



e-ISSN: 2319-8753 | p-ISSN: 2320-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 6, June 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.512

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Implementation of Face Sketch Recognition Using MTCNN

Ajay Ranaware¹, Abhishek Bhamare², Prashant Chafle³, Sanket Matte⁴, Poonam Sawdekar⁵

U.G. Student, Department of Computer Engineering, Dr. D. Y. Patil Institute of Technology College, Pimpri,

Maharashtra, India^{1,2,3,4}

Assistant Professor, Department of Computer Engineering, Dr. D. Y. Patil Institute of Technology College, Pimpri,

Maharashtra, India⁵

ABSTRACT: People identification has become today a central activity in order to provide custom services to the community and face recognition is one of the most widely used methods to perform such tasks. The wrong prediction of the face leads to be dangerous in the law enforcement field. Face sketch recognition can be used to automatically classify hand-drawn face sketches into known categories. Broad examinations are directed on a face sketch information base including 606 faces, which can be downloaded from Kaggle. MTCNN is used for face detection and FaceNet is used for feature extraction. Finally, the extracted features are to be classified using a KNN.

KEYWORDS: Face Sketch Recognition; Face Sketch Synthesis; MTCNN; Face Net; K-Nearest Neighbors (KNN).

I. INTRODUCTION

Face recognition focuses on still photographs, which can be commonly classified into pictorial and Functional methods. In applications like human-machine interface, automated access control and forensic investigation systems, and so forth, face recognition is popular. Comparing the face picture to the stored faces database to recognize the face in the photo. As demands in areas such as law enforcement, video monitoring is increasing, Authentication, automated face recognition, banking, and security device access attraction. Moreover, the use of face recognition systems is simpler and cheaper, since identification findings can be reversed by people in unclear situations. Without intensive training, without extensive training. The support of law is an essential application of face recognition Implementation. So in short the primary work is to extract the sketch from a given image and compare it to the already hand-made sketch and check the accuracy.

II. RELATED WORK

The generative or intra-modality approach attempts to synthesize photos from input sketches for matching sketches and photos in the same modality or vice versa. Tang and Wang [1] have given the first learning algorithm that transforms the face photos to pseudo- sketches using geometrical measures with the traditional eigenface analysis, and then perform matching techniques by projecting the resulting pseudo sketch into query sketch eigenspace. On the other way around, the authors could also metamorphose the sketch into a pseudo-photo, and then compared the outcome with the face photo. The algorithm has been tested successfully on a gallery of 188 pairs of faces and sketches from the Chinese University of Hong Kong student database or CUHK [2]. In [3], Liu et al synthesized sketches based on the locally linear embedding method. The nonlinear discriminate analysis is used to match the sketch from the synthesized pseudo-sketches and tested over 600 photo-sketch pairs. In [4], Tang and Wang developed an alternate approach synthesizing local face structures at different scales using a learning Markov Random Fields model. A random sampling LDA classifier is utilized to match the pseudo synthesized photo from a sketch with original photos under different conditions of lightings, expressions, and occlusions. In their experimental stage, they build photo-sketch datasets that include 188 faces from the CUHK database, 123 faces from the AR database [5], 295 faces from the XM2VTS database. In [6], Gao et al. proposed Embedded Hidden Markov Model to synthesize a sketch from a photo. For matching purposes, he used the classical eigenface analysis.

III. METHODOLOGY

- Algorithms Used

1] MTCNN

Perform various tasks Cascaded Convolutional Networks (MTCNN) is a system created as an answer for both face discovery. The Network's errand is to yield three things: face/non-face order, jumping box relapse, and facial milestone restriction. There are four steps that are responsible for detecting face:

- 1]Face classification: this is a binary classification problem that uses cross-entropy loss;
- 2]Bounding box regression: the learning objective is a regression problem. For each candidate window, the offset between the candidate and the nearest ground truth is calculated. Euclidean loss is employed for this task.
- 3]Facial Landmark localization: the localization of facial landmarks is formulated as a regression problem, in which the loss function is Euclidean distance.
- 4] There are five landmarks: left eye, right eye, nose, left mouth corner and right mouth corner.

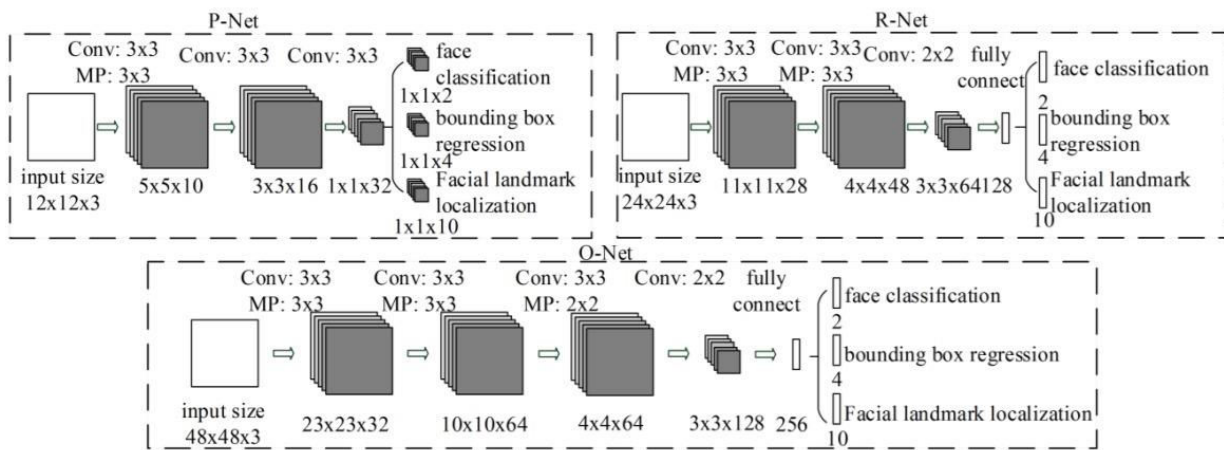


Fig 2. MTCNN Architecture

P-Net stage is responsible for proposing different regions of faces to the next stage.

R-Net (Refine Net) is responsible for refining the proposed bounding boxes from predicted from P-Net.

O-Net uses both P-Net & R-Net before taking the output of R-Net as input for final prediction.

The MTCNN architecture is reasonably complex to implement. Thankfully, there are open source implementations of the architecture that can be trained on new datasets, as well as pre-trained models that can be used directly for face detection.

2] FaceNet

Face Net is a deep neural network used for extracting features from an image of a person's face by converting it to a vector of 128 bytes which represents important features of a face. This vector is called embedding. FaceNet uses the Siamese network. These Embeddings are vectors and we can represent these points in the Euclidean Space. Embeddings can be used to plot an image of a face in the Euclidean system. One possible way of recognizing a person on an unseen image would be to calculate its embedding, calculate distances between two embeddings and if the face embeddings are close enough we say that image contains a person with similar embeddings.

To improve the quality of the embeddings FaceNet uses Triplet loss function which uses three images :

1. Anchor image - An image of a specific person.
2. Negative image - Different person than anchor image.
3. Positive image - Different image of same person used for anchor image

$$l(A,P,N) = \max(0, \|f(A) - f(P)\|^2 - \|f(A) - f(N)\|^2 + a)$$

It trains the neural network such that the difference between the anchor image and positive image is smaller than the difference between anchor image and negative image. It tries to minimize the distance between the anchor image and the positive image.

The FaceNet utilizes a profound convolutional network to straightforwardly streamline the actual installing, instead of a moderate hidden layer as in past profound learning draws near. In the preparation system, the trios of generally adjusted coordinating/non-coordinating with face patches produced by a novel online trio mining technique were embraced. Fig. 3 shows the model construction of the FaceNet. It comprises a group input layer and a profound CNN followed by L2 standardization, which brings about face implanting. The FaceNet straightforwardly prepares its yield to be a conservative 128-measurement installing utilizing a trio put together misfortune work based with respect to LMNN. It comprises of two coordinating with face thumbnails and a non-coordinating with face thumbnail and the misfortune expects to isolate the positive pair from the negative by a distance edge. The thumbnails are tight yields of the face territory, no 2D or 3D arrangement, other than scale and interpretation is performed.

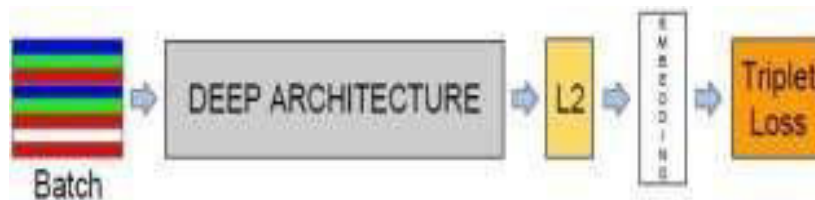


Fig.3 Face Net architecture

3] KNN

A refinement of the k-NN classification algorithm is to weigh the contribution of each of the k neighbors according to their distance to the query point, giving greater weight to closer neighbors. In this method, we are using Euclidean distance as a design metric as shown in equation (1). The KNN classifier suggesting the factory details for every crop based on the nearest distance.

$$euc(x^p, y^p) = \sqrt{\sum_{i=1}^f (x_i - y_i)^2} \quad (1)$$



IV. RESULTS OF THE SYSTEM

In our experimental setup, as shown in table 1, the total numbers of 192 trained images for five faces and 22 new images were tested. These images go through facenet framework by following feature extraction using our image processing module. Then our trained model of classification of sketches using kNN algorithm.

In our experimental setup, we are shown in table, the total numbers of images were 214.

Sr. No.	Category	Number of Images
1	Positive Images	198
2	Negative Images	16

Table 1. Classification of Image



Figure 4: Output Image

These images were then divided into two subcategories; among which 192 predicted and 22 not predicted respectively. The figure shows the predicted image with respect to sketches and recommends the image for it. We classified image data into predicted and not predicted categories based on the accuracy factor which is our main motive.

V. CONCLUSION

With this system, we are provided a user-friendly application that covers aspects like the name of that sketch. Using the challenging face database in which the photos are taken under different illumination conditions and sketches have shape exaggeration as compared to photos, the proposed method still can obtain the best performance. FaceNet was developed to generate feature vectors for every image which can be helpful to quantify face resemblance. FaceNet uses only 128 bytes per face which increases representation efficiency and produces significantly better results over other sketch recognition methods. At the front end, i.e. the User Interface, the input will be the hand drawn sketch by the user. The output will be in the form of a Suggested face of that sketch with the image of that person with the person identity.



REFERENCES

- [1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting", in Proc. SIGGRAPH, pp. 417–424, 2000.
- [2] A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting.", IEEE Transactions on Image Processing, vol. 13, no. 9, pp. 1200–1212, 2004.
- [3] Marcelo Bertalmio, Luminita Vese, Guillermo Sapiro, Stanley Osher, "Simultaneous Structure and Texture Image Inpainting", IEEE Transactions On Image Processing, vol. 12, No. 8, 2003.
- [4] Yassin M. Y. Hasan and Lina J. Karam, "Morphological Text Extraction from Images", IEEE Transactions On Image Processing, vol. 9, No. 11, 2000
- [5] Eftychios A. Pnevmatikakis, Petros Maragos "An Inpainting System For Automatic Image Structure-Texture Restoration With Text Removal", IEEE trans. 978-1-4244-1764, 2008
- [6] S. Bhuvaneswari, T. S. Subashini, "Automatic Detection and Inpainting of Text Images", International Journal of Computer Applications (0975 – 8887) Volume 61– No. 7, 2013
- [7] Aria Pezeshk and Richard L. Tutwiler, "Automatic Feature Extraction and Text Recognition from Scanned Topographic Maps", IEEE Transactions on geosciences and remote sensing, VOL. 49, NO. 12, 2011
- [8] Xiaoqing Liu and Jagath Samarabandu, "Multiscale Edge-Based Text Extraction From Complex Images", IEEE Trans., 1424403677, 2006
- [9] Nobuo Ezaki, Marius Bulacu Lambert, Schomaker, "Text Detection from Natural Scene Images: Towards a System for Visually Impaired Persons", Proc. of 17th Int. Conf. on Pattern Recognition (ICPR), IEEE Computer Society, pp. 683-686, vol. II, 2004
- [10] Mr. Rajesh H. Davda, Mr. Noor Mohammed, "Text Detection, Removal and Region Filling Using Image Inpainting", International Journal of Futuristic Science Engineering and Technology, vol. 1 Issue 2, ISSN 2320 – 4486, 2013
- [11] Uday Modha, Preeti Dave, "Image Inpainting-Automatic Detection and Removal of Text From Images", International Journal of Engineering Research and Applications (IJERA), ISSN: 2248-9622 Vol. 2, Issue 2, 2012
- [12] Muthukumar S, Dr. Krishnan .N, Pasupathi.P, Deepa. S, "Analysis of Image Inpainting Techniques with Exemplar, Poisson, Successive Elimination and 8 Pixel Neighborhood Methods", International Journal of Computer Applications (0975 – 8887), Volume 9, No. 11, 2010



INNO  **SPACE**
SJIF Scientific Journal Impact Factor

**Impact Factor:
7.512**

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details