



**HAL**  
open science

# Effectiveness of Blind Face Restoration to Boost Face Recognition Performance at Low-Resolution Images

Yoanna Martínez-Díaz, Luis S. Luevano, Heydi Méndez-Vázquez

► **To cite this version:**

Yoanna Martínez-Díaz, Luis S. Luevano, Heydi Méndez-Vázquez. Effectiveness of Blind Face Restoration to Boost Face Recognition Performance at Low-Resolution Images. IWAIPR 2023 - International Workshop on Artificial Intelligence and Pattern Recognition, Oct 2023, Varadero, Cuba. pp.455–467, 10.1007/978-3-031-49552-6\_39 . hal-04393649

**HAL Id: hal-04393649**

**<https://hal.science/hal-04393649v1>**

Submitted on 23 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Effectiveness of Blind Face Restoration to Boost Face Recognition Performance at Low-Resolution Images

Yoanna Martínez-Díaz<sup>1</sup>, Heydi Méndez-Vázquez<sup>1</sup>, and  
Luis S. Luevano<sup>2,3</sup>

<sup>1</sup> Advanced Technologies Application Center (CENATAV), P.C. 12200 Havana, Cuba.  
{ymartinez, hmendez}@cenatav.co.cu

<sup>2</sup> Tecnológico de Monterrey, School of Engineering and Sciences, Nuevo León, P.C. 64700 Mexico  
luis.s.luevano@tec.mx

<sup>3</sup> Univ Rennes, Inria, CNRS, IRISA  
luis-santiago.luevano-garcia@inria.fr

**Abstract.** This paper study the effectiveness of Blind Face Restoration methods to boost the performance of face recognition systems on low-resolution images. We investigate the use of three blind face restoration techniques, which have demonstrated impressive results in generating realistic high-resolution face images. Three state-of-the-art face recognition methods were selected to assess the impact of using the generated high-resolution images on their performance. Our analysis includes both, synthesized and native low-resolution images. The conducted experimental evaluation show that this is still an open research problem.

**Keywords:** low-resolution · face recognition · blind face restoration.

## 1 Introduction

Recent developments in deep neural network architectures have allowed to achieve impressive performance improvements on several computer vision tasks, including face recognition [24]. However, the accuracy of current face recognition methods remains a significant challenge on unconstrained and native low-resolution (LR) images such as those acquired in surveillance scenarios or from a wide field of view at distance [5]. Existing methods to deal with LR image constraints can be grouped into two main categories: (1) resolution-invariant learning and (2) image face restoration. Resolution-invariant learning methods try to develop a single feature space for both LR and high-resolution (HR) face images. These methods require a large number of LR and HR image pairs from the same subject for training, which are usually not available. As an alternative approach, image face restoration uses super-resolution and deblurring techniques to recovery a high-quality face image from its degraded low-quality counterpart. An additional advantage of these kind of methods is that they can be used for face detection, face recognition and many other vision tasks [25].

Among existing algorithms for face image restoration, traditional methods predefine the degradation type before training, leading to poor generalization ability. Thus, in the last years, researchers pay more attention to the so-called blind face restoration (BFR) problem [32] taking into account that, mostly we cannot know the degradation type when a real-world image is captured. In this paper, we aim at studying the effectiveness of BFR to enhance the face recognition performance at LR face images. Specifically, we selected three state-of-the-art blind face restoration methods that have demonstrated excellent and remarkable abilities to generate realistic and faithful HR images. In addition, we

choose three accurate face recognition methods to evaluate the impact of using the generated HR images on their performance. Both, synthesized and native low-resolution benchmarks are used for the experimental evaluation.

The main contributions of this work are as follows:

- We investigate the effectiveness of three state-of-the-art blind face restoration methods to improve the performance of three face recognition models at low-resolution images.
- We provide an extensive experimental evaluation by testing different combining methods of BFR and face recognition models on synthetic and native LR face datasets.

The remainder of this paper is organized as follows. Section 2 takes an overview of low-resolution face recognition and deep face restoration methods. In Section 3, we present the experimental setup used for evaluating blind face restoration methods on low-resolution face recognition scenarios. Section 4 presents the experimental evaluation carried out and Section 5 exposes some discussion regarding the results obtained on this study. Finally, Section 6 summarizes the conclusions of this work.

## 2 Related Work

There is not a unique accepted criterion for classifying a face image as low-resolution. Most of the authors consider that a face image with bounding box smaller than  $32 \times 32$  pixels, represents a great challenge for both humans and systems, but others accept a minimal face resolution of  $16 \times 16$  pixels [14]. In any case, the minimum resolution required should ensure the performance of the face recognition method and thus, depends on the algorithm at hand.

Several methods have been proposed to deal with LR face images. They can be divided into two categories: (1) resolution-invariant learning [18, 31] and (2) image face restoration [25, 32]. The first category aims at extracting resolution-invariant features and learning a common space from both LR probe images and HR gallery images. The second category consist of enhancing the quality of the LR face images as an input of the face recognition system. This is the most intuitive way to deal with LR and quality degradation of the face images and is the focus of this paper.

### 2.1 Deep Face Restoration

According to degradation types, the face restoration can be divided into different tasks such as face deblurring, face super-resolution, face artifact removal and face denoising. Due to its great advances, deep learning-based methods have been adopted for face restoration tasks, including the development of basic Convolution Neural Networks (CNN) or Generative Adversarial Networks (GAN).

Face deblurring aims to recover a latent sharp face image from a blurred face image caused by various factors such as camera shake or object motion.

Most of existing face deblurring methods take advantage of various facial priors including face landmarks [6], 3D facial model [20], semantic clues [29] and deep features that contains both the geometric and texture information [13]. Face super-resolution, also known as face hallucination, focuses on enhancing quality and resolution of low-resolution facial images. Current deep learning approaches [28, 12] have investigated the use of GANs, wavelet coefficients prediction into deep networks and facial prior knowledge to super-resolve low-resolution face images. Face artifact removal bases on recovering high-quality face images from the low-quality face images that contain artifacts caused by lossy compression in the process of image storage and transmission. For example, in real-world applications, lossy compression techniques (e.g., JPEG, Webp, and HEVC-MSP) are widely adopted for saving storage space and bandwidth. which lead to information loss and introduces undesired artifacts for recorded face images. Most of existing methods [16] aim at alleviating these undesirable artifacts such as blockiness, ringing, and banding caused by quantization and approximation in the compression process. Face denoising refers to removing the noise contained in the face image. Some CNN models [30, 1] have improved denoising performance, due to their modeling capacity, network training, and design. However, the performance of current learning models is limited and tailored for a specific level of noise.

The main problem of these face restoration methods designed for one specific and known degradation is that they perform with limited generalization in real-world LR scenarios, where face images are often degraded owing to more than one factor such as compression, blur, and noise. That’s why, in recent years, blind face restoration (BFR) [32], which recovers high-quality (HQ) face images from low-quality (LQ) inputs with unknown degradations, has become more practical and attracted increasing attention. Existing works typically exploit face-specific priors, including geometric priors [2], reference priors [15], and generative priors [26, 27, 33]. In recent years, those methods based on generative priors in face restoration have shown superior performance.

On the other hand, although the existing face restoration methods have demonstrated excellent and remarkable abilities to generate realistic and faithful images, their effectiveness for increasing the face recognition rates is still challenge. Many face restoration methods generate visually pleasant face images, but face recognition accuracy is lower than expected with their generated images because they produce new details and the original structure of the face images are not well preserved. In this sense, some works have proposed specific approaches to deal with this issue. For example, a super-resolution-based approach was presented in [23] to overcome the LR image constraints and improve the accuracy of a face recognition system based on pioneering descriptor-based techniques (Eigenface and BRISK). Evaluations on synthesized images from ORL, Caltech, and Chokepoints datasets showed an increase in recognition rates, where face images did not contain pose expressions and scale variations. A GAN was used in [22] to increase the image resolution at the feature level and empower face recognition in low-resolution images. The introduced network considers image edges

and reconstructs high-frequency details to preserve the face structure. Experiments with recent face recognition methods on FERET dataset indicate that, by using super-resolving facial images generated with the proposed method improve the accuracy of these face recognition methods. Most of these works have evaluated the effectiveness of their proposals on non-native low-resolution datasets. Usually, artificially down-sampled images from HR images are used for testing. It is still unclear how effective recent BFR methods are for enhancing the face recognition performance on native low-resolution face images.

### 3 Experimental Setup

In this section, we introduce the blind face restoration and face recognition methods used for the benchmark evaluations. In addition, we describe the datasets employed for testing as well as some implementation details of selected methods.

#### 3.1 Blind Face Restoration

For obtaining HQ face images from LQ inputs, we selected three state-of-the-art BFR methods including GPEN [27], GFP-GAN [26] and CodeFormer [33].

**GPEN** [27]: is a GAN prior embedded network learned to generate high quality face images from severely degraded ones. First, a GAN is pretrained for high-quality face image generation and embedded it into a U-shaped DNN as a prior decoder. Then, the GAN prior embedded DNN is fine-tuned on synthesized LQ-HQ face image pairs. The GAN blocks are designed to ensure that the latent code and noise input to the GAN can be respectively generated from the deep and shallow features of the DNN, controlling the global face structure, local face details and background of the reconstructed image.

**GFP-GAN** [26]: is framework that leverages rich and diverse priors encapsulated in a pretrained face GAN for blind face restoration. Specifically, GFP-GAN consists of a degradation removal module (U-Net) and a pre-trained face GAN as facial prior, which are connected by a direct latent code mapping and several Channel-Split Spatial Feature Transform (CS-SFT) layers in a coarse-to-fine manner. During training, intermediate restoration losses to remove complex degradation, a facial component loss with discriminators to enhance facial details, and a identity preserving loss to retain face identity, are introduced.

**CodeFormer** [33]: is a Transformer-based prediction network that casts blind face restoration as a code prediction task, showing superior robustness to degradation as well as rich expressiveness. Specifically, The codebook is learned by self-reconstruction of HQ faces using a vector-quantized autoencoder, which along with decoder stores the rich HQ details for face restoration. Taking the LQ features as input, the Transformer module predicts the code token sequence which is treated as the discrete representation of the face images in the codebook space. In addition, a controllable feature transformation module is introduced to enhance the adaptiveness for different degradations, that allows a flexible trade-off between restoration quality and fidelity.

### 3.2 Face recognition

To assess the accuracy of face recognition on the HQ generated by the selected BFR methods, we choose three recent face recognition models such as MobileFaceNet [3], ShuffleFaceNet [19] and R100-ArcFace [7].

**MobileFaceNet** [3]: is an extremely efficient CNN model tailored for high-accuracy real-time face verification on mobile and embedded devices. It uses the residual bottlenecks proposed in MobileNetV2 [21] as their main building blocks, but with expansion factors much smaller. Moreover, the authors replace the Global Average Pooling layer for a Global Depth-wise Convolution (GDC) layer to obtain a more discriminative face representation, and use the Parametric Rectified Linear Unit (PReLU) as non-linear activation function due to its accuracy improvement over the Rectified Linear Unit (ReLU) function for face recognition.

**ShuffleFaceNet** [19]: is an efficient and lightweight architecture built upon ShuffleNetv2 [17] for the face recognition domain. It introduces several modifications for improving accuracy demonstrating its feasibility in real-time applications or computationally limited platforms. Similar to MobileFaceNet, ShuffleFaceNet adopts PReLU activation function as non-linearity and uses a linear  $1 \times 1$  convolution layer following a GDC layer as the feature output.

**R100-ArcFace** [7]: is a face recognition model that employs the widely used ResNet100 CNN architecture [10] as embedding network. In order to enhance the discriminative power of this model and to stabilize the training process, an Additive Angular Margin Loss (ArcFace) function was proposed by the authors.

### 3.3 Databases

For the experimental evaluation, we use four LR face databases. Although our main goal is to assess the performance on native LR datasets, we include experiments on artificially synthesized LR images from the LFW database.

**Synthesized LFW**: Labeled Face in the Wild (LFW) [11] is a hR face verification benchmark that contains 13,233 web-collected images from 5,749 different identities, with large variations in pose, expression and illuminations. Given LFW images, we synthesize the corresponding low-resolution images employing down-sampling operation to artificially simulate the degradation process. Specifically, we use bicubic interpolation to generate LR images at five different

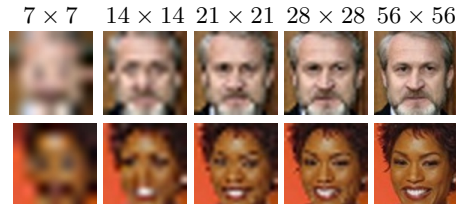


Fig. 1: Examples of synthetic LR face images from LFW database.

resolutions:  $7 \times 7$ ,  $14 \times 14$ ,  $21 \times 21$ ,  $28 \times 28$  and  $56 \times 56$ , as is illustrated in Fig. 1. The evaluation protocol provides 6,000 face pairs, divided into ten subsets, each having 300 positive pairs and 300 negative pairs. For benchmarking, we implement two experimental settings. The first one, LR-to-HR pairs matching, is composed of pairs of low-resolution and high-resolution facial images. The second one, LR-to-LR pairs matching, consists of low-resolution pair images.

**SCface** [8]: is designed for testing face recognition algorithms in real-world surveillance scenarios. In order to achieve a realistic setup, 4,160 images of 130 subjects were acquired in uncontrolled indoor environment using commercially available surveillance cameras of varying quality. For each subject, there are 15 images in the visible spectrum taken at three different distances (d): d1 (4.2m), d2 (2.6m) and d3 (1.0m), by five surveillance cameras (5 images at each distance), and one high quality frontal mugshot image taken in controlled indoor illumination conditions environment by a digital camera. Fig. 2 shows the LR images corresponding to one subject from the SCface database. For experimental settings, HR frontal mugshot images are employed as gallery, while images taken by surveillance cameras at distance 1, 2 and 3 are used as probes.

**TinyFace** [4]: is a large-scale face recognition benchmark containing native unconstrained low-resolution face images. It consists of 169,403 native LR face images (average  $20 \times 16$  pixels) from 5,139 labelled facial identities designed for 1:N recognition test. All the images are collected from public web data across a large variety of imaging scenarios, captured under uncontrolled viewing conditions in pose, illumination, occlusion and background. Fig. 3(a) shows some examples of LR face images of the TinyFace dataset.

**QMUL-SurvFace** [5]: is a very challenging dataset, that contains 463,507 real-world native surveillance face images of 15,573 unique subjects. Moreover, the LR face images present uncontrolled appearance variations in pose, illumination, motion blur, occlusion and background clutter (See Fig. 3(b)). The face spatial resolution ranges from  $6/5$  to  $124/106$  pixels in height/width, and the average is  $24/20$  pixels. For the benchmark evaluation, we follow the verification protocol of QMUL-SurvFace.

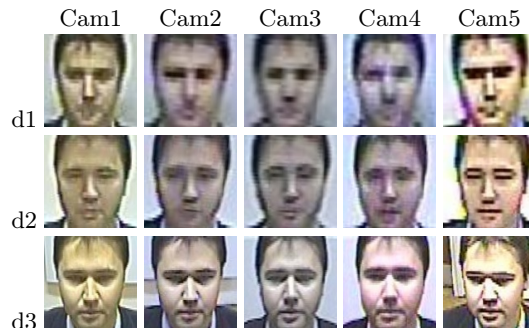


Fig. 2: LR images from one subject of SCface database.



Fig. 3: Examples of LR face images from a) TinyFace database and b) QMUL-Surv Face database.

### 3.4 Implementation Details

For blind face restoration, we used the released codes and models by the authors for GPEN<sup>1</sup>, GFP-GAN<sup>2</sup> and CodeFormer<sup>3</sup> methods. All models were trained on synthesized LR databases. In the case of face recognition models, we use pre-trained models on the cleaned MS1M dataset [9] that includes 5.1 million photos from 93K face identities. The faces are cropped and resized to  $112 \times 112$ , and each pixel (ranged between  $[0; 255]$ ) in RGB images is normalised by subtracting 127.5 then divided by 128. All experiments in this paper are implemented on PyTorch.

## 4 Experiments and Discussion

In this section, we present and discuss the performance achieved by the selected face recognition methods: ShuffleFaceNet, MobileFaceNet and R100-ArcFace on both synthetic and native LQ face images. To examine the impact of blind face restoration, we compare the performance of face recognition models on the generated HQ images with those obtained on the original LQ images.

### 4.1 Results on Synthesized LR Images

Following the standard protocol of unrestricted with labeled outside data, Table 1 and Table 2 show the verification accuracy on the synthesized LFW database for LR-to-HR and LR-to-LR pairs matching at different resolutions, respectively. As it can be seen, using blind face restoration degrades the performance of face recognition models. The best results are achieved by MobileFaceNet without employing any BFR method previously. In general, the verification accuracy for matching LR-to-HR pairs are lower than those obtained by matching LR-to-LR pairs, especially for very low-resolution sizes like  $7 \times 7$ . Although the usage of blind face restoration methods do not improve the performance of face recognition models, among the them, CodeFormer obtains higher accuracy values in the case of LR-to-HR pairs matching and, GFP-GAN for LR-to-LR pairs matching.

<sup>1</sup> <https://github.com/yangxy/GPEN>

<sup>2</sup> <https://github.com/TencentARC/GFPGAN/>

<sup>3</sup> <https://shangchenzhou.com/projects/CodeFormer/>



Table 1: Verification accuracy (%) on synthesized LFW database for LR-to-HR pairs matching at different resolution sizes.

Method	7×7	14×14	21×21	28×28	56×56
ShuffleFaceNet	56.7	<b>82.1</b>	<b>98.5</b>	<b>99.4</b>	<b>99.4</b>
GPEN + ShuffleFaceNet	56.1	74.3	89.8	96.5	99.3
GFP-GAN + ShuffleFaceNet	59.4	78.7	91.2	97.0	99.3
CodeFormer + ShuffleFaceNet	<b>59.6</b>	81.5	92.9	97.6	99.3
MobileFaceNet	<b>62.1</b>	<b>85.3</b>	<b>94.8</b>	<b>98.5</b>	99.6
GPEN + MobileFaceNet	55.0	74.1	88.7	96.5	99.5
GFP-GAN + MobileFaceNet	57.6	75.7	89.9	96.2	99.2
CodeFormer + MobileFaceNet	61.0	81.6	92.9	97.9	99.5
R100-ArcFace	<b>58.2</b>	<b>87.5</b>	<b>97.6</b>	<b>99.4</b>	<b>99.6</b>
GPEN + R100-ArcFace	54.3	71.4	89.3	96.6	99.5
GFP-GAN + R100-ArcFace	55.4	71.4	88.1	96.5	99.4
CodeFormer + R100-ArcFace	57.2	80.3	92.9	97.7	99.5

Table 2: Verification accuracy (%) on synthesized LFW database for LR-to-LR pairs matching at different resolution sizes.

Method	7×7	14×14	21×21	28×28	56×56
ShuffleFaceNet	<b>63.1</b>	<b>77.8</b>	<b>89.9</b>	<b>96.7</b>	<b>99.4</b>
GPEN + ShuffleFaceNet	62.3	73.6	85.9	95.1	99.2
GFP-GAN + ShuffleFaceNet	63.0	73.5	86.3	94.7	99.2
CodeFormer + ShuffleFaceNet	61.6	76.6	88.9	96.3	99.2
MobileFaceNet	<b>67.3</b>	<b>79.8</b>	<b>91.5</b>	<b>97.2</b>	<b>99.6</b>
GPEN + MobileFaceNet	64.5	74.1	85.8	95.1	<b>99.6</b>
GFP-GAN + MobileFaceNet	65.7	74.2	85.8	94.6	99.5
CodeFormer + MobileFaceNet	62.9	78.3	89.1	96.2	99.5
R100-ArcFace	<b>65.8</b>	<b>82.5</b>	<b>94.8</b>	<b>98.6</b>	<b>99.7</b>
GPEN + R100-ArcFace	63.0	73.6	87.1	96.0	<b>99.7</b>
GFP-GAN + R100-ArcFace	62.3	73.1	85.1	94.6	99.6
CodeFormer + R100-ArcFace	58.3	77.2	89.7	96.7	99.6

## 4.2 Results on Native LR Images

Tables 3, 4 and 5 present the performance on native low-resolution images from the SCface, TinyFace and QMUL-SurvFace databases, respectively. Similar to the results obtained on the synthesized LFW database, using BFR methods also decrease the performance of deep face recognition models on native surveillance LR images. We can observe that, in this scenario, the accuracy drops are greater, which reflect that BFR methods trained on artificially synthesized low-resolution images do not generalize well in front the true challenges of native surveillance images. These methods suffer the lack of native surveillance pairs of low-quality and high-quality facial images which are necessary for model training. Thus, when the training and test data distributions are very different, BFR becomes extremely challenging due to an extra need for domain adaptation. Moreover,

in the surveillance scenarios, the captured images have additional variations such as occlusions, pose and illumination changes, that add complexity to the generation process. As result, several artifacts and noise are introduced in the generated images (see Figure 4), which impact negatively in the performance of face recognition models.

From the quantitative results presented on Tables 3, 4 and 5, and the qualitative results shown in Figure 4, we find that GPEN performs better than GFP-GAN and CodeFormer, as well as the images recovered by GPEN are more satisfying to human visual perception. Nonetheless, the best performance is achieved by MobileFaceNet model without using any blind face restoration method.

Table 3: Recognition Rates (%) at Rank-1 on SCface database.

Method	d1 (4.2m)	d2 (2.6m)	d3 (1.0m)
ShuffleFaceNet	<b>48.8</b>	<b>94.0</b>	<b>99.5</b>
GPEN + ShuffleFaceNet	32.8	87.2	99.4
GFP-GAN + ShuffleFaceNet	28.0	86.9	98.8
CodeFormer + ShuffleFaceNet	32.6	87.8	98.8
MobileFaceNet	<b>60.9</b>	<b>97.1</b>	99.7
GPEN + MobileFaceNet	35.1	89.5	<b>99.8</b>
GFP-GAN + MobileFaceNet	26.3	86.8	99.2
CodeFormer + MobileFaceNet	32.5	88.2	99.2
R100-ArcFace	<b>67.7</b>	<b>99.2</b>	<b>99.8</b>
GPEN + R100-ArcFace	31.5	87.2	99.5
GFP-GAN + R100-ArcFace	24.5	81.5	99.2
CodeFormer + R100-ArcFace	30.2	84.2	99.1

Table 4: Face identification (%) results on TinyFace database.

Method	Rank-1	Rank-10	Rank-20	Rank-50	mAP
ShuffleFaceNet	<b>22.3</b>	<b>35.6</b>	<b>40.9</b>	<b>49.5</b>	<b>16.1</b>
GPEN + ShuffleFaceNet	15.4	26.1	31.1	39.3	11.1
GFP-GAN + ShuffleFaceNet	10.9	20.0	23.9	31.1	8.2
CodeFormer + ShuffleFaceNet	12.1	23.0	28.4	36.3	8.9
MobileFaceNet	<b>26.2</b>	<b>39.2</b>	<b>44.0</b>	<b>51.7</b>	<b>18.9</b>
GPEN + MobileFaceNet	16.9	27.4	31.9	40.4	12.1
GFP-GAN + MobileFaceNet	10.9	20.6	24.8	32.8	8.1
CodeFormer + MobileFaceNet	14.3	25.6	30.4	40.4	10.4
R100-ArcFace	<b>25.9</b>	<b>37.5</b>	<b>42.2</b>	<b>50.5</b>	<b>18.8</b>
GPEN + R100-ArcFace	14.7	25.8	30.5	37.7	10.6
GFP-GAN + R100-ArcFace	8.6	15.8	20.1	26.9	6.7
CodeFormer + R100-ArcFace	10.7	20.3	24.9	32.4	7.7

Table 5: Face verification (%) results on QMUL-SurvFace.

Method	TAR@FAR				AUC	Mean Accuracy
	30%	10%	1%	0.1%		
ShuffleFaceNet	<b>57.0</b>	<b>33.1</b>	<b>10.6</b>	<b>4.6</b>	<b>68.7</b>	<b>63.6</b>
GPEN + ShuffleFaceNet	51.1	26.3	6.2	1.8	64.3	60.6
GFP-GAN + ShuffleFaceNet	47.9	23.1	5.2	1.1	62.2	58.7
CodeFormer + ShuffleFaceNet	51.1	24.7	6.2	1.4	64.6	60.3
MobileFaceNet	<b>58.8</b>	<b>34.3</b>	<b>12.1</b>	<b>4.3</b>	<b>70.2</b>	<b>64.8</b>
GPEN + MobileFaceNet	51.3	25.9	6.5	1.3	64.5	60.4
GFP-GAN + MobileFaceNet	47.5	24.6	5.3	0.8	62.2	58.9
CodeFormer + MobileFaceNet	52.1	26.5	7.0	3.1	65.0	60.9
R100-ArcFace	<b>54.5</b>	<b>29.0</b>	<b>8.7</b>	<b>2.4</b>	<b>66.7</b>	<b>61.9</b>
GPEN + R100-ArcFace	49.1	25.1	6.7	1.5	63.3	59.2
GFP-GAN + R100-ArcFace	45.5	20.1	4.8	1.1	60.8	57.7
CodeFormer + R100-ArcFace	44.9	20.1	3.7	0.9	60.9	57.3

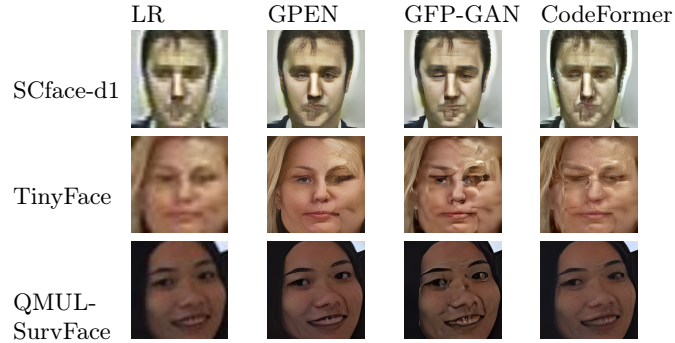


Fig. 4: Visual results of BFR methods: GPEN, GFP-GAN and CodeFormer on the SCface, TinyFace and QMUL-SurvFace datasets.

## 5 Conclusion

In this paper, we have presented a comprehensive evaluation of the impact of blind face restoration on the face recognition performance at LR imagery. Specifically, we use three state-of-the-art blind face restoration methods: GPEN, GFP-GAN and CodeFormer, to generate high-quality face images from synthesized and native low-resolution images. Then, we assess the performance of three well-established deep face recognition models: MobileFaceNet, ShuffleFaceNet and ResNet100-ArcFace on the generated high-quality images. Experimental results, on both synthetic and native LR images, show that state-of-the-art BFR algorithms are not effective to improve the accuracy of face recognition methods in low-resolution face images. This is because most of existing BFR methods have no access to native surveillance pairs of low-quality and high-quality images re-

quired in the training process. Mostly, face recognition and blind face restoration researches advance independently, which does not offer any benefit.

## References

1. Anwar, S., Barnes, N.: Real image denoising with feature attention. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 3155–3164 (2019)
2. Chen, C., Li, X., Yang, L., Lin, X., Zhang, L., Wong, K.Y.K.: Progressive semantic-aware style transformation for blind face restoration. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 11896–11905 (2021)
3. Chen, S., Liu, Y., Gao, X., Han, Z.: Mobilefacenet: Efficient cnns for accurate real-time face verification on mobile devices. In: Biometric Recognition. pp. 428–438 (2018)
4. Cheng, Z., Zhu, X., Gong, S.: Low-resolution face recognition. In: Asian Conf. on Computer Vision. pp. 605–621. Springer (2018)
5. Cheng, Z., Zhu, X., Gong, S.: Surveillance face recognition challenge. arXiv preprint arXiv:1804.09691 (2018)
6. Chrysos, G.G., Zafeiriou, S.: Deep face deblurring. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 69–78 (2017)
7. Deng, J., Guo, J., Xue, N., Zafeiriou, S.: Arcface: Additive angular margin loss for deep face recognition. In: IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). pp. 4690–4699 (2019)
8. Grgic, M., Delac, K., Grgic, S.: Scface—surveillance cameras face database. *Multimedia tools and applications* **51**(3), 863–879 (2011)
9. Guo, Y., Zhang, L., Hu, Y., He, X., Gao, J.: Ms-celeb-1m: A dataset and benchmark for large-scale face recognition. In: Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part III 14. pp. 87–102. Springer (2016)
10. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). pp. 770–778 (2016)
11. Huang, G.B., Ramesh, M., Berg, T., Learned-miller, E.: Labeled faces in the wild: A database for studying face recognition in unconstrained environments (2007)
12. Jiang, J., Wang, C., Liu, X., Ma, J.: Deep learning-based face super-resolution: A survey. *ACM Computing Surveys (CSUR)* **55**(1), 1–36 (2021)
13. Jung, S.H., Lee, T.B., Heo, Y.S.: Deep feature prior guided face deblurring. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 3531–3540 (2022)
14. Li, P., Prieto, L., Mery, D., Flynn, P.: Face recognition in low quality images: a survey. arXiv preprint arXiv:1805.11519 (2018)
15. Li, X., Chen, C., Zhou, S., Lin, X., Zuo, W., Zhang, L.: Blind face restoration via deep multi-scale component dictionaries. In: Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IX 16. pp. 399–415. Springer (2020)
16. Liu, J., Liu, D., Yang, W., Xia, S., Zhang, X., Dai, Y.: A comprehensive benchmark for single image compression artifact reduction. *IEEE Transactions on image processing* **29**, 7845–7860 (2020)

17. Ma, N., Zhang, X., Zheng, H.T., Sun, J.: Shufflenet v2: Practical guidelines for efficient cnn architecture design. In: Proceedings of the European conference on computer vision (ECCV). pp. 116–131 (2018)
18. Makwana, P., Kumar Singh, S., Ram Dubey, S.: Resolution invariant face recognition. In: Computer Vision and Machine Intelligence: Proceedings of CVMI 2022, pp. 733–745. Springer (2023)
19. Martínez-Díaz, Y., Luevano, L.S., Méndez-Vázquez, H., Nicolás-Díaz, M., Chang, L., González-Mendoza, M.: Shufflefacenet: A lightweight face architecture for efficient and highly-accurate face recognition. In: IEEE Int. Conf. on Computer Vision Workshops (Oct 2019)
20. Ren, W., Yang, J., Deng, S., Wipf, D., Cao, X., Tong, X.: Face video deblurring using 3d facial priors. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 9388–9397 (2019)
21. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C.: Mobilenetv2: Inverted residuals and linear bottlenecks. In: IEEE Conf. on Computer Vision and Pattern Recognition. pp. 4510–4520 (2018)
22. Shahbakhsh, M.B., Hassanpour, H.: Empowering face recognition methods using a gan-based single image super-resolution network. *International Journal of Engineering* **35**(10), 1858–1866 (2022)
23. Singh, N., Rathore, S.S., Kumar, S.: Towards a super-resolution based approach for improved face recognition in low resolution environment. *Multimedia Tools and Applications* **81**(27), 38887–38919 (2022)
24. Wang, M., Deng, W.: Deep face recognition: A survey. *Neurocomputing* **429**, 215–244 (2021)
25. Wang, T., Zhang, K., Chen, X., Luo, W., Deng, J., Lu, T., Cao, X., Liu, W., Li, H., Zafeiriou, S.: A survey of deep face restoration: Denoise, super-resolution, deblur, artifact removal. *arXiv preprint arXiv:2211.02831* (2022)
26. Wang, X., Li, Y., Zhang, H., 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. pp. 9168–9178 (2021)
27. Yang, T., Ren, P., Xie, X., Zhang, L.: Gan prior embedded network for blind face restoration in the wild. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 672–681 (2021)
28. Yang, W., Zhang, X., Tian, Y., Wang, W., Xue, J.H., Liao, Q.: Deep learning for single image super-resolution: A brief review. *IEEE Transactions on Multimedia* **21**(12), 3106–3121 (2019)
29. Yasarla, R., Perazzi, F., Patel, V.M.: Deblurring face images using uncertainty guided multi-stream semantic networks. *IEEE Transactions on Image Processing* **29**, 6251–6263 (2020)
30. Yue, Z., Yong, H., Zhao, Q., Meng, D., Zhang, L.: Variational denoising network: Toward blind noise modeling and removal. *Advances in neural information processing systems* **32** (2019)
31. Zeng, D., Chen, H., Zhao, Q.: Towards resolution invariant face recognition in uncontrolled scenarios. In: 2016 Int. Conf. on Biometrics (ICB). pp. 1–8. IEEE (2016)
32. Zhang, P., Zhang, K., Luo, W., Li, C., Wang, G.: Blind face restoration: Benchmark datasets and a baseline model. *arXiv preprint arXiv:2206.03697* (2022)
33. Zhou, S., Chan, K., Li, C., Loy, C.C.: Towards robust blind face restoration with codebook lookup transformer. *Advances in Neural Information Processing Systems* **35**, 30599–30611 (2022)