Webology (ISSN: 1735-188X)
19, Number 3, 2022

Sketch-Based Face Recognition Using Deep Learning

Mubashir Hassan[1] , Muhammad Munawar Iqbal[2] , Huma Qayyum[3] , Marriam Nawaz[4] , Farooq Ali*[5]

[1,2,4,5]Department of Computer Science, University of Engineering and Technology-Taxila, 47050, Pakistan.
[3]Department of Software Engineering, University of Engineering and Technology-Taxila, 47050, Pakistan.

Received: November 03,2021; Accepted: March 28,2022

Abstract- Sketch-based face recognition is extremely useful for law enforcement and security services in the identification of suspects and the resolution of criminal investigations. Real-time face matching versus face sketches is a tough and essential issue in the sketch-based face recognition field because of the extensive modality gap between actual pictures and sketches created by artists or machines. In this study, we offered a deep learning-based solution to deal with the above-mentioned problems. More clearly, we have developed a new CNN framework and modified two existing deep learning frameworks namely the Extended AlexNet (E-AlexNet), and LetNet (E-LeNet) for face recognition using sketches. The performance of the introduced algorithms is evaluated by using two distinct types of sketches which include both hand-drawn sketches (viewed, forensic) and computer-generated (composite) sketches. The experimental findings suggest that the proposed technique namely the E-AlexNet model performs satisfactorily against both types of sketches and achieved a higher accuracy of 92.22%, and 95.18% on the CUHK and AR datasets respectively.

Keywords: CNN, Face Recognition, Sketches, LeNet, AlexNet.

1. Introduction

Identifying people from their face sketches is a difficult endeavor to accomplish. Photos or videos captured by security surveillance cameras in the real world are of poor quality and are unable to provide any useful information or clues to the authorities [1]. For this reason, investigators rely on forensic sketch artists [2] or sketching software [3] to sketch the faces of suspects in the course of their inquiry. Between an actual photograph and a sketch, however, there is a significant difference in modality. As a result of this void, sketch-based facial recognition was identified as a difficult topic. Even though face sketch recognition appears to be quite simple for a human, however, it is highly dependent on the availability of experts. Therefore, there is a requirement for an automated system capable of accurately and quickly matching the sketches against the suspect images.

Several problems may encounter during the matching process of a sketch with a photo. During the sketch generation, several necessary details can be omitted like the information about the skin tone, hair color, etc. [4] Another problem is the environmental condition gap that exists between the sketch and a photo, i.e. brightness conditions, body expressions, background information, light variations, etc [5]. Moreover, the availability of a single sketch of the suspect further complicates the face recognition process and can generate false results. Hence, such systems are required which can generate reliable results with the minimum amount of information. Forensic sketches are broadly divided into three categories namely the viewed, forensic and composite sketches respectively. In the first type namely the viewed sketches, the sketch of a target is produced by directly using his/her picture, while for the forensic sketches, the sketches are created based on the description provided by a witness [6, 7]. Whereas, the composite sketches employ computer-based systems to produce the sketches of targets by using the different regions of human faces. Globally, 80% of law-making organizations utilize software-based tools for sketch generation due to their high speed and efficiency [8].

In order to overcome the challenge of matching a face sketch picture to its real-world counterpart, algorithms that are focused on sketches have been utilized in face detection technology. The fundamental categories of recognition methods are generative and discriminative. Generative methods are the more common kind. In the generative method, an image of a sketch is first transformed into an image of a real face, and then this real face picture is compared to an actual image. To do the matching in a discriminatory manner, one can employ feature extraction from a sketch in conjunction with stored databases of actual face photos. In this particular work, we applied a discriminative method for the purpose of face recognition. There have been several new approaches to facial recognition based on sketches in the last few years. Two types of face sketch datasets have been used in the literature: viewed sketches [1, 9]and forensic sketches [10-12]. Sketches produced by artists based on mugshot images have been shown to have a high level of accuracy [13-18]. Although they are based on the memories of eyewitnesses and artists, forensic and software-generated sketches generally differ from genuine or mug shot images in style [19]. Real-world photos and computer-generated or forensic sketches can't always be perfectly matched. Therefore, there is a need to propose such a model which can work well for both forensic and composite sketches.

Several studies from history have attempted to recognize the faces from the sketches either hand or software generated. Sketch-to-face detection has seen far less research than facial recognition using a real image. To top it all off, almost all of the work was done on hand-drawn view sketches. As an example of what can be achieved with a modality, invariant descriptor-based method: [20] Using several descriptors, Peng et al. [1] developed benchmarks for the sketch to face identification. They've used a variety of fusion algorithms as a starting point, including pixel-level

fusion, feature-level fusion, and score-level fusion, among others, on datasets before applying various descriptors, including LBP, Fisherface, P-RS, and VGG Face. Fusion approaches, on the other hand, have their own set of problems. The enhanced Evolutionary-Based Face Recognition Model was proposed by Salem [21]. Images were matched using the HOG characteristics. The dataset in this study, on the other hand, is limited to hand-drawn sketches.
Bhatt et al. [22] proposed an automated technique that extracts features along with minute details from local facial areas of both sketch images and real face photographs. Structures in local regions can be identified using Weber's descriptors. Evolutionary metric optimization, on the other hand, was offered as a way to apply weights to each local facial region in order to improve recognition. An approach known as locality-constrained feature space learning (LCFSL) was proposed by Gao et.al. [23] for matching sketch face images against mugshot images of various resolutions. An emphasis on the latent feature space and dictionary learning are combined in this strategy. For cross-resolution heterogeneous image data, we can use the observed dictionary pair to learn a latent feature space during optimization. This strategy allows us to simply deal with the challenge of cross-resolution recognition. Recognition of sketch-photo faces with varied resolutions is also possible using the nearest neighbor classifier.

Face sketch identification was proposed by Wan et al., [24] using transfer learning on VGG-Face networks. To compute the reliable face features and reduce intra-class variability, the authors carefully employ triplet loss because of the shortage of training data. Hard triplet sample selection has been used by the authors to improve the training triplet images while also preventing delayed convergence. Facial images and sketches can be converted to 1024 sized embedding due to the fine-tuned deep network, and this can be utilized to directly quantify face similarity. Comparing the results of the experiments with other popular state-of-the-art methods, the deep modelextracted facial features had the highest recognition accuracy. For cross-modality face recognition, Peng et al. [7] presented DLFace, a deep local descriptor learning system. As a result of this discriminant, local data from suspected sample patches may readily be learned. For cross-modality face recognition, they first look at deep local descriptors. Local patch-level modality gaps are eliminated by using an enumeration loss function. EnumerateNet is a CNN that incorporates the proposed loss function and compactness requirement. Any typical face recognition system can benefit from the use of the deep local descriptor. In a series of extensive cross-modality facial recognition trials, DLFace outperforms current state-of-the-art approaches. Wan et al. [25] provide a GAN-based face sketch generation and detection approach via using a single process. The authors developed an efficient generator that exploits the benefits of U-Net structure together with residual blocks to produce the high-resolution sketches. The authors created a multi-task deep network that serves as both a generator training discriminator and a face sketch identification feature extractor. The performance of the work suggest that this method can obtain higher performance on both the synthesis and recognition of face sketches.

Researchers have been focusing on viewed sketches rather than other categories like forensic or computer-generated sketches as there are difficult to recognize. In this work, we have covered the limitations of existing work by proposing a solution to tackle both hand-drawn and softwaregenerated sketch modalities. We have employed three types of deep learning models namely CNN, E-AlextNet, and E-LeNet for recognizing the human face from sketches. Initially, we developed a new CNN network to match human faces with sketches. Then we modified two existing deep learning frameworks namely E-LeNet, and E-AlexNet. Following are the major contributions of our work:

- A new CNN model is introduced to recognize human faces with sketches.
- We have modified two models namely E-LeNet, and E-AlexNet to categorize if an input pair of sketch and a digital sample is from the same class or not.
- We have presented a lightweight solution to face recognition from sketches.
- An extensive evaluation is conducted by using both the hand-drawn and softwaregenerated sketches to exhibit the efficacy of the proposed solution for face recognition.

2. Proposed Method

The proposed method is shown in Fig.1. First of all the real images or sketches are resized into 126 x126 dimensions. Then the feature vector of both images is extracted. In the end, neural network models i.e. CNN, E-LeNet, and E-AlexNet are trained on the processed samples to recognize the human faces from sketches. The distinguishing part of all employed networks is the detection of the matched human face with the given sketches. So, for all three networks, we have trained both the facial images and sketches separately in the classification section. Then, the absolute difference among the input image and sketch is computed which is treated as the final output score as illustrated in Eq. 1.

$$\text{Absolute}_{\text{Diff}} = \text{abs}(I_s - I_k) \qquad (1)$$

Here, I_s is the input face image, and I_k is a sketch sample. Then the computed score is checked across a defined threshold to take the final decision whether a face sample or sketch is matched or not. We have set the value of threshold 0.5 for all three networks. The description of all deep learning methods is elaborated in the subsequent sections.

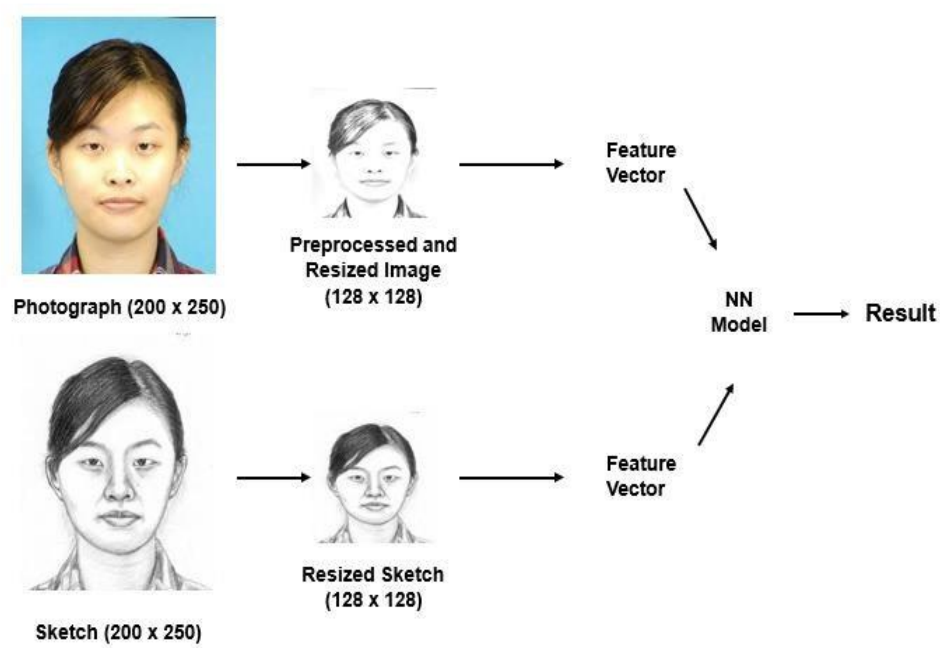

Figure 1 Proposed Method

2.1. CNN

CNNs are biologically inspired derivatives of MLP NN, extensively used for 2D pattern recognition research problems i.e. object detection and Optical Character Recognition (OCR), etc [26]. Unlike MLP which contains fully connected layers, CNN consists of convolution and subsampling layers. The basic CNN contains 3 components named the convolution, pooling, and output layers where the pooling layer is optional in some cases.

The traditional architecture of CNN with three convolutional layers is well suited to face recognition problems. It comprises the input layer, multiple hidden layers (repetitive convolution layers, normalization, pooling), and a fully connected output classification layer. The input layer receives and stores the input data. This layer defines the height, width, and channel information of the input sample. While the hidden layer is the fundamental part of the CNN framework which consists of convolution, pooling, and activation functions. At this stage, the keypoints for face recognition are computed by employing a combination of convolution and pooling layers together with activation functions to have a representative set of features for numeral classification. The convolutional layer is put over the input image to generate the feature maps. The computed feature maps consist of neurons that receive input from a local receptive field. Each feature map shares the neuron's weight which enables CNN to use repeating units with the same configuration. This distinguishing characteristic of CNN allows it to detect the features irrespective of their position in the visual field and increases the learning accuracy by minimizing the number of learned parameters. Furthermore, to reduce the computational complexity, pooling layers are added which perform a data reduction step. This data reduction step is performed to the convolution layer result by employing local averaging over a predefined window. It divides the input sample into a series of non-overlapping regions and then for each window computes the maximum value. The pooling operation is significant as it assists to remove non-maximal values and generates a translational invariant form of data. The output layers produce the classification result. The output layers are fully connected and have a distinctive set of weights that enable them to compute the complex keypoints and perform classification. The architecture of the introduced CNN network for human face recognition from the sketches is elaborated in Table 1.

Table 1: Architectural description of introduced CNN framework.

No.	Layer	Details
1.	Input	input samples with the resolution of 128 x128 x3
2.	Con_L1	64 3x3x3 convolutions with stride rate of 1x 1 and 1 x1 padding
3.	ReLu	Activation function
4.	Pool_Layer1	Average pooling with the size of 2x2 having output shape of 63, 63, 6
5.	Con_L2	128 3x3x64 convolutions with a stride rate of 1x1 and 1 x1 padding and output shape 61, 61, 16
6.	ReLu	Activation function
7.	Pool_Layer2	Average pooling with the size of 3x3 having stride rate of 2x 2 and 0 padding
8.	Drop out	Drop out layer illustrates an output shape of 30, 30, 16
9.	Flatten	It will flatten the output features into 14400
10.	Dense	It will change the dimension of the vector and shows an output shape of 120
11.	ReLu	Activation function
12.	Dense	It will change the dimension of the vector and shows an output shape of 311
13.	ReLu	Activation function
14.	Softmax_layer	
15.	Classification_Layer	

2.2. E-Alex Net

We have modified the AlexNet model for the face recognition task from the sketches. The AlexNet is a well-known CNN framework due to its high recognition ability and low time complexity [27]. The generic architecture of the AlexNet is demonstrated in Fig. 2. The AlexNet model comprises a total of 5 convolutional and three fully-connected layers where the last layer is concerned to perform the classification task. The model computes a total of 4096 reliable sets of image features which assist it in better recognizing the target class.

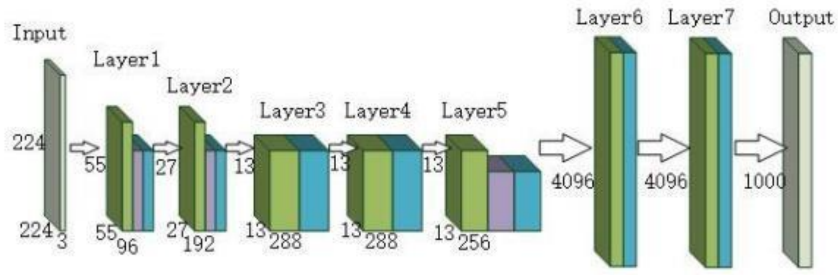

Figure 2: AlexNet structure [28]

We have modified the existing structure of the AlexNet and fine-tuned it for the face recognition task from the sketches. We have used the pertained AlexNet architecture and the feature computation layers are retrained for our datasets to perform the classification task. For this purpose, we have used the concept of transfer learning and replaced the last three layers of the pertained AlexNet by introducing additional dense layers at the end of the model architecture. We have added three fully connected (FC) layers with the filter window of 2048x2048, 1024x1024, and1024x1024 respectively from layer 21 of the fine-tuned AlexNet and named it the E-AlexNet. We have used the Rectified Linear Unit (ReLU) activation function with the added layers to enhance the non-linear problem-solving power of the E-AlexNet as mentioned in [24]. The entire structure of the E-AlexNet is demonstrated in Table 2.

Table 2: Architectural details of E-AlexNet.

No.	Layer	Details
1.	Input	input samples with the resolution of 126 x126 x3
2.	Con_L1	96 11x11x3 convolutions with stride rate of 4 x 4 and 0 padding
3.	ReLu	Activation function
4.	Norm_Layer1	Normalization having 5 channels per element
5.	Pool_Layer1	max pooling with the size of 3x3 having a stride rate of 2x 2 and 0 padding
6.	Con_L2	256 5x5x48 convolutions with stride rate of 1x1 and 2 x2 padding
7.	ReLu	Activation function
8.	Norm_Layer2	Normalization having 5 channels per element
9.	Pool_Layer2	max pooling with the size of 3x3 having a stride rate of 2x 2 and 0 padding

10.	Con_L3	384 3x3x256 convolutions with stride rate of 1x1 and 1 x1 padding
11.	ReLu	Activation function
12.	Con_L4	384 3x3x192 convolutions with stride rate of 1x1 and 1 x1 padding
13.	ReLu	Activation function
14.	Con_L5	256 3x3x192 convolutions with stride rate of 1x1 and 1 x1 padding
15.	ReLu	Activation function
16.	Pool_Layer3	max pooling with the size of 3x3 having a stride rate of 2x 2 and 0 padding
17.	FC_Layer1	FC layer with the size of 4096
18.	ReLu	Activation function
19.	FC_Layer2	FC layer with the dimension of 4096
20.	ReLu	Activation function
21.	Additional_Layer1	FC layer with the dimension of 2048
22.	ReLu	Activation function
23.	Additional_Layer2	FC layer with the dimension of 1024
24.	ReLu	Activation function
25.	Additional_Layer3	FC layer with the dimension of 1024
26.	Softmax_layer	
27.	Classification_Layer	

2.3. E-Le Net

The traditional LeNet model is a gradient-oriented deep learning approach that was originally developed for classifying the numerals [29]. The generic structure of the LeNet model is depicted in Fig. 3 which accepts the image containing digits from 0 to 9 and classifies them in their respective classes. The original LeNet model comprises 5 convolutional layers along with the two pooling layers and one FC layer. The window size in the convolutional layer is set to 5×5 which is 2×2 for the pooling layer. The window size of 5×5 is used in the convolutional layer which is 2×2 in the pooling layer.

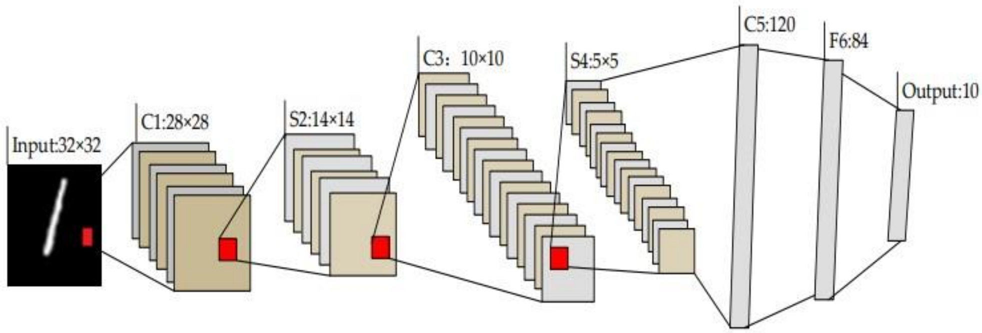

Figure 3: LeNet structure [30]

We have extended the original LeNet model by introducing the additional three dense layers at the end of the model architecture. The modified architecture of the LeNet model is given in Table 3. The added dense layer assists the E-LeNet model to better nominate the reliable set of image features which in turn assist the model to better recognize the human face from the sketches and improve the model accuracy.

Table 3: Architectural details of E-LeNet.

No.	Layer	Details
16.	Input	input samples with the resolution of 126 x126 x3
17.	Con_L1	64 3x3x3 convolutions with stride rate of 1x 1 and 1 x1 padding
18.	ReLu	Activation function
19.	Pool_Layer1	max pooling with the size of 2x2 having a stride rate of 2x 2 and 0 padding
20.	Con_L2	128 3x3x64 convolutions with stride rate of 1x1 and 1 x1 padding
21.	ReLu	Activation function
22.	Pool_Layer2	max pooling with the size of 3x3 having a stride rate of 2x 2 and 0 padding
23.	Con_L3	256 3x3x128 convolutions with stride rate of 1x1 and 1 x1 padding
24.	ReLu	Activation function
25.	Pool_Layer3	max pooling with the size of 3x3 having a stride rate of 2x 2 and 0 padding
26.	Con_L4	512 3x3x256 convolutions with stride rate of 1x1 and 1 x1 padding
27.	ReLu	Activation function

28.	Pool_Layer4	max pooling with the size of 3x3 having a stride rate of 2x 2 and 0 padding
29.	Con_L5	512 3x3x512 convolutions with stride rate of 1x1 and 1 x1 padding
30.	ReLu	Activation function
31.	Pool_Layer5	max pooling with the size of 3x3 having a stride rate of 2x 2 and 0 padding
32.	FC_Layer1	FC layer with the dimension of 4096
33.	Additional_Layer1	FC layer with the dimension of 2048
34.	ReLu	Activation function
35.	Additional_Layer2	FC layer with the dimension of 1024
36.	ReLu	Activation function
37.	Additional_Layer3	FC layer with the dimension of 1024
38.	Softmax_layer	
39.	Classification_Layer	

3. Results

In this part, we have discussed the description of used datasets for face recognition from sketches along with the performance metrics. Besides, a detailed experimental analysis is conducted to show the efficacy of the proposed approach in recognizing the human faces from both hand-drawn and computer-generated sketches.

3.1. Dataset

To test the performance of the introduced approaches we have employed two standard datasets namely the CUHK [22] and AR [31] sketch datasets. The AR dataset comprises a total of 123 samples. While the CUHK database, often known as CUFS, is intended for use in studies pertaining to face sketch recognition and face sketch synthesis. It is comprised of 295 faces from the XM2VTS dataset, 123 faces from the AR dataset, and 188 faces from the student dataset of the Chinese University of Hong Kong (CUHK). In all, there are 606 different people's faces. An artist has created a drawing of each face, and that sketch is based on a photograph of that face taken in a frontal stance, with typical lighting conditions, and with the subject displaying a neutral expression [2, 15]. The selected datasets are quite complex in nature as these contain samples for persons with different age groups, regions, and complexions and samples are suffering from intense light and color variations. Few samples from the AR and CUHK datasets are presented in Fig. 4, and Fig. 5 respectively.

Figure 4: Samples from the AR dataset

Figure 5: Samples from the CUHK dataset

3.2. Evaluation measures

To assess the face recognition ability of the proposed approaches from the sketches, we have used the standard metrics namely accuracy and recall. The mathematical demonstration of the accuracy and recall metrics can be found in Eq.2, and Eq.3 respectively.

$$\text{Accuracy} = \frac{Tp+Tn}{Tp+Fp+Tn+Fn} \quad (2)$$

$$\text{Recall} = \frac{Tp}{Tp+Fn} \quad (3)$$

3.3. Evaluation of models

In this part, we have discussed the results obtained by the three networks namely CNN, E-AlexNet, and E-LeNet models for matching the human faces with sketches via using an experiment. We have computed the accuracy for all three networks and attained results are shown in Table. 4. It can be seen from the reported results in Table 4, that the proposed E-AlexNet model outperformed the other two models for both datasets CUHK, and AR with the values of 92.22%, and 95.18% respectively. While the second better results are shown by the E-LeNet model with the values of 92.18%, and 94.59% for the CUHK and AR datasets respectively. While the CNN exhibits the lowest result for both datasets namely the CUHK and AR with the values of 90.79%, and 91.24% respectively. The graphical comparison for both datasets can be found in Fig. 6.

Table 4: Performance comparison of proposed algorithms.

Model	CUHK	AR
CNN	90.79%	91.24%
E-AlexNet	92.22%	95.18%
E-LetNet	92.18%	94.59%

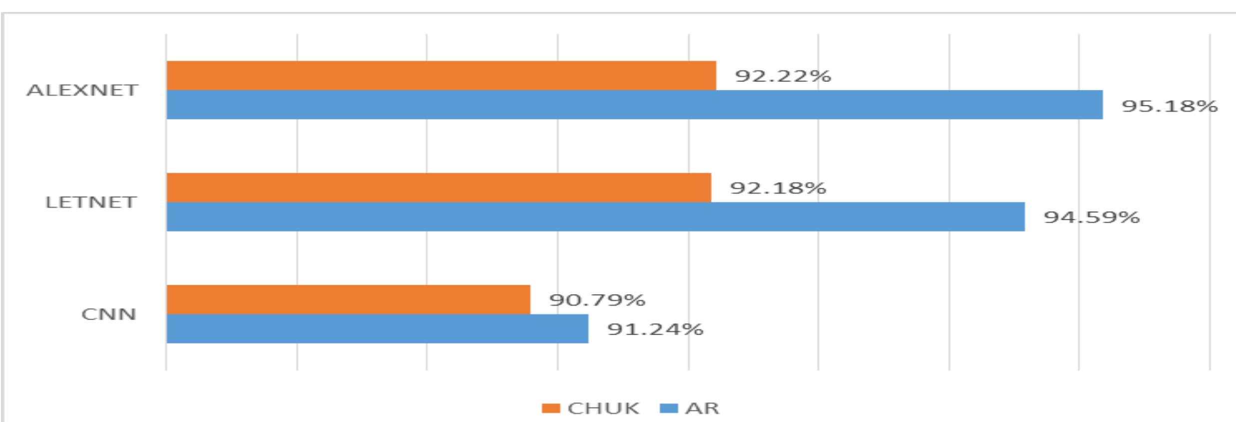

Figure 6: Comparison of proposed methods on the CUHK and AR databases

We have further plotted the confusion matrix for all three proposed models over both datasets namely the CUHK and AR datasets which are reported in Fig. 7 and Fig. 8 respectively. The confusion matrix assists the readers in better understanding the recall rate of a model. It can be seen from the confusion matrixes reported in Fig. 7 and Fig. 8 that the E-AlexNet model performs well for both datasets in comparison to the other two techniques. Hence, it can be said that the EAlexNet approach is more robust to identify the human faces against the sketches due to its high recall rate and can be utilized in automated human face recognition systems from the sketches with high reliability.

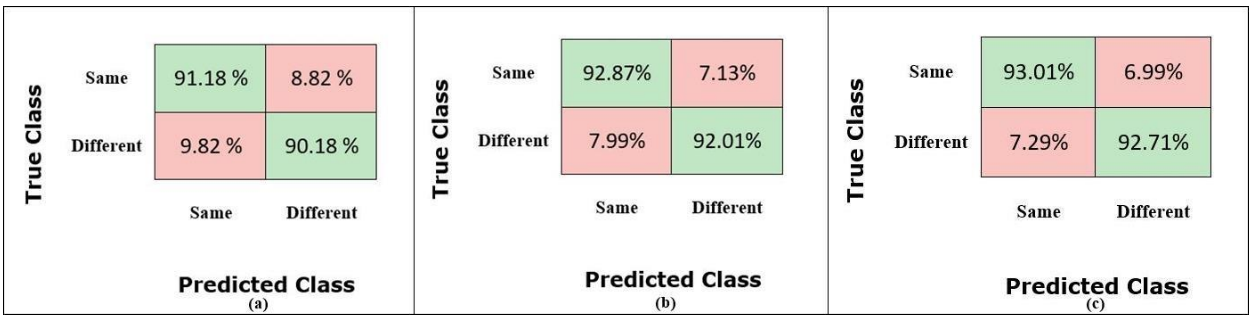

Figure 7: Confusion matrixes over the CUHK dataset, where part (a) is showing the results for the CNN model, part (b)is exhibiting the results for the E-LeNet model while part (c) contains the result values for the E-AlexNet model respectively

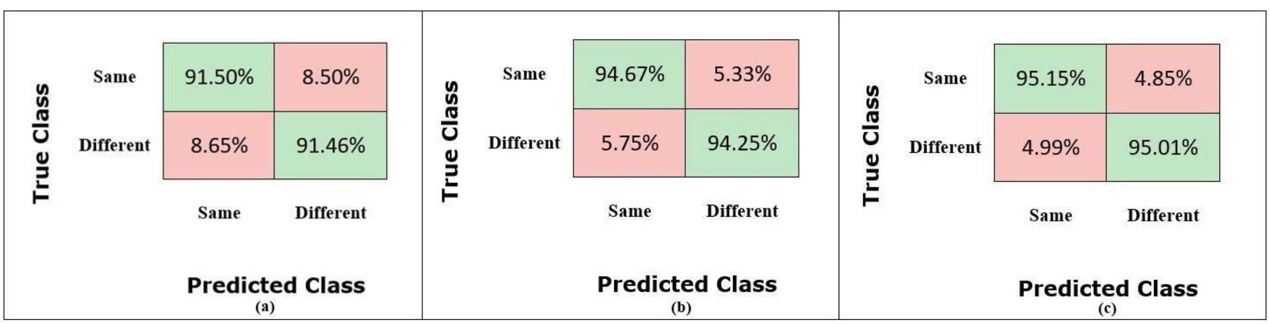

Figure 8:Confusion matrixes over the AR dataset, where part (a) is showing the results for the CNN model, part (b)is exhibiting the results for the E-LeNet model while part (c) contains the result values for the E-AlexNet model respectively

3.4. Comparison with latest approaches

In this part, we have compared the performance of the proposed approach with some latest techniques to show the robustness of our approach to recognizing the human face from the sketches. As we have attained the highest results for the E-AlexNet model over both datasets namely the CUHK and AR datasets, so we have only taken this method for comparison. We have taken many state-of-the-art approaches like locality-constrained representation (LcR) [32], smooth

sparse representation (SSR) [33], thresholding LcR (TLcR) [34], multidimensional scaling (MDS) [35], DCA [36], semi-coupled LR (SLR) [37], locality-constrained feature space learning (LCFSL) [23] as reported in [23] and obtained results are shown in Table 5. One can clearly visualize from the values shown in Table 5 that for both datasets, the E-AlexNet approach has attained the highest performance. More clearly, for the CHUK dataset, the comparative techniques have attained the average value of 84% which is 92.22% for our technique. So, for the CUHK dataset, we have attained an average performance gain of 7.93%. Similarly, for the AR dataset, the respective techniques have acquired the average value of 85%, while in comparison the proposed E-AlexNet technique has obtained the average value of 95.18%. So, we have acquired the average performance gain of 9.75% for the AR dataset. The major reason for the better human face recognition ability of our approach from the sketches is that the added dense layers at the end of the model assist to better nominate the more reliable set of features which in turn improves the recall ability of the proposed approach. While in comparison the competent methods are employing very deep structures which eventually lead to the model over-fitting problem.
Therefore, we can say that our work is more efficient to human face recognition from the sketches.

Table 5: comparison with latest techniques.

Sr. No.	**Method**	**Accuracy**	
		CUHK Dataset	**AR Dataset**
1.	LcR [32]	80%	82%
2.	SSR [33]	81%	81%
3.	TLcR [34]	82%	83%
4.	MDS [35]	85%	86%
5.	DCA [36]	85%	87%
6.	SLR [37]	87%	88%
7.	LCFSL [23]	90%	91%
8.	**Proposed**	92.22%	95.18%

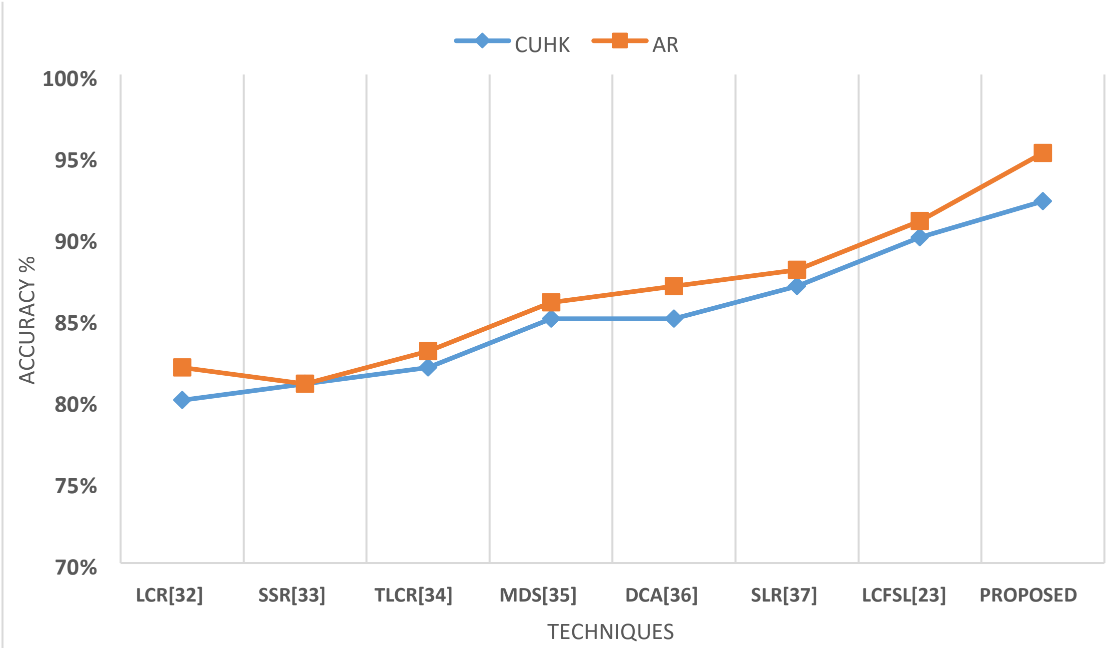

Figure 8: Comparison of results with state of the art techniques

4. Conclusion

In this work, we have adopted a deep-learning-based strategy for identifying the human faces from the sketches. In terms of emotion, texture, color, and projection details, the sketches are vastly different from the actual photos. So to overcome the challenges of existing techniques, we have developed three models namely CNN, E-AlexNet, and E-LeNet model. We have evaluated the proposed solutions on two challenging databases named CUHK and AR to show the effectiveness of the proposed techniques. The comparison shows that our proposed algorithms performed well as compared to the latest techniques. In the future, we plan to look into new methods for face recognition. Additionally, we will use ensemble Deep learning approaches to further improve the performance.

References

Peng, C., X. Gao, N. Wang and J. Li, Face recognition from multiple stylistic sketches: Scenarios, datasets, and evaluation. Pattern Recognition, 2018. **84**: p. 262-272.

Wang, X. and X. Tang, Face photo-sketch synthesis and recognition. IEEE transactions on pattern analysis machine intelligence, 2008. **31**(11): p. 1955-1967.

Han, H., B.F. Klare, K. Bonnen and A.K. Jain, Matching composite sketches to face photos: A component-based approach. IEEE Transactions on Information Forensics Security, 2012. **8**(1): p. 191-204.

Ibrahim, M.R., S.M. Youssef and K.M. Fathalla, Abnormality detection and intelligent severity assessment of human chest computed tomography scans using deep learning: a
case study on SARS-COV-2 assessment. Journal of Ambient Intelligence Humanized Computing, 2021: p. 1-24.

Tao, Q. and R. Veldhuis. Hybrid fusion for biometrics: Combining score-level and decision-level fusion. in 2008 IEEE Computer society conference on computer vision and pattern recognition workshops. 2008. IEEE.

Zhang, J., J. Yang, J. Qian and J. Xu, Nearest orthogonal matrix representation for face recognition. Neurocomputing, 2015. **151**: p. 471-480.

Peng, C., N. Wang, J. Li and X. Gao, DLFace: Deep local descriptor for cross-modality face recognition. Pattern Recognition, 2019. **90**: p. 161-171.

Liu, Q., X. Tang, H. Jin, H. Lu and S. Ma. A nonlinear approach for face sketch synthesis and recognition. in 2005 IEEE Computer Society conference on computer vision and pattern recognition (CVPR'05). 2005. IEEE.

Tang, X. and X. Wang. Face photo recognition using sketch. in Proceedings. International Conference on Image Processing. 2002. IEEE.

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

Ouyang, S., T.M. Hospedales, Y.-Z. Song and X. Li. Forgetmenot: Memory-aware forensic facial sketch matching. in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.

Klare, B., Z. Li and A.K. Jain, Matching forensic sketches to mug shot photos. IEEE transactions on pattern analysis machine intelligence, 2010. **33**(3): p. 639-646.

Peng, C., X. Gao, N. Wang and J. Li, Graphical representation for heterogeneous face recognition. IEEE transactions on pattern analysis machine intelligence, 2016. **39**(2): p. 301-312.

Peng, C., X. Gao, N. Wang and J. Li, Superpixel-based face sketch–photo synthesis. IEEE Transactions on Circuits Systems for Video Technology, 2015. **27**(2): p. 288-299.

Zhang, W., X. Wang and X. Tang. Coupled information-theoretic encoding for face photosketch recognition. in CVPR 2011. 2011. IEEE.

Galoogahi, H.K. and T. Sim. Face sketch recognition by local radon binary pattern: Lrbp. in 2012 19th IEEE International Conference on Image Processing. 2012. IEEE.

Zhang, S., X. Gao, N. Wang, J. Li and M. Zhang, Face sketch synthesis via sparse representation-based greedy search. IEEE transactions on image processing, 2015. **24**(8): p. 2466-2477.

Zhang, M., J. Li, N. Wang and X. Gao, Compositional model-based sketch generator in facial entertainment. IEEE transactions on cybernetics, 2017. **48**(3): p. 904-915.

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

Zhu, J.-Y., W.-S. Zheng, F. Lu and J.-H. Lai, Illumination invariant single face image recognition under heterogeneous lighting condition. Pattern Recognition, 2017. **66**: p. 313327.

Sama, A.B., Face sketch recognition by enhancing evolutionarybased model. International Journal of Computer Technology Applications, 2017. **8**(4): p. 494-498.

Bhatt, H.S., S. Bharadwaj, R. Singh and M. Vatsa. On matching sketches with digital face images. in 2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS). 2010. IEEE.

Gao, G., Y. Wang, P. Huang, H. Chang, H. Lu and D. Yue, Locality-constrained feature space learning for cross-resolution sketch-photo face recognition. Multimedia Tools Applications, 2020. **79**(21): p. 14903-14917.

Wan, W., Y. Gao and H.J. Lee, Transfer deep feature learning for face sketch recognition. Neural Computing Applications, 2019. **31**(12): p. 9175-9184.

Wan, W. and H.J. Lee. Generative adversarial multi-task learning for face sketch synthesis and recognition. in 2019 IEEE International Conference on Image Processing (ICIP). 2019. IEEE.

Roska, T. and L.O. Chua, The CNN universal machine: an analogic array computer. IEEE Transactions on Circuits Systems II: Analog Digital Signal Processing, 1993. **40**(3): p. 163173.

Abd Almisreb, A., N. Jamil and N.M. Din. Utilizing AlexNet deep transfer learning for ear recognition. in 2018 Fourth International Conference on Information Retrieval and Knowledge Management (CAMP). 2018. IEEE.

Yuan, Z.-W. and J. Zhang. Feature extraction and image retrieval based on AlexNet. in Eighth International Conference on Digital Image Processing (ICDIP 2016). 2016. International Society for Optics and Photonics.

El-Sawy, A., H. El-Bakry and M. Loey. CNN for handwritten arabic digits recognition based on LeNet-5. in International conference on advanced intelligent systems and informatics. 2016. Springer.

LeCun, Y., L. Bottou, Y. Bengio and P. Haffner, Gradient-based learning applied to document recognition. Proceedings of the IEEE, 1998. **86**(11): p. 2278-2324.

Peng, C., X. Gao, N. Wang, D. Tao, X. Li, J.J.I.t.o.n.n. Li and l. systems, Multiple representations-based face sketch–photo synthesis. IEEE transactions on neural networks learning systems, 2015. **27**(11): p. 2201-2215.

Jiang, J., R. Hu, Z. Wang and Z. Han, Noise robust face hallucination via localityconstrained representation. IEEE Transactions on Multimedia, 2014. **16**(5): p. 1268-1281.

Jiang, J., J. Ma, C. Chen, X. Jiang and Z. Wang, Noise robust face image super-resolution through smooth sparse representation. IEEE transactions on cybernetics, 2016. **47**(11): p. 3991-4002.

Liu, L., C.P. Chen, S. Li, Y.Y. Tang and L. Chen, Robust face hallucination via localityconstrained bi-layer representation. IEEE transactions on cybernetics, 2017. **48**(4): p. 1189-1201.

Mudunuri, S.P. and S. Biswas, Low resolution face recognition across variations in pose and illumination. IEEE transactions on pattern analysis machine intelligence, 2015. **38**(5): p. 1034-1040.

Haghighat, M. and M. Abdel-Mottaleb. Low resolution face recognition in surveillance systems using discriminant correlation analysis. in 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017). 2017. IEEE.

Asif Raza et.al., 2021. Comparative Analysis of Machine Learning Algorithms for Fake Review Detection. International Journal of Computational Intelligence in Control (13).

Lu, T., X. Chen, Y. Zhang, C. Chen and Z. Xiong, SLR: Semi-coupled locality constrained representation for very low-resolution face recognition and super resolution. IEEE Access, 2018. **6**: p. 56269-56281.