Transfer Learning based Evolutionary Algorithm for Composite Face Sketch Recognition

Tarang Chugh[1], Maneet Singh[2], Shruti Nagpal[2], Richa Singh[2] and Mayank Vatsa[2]
Michigan State University, USA[1], IIIT Delhi, India[2]
chughtar@msu.edu, {maneets, shrutin, rsingh, mayank}iiitd.ac.in

Abstract

Matching facial sketches to digital face images has widespread application in law enforcement scenarios. Recent advancements in technology have led to the availability of sketch generation tools, minimizing the requirement of a sketch artist. While these sketches have helped in manual authentication, matching composite sketches with digital mugshot photos automatically show high modality gap. This research aims to address the task of matching a composite face sketch image to digital images by proposing a transfer learning based evolutionary algorithm. A new feature descriptor, Histogram of Image Moments, has also been presented for encoding features across modalities. Moreover, IIITD Composite Face Sketch Database of 150 subjects is presented to fill the gap due to limited availability of databases in this problem domain. Experimental evaluation and analysis on the proposed dataset show the effectiveness of the transfer learning approach for performing cross-modality recognition.

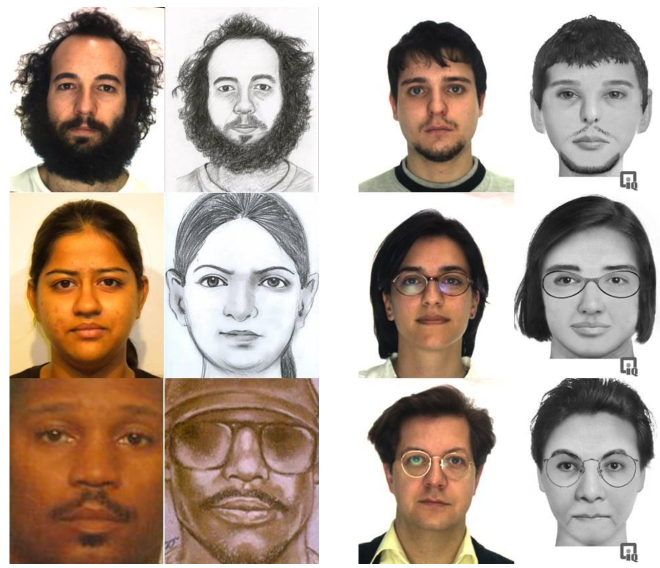


Figure 1. Sample sketch-digital image pairs. The first two columns contain digital images and their corresponding hand-drawn sketch images, while the next two columns refer to digital image and composite sketch pairs.

1. Introduction

A recent report on *Crime and Operation of Criminal Justice Systems* by United Nations Office on Drugs and Crime (UNODC) shows that there is an increase in criminal activities such as assault, robbery, kidnapping, and sexual violence [3]. To assist the law enforcement agencies, prolific research has been undertaken for improving the speed and accuracy of various processes, such as biometrics based matching for searching the list of suspects. For example, next generation latent fingerprint matching and face recognition against large databases is regularly being utilized in the criminal investigation. One of the applications where recent advancements in technology have been helping the agencies is creating sketches and matching them against a gallery of digital face (mugshot) photos.

In traditional sketch to photo matching, sketch artists play an important role in creating hand-drawn sketches based on the description provided by an eyewitness. However, lack of large number of skilled (forensics) artists and the disparity between hand-drawn forensic sketches and original faces are some of the major challenges associated with this problem. The disparity between the sketch and digital image depends on two factors: (i) facial description provided by an eye-witness and (ii) the skill-set of a sketch artist. Zhang et al. [33] have demonstrated that there can be a difference as high as 31% in face identification accuracy across different sketch artists. This difference can be attributed to the fact that hand-drawn sketches represent facial information as a combination of soft and prominent edges, whereas digital mugshot images are rich in texture and edge information [5]. Due to this heterogenous differences, it is an arduous task to match hand drawn sketches with digital images.

Table 1. Literature review of digital face to sketch recognition.

Authors	Year	Approach
Hand-Drawn Sketches (Discriminative Approaches)		
Uhl and Lobo [28]	1996	Photometric standardization of sketches and use of Eigen Analysis for matching
Zhang et al. [33]	2010	Humans vs. Automated System based on Principal Component Analysis
Klare et al. [14]	2011	SIFT and MLBP based local feature discriminant analysis (LFDA)
Bhatt et al. [5]	2012	Genetic optimization based Multiscale Circular Weber Local Descriptor (MCWLD)
Hand-Drawn Sketches (Generative Approaches)		
Tang et al. [26]	2002	Converting digital images into sketches using Eigen Analysis
Liu et al. [16]	2005	Separate texture and shape information and use of Bayesian classifier for matching
Gao et al. [9]	2008	Synthesizing pseudo-sketches using Hidden Markov Models
Wang et al. [30]	2009	Generating sketches from digital images using Markov Random Fields
Ouyang et al. [21]	2016	Model to reverse the forgetting process of the eye-witness
Composite Sketches		
Han et al. [12]	2013	Fusion of multi-scale local binary pattern encoding of key facial regions
Chugh et al. [7]	2013	Fusion of HOG and Image Moments
Mittal et al. [19]	2013	Daisy descriptor and Gentle boost classifier
Mittal et al. [17]	2014	Local multi-resolution self similarity descriptor based bag of words
Klum et al. [15]	2014	Holistic and Component based approach for scalable operational system
Mittal et al. [20]	2015	Transfer learning approach using deep network
Mittal et al. [18]	2017	Visual saliency, combination of texture features, and attributes
Proposed	2017	HOG + HIM + Evolutionary Algorithm + Transfer Learning

In the last few years, researchers have developed automated tools to create *composite sketches* based on the description of eyewitness. These sketch generation tools, e,g. FACES [2], are fast and the law enforcement agencies now often utilize them for creating the composite sketches. In contrast to hand-drawn sketches, composite sketches enable a quick feedback from the eye-witness during the process of sketch generation. Figure 1 presents sample digital face images and their corresponding hand-drawn images, along with sample digital face images and their corresponding composite sketches. It can be observed that composite sketches appear like a *synthetic* image because unlike hand-drawn sketches, they have limited fine (soft) edges and texture information. Therefore, it is our assertion that algorithms developed for hand-drawn sketches may not be directly applicable but these sketches can help in building automated algorithm for matching composite sketches with mugshot photos.

Table 1 summarizes the literature of facial sketch recognition, both with hand-drawn and composite sketches. It can be observed that while researchers have focused on addressing the problem of hand-drawn or composite face sketch recognition independently, no work aims to utilize the availability of hand-drawn sketches for developing composite sketch to photo recognition algorithms. This can serve as an important stepping stone for composite sketch recognition because most of the legacy sketches are in hand-drawn format, and having an efficient mechanism to utilize them for training can be very useful.

In this research, we address the problem of composite sketch to photo recognition by proposing a transfer learning based approach for incorporating knowledge learned from hand-drawn sketch recognition. Specifically, this research presents three fold contributions: (a) a feature descriptor termed as Histogram of Image Moments (HIM) is presented, (b) a transfer learning approach for evolutionary algorithm is proposed to match composite sketches and digital face images, and (c) IIITD Composite Face Sketch Database of 150 individuals is presented.

2. Proposed Algorithm

The proposed algorithm is developed using Histogram of Oriented Gradient (HOG) [8] feature descriptor and the proposed Histogram of Image Moments (HIM) descriptor. We utilize a transfer learning approach in a genetic learning framework for matching composite sketch to digital photos. Transfer learning focuses on utilizing the knowledge learned by solving a problem in one domain, known as the source domain, into another different but related domain, known as the target domain [22]. In the proposed approach, transfer learning helps in utilizing different kinds of sketches available in literature to learn better matcher that improves the recognition performance. Figure 2 summarizes the steps involved in the proposed algorithm. The individual components of the proposed algorithm are explained in the following subsections.

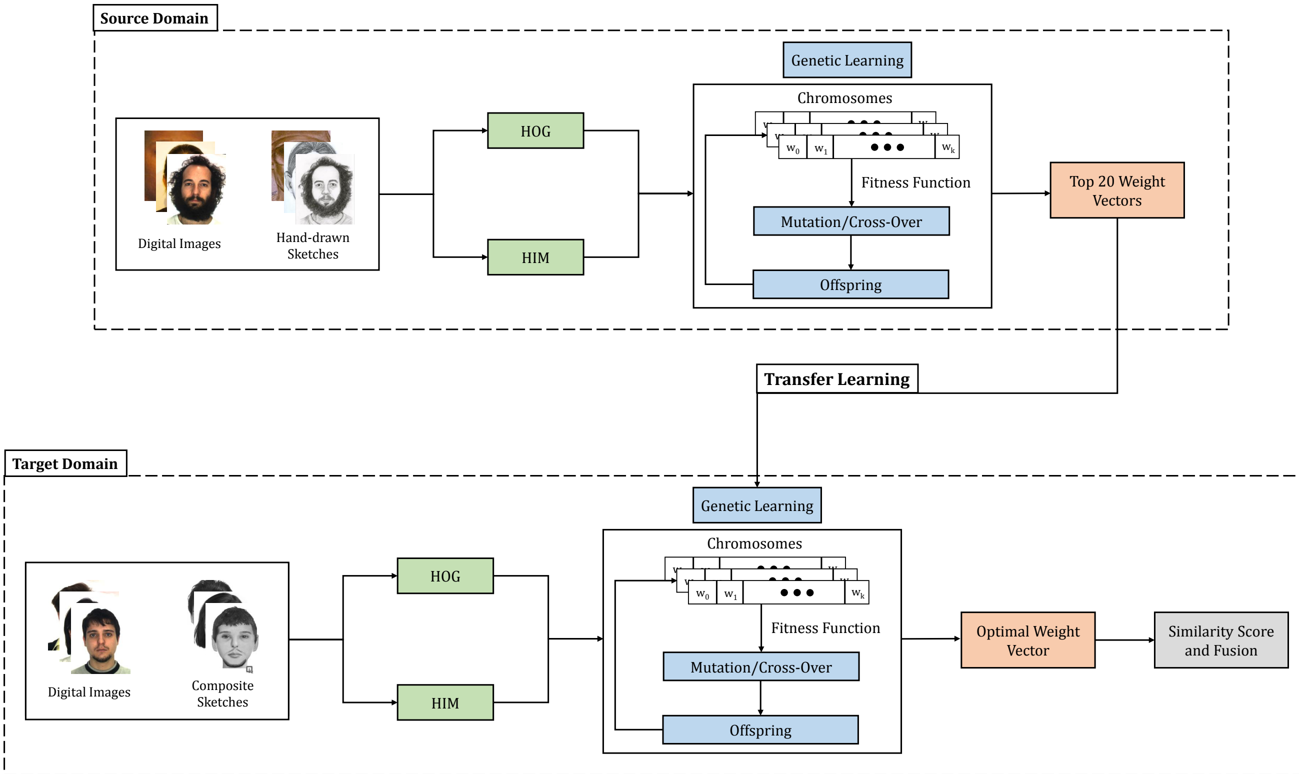


Figure 2. Block diagram illustrating the proposed algorithm for performing composite sketch to digital photo matching.

2.1. Pre-Processing

Digital photos and sketches may be misaligned; therefore, geometric normalization is performed to register the images. The eye-coordinates are detected using the OpenCV's boosted cascade of Haar-like features [29], and variations in rotation are normalized with respect to the horizontal axis using the eye-coordinates and inter-eye distance is fixed to 100 pixels. Finally, all images are resized to 192×224 pixels.

2.2. Feature Extraction

In this research (handcrafted) texture features have been used for encoding the image representations. Two kinds of feature extractors, the proposed Histogram of Image Moments (HIM) and Histogram of Oriented Gradient (HOG) [8], have been used for the same.

Histogram of Image Moments: Hu [13] proposed a set of regular image moments which have shown invariance to changes in rotation, translation, and scale. Image moments provide details such as the orientation information of an image, average localized weighted pixel intensity, and the centroid. In case of face images pertaining to different modalities, exact shape of facial structures may change with input modalities [24]; however, the variation in orientation of facial features is relatively less or varies consistently [6]. Such features can therefore be used to encode discriminative information in local regions across different modalities. Inspired from these findings, this research utilizes and builds upon image moments for performing face sketch recognition.

The image moments are based on the normalized central moment n_{pq}, which is calculated as follows:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}}; \gamma = \left(\frac{(p+q)}{2}\right) + 1 \tag{1}$$

$$\mu_{pq} = \sum_{x}\sum_{y}(x-\bar{x})^p(y-\bar{y})^q I(x,y) \tag{2}$$

$$\bar{x} = \frac{m_{10}}{m_{00}}, \qquad \bar{y} = \frac{m_{01}}{m_{00}} \tag{3}$$

where, μ_{pq} is the central moment. Central image moments correspond to the regular image moments with the origin as the centroid of the image $(\bar{x}, \bar{y})$. This is performed to eliminate the translation effect in images, since the central moments encode information related to the structure and spread of the image around the image centroid, as opposed to the actual origin. Out of the seven image moments proposed by Hu [13], concatenation of the first, third and fourth image moments $[M_1, M_3, M_4]$ is empirically shown to encode the features consistent across the two modalities for local patches in face images. The first, third, and fourth image moments are calculated as follows:

$$M_1 = \eta_{20} + \eta_{02} \tag{4}$$

$$M_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \tag{5}$$

$$M_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \tag{6}$$

In case of a gray-scale image I with dimensions $M \times N$, the regular moments of order $(p+q)$ are calculated as shown below:

$$m_{pq} = \sum_{x=1}^{M} \sum_{y=1}^{N} x^p y^q I(x, y) \tag{7}$$

where $I(x, y)$ correspond to the pixel intensity value at the image coordinates (x, y) and p, q ranges from $0, 1, 2, ...n$.

To compute a HIM feature image signature, each image is tessellated into 4×4 non-overlapping patches. The first, third and fourth image moment invariants are calculated for each of the 16 local patches using Eqs. 1-7. Concatenation of the three image moments for each of the 16 local patches constitutes the image signature of length 48.

Histogram of Oriented Gradients (HOG): Histogram of Oriented Gradients (HOG) was proposed by Dalal and Triggs [8] for performing the task of pedestrian detection. The feature descriptor makes use of the image gradients, in order to encode the intensity variations in local regions. This results in a feature descriptor of length 81, capable of encoding the object's shape and direction of pixel flow.

Feature Matching: Both the above features are histogram based approaches, therefore χ^2 distance metric can be used to calculate the distance between two given feature vectors. χ^2 distance between histogram vectors X and Y is calculated using the following expression:

$$\chi^2(X, Y) = \frac{1}{2} \sum_k \frac{(X_k - Y_k)^2}{X_k + Y_k}$$

where, k represents the local facial region in the form of a histogram bin.

2.3. Fusion via Evolutionary Algorithm

The tessellation and feature extraction process yields 16 patches with two feature extractors, thus providing 32 feature sets. Many studies including Bhatt et al. [4], Shan et al. [25], and Zhan et al. [32] have shown that all facial regions are not equally discriminative and contribute disparately in the task of face recognition. Since the feature sets of Histograms of Image Moments and Histogram of Oriented Gradients represent the local regions in a face, the histogram bins representing a specific local region should be assigned weights relative to their significance. Since the amount of inconsistency and error in individual facial regions vary, having uniform weight for all the regions and both the feature vectors is not suitable. However, the feature vectors of HIM and HOG, of length 48 and 81 respectively, have to be combined using an optimized weight vector for best performance. Considering the above mentioned factors, it is important to utilize an intelligent and adaptive weighting scheme for feature fusion.

2.3.1 Evolutionary Algorithm

The steps involved in the evolutionary approach of weight optimization are as follows:

Encoding: In evolutionary approaches, a single weight vector which needs to be optimized is known as a chromosome. Length of each chromosome is equal to the size of the feature vector, which in this case is equal to 129 (48+81). The weights in a chromosome are real valued numbers which are normalized to keep the sum of these weights equal to one.

Initialization: The pool size of the chromosomes is fixed to 100 units. A particular set of chromosomes is called a generation. The first generation of chromosomes is randomly initialized, and then used in subsequent iterations. In order to remove any initialization based bias, the complete process of weight optimization is repeated five times and the average accuracies are reported.

Fitness Function: Each chromosome is a possible solution of weight vectors, and only those chromosomes survive to the next generation which are better than the others. In the proposed algorithm, the measure of rank-10 identification accuracy is used as the fitness function to evaluate the chromosomes. Initially, 20 best performing chromosomes are retained and the rest are discarded. In the following iterations, the number of chromosomes retained depends on the change in the best identification accuracy obtained in the new generation as compared to the previous generation. Therefore, an improvement in identification accuracy results into more number of survivor chromosomes and no change in accuracy results in decrease of survivors. The percent change in the accuracy determines the actual number of change in survivors. The maximum limit of survivors is 30 whereas minimum limit is of 10 chromosomes. A weighted χ^2 distance measure is used for calculating the distance between a given pair of feature vectors. The formula for weighted χ^2 distance measure can be expressed as:

$$\chi^2(X, Y, W) = \frac{1}{2} \sum_k w_k \frac{(X_k - Y_k)^2}{X_k + Y_k}$$

where, X and Y are the two feature vector histograms which need to be matched, k represents the k^{th} bin of the histogram and w_k corresponds to the weight for the k^{th} histogram bin. The weight vector W is learnt by the evolutionary algorithm.

Crossover and Mutation: The knowledge of already found best performing survivor chromosomes, which now act as parents, is exploited by performing crossover and mutation operations. A set of uniform crossover operations are applied on the survivor chromosomes in pairs to generate a new set of 25 chromosomes. The top rank chromosomes participate in more cross-overs than low rank chromosomes. Also, a set of mutation operations are applied on the survivors to generate 25 new chromosomes. The mutation operations include swapping of two randomly selected weights and reversing the order of weights between two randomly selected positions.

Random Offspring Generation: In each iteration, a set of 20-40 randomly generated chromosomes are injected in the new generation to maintain the chromosome pool size as 100. This is performed in order to explore the solution space and avoid local optima.

2.4. Transfer Learning

Transfer Learning can be categorized into different settings depending on the similarity between source and target domains and tasks, type of available data (labeled or unlabeled) and the kind of modifications required to transfer the knowledge. Broadly, transfer learning can be categorized in three sub-settings, namely Inductive, Transductive, and Unsupervised Transfer Learning [22]. Our experiment setting is analogous to that of the inductive transfer learning setting where the source domain D_S and target domain D_T, of face recognition, are similar. However, the source task T_S of matching hand-drawn sketches/digital images with digital images and target task T_T, of matching composite sketches with digital images, are different.

"In Inductive Transfer Learning, the source domain D_S and source task T_S, the target domain D_T and target task T_T, we aim to improve the target predictive function $f_T(.)$ in D_T using the knowledge in D_S and T_S where $T_S \neq T_T$" [22]. In addition to the knowledge from source data, a small labeled dataset in target domain is also required to induce the predictive function which is used in target task. This dataset is referred to as the validation set.

There are different approaches in each setting to transfer the knowledge. In Inductive Transfer Learning setting, knowledge from source domain can be transferred in four possible approaches, namely instance transfer, feature-representation transfer, parameter transfer and relational knowledge transfer. In our application, as shown in Figure 2, the source and target task data use exactly the same set of features but the distributions of the data are different. Therefore, in the first stage (Figure 2), evolutionary learning is applied on the source domain (consisting of pairs of digital and hand drawn sketch images). The top performing chromosomes are then utilized to initialize the pool of chromosomes in the target domain (composite sketch and digital image pairs). Due to the difference in distributions and image representation between source and target tasks, there would be some of the source data which would help in learning the task and some data may negatively affect the learning for the target task. Using the iterative approach in evolutionary algorithm, we try to re-weight the source data to minimize the effects of "bad" source data and emphasize on "good" source data to contribute more for the target task. This inductive transfer learning approach helps in learning the weight parameters with limited training samples in the composite sketch database.

3. Databases and Experimental Protocols

In this research, we propose a new database, termed as the IIITD Composite Face Sketch Database containing multiple digital images for a single sketch image of a particular subject. The characteristics of this database are described in Section 3.1. Along with this, the algorithm also utilizes different kinds of sketch images from publicly available databases for training and transfer learning. Description of these existing databases are presented in Section 3.2.

3.1. IIITD Composite Face Sketch Database

In order to facilitate further research in face sketch recognition, the authors present IIITD Composite Face Sketch Database. The database contains face images pertaining to 150 subjects, such that each subject has a composite face sketch and two digital face images. Since composite sketches are software-generated sketches, in this research, FACES software [2] has been used to prepare the sketch images. FACES is a professional tool containing large number of facial features (more than $4,400$) and markings such as piercings, scars, moles, and tattoos. Depending upon the description of the eye-witness, these components can be used with a drag-and-drop feature to generate different face images. The proposed IIITD Composite Face Sketch Database contains 52 subjects from the FG-Net aging database [23], 82 subjects from the IIIT-D Aging Database [31], and the remaining 16 subjects are obtained from the Internet [1]. In the proposed database, for each subject, an eyewitness observes the first digital image for a few minutes and then creates the composite sketch image using the FACES tool. The second digital image is used during matching (first image is only used in sketch generation) as a gallery image. This is performed to imitate the real world scenario of variations observed in the suspect's face - at the time of crime, a witness sees an impression of the suspect which is different when compared with digital mugshot image. Over-

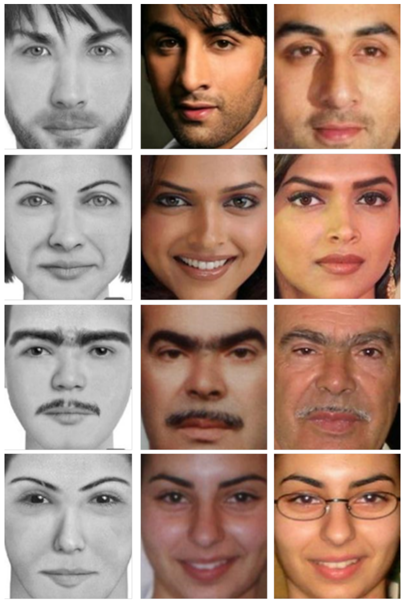


Figure 3. Sample images from the proposed IIITD Composite Face Sketch Database.

all, the proposed IIITD Composite Face Sketch Database contains 300 digital images and 150 composite sketch images. Figure 3 presents some sample images of the proposed database.

3.2. Existing Databases

Extensive research has been performed in the domain of matching hand-drawn sketches and digital face images. The hand drawn sketch databases have been classified into the following three categories: forensic, semi-forensic, and viewed [5]. In this research, we use all three kinds of hand-drawn databases along with the digital images for training.

Digital Image Database: In order to learn features on digital images, 250 digital-digital image pairs from the CMU Multi-PIE Dataset [11] have been used in the experiments. This subset consists of image pairs incorporating illumination and expression variations.

Forensic Sketch Database: It comprises of 190 forensic sketch-digital image pairs from the IIIT-Delhi Sketch Database [5] which were obtained from works by Lois Gibson [10], Karen Taylor [27] and few sources from Internet. Forensic sketches are drawn by a professional sketch artist based on the facial description provided by an eye-witness. The memory of eye-witness and ability to convey that information to the sketch artists as well as the artists' experience and interpretation of the description, holds important role in the quality of sketches.

Semi-Forensic Database: It consists of 106 semi-forensic sketch-digital image pairs from the IIIT-Delhi Sketch Database [5]. These sketches are drawn by a professional sketch artist. Unlike the forensic sketches where the description provided by the eye-witness holds utmost importance, semi-forensic sketches are drawn based on the recollection of face by sketch artist. The artist is shown a digital image and has to draw the sketch after a defined time interval.

Viewed Sketch Database: It comprises 482 sketch-digital image pairs from the IIIT-Delhi Viewed Sketch Database [5]. This database consists of sketches drawn by a professional sketch artist for digital images collected from different sources.

3.3. Experimental Protocol

The effectiveness of the proposed algorithm is demonstrated using five different experiments. The protocol of each experiment is illustrated in Table 2. The protocols are divided based on the images in the source domain; however, in each of the experiments, the target task is to match composite sketches with digital images. These experiments aim to identify which type of sketches provide the best knowledge that can be transferred to the target domain task of matching composite sketches and digital images.

For each experiment, the testing database of the target domain comprises of 100 image pairs from the proposed dataset. A set of 25 composite-digital image pairs are selected as the validation set. Further, for experiment 5 (discussed below), 25 pairs are selected as source domain. It is ensured that for each fold of the experiment, the 25 pairs of source domain (experiment 5), 25 pairs of validation set, and 100 testing pairs of target domain have no overlap among them. The results are reported with 5-fold cross validation. In order to select the best performing chromosome for target domain, a set of 20 top performing weight vectors obtained from the genetic algorithm are validated using the validation set. The details of each experiment are provided below.

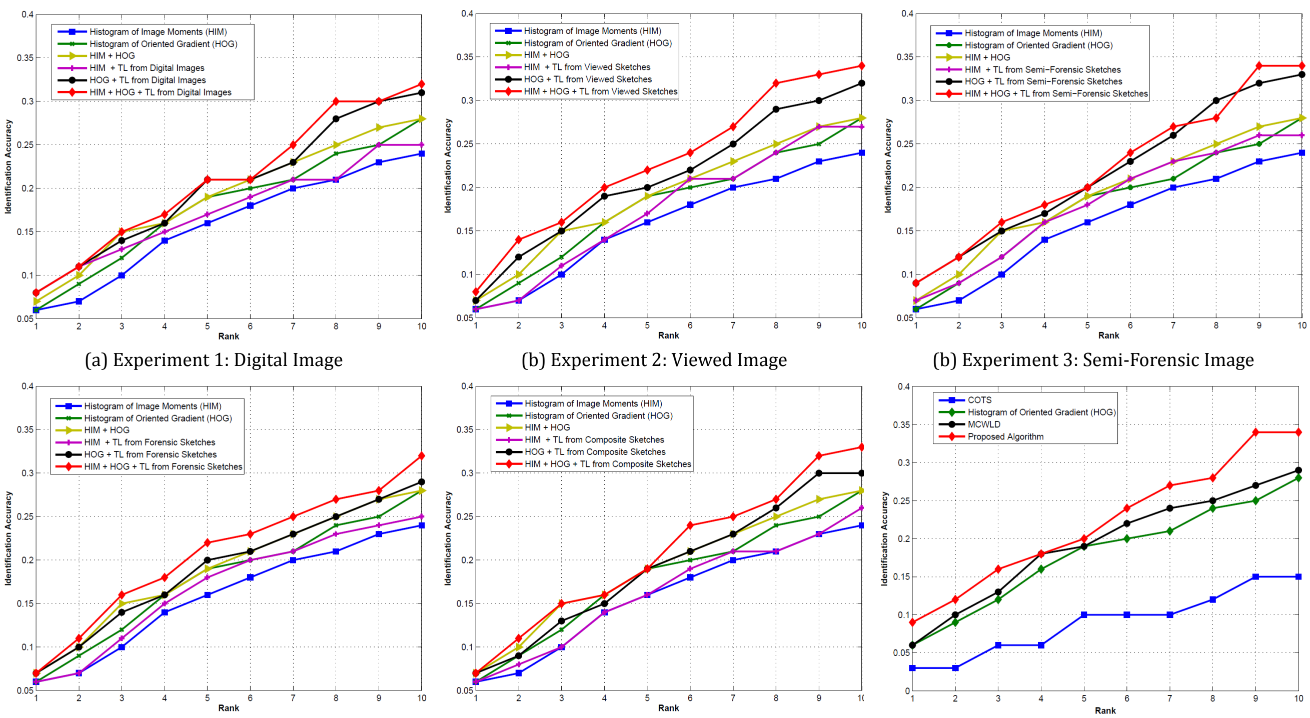
Experiment 1 - Digital Images: Experiment 1 makes use of the knowledge learned from the source domain of digital face images which are considered to be rich in information content than any type of sketches. The training database comprises of 250 digital image pairs (frontal images) from CMU-PIE dataset [11].

Experiment 2 - Viewed Hand-Drawn Sketches: It uses the knowledge learned in source domain of matching viewed hand-drawn sketches and digital face image pairs. The source domain comprises of 482 pairs of viewed hand-drawn sketches and digital face images.

Experiment 3 - Semi-Forensic Hand-Drawn Sketches:

Table 2. Details of the experimental protocols.

Experiment	Source Domain / Training Data	Training Set Size	Target Domain / Testing Data
Experiment - 1	Digital Image	250	Composite Sketch
Experiment - 2	Viewed Hand Drawn Sketch	482	Composite Sketch
Experiment - 3	Semi-Forensic Hand Drawn Sketch	106	Composite Sketch
Experiment - 4	Forensic Hand Drawn Sketch	190	Composite Sketch
Experiment - 5	Composite Sketch	25	Composite Sketch



(a) Experiment 1: Digital Image (b) Experiment 2: Viewed Image (b) Experiment 3: Semi-Forensic Image

(d) Experiment 4: Forensic Image (e) Experiment 5: Composite Image (f) Comparison with Existing Algorithms

Figure 4. CMC curves for all five experimental protocols, and comparison with existing algorithms.

The task in source domain is to match semi-forensic hand-drawn sketches and digital face image pairs. The source domain consists of 106 pairs of semi-forensic and digital face images from the IIIT-Delhi Sketch Database [5].

Experiment 4 - Forensic Hand-Drawn Sketches: The source dataset consists of 190 pairs of forensic sketches and digital face images from the IIIT-Delhi Sketch Database [5].

Experiment 5 - Composite Sketches: In this experiment, both source and target domain correspond to matching composite sketches and digital face image pairs. The database used in this experiment for source task is 25 pairs of composite sketch and digital face images taken from the proposed IIITD Composite Face Sketch Database.

4. Results and Discussion

Performance of the proposed algorithm is compared with its constituent elements and existing algorithms. Figure 4 presents the Cumulative Match Characteristic (CMC) curves for all the experiments. It can be observed that for all five experiments, the proposed algorithm yields improved accuracies than both the feature vectors individually. In all cases, rank-10 identification accuracy of both HIM and HOG features improve by 5 - 10%, after applying transfer learning. This clearly demonstrates the importance of transfer learning when data characteristics of source and target domains are not exactly same. With viewed sketch images as the source domain, the proposed algorithm yields a rank-10 accuracy of 34%. This is similar to the results obtained with semi-forensic images. It is interesting to note that training with forensic sketches yield lower accuracy than semi-forensic sketches, even when the number of forensic hand-drawn training images is higher than semi-forensic images. It is our assertion that the difference in accuracy is because there is a large difference between forensic sketch and corresponding digital images, whereas, the difference between composite test and digital image is not that large. Similarly, lower accuracy with digital images as the training set can be attributed to the fact that source domain does not contain inter-modality differences but the target domain has significant variations. Moreover, target domain training with a smaller number of images might not be sufficient to learn the modality variations. We also observed that training with composite sketches in the source domain yields around 5% rank-1 accuracy. This is less than the results obtained with semi-forensic image and it is our assertion that the lower accuracy may be due to smaller number of training images in the source domain.

Comparison has also been performed with the state-of-the-art algorithm MCWLD [5] on hand-drawn sketches and a commercial system, FaceVACS[1]. For this experiment, we selected the semi-forensic database as source domain, as it yields the best results among the five protocols. The results show that compared to state-of-the-art algorithm and COTS, the proposed algorithm yields significantly better results. At rank-10, accuracy of the proposed algorithm is almost 20% higher than the commercial system and 7% higher than MCWLD.

5. Conclusion

This research aims to address the problem of composite face sketch recognition by proposing a novel transfer learning based evolutionary algorithm. Owing to the limited availability of composite sketch datasets, the proposed approach utilizes *similar* data (hand-drawn and digital image pairs) in order to learn an initial model in the source domain. The learned parameters of the initial model are then *transferred* to the target domain for matching composite face sketches with digital images. Therefore, this research presents multiple contributions: (i) a transfer learning based evolutionary algorithm is presented for face sketch recognition, (ii) a new feature descriptor, termed as Histogram of Image Moments is proposed for encoding facial information, and (iii) IIITD Composite Face Sketch Database is proposed which consists of composite face sketch and digital images for a single subject. The proposed algorithm showcases improved identification performance on the IIITD Composite Face Sketch Databases when compared with existing state-of-the-art algorithm and commercial system.

[1] http://www.cognitec.com/

References

[1] Face Sketch Software : Famous Faces. https://www.facebook.com/FACESketchSoftware. 5

[2] FACES: The ultimate composite software. http://www.facesid.com. 2, 5

[3] United Nations Office on Drugs and Crime, CTS-2013 crime statistics report. www.unodc.org/unodc/en/data-and-analysis/statistics/crime.html. 1

[4] H. S. Bhatt, S. Bharadwaj, R. Singh, and M. Vatsa. On matching sketches with digital face images. In *IEEE International Conference on Biometrics: Theory, Applications and Systems*, pages 1–7, 2010. 4

[5] H. S. Bhatt, S. Bharadwaj, R. Singh, and M. Vatsa. Memetically optimized MCWLD for matching sketches with digital face images. *IEEE Transactions on Information Forensics and Security*, 7(5):1522–1535, 2012. 1, 2, 6, 7, 8

[6] S. Biswas, G. Aggarwal, N. Ramanathan, and R. Chellappa. A non-generative approach for face recognition across aging. In *IEEE Second International Conference on Biometrics: Theory, Applications and Systems*, pages 1–6, 2008. 3

[7] T. Chugh, H. S. Bhatt, R. Singh, and M. Vatsa. Matching age separated composite sketches and digital face images. In *IEEE International Conference on Biometrics: Theory, Applications and Systems*, pages 1–6, 2013. 2

[8] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 886–893, 2005. 2, 3, 4

[9] X. Gao, J. Zhong, J. Li, and C. Tian. Face sketch synthesis algorithm based on E-HMM and selective ensemble. *IEEE Transactions on Circuits and Systems for Video Technology*, 18(4):487–496, 2008. 2

[10] L. Gibson. *Forensic Art Essentials*. Academic Press-Elsevier, 2008. 6

[11] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker. Multi-PIE. *Image and Vision Computing*, 28(5):807–813, 2010. 6

[12] H. Han, B. F. Klare, K. Bonnen, and A. K. Jain. Matching composite sketches to face photos: A component-based approach. *IEEE Transactions on Information Forensics and Security*, 8(1):191–204, 2013. 2

[13] M.-K. Hu. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, 8(2):179–187, 1962. 3, 4

[14] B. Klare, Z. Li, and A. K. Jain. Matching forensic sketches to mug shot photos. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(3):639–646, 2011. 2

[15] S. J. Klum, H. Han, B. F. Klare, and A. K. Jain. The FaceSketchID system: Matching facial composites to mugshots. *IEEE Transactions on Information Forensics and Security*, 9(12):2248–2263, 2014. 2

[16] H. Lu, S. Ma, X. Tang, Q. Liu, and H. Jin. A nonlinear approach for face sketch synthesis and recognition. *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 1:1005–1010, 2005. 2

[17] P. Mittal, A. Jain, G. Goswami, R. Singh, and M. Vatsa. Recognizing composite sketches with digital face images via SSD dictionary. In *IEEE International Joint Conference on Biometrics*, pages 1–6, 2014. 2

[18] P. Mittal, A. Jain, G. Goswami, M. Vatsa, and R. Singh. Composite sketch recognition using saliency and attribute feedback. *Information Fusion*, 33:86 – 99, 2017. 2

[19] P. Mittal, A. Jain, R. Singh, and M. Vatsa. Boosting local descriptors for matching composite and digital face images. In *IEEE International Conference on Image Processing*, pages 2797–2801, 2013. 2

[20] P. Mittal, M. Vatsa, and R. Singh. Composite sketch recognition via deep network - a transfer learning approach. In *International Conference on Biometrics*, pages 251–256, 2015. 2

[21] S. Ouyang, T. M. Hospedales, Y. Z. Song, and X. Li. Forgetmenot: Memory-aware forensic facial sketch matching. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 5571–5579, 2016. 2

[22] S. J. Pan and Q. Yang. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345–1359, 2010. 2, 5

[23] G. Panis, A. Lanitis, N. Tsapatsoulis, and T. F. Cootes. Overview of research on facial ageing using the FG-NET ageing database. *IET Biometrics*, 5(2):37–46, 2016. 5

[24] U. Park, Y. Tong, and A. K. Jain. Age-invariant face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(5):947–954, 2010. 3

[25] C. Shan and T. Gritti. Learning discriminative LBP-histogram bins for facial expression recognition. In *British Machine Vision Conference*, 2008. 4

[26] X. Tang and X. Wang. Face photo recognition using sketch. In *International Conference on Image Processing*, volume 1, pages I–257–I–260, 2002. 2

[27] K. Taylor. *Forensic Art Essentials*. CRC Press, 2000. 6

[28] R. G. Uhl and N. da Vitoria Lobo. A framework for recognizing a facial image from a police sketch. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 586–593, 1996. 2

[29] P. Viola and M. J. Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154, 2004. 3

[30] Y. Wang, L. Zhang, Z. Liu, G. Hua, Z. Wen, Z. Zhang, and D. Samaras. Face relighting from a single image under arbitrary unknown lighting conditions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(11):1968–1984, 2009. 2

[31] D. Yadav, R. Singh, M. Vatsa, and A. Noore. Recognizing age-separated face images: Humans and machines. *PLoS ONE*, 9, 2014. 5

[32] C. Zhan, W. Li, and P. Ogunbona. Finding distinctive facial areas for face recognition. In *International Conference on Control Automation Robotics Vision*, pages 1848–1853, 2010. 4

[33] Y. Zhang, C. McCullough, J. R. Sullins, and C. R. Ross. Hand-drawn face sketch recognition by humans and a PCA-based algorithm for forensic applications. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 40(3):475–485, 2010. 1, 2