

Matching Forensic Composite Sketches with Digital Face Photos: A Bidirectional Local Binary Pattern-Based Approach



H. T. Chethana and Trisiladevi C. Nagavi

Abstract Facial sketches are extensively used by investigators in order to identify the suspects involved in criminal activities. The manual method of identifying suspects is slow and complex. To make the process automated, proposed method attempts to map the computer created composite sketches to face photos automatically. This research work focuses on searching for missing and wanted persons who are involved in criminal activities that in turn assist investigative agencies in locating suspects in a timely manner. Proposed method attempts to address the challenge of mapping composite sketch to facial photos using bidirectional local binary pattern (BLBP). In the proposed method, Viola–Jones algorithm is used to detect composite sketch; feature extraction is done using BLBP; classification and recognition are done using two-dimensional convolution neural networks (2D-CNNs). The experimental results show that BLBP and 2D-CNN combined approach achieves recognition accuracy of 90% in comparison with other existing methods (Han et al. in *IEEE Trans. Inf. Forensics Secur.* 8, 191–204, 2013; Hochreiter et al. in *Neural Comput.* 9, 1735–1780, 1997; Paritosh et al.: in *International Conference on Biometrics*, Phuket, Thailand, pp. 251–256, 2015; Roy, H., Bhattacharjee, D.: *Adv. Intell. Syst. Comput.* 883, 2019).

Keywords Two-dimensional convolution neural networks (2D-CNNs) · Bidirectional local binary pattern (BLBP) · Viola–Jones algorithm · Composite sketch · Digital face photos · Composite sketch with age variations dataset · Feature extraction · AdaBoost algorithm

H. T. Chethana (✉) · T. C. Nagavi

Department of Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysore, Karnataka, India
e-mail: chethanaht@vce.ac.in

T. C. Nagavi

e-mail: trisiladevi@sjce.ac.in

H. T. Chethana

Department of Computer Science and Engineering, S. J. College of Engineering, JSS Science and Technology University, Mysore, Karnataka, India

1 Introduction

Facial composite sketches are extensively employed to assist identification of criminals when facial image of the suspect is not available at the time of the crime. This system is significant because in most cases the facial photograph of a suspect is not available [4, 20, 32]. Sketches are of two kinds: composite and hand drawn. Composite sketches are generated by taking the description from eyewitness with the help of software tools such as FACES and EvoFIT [18]. Alternatively, hand drawn sketches are generated by professional face drawing artists based on eye witness description. Preparation of composite sketches takes less time, effort, and cost when compared to hand drawn sketches [7].

The real-world examples of composite sketches with digital face photos are shown in Fig. 1. Figure 1a and c shows composite sketches of sample face 1 and 2, whereas Fig. 1b and d demonstrate the corresponding digital face photos. Composite sketches do not contain miniature feature details and lacks in texture information; as a result, these sketches look artificial. But, in digital images, representation of information is very rich. Due to this diversity, it is a highly challenging task to match composite sketch with corresponding digital face photos [5, 6].

Once a sketch of suspect's face is created, authorities will think that using these sketches someone will recognize the individual and provide relevant information which helps the investigating agencies to find the suspects quickly. The eyewitness or victim's description becomes the only form of evidence available [18]. But, this process is inefficient and does not leverage. In particular, law enforcement agencies maintain extensive mug shot databases.

Fig. 1 Sample pair of composite sketch with digital faces photos [8, 18]



Sample 1: a) Composite sketch b) Digital face photo



Sample 2: c) Composite sketch d) Digital face photo

In accordance to that proposed system focuses on automatic system of mapping composite sketch with digital face photo using BLBP with 2D-CNN. The body of the paper is structured as follows. The exhaustive survey of literature available on the proposed method is discussed in Sect. 2. Further, proposed methodology is illustrated in Sect. 3. Experimental results and analysis is presented in Sect. 4. Finally, the paper is concluded with concluding remarks and future scope of the work.

2 Literature Survey

In this section, the exhaustive literature survey on various methods implemented for matching composite sketch to digital face photos is discussed. Plenty of research works are observed in the field of facial forensic image mapping.

An approach based on fine-tuning dual streams deep network (FTSDSN) with multi-scale pyramid decision (MsPD) was proposed by authors [14] for solving heterogeneous face recognition. The approach which combines CNN and FTSDSN removes nonlinear information and retains discriminative data by Rayleigh quotient. The MsPD adaptively regulates sub-structure weight and obtains strong classification performance. Experimental analysis is performed on CUHK FERET and CASIA NIR-VIS 2.0 databases.

In another work, authors [9, 31] have proposed heterogeneous face recognition (HFR) to recognize the person from facial visible and near-infrared images. The task is challenging as it contains very few training samples. Mutual component convolutional neural network (MC-CNN) is used to tackle these two issues. The MC-CNN incorporates mutual component analysis (MCA) and deep CNN having special fully connected (FC) layer [34]. This FC layer extracts modal-independent hidden factors and updates according to maximum likelihood rather than back propagation which reduces overfitting from limited data [35].

An automatic recognition of photo face from a sketch image was proposed by authors [27] for criminal investigation. Using nearest neighbor algorithm, the feature vectors of both composite sketch and photo were compared. The 'n' most similar photos were retrieved and matched using the L1-distance measure. The experiment was carried out on three datasets, namely CUHK, CUFS, and FERET. Proposed method provides promising results compared with other existing methods.

Similarly, a binary method for matching images between sketch and photos of heterogeneous faces was proposed by the authors [24]. The robust binary pattern of local quotient (RBPLQ) is used to extract illumination and noise invariant features. Local quotient (LQ) extracts illumination invariant information. Robust local binary pattern (RLBP) captures LQ variations. The experiments are carried out with NIR-VIS benchmark database, and recognition accuracy of 60.72% is achieved.

An approach based on counter propagation network (CPN) using biogeography particle swarm optimization (BPSO) was proposed by the authors [1] for face recognition based on sketches. Proposed method was used to calculate the mean square error (MSE) between the feature vectors of facial sketch and photo. The BPSO-CPN

method is demonstrated on CUHK and IIITD dataset with 1000 sketches and photos. The experimental results are promising and of high precision nature compared to other existing methods.

A face sketch synthesis method was proposed by the authors [30] to perform the representation of nonlinear mapping between face photos and sketches. The sketches synthesized from exemplar methods are blurred and affected with block artifacts. To improve synthesis performance, joint training scheme is proposed by considering sketches. A joint training is performed with a photo and sketch. The sample is constructed first by combining photo and its sketch using high-pass filtered (HPF) image. A random sampling approach is adopted for each test photo to choose the joint training photo and sketch in the neighboring region. The results of the experiments which are performed on public datasets reveal the superiority of the joint training model.

In another work, authors [22] have proposed the concept of multi-scale Markov random fields RF model and facial landmarks (MRF-FLs) method. The combined use of MRF with FL helps in reducing the distortions at the lower part of face contour. Experimental results are evaluated on CUHK and AR face sketch databases [28].

Most of the works [1, 9, 24] have been carried out on heterogeneous face recognition. As per the observation, few works are noted in the area of matching composite sketch with digital face photos on IIITD dataset. Even though these research works [22, 30] are based on matching composite sketch with digital face photos, they are addressing traditional features like illumination and pose variation.

The works proposed in the literature review have focused on matching composite sketch with digital face photos by considering noisy and synthetic sketches. It is observed that none of the works [24, 33] have employed BLBP approach with 2D-CNN. In the current work, an attempt is made to improve the face photo recognition results with composite sketches by employing BLBP approach with 2D-CNN.

3 Proposed Methodology

The proposed work aims to address the mapping of composite sketch with digital face photos using supervised learning approach. It uses Viola–Jones algorithm for detecting faces from composite sketches, BLBP for salient feature extraction, classification, and recognition using 2D-CNN. The reasons for adopting these methods are stated below.

Viola–Jones facilitates the detection of faces from the given image. It has high face detection rate, robust and can be used in real-time applications. It is one of the good and popular face detection algorithms. Hence, it is employed in the proposed work for detecting the face of a given composite sketch.

Local binary pattern (LBP) is an effective texture descriptor and helps to capture prominent features using spatial information. Further, the process of feature extraction is made robust by using bidirectional LBP coding [2].

CNNs are type of MLPs which are fully connected networks. The fully connected nature of these networks makes them susceptible to overfitting with the data. Typically, overfitting is avoided by adding some form of magnitude measurement of weights to the loss function. But, CNNs follow a distinctive means to avoid overfitting. They exploit the hierarchical patterns in the input data and congregate multifaceted patterns through simpler and smaller patterns.

Also, the connectivity mode between neurons has resemblance to connectivity between cortical neurons. Here, neurons respond to stimuli only through restricted region of the visual field known as the receptive field. The receptive fields of every neuron participating in CNN partially overlap to cover the whole visual field.

This nature of CNNs makes them simple and efficient to adopt. In the proposed system, two-dimensional convolutional kernels are applied to leverage the spatial and temporal features of the slice so that the classification accuracy improves.

The proposed architecture has training and testing phase. During training phase, preprocessed facial composite sketch is given as input for extracting facial information. Viola–Jones algorithm is one of the most suitable algorithms for facial region of interest extraction. Algorithm detects the face in the frame sequence from preprocessed facial composite sketch. In order to extract the features from the detected composite sketch, BLBP-based approach is used. These extracted features are fed as input to 2D-CNN. Later on, the generated features from the network are stored in database. Testing phase follows all the steps of the training phase on test image sequence. It makes use of the generated features generated during training for classifying the test sample. The architecture of proposed system is shown in Fig. 2.

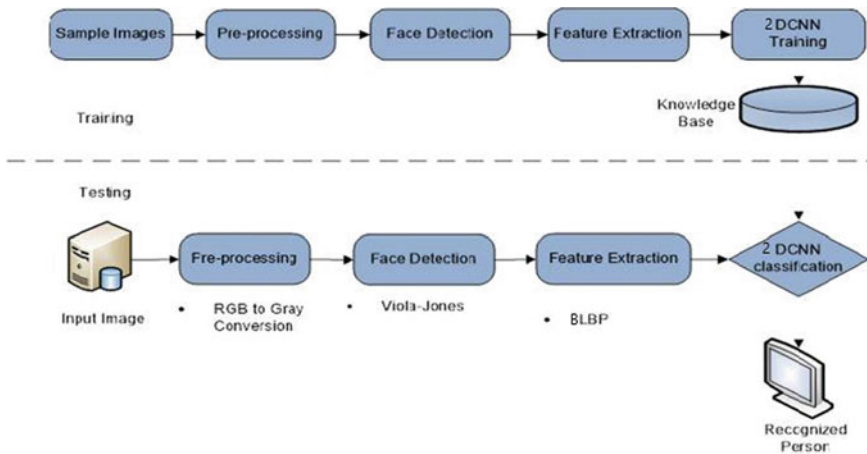


Fig. 2 Architecture of the proposed system

Algorithm to Match Composite Sketch with Digital Face Photos

Step 1: Preprocessing technique is applied to composite sketch (input image) and corresponding digital face photo. In preprocessing stage, threshold is set by varying the illumination.

Step 2: Digital face photos of corresponding sketches are trained by using a combination of BLBP—a feature extraction technique and 2D-CNN.

Step 3: Composite sketch is fed as input to the proposed system in the testing phase.

Step 4: For every composite sketch, corresponding match is computed and digital face photo bearing highest accuracy is displayed.

3.1 Preprocessing

Preprocessing is the series of operations performed on image data which reduces noise, removes unwanted distortions, and enhances some image features which improves the quality of the image. This step is very important for further processing, and it is needed on color, gray level, or binary images. In this stage, each facial photo and composite sketch are resized and transformed in to two dimensional grayscale images. In the proposed methodology, simple preprocessing is used because CNN learns through filters. This is the salient feature of the CNN, and it is a major advantage.

3.2 Face Detection

Viola–Jones algorithm [29] is used to detect face from preprocessed digital face image and composite sketch. This algorithm exhibits high detection accuracy. Even in composite sketch images, face is identified irrespective of the texture, color, motion, position, and size. The main goal of this approach is to simply differentiate faces from non-faces. It consists of four stages: integral image conversion, Haar feature selection, AdaBoost training, and cascading classifiers.

3.2.1 Integral Image Conversion

It is the first step in Viola–Jones algorithm which converts the original image in to an integral image. Integral image facilitates to calculate the Haar features quickly. Figure 3a, b, and c depicts input image, integral image, and area of rectangle, respectively. An integral image shown in Fig. 3b is obtained by converting all pixel values of input image shown in Fig. 3a equal to sum of the pixels left and above of the associated pixel values. It is depicted in Fig. 3b, and it helps to evaluate the rectangle

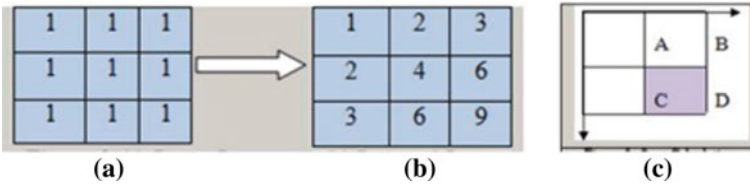


Fig. 3 **a** Input image, **b** integral image representation, **c** area of the rectangle

features. Figure 3c denotes the overlapping of rectangle corners of the input image with the integral image pixel values.

Each rectangle area is computed by using the integral image. Figure 3c indicates the purple rectangular area = $D + A - B - C$, where rectangle B, C includes rectangle A ; hence, A is added in the calculation.

3.2.2 Haar Feature Selection

Haar features are evaluated in two ways (i) by computing each rectangle area, multiplying each rectangle area with their individual weights, then results are combined (ii) by taking the black and white rectangles difference. The obtained Haar feature values [17] are given as input to the cascaded classifier. Haar filters are applied onto one special area, and Haar feature values are calculated by taking the difference of pixels from black and white area as shown in Eq. (1).

$$\text{Haar Feature value} = \sum (\text{black pixels area}) - \sum (\text{white pixels area}) \quad (1)$$

3.2.3 AdaBoost Algorithm

In this step, weighted weak classifiers are combined to construct AdaBoost. A weak classifier is mathematically represented in Eq. (2).

$$h(s, f, p, \theta) = \begin{cases} 1 & \text{if } pf(s) > p\theta \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where s denotes 24×24 sub-window, f indicates the Haar feature, p denotes the polarity, and θ indicates the threshold to classify whether s is face or non-face. From the obtained Haar feature values, best features are selected by modifying AdaBoost algorithm.

3.2.4 Cascading Classifier

Here, the obtained Haar feature values are given as input to the cascaded classifier. The cascading classifier is constructed with many stages, and each stage has a strong classifier whose work is to compute and classify whether the given input sub-window is positive (face) or negative (non-face). If the input sub-window for a given stage is classified as non-face, then immediately, the given sub-window is discarded. Conversely, if input sub-window is categorized as a face, then this face sub-window is passed to a subsequent stage of cascaded classifier.

Initially, frame sequences are taken as input than a 24×24 sub-window which slides over the entire frame to calculate the Haar feature values. Using cascading classifier, each Haar feature is classified and used to detect whether it is composite face or not. If it is a composite face, then the sub-window is passed to the next location, otherwise sub-window is discarded. The flow diagram of face detection is shown in Fig. 4.

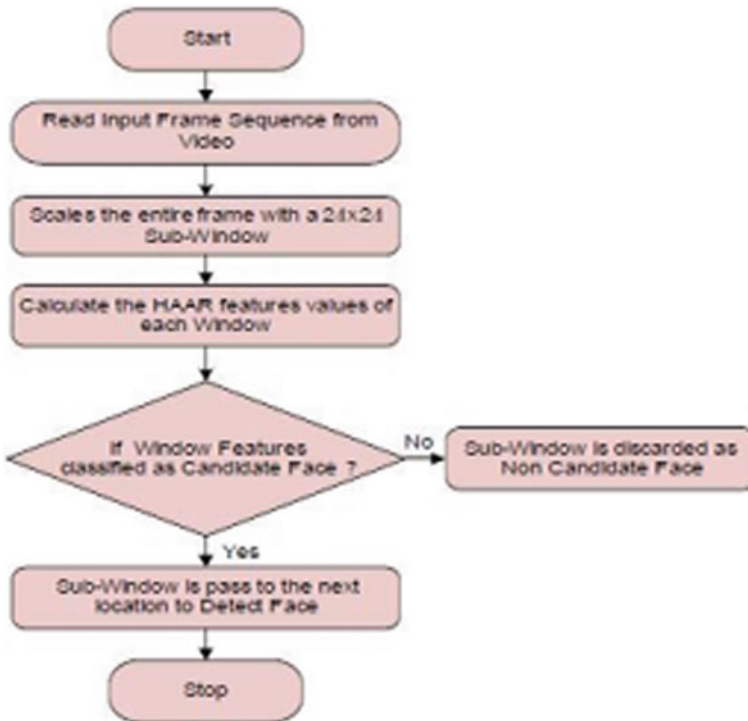


Fig. 4 Flow diagram of face detection

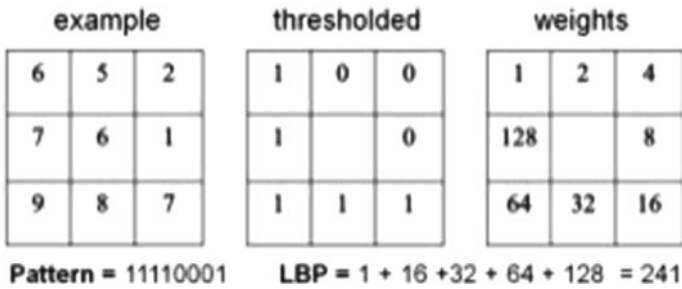


Fig. 5 Local binary pattern

3.3 Feature Extraction

The proposed approach uses BLBP to extract the features. Unidirectional LBP remembers only the past information from the input but to predict both past and future information a bidirectional LBP is implemented [11, 15, 16].

Spatial structure features of the local image such as texture are extracted by using LBP. For a center pixel value t_c , the neighboring pixels t_i are equally spaced. LBP works by thresholding the neighbors t_i with the center pixel value t_c to generate a bit binary pattern. The output of LBP for t_c is expressed in decimal form. Output binary pattern is obtained as $d_i = t_i - t_c$ the difference between neighboring and center pixel, $d_i = 1$ if $d_i \geq 0$ and $d_i = 0$ if $d_i < 0$. The LBP uses only the magnitude information while ignoring sign information [10]. The pictorial representation of LBP is given in Fig. 5.

3.4 Two-Dimensional Convolutional Neural Networks

The basic CNN is used to collect only spatial information but ignores the temporal information. To overcome this problem, the 2D-CNN approach uses a 2D object as input which collects both spatial information and temporal relations [25, 26].

The 2D-CNN filter/kernel slides on the temporal and spatial dimension of input which enables the neural network to collect more feature vectors. The extracted features map contains the temporal information essential for labeling the frame sequence. It contains 2D feature values for input data. The convolution and pooling layers of 2D-CNN work in cubic style [19]. Figure 6 depicts the 2D convolution of multiple frames. Height (H) represents the height of each frame; length (L) represents the depth of stack of frames, and W represents width of each frame.

The input layers of 2D-CNN accept the image as input to represent it in the matrix form. The array of the pixel represented in matrix form is given as input to the convolution layer. This layer uses a convolutional function defined as $y = \text{conv}(x, w)$ where x denotes input data, w denotes the convolution filter, and y represents the

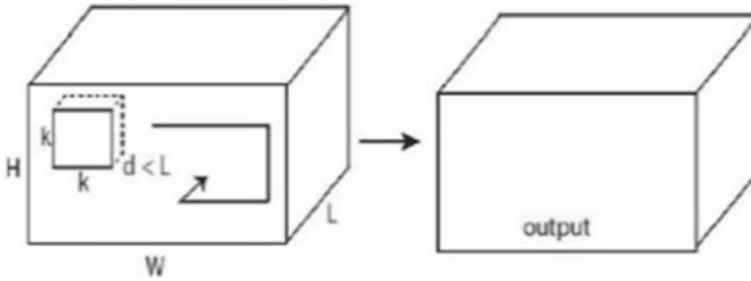


Fig. 6 2D convolution on multiple frames

convolution filter output. Convolution filters are also referred to as kernel or neuron. For 2D input data ‘ x ’, the dimensions are $Z \times M \times N \times K \times S$, where Z defines the feature length values, M depicts height, N depicts width of the feature maps, K represents the image channel number, and S defines the batch size. These dimensions are same for convolution filter with height, width, and length.

The feature maps are extracted by applying these convolution filters. Rectified linear unit (ReLU) is an activation function in neural network which will not trigger all the neurons at the same time [3, 23]. It works six times faster when generally compared with other activation functions such as sigmoid and tangent hyperbolic (tanh). The rectified output features are given as input to pooling layer to decrease the amount of parameters, spatial size, and computation of the network [36].

4 Experimental Results and Analysis

In this section, a discussion is made on the results of the experiments conducted to study the performance of the proposed method based on composite sketch with age variations (CSA) dataset [8, 18]. The dataset is collected from IIIT Delhi which consists of three age variations: same, old, and young. Fifty pairs of composite sketch with same age variations and its corresponding digital face photos are considered for experimentation. The training and testing set consists of digital face photos and composite sketches, respectively. Few sample composite sketches are shown in Fig. 7.

During testing phase, composite sketches are fed as input to the proposed system. It is compared with all the digital face photos stored in the training set. Proposed system extracts the matching digital face photo having highest percentage of matching. Sample composite sketch, its matching digital photo, and corresponding recognition accuracy are shown in Figs. 8 and 9. The plot of recognition accuracy for matching the given input composite sketch with all the digital face photos is shown in Fig. 10. In this plot, values along X-axis denote various digital face photos, and Y-axis values denote the matching accuracy in percentage.

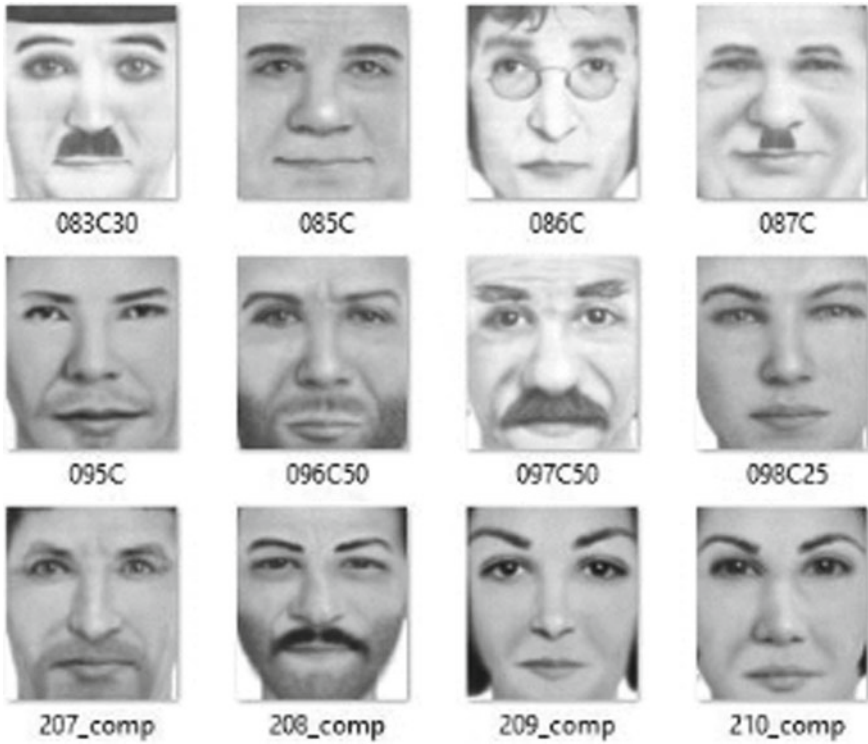


Fig. 7 Sample composite sketches used for testing



Fig. 8 Sample input and its corresponding output image

```
Select Query
DCNN Classification
Percentage of Matching:
77.3639
```

Fig. 9 Sample percentage of matching

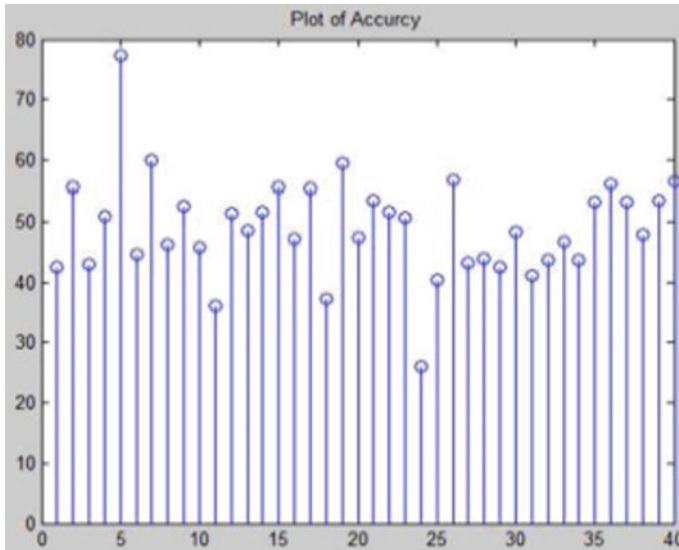


Fig. 10 Sample plot of accuracy in percentage

The performance analysis of the proposed method is carried out using the confusion matrix. Values are shown in Table 1. The standard yardsticks used for analyzing the results are true positive (TP) and false positive (FP) which are shown in Table 1.

Fifty pairs of composite sketches with digital face photos were considered for experimentation. Out of 50 pairs, 45 pairs were correctly matched and predicted as positive class (TP). Remaining five pairs were incorrectly matched and predicted as negative class (FP). Results of the sample cases are portrayed in Fig. 11 for correct and incorrect matches.

The accuracy of our proposed system is 90%, and it is calculated using Eq. (3).

Table 1 Confusion matrix

	True positive	False positive	Accuracy
Results	45	05	90%



Fig. 11 Sample results using the proposed method **a** correct match **b** incorrect match

$$\text{Accuracy} = \text{TP}/(\text{TP} + \text{FP}) \tag{3}$$

Details of recognition accuracy for fifty pairs of composite sketch and digital face photo are shown in Table 2. Same is pictorially represented in the graph presented in Fig. 12. Here, values along X-axis represent sample numbers, and values across Y-axis denote corresponding accuracy. In Fig. 10, the plot is drawn based on the values obtained for matching one composite sketch with all the digital face photos. Conversely, in Fig. 12, values in X-axis denote various composite sketches, and Y-axis represents recognition accuracy obtained for matching its corresponding digital face photos.

Comparative study is conducted on existing methods and proposed method. It is observed that proposed work provides a recognition accuracy of 90% compared with other existing methods is indicated in Table 3. Same is pictorially represented in Fig. 13.

The system performance is appraised through exhaustive testing and results analysis. The experimental results prove that BLBP and 2D-CNN combined approach achieves recognition accuracy of 90% in comparison with other existing methods as shown in Fig. 13. Here, x denotes names of various methods, and y -axis portrays recognition accuracy. The system can be made scalable for large database of composite sketch-based recognition.

5 Conclusion

This research paper aims to solve the problem faced by forensic facial experts in manual identification of suspects. To address this issue, a new approach for facial photo identification based on composite sketch is proposed which is helpful for forensic study. The approach uses BLBP and 2D-CNN for feature extraction and classification, respectively. The experimental results of the proposed approach envisage an appreciable recognition accuracy of 90% in comparison with other existing methods [12, 13, 21, 24]. The task of suspect’s face analysis based on the composite

Table 2 Details of recognition accuracy for fifty pairs of composite sketch and digital face photo

S. No.	Image No.	Accuracy	Output image	S. No.	Image No.	Accuracy	Output image
1	85	79.53	085_Same	26	211	82.5516	211_Same
2	86	81.84	086_Same	27	212	84.6122	212_Same
3	87	83.2013	087_Same	28	213	83.9789	213_Same
4	88	82.9092	088_Same	29	214	84.2316	214_Same
5	89	77.3639	089_Same	30	215	83.6245	215_Same
6	90	83.1291	090_Same	31	216	82.545	216_Same
7	91	81.0389	091_Same	32	217	84.6122	217_Same
8	92	78.2006	092_Same	33	218	82.2792	218_Same
9	93	82.9092	093_Same	34	219	84.0183	219_Same
10	94	84.2316	094_Same	35	220	83.093	220_Same
11	95	79.1883	095_Same	36	221	82.9256	221_Same
12	96	82.3481	096_Same	37	222	81.3014	222_Same
13	97	83.6245	097_Same	38	223	82.1709	223_Same
14	98	82.3481	098_Same	39	224	83.9986	224_Same
15	99	79.3064	099_Same	40	225	82.0233	225_Same
16	201	82.6664	201_Same	41	226	82.6205	226_Same
17	202	80.3827	202_Same	42	227	82.0987	227_Same
18	203	81.3539	203_Same	43	228	82.2891	214_Same
19	204	83.093	204_Same	44	229	83.1061	229_Same
20	205	82.5319	225_Same	45	230	82.6205	230_Same
21	206	82.1709	206_Same	46	231	82.8305	204_Same
22	207	81.4064	207_Same	47	232	83.2537	204_Same
23	208	81.5409	208_Same	48	233	81.8723	233_Same
24	209	76.3828	209_Same	49	234	82.0528	211_Same
25	210	83.9789	210_Same	50	235	81.8297	214_Same

sketch is addressed and explored for the purpose of criminal identification. The proposed BLBP and 2D-CNN combined algorithm significantly outperforms the state-of-the-art criminal face analysis techniques. The system can be made scalable for large database of composite sketch-based face recognition.

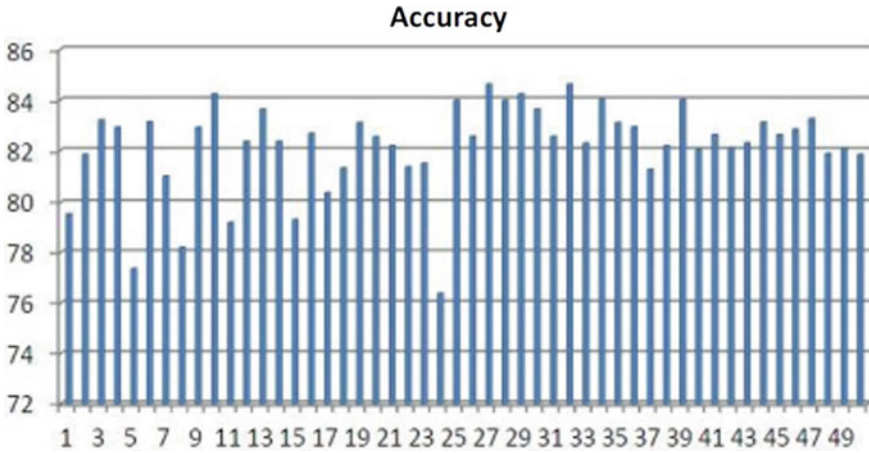


Fig. 12 Sample plot of accuracy for 50 samples pairs

Table 3 Comparison of proposed method with existing approaches

S. No.	Author	Feature extraction	Classifier	Accuracy (%)
1	Hochreiter et al. [11]	LBP	Haar feature	87
2	Paritosh et al. [19]	Single shot detector	Gentle boost KO	58.6
3	Roy et al. [21]	RBPLQ with LQ	Neural networks	60.72
4	Han et al. [10]	MLBP	Support vector machine	89
5	Proposed work	BLBP	2D-CNN	90

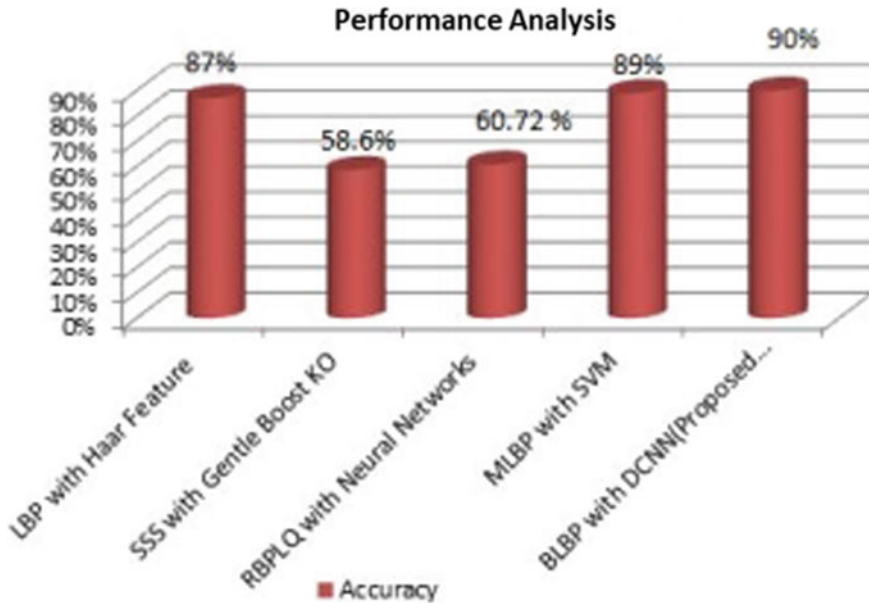


Fig. 13 Comparison of proposed method with other existing approaches

References

1. Agrawal, S., Singh, R.K., Singh, U.P., Jain, S.: Biogeography particle swarm optimization based counter propagation network for sketch based face recognition. *Multimedia Tools Appl.* **78**, 9801–9825 (2018)
2. Ahonen, T., Hadid, A., Pietikainen, M.: Face description with local binary patterns: application to face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**, 2037–2041 (2006)
3. Cambria, E., Hazarik, D., Poria, S., Hussain, A., Subramanyam, R.B.V.: *Benchmarking Multimodal Crime Analysis*, pp. 166–179. Springer Nature (2017)
4. Chethana, H.T., Trisiladevi, C.N.: Face recognition methods for facial image analysis in forensics. In: *Proceedings of 3rd International Conference on Electrical, Electronics, Communication, Computer Technologies & Optimization Techniques*, p. 56. Mysuru, India (2018)
5. Chethana H.T., Nagavi, T.C.: Face recognition for criminal analysis using Haar Classifier. *i-Manager's J. Comput. Sci.* **8**(1) (2020)
6. Chethana, H.T., Nagavi, T.C.: A new framework for matching forensic composite sketches with the digital images. *IJDCF. Special Issue Submission: Advanced Digital Forensic Techniques for Digital Traces*, vol. 13, Issue 5, Article 1 (2021)
7. Chugh, T., Bhatt, H.S., Singh, R., Vatsa, M.: Matching age separated composite sketches and digital face images. In: *Proceedings of 6th International Conference on Biometrics: Theory, Applications & Systems*. Arlington, VA, USA (2018)
8. Chugh, T., Singh, M., Nagpal, S., Vatsa, M.: Transfer learning based evolutionary algorithm for composite face sketch recognition. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI, USA (2017)
9. Deng, Z., Peng, X., Li, Y., Qiao, Y.: Mutual component convolutional neural networks for heterogeneous face recognition. *IEEE Trans. Image Process.* **28**, 3102–3114 (2019)

10. Frinken, V., Uchida, S.: Deep BLBP neural networks for unconstrained continuous handwritten text recognition. In: Proceedings of 13th International Conference on Document Analysis and Recognition (ICDAR), pp. 911–915. IEEE, NW Washington, DC, United States (2015)
11. Graves, A., Jaitly, N., Mohamed, A.R.: Hybrid speech recognition with deep bidirectional LSTM. In: Proceedings of IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU), pp. 273–278 (2013)
12. Han, H., Klare, B.F., Bonnen, K., Jain A.K.: Matching composite sketches to face photos: a component based approach. *IEEE Trans. Inf. Forensics Secu.* **8**, 191–204 (2013)
13. Hochreiter, S., Schmidhuber, J., Mehal, K.: Long short-term memory neural computation. *Neural Comput.* **9**, 1735–1780 (1997)
14. Hu, W., Hu, H.: Fine tuning dual streams deep network with multi-scale pyramid decision for heterogeneous face recognition. *Neural Process. Lett.* **50**, 1465–1483 (2019)
15. Karim, F., Majumdar, S., Darabi, H., Chen, S.: LBP fully convolutional networks for time series classification. *IEEE Access* **6**, 1662–1669 (2018)
16. KaaeSonderby, S., KaaeSonderby, C., Nielsen, H., Winther, O.: Convolutional LBP networks for subcellular localization of proteins. In: Proceedings of International Conference on Algorithms for Computational Biology, pp. 68–80. Springer, Cham (2015)
17. Ma, S., Bai, L.: A face detection algorithm based on Adaboost and new Haar-like feature. In: Proceedings of 7th IEEE International Conference on Software Engineering and Service Science (ICSESS), pp. 651–654. Beijing (2016)
18. Nagpal, S., Singh, M., Singh, R., Noore, A., Majumder: A face sketch matching via coupled deep transform learning. In: Proceedings of International Conference on Computer Vision (ICCV), pp. 5419–5428. Venice, Italy (2017)
19. Ogawa, A., Hori, T.: Error detection and accuracy estimation in automatic speech recognition using deep bidirectional recurrent neural networks. *Speech Commun.* **89**, 70–83 (2017)
20. Patil, S., Shibhangi, D.C.: Composite sketch based face recognition using ANN classification. *Int. J. Sci. Technol. Res.* **9**, 42–50 (2020)
21. Paritosh, M., Vatsa, M., Singh, R.: Composite sketch recognition via deep network—a transfer learning approach. In: International Conference on Biometrics, pp. 251–256. Phuket, Thailand (2015)
22. Radman, A., Suandi, S.A.: Markov random fields and facial landmarks for handling uncontrolled images of face sketch synthesis. *Pattern Anal. Appl.* **22**, 259–271 (2019)
23. Rosas, V.P., Mihalcea, R., Morency, L.P.: Multimodal crime analysis of Spanish online images. *IEEE Intell. Syst.* **28**, 38–45 (2013)
24. Roy, H., Bhattacharjee, D.: Heterogeneous face matching using robust binary pattern of local quotient: RBPLQ. *Adv. Intell. Syst. Comput.* 883 (2019)
25. Sainath, T.N., Vinyals, O., Senior, A., Sak, H.: Convolution long short-term memory, fully connected deep neural networks. In Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 4580–4584. United States (2015)
26. Salama, S.E., Shoman, M.E., WahbyShalaby, M.A.: EEG-based emotion recognition using 2D convolutional neural networks. *Int. J. Adv. Comput. Sci. Appl.* **9**, 329–337 (2018)
27. Setumin, S., Suandi, S.A.: Cascaded static and dynamic local feature extractions for face sketch to photo matching. *IEEE Access* **7**, 27135–27145 (2019)
28. Trisiladevi, C.N., Bhajantri, N.U.: Overview of automatic Indian music information recognition, classification and retrieval systems. In: Proceedings of International Conference on Recent Trends in Information Systems (ReTIS). Kolkata, India (2011)
29. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. I-I. Kauai, HI, USA (2011)
30. Wan, W., Lee, H.J.: A joint training model for face sketch synthesis. *Appl. Sci.* **9**, 1731 (2019)
31. Wang, J., Yang, Y., Mao, J., Haug, Z., Haug, C., Xu, W.: CNN-RNN a unified framework for multi-label image classification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2285–2294 (2016)

32. Xu, J., Xue, X., Wu, Y., Mao, X.: Matching a composite sketch to a photographed face using fused HOG and deep feature models. *The Visual Computer* (2020). <https://doi.org/10.1007/s00371-020-01976-5>
33. Xu, X., Li, Y., Jin, Y.: Hierarchical discriminant feature learning for cross-modal face recognition. *Multimedia Tools Appl.* (2019)
34. Zhao, F.P., Li, Q.N., Chen, W.K., Liu, Y.F.: An efficient sparse quadratic programming relaxation based algorithm for large-scale MIMO detection. arXiv e-prints, [arXiv:2006.12123](https://arxiv.org/abs/2006.12123) (2016)
35. Zhang, Y., Gao, S., Xia, J., Liu, Y.F.: Hematopoietic hierarchy: an updated roadmap. *Trends Cell Biol.* **28**, 976–986 (2018)
36. Zhang, M., Wang, N., Li, Y., Gao, X.: Neural probabilistic graphical model for face sketch synthesis. *IEEE Trans. Neural Netw. Learn. Syst.* **31**, 2623–2637 (2019)