International Journal of Engineering

Journal Homepage: www.ije.ir

Empowering Face Recognition Methods using a GAN-based Single Image Super-Resolution Network

M. B. Shahbakhsh, H. Hassanpour*

Computer Engineering and IT Department, Shahrood University of Technology, Shahrood, Iran

PAPER INFO

Paper history:
Received 30 April 2022
Received in revised form 06 June 2022
Accepted 07 June 2022

Keywords:
Face Recognition
Super-Resolution
Face Hallucination
Residual Self-attention
Generative Adversarial Network
Identity Preserving

ABSTRACT

Face recognition is one of the most common authentication techniques widely used due to its easy access. In many face recognition applications, captured images are often of low resolution. Face recognition methods perform poorly on low resolution images because they are trained on high resolution face images. Although existing face hallucination methods may generate visually pleasing images, they cannot improve the performance of face recognition methods at low resolution as the structure of the face image and high-frequency details are not sufficiently preserved. Recent advances in deep learning have been used in this paper to propose a new face super-resolution approach to empower face recognition methods. In this paper, a Generative Adversarial Network is used to empower face recognition in low-resolution images. This network considers image edges and reconstructs high-frequency details to preserve the face structure. The proposed technique to generate super-resolved features is usable in any face recognition method. We have used some state-of-the-art face recognition methods to evaluate the proposed method. The results showed a significant impact of the proposed method on the accuracy of face recognition of low resolution images.

doi: *10.5829/ije.2022.35.10a.05*

1. INTRODUCTION

Among the biometric techniques available for authentication, face recognition is widely used due to its simplicity of use, easy access, and the possibility of using it at long distances and secretly [1]. Face recognition is one of the primary research fields in computer vision and plays an essential role in machine learning. Video surveillance, access control of buildings, criminal identification, and self-driving cars are just a few applications of face recognition systems.

Despite recent advances, face recognition still has challenges such as obstruction, angle and brightness changes, resolution, and other factors. For example, in areas where Closed-Circuit Television (CCTV) cameras are used for surveillance, cropped face images usually have breakdowns such as blur and noise [2]. Also, the extracted face images may not have a suitable resolution due to the person's distance from the camera. In this problem, we encounter the challenge of recognizing faces in uncontrolled conditions and using low-resolution (LR) images.

Face recognition techniques based on deep neural networks have recently attained accuracy nearby human performance. Some are even more accurate than humans [3-5]. Advanced convolutional neural networks (CNNs) such as FaceNet [5], DeepFace [4], ArcFace [6], and VGGFace [7] have achieved extraordinary recognition performance in the Wild (LFW) database on Labeled Faces [8]. Alexnet-based Deep-Face [4] achieved 97.35% accuracy for labeled face images in the Wild (LFW) database. After that, VGGFace, based on VGG-16 [7], achieved 98.95% accuracy, and in 2018, Arc Face [6] reached the highest accuracy, 99.83%. Recently, state-of-the-art methods have used ResNet instead of visual geometry (VGG) because of its better performance and lower memory consumption [9]. These neural networks are trained on large datasets containing high-resolution (HR) images [10-14], including facial expressions and pose changes. Several factors affect the accuracy of face recognition methods,

*Corresponding Author Institutional Email: h.hassanpour@shahroodut.ac.ir (H. Hassanpour)

Please cite this article as: M. B. Shahbakhsh, H. Hassanpour, Empowering Face Recognition Methods using a GAN-based Single Image Super-Resolution Network, ***International Journal of Engineering, Transactions A: Basics***, Vol. 35 No. 10, (2022), 1123-1132

such as poses, brightness, and image resolution. It has been shown that one of the factors influencing the performance of face recognition is image resolution. [15, 16]. Face recognition accuracy decreases rapidly with reducing image resolution [17, 18]. Hence, face recognition with LR images remains a demanding problem [19].

Smaller images with lower resolution have less information than higher resolution images. Image super-resolution (SR) methods have been presented to increase image resolution and details [20, 21]. The LR image is mapped to an HR space for training the recognition model. For example, a deep SR method C-SRIP (cascaded SR and identity priors) was proposed by Grm *et al.* [20], with three cascading SR networks by a 2x SR factor including identity priorities in learning. Cascading SR networks are based on CNNs. Banerjee and Das [22] introduced a deep neural network based on generative adversarial networks (GANs) [23] to reconstruct more realistic face images from LR face inputs. Likewise, Zangeneh *et al.* [24] proposed a two-section deep convolutional model, Feature extraction convolutional neural network (FECNN) and SR-FECNN, for mapping HR and LR images into joint space for recognition. The SR network contains a five-layer CNN to enhance the LR input image, and FECNN is a VGG-16 pre-trained without the last two fully connected layers. These two pre-trained networks are connected and jointly trained again to minimize the loss function and reduce the distance between LR and HR paired in the common space. The LR image needs to have a fixed input size in this method. Another type of SR method is feature super-resolution (FSR) [25]. This kind of SR method increases the image resolution on the feature level. In 2018, Tan *et al.* [25] proposed an FSR-GAN neural network to enhance recognition accuracy and reduce training costs. They designed a GAN network to generate HR features from the LR features. This module has a simple structure with low training costs.

Face recognition accuracy will be severely reduced if the input image is distorted and noisy or the target is far from the camera. Especially in a large database, the face recognition model errs in distinguishing features. Previous works have used various methods like clustering [26] and hashing [27] to increase face recognition accuracy in uncontrolled conditions. These methods work well on a large dataset problem and the angularity of faces. Nevertheless, the remaining challenge in these methods is the LR input image as the features and facial components will not be distinguishable [28, 29].

In this paper, we proposed a GAN-based network to increase the resolution at the feature level. The face recognition methods can distinguish facial components more accurately using the proposed method. Many face hallucination methods generate visually desirable face images, but the accuracy of face recognition is lower than anticipated with the generated images. Because the face structure and its high-frequency details are not well preserved [30]. This paper proposes a new generative adversarial network with self-attention to reconstruct the high-frequency details. The proposed method to generate SR features is applicable in any face recognition method.

The main contributions of this paper include:

- A GAN structure is proposed to empower face recognition at LR images. This network considers image edges and restores high-frequency details to preserve the face structure.
- A self-attention mechanism is used residually to employ both global and local features. Super-resolving global and local features improve the accuracy of face recognition methods.

2. RELATED WORKS

After Dong *et al.* introduced the Super-Resolution Convolutional Neural Network (SRCNN) [31], the use of deep learning techniques for single image SR became widely considered. Further advances in network architectures [32] and upsampling techniques [33] help others improve quantitative and qualitative results. Wang *et al.* [34] expressed a deeper study of the developments in this area.

Deeper models with more layers have been shown to perform sufficiently in several tasks, including super-resolving [35]. The first extremely deep model for SR tasks is the VDSR model [36]. This model employs VGG architecture with 20 layers. This network also uses multi-scale and SR in a residual way. LR images are super-resolved in several scales by reconstructing the high-frequency details and adding them to the output of bicubic interpolation. In order to decrease the VDSR parameters, the DRCN model was introduced [37]. A recursive convolutional layer has been used 16 times in the DRCN model. A multi-supervised learning strategy has been deployed to devastate the tribulation of training this network. Since ResNet has surpassed the VGG model in several tasks, it was a good choice for the SR task [9]. SRResNet has been proposed for SR as the first ResNet model [32]. This network uses 16 residual blocks with batch normalization for stabilizing the training process. Lee *et al.* proposed EDSR, one of the most advanced models for general SR [38]. EDSR and SRResNet have several primary differences. In the EDSR, the batch normalization layers are deleted, the weights are initialized for a very large-scale SR, and the number of output features is increased.

Although general SR methods can be used on face images, there are techniques to take advantage of the unique features of face images [39]. The face components distortion is the main obstacle in a face SR process. Several methods utilized facial landmarks or used facial priors to address this problem. For example, a network has been proposed in FSRNet [39] to extract the facial landmark heatmaps and estimate the priors. In the FSRNet method, the prior details are extracted using an Hour Glass (HG) [40] model to produce roughly super-resolved face images. Then, in the encoder network, the prior heatmaps are joined with feature maps. Eventually, a decoder network uses these feature maps to produce a super-resolved face image.kim et al. [41], proposed a face alignment model to compare the super-resolved and ground-truth images' facial heatmaps. A pre-trained HG network is used with an MSE (Mean Squared Error) loss function. However, landmarks must be labeled in the HG network, which has a high computational cost.

3. PROPOSED METHOD

One of the shortcomings in most SR methods is their inability to preserve the overall structure of the face image as they focus on the local features using interconnected convolutional layers. In this paper, we intend to use both global and local features to properly reconstruct high-frequency details and essential features of the face image. Extracting the edges of the face image using the Local Binary Pattern (LBP) [42] and Unsharp Masking (UM) [43] techniques and then adding them to the input image as a preprocessing can help the proposed model to accurately extract the high-frequency information of the LR image. In addition, we use the Self-Attention mechanism residually to retrieve face components correctly. We also use an edge extraction network and its associated loss function to preserve the structure of the face image. Finally, the super-resolved features generated by the proposed model are transferred to the face recognition method to perform the recognition operation.

3. 1. Preprocessing Edge information plays an essential role in preserving the structure of the face image and retrieving high-frequency information. To better retrieve the high-frequency details of the face image, we must properly use the edges from the LR image. In addition, the use of LR image texture and its delicate edges are essential in preserving the structure. In this paper, we use some widely used methods such as UM and LBP to enhance the edges of LR images in preprocessing. In this section, the delicate edges and texture of the image are extracted by the LBP method and then added to the input image. Finally, all the edges of the LR image are sharpened using UM. Figure 1 shows an example of adding LBP output to an input image. The effect of λ, the gain factor of the filter in UM, is shown in Figure 2. The extracted high-frequency details are scaled with λ. Then, these details are added to the input face image.

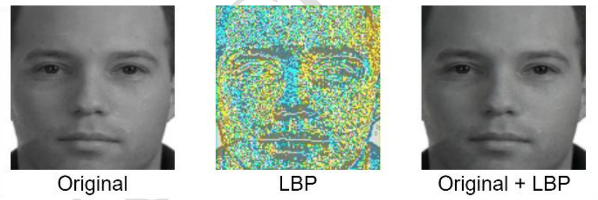

Figure 1. Adding LBP output to an input image

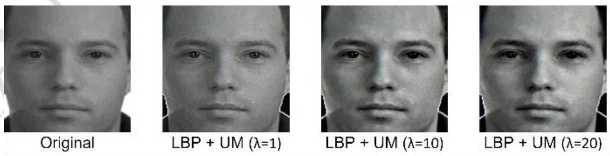

Figure 2. The effects of applying LBP and UM with different values of λ

Following the results shown in Figure 2, λ = 20 is chosen because the image details are more enhanced.

3. 2. Network Structure In this paper, we used a GAN model for super-resolving face images. GAN is a deep learning network that includes two networks called generator and discriminator. These two networks are in a zero-sum game [44].

3. 2. 1. Generator Network The overall structure used for training the generator network is shown in Figure 3. The proposed generator structure supper-

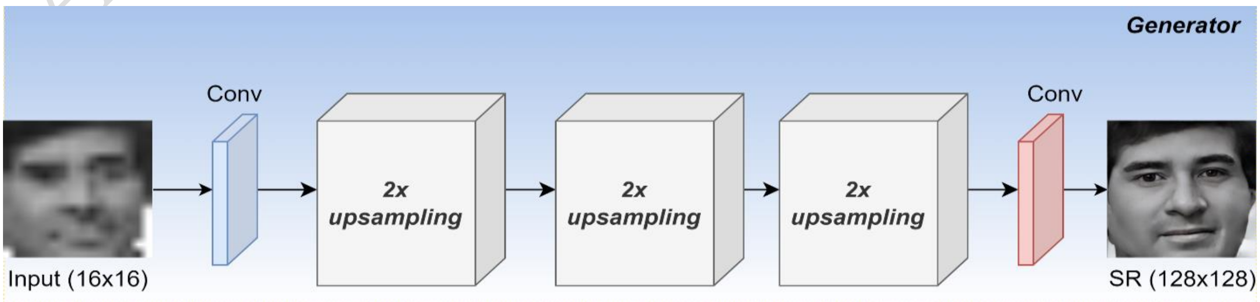

Figure 3. The overall structure used for training the proposed generator network

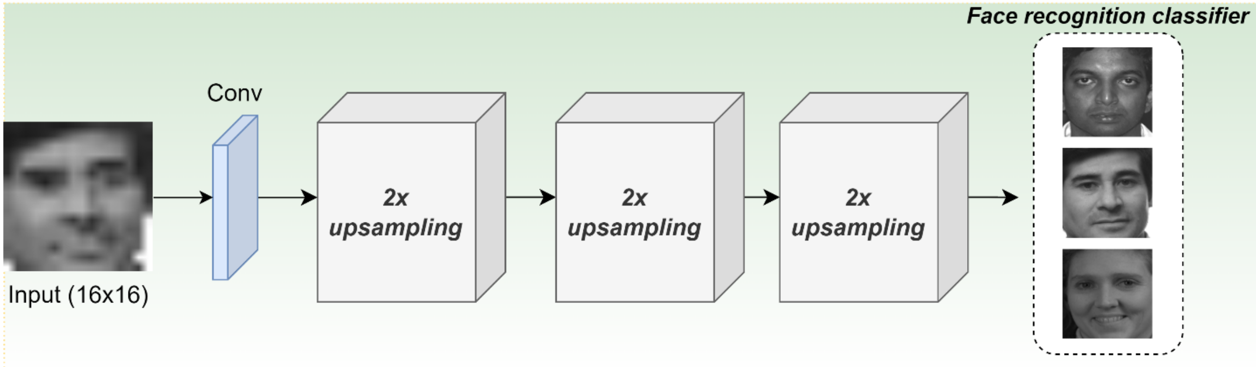

Figure 4. The combination of the face recognition classifier with the proposed method

resolves the input image up to 8x. Two convolutional layers are placed at the end and beginning of the proposed generator structure. The first convolution layer has a 3×3 filter with stride one, and 512 output channels. The final 3×3 convolutional layer generates the output image.

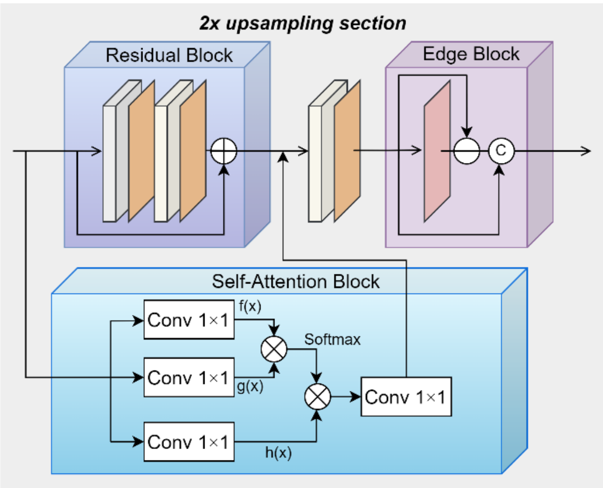

Figure 5. The proposed structure for the generator 2x upsampling in Figure 4

Since this research aims not to generate an RGB super-resolved face image, the last 3×3 convolutional layer is only in the training section. This layer is removed from the network when combining the proposed method with the face recognition classifier. The combination of the face recognition classifiers with the proposed method is shown in Figure 4. The structure of the “2x upsampling” section is shown in Figure 5.

As shown in Figure 5, the 2x upsampling structure consists of a residual self-attention block, a residual block, an edge block, and a transposed convolutional block. The convolution kernel and original features are used in convolutional layers to compute the output features. The output feature map is usually local and limited. Therefore, we have used a residual self-attention mechanism to extract global and context information. Self-attention mechanism enhances the valuable features and learns each feature channel's weights by data-driven learning.

In summary, in the proposed 2x upsampling structure in Figure 5, the residual self-attention block extracts global features, and simultaneously, local features are extracted in the residual block. Finally, all global and local features are combined. The residual block has two convolution layers followed by two Rectifier Linear Unit (ReLU) activation functions. The convolutional kernel in all residual blocks is 3×3, and in the first 2x upsampling block, the number of filters is set to 512, in the second one 256, and 64 for the third one.

In the residual self-attention block, the combined feature map received from previous layers is first converted into two new feature spaces called f and g. Then the attention map is calculated by a softmax operation. The h channel normalizes the correlation between each pixel and all other position pixels in the feature maps. Finally, a 1×1 convolution layer is used. The output feature maps from the residual self-attention block and residual block are combined cascade-wise, and the transposed convolution layer is used for the 2x upsampling operation. We also used a ReLU activation function in transposed convolution layer.

After upsampling feature maps using the transposed convolutional layer, to preserve the structure of the face image, the edge block is used to extract high-frequency details and sharpen edge information. We have used an average pooling layer in the edge block, followed by a 1×1 convolution layer. The smooth features extracted by the average pooling layer are subtracted from the main feature map to extract high-frequency details. Then, the extracted high-frequency information are concatenated with the original feature map to enhance the result and preserve the structure. As shown in Figure 3, the 2x upsampling operation is performed three times on an LR input image.

3. 2. 2. Discriminator Network In the proposed discriminator network, we used seven convolution layers with Leaky Relu followed by a fully connected layer. The kernel size for convolution layers is 3×3, and filters are set to 128, 256, and 512.

3. 3. Loss Functions In this research, we have used several loss functions to train the proposed generator network. Each loss function compares the generated face image with the ground-truth face image from different aspects, and finally, the sum of these functions is applied to the gradient as a total loss.

3. 3. 1. Context Loss We have used the pixel-wise MSE function to decrease the distance between the ground truth and generated image. The MSE loss function is defined as:

$$L_{context} = \frac{1}{wh}\sum_{c=1}^{RGB}\sum_{x=1}^{w}\sum_{y=1}^{h}\|I_{GT}(x,y) - G(I_{LR})(x,y)\|^2 \tag{1}$$

where I_{GT} is the ground truth image, I_{LR} is the LR image, and c is RGB channels.

3. 3. 2. Edge Loss As mentioned in the previous sections, edge information is essential in preserving image structure. We have dealt with this issue both in preprocessing and the generator network. To train the model, preserve the structure, and optimal use of edges and high-frequency information, in this section, we define a loss function that compares the edges obtained from the edge block with the edges extracted by the Canny method. The edge block loss function is defined as:

$$L_{edge} = \frac{1}{r^2wh}\sum_{x=1}^{rw}\sum_{y=1}^{rh}\|C(I_{GT})(x,y) - E(I_{LR})(x,y)\|^2 \tag{2}$$

where r is the scaling factor, set to be 1, 2, and 4, C is the Canny edge detector method, and E is our edge extraction block.

3. 3. 3. GAN Loss As mentioned before, we utilize a GAN structure to generate more realistic super-resolved images. The GAN loss is defined as:

$$L_{GAN} = \mathbb{E}[\log D(I_{GT})] - \mathbb{E}[\log(1 - D(G(I_{LR})))] \tag{3}$$

where $\mathbb{E}$ is the probability distribution expectation, and D is the discriminator network.

3. 3. 4. Luminance Loss RGB domain is the most common image representation format used in image application researches. In face hallucination, several state-of-the-art methods have used RGB format with MSE loss function to compare the distance between output image and ground truth. However, the YUV domain is another popular representation format widely used in image processing; and following the results reported in literature [45], the YUV color space is better than the RGB format in terms of perceptually quality. In addition, the RGB format has details that combine chrominance and luminance components, which yields redundancy to color shift, color distortion, and channel information. We used the Luminance distance to reduce the above effects and generate enhanced SR images, which converted RGB format to YUV format to independently distribute illumination and color components.

3. 3. 5. Identity Loss Pixel distance is unreliable for comparing images on fidelity and diversity. For example, suppose a pixel is moved in the generated image. The difference is not evident in an HR image with millions of pixels. However, the absolute difference will be huge based on the pixel distance. One alternative is to compare images at the higher-level features instead of the pixels. Higher-level semantic information is less sensitive to small changes.

In this paper, to compare the high-level features of generated and ground truth images, we used the pre-trained Inception V3 network to extract class-encoded vectors. We used the last pooling layer of the Inception V3 network prior to the output classification to extract features. The covariance and mean of the images are calculated, and activations are summarized as multivariate Gaussian. Then the statistics are computed for activations across generated and ground truth images. Finally, the distance between two obtained distributions is calculated using the Frechet distance. The identity loss is defined as:

$$Identity\ loss = \|\mu_x - \mu_y\|^2 + Tr(\textstyle\sum x + \sum y - 2\sqrt{\sum x \sum y}) \tag{4}$$

where μ_x and μ_y are the feature-wise mean of the ground truth and generated images, respectively. Including 512 element vectors where each element is the mean feature perceived across the images, $\sum x$ and $\sum y$ are the covariance matrices for the generated and ground truth feature vectors. $Tr()$ represents the trace linear algebra operation.

3. 3. 6. Total Loss Eventually, all loss functions are aggregated. Therefore, the total loss is as follows:

$$L_T = L_{Context} + \alpha L_{Edge} + \beta L_{GAN} + L_{Luminance} + \gamma L_{Identity} \tag{5}$$

where α, β, and γ are weight coefficients.

4. EXPERIMENTS AND DISCUSSION

In this section, we first introduce the datasets used to train the proposed SR network. The implementation details of the proposed method are then explained.

Finally, we will discuss the effect of the proposed approach on face recognition accuracy.

4. 1. Datasets In this paper, for the training section, 990 face images from the FERET [46] Frontal (fa) collection are employed, along with images from several standard datasets such as MUCT [47], FEI [48], Face94, Psychological Image Collection at Stirling (PICS), and in total there are 1684 images in this part of training data. The CelebA [49] is a celebrities' face dataset with more than 0.2 million face images. We used 174,400 face images in the training set and 21,230 images in the test set.

To evaluate the impact of the proposed method on face recognition accuracy, we used 990 images from the frontal set (Fa) of the FERET dataset for the training section, and their related Fb images were used for testing.

4. 2. Implementation Details As mentioned in the preprocessing section, first, LBP features are extracted from the LR image and added to the input image. Since we intend to extract high-frequency details, texture, and delicate edges of the input image, the LBP technique is used in "uniform" mode with 8 points and a radius of 1. After extracting these details and adding them to the input image, we used the UM technique to sharpen all the edges of the input image. To improve the network performance and make the high-frequency details clearer, based on the results presented in Figure 2, we employed λ=20 in the UM method.

As shown in Figure 3, the input of the proposed generator network is a 16×16 LR image, which is super-resolved to 128 × 128. For training the network, we used Adam optimizer with a learning rate of 0.0001, epsilon = 1e-07, $\beta_1 = 0.5$, and $\beta_2 = 0.999$ in both generator and discriminator networks. All codes were written in Python 3.7, and we used Tensorflow 2. We have used a GeForce GTX 3060 GPU for network training and evaluation. As mentioned before, we do not intend to generate the final RGB super-resolved image to recognize the face. Therefore, the final layer is removed in combining the proposed method with a face recognition method, and the super-resolved features are given to the face recognition method.

4. 3. Super-resolution Results This research aims not to generate HR images at the output but to demonstrate the performance of the proposed method in super-resolving facial images. We compared the proposed method with several state-of-the-art face SR methods such as EIPNet [50] and GFP-GAN [51]. Qualitative comparison results are shown in Figure 6.

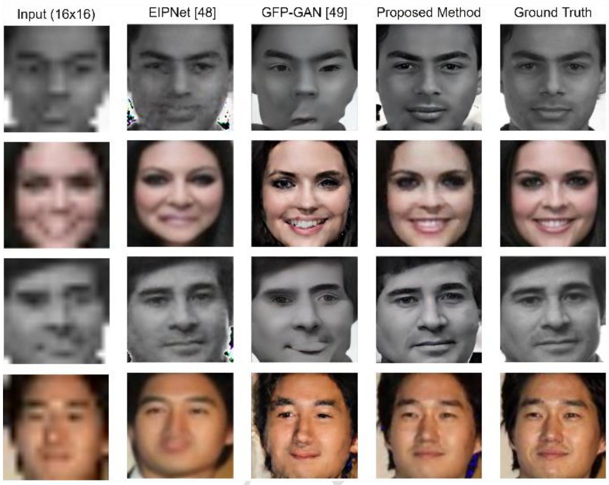

Figure 6. Qualitative comparison of SR images obtained using several state-of-the-art methods

As shown in Figure 6, the structure of the generated images with the proposed method is well preserved due to the use of high-frequency details in different stages.

4. 4. Face recognition Results The purpose of face hallucination is to improve facial recognition accuracy. Many face SR methods generate eye-catching face images. However, when they are used in a face recognition method, the accuracy is not as expected. Because they produce new details instead of preserving the original structure of the LR image. In this section, we intend to examine the effect of LR face images on some of the newest face recognition methods, such as Nikan [52], Face Net [53], VGG Face [54], and Arc Face [6]. Then we showed the effect of using the proposed method compared to several state-of-the-art super-resolution methods on increasing face recognition accuracy. Therefore, we perform our experiments on several state-of-the-art SR methods, such as DFDNet [55], EIPNet [50], and GFP-GAN [51]. Table 1 compares the accuracy of some face recognition methods using the proposed method versus other SR methods. As shown in Table 1, the proposed method performs better than other SR methods in face recognition.

TABLE 1. Face recognition accuracy comparison of the proposed method with state-of-the-art methods on LR images

		Super-resolution (%)						
		Original (128×128)	**LR (16×16)**	**Bicubic**	**DFDNet** [55]	**EIPNet** [50]	**GFPGAN** [51]	**Proposed method**
Face Recognition	**Nikan** [52]	*92.7*	3.1	9.32	41.8	49.04	26.32	**71.84**
	Face Net [53]	*98.93*	0.11	0.05	42.40	56.92	28.18	**75.09**
	VGG Face [54]	*99.10*	8.31	10.15	64.41	66.13	34.07	**79.51**
	Arc Face [6]	*98.82*	2.85	3.02	52.96	63.22	27.28	**76.89**

5. CONCLUSION

In this paper, we proposed a GAN-based network to increase the image resolution at the feature level and improve the accuracy of face recognition methods. The results provided in this paper indicate that a face recognition method can distinguish facial components more accurately using the proposed method. Many face super-resolution methods generate visually pleasant face images, but face recognition accuracy is lower than expected with their generated images. Because the structure of the face image and its high-frequency details are not well preserved. The proposed network considers the image's edges and restores high-frequency details to preserve the face's structure. The generated super-resolved features can be used in any face recognition method to improve the accuracy of recognition.

6. REFERENCES

1. Kortli, Y., Jridi, M., Al Falou, A. and Atri, M., "Face recognition systems: A survey", ***Sensors***, Vol. 20, No. 2, (2020), 342. DOI: 10.3390/s20020342.
2. Annalakshmi, M., Roomi, S. and Naveedh, A.S., "A hybrid technique for gender classification with slbp and hog features", ***Cluster Computing***, Vol. 22, No. 1, (2019), 11-20. DOI: 10.1007/s10586-017-1585-x.
3. Yang, W., Gao, H., Jiang, Y., Yu, J., Sun, J., Liu, J. and Ju, Z., "A cascaded feature pyramid network with non-backward propagation for facial expression recognition", ***IEEE Sensors Journal***, Vol. 21, No. 10, (2020), 11382-11392. DOI: 10.1109/JSEN.2020.2997182.
4. Taigman, Y., Yang, M., Ranzato, M.A. and Wolf, L., "Deepface: Closing the gap to human-level performance in face verification", in Proceedings of the IEEE conference on computer vision and pattern recognition. (2014), 1701-1708.
5. Schroff, F., Kalenichenko, D. and Philbin, J., "Facenet: A unified embedding for face recognition and clustering", in Proceedings of the IEEE conference on computer vision and pattern recognition, (2015), 815-823.
6. Deng, J., Guo, J., Xue, N. and Zafeiriou, S., "Arcface: Additive angular margin loss for deep face recognition", in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, (2019), 4690-4699.
7. Parkhi, O.M., Vedaldi, A. and Zisserman, A., "Deep face recognition", (2015).
8. Huang, G.B., Mattar, M., Berg, T. and Learned-Miller, E., "Labeled faces in the wild: A database forstudying face recognition in unconstrained environments", in Workshop on faces in'Real-Life'Images: detection, alignment, and recognition, (2008).
9. He, K., Zhang, X., Ren, S. and Sun, J., "Deep residual learning for image recognition", in Proceedings of the IEEE conference on computer vision and pattern recognition, (2016), 770-778.
10. Taigman, Y., Yang, M., Ranzato, M.A. and Wolf, L., "Web-scale training for face identification", in Proceedings of the IEEE conference on computer vision and pattern recognition, (2015), 2746-2754.
11. Cao, Q., Shen, L., Xie, W., Parkhi, O.M. and Zisserman, A., "Vggface2: A dataset for recognising faces across pose and age", in 2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018), IEEE. (2018), 67-74. DOI: 10.1109/FG.2018.00020.
12. Guo, Y., Zhang, L., Hu, Y., He, X. and Gao, J., "Ms-celeb-1m: A dataset and benchmark for large-scale face recognition", in European conference on computer vision, Springer. (2016), 87-102. DOI: 10.1007/978-3-319-46487-9_6.
13. Yi, D., Lei, Z., Liao, S. and Li, S.Z., "Learning face representation from scratch", ***arXiv preprint arXiv:1411.7923***, (2014). DOI: 10.48550/arXiv.1411.7923.
14. Lui, Y.M., Bolme, D., Draper, B.A., Beveridge, J.R., Givens, G. and Phillips, P.J., "A meta-analysis of face recognition covariates", in 2009 IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems, IEEE. (2009), 1-8. DOI: 10.1109/BTAS.2009.5339025.
15. Zou, W.W. and Yuen, P.C., "Very low resolution face recognition problem", ***IEEE Transactions on image processing***, Vol. 21, No. 1, (2011), 327-340. DOI: 10.1109/TIP.2011.2162423.
16. Van Ouwerkerk, J., "Image super-resolution survey", ***Image and vision Computing***, Vol. 24, No. 10, (2006), 1039-1052. DOI: 10.1016/j.imavis.2006.02.026.
17. Wu, J., Ding, S., Xu, W. and Chao, H., "Deep joint face hallucination and recognition", ***arXiv preprint arXiv:1611.08091***, (2016). DOI: 10.48550/arXiv.1611.08091.
18. Lu, Z., Jiang, X. and Kot, A., "Deep coupled resnet for low-resolution face recognition", ***IEEE Signal Processing Letters***, Vol. 25, No. 4, (2018), 526-530. DOI: 10.1109/LSP.2018.2810121.
19. Luevano, L.S., Chang, L., Méndez-Vázquez, H., Martínez-Díaz, Y. and González-Mendoza, M., "A study on the performance of unconstrained very low resolution face recognition: Analyzing current trends and new research directions", ***IEEE Access***, Vol. 9, No., (2021), 75470-75493. DOI: 10.1109/ACCESS.2021.3080712.
20. Grm, K., Scheirer, W.J. and Štruc, V., "Face hallucination using cascaded super-resolution and identity priors", ***IEEE***

Transactions on image processing, Vol. 29, (2019), 2150-2165. DOI: 10.1109/TIP.2019.2945835.

21. Seyyedyazdi, S. and Hassanpour, H., "Super-resolution of defocus blurred images", ***International Journal of Engineering, Transactions A: Basics***, Vol. 33, No. 4, (2020), 539-545. DOI: 10.5829/ije.2020.33.04a.04.
22. Banerjee, S. and Das, S., "Lr-gan for degraded face recognition", ***Pattern Recognition Letters***, Vol. 116, No., (2018), 246-253. DOI: 10.1016/j.patrec.2018.10.034.
23. Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B. and Bharath, A.A., "Generative adversarial networks: An overview", ***IEEE Signal Processing Magazine***, Vol. 35, No. 1, (2018), 53-65. DOI: 10.1109/MSP.2017.2765202.
24. Zangeneh, E., Rahmati, M. and Mohsenzadeh, Y., "Low resolution face recognition using a two-branch deep convolutional neural network architecture", ***Expert Systems with Applications***, Vol. 139, No., (2020), 112854. DOI: 10.1016/j.eswa.2019.112854.
25. Tan, W., Yan, B. and Bare, B., "Feature super-resolution: Make machine see more clearly", in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2018), 3994-4002.
26. Abbaspoor, N. and Hassanpour, H., "Face recognition in a large dataset using a hierarchical classifier", ***Multimedia Tools and Applications***, (2022), 1-19. DOI: 10.1007/s11042-022-12382-5
27. Hassanpour, H. and Ghasemi, M., "A three-stage filtering approach for face recognition", ***International Journal of Engineering***, ***Transactions B: Applications*** Vol. 34, No. 8, (2021). DOI: 10.5829/ije.2021.34.08b.06.
28. Dey, A. and Dasgupta, K., "Emotion recognition using deep learning in pandemic with real-time email alert", in Proceedings of Third International Conference on Communication, Computing and Electronics Systems, Springer. (2022), 175-190. DOI: 10.1007/978-981-16-8862-1_13.
29. Chaabane, S.B., Hijji, M., Harrabi, R. and Seddik, H., "Face recognition based on statistical features and svm classifier", ***Multimedia Tools and Applications***, Vol. 81, No. 6, (2022), 8767-8784. DOI: 10.1007/s11042-021-11816-w.
30. Shahbakhsh, M.B. and Hassanpour, H., "Enhancing face super-resolution via improving the edge and identity preserving network", in 2021 7th International Conference on Signal Processing and Intelligent Systems (ICSPIS), IEEE, (2021), 1-4. DOI: 10.1109/ICSPIS54653.2021.9729372.
31. Dong, C., Loy, C.C., He, K. and Tang, X., "Learning a deep convolutional network for image super-resolution", in European conference on computer vision, Springer, (2014), 184-199. DOI: 10.1007/978-3-319-10593-2_13.
32. Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J. and Wang, Z., "Photo-realistic single image super-resolution using a generative adversarial network", in Proceedings of the IEEE conference on computer vision and pattern recognition, (2017), 4681-4690.
33. Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A.P., Bishop, R., Rueckert, D. and Wang, Z., "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network", in Proceedings of the IEEE conference on computer vision and pattern recognition, (2016), 1874-1883.
34. Wang, Z., Chen, J. and Hoi, S.C., "Deep learning for image super-resolution: A survey", ***IEEE Transactions on pattern analysis and machine intelligence***, Vol. 43, No. 10, (2020), 3365-3387. DOI: 10.1109/TPAMI.2020.2982166.
35. Montufar, G.F., Pascanu, R., Cho, K. and Bengio, Y., "On the number of linear regions of deep neural networks", ***Advances in neural information processing systems***, Vol. 27, (2014).
36. Kim, J., Lee, J.K. and Lee, K.M., "Accurate image super-resolution using very deep convolutional networks", in Proceedings of the IEEE conference on computer vision and pattern recognition, (2016), 1646-1654.
37. Kim, J., Lee, J.K. and Lee, K.M., "Deeply-recursive convolutional network for image super-resolution", in Proceedings of the IEEE conference on computer vision and pattern recognition, (2016), 1637-1645.
38. Lim, B., Son, S., Kim, H., Nah, S. and Mu Lee, K., "Enhanced deep residual networks for single image super-resolution", in Proceedings of the IEEE conference on computer vision and pattern recognition workshops, (2017), 136-144.
39. Chen, Y., Tai, Y., Liu, X., Shen, C. and Yang, J., "Fsrnet: End-to-end learning face super-resolution with facial priors", in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, (2018), 2492-2501.
40. Newell, A., Yang, K. and Deng, J., "Stacked hourglass networks for human pose estimation", in European conference on computer vision, Springer, (2016), 483-499. DOI: 10.1007/978-3-319-46484-8_29.
41. Kim, D., Kim, M., Kwon, G. and Kim, D.-S., "Progressive face super-resolution via attention to facial landmark", ***arXiv preprint arXiv:1908.08239***, (2019). DOI: 10.48550/arXiv.1908.08239.
42. Pavithra, L. and Sharmila, T.S., "An efficient framework for image retrieval using color, texture and edge features", ***Computers & Electrical Engineering***, Vol. 70, (2018), 580-593. DOI: 10.1016/j.compeleceng.2017.08.030
43. Mortezaie, Z., Hassanpour, H. and Asadi Amiri, S., "An adaptive block based un-sharp masking for image quality enhancement", ***Multimedia Tools and Applications***, Vol. 78, No. 16, (2019), 23521-23534. DOI: 10.1007/s11042-019-7594-4
44. Liu, H., Zheng, X., Han, J., Chu, Y. and Tao, T., "Survey on gan-based face hallucination with its model development", ***IET Image Processing***, Vol. 13, No. 14, (2019), 2662-2672. DOI: 10.1049/iet-ipr.2018.6545.
45. Podpora, M., Korbas, G.P. and Kawala-Janik, A., "Yuv vs rgb-choosing a color space for human-machine interaction", in FedCSIS (Position Papers), (2014), 29-34. DOI: 10.15439/2014F206.
46. Phillips, P.J., Moon, H., Rizvi, S.A. and Rauss, P.J., "The feret evaluation methodology for face-recognition algorithms", ***IEEE Transactions on pattern analysis and machine intelligence***, Vol. 22, No. 10, (2000), 1090-1104. DOI: 10.1109/34.879790.
47. Milborrow, S., Morkel, J. and Nicolls, F., "The muct landmarked face database", ***Pattern recognition association of South Africa***, Vol. 201, (2010).
48. Thomaz, C.E. and Giraldi, G.A., "A new ranking method for principal components analysis and its application to face image analysis", ***Image and vision Computing***, Vol. 28, No. 6, (2010), 902-913. DOI: 10.1016/j.imavis.2009.11.005
49. Liu, Z., Luo, P., Wang, X. and Tang, X., "Large-scale celebfaces attributes (celeba) dataset", ***Retrieved August***, Vol. 15, No. 2018, (2018), 11.
50. Kim, J., Li, G., Yun, I., Jung, C. and Kim, J., "Edge and identity preserving network for face super-resolution", ***Neurocomputing***, Vol. 446, (2021), 11-22. DOI: 10.1016/j.neucom.2021.03.048
51. Wang, X., Li, Y., Zhang, H. and Shan, Y., "Towards real-world blind face restoration with generative facial prior", in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, (2021), 9168-9178.
52. Nikan, F. and Hassanpour, H., "Face recognition using non-negative matrix factorization with a single sample per person in a large database", ***Multimedia Tools and Applications***, Vol. 79, No. 37, (2020), 28265-28276. DOI: 10.1007/s11042-020-09394-4.

53. Jose, E., Greeshma, M., Haridas, M.T. and Supriya, M., "Face recognition based surveillance system using facenet and mtcnn on jetson tx2", in 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), IEEE, (2019), 608-613. DOI: 10.1109/ICACCS.2019.8728466.
54. Prasad, P.S., Pathak, R., Gunjan, V.K. and Ramana Rao, H., Deep learning based representation for face recognition, in Iccce 2019. 2020, Springer.419-424. DOI: 10.1007/978-981-13-8715-9_50.
55. Li, X., Chen, C., Zhou, S., Lin, X., Zuo, W. and Zhang, L., "Blind face restoration via deep multi-scale component dictionaries", in European Conference on Computer Vision, Springer, (2020), 399-415. DOI: 10.1007/978-3-030-58545-7_23.

Persian Abstract

چکیده

شناسایی چهره یکی از رایج‌ترین تکنیک‌های احراز هویت است که به دلیل دسترسی آسان آن به طور گسترده مورد استفاده قرار می‌گیرد. در بسیاری از کاربردهای تشخیص چهره، تصاویر ضبط شده وضوح پایینی دارند. روش‌های شناسایی چهره روی تصاویر وضوح پایین عملکرد ضعیفی دارند زیرا این روش‌ها عموماً با استفاده از تصاویر چهره وضوح بالا آموزش داده می‌شوند. اگرچه روش‌های بزرگنمایی چهره موجود ممکن است تصاویر بصری چشم‌نوازی تولید کنند، اما نمی‌توانند عملکرد روش‌های تشخیص چهره را در وضوح پایین بهبود بخشند زیرا ساختار تصویر و جزئیات فرکانس بالا به خوبی حفظ نمی‌شوند. در این مقاله، پیشرفت‌های اخیر در حوزه یادگیری عمیق با هدف ساخت یک روش بزرگنمایی چهره جدید جهت توانمندسازی روش‌های شناسایی چهره استفاده شده‌اند. در این مقاله، یک شبکه متخاصم مولد جهت توانمندسازی شناسایی چهره با تصاویر چهره وضوح پایین پیشنهاد شده است. این شبکه با تمرکز بر لبه‌ها و بازسازی جزئیات فرکانس بالا، ساختار تصویر چهره وضوح پایین را به خوبی حفظ می‌کند. روش پیشنهادی جهت بزرگنمایی ویژگی‌ها، قابل استفاده در هر روش شناسایی چهره است. ما از برخی از جدیدترین روش‌های شناسایی چهره جهت ارزیابی روش پیشنهادی استفاده کرده‌ایم. نتایج حاکی از تأثیر قابل توجه روش پیشنهادی بر دقت شناسایی چهره با تصاویر وضوح پایین است.