



Face illumination normalization based on generative adversarial network

Dequan Guo¹ · Lingrui Zhu¹ · Shenggui Ling^{2,3} · Tianxiang Li⁴ · Gexiang Zhang¹ · Qiang Yang¹ · Ping Wang⁵ · Shiqi Jiang¹ · Sidong Wu¹ · Junbao Liu¹

Accepted: 7 May 2022 / Published online: 20 July 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

Face recognition technology has been widely used in the field of artificial intelligence. The technology needs to be carried out normally under the appropriate light, however, there is not ideal light, even poor-lighted for the face recognition device, and with the head in deflect angle. The poor-lighted under various head poses will influence the face recognition significantly. To address the issue, we present a novel and practical architecture based on deep fully convolutional neural network and generative adversarial networks for illumination normalization of facial images. The proposed method is termed as illumination normalization generative adversarial network. Compared with previous methods based on deep learning, our approach does not require identity and illumination label as input. We preserve identities of faces by an elaborately-designed generator together with content loss. Moreover, the framework of our scheme is simpler than previous methods based on deep learning. It can address the illumination of frontal and non-frontal face. In order to fairly evaluate the proposed method against state-of-the-art models, the peak signal to noise ratio is employed to estimate the performance of illumination normalization algorithm. Experimental results show that the proposed method achieves favorable normalization results against previous models under various head poses and illumination challenges.

Keywords Face recognition · Illumination normalization · Generative adversarial networks · Loss function

1 Introduction

It is well known that illumination is an important factor to perform computer vision task. As is shown in Fig. 1, some reasons, such as the excessive exposure and the lack of exposure of the camera, the intensity and direction of the

light, could make the lighting conditions complicated. Face appearances can change dramatically due to illumination variations. This fact will cause: the variations between the images of the same face since illumination are almost always larger than image variations due to changes in face identities. Zhang et al. (2021) proposed a recursive copy and paste generative adversarial network (Re-CPGAN) to recover authentic high-resolution face images while compensating for non-uniform illumination. A depth algorithm was proposed for a single-lens occluded portrait to estimate the actual portrait distance for different poses, angles of view and obscuration (Tsai et al. 2020). A network analysis approach was carried out, depicting sub-communities of nodes related to face (Moret-Tatay et al. 2020). There are some still challenges in illumination (Adini et al. 1997). Therefore, illumination normalization of face is essential and valuable. In this study, we focus on the task of illumination normalization of facial images under different illumination conditions and a variety of head poses.

✉ Shenggui Ling
280573298@qq.com

¹ School of Automation, Chengdu University of Information Technology, Chengdu 610225, Sichuan, China

² Neijiang Vocational and Technical College, Neijiang 641000, Sichuan, China

³ National Key Laboratory of Fundamental Science on Synthetic Vision, Sichuan University, Chengdu 610065, Sichuan, China

⁴ Chengdu Industry and Trade College (Chengdu Technician College), Chengdu 611731, Sichuan, China

⁵ Communication Engineering School, Chengdu Technological University, Chengdu 611730, Sichuan, China

Fig. 1 Some poor-lighted faces under various head poses



As a pioneering work, Faisal et al. (2006) combine Phong's lighting model and a 3D face model to normalize illumination of color face. Unfortunately, due to the requirement for 3D point clouds and a large amount of computation, this method has limited practical application. With the developments of hardware and neural networks, illumination normalization is gradually evolved from traditional ways to deep learning-based techniques. So far, there are a few of the methods to process illumination with deep learning, for which there remain two challenges: (1) The key challenge is identity preservation. (2) It is difficult to deal with the illumination of color face. Ma et al. (2018) first used generative adversarial nets to process illumination of facial images. Han et al. (2019) put forward asymmetric joint generative adversarial networks (GAN) to normalize facial illumination with lighting labels. Therefore, in view of the shortcomings of the methods, we propose a new method to normalize illumination of color facial image without any identity label and illumination label as input. Moreover, our method which contains a generator, a discriminator and one feature extractor, is simpler than the previous methods based on deep learning.

Inspired by the success of GAN on image denoising, image synthesis and transfer learning, we reformulate the illumination normalization problem like the way of dealing with the tasks above. Our goal is to learn a GAN mapping from any poor-lighted images to well-lighted images, and the latter is called standard illumination cases in this study. In summary, our main contributions are listed as follows:

1. A new scheme is proposed for the illumination normalization of color face images. Unlike previous deep learning methods, we reduce the number of discriminators and not utilize reconstruction computation, which improve the computational speed.
2. We use content loss and elaborately designed generator for preserving identity. Experimental results demonstrate that the combination of content and L_1 loss makes our method achieve good performance. The

proposed method can not only process the illumination of frontal face but also the non-frontal face.

3. Though our model is trained on the faces under well-controlled lighting variations, it generalizes to face images with less-controlled lighting variations well, meanwhile preserving its identity effectively.

The reminder of the paper is organized as follows. Section 2 describes related work on illumination normalization. Section 3 describes the proposed Illumination Normalization GAN (IN-GAN) in details. Experimental results, evaluation and comparisons are included in Sect. 4. Section 5 describes validation for identity preserving. Finally, conclusions are drawn in Sect. 6.

2 Related work

2.1 Illumination normalization

(1) Traditional methods

To deal with the illumination variation problem, numerous works have been put over the past decades. In 1987, Pizer et al. (1987) proposed adaptive histogram equalization to enhance image contrast. Afterward, many researchers extended the histogram equalization algorithm. For instance, Shan et al. (2003) propose region-based histogram equalization to deal with illumination. Xie et al. (2005) proposed block-based histogram equalization for illumination processing. Orientated local histogram equalization that compensates illumination while encoding rich information on the edge orientations is presented by Lee et al. (2012). In 1999, Shashua and Riklin-Raviv (1999) proposed the quotient image method that provided an invariant approach to deal with the illumination variation. Afterward, many researchers extended the quotient image algorithm. Shan et al. (2003) developed gamma intensity correction for normalizing the overall image intensity at the given illumination level by introducing an intensity

mapping and quotient image relighting. Wang et al. (2004) put self-quotient image. Chen et al. (2005) came up with TV-based quotient image model for illumination normalization. Srisuk et al. (2008) proposed Gabor quotient image to extend from the self-quotient image by which the 2D Gabor filter is applied instead of weighted Gaussian filter. An et al. (2010) proposed a decomposed image under L_1 and L_2 norm constraint, then obtained illumination invariant large-scale part by region-based histogram equalization and got illumination invariant small-scale part by self-quotient image.

Adini et al. (1997) proposed logarithmic transformation, directional gray-scale derivation, and Laplacian of Gaussian for illumination normalization. Single-scale Retinex was put by Jobson et al. (1997) for processing illumination. Fitzgibbon et al. (2002) proposed Gaussian high-pass to process illumination. Local normalization technology proposed by Xie et al. (2006). It could effectively eliminate the adverse effect of uneven illumination while keeping the local statistical properties of the processed image the same as in the corresponding image under normal illumination condition. Chen et al. (2005) came up with a lighting normalization method based on the generic intrinsic illumination subspace, which was used as a bootstrap subspace for novel images. Du et al. (2005) presented wavelet-based illumination normalization. Chen et al. (2006a, b) proposed logarithmic total variation for processing illumination. Chen et al. (2006a, b) put a new method named logarithmic discrete cosine transformation for illumination compensation and normalization. Tan and Triggs (2010) processed illumination by combination of gamma correction, difference of Gaussian filtering, masking, and contrast equalization, which was called TT in literature (2013). Fan et al. (2011) proposed a method named homomorphic filtering-based illumination normalization. The filter's key component was a difference of Gaussian.

Wang et al. (2011) came up with illumination normalization based on Weber's Law. Zhao et al. (2012) proposed a self-lighting ratio to suppress illumination differences in the frequency domain. A linear representation-based face illumination normalization method was put forward by Li et al. (2012). BimaSenaBayu et al. (2013) proposed an adaptive contrast ratio based on fuzzy by considering two models of individual face as input, appearance estimation model and shadow coefficient model. Goel et al. (2013) put forward an approach for illumination normalization based on discrete wavelet transformation and discrete cosine transformation. Discrete wavelet transformation was performed on the image and discrete cosine transformation was employed on low frequency sub band. Then low frequency discrete cosine transformation coefficients were modified to suppress the illumination variations. Vishwakarma (2015) proposed a fuzzy filter applied over the

low-frequency discrete cosine transformation coefficients method for illumination normalization. With the development of 3D technologies, physical lighting models became a mainstream. Zhao et al. (2014) decomposed lighting effect by ambient, diffuse, and specular lighting maps and estimated the albedo for face images with drastic lighting conditions. Tu et al. (2017) presented a new and efficient method for illumination normalization with an energy minimization framework. Ahmad et al. (2017) used independent component analysis and filtering to process illumination. Zhang et al. (2018) presented a novel patch-based dictionary learning framework for face illumination normalization. Zheng et al. (2019) proposed a local texture enhanced illumination normalization method based on fusion of difference of Gaussian filters and difference of bilateral filters. Zhang et al. (2019) first combined Phong's model and Lambertian model, then generated the chromaticity intrinsic image (CII) in a log chromaticity space that was robust to illumination variations. The largest matching area was helpful to perform lighting normalization, occlusion de-emphasis and finally face recognition (McLaughlin et al. 2017).

(2) Deep learning-based methods

The essence of deep learning is to solve a function to realize the mapping from input to output. Ma et al. (2018) used GAN to process illumination of color faces. Though their method can generate vivid and well-lighted facial images based on illumination label. However, reconstruction and discriminators were used, it took more time to complete its computation. Han et al. (2020) put forward asymmetric joint GAN to normalize face illumination. Their method contained two GANs, one is to normalize illumination, the other is to maintain personalized facial structures. Moreover, their method needs lighting labels.

2.2 GAN and their applications

The GAN (Goodfellow 2014) brings extraordinary vitality to the image generation, even expand to Fourier series (González-Prieto et al. 2021). With the help of the combination of GAN and convolutional neural network (CNN), deep convolutional generative adversarial (Radford et al. 2016) makes a great leap in the ability of image generation. By specifying the input conditions, conditional GAN (Mirza et al. 2014) can generate the specific target photos. At the same time, GAN leads to a series of breakthroughs in image inpainting (Demir et al. 2018), super-resolution (Wang et al. 2018), style transfer (Yang et al. 2019), image translation (Zhu et al. 2017; Lin et al. 2018) and so forth. In the field of face recognition, the training data is expanded by using GAN to generate some face photos (Shrivastava et al. 2017), different expressions (Pumarola et al. 2018),

different age faces (Zhu et al. 2017), etc. For the application of generating frontal photos for FR, TP GAN (Huang et al. 2017), DR-GAN (Tran et al. 2017) and others synthesize large pose face images into frontal photos, and obtain good FR results.

3 Proposed method

3.1 Overall framework

In this section, we detail the proposed IN-GAN for illumination normalization. Figure 2 shows the block diagram of IN-GAN, which takes a set of poor-lighted face images and corresponding standard illumination images as input and outputs a set of well-lighted face images in an end-to-end way. As is illustrated in Fig. 2, the core of our approach consists of a generator, a discriminator, and a feature extractor. Our generator is composed of an encoder network and a decoder network. The discriminator is employed to judge its input whether real images or fake images generated by the generator. The feature extractor is to extract face features for every face image. In the testing phase, we just use a generator to transform poor-lighted face images into well-lighted images. We utilize 3 loss items: adversarial loss, content loss and L_1 loss.

3.2 Generator and discriminator architecture

The generator of our IN-GAN is inspired by the components of Pix2Pix (Isola et al. 2017) and U-net (Ronneberger et al. 2015). Our generator consists of 11 convolutional

layers and 11 deconvolutional layers, each of which is equipped with a LeakyReLU as activation. Details of the generator are shown in Fig. 3. As is shown in Fig. 3, the input size of the generator is designed to be a 128×128 color image. The output resolution of our generator is 128×128 pixels in size. The dotted lines in Fig. 3 are skip connections that are conducive to feature retention. In the middle 6 convolutional layers, we utilize dropout to avoid overfitting and special deconvolutional and convolutional layers at the end of the generator for enhancing the synthetic ability of our model. For this special design, which further enhances the ability of feature retention. Because InstanceNorm (Ulyanov et al. 2016) has the characteristics of preventing instance-specific mean and covariance shift simplifying the learning process, we utilize InstanceNorm after each convolutional layer. Moreover, experiments demonstrate that InstanceNorm makes our model converge fast, which means that our model can obtain higher recognition rate, good visual effect, and fine illumination normalization results after about 14 epochs. InstanceNorm can be computed by:

$$y_{ijk} = \frac{x_{ijk} - \mu_{ii}}{\sqrt{\sigma_{ii}^2 + \epsilon}} \tag{1}$$

$$\mu_{ii} = \frac{1}{HW} \sum_{l=1}^W \sum_{m=1}^H x_{tilm} \tag{2}$$

$$\sigma_{ii}^2 = \frac{1}{HW} \sum_{l=1}^W \sum_{m=1}^H (x_{tilm} - \mu_{ii})^2 \tag{3}$$

where $x \in R^{T \times C \times W \times H}$ is an input tensor with a batch of T images. Let x_{ijk} mean its t_{ijk} -th element, where k and j are

Fig. 2 Our overall generative adversarial network framework

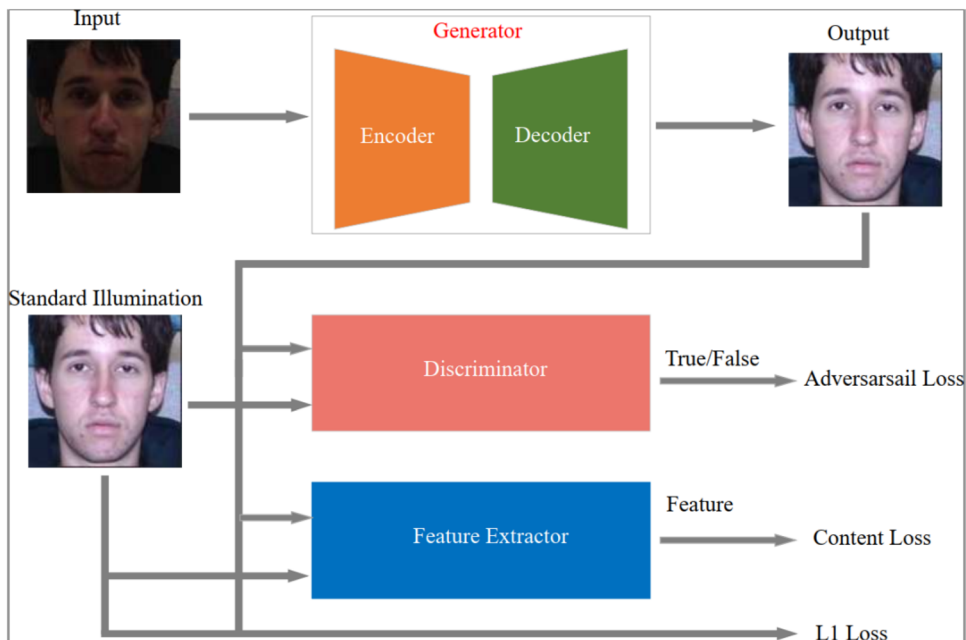
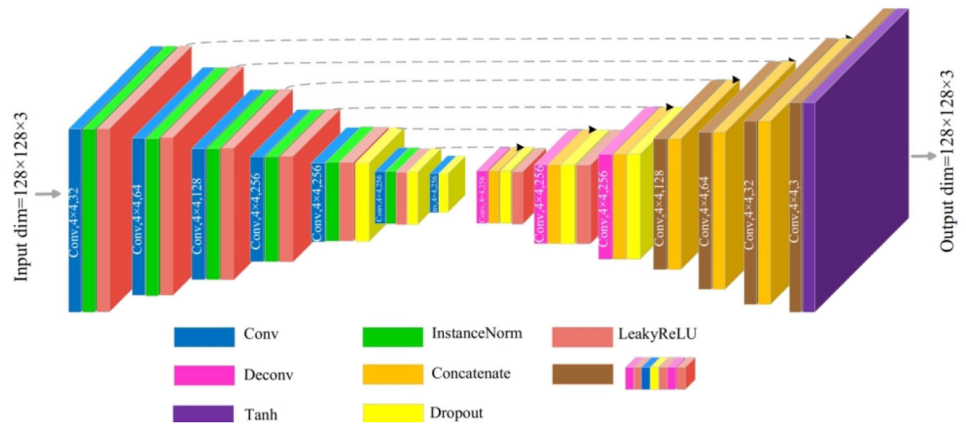


Fig. 3 The detailed structure of our generator



span spatial dimensions, i denotes the feature channel (color channel if the input is a RGB image), t is the index of the image in the batch. The discriminator of our IN-GAN is inspired by the compoents of Pix2Pix (Huang et al. 2017) and consists of 5 convolutional layers, each of which is equipped with a LeakyReLU as activation. Details of the discriminator are shown in Fig. 4. As is illustrated in Fig. 4, the input size of the discriminator is designed to be a 128×128 color image. We use InstanceNorm after each convolutional layer.

3.3 Objective

Our method uses one discriminator D and a generator G , which constitutes a set of adversarial training processes respectively and optimizes their min–max problems. Our full objective is:

$$L(G, D) = L_{adversarial}(G, D) + \lambda_1 \times L_{content}(G) + \lambda_2 \times L_{l1}(G) \tag{4}$$

where λ_1, λ_2 are weight parameters respectively, $L_{adversarial}$, $L_{content}$ and L_{l1} are as follows:

$$L_{adversarial}(G, D) = E_x[\log D(x)] + E_{G(x)}[\log(1 - D(G(x)))] \tag{5}$$

$$L_{content}(G) = \|F(y) - F(G(x))\|_1 \tag{6}$$

$$L_{l1}(G) = \|y - G(x)\|_1 \tag{7}$$

where x denotes input image, whereas y is the target image (standard illumination), F means feature extractor such as VGG-19 (Simonyan et al. 2014), ResNet-50 (He et al. 2016) for extracting feature. In this study, ResNet-50 is trained on VGGFace2 (Cao et al. 2018).

4 Experiments

4.1 Datasets

As far as the dataset, we choose Multi-PIE (Gross et al. 2010) that is with 15 poses ranging from -90° to $+90^\circ$. Each pose includes 20 illuminations and up to 6 expressions. All the dataset is 337 identities. We select $0^\circ, -15^\circ, -30^\circ, -45^\circ$ faces, 20 illuminations and natural expression, without glasses from session 1 of Multi-PIE as our dataset. We detect and crop faces from the dataset with single shot scale-invariant face detector (S3FD) (Zhang et al. 2017) and resize to 128×128 as our training and test sets. The 07 illumination faces are chosen as standard face (standard illumination) and the rest are selected as poor lighted facial images. The total identity number is 129. 30 identities are chosen as our test dataset and the rest 99 identities as our training dataset. As to the test dataset, we organize 4 settings. The setting 1 contains only frontal facial image, which is 30 identities and includes 19 illuminations. Regarding settings 2, it is the same identities as setting 1 but with large head poses such as $[-15^\circ, -30^\circ, -45^\circ]$ under 19 illuminations. With respect to setting 3, we choose -15° facial images and convert RGB to gray. Regarding setting 4, we obtain some poor-lighted faces using search engines from the internet

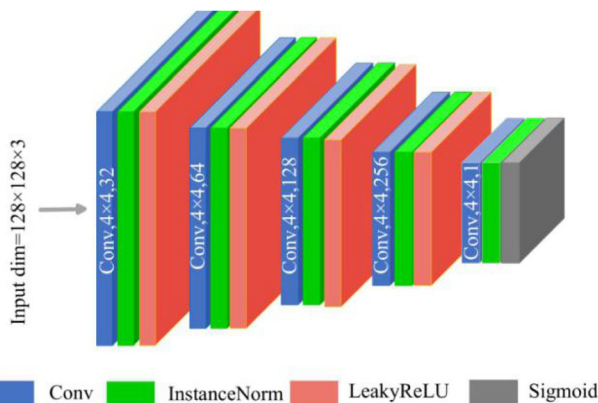


Fig. 4 The detailed structure of our discriminator

and face recognition grand challenge (FRGC) database (Phillips et al. 2005). For the purpose of verifying the performance of our algorithm, we also add people who wear glasses and faces under large head poses to our test dataset. All the faces of setting 4 are under less controlled lighting variations.

4.2 Implementation details

As to the encoder, LeakyReLU with a slope of 0.2 is used for activation firstly. Regarding the decoder, LeakyReLU is a slope of 0.2. For gradient descent, we use Adam (Kingma et al. 2015) optimizer, and choose a learning rate of 0.0002 with momentum parameters $\beta_1 = 0.5$, $\beta_2 = 0.999$, weight decay = 0.0001. One 1080Ti graphics cards with batch size of 16 is used for training. The time of 20 epochs is about 25 min. Besides, during the training period, without any data enhancement methods. By setting different values for λ_1 and λ_2 , we get 3 combinations of loss items. Set $\lambda_1 = 1.0$, $\lambda_2 = 0$ to train a model with content loss item only. We choose $\lambda_1 = 0.0$, $\lambda_2 = 1.0$ to train a model with L_1 loss item only. When $\lambda_1 = 1.0$ and $\lambda_2 = 0.1$, we train a model with content and L_1 loss items.

4.3 Metrics

At present, most literatures evaluate the performance of illumination normalization algorithms from two aspects: one is recognition rate, the other is illustrated by some face images before and after processing by various methods. Cosine similarity of feature vectors for face recognition is selected. For the purpose of more comprehensively estimating the performance of various illumination normalization algorithms, except for comparing recognition rate and illustrating examples, peak signal to noise ratio (PSNR) is used to evaluate the performance of various illumination normalization methods. Cosine similarity and PSNR are briefly introduced as follows:

- (1) Cosine similarity is defined as:

$$\cos(\theta) = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \quad (8)$$

where A and B are the feature vectors obtained from ResNet-50 (Ulyanov et al. 2016) trained on VGGFace2 (Simonyan et al. 2014), A_i denotes i -th element of vector A , B_i denotes i -th element of vector B .

- (2) $PSNR$ is defined as:

$$PSNR = 10 \times \log_{10} \left[\frac{(2^n - 1)^2}{MSE} \right] \quad (9)$$

where n is the maximum bits of each pixel, MSE means mean square error of two images, which can be computed by:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2 \quad (10)$$

where m and n denote the width and height of image, I and K are two gray images that are equal in width and height.

4.4 Quantitative evaluation

In this section, the performance of other existed methods and our algorithm are evaluated from two aspects: one is face recognition rate, the other is $PSNR$. In terms of recognition rate, we evaluate the performance of directional gray-scale derivation X (DGDx) (Adini et al. 1997), directional gray-scale derivation Y (DGDy) (Adini et al. 1997), Gaussian high-pass (GHP) (Fitzgibbon et al. 2002), histogram equalization (HE) (Pizer et al. 1987), logarithmic discrete cosine transform (LDCT) (Chen et al. 2006a, b), local normalization technology (LN) (Xie et al. 2006), Laplacian of Gaussian (LoG) (Adini et al. 1997), logarithmic transform (LT) (Adini et al. 1997), logarithmic total variation (LTV) (Du et al. 2005), self-quotient image (SQI) (Wang et al. 2004), single-scale retinex (SSR) (Jobson et al. 1997), TT (Tan et al. 2010) and ours. In case of $PSNR$, we only evaluate GHP, GIC, HE, LN, LT, SQI, SSR, TT and ours. All the methods have higher recognition rate and better visual effect than the rest when they are used to process frontal faces.

As is shown in Table 1, the recognition rate of GIC, HE, LN, SQI, SSR and ours is 100% when the probe images are frontal faces under various illuminations processed by a variety of methods. GHP, LT, LTV and TT also obtain high recognition rate. All of them are greater than 92%. In general, all the methods can preserve details or identity information in some sense. But GIC, HE, LN, SQI, SSR, and ours obtain the maximal recognition rate. In qualitative comparisons section, we illustrate contrast face images of the aforesaid algorithms.

In Table 2, the $PSNR$ of GHP, GIC, HE, LN, LT, SQI, SSR, TT and ours under 19 illuminations are illustrated. As is shown in Table 2, it can be concluded that the $PSNR$ of original images is different, but after illumination normalization, GHP, GIC, HE, LN, LT, SQI, SSR, TT and ours narrow the difference of $PSNR$ and enhance the quality of poor-lighted images. It is obvious that our method is the best of all the illumination normalization

Table 1 Recognition rate on setting 1

Method	DGDx	DGDy	GHP	GIC	HE	LDCT	LN	LoG	LT	LTV	SQI	SSR	TT	Ours
Recognition rate	30.88%	21.75%	92.81%	100%	100%	70.00%	100%	2.98%	99.47%	96.32%	100%	100%	94.91%	100%

Bold values indicate the maximum value

Table 2 PSNR of original and processed images by various illumination normalization methods

Methods	Illumination	Original	GHP	GIG	HE	LN	LT	SQI	SSR	TT	Ours
00		7.17	10.12	14.85	13.27	10.12	15.90	13.63	14.48	13.29	20.57
01		9.17	11.28	13.12	11.73	9.30	14.12	13.72	13.33	13.11	20.49
02		10.10	11.70	13.51	12.13	10.35	14.69	14.11	13.62	13.47	21.11
03		10.36	11.65	13.63	12.39	11.10	14.78	14.14	13.67	13.62	20.82
04		11.59	11.63	14.03	12.86	12.19	14.82	14.34	13.90	13.78	20.92
05		14.27	11.81	14.37	13.45	13.76	14.44	14.57	13.61	13.94	20.54
06		15.84	11.51	14.72	14.42	14.97	14.80	14.57	13.53	14.02	20.43
08		16.62	12.04	14.67	14.16	15.75	13.79	15.25	13.19	14.11	20.77
09		13.12	12.28	13.68	13.27	13.98	14.14	15.22	13.62	14.05	20.83
10		11.27	12.04	13.55	12.66	12.82	14.23	14.85	13.82	13.79	21.06
11		10.72	12.30	13.06	12.15	11.78	14.05	14.70	13.57	13.68	20.78
12		9.74	11.83	13.20	11.89	10.71	14.28	14.08	13.58	13.38	20.90
13		9.59	11.58	12.59	11.46	9.68	13.82	13.89	13.32	13.08	20.55
14		10.58	11.84	14.04	12.99	12.20	15.12	14.39	14.21	13.82	21.06
15		12.79	12.15	14.56	13.55	13.47	15.18	14.78	14.17	13.99	21.22
16		13.86	11.37	15.51	14.33	13.82	16.42	14.58	14.22	13.78	21.05
17		12.19	12.37	14.26	13.30	13.49	15.04	15.20	14.03	13.98	21.29
18		10.54	11.98	14.17	12.89	12.59	15.01	14.93	14.12	13.83	21.11
19		7.28	10.13	14.89	13.20	10.15	15.72	13.65	14.45	13.29	20.59
Average		11.41	11.67	14.02	12.95	12.22	14.75	14.45	13.81	13.68	20.85

Bold value indicates the maximum value

algorithms illustrated in Table 2 and our approach archive the best result of illumination normalization.

4.5 Qualitative comparisons

Because most of the previous methods focused on the illumination normalization of gray faces. For the sake of fairness, the comparative experiment on gray faces are conducted. According to Fig. 5, the faces from 2 to 7 columns are -15° , from 8 to 13 are -30° , from 14 to 19 are -45° . The first row is the original poor-lighted facial images. The second and third rows are the results of GHP and GIC separately. The fourth row is obtained from HE method. Next, the experimental results of LN, LT, SQI, SSR, TT, DGDx, DGDy, LDCT, LoG, and LTV are shown. The final row is the output results of our method.

As can be observed from the Fig. 5, in general, these methods such as SQI, GIC, HE, LN, SSR, LT and ours achieve good performance, whereas the other methods

preserve only part of the details and have poor visual effect. Although SQI, GIC, HE, LN, SSR, LT achieve good results, there are some defects. For instance, though the SQI method has a high recognition rate, its visual effect is a little worse. From the fourth row of Fig. 5, HE method is with a lot of noise and can't handle shadows effectively. Although LN method improves the illumination of the original facial images, it cannot deal with the shadows. From the third, ninth, eleventh and fifteenth image of the third row of Fig. 5, GIC cannot handle light and cast shadow effectively.

There is a large difference in light intensity on the whole face region. Though SSR method is with good visual effect, some images are over-saturated and cannot process cast shadow effectively. Although the LT method can deal with the illumination of each facial image effectively and with fine visual effect, it cannot process cast shadow effectively and some images are over-saturated. It is obvious in the final row in Fig. 5. Our method can not only

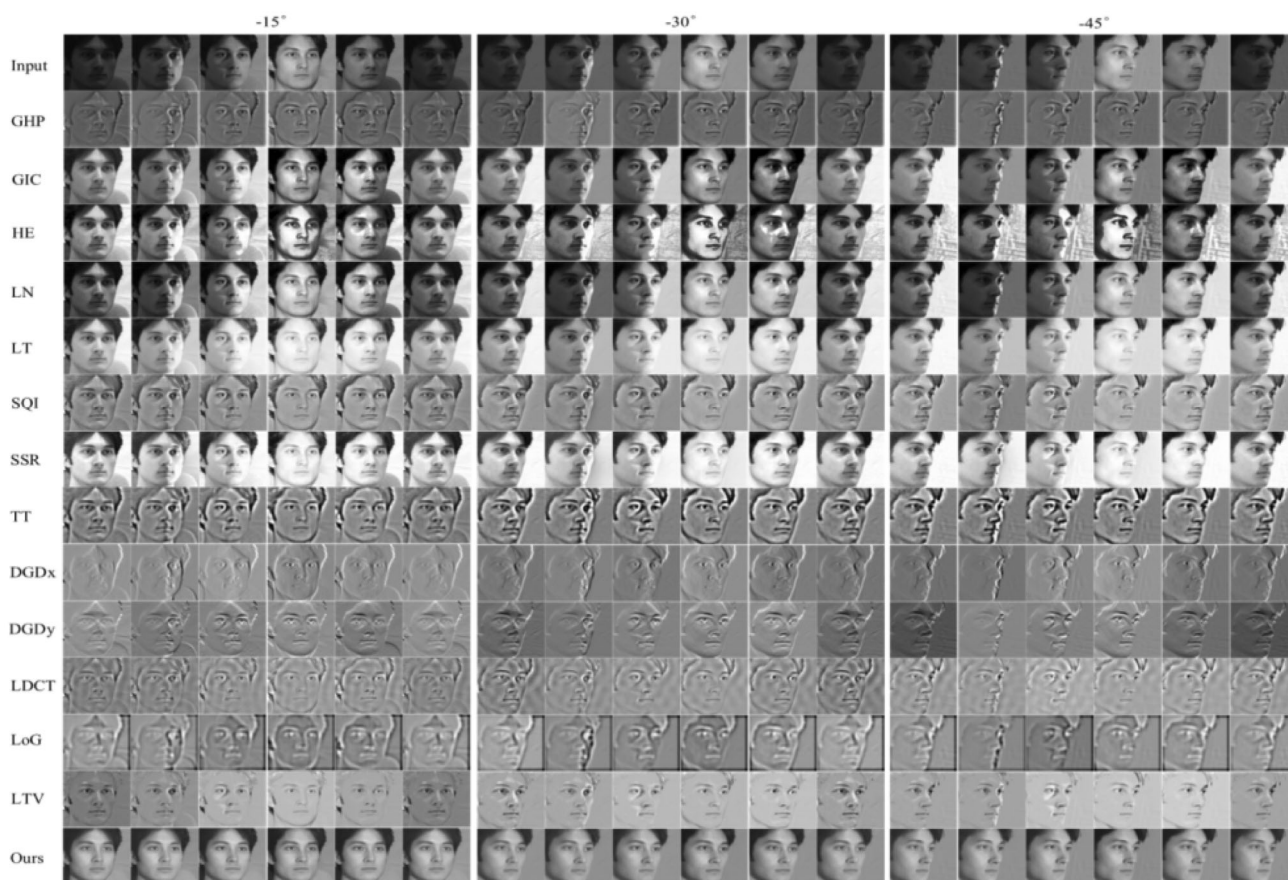


Fig. 5 Some gray faces under various illuminations and large head poses before and after processing by various methods. (Color figure online)

deal with the illumination of each image but also keep the corresponding identity of the images under various illuminations effectively. In summary, GIC, SSR, LT and ours perform better than other illumination normalization algorithms. From three aspects: recognition rate, illumination normalization effect and visual effect, our method is the best of all the approaches.

From the previous parts, the methods such as GHP, GIC, HE, LN, LT, SQI, SSR, TT and ours have better performance than others. To verify their performance under less-controlled lighting variations, another set of experiments illustrate the comparison results are shown in Fig. 6. It can be concluded that GIC, HE, LN, LT, SSR and ours have better visual effect than others. From the third image of the third row, GIC cannot process the light. From the third image of the fourth row, HE encounters same problem and there is some noise. In line 5 of Fig. 6, it is obvious that all the face images still have severe shadows after processing by LN. LT and SSR are the best if only consider from the aspect of visual effect. But some images are over-saturated. Our algorithm is the best of all the approaches from two aspects: illumination normalization and visual effect.

4.6 Ablation studies

Since our method is without illumination label and identity label, we compare it only to algorithms that do not require any label. Because there is no label free algorithm in deep learning, we only illustrate some results of our method. According to Fig. 7, the first row is the original facial images under various illuminations. The second row is the synthetic facial images using our approach training with the content loss item only. The third row is the synthetic faces using our method trained with the L_1 loss item only. The final row of Fig. 7 is the result obtained from our algorithm trained with the content and L_1 loss items. According to Fig. 7, our method can not only make the poor illumination become well-lighted and unanimous but also preserves the identity information of the original poor-lighted facial images effectively. Since content and L_1 loss items can preserve identity well, the model trained with the combination of content and L_1 loss items can not only keep identity effectively but also obtain better visual effect.

In order to further verify the performance of our algorithm, we perform illumination normalization experiment on the non-frontal facial images under various lighting

Fig. 6 Some gray faces under less-controlled illuminations by various methods. (Color figure online)



Fig. 7 Some frontal color faces under various loss items by our method. (Color figure online)



Fig. 8 Some color faces under large head poses and various illuminations before and after processing by our method. (Color figure online)

conditions. As is shown in Fig. 8, we can see that the first and third rows are -15° , -30° and -45° faces under various illuminations, the second row is the output of the

first row. The fourth row is the corresponding synthetic results of the third row. It is obvious that our method can not only process illumination of color faces under large-

poses and poor-lighted conditions but also keep corresponding identities effectively.

In practical application, it is basically less-controlled illumination. Therefore, it is necessary to verify the performance of our algorithm under the less-controlled lighting variations. According to Fig. 9, the first row is the original input with poorly lighted faces and the second row is the corresponding output. It is apparent that the input images are not aligned and under various less-controlled lighting variations. The results of Fig. 9 indicate that our method can not only process the illumination of color face under less-controlled lighting variations with various head poses but also keep its identity effectively. Though no one wore glasses in our training set, our method still has the synthetic ability to generate glasses, which demonstrates our approach has a strong ability of feature retention.

As is illustrated in the Fig. 10, the first row is the original images, and the next 3 rows are the corresponding output of epoch 1, 5 and 13. According to Fig. 10, we can conclude that our method trained with content loss, L_1 loss, the combination of content and L_1 loss has good feature retention ability. When we combine content and L_1 loss, our method converges faster than L_1 loss and content loss used separately. Although content loss also makes our algorithm converge fast, its visual effect is not as good as the combination of content and L_1 loss. It is apparent in the Fig. 10 that our method has the advantages such as rapid convergence, good feature retention, favorable illumination normalization results.

To evaluate our method quantitatively in Figs. 11 and 12. They are under 3 loss items from epoch 1 to 14. Figure 11 illustrates the recognition rate of frontal faces under 3 loss items from epoch 1 to 14. From Fig. 11, it is known that the combination of content and L_1 loss makes our model converge fast and obtain higher recognition rate. Figure 12 shows the recognition rate of -15° faces under 3 loss items from epoch 1 to 14, which demonstrates that our method can get better performance after combining content loss and L_1 loss than using content loss or L_1 loss separately. From Figs. 11 and 12, It can be concluded that our method converges fast and obtains high recognition rate after combining content loss and L_1 loss.

5 Validation for identity preserving

In this section, we discuss our algorithm's ability to keep identity information by conducting face recognition experiment. Table 3 shows the recognition results of our algorithm. In line 2 of Table 3, it is the recognition rates of color faces under 0° , -15° , -30° , -45° . It is obvious that our method under content and L_1 loss items can obtain high recognition rate. As is illustrated in Table 3, there is no any input such as identity label, illumination label, our method still preserves the corresponding identities of the original faces effectively.

6 Conclusion

In this study, we put a novel and practical deep fully convolutional neural network architecture for illumination normalization of color face termed IN-GAN. Our method can process not only the illumination of color face images but also the illumination of gray face images. Furthermore, all the existed methods mainly focus on processing the illumination of frontal or near frontal face. Our scheme can not only process the illumination of frontal face but also the non-frontal face. Moreover, our method can normalize illumination of face image, and retain identity information effectively. Finally, though our trained model on faces under well-controlled lighting variations, it can process face under less-controlled lighting variations and preserve identity information effectively.

In our further researches, other features, and geometric structure (Pareja-Corcho et al. 2020) need to be considered. We find that the number of layers and types of connection are important in CNN (Guo et al. 2021a). Even the discriminate model is considered in the generative adversarial network (Guo et al. 2021b) and multiple features (Guo et al. 2022). The authors showed that the six-convolution layer and three fully-connected layer CNN, nine-layers in total, achieved better performance in sensitivity, specificity, accuracy, and precision (Wang et al. 2020a). In Wang et al. (2020b), the performance of parametric rectified linear unit is better than ordinary ReLU. They also verify that batch normalization overcomes the internal

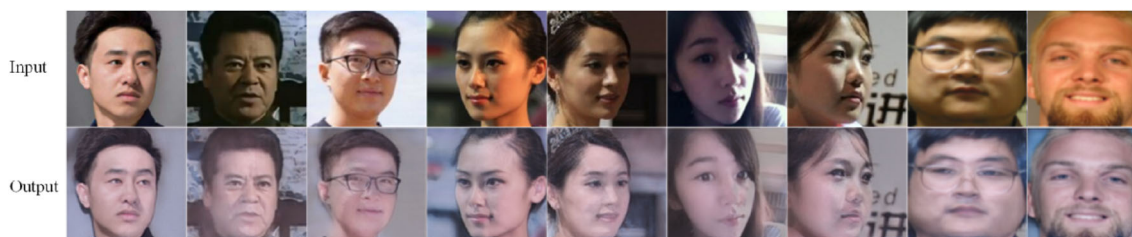


Fig. 9 Verify our algorithm to process the illumination of color faces

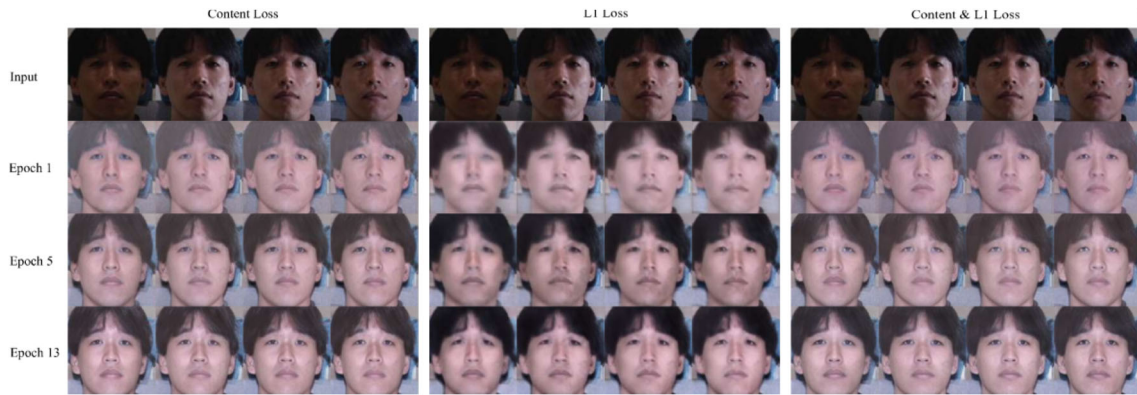


Fig. 10 Some color face after illumination normalization by our method under various epochs and loss items. (Color figure online)

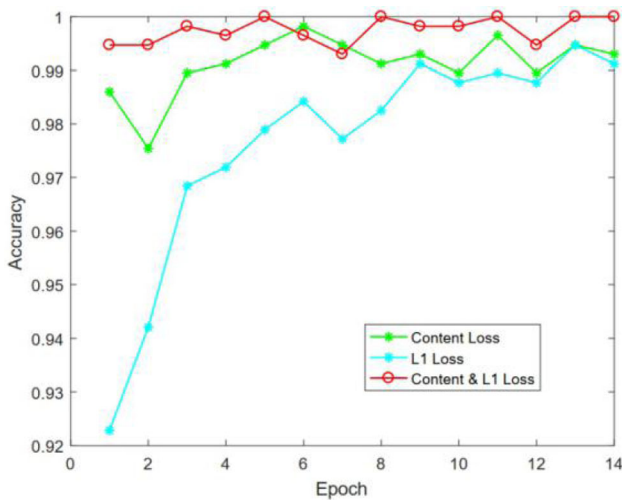


Fig. 11 Recognition rate of frontal faces

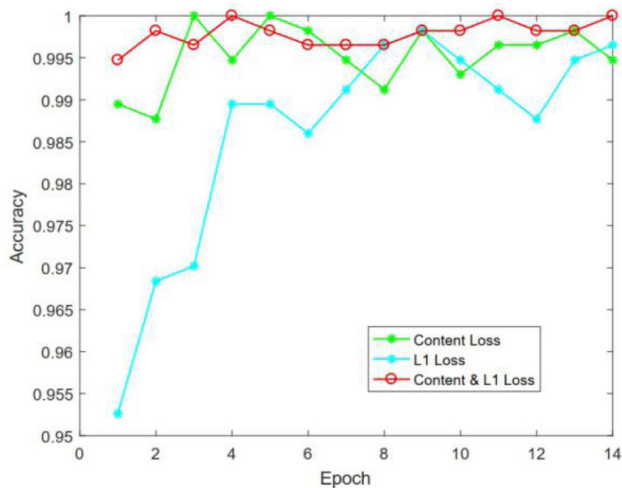


Fig. 12 Recognition rate of -15°

covariate shift and dropout got over the overfitting. We shall try different layer CNN model, and various types of connection in the future. There is no gradient in ReLU,

Table 3 Recognition rate of faces under various head poses and lighting variations after doing illumination normalization by our method

Angle	0°	-15°	-30°	-45°
Recognition rate	100%	100%	100%	97.54%

when values are less than zero. However, there is a small gradient in LeakyReLU, while values are less than zero. Therefore, LeakyRelu is selected in our experiments, which can get good results.

Although our illumination normalization algorithm achieves preferable results from qualitative and quantitative comparisons. Compared with other algorithms, our algorithm has gained advantages. There is a lot of future work here worth continuing to study:

- To improve our network structure for preserving more texture details.
- To train a feature extractor and classifier for the facial images after normalizing illumination by our method.
- To process illumination normalization of other image types, such as landscape and medical images.
- To the preprocessing stage of other visual analysis tasks, such as facial landmark detection and face alignment.

Acknowledgements The authors would like to thank the editor and anonymous reviewers for their insight and suggestions. We also thank Dr. Ferrante Neri (University of Surrey), Prof. DongC. Liu and Dr. Paul Liu (Stork Healthcare) for value suggestions. This work was supported in part by the National Natural Science Foundation of China under Grant 61806028, Grant 61672437, Grant 62103064 and Grant 61702428, Sichuan Science and Technology Program under Grant 21ZDYF2484, Grant 2021YFN0104, Grant 21GJHZ0061, Grant 21ZDYF3629, Grant 21ZDYF0418, Grant 21YYJC1827, Grant 2021YJ0086, Grant 2021YFG0295, Grant 21ZDYF3537, Grant 21ZDYF3598, Grant 2020YFG0177, Grant 2022YFN0020, Chinese Scholarship Council under Grant 202008510036, Opening Project of International Joint Research Center for Robotics and Intelligence

System of Sichuan Province under Grant JQZN2021-003, Department of Science and Technology of Sichuan Province under Grant 2019YFSY0043, AECC Sichuan Gas Turbine Establishment, Key Laboratory on Aero-engine Altitude Simulation Technology, and Intelligent Control Education Reform Project of Chengdu University of Information Technology under Grant JYJG2021044. Program of Chengdu Technological University under Grant 2019ZR005, Program Name: Intelligent Sensing of Complex Power Environments Key Technologies and Applications.

Declarations

Conflict of interest We promise that this manuscript is the authors' original work and has not been published nor has it been submitted simultaneously elsewhere. All authors have checked the manuscript and have agreed to the submission.

References

- Adini Y, Moses Y, Ullman S (1997) Face recognition: the problem of compensating for changes in illumination direction. *IEEE Trans Pattern Anal Mach Intell* 19:721–732
- Ahmad F, Khan A, Islam IU, Uzair M, Ullah H (2017) Illumination normalization using independent component analysis and filtering. *Imaging Sci J* 65(5):308–315
- Al-Osaimi FR, Bennamoun M, Mian AS (2006) Illumination normalization for color face images. In: *International Symposium on Visual Computing, Advances in Visual Computing*, pp 90–101.
- An G, Wu J, Ruan Q (2010) An illumination normalization model for face recognition under varied lighting conditions. *Pattern Recogn Lett* 31:1056–1067
- BimaSenaBayu D, Miura J (2013) Fuzzy-based illumination normalization for face recognition. In: *2013 IEEE Workshop on Advanced Robotics and its Social Impacts*, pp 131–136
- Cao Q, Shen L, Xie W, Parkhi OM, Zisserman A (2018) Vggface2: a dataset for recognizing faces across pose and age. In: *13th IEEE International Conference on Automatic Face and Gesture Recognition*, pp 67–74
- Chen T, Yin W, Zhou XS, Comaniciu D, Huang TS (2005) Illumination normalization for face recognition and uneven background correction using total variation-based image models. *IEEE Comput Soc Conf Comput vis Pattern Recogni* 2:532–539
- Chen T, Yin W, Zhou XS, Comaniciu D, Huang TS (2006a) Total variation models for variable lighting face recognition. *IEEE Trans Pattern Anal Mach Intell* 28(9):1519–1524
- Chen W, Er MJ, Wu S (2006b) Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. *IEEE Trans Syst Man Cybern B Cybern* 36(2):458–466
- Demir U, Lnal GB (2018) Patch-based image in painting with generative adversarial networks. Accessed from <https://arxiv.org/abs/1803.07422>
- Du S, Ward RK (2005) Wavelet-based illumination normalization for face recognition. *IEEE Int Conf Image Process* 2:II-954
- Fan C-N, Zhang F-Y (2011) Homomorphic filtering-based illumination normalization method for face recognition. *Pattern Recogn Lett* 32:1468–1479
- Fitzgibbon AW, Zisserman A (2002) On affine invariant clustering and automatic cast listing in movies. *Computer vision*. Springer, Berlin, pp 304–320
- Goel T, Nehra V, Vishwakarma VP (2013) Illumination normalization using down-scaling of low-frequency dct coefficients in dwt domain for face recognition. In: *2013 Sixth International Conference on Contemporary Computing*, pp 295–300
- González-Prieto Á, Mozo A, Talavera E, Gómez-Canaval S (2021) Dynamics of Fourier modes in torus generative adversarial networks. *Mathematics* 9:325. <https://doi.org/10.3390/math9040325>
- Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B (2014) Generative adversarial nets. *Adv Neural Inf Process Syst* 27:2672–2680
- Gross R, Matthews I, Cohn J, Kanade T, Baker S (2010) Multi-pie. *Image vis Comput* 28(5):807–813
- Guo DQ, Yang Q, Zhang Y, Jiang T, Yan H (2021a) Classification of domestic refuse in medical institutions based on transfer learning and convolutional neural network. *Comput Model Eng Sci* 127(2):599–620
- Guo DQ, Yang Q, Zhang Y, Zhang G, Zhu M, Yuan J (2021b) Adaptive object tracking discriminate model for multi-camera panorama surveillance in airport apron. *Comput Model Eng Sci* 129(1):191–205
- Guo DQ, Zhang GX, Neri F, Peng S, Yang Q, Liu P (2022) An adaptive kernelized correlation filters via multiple features in the tracking application. *J vis Commun Image Represent* 84:1–14
- Han H, Chen SSX, Gao W (2013) A comparative study on illumination preprocessing in face recognition. *Pattern Recogn* 46:1691–1699
- Han X, Yang H, Xing G, Liu Y (2020) Asymmetric joint gans for normalizing face illumination from a single image. *IEEE Trans Multimed* 22(6):1619–1633. <https://doi.org/10.1109/TMM.2019.2945197>
- He K, Zhang X, Ren S and Sun, J (2016) Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp 770–778
- Huang R, Zhang S, Li T, He R (2017) Beyond face rotation: global and local perception GAN for photorealistic and identity preserving frontal view synthesis. In: *2017 IEEE International Conference on Computer Vision (ICCV)*, pp 2458–2467
- Isola P, Zhu J-Y, Zhou T, Efros AA (2017) Image-to-image translation with conditional adversarial networks. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp 5967–5976
- Jobson D, Rahman Z, Woodell G (1997) Properties and performance of a center/surround retinex. *IEEE Trans Image Process* 6(3):451–462
- Kingma DP, Ba JL (2015) Adam: a method for stochastic optimization, international conference on learning representations. *Computer science*. Accessed from <https://arxiv.org/abs/1412.6980>
- Lee PH, Wu SW, Hung YP (2012) Illumination compensation using oriented local histogram equalization and its application to face recognition. *IEEE Trans Image Process* 21:4280–4289
- Li Y, Meng L, Feng J (2012) Lighting coefficients transfer based face illumination normalization. *Chinese Conference on Pattern Recognition*, pp 268–275
- Lin J, Xia Y, Qin T, Chen Z, Liu T Y (2018) Conditional image-to-image translation. In: *2018 IEEE Conference on Computer Vision and Pattern Recognition*, pp 5524–5532
- Ma W, Xie X, Yin C, Lai JH (2018) Face image illumination processing based on generative adversarial nets. *2018 24th International Conference on Pattern Recognition (ICPR)*, pp 2558–2563
- McLaughlin N, Ming J, Crookes D (2017) Largest matching areas for illumination and occlusion robust face recognition. *IEEE Trans Cybern* 47(3):796–808. <https://doi.org/10.1109/TCYB.2016.2529300>
- Mirza M, Osindero S (2014) Conditional generative adversarial nets. Accessed from <https://arxiv.org/abs/1411.1784>
- Moret-Tatay C, Baixauli-Fortea I, Grau Sevilla MD, Irigaray TQ (2020) Can you identify these celebrities? A network analysis on

- differences between word and face recognition. *Mathematics* 8(5):699. <https://doi.org/10.3390/math8050699>
- Pareja-Corcho J, Betancur-Acosta O, Posada J, Tammara A, Ruiz-Salguero O, Cadavid C (2020) Reconfigurable 3D CAD feature recognition supporting confluent n-dimensional topologies and geometric filters for prismatic and curved models. *Mathematics* 8(8):1356. <https://doi.org/10.3390/math8081356>
- Phillips P J, Flynn P J, Scruggs T, Bowyer K W, and Worek W (2005) Overview of the face recognition grand challenge. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*. Vol 1, pp 947–954
- Pizer SM, Amburn EP, Austin JD, Cromartie R, Zuiderveld K (1987) Adaptive histogram equalization and its variations. *Comput vis Graphics Image Process* 39(3):355–368
- Pumarola A, Agudo A, Martinez AM, Sanfeliu A, Moreno-Noguer F (2018) Ganimation: anatomically-aware facial animation from a single image. *Eur Conf Comput vis* 8:835–851
- Radford A, Metz L, Chintala S (2016) Unsupervised representation learning with deep convolutional generative adversarial networks. *International conference on learning representations*. Accessed from <https://arxiv.org/abs/1511.06434>
- Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. *Medical image computing and computer assisted intervention*. Springer, Cham, pp 234–241
- Shan, S, Gao, W, Cao B, Zhao D (2003) Illumination normalization for robust face recognition against varying lighting conditions. In: *2003 IEEE International SOI Conference. Proceedings (Cat. No.03CH37443)*, pp 157–164
- Shashua A, Riklin-Raviv T (1999) The quotient image: class-based rerendering and recognition with varying illuminations. *IEEE Trans Pattern Anal Mach Intell* 23:129–139
- Shrivastava A, Pfister T, Tuzel O, Susskind J, Webb R (2017) Learning from simulated and unsupervised images through adversarial training. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp 2242–2251
- Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. Accessed from <https://arxiv.org/abs/1409.1556>
- Srisuk S, Petpon A (2008) A gabor quotient image for face recognition under varying illumination. *ISVC advances in visual computing*. Springer, Berlin, pp 511–520
- Tan X, Triggs B (2010) Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Trans Image Process* 19:1635–1650
- Tran L, Yin X, Liu X (2017) Disentangled representation learning gan for pose-invariant face recognition. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp 1283–1292
- Tsai Y, Hsu L, Hsieh Y, Lin S (2020) The real-time depth estimation for an occluded person based on a single image and openpose method. *Mathematics* 8(8):1333. <https://doi.org/10.3390/math8081333>
- Tu X, Yang F, Xie M, Ma Z (2017) Illumination normalization for face recognition using energy minimization framework. *IEICE Trans* 100-D:1376–1379
- Ulyanov D, Vedaldi A, Lempitsky VS (2016) Instance normalization: The missing ingredient for fast stylization. Accessed from <https://arxiv.org/abs/1607.08022>
- Vishwakarma VP (2015) Illumination normalization using fuzzy filter in dct domain for face recognition. *Int J Mach Learn Cybern* 6:17–34
- Wang H, Li SZ, Wang Y (2004) Generalized quotient image. In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2, II
- Wang B, Li W, Yang W, Liao Q (2011) Illumination normalization based on weber's law with application to face recognition. *IEEE Signal Process Lett* 18:462–465
- Wang X, Yu K, Wu S, Gu J, Liu Y (2018) Esrgan: enhanced super-resolution generative adversarial networks. In: *Proceedings of the European Conference on Computer Vision (ECCV)*, pp 63–79
- Wang SH, Sun JD, Mehmood I, Pan CC, Chen Y, Zhang YD (2020a) Cerebral micro-bleeding identification based on a nine-layer convolutional neural network with stochastic pooling. *Concurrency Comput* 32:e5130.1-e5130.16
- Wang SH, Muhammad K, Hong J, Sangaiyah AK, Zhang YD (2020b) Alcoholism identification via convolutional neural network based on parametric ReLU, dropout, and batch normalization. *Neural Comput Appl* 32:665–680
- Xie X, Lam K-M (2005) Face recognition under varying illumination based on a 2d face shape model. *Pattern Recogn* 38:221–230
- Xie X, Lam K-M (2006) An efficient illumination normalization method for face recognition. *Pattern Recognit Lett* 27:609–617
- Yang FW, Lin HJ, Yen S-H, Wang C-H (2019) A study on the convolutional neural algorithm of image style transfer. *Int J Pattern Recognit Artif Intell* 33(5):1954020
- Zhang S, Zhu X, Lei Z, Shi H, Wang X, and Li S Z (2017) S3fd: Single shot scale-invariant face detector. *2017 IEEE International Conference on Computer Vision (ICCV)*, pp 192–201
- Zhang Y, Wang L, Guan X, Wei H (2018) Illumination normalization for face recognition via jointly optimized dictionary-learning and sparse representation. *IEEE Access* 6:66632–66640
- Zhang W, Zhao X, Morvan JM, Chen L (2019) Improving shadow suppression for illumination robust face recognition. *IEEE Trans Pattern Anal Mach Intell* 41(3):611–624
- Zhang Y, Tsang I, Luo Y, Hu C, Lu X, Yu X (2021) Recursive copy and paste GAN: face hallucination from shaded thumbnails. *IEEE Trans Pattern Anal Mach Intell*. <https://doi.org/10.1109/TPAMI.2021.3061312>
- Zhao X, Evangelopoulos G, Chu D, Shah SK, Kakadiaris IA (2014) Minimizing illumination differences for 3d to 2d face recognition using lighting maps. *IEEE Trans Cybern* 44:725–736
- Zhao X, Shah S K and Kakadiaris IA (2012) Illumination normalization using self-lighting ratios for 3d2d face recognition. In: *ECCV Workshops. European Conference on Computer Vision, Workshops and Demonstrations*, pp 220–229
- Zheng C, Wu S, Xu W, Xie S (2019) Illumination normalization via merging locally enhanced textures for robust face recognition. Accessed from <https://arxiv.org/abs/1905.03904>
- Zhu J-Y, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *2017 IEEE International Conference on Computer Vision (ICCV)*, pp 2242–2251

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.