLCM descriptors. Then features from both the techniques are concatenated, so that a feature vector can be formed, that will further be used for matching of sketch to photo.

2.2.1. *Feature extraction using Histogram of Oriented gradient (HoG)*

The HOG features are calculated for every image and hence feature vector for that image will be formed for matching purpose. The HOG features are extracted to compute the difference of the pixel values in m and n directions as shown in the equations:

$$f_m(m,n) = f(m+1,n) - f(m-1,n)$$
$$f_n(m,n) = f(m,n+1) - f(m,n-1) \tag{1}$$

where, $f(m,n)$ is the pixel intensity value of the image for the pixel (m,n). After that the magnitude (Arg) and direction (θ) are calculated:

$$Arg(m,n) = \sqrt{f_m(m,n)^2 + f_n(m,n)^2} \tag{2}$$

$$\theta(m,n) = arctan\frac{f_m(m,n)}{f_n(m,n)} \tag{3}$$

The range of values obtained in equation (2) and (3) is wide because of the variation illumination conditions and background of the images. Hence, normalization should be done for the accurate matching after the feature extraction. This is done by L2-norm,

$$\nu \rightarrow \frac{\nu}{\sqrt{\|\nu\|_2^2+\varepsilon^2}} \tag{4}$$

Where ε is a small constant restraining division by zero which is generally 1. ν is normalized vector.

2.2.2. *Feature extraction using Gray Level Co-Occurrence Matrix (GLCM)*

GLCM is one of the techniques which are used to extract information of texture of an image. Texture is property of the surface which is defined as something that consists of mutually related elements. By repetitive appearance of primitive elements either periodically or semi-periodically, texture is obtained. Texture provides very important information about roughness or smoothness of surface and also gives regularity in the surface. Advantage of computing GLCM is that it not only gives the variation in the pixel values but also provides relative positions of the pixels. Say, there is a position operator '*Q*' in GLCM which helps in providing relative position of the pixels. To generate GLCM feature a matrix '*X*' of dimension $M \times M$ is generated using constraint specified by the position operator '*R*' where '*M*' is the gray level values in the original image. Elements '*x*' in matrix '*X*' indicate number of times points with intensity value z_j occur at a position determined by '*R*' in relation to points with intensity z_i. Let us assume an image '*I*' whose co-occurrence matrix has to be calculated and position operator '*R*' is defined as one pixels to right. Image '*I*' has only 3 different intensity levels i.e. $z_0 = 0, z_1 = 1 \; and \; z_2 = 2$ which will make matrix '*X*' of dimension 3 × 3.

$$I = \begin{bmatrix} 0 & 0 & 0 & 1 & 2 \\ 1 & 1 & 0 & 1 & 1 \\ 2 & 2 & 1 & 0 & 0 \\ 1 & 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

Now matrix '*X*' is calculated from matrix '*I*' following position operator '*R*'.

$$X = \begin{bmatrix} 4 & 4 & 1 \\ 4 & 3 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

If sum of the elements of matrix '*X*' is considered, which comes to be 20 and then divide matrix '*X*' with 20 then the matrix formed i.e. '*Q*' is the co-occurrence matrix which is used to calculate features.

$$Q = \frac{1}{20}[X]$$

Various parameters of GLCM are calculated using the equation's given below which are further used to calculate features:

$$\mu_x = \sum_{i=0}^{M-1} i \, Q_i(i) \text{ and } \mu_y = \sum_{j=0}^{M-1} j Q_j(j) \tag{5}$$

$$\sigma_x^2 = \sum_{i=0}^{M-1} (Q_x(i) - \mu_x(i))^2 \tag{6}$$

$$\sigma_y^2 = \sum_{i=0}^{M-1} (Q_y(j) - \mu_y(j))^2 \tag{7}$$

where *M* is the number of gray level in the original image, μ is the mean of *Q* and $\mu_x, \mu_y, \sigma_x \; and \; \sigma_y$ are the means and standard deviations of $Q_x \; and \; Q_y$ respectively. $Q_x(i)$ is the i_{th} entry obtained by adding the rows of $Q(i, j)$. By using the above equations, different texture features can be calculated for classification or matching purposes. The details of the GLCM features are computed in the form of four parameters. These are as follows: Contrast, correlation, energy and homogeneity.

- Contrast gives the difference in the luminance or color which makes an image distinguishable. Contrast is calculated using the following equation:

$$\text{Contrast} = \sum_{i,j=0}^{M-1} (Q_{i,j}(i-j))^2 \tag{8}$$

Contrast increases with increase in the value of (*i-j*) and there is no contrast if *i* becomes equal to *j* i.e. *i-j*=0 and it decreases with decreasing value of (*i-j*).

- Correlation gives mutual connection between two pixels in an image. For the sake of a perfectly positively or negatively associated picture, the correlation value is 1 and - 1. For the pictures with constant pattern of pixels, its value is NaN. Its range of value is [-1,1] and it is calculated using the following equation:

$$\text{Correlation} = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \frac{\{i,j\} \times Q(i,j) - \{\mu_x \times \mu_y\}}{\sigma_x \times \sigma_y} \tag{9}$$

- Energy is defined as the sum square of the pixels in an image. Its range is [0 1]. For image with constant pattern of pixels, its value is 1. It can be calculated using the following equation:

$$\text{Energy} = \sum_{i,j=0}^{M-1} \left(Q(i,j)\right)^2 \tag{10}$$

- Homogeneity gives the calculation of the same kind of pixels in an image. It can be calculated by using the following equation:

$$\text{Homogeneity} = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \frac{1}{1+(i-j)^2} Q(i,j) \tag{11}$$

Features calculated using both HOG descriptors and GLCM descriptors are then concatenated to form a single feature vector for matching purpose.

2.3. Matching of sketch with photo

It is troublesome and inefficient to match sketch with photograph directly, that is the reason feature extraction becomes possibly the most important factor. Features help in giving great outcomes as well as it diminishes the time utilization and complexity cost.

Feature vector calculated from the above techniques are used to match sketch with correct photo and Euclidean Distance has been used for that purpose as it is the simplest way of matching two photos with minimum complexity cost. Euclidean Distance can be calculated using the following equation:

$$\text{Distance} = \sqrt{\sum_{i=1}^{n} (v_{si} - v_{pi})^2} \tag{12}$$

where v_s and v_p are sketch feature vector and photo feature vector respectively and *n* is the number of features in both the feature vector [17].

Then the feature vectors having minimum distance are considered to be matched and photo corresponding to given sketch having minimum distance has been declared as the correct match.

3. Experimental Results and Analysis

This method has been implemented using MATLAB (Math Works Inc.) software tool, the numerical computing environment and programming language software for modeling complex algorithms. MATLAB is a problem-oriented language that allows performing computationally complex tasks quicker than the other programming language like C and C++.

This method has been implemented using MATLAB (Math Works Inc.) software tool, the numerical computing

environment and programming language software for modeling complex algorithms. MATLAB is a problem-oriented language that allows performing computationally complex tasks quicker than the other programming language like C and C++.

3.1. Database Used

Viewed sketches have been used for the experiment. Viewed sketches are available in an open store (accessible free on the web), from where the dataset is downloaded and tested. The viewed sketches, as a cluster of sketch-photographic sets, were gathered from CUHK face sketch database [2] where 188 sets were gathered. Therefore, 188 sets of viewed sketches are there which are accessible for experimentation. Along with the sketches, the photographic sets of the relating sketches are present for the image matching.

Database Used

Viewed sketches have been used for the experiment. Viewed sketches are available in an open store (accessible free on the web), from where the dataset is downloaded and tested. The viewed sketches, as a cluster of sketch-photographic sets, were gathered from CUHK face sketch database [2] where 188 sets were gathered. Therefore, 188 sets of viewed sketches are there which are accessible for experimentation. Along with the sketches, the photographic sets of the relating sketches are present for the image matching.

Fig. 2. Viewed Sketch and its corresponding photograph

3.2. Preprocessing

The images available in database i.e. both sketches and photos are already cropped and having size 250×200. As both the images are not in the same domain of intensity, coloured images have been converted into grayscale images as shown in Fig. 3. It also gives the output after processing images with media filter.

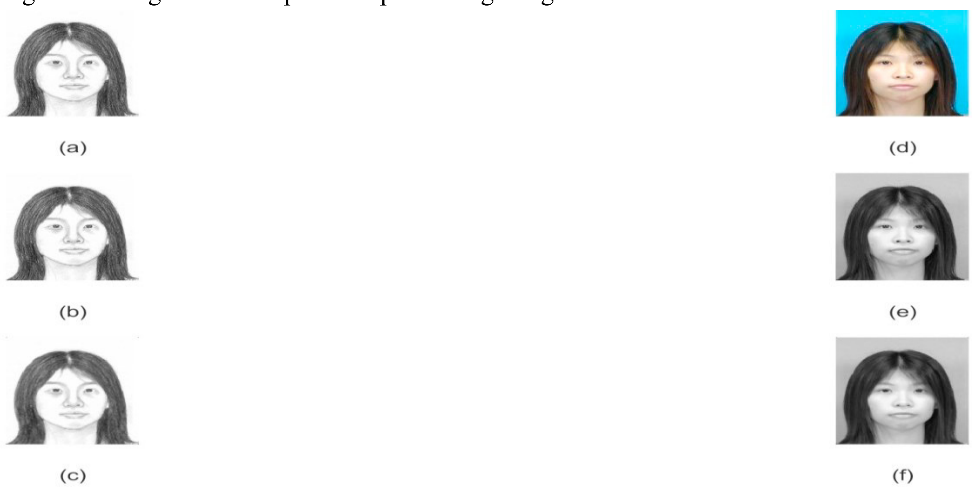

Fig. 3. (a), (b) and (c) are showing original sketch, grayscale converted sketch and median filtered image respectively and (d), (e) and (f) showing original photo, grayscale converted photo and median filtered image respectively.

3.3. Feature Calculation

Features have been calculated using two different feature descriptors i.e. HOG and GLCM descriptors and features calculated have been discussed as follows:

3.3.1. HOG features

Hog features in our experiment have been calculated block wise by using different block sizes i.e. 32x32, 16x16 and 8x8 block sizes respectively for both the sketches and photos which are giving different matching accuracies which will be discussed in the next section and it has been found that maximum matching accuracy has been given with 8x8 block size. Number of features calculated for different block sizes are shown in Table 1.

Table 1. HOG features calculated for various block size

Block Size	32×32	16×16	8×8
Number of features	1,080	5,544	25,920

3.3.2. GLCM features

GLCM features gives information of the texture of an image and in a face image, the texture varies mainly in key features of the face e.g. eyes, nose and mouth. So, in our experiment key features have been extracted first from the images and then GLCM features discussed in section 3.2.2 (b) have been calculated. 3 key features from the face i.e. eyes, nose and mouth have been extracted as shown in Fig. 3 and 4 features for each key feature have been calculated which gives a feature vector of total 12 features.

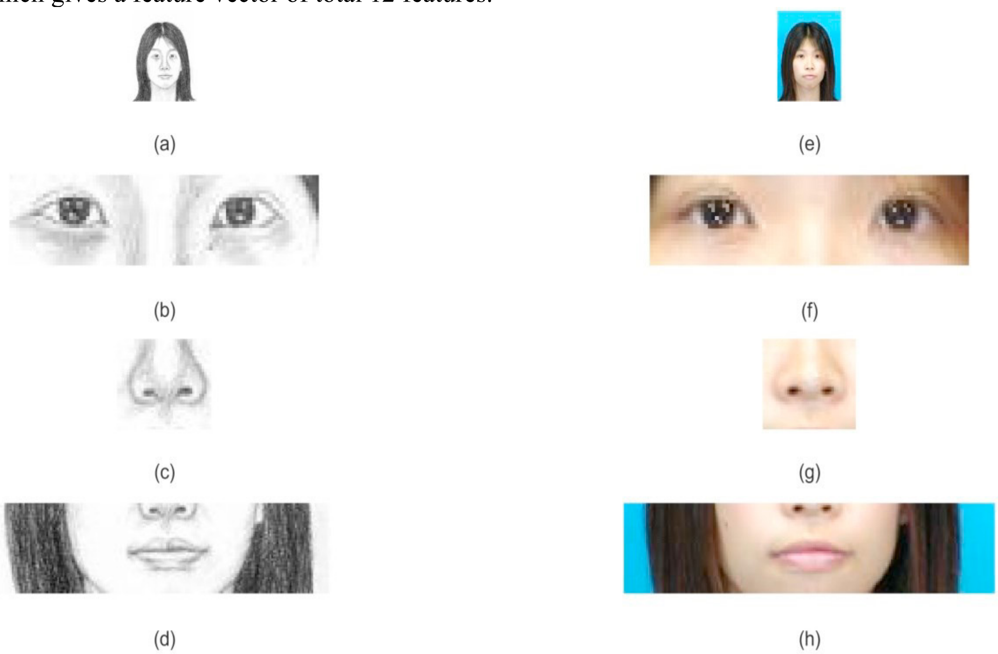

Fig. 4. (a), (b), (c) and (d) showing original sketch, eyes, nose and mouth extracted respectively and (e), (f), (g) and (h) showing original photo, eyes, nose and mouth extracted respectively.

Both HOG features and GLCM features have been concatenated to form a single feature vector of length 25932 (i.e. 25920 HOG plus 12 GLCM features) for each image.

3.4. Sketch to face matching

Sketches have been matched with photos using above feature vectors and it has been found that accuracy varies with the block size used in HOG which has been shown in Table 2. Also, computed accuracy has been compared with some state-of-art methods and it has been found that 181 out of 188 images are correctly matched and 7 are wrongly matched.

Table 2. Classification Accuracy of proposed method for various block size used in HOG feature calculation

Block size in HOG feature calculation	Number of features	Classification Accuracy (%)
32×32	1,080	82.98
16×16	5,544	94.7
8×8	25,920	96.3

Table 3. Comparison of classification accuracy of proposed method with other methods

Methods	Classification Accuracy (%)
Proposed method	**96.3**
PCA/LDA+ Geometric information (wavelet method) [11]	94
Canonical correlation analysis (CCA) [18]	94.6
Partial least squares (PLS) [19]	93.6

Table 3 shows the comparison of the proposed approach with other state-of-art approaches in terms of classification accuracy. It can be seen from the comparison that the proposed approach gives better results than the existing methods, even it is only rank-1 classification, while some of the existing approaches used rank-5 and rank-10 classification approach as well.

4. Conclusion and Future scope

A novel method for the matching of the sketches with photos has been introduced. In most of the previous work on this field, only one feature descriptor has been used for matching sketches and photos. But, to overcome their performances, hybrid of two different types of feature descriptors i.e. HOG descriptors and GLCM descriptors have been utilized here. The proposed method is giving better results compared with other state-of-art methods. In future, it can be extended by using a good feature reduction method to further reduce time consumption and computational cost. Work to improve the accuracy can also be done. This method can also be implemented on forensic sketches if its database is available online.

References

1. Uhl RG, da Vitoria Lobo N, Kwon YH (1994) Recognizing a facial image from a police sketch. In: Proceedings of 1994 IEEE Workshop on Applications of Computer Vision. pp 129–137
2. Wang X, Tang X (2009) Face photo-sketch synthesis and recognition. IEEE Trans Pattern Anal Mach Intell 31:1955–1967
3. Han H, Klare BF, Bonnen K, Jain AK (2013) Matching composite sketches to face photos: A component-based approach. IEEE Trans Inf Forensics Secur 8:191–204
4. Klum S, Han H, Jain AK, Klare B (2013) Sketch based face recognition: Forensic vs. composite sketches. In: 2013 international conference on biometrics (ICB). pp 1–8
5. Liu J, Bae S, Park H, et al (2015) Face photo-sketch recognition based on joint dictionary learning. In: 2015 14th IAPR International Conference on Machine Vision Applications (MVA). pp 77–80
6. Roy H, Bhattacharjee D (2016) Face sketch-photo matching using the local gradient fuzzy pattern. IEEE Intell Syst 31:30–39
7. Wang N, Gao X, Sun L, Li J (2018) Anchored neighborhood index for face sketch synthesis. IEEE Trans Circuits Syst Video Technol 28:2154–2163
8. Galea C, Farrugia RA (2017) Forensic face photo-sketch recognition using a deep learning-based architecture. IEEE Signal Process Lett 24:1586–1590
9. Setumin S, Suandi SA (2018) Difference of Gaussian oriented gradient histogram for face sketch to photo matching. IEEE Access 6:39344–39352
10. Chugh T, Singh M, Nagpal S, et al (2017) Transfer learning based evolutionary algorithm for composite face sketch recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp 117–125
11. Patil S, Shubhangi DC (2017) Forensic sketch based face recognition using geometrical face model. In: 2017 2nd International

Conference for Convergence in Technology (I2CT). pp 450–456
12. Li X, Yang X, Su H, et al (2018) Recognizing Facial Sketches by Generating Photorealistic Faces Guided by Descriptive Attributes. IEEE Access 6:77568–77580
13. Rajput GG, Geeta B, others (2018) Face Photo Recognition from Sketch Images Using HOG descriptors. In: 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT). pp 555–558
14. Zhang M, Wang R, Gao X, et al (2019) Dual-transfer face sketch--photo synthesis. IEEE Trans Image Process 28:642–657
15. Shu C, Ding X, Fang C (2011) Histogram of the oriented gradient for face recognition. Tsinghua Sci Technol 16:216–224
16. Shen X, Shi Z, Chen H (2016) Splicing image forgery detection using textural features based on the grey level co-occurrence matrices. IET Image Process 11:44–53
17. Danielsson P-E (1980) Euclidean distance mapping. Comput Graph image Process 14:227–248
18. Hotelling H (1992) Relations between two sets of variates. In: Breakthroughs in statistics. Springer, pp 162–190
19. Sharma A, Jacobs DW (2011) Bypassing synthesis: PLS for face recognition with pose, low-resolution and sketch. In: CVPR 2011. pp 593–600