Occluded Face Recognition by Identity-Preserving Inpainting

Chenyu Li[1,2], Shiming Ge[1⋆], Yingying Hua[1,2], Haolin Liu[1,2], Xin Jin[3]

[1]Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China
[2]School of Cyber Security, University of Chinese Academy of Sciences
[3]Department of Cyber Security, Beijing Electronic Science and Technology Institute

Abstract. Occluded face recognition, which has an attractive application in the visual analysis field, is challenged by the missing cues due to heavy occlusions. Recently, several face inpainting methods based on generative adversarial networks (GANs) fill in the occluded parts by generating images fitting the real image distributions. They can lead to a visually natural result and satisfy human perception. However, these methods fail to capture the identity attributes, thus the inpainted faces may be recognized at a low accuracy by machine. To enable the convergence of human perception and machine perception, this paper proposes an Identity Preserving Generative Adversary Networks (IP-GANs) to jointly inpaint and recognize occluded faces. The IP-GANs consists of an inpainting network for regressing missing facial parts, a global-local discriminative network for guiding the inpainted face to the real conditional distribution, a parsing network for enhancing structure consistence and an identity network for recovering missing identity cues. Especially, the novel identity network suppresses the identity diffusion by constraining the feature consistence from the early subnetwork of a well-trained face recognition network between the inpainted face and its corresponding ground-true. In this way, it regularizes the inpaintor, enforcing the generated faces to preserve identity attributes. Experimental results prove the proposed IP-GANs capable of dealing with varieties of occlusions and producing photorealistic and identity-preserving results, promoting occluded face recognition performance.

Keywords: Image inpainting, Occluded face recognition, Generative adversarial networks (GANs), Identity preserving.

1 Introduction

With the rapid development of deep learning in intelligent image recognition, face recognition has achieved impressive accuracy under un-occluded conditions. However in wild scenarios, it is common to encounter occlusions. Occlusions caused by facial accessories, objects in front of the faces, extreme illumination, self-occlusion or poor image quality exist everywhere, which results in sharp drop

⋆ Corresponding author: Shiming Ge (email: geshiming@iie.ac.cn)

in recognition accuracy. Occlusion has become one of the strongest hinderance of face recognition in the wild.

Existing occluded face recognition methods mainly depend on partial feature extraction [12], inpainting [2] or sparse representation [21]. They had good performance in handling light occluded faces but were difficult to work effectively under heavy occlusions. Recently, with the help of deep generative models, face inpainting or completion approaches [5, 7, 14] have achieved remarkable improvements. These approaches are capable of extracting high-level context features and generating photorealistic results, making them a feasible fit for occluded face recognition. In general, they mainly aim to obtain consistent looking, neglecting to capture the identity attributes. This leads to a gap of recognition accuracy compared to the ground truth, though winning the occluded one.

In this work, we devise a novel model named Identity Preserving Generative Adversarial Networks (IP-GANs) to improve occluded face recognition. The networks consist of an inpainting network for generating the missing contents, local-global discriminators and a semantic parsing network are incorporated to urge the inpainted result to be photorealistic, and an identity network is introduced to ensure the correct diffusion of identity information along the border. In this way, the inpainting network is forced to narrow the distance with ground truth in both pixel and feature subspaces. Toward this end, occluded face inpainting and recognition can be jointly solved.

Our main contributions can be summarized as three folds:

- We propose a novel model named Identity Preservation Generative Adversarial Networks (IP-GANs) for photorealistic and identity preserving face completion, under the circumstance where large region of contents are missing.
- We propose an identity network to suppress the identity diffusion, enabling the convergence of human and machine perception.
- We conduct qualitative and quantitative experiments to show the strength of IP-GANs to generate photorealistic and identity-preserving completed faces.

2 Related Work

2.1 Image Inpainting

Since the concept of digital image inpainting was introduced by Bertalmio et al. [1], many approaches have been proposed in this field. Early inpainting approaches based on information diffusion and texture synthesis exhibited fine performance when missing region is relatively small, but blurring when missing region is large. Exemplar-based methods applied texture synthesis by seeking similar patches as reference, maybe perfectly fitting background completing task. But when completing objects with unique textures like faces, things get tough. In the last few years, as deep models, GANs especially, have made remarkable progress in image generation [15], enhancement [16], cross-modal retrival [23], widely applied in fields including safety [11], medical [9] and so on, the field of image inpainting also witness an evolution. In [22], Xie et al. first applied deep networks into image denoising and reconstruction tasks. Gaining knowledge from big data, the model could handle inpainting problems more flexibly than traditional ones. However when it comes to cases where large region of contents are missing, higher semantic knowledge is needed. Pathak et al. [14] proposed context encoder to learn feature representation to capture both appearance and semantics by introducing adversary loss. Iizuka et al. [5] employed two discriminators to jointly enforce global and local consistency. In their approach, adversary training allows generators to better learn the real image distribution, therefore leading to a visually natural result. Wang et al. [20] employed perceptual loss to increase semantic similarity. Targeting at GAN based networks' huge scale and resulting high training and executing cost, some approaches aimed at fast and efficient inpainting, by introducing new structure of deep models, such as fully connected networks, dilated convolution layers [24]. Some researchers also made effort to simulate the way human learn [10].

2.2 Occluded Face Recognition

Existing occluded face recognition approaches can be divided into two types: representation-based and reconstruction-based. The "representation" idea aims to represent the occluded face directly by decreasing or excluding the influence of missing regions and tapping the useable information hidden in remaining pixels. Zhang et al. [26] introduced local Gabor binary patterns (LGBP) to form face representation, then estimated the probability of occlusion by using KullbackLeibler divergence of the LGBP features between the local region and the non-occluded local region. In [12], Oh et al. proposed selective local non-negative matrix factorization to segment the image into multiple patches and detect the unoccluded parts first, and then map them into local non-negative matrix factorization space and perform the matching. However, they are not as convenient to execute as they sound. There still lies challenges like weights setting, especially when missing area is large.

Different from the "representation" idea, the "reconstruction" idea utilizes the redundancy of image data instead by recognizing after completing. Deng et al. [2] proposed exemplar-based graph Laplace algorithm to complete occluded faces. In this way, the approach can use completed faces to boost recognition accuracy. It performs well when similar appearance and expression can be found in the library. However, the type and shape of the occlusions are unenumerable and unpredictable in real scenarios, which limits its applications. More recently, novel methods based on sparse representation and deep learning were proposed. Zhang et al. [25] proposed DeMeshNet to enforce pixel as well as feature level similarity between input MeshFaces and output inpainted faces. It can recover the missing contents with little pixel difference, greatly benefiting recognition task at the same time. While it can only deal with small occluded region, the model proposed in [7] can perform remarkable face completion when a large part of image content is missing, and to some extent increase the identity matching accuracy.

3 Identity Preserving GANs

It is a natural idea to apply state-of-the-art inpainting methods to help promote occluded face recognition performance. Intuitively, an ideal image inpainting model which could truly benefit occluded face recognition should follow two rules: photorealistic and identity preserving. Inspired by that, we propose IP-GANs to meet these two rules for occluded face recognition, which consist of five networks (as shown in Fig 1), including: an inpainting network, global and local discriminate networks, a parsing network and an identity network. The inpainting network acts as generator, while the global and local discriminate networks play for the discriminator. Moreover, the parsing network and the identity network are used to regularize the generator by enforcing the semantic and identity consistencies, respectively.

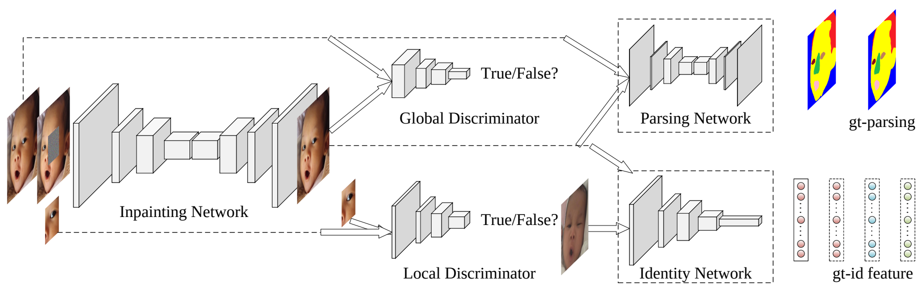

Fig. 1. The structure of IP-GANs. They consist of five networks: an inpainting network acting as the generator, two discriminate networks as the discriminator, a parsing network for facial harmony, and an identity network for recovering identity cues.

3.1 Generator for Inpainting

The inpainting network takes a face image with missing regions (generally filled with random noises) as input and generate a photorealistic and identity-preserving completed face image. Let $\mathbf{x}$ and $\mathbf{y}$ denote the face image with size $m \times n$ to be inpainted and the complete ground truth, respectively. The binary mask $\mathbf{M}$ with the same size as $\mathbf{x}$ is used to label the missing region, marking with 1 and 0 for inside and outside the missing region, respectively. Then, the generator takes $\mathbf{x}$ and $\mathbf{M}$ as input and is defined as $\mathcal{G}(\mathbf{x}, \mathbf{M})$. Following traditional GANs [3], the reconstruction loss could be defined with mean square error (MSE) between the inpainted image $\mathbf{x}$ and the ground truth $\mathbf{y}$:

$$\mathcal{L}_{rec} = \frac{1}{mn} \|\mathcal{G}(\mathbf{x}, \mathbf{M}) - \mathbf{y}\|^2 \tag{1}$$

where $\| * \|$ is ℓ_2 norm operator. In this way, MSE loss can benefit the stability during training although it tends to bring blurring.

The architecture follows an encoder-decoder framework. It first maps the input image into a hidden feature space, then reconstructs it based on the feature representation. Following [7], we use the architecture from "conv1" to "pool3" of the VGG-19 network [17], stack two more convolution layers and one more pooling layer on top of that, and add a fully-connected layer after that as the encoder. The decoder is symmetric to the encoder with unpooling layers accordingly.

3.2 Discriminators in Global and Local

Inspired by recent success of inpainting using GANs [5, 7], we adopt two discriminate networks to preserve global and local structures, respectively. Toward this end, the global discriminator takes the whole image as input, while the local discriminator uses only the inpainted region. During the training, the discriminators learn to identify whether input images are real or fake. Contextual information from local and global regions compensate each other, eventually reaching a balance between global consistency and local details. Two discriminators regularize the inpainting network via local and global adversary loss by solving a min-max optimization problem:

$$\mathcal{L}_{D_i} = \min_{\mathcal{G}} \max_{\mathcal{D}_i} \mathbb{E}[\log \mathcal{D}_i(\mathbf{y}) + \log(1 - \mathcal{D}_i(\mathcal{G}(\mathbf{x}, \mathbf{M})))], i \in \{g, l\} \tag{2}$$

where $\mathcal{D}_i$ is the global discriminator when $i = g$ and the local discriminator when $i = l$. At each iteration, the generator and discriminator are alternatively optimized and updated.

The local and global discriminate networks follow the similar architecture, except that the input is the inpainted region and entire image respectively. The architecture is similar to that in DCGAN [15]. It consists of ten convolutional layers and a single fully-connected layer.

3.3 Regularizers for Semantic and Identity

Semantic Regularization. As Fig 2(e) shows, introduction of the two discriminators enable the inpainting network to generate realistic face components. However we noted that the generated contents sometimes lack certain consistency with existing parts in size and shape, leading to unnatural expression. Same to [7], we adopt a semantic parsing network to encourage facial harmony. The parsing network functions as a multi-class classifier and assigns a label to every pixel, semantically segmenting the image into 11 parts representing eyes, mouth etc. The parsing network is pre-trained on Helen[5] dataset, which contains 2,330 images, labeled by [18]. The parsing loss is defined as the pixel-wise softmax loss:

$$\mathcal{L}_p = -\frac{1}{mn} \sum_{i=1}^{mn} \log\left(\frac{e^{f_i^{l_i}}}{\sum_{j=1}^{k} e^{f_i^j}}\right) \tag{3}$$

where k is the number of classes, f_i denotes the feature vector of the ith sample and f_i^j denotes the predicted probability of the ith sample belonging to class j, and l_i is the ground true class label.

Identity Regularization. Recent researches have found deep networks to be fragile when facing adversarial examples, which are indistinguishable to human eyes while causing a sharp drop in performance of networks [19]. Inpainting, as an image preprocess method, should provide convenience for later high-level tasks like recognition. Toward this end, both visual and identity information need rightly diffuse into the region to complete. Existing methods tend to neglect the huge divergence between human and machine perception. During the inpainting process, identity information diffuse into the inpainted region with less restraint, leaving a gap of recognition accuracy between the inpainted face images and the ground truths. In order to truly benefit recognition performance, we introduce a novel identity network to narrow the gap in feature space. We use the VGGFace [13] and extract the identity feature representations from its $fc7$ layer. We enforce the effective diffusion of identity information by demanding the features to be as close to the ground truth as possible. Different from perceptual loss, we use another two images with same identity to further restrain the diffusion. We define a corresponding identity content loss $\mathcal{L}_{ip}$ as the ℓ_1 distance between features of inpainted face and realistic face, formulated as below:

$$\mathcal{L}_{ip} = \frac{1}{3}\sum_{i=1}^{3}\left| f_{\mathcal{G}(\mathbf{x},\mathbf{M})} - f_{Y_i} \right| \tag{4}$$

where $|*|$ is ℓ_1 norm operator, $f_{\mathcal{G}(\mathbf{x},\mathbf{M})}$ and f_y denotes the feature representation of the completed face and realistic face respectively, Y_i the ground truth $\mathbf{y}$ when $i = 1$, and images with same identity when $i = 2, 3$.

3.4 Total Loss

Intergrating the terms above, the total loss function can be formulated as

$$\mathcal{L} = \mathcal{L}_{rec} + \lambda_{D_l}\mathcal{L}_{D_l} + \lambda_{D_g}\mathcal{L}_{D_g} + \lambda_p\mathcal{L}_p + \lambda_{ip}\mathcal{L}_{ip} \tag{5}$$

where λ_{D_l}, λ_{D_g}, λ_p and λ_{ip} are weights, the losses $\mathcal{L}_{rec}$, $\mathcal{L}_{D_l}$, $\mathcal{L}_{D_g}$, $\mathcal{L}_p$ and $\mathcal{L}_{ip}$ are defined in Equation 1, 2, 3 and 4, respectively. The optimization is implemented in Caffe [6] with ADAM algorithm.

4 Experiment

4.1 Dataset

We use the CelebA [8] dataset to train our model. It consists of 202,599 face images covering 10,177 celebrities. Each image is aligned and cropped, via similarity transform based on the two eyes and mouth. As the provided images are rectangle, we do simple cropping and resize them into 128×128 pixels. We follow the standard split with 162,770 images for training and 19,867 for validation. As our aim is to improve the recognition performance, we apply the LFW [4] dataset for testing. As one of the most authoritative dataset for face verification in the unconstrained environment, the LFW dataset consists of 13,323 images of 5,749 identities. To show justice, we apply the same process as in CelebA dataset to prepare the data. We set the mask size to be 64 × 64 for training. To prevent over-fitting and improve the generalization ability of the model, we do data augmentation that includes flipping, shift and rotation (±15 degrees). During the training, the size of the mask is fixed but the position is random, preventing the model from latching on certain contents only.

4.2 Implementation Details

The masked region of the image is pre-filled with random noise before inputting into the generator. In this experiment, we set $\lambda_{D_l} = \lambda_{D_g} = 300$, $\lambda_p = 0.05$ and $\lambda_{ip} = 2 \times 10^5$, other hyper parameters like learning rate refer to [15].

During the training, joint optimization of the generator and two discriminators brings great instability. To settle this, we split the training process into three stages and gradually increase the difficulty. At the first stage, only reconstruction loss is used. Then at the second stage, we finetune the model with local adversary loss. Finally we introduce the global adversary loss, the parsing loss and the identity loss. Each stage serves as a pre-training for the next one, allowing more efficient and stable training. When training with adversary loss, we further split it into two phases to ensure stable training. At the local and global adversary stage, we fixed the generator and train the discriminator for some time to initiate, then jointly train the generator and discriminators.

4.3 Qualitative Results

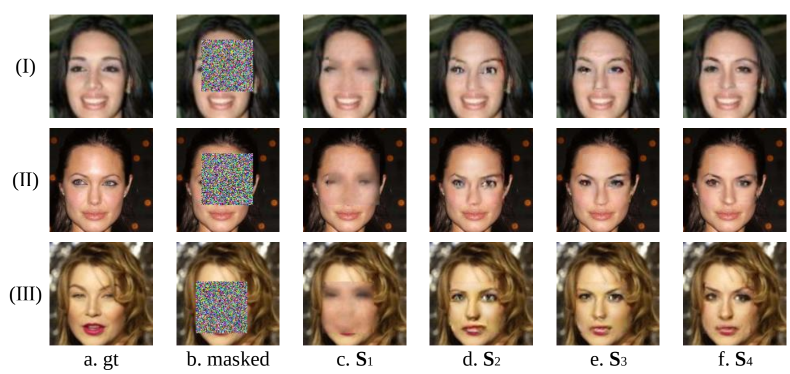

Fig. 2. Completion results under different settings of our models. The first two column are original images and masked input. The others are completion results of models trained with $\mathbf{S}_1 : \{\mathcal{L}_{rec}\}$, $\mathbf{S}_2 : \{\mathcal{L}_{rec}, \mathcal{L}_{D_l}\}$, $\mathbf{S}_3 : \{\mathcal{L}_{rec}, \mathcal{L}_{D_l}, \mathcal{L}_{D_g}\}$, $\mathbf{S}_4 : \{\mathcal{L}_{rec}, \mathcal{L}_{D_l}, \mathcal{L}_{D_g}, \mathcal{L}_p\}$, from left to right.

First we investigate the influence of different regularization terms, by only use incomplete combination of them and make comparations. The completion results by models of different stages are present in Fig 2, where the models are trained with $\mathbf{S}_1 : \{\mathcal{L}_{rec}\}$, $\mathbf{S}_2 : \{\mathcal{L}_{rec}, \mathcal{L}_{D_l}\}$, $\mathbf{S}_3 : \{\mathcal{L}_{rec}, \mathcal{L}_{D_l}, \mathcal{L}_{D_g}\}$, $\mathbf{S}_4 : \{\mathcal{L}_{rec}, \mathcal{L}_{D_l}, \mathcal{L}_{D_g}, \mathcal{L}_p\}$, respectively. As Fig 2 shows, first at the reconstruction stage, the generator reconstruct the general shape of facial components, though blurry. More details are restored as local adversary loss is invited, which greatly help the realism. The global adversary loss then make some adjustment to keep consistency around the borders of region to inpaint. On top of that, the semantic parsing loss refines the synthesised content to keep in harmony with existing contents. Finally, the identity-preserving term further refines the image, narrowing the gap in feature space, therefore truly benefiting the field of occluded face recognition.

Fig 3 shows our completion results on LFW dataset. It is worth noting that in the second and third column at (II) row, given different definition of mask, the model presents different inpainting accordingly. This proves our model capable of utilize semantic context and do intelligent inpainting.

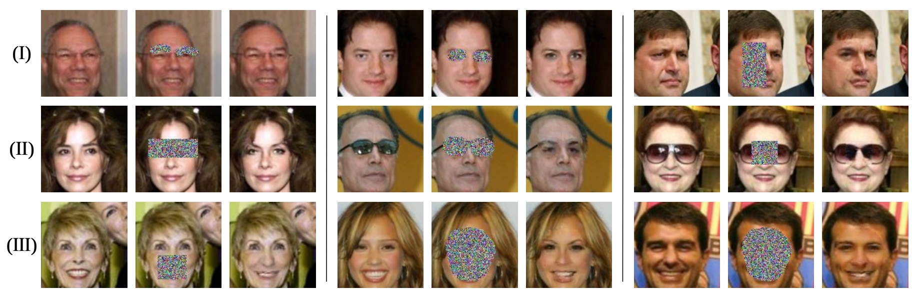

Fig. 3. Inpainting Results with irregular mask. In each panel, from left to right: ground truth, masked face, faces completed by IP-GANs.

4.4 Quantitative Results

Beyond the subjective visual results, we further perform quantitative evaluation using metrics with PSNR and SSIM (Structural Similarity Index) following [7, 14]. PSNR directly measures the pixel-level difference between two pictures, and SSIM measures the similarity between images from a global structural view. Moreover, the accuracy is applied for evaluating the performance of occluded face recognition. These three metrics are computed between every completed image and its corresponding original one. The results are shown in Table 1.

Table 1. Quantitative results of un-inpainted face images, face images inpainted by GFC [7] and the our proposed models trained with different loss setting. Specifically, model $\mathbf{S}_{1-4}$ are trained with $\mathbf{S}_1 : \{\mathcal{L}_{rec}\}$, $\mathbf{S}_2 : \{\mathcal{L}_{rec}, \mathcal{L}_{D_l}\}$, $\mathbf{S}_3 : \{\mathcal{L}_{rec}, \mathcal{L}_{D_l}, \mathcal{L}_{D_g}\}$, $\mathbf{S}_4 : \{\mathcal{L}_{rec}, \mathcal{L}_{D_l}, \mathcal{L}_{D_g}, \mathcal{L}_p\}$, and the final IP-GANs $\mathbf{S}_5 : \{\mathcal{L}_{rec}, \mathcal{L}_{D_l}, \mathcal{L}_{D_g}, \mathcal{L}_p, \mathcal{L}_{ip}\}$

Metrics	un-inpainted	GFC [7]	$\mathbf{S}_1$	$\mathbf{S}_2$	$\mathbf{S}_3$	$\mathbf{S}_4$	$\mathbf{S}_5$
SSIM	0.7285	0.9335	0.9384	0.9337	0.9405	0.9408	0.9442
PSNR(dB)	5.3952	16.9467	18.088	17.2232	18.7169	18.7346	19.301
Accuracy	0.7998	0.819	0.8417	0.8335	0.842	0.844	0.855

5 Conclusion

In this work, we propose a novel Identity Preserving GANs (IP-GANs) for occluded face recognition tasks. The model can successfully synthesize semantically valid and visually plausible contents for the missing facial key parts from random noise and greatly promote recognition performance. Thorough experiments show that our model is capable of handling occlusions with varieties of shapes and sizes, providing general and effective solution for occluded face recognition.

Acknowledgment

This work is supported in part by the National Natural Science Foundation of China (61772513 & 61402463), the National Key Research and Development Plan (2016YFC0801005) and the International Cooperation Project of Institute of Information Engineering, Chinese Academy of Sciences (Y7Z0511101). Shiming Ge is also supported by Youth Innovation Promotion Association, CAS.

References

1. Bertalmio, Marcelo, Sapiro, Guillermo, Caselles, Vincent, Ballester, Coloma: Image inpainting. Siggraph **4**(9), 417–424 (2005)
2. Deng, Y., Dai, Q., Zhang, Z.: Graph laplace for occluded face completion and recognition. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society **20**(8), 2329–2338 (2011)
3. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: International Conference on Neural Information Processing Systems. pp. 2672–2680 (2014)
4. Huang, G.B., Ramesh, M., Berg, T., Learned-Miller, E.: Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Tech. Rep. 07-49, University of Massachusetts, Amherst (October 2007)
5. Ishikawa, H., Ishikawa, H., Ishikawa, H.: Globally and locally consistent image completion. ACM (2017)
6. Jia, Y., Shelhamer, E., Donahue, J., et al: Caffe: Convolutional architecture for fast feature embedding. arXiv preprint arXiv:1408.5093 (2014)
7. Li, Y., Liu, S., Yang, J., Yang, M.H.: Generative face completion. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (July 2017)

8. Liu, Z., Luo, P., Wang, X., Tang, X.: Deep learning face attributes in the wild. In: IEEE International Conference on Computer Vision. pp. 3730–3738 (2015)
9. Lu, H., Li, B., Zhu, J., Li, Y., Li, Y., Xu, X., He, L., Li, X., Li, J., Serikawa, S.: Wound intensity correction and segmentation with convolutional neural networks. Concurrency & Computation Practice & Experience **29**(6) (2016)
10. Lu, H., Li, Y., Chen, M., Kim, H., Serikawa, S.: Brain intelligence: Go beyond artificial intelligence. Mobile Networks & Applications **23**(2), 368–375 (2017)
11. Lu, H., Li, Y., Mu, S., Wang, D., Kim, H., Serikawa, S.: Motor anomaly detection for unmanned aerial vehicles using reinforcement learning. IEEE Internet of Things Journal **5**(4), 2315–2322 (2018)
12. Oh, H.J., Lee, K.M., Sang, U.L.: Occlusion invariant face recognition using selective local non-negative matrix factorization basis images. Image & Vision Computing **26**(11), 1515–1523 (2008)
13. Parkhi, O.M., Vedaldi, A., Zisserman, A.: Deep face recognition. In: British Machine Vision Conference (2015)
14. Pathak, D., Krahenbuhl, P., Donahue, J., Darrell, T., Efros, A.A.: Context encoders: Feature learning by inpainting. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2536–2544 (2016)
15. Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks. CoRR **abs/1511.06434** (2015), http://arxiv.org/abs/1511.06434
16. Serikawa, S., Lu, H.: Underwater image dehazing using joint trilateral filter. Pergamon Press, Inc. (2014)
17. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. CoRR **abs/1409.1556** (2014), http://arxiv.org/abs/1409.1556
18. Smith, B.M., Zhang, L., Brandt, J., Lin, Z., Yang, J.: Exemplar-based face parsing. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 3484–3491 (2013)
19. Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., Fergus, R.: Intriguing properties of neural networks. Computer Science (2013)
20. Wang, C., Xu, C., Wanga, C., Tao, D.: Perceptual adversarial networks for image-to-image transformation. IEEE Transactions on Image Processing **27**(8), 4066–4079 (2018)
21. Wright, J., Ganesh, A., Zhou, Z., Wagner, A., Ma, Y.: Demo: Robust face recognition via sparse representation. In: IEEE International Conference on Automatic Face & Gesture Recognition. pp. 1–2 (2009)
22. Xie, J., Xu, L., Chen, E.: Image denoising and inpainting with deep neural networks. In: International Conference on Neural Information Processing Systems. pp. 341–349 (2012)
23. Xu, X., He, L., Lu, H., Gao, L., Ji, Y.: Deep adversarial metric learning for cross-modal retrieval. World Wide Web-internet & Web Information Systems pp. 1–16 (2018)
24. Yu, J., Lin, Z., Yang, J., Shen, X., Lu, X., Huang, T.S.: Generative image inpainting with contextual attention. CoRR **abs/1801.07892** (2018), http://arxiv.org/abs/1801.07892
25. Zhang, S., He, R., Sun, Z., Tan, T.: Demeshnet: Blind face inpainting for deep meshface verification. IEEE Transactions on Information Forensics & Security **13**(3), 637–647 (2017)
26. Zhang, W., Shan, S., Chen, X., Gao, W.: Local gabor binary patterns based on kullbackleibler divergence for partially occluded face recognition. IEEE Signal Processing Letters **14**(11), 875–878 (2007)