Journal of Advances in Mathematics and Computer Science

34(6): 1-14, 2019; Article no.JAMCS.53845

ISSN: 2456-9968

(Past name: British Journal of Mathematics & Computer Science, Past ISSN: 2231-0851)

Proposing an Image Enhancement Algorithm Using CNN for Applications of Face Recognition System

Phat Nguyen Huu[1*], Loc Hoang Bao[1] and Hoang Lai The[2]

[1]School of Electronics and Telecommunication, Hanoi University of Science and Technology, 1 Dai Co Viet Road, Hanoi, Vietnam.
[2]Department of Research and Development, Lumi Company, 38 Do Duc Duc, Nam Tu Liem District, Hanoi, Vietnam.

Authors' contributions

This work was carried out in collaboration among all authors. Author PNH designed the study, performed the statistical analysis, wrote the protocol and wrote the first draft of the manuscript. Authors LHB and HLT managed the analyses of the study. Author LHB managed the literature searches. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JAMCS/2019/v34i630230
Editor(s):
(1) Dr. Doina Bein, Assistant Professor, Department of Computer Science, California State University, USA.
Reviewers:
(1) Saad Albawi, University of Diyala, Iraq.
(2) Azween Abdullah, Taylors University, Malaysia.
(3) Raheel Muzzammel, The University of Lahore, Pakistan.
Complete Peer review History: http://www.sdiarticle4.com/review-history/53845

Received: 06 November 2019
Accepted: 12 January 2020
Published: 17 January 2020

Method Article

Abstract

Many researches have been going on since last two decades for object recognition, shape matching, and pattern recognition in the field of computer vision. Face recognition is one of the important issues in object recognition and computer vision. Many face image datasets, related competitions, and evaluation programs have encouraged innovation, producing more powerful facial recognition technology with promising results. In recent years, we have witnessed tremendous improvements in face recognition performance from complex deep neural network architectures trained on millions of face images. Face recognition is the most important biometric and stills many challenges such as pose variation, illumination variation, etc. In order to achieve the desired performance when deploying in reality, the methods depend on many factors. One of the main factors is quality of input image. Therefore, facial recognition systems is installed outdoors which are always affected by extreme weather events such as haze, fog. The existence of haze dramatically degrades the visibility of outdoor images captured in inclement weather and affects many high-level computer vision tasks such as detection and recognition system. In this paper, we propose a preprocessing method to remove haze from input images that

**Corresponding author: E-mail: nhphat@gmail.com, phat.nguyenhuu@hust.edu.vn;*

enhances their quality to improve effectiveness and recognition rate for face identification based on Convolutional Neural Network (CNN) based on the available datasets and our self-built data. To perform the proposed method for outdoor face recognition system, we have improved the system accuracy from 90.53% to 98.14%. The results show that the proposed method improves the quality of the image with other traditional methods.

Keywords: Face recognition; convolutional neural networks; FaceNet; dark channel prior; histogram equalization.

1 Introduction

A facial recognition system (FRS) is capable of identifying or verifying a person from a digital image or a video frame from a video source. There are multiple methods in which facial recognition systems. Face recognition (FR) has been the prominent biometric technique for identifying authentication and widely used in many areas such as military, finance, public security and daily life. Recently progress in this area has been due to two factors: (i) end to end learning for the task using a convolutional neural network (CNN), and (ii) the availability of very large scale training datasets.

CNNs have taken the computer vision community by storm, significantly improving the state of the art in many applications such as FRS. One of the most important ingredients for the success of such methods is the availability of large quantities of training data. In the world of facial recognition system, recognition of face images acquire in an outdoor environment with changing in illumination and/or pose, etc. Besides, there are extreme weather events such as haze, fog, and smoke.

The difficulties in face recognition mainly come from two aspects:

1) Environmental conditions affect the quality of the input image.

2) The method of implementation: Images of outdoor scenes are usually degraded by the turbid medium (e.g., particles, water-droplets) in atmosphere. Haze, fog, and smoke are such phenomena due to atmospheric absorption and scattering. The irradiance receiving by camera from scene point is attenuated along line of sight. Furthermore, the incoming light is blended with air light (ambient light reflected into the line of sight by atmospheric particles). The degraded images lose their contrast and color fidelity. Since the amount of scattering depends on distances of scene points from camera, the degradation is spatial-variant. Haze removal (or dehazing) is highly desired in both consumer/computational photography and computer vision applications. Firstly, removing haze can significantly increase the visibility of scene and correct color that are caused by air light. In general, the haze-free image is more visually pleasing. Secondly, computer vision algorithms perform on from low-level image analysis to high-level object recognition. Therefore, the haze removal can produce depth information and benefit many vision algorithms and advanced image editing.

In the past, the field of machine learning has undergone several major developments. One important advancement is a technique known as “deep learning” that aims to model the high-level data abstractions by employing deep networked architectures composing of multiple linear/non-linear transformations. It has made a remarkable impact in computer vision performance previously unattainable on many tasks such as image classification, object detection, and especially face recognition. Despite significant recent advances in the field of face recognition, there are different approaches for using the CNN. However, implementing face verification and recognition efficiently at scale presents serious challenges to current approaches. In this paper, we employ a system called FaceNet based on [1] to perform face recognition. FaceNet is a face recognition system that was described by Florian Schroff, et al. at Google in 2015. It is a system that will extract high-quality features from the face and predict a 128 element vector representation (these features are called a face embedding) and inserts them into a feature space. The model is a deep CNN trained via a triplet loss function that encourages vectors for the same identity to become more similar (smaller distance), whereas vectors for different identities are expected to become less similar (larger distance).

A facial recognition system that can identify people at significantly higher accuracy than a human is very value. It can generate a shortlist for the human who can then make a final decision for further action. The accuracy of such a system will depend on the main factors such as quality of images. Consequently, the purpose of this paper is to propose a solution which improves performance of systems. By combining the two main algorithms are histogram equalization and dark channel prior [1] to remove haze from a single input image, we can recover a high quality haze -free image and produce a good depth map to process image input of systems which aims at a facial recognition system.

2 Related Work

In the context of computational photography, there is an increasing focus on developing methods that restore images as well as extracting other quantities at minimal requirements in terms of input data, user intervention, and sophistication of acquisition hardware. Therefore, preprocessing to eliminate noise, recover images, improve and enhance the quality of input images is often studied and proposed in the facial recognition system to aims increase in accuracy or minimize calculations to deploy in real time in the most optimal way. Haze removal is a challenging problem since it is dependent on the unknown depth information.

There are many methods that have been proposed by using multiple images or additional information. Polarization [1] removes the haze effect through two or more images taken with different degrees. In [4-6] more constraints are obtained from multiple images of the same scene under different weather conditions. The dark-object subtraction method [7] subtracts a constant value, corresponding the darkest object in scene from each band. Recently, single image haze removal [7][7] has made significant progresses. The success of these methods lies in using a stronger prior or assumption. Tan [7] shows that the haze-free image must have higher contrast comparing with input haze image. Therefore, they removes the haze by maximizing local contrast of restored image. Coming up with contrast enhancement can be obtained by using histogram equalization.

Deep CNN has recently demonstrated outstanding performance in variety of vision tasks such as face recognition object detection and classification [11],[21],13]. Deep learning applies multiple processing layers to learn representation of data with multiple levels of feature extraction. This emerging technique has reshaped the research landscape of face recognition (FR) and launched by the breakthroughs of Deep face method. Therefore, deep FR technique, which leverages hierarchical architecture to stitch together pixels into invariant face presentation, has dramatically improved the state-of-the-art performance and fostered successful real-world applications. Two deep network architectures that have been recently are used to great success in the computer vision community. Both are deep convolutional network [1],[14].

The first architecture is based on model consists of multiple interleaved layers of convolutions, non-linear activations, local response normalizations, and max pooling layers [14]. We additionally add several $1\times1\times d$ convolution layers inspired by the work of [14]. The second architecture is based on the Inception model of Szegedy et al. which was recently used as for Image Net 2014 [7]. The works of [3],[14],[17] all employ complex system of multiple stages that combine the output of deep convolutional network with PCA for dimensionality reduction and an SVM for classification.

Sun et al. [5,14] propose a compact for compute network. They use an ensemble of 25 of these networks and each operating on a different face patch. For their final performance on LFW (99:47%) [9] the authors combine 50 responses (regular and flipped). Both PCA and a Joint Bayesian model [17] that effectively correspond to a linear transform in the embedding space are employed. Their method does not require explicit 2D/3D alignment. The networks are trained by using a combination of classification and verification loss. The verification loss is similar to the triplet loss. We employ [6],[14],[7] in that and it minimizes the

L2-distance between faces of the same identity and enforces a margin between the distance of faces of different identities. The main difference is that only pairs of images are compared, whereas the triplet loss encourages a relative distance constraint. The networks are trained by using a combination of classification and verification loss. The verification loss is similar to the triplet loss we employ [6] that it minimizes the L2-distance between faces of the same identity and enforces a margin between the distance of faces of different identities. The main difference is that only pairs of images are compared whereas the triplet loss encourages a relative distance constraint which used in systems.

This paper is organized as follows. In section III, we present approach for image restoration, haze removal and face recognition using FaceNet. The results are present in Section IV. In Section V we summarize our approach and discuss its limitations.

3 Methodology

Based on the requirements and objectives above, we propose the system which consists of two main blocks:

(i) Preprocessing image (Module A),
(ii) Face recognition (Module B).

System structure diagram is illustrated in Fig. 1. More detail the blocks is described in Fig. 1.

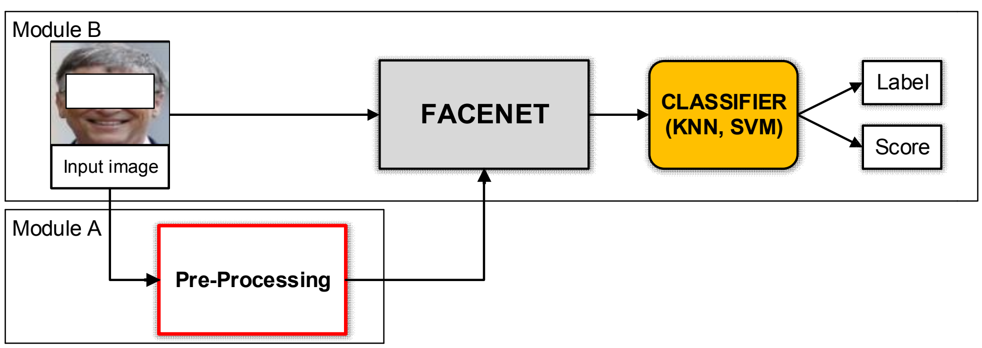

Fig. 1. Structure of proposal system

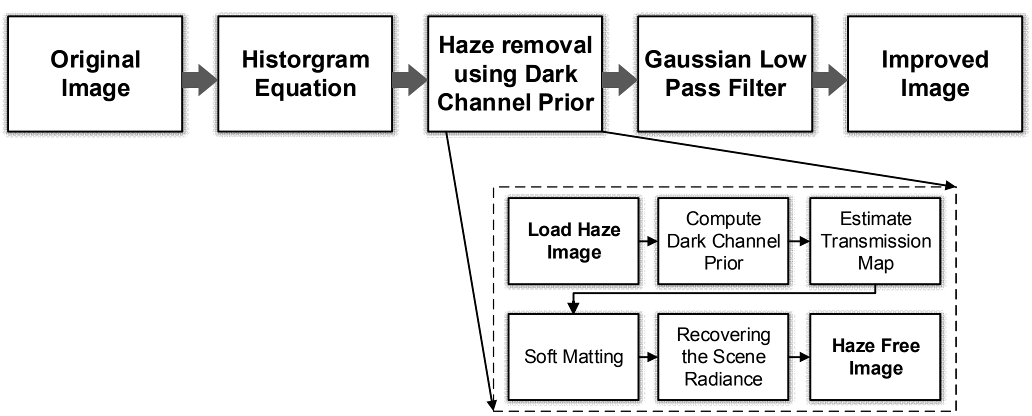

Fig. 2. Diagram of preprocessing steps

3.1 Module A: Preprocessing image

The pre-processing of input image will be made through the following steps as shown in Fig. 2 with the initially established purpose.

By combining image processing algorithms sequentially, we improve input image for facial system. The algorithms use histogram equation, Gaussian low pass filter, and Dark channel prior (DCP). In this paper, we focus on haze removal using the Dark channel prior algorithm, and propose an improvement by increasing patch size of dark channel prior which is demonstrated below.

Haze (or fog, mist, and other atmospheric phenomena) is a main degradation of outdoor image that is weak both colors and contrasts. In computer vision and computer graphics, the model widely used to describe the formation of a haze image is as follow [4,16,19,20]

$$I(x) = J(x)t(x) + A(1-t(x)) \tag{1}$$

where I is the observed intensity, J is the scene radiance, A is the global atmospheric light, and t is the medium transmission describing the portion of light that is not scattered and reaches by camera. The goal of haze removal is to recover J and A.

The first parameter (J(x)t(x)) of Equation (1) is called direct attenuation [7], and the second parameter (A(1-t(x))) is called air light [14], [17]. Direct attenuation describes the scene radiance and its decay in the medium. This model (model A) explains the loss of contrasts due to haze as the average result of image with constant color. For a patch with uniform transmission t, the visibility (sum of gradient) of input image is reduced by the haze, since t<1:

$$\sum_x \| \nabla I(x) \| = t \sum_x \| \nabla J(x) \| < \sum_x \| \nabla I(x) \| \tag{2}$$

The transmission t in local patch is estimated by maximizing the visibility of patch and satisfying a constraint where the intensity of J(x) is less than that of A.

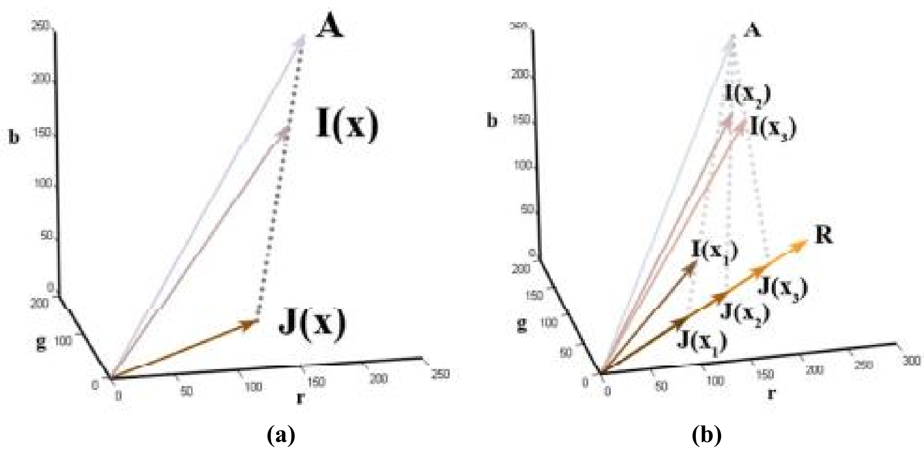

Fig. 3. (a) Left: Haze image formation model. (b) Right: Constant albedo model used in Fattal's work [17]

The dark channel prior is based on the following observation on haze-free outdoor images. In most of non-sky patches, at least one color channel has very low intensity at several pixels. The formation of an image J is calculated as shown in Fig. 3 as [7]:

$$J^{dark} = \min_{c\in\{R,G,B\}}\left(\min_{y\in\Omega(x)}(J^c(x))\right) \quad (3)$$

where J^c is a color channel of J and $\Omega(x)$ is a local patch at *x*. Based on the statistical observation which is collected an outdoor, images are selected from [21] and 150 most popular tags annotated by users. We randomly select 5000 images and manually cut out the sky regions. According to the observation, the intensity of J^{dark} is low and tends to be zero excepting sky region, if J is haze-free outdoor image [7]. Therefore, J^{dark} is called the dark channel of J. In other word, the pixel value at the dark channel can be approximated as follows:

$$J^{dark} \approx 0 \quad (4)$$

Due to the additive air light, a haze image is brighter than its haze-free version in where the transmission t is low. Therefore, the dark channel of haze image will have higher intensity in regions with denser haze.

The minimum intensity in local patch of each color channel is taken after dividing Eq. (1) by A^c as follows:

$$\min_{y\in\Omega(x)} \frac{I^c(y)}{A^c} = \tilde{t}(x) \min_{y\in\Omega(x)} \frac{J^c(y)}{A^c} + (1-\tilde{t}(x))$$
(5)

where the transmission in local patch $\Omega(x)$ is assumed to be constant and is represented as $\tilde{t}(x)$ [5]. The min operator of the three color channels can then be applied to Eq. (2) as follows:

$$\min_{y\in\Omega(x)}\left(\min_c \frac{I^c(y)}{A^c}\right) = \tilde{t}(x) \min_{y\in\Omega(x)}\left(\min_c \frac{J^c(y)}{A^c}\right) + (1-\tilde{t}(x)) \quad (6)$$

According to DCP approximation of Eq. (3), $\tilde{t}(x)$ can be represented as:

$$\tilde{t}(x) = 1 - \min_{y\in\Omega(x)}\left(\min_c \frac{I^c(y)}{A^c}\right) \quad (7)$$

The atmospheric light A needs to be estimated in order to obtain the transmission map t. The pixel value of dark channel is first selected, and the color with the highest intensity value among selecting pixels is then used as value for A. It is noted in [11] that DCP is not reliable in the sky region. Therefore, the color of sky is close to A in haze images, and is calculated by:

$$\min_{y\in\Omega(x)}\left(\min_c \frac{I^c(y)}{A^c}\right) \approx 1 \text{ and } \tilde{t}(x) \approx 0 \quad (8)$$

In [7], they estimate transmission map from an input haze image using the patch size 15x15. It is reasonably good but containing several block effects since the transmission is not always constant in a patch. Therefore, this map is a soft-matting method [21]. The original DCP based on dehazing algorithm adopts soft matting to refine the transmission map. The degradation model in Eq. (1) is similar to the soft-matting. The following cost function is calculated by:

$$E(t) = t^T Lt + \lambda(t-\tilde{t})^T(t-\tilde{t}) \quad (9)$$

where L is the matting Laplacian matrix proposed by [21], and $\lambda$is a regularization parameter. The first parameter is the smooth and the second is the data.

The optimal t can be obtained by solving the following sparse linear system as:

$$(\mathrm{L}+\lambda U)\,\mathrm{t}=\lambda\,\tilde{t} \tag{10}$$

where U is an identity matrix of same size as L and $\lambda = 10^{-4}$ is used in our experiments.

In our experiment instead of using large size, we're going to use small size such as 7x7. This proposal will significantly reduce block effect. In several cases, we have recovered the scene radiance with transmission without needing through soft matting. Therefore, we decrease processing time which doesn't affect the results. Fig. 4 is example of image after performing Haze removal. The results will present in the simulation section.

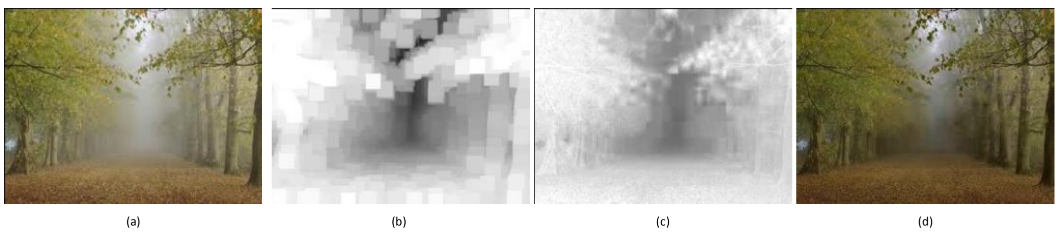

Fig. 4. Haze removal result: (a) input haze image [21], (b) estimated transmission map, (c) refined transmission map after soft matting, (d) final haze-free image [21]

3.2 Module B: Face recognition

A FRS consists of four modules as depicted in Fig. 5: detection, alignment, feature extraction, and matching where face detection segments the areas from background. Face alignment is aimed at achieving more accurate localization and at normalizing faces where face detection provides coarse estimates of location and scale of each detected face. After normalizing face geometrically and photometrically, feature extraction is performed to provide effective information that is useful for distinguishing between faces of different persons and stable with respect to the geometrical and photometrical variations and face matching. The extracting feature vector of input face is matched against those of enrolled faces in database. It output the identity of face when a match is found with sufficient confidence or indicates an unknown face.

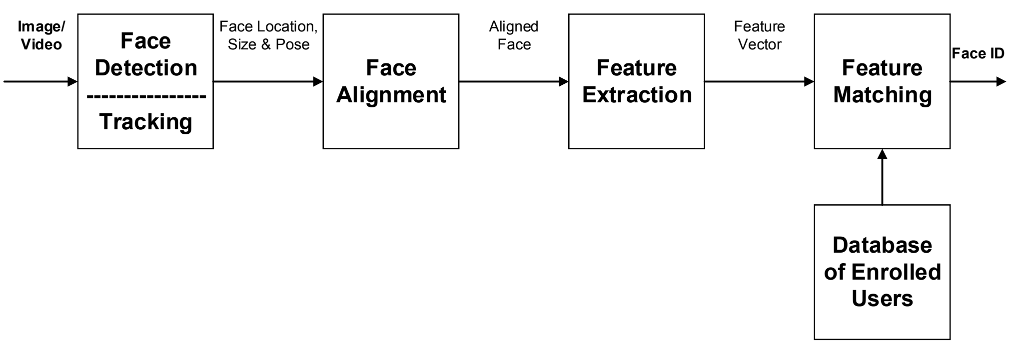

Fig. 5. Face recognition processing flow based on [21]

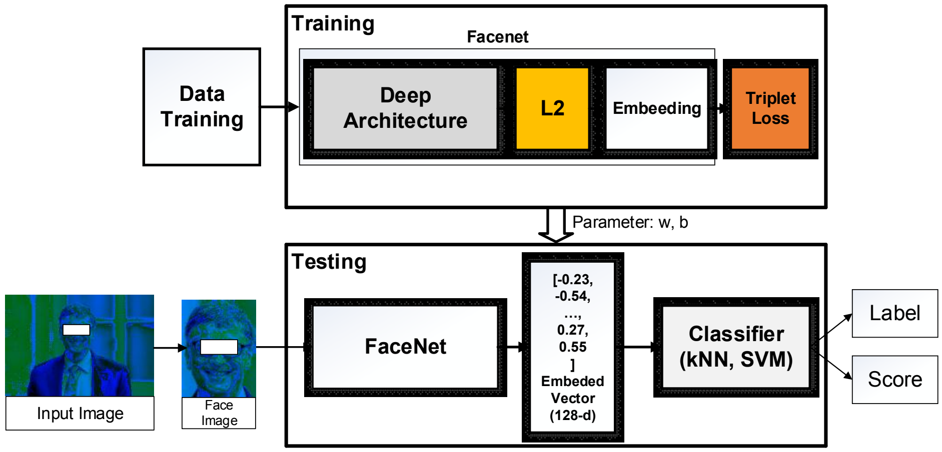

Fig. 6. Face recognition structure diagram

In this paper, we will develop a face recognition system as shown in Fig. 6. We will use the Multi-Task Cascaded Convolutional Neural Network (MTCNN) for face detection. FaceNet model will be used to create a face embedding for each detected face. We then will develop a linear Support Vector Machine (SVM) classifier model to predict a face.

3.2.1 Face detection and alignment

MTCNN is state-of-the-art deep learning model for face detection describing in [23]. The network uses a cascade structure with three networks. Firstly, the image is rescaled to a range of different sizes (called an image pyramid). The first model (Proposal Network or P-Net) then selects candidate facial regions, and the second model (Refine Network or R-Net) filters bounding boxes, and the third model (Output Network or O-Net) proposes facial landmarks. It is described in Fig. 7. The three models are not connected directly. Outputs of previous stage are fed as input to next stage. This allow additional processing to be performed between stages. The output of MTCNN is a list of "dict" object. Each of them provides a number of keys for detecting face that includes:

1) 'box' provides x and y of the bottom left of bounding box as well as its width and height
2) 'confidence' is the probability confidence of the prediction
3) 'keypoints' provides a "dict" with their features such as 'left_eye', 'right_eye', 'nose', 'mouth_left', and 'mouth_right'.

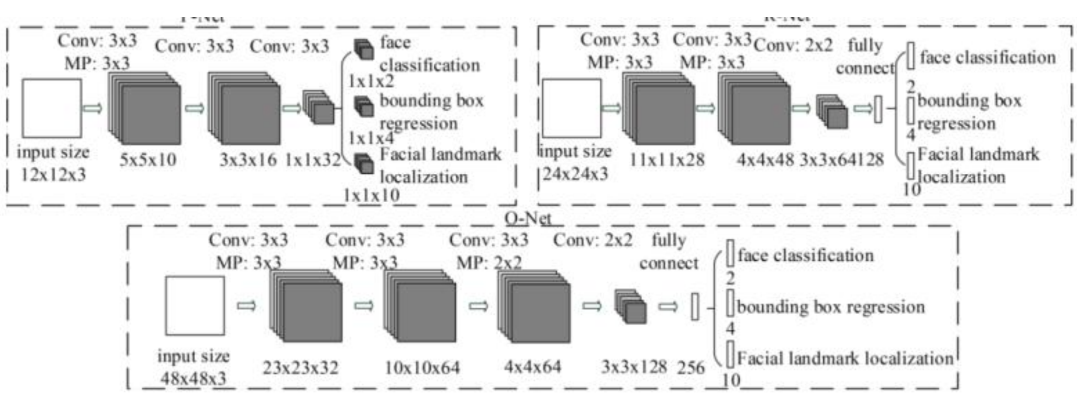

Fig. 7. The architectures of P-Net, R-Net and O-Net [21]

3.2.2 FaceNet

FaceNet is a deep CNN trained via a triplet loss function that encourages vectors of the same identity to become more similar (smaller distance). The vectors with different identities are expected to become loess similar (larger distance). The training model is an important innovation. The main difference between FaceNet and other techniques is mapping from images and creating embeddings rather than using any bottleneck layer for recognition or verification tasks. The embeddings are created all the other tasks like verification, recognition, etc.

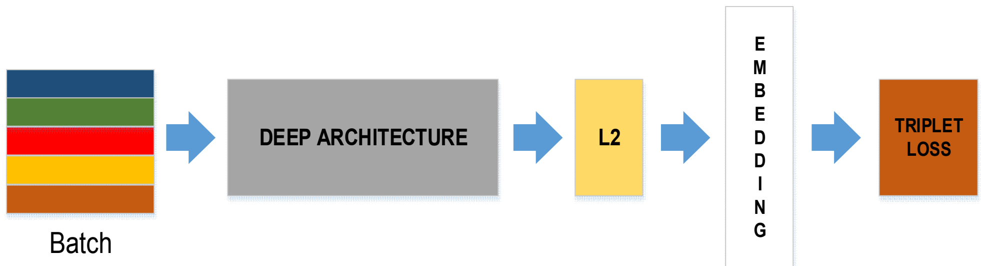

Fig. 8. Model structure of FaceNet [1]

This system employs a particular loss function called the triplet loss. The triplet loss minimizes the L_2 distance between images of same identity and maximizes L_2 distance between different characters of face images. It is described in Fig. 8. The creator devises an efficient triple selection mechanism which smartly selects three images at a time. These images consist of three types, namely anchor (an image of a random person), positive image (another image of the same person), and negative (an image of a different person). Choosing the correct image pairs is extremely important. In order to ensure fast convergence, it is crucial to select triplets that violate their constraint.

The loss that is being minimized as:

$$\sum_{i}^{N}[\| f(x_i^a)-f(x_i^p)\|_2^2 - \| f(x_i^a)-f(x_i^n)\|_2^2 + \alpha] \quad (11)$$

One of the network architectures used in facenets is based on GoogLeNet style Inception models [1]. These models have around 6.6M to 7.5M parameters. According to [1], authors train the CNN using Stochastic Gradient Descent (SGD) with standard backprop and AdaGrad. In this paper, the authors begin to train with a learning rate of 0.05 and descending. Facenet results achieved an accuracy of about 99.63% with LFW dataset and 95.13% with Youtube Faces DB [1].

3.3 Peak Signal - to- Noisy Ration (PSNR) and Mean Squared Error (MSE)

We try to improve maximum possible power of a signal and corrupting noise that effects the fidelity of its representation. Peak Signal- to- Noisy Ration (PSNR) is usually expressed in term of logarithmic.

PSNR is defined via mean squared error (MSE). Given a noise-free m $\times$n monochrome image/ and its noisy approximation K, MSE is defined as:

$$MSE = \frac{1}{mxn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[\mathrm{I}(i,j)-\mathrm{K}(i,j)]^2 \quad (12)$$

The PSNR (in dB) is defined is:

$$PSNR = 10\log_{10}\left(\frac{MAX_I^2}{MSE}\right) \tag{13}$$

where, MAX_I is the maximum pixel value of image. When the pixels are represented using 8 bits per sample, it is 255. When samples are presented using linear PCM with B bits per sample, MAX_I is $2^B - 1$. The image input of system is color images. For color images with three RGB values per pixel, MSE is the sum over all squared value differences divided by image size and by three.

PSNR and MSE are used to comparing the squared error between the original image and reconstructed image. There is an inverse relationship between PSNR and MSE. A high PSNR value indicates the good quality of image.

4 Experiments

To evaluate the effectiveness of proposal system, we calculate two parameters for two models:

1) As for image quality improvement, we use MSE and PSNR.
2) About face recognition, we perform through the corresponding top-k scores.

4.1 Evaluations on dehazing

Fig. 9 shows our haze removal results. In Table 1, we compare the result of proposal with other methods. Experiments were carried out on image data sets collected from NYU2 Depth Database [24] and Middle-bury stereo database [25]. The data are scenes of nature, streets, construction sites, aerial images etc. that are affected by haze and fog. We set different values of atmospheric light A in the range [0.6, 1.0] and select values [0.4, 1.4]. In addition, we added an outdoor dataset that we collected from the internet ourselves to evaluate the effectiveness of the method.

We then compared the proposed method with several state-of-the-art dehazing methods such as Fast Visibility Restoration (FVR) [26], Boundary Constrained Context Regularization (BCCR) [27], Automatic Atmospheric Light Recovery (ATM) [7]. The synthesized hazing images are accompanied with ground-truth images. Table 1 displays the average of PSNR and MSE.

Based on results, we found that the proposed method has greatly improved the accuracy (with PSNR value increasing by nearly 10dB) without changing algorithm complexity and additional preprocessing module.

4.2 Evaluate effectiveness of proposed method

To evaluate the effectiveness of the proposed pretreatment method with the facial recognition system, we use face image data. We use LFW dataset in facenet [1]. This dataset contains 13233 photos of 5749 people. The images are then randomly generated with fog and smoke effects.

In addition, we added an image dataset of 100 self-built people. Data containing the human face was taken in the context of fog, haze like nature. In this paper, we focus on improving the accuracy of the system after pre-processing the input image for the recognition system.

Evaluation results for model B are shown in Fig. 10. Top-k scores results in two cases with and without pretreatment are 98.14 and 90.53, respectively. In the case of preprocessing, the coefficient *k* is improved up to 9%.

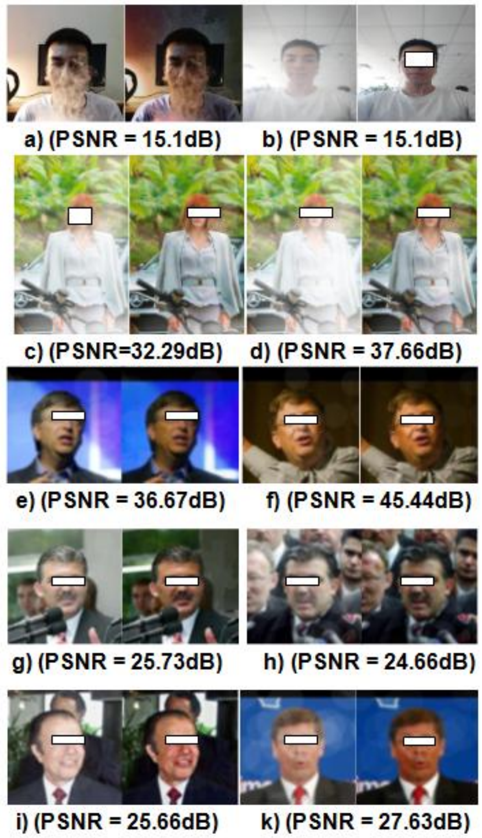

Fig. 9. Results of quality image after recovering

Table 1. Comparing result of the proposal method with other methods

Value	Proposal Method	ATM [7]	NLD [28]	BCCR [27]	FVR [26]
MSE	650.25	4794.4	2130.6	917.2	849.23
PSNR	20.19	11.33	14.8	18.51	18.84

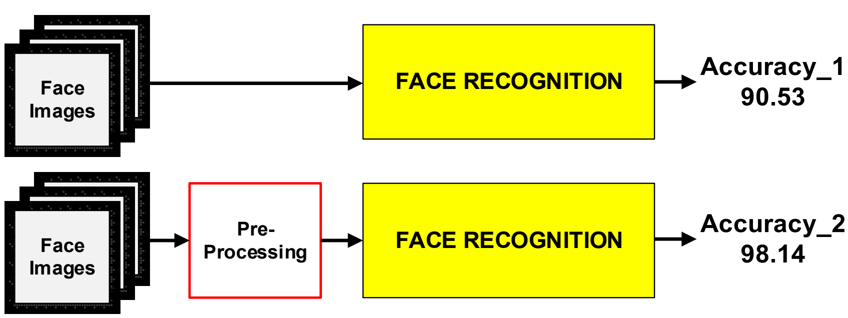

Fig. 10. Before and after preprocessing results

5 Conclusion

In this paper, we proposed a preprocessing solution that enhances the input image quality for outdoor systems. However, the input images always have challenges of lighting conditions, diverse angles, it loses details. In particular, the faces in these images are usually small with low resolution. In the future, we will improve the resolution by applying state-of-the-art methods developed in Deep Learning or GAN structures to solve the task.

Disclaimer

The products used for this research are commonly and predominantly use products in our area of research and country. There is absolutely no conflict of interest between the authors and producers of the products because we do not intend to use these products as an avenue for any litigation but for the advancement of knowledge. Also, the research was not funded by the producing company rather it was funded by personal efforts of the authors.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Schroff, Florian, Dmitry Kalenichenko, and James Philbin. “FaceNet: A Unified Embedding for Face Recognition and Clustering.” 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR): n. pag. Crossref. Web; 2015.

[2] Szegedy C, et al. Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA. 2015;1-9.
DOI: 10.1109/CVPR.2015.7298594
Available:http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7298594&isnumber=7298593

[3] Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. CoRR, abs/1311.2901; 2013.

[4] Narasimhan SG, Nayar SK. Contrast restoration of weather degraded images," in IEEE Transactions on Pattern Analysis and Machine Intelligence. 2003;25(6):713-724.
DOI: 10.1109/TPAMI.2003.1201821
Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1201821&isnumber=27059

[5] Schechner YY, Narasimhan SG, Nayar SK. Instant dehazing of images using polarization," Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, Kauai, HI, USA. 2001; I-I.
DOI:10.1109/CVPR.2001.990493
Available:http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=990493&isnumber=21353

[6] Taigman Y, Yang M, Ranzato M, Wolf L. DeepFace: Closing the gap to human-level performance in face verification," 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH. 2014;1701-1708.
DOI: 10.1109/CVPR.2014.220
Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6909616&isnumber=6909393

[7] Sulami M, Glatzer I, Fattal R, Werman M. Automatic recovery of the atmospheric light in hazy images," in 2014 IEEE International Conference on Computational Photography (ICCP). 2014; 1–11.

[8] He K, Sun J, Tang X. Single image haze removal using dark channel prior," in IEEE Transactions on Pattern Analysis and Machine Intelligence. 2011;33(12):2341-2353.
DOI:10.1109/TPAMI.2010.168
Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5567108&isnumber=6047995

[9] Rumelhart DE, Hinton GE, Williams RJ. Neurocomputing: Foundations of research," J. A. Anderson and E. Rosenfeld, Eds. Cambridge, MA, USA: MIT Press, ch. Learning Representations by Back-propagating Errors. 1988;696–699. \
Available: http://dl.acm.org/citation.cfm?id=65669.104451

[10] Tan RT. Visibility in bad weather from a single image," in 2008 IEEE Conference on Computer Vision and Pattern Recognition. 2008;1–8.

[11] Kilian Q. Weinberger and Lawrence K. Saul. Distance Metric Learning for Large Margin Nearest Neighbor Classification. J. Mach. Learn. Res. 2009;10:207-244.

[12] Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7553523&isnumber=7539715

[13] Albawi, Saad, Abed Mohammed, Tareq and Alzawi, Saad. Understanding of a Convolutional Neural Network; 2017.
DOI: 10.1109/ICEngTechnol.2017.8308186.

[14] Sun Y, Wang X, Tang X. Deeply learned face representations are sparse, selective, and robust. CoRR. abs/1412.1265; 2014.

[15] LeCun Y, et al. Backpropagation applied to handwritten zip code recognition," in Neural Computation. 1989;1(4):541-551.
DOI: 10.1162/neco.1989.1.4.541
Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6795724&isnumber=6795285

[16] Sun Y, Wang X, Tang. X Deep learning face representation by joint identification-verification. CoRR, abs/1406.4773; 2014.

[17] Tarel J, Hautire N. Fast visibility restoration from a single color or gray level image," in 2009 IEEE 12th International Conference on Computer Vision. 2009;2201–2208.

[18] Chen D, Cao X, Wang L, Wen F, Sun J. Bayesian face revisited: A joint formulation. In Proc. ECCV; 2012.

[19] Narasimhan SG, Nayar SK. Chromatic framework for vision in bad weather, in Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662). 2000;1:598–605.

[20] Fattal R. Single image dehazing," ACM Trans. Graph. 2008;27(3):72:1–72:9.
Available: http://doi.acm.org/10.1145/1360612.1360671

[21] Stan Z. Li, Anil K. Jain. Handbook of Face Recognition (2nd ed.). Springer Publishing Company, Incorporated; 2011.

[22] A Levin, D Lischinski, Y Weiss. A closed-form solution to natural image matting," in IEEE Transactions on Pattern Analysis and Machine Intelligence. 2008;30(2):228-242.
DOI: 10.1109/TPAMI.2007.1177
Available:http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4359322&isnumber=4407426

[23] Zhang K, Zhang Z, Li Z, Qiao Y. Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks," in IEEE Signal Processing Letters. 2016;23(10):1499-1503.
DOI: 10.1109/LSP.2016.2603342

[24] Tarel JP, Hautiere N. Fast visibility restoration from a single color or gray level image". In Computer Vision, IEEE 12th International Conference. 2009;2201–2208.

[25] Scharstein D, Szeliski R. High-accuracy stereo depth maps using structured light. In Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference. 2003;1:I–I.

[26] About Flickr.
Available: https://flickr.com/

[27] Meng G, Wang Y, Duan J, Xiang S, Pan C. Efficient image dehazing with boundary constraint and contextual regularization," in 2013 IEEE International Conference on Computer Vision. 2013;617–624.

[28] Berman D, Treibitz T, Avidan S. Non-local image dehazing," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016;1674–1682.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://www.sdiarticle4.com/review-history/53845