

# Supplementary Material

As a supplementary asset to the primary survey, we offer extra visual representations and informative material related to the discussed subjects. This enhances readers' understanding of the main concepts. First, we display a more comprehensive visualization result about the blind methods and give experimental comparisons of the non-blind/blind methods on the deblurring task (not given because there are too few methods in the face denoising task). Next, we showcase crucial data and visualizations of outcomes related to different methodologies used in various collaborative tasks. Furthermore, we include examples of the various technical frameworks described in the main text, including GAN architectures, Prior bootstrapping structures, and more. Lastly, we provide comprehensive information about commonly used facial restoration datasets to improve readers' understanding of the dataset tables presented in the main text.

## VISUALIZATION OF BLIND TASKS.

Owing to spatial constraints, only a subset of the visualization results is presented in the main text. Thus, we provide additional visualization outputs of the blind methods using the CelebA-HQ test sets in Fig. 3, along with visualization outputs of the blind methods applied to a broader range of real-world datasets in Fig. 4. It can be seen that the face restored by the method based on generating prior has higher clarity.

## BLIND/NON-BLIND FACE DEBLUR.

In addition to face super-resolution and blind face restoration described in the main text, some blind/non-blind methods also focus on the face deblurring task, but mainly pre-2020 methods. Given that the face deblurring task has rarely been explored in face restoration in recent years, we place it within the supplemental material to compare simple discussions. In addition, since there are too few methods for face denoising, this task is not described here.

Among them, Tab. 1 and Fig. 1 give each method's quantitative and qualitative results for blind face deblurring, respectively. It can be seen that the method proposed by Yasarla *et al.* performs the best both in terms of quantitative results and the quality of the recovered face image. Tab. 2 and Fig. 2 give each method's quantitative and qualitative results for blind face deblurring, respectively. FSGN [51] stands out among these methods in that it substantially outperforms the other methods in terms of quantitative results, and the eye region of the recovered face image is closest to the ground truth. For face deblurring, blind face deblurring has been studied more in deep learning. However, with the rise of the blind face restoration task, the blind face deblurring task has been gradually replaced by the blind face restoration task.

---

Author's address:

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© Association for Computing Machinery.

0360-0300/2024/2-ART \$15.00

<https://doi.org/>

Table 1. PSNR and SSIM comparison of state-of-the-art methods on **blind face deblur** task.

Deblurring Method	Helen [27]		CelebA [38]	
	PSNR	SSIM	PSNR	SSIM
Xu <i>et al.</i> [57]	20.11	0.711	18.93	0.685
Zhong <i>et al.</i> [80]	16.41	0.614	17.26	0.695
Pan <i>et al.</i> [44]	20.93	0.727	18.59	0.677
Mah <i>et al.</i> [42]	24.12	0.823	22.43	0.832
Shen <i>et al.</i> [49] (w/GAN)	25.58	0.861	24.34	0.860
Shen <i>et al.</i> [49]	25.99	0.871	25.05	0.879
Kupyn <i>et al.</i> [26]	26.45	0.880	25.42	0.884
Yasarla <i>et al.</i> [64]	<b>27.75</b>	<b>0.897</b>	<b>26.62</b>	<b>0.908</b>

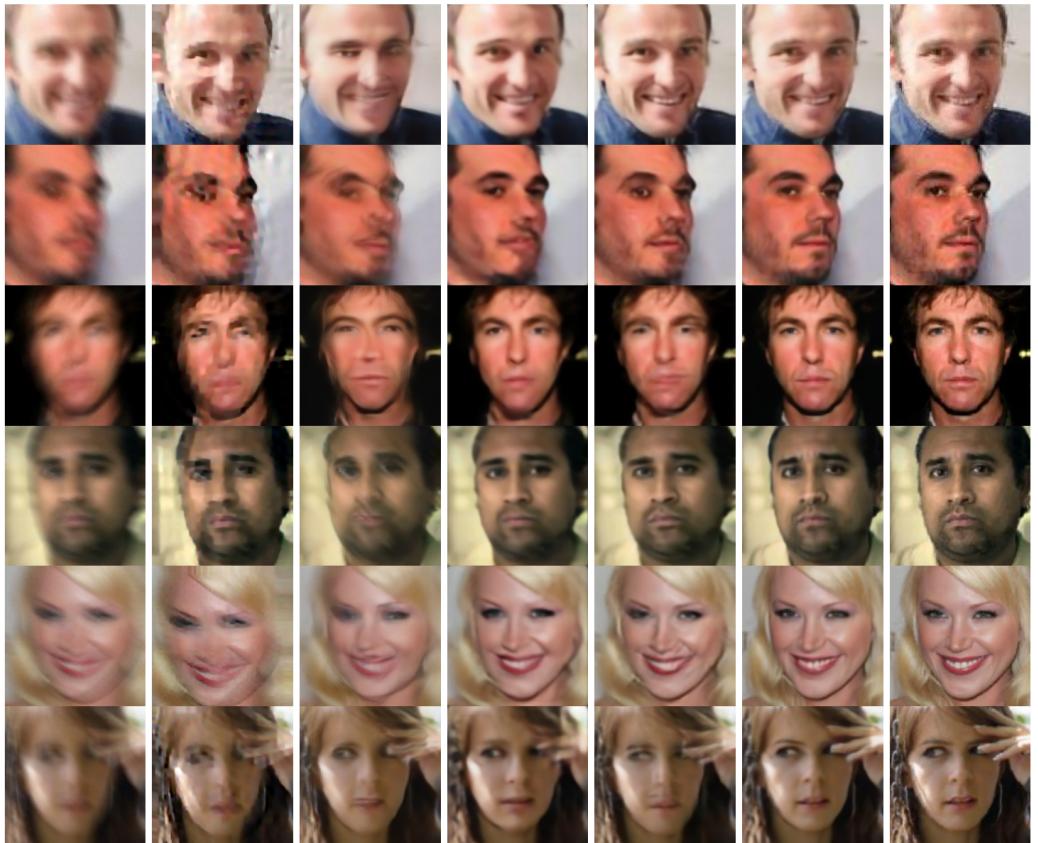
Fig. 1. Qualitative comparisons of some state-of-the-art **blind face deblur** methods. From left to right, we demonstrate LQ input, and results from Xu *et al.* [57], Zhong *et al.* [80], Shen *et al.* [49], Kupyn *et al.* [26], Yasarla *et al.* [64] and ground-truth face image, respectively.

Table 2. The evaluation of the Multi-PIE dataset with state-of-the-art methods on **non-blind face deblur** task.

	PSNR	SSIM	Similarity
	SR / Deblur	SR / Deblur	SR / Deblur
Bicubic	32.43 / 23.58	0.89 / 0.81	0.92 / 0.87
SFH [59]	31.60 / 23.65	0.86 / 0.79	0.91 / 0.86
DFE [44]	31.53 / 25.26	0.87 / 0.85	0.94 / 0.94
SRCNN [9]	33.89 / 23.73	0.90 / 0.82	0.94 / 0.90
SRCSC [13]	33.95 / 23.82	0.90 / 0.82	0.95 / 0.91
SRResNet [28]	34.10 / 23.95	0.90 / 0.81	0.96 / 0.92
VDSR [24]	34.62 / 24.33	0.91 / 0.81	0.97 / 0.92
RBF [58]	30.05 / 24.73	0.86 / 0.77	0.93 / 0.94
FSGN [51]	<b>34.93 / 25.75</b>	<b>0.92 / 0.86</b>	<b>0.98 / 0.96</b>



Fig. 2. Qualitative comparisons of some state-of-the-art **non-blind face deblur methods**. From left to right, we demonstrate LQ input, and results from SFH [59], SRCNN [9], SRCSC [13], DFE [44], RBF [58], FSGN [51] and ground-truth face image, respectively.

## JOINT TASKS.

In this section, we give examples of visual comparisons of joint face restoration tasks, which are summarized in the main text. Fig. 6 provides a visual representation of the outcomes yielded by certain approaches concerning the Joint Face Completion task. Moving to Fig. 7, it showcases the outcomes achieved by various methods in the context of the Joint Face Frontalization task. Fig. 8 offers visual insights into the results obtained by some techniques for the joint face alignment task. Similarly, Fig. 9 illustrates the results obtained by some methodologies for the joint face recognition task. Furthermore, we introduce TABLE 3 and TABLE 4 as supplementary resources, aiding readers in better comprehending the assistance rendered by these methods in enhancing face recognition rates. Shifting to Fig. 10, it presents visual depictions of the outputs generated by several techniques targeting the Joint Face Illumination Compensation. Lastly, Fig. 12 offers visual results of some methods concerning the Joint Face Fairness.

## TECHNICAL FRAMEWORKS

### Prior Guide Approach.

In this section, we first delve into the examination of bootstrapping methods aimed at enhancing the validity and robustness of prior knowledge. Inspired by [20], we categorize priors bootstrapping can be categorized into the following five categories: pre-prior, in-prior, parallel-prior, input-prior,

and post-prior. Each category represents different strategies for incorporating face prior to the restoration, potentially leading to varying outcomes.

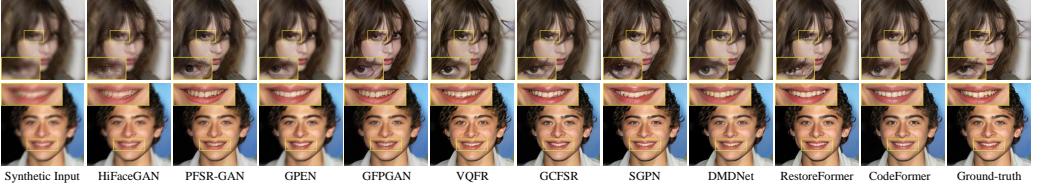


Fig. 3. Visual comparison of different blind methods on the CelebA-HQ test set for blind face restoration.

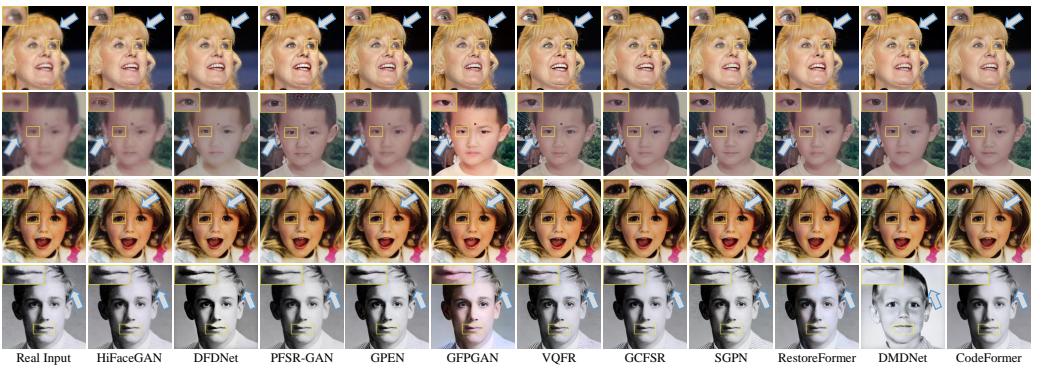


Fig. 4. Qualitative comparison of restoration for the real test sets, including LFW-Test (first row), WebPhoto (second row), Celeb-Child (third row), and Celeb-Adult (fourth row).

- **Pre-prior.** As depicted in Fig. 5 (a), pre-prior methods [5, 21, 49, 50, 52] typically involve estimating the prior of the LQ face image before the restoration process. They guide the restoration by utilizing extracted prior information as inputs to the restoration network. It enables the network to produce more accurate and contextually relevant results based on prior knowledge. However, this type of method extracts the prior directly from the LQ face image, which can limit the accuracy of the prior extraction and subsequently restrict the overall performance of the restoration.

- **In-prior.** In-prior approach [6, 39, 50, 77] is developed to address the disadvantage of inaccurate prior estimation. As depicted in Fig. 5 (b), it involves adding a restoration network before the prior estimation network. This initial restoration network is responsible for roughly recovering the LQ face image. Then, the prior information is extracted from this intermediate feature. Finally, an additional restoration network is used to complete the fine face restoration, utilizing the extracted prior information. While the in-prior approach brings performance gains in FR tasks, it also brings a significant increase in computational consumption.

- **Parallel-prior.** The above structure overlooks the correlation between prior estimation and face recovery. In the Parallel-prior approach [30, 33, 66, 81], as shown in Fig. 5 (c), the method first extracts the main features of the face using a shared feature extraction network. Then, it jointly trains both the recovery network and the prior estimation network for the two tasks. Some methods [33] in this category also feed features from the prior estimation branch into the recovery network to leverage the correlation and achieve better results. However, this type of method may fail to fully utilize the power of prior guidance.

• **Input-prior.** To address the challenge of estimating prior directly or indirectly from LQ faces, the input-prior approach [8, 35–37], as shown in Fig. 5 (d), takes both an LQ face and another HQ face image of the same person as inputs, utilizes a reference HQ face as prior to aid the restoration process. This approach leverages more bootstrapping features, which can help restore and alleviate computational consumption. However, a drawback is that it requires different face images of the same person, which increases the difficulty of training and inference.

• **Post-prior.** As shown in Fig. 5 (e), post-prior [3, 23, 32] approach estimates the prior from the recovered face images and utilizes the feedback of the estimation accuracy to regulate the recovery process. This approach mainly focuses on designing the loss function for the prior estimation. This approach has the advantage of eliminating the prior estimation process during inference, making it faster in inference than other methods. However, the indirect feedback process may lead to untimely updates and potentially result in suboptimal restoration quality.

## Frameworks.

Then, we will give some examples of the methodological structure of the framework, which are summarized in the main text. Regarding the GAN structure, Fig. 13, Fig. 14, and Fig. 15 show examples of using the general GAN structure, the Pre-trained GAN structure, and using the unpaired images GAN structure, respectively. In addition, regarding the prior guided structure, Fig. 16, Fig. 17, Fig. 18 Fig. 19 and Fig. 20 show examples of the use of the Pre-prior framework, the In-prior framework, the Parallel-prior framework, the Input-prior framework, and the Post-prior framework, respectively.

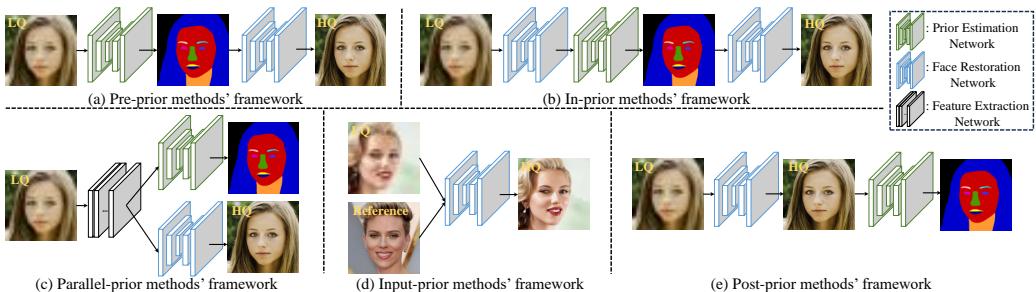


Fig. 5. Summary of prior-guided methods' architecture used for face restoration.

## DATASETS

This section provides a detailed overview of the dataset utilized for face restoration.

- **LFW [16]** contains an extensive collection of labeled face images collected from the internet, focusing on unconstrained, real-world conditions. It consists of more than 13,000 images of over 5,700 individuals, capturing a wide range of pose, lighting, expression, and occlusion variations.
- **Multi-PIE [12]** features an extensive collection of face images captured under controlled lighting and from multiple viewing angles. The dataset contains images of over 337 subjects, each captured under 15 different viewpoints and 20 different lighting conditions.
- **AFLW [25]** contains an extensive collection of facial images that are taken "in the wild", and includes images of various ethnicities, ages, and gender, making it suitable for training and evaluating algorithms in real-world scenarios.

- **SCFace** [11] is a set of 4,160 still face images containing 130 subjects, and the face images were captured in an uncontrolled indoor environment using five video surveillance cameras of different quality.
- **Helen** [27] contains 2,330 face images and is specifically designed for the task of intensive landmark annotation with landmarks covering a variety of facial features, including eyes, eyebrows, nose, mouth, and jawline.
- **300W** [48] is a widely used benchmark dataset in facial landmark detection and face alignment.
- **CASIA-WebFace** [65] consists of 494,414 high-quality, aligned face images of over 10,000 unique subjects; each subject has multiple images taken under different conditions.
- **CelebA** [38] consists of more than 200,000 celebrity images, known for their diversity in terms of gender, age and ethnicity. Each image is labeled with 40 different attribute annotations and bounding box annotations around the face.
- **Widerface** [61] contains 32,203 images containing bounding boxes of face locations and encompasses different scenes with faces of different scales and orientations.
- **IMDB-WIKI** [47] consists of over 500,000 face images collected from IMDb and Wikipedia databases, with name, age and gender. However, it may contain some bias, as these images mainly represent individuals from specific regions.
- **VGGFace** [45] contains 2.6 million images of over 2,600 people, including celebrities, public figures and ordinary people. These images vary in terms of pose, lighting conditions and facial expressions.
- **LS3D-W** [2] consists of a large number of 2D face images and corresponding 3D facial landmark annotations, with the primary purpose of facilitating the development and evaluation of face alignment and 3D face reconstruction algorithms.
- **VGGFace2** [4] consists of over 3.3 million images from over 9,000 people and is one of the most extensive publicly available face recognition datasets.
- **IJB-C** [40] consists of over 138,000 still images and 1,400 video clips covering a wide range of variations in pose, expression, lighting conditions, occlusion, and resolution.
- **FFHQ** [22] contains over 70,000 face images with a resolution of 1024×1024 pixels, covering a diverse range of individuals, including people of different ages, genders, races and facial expressions.
- **CelebAMask-HQ** [29] is an extension of the CelebA [38] and contains over 30,000 celebrity face images with a resolution of 512×512 pixels. Each of its images is annotated with pixel-level semantic segmentation masks, which provide detailed information about different facial regions and attributes.
- **EDFace-Celeb-1M** [75] is a benchmark dataset dedicated to face restoration, containing 1.7 million pairs of low-quality and high-quality face images, covering faces of different ethnicities from different regions.
- **CelebRef-HQ** [37] contains 10,555 images of 1,005 identities covering various ages, genders, ethnicities, backgrounds, poses and expressions.



Fig. 6. Qualitative comparison of some state-of-the-art **joint face completion methods** on the CelebA-HQ dataset. From left to right, we demonstrate the ground-truth image, the masked image, and the inpainting results from EC [43], RFR [31], Lafin [62], and FT-TDR [54] with the predicted mask, respectively.

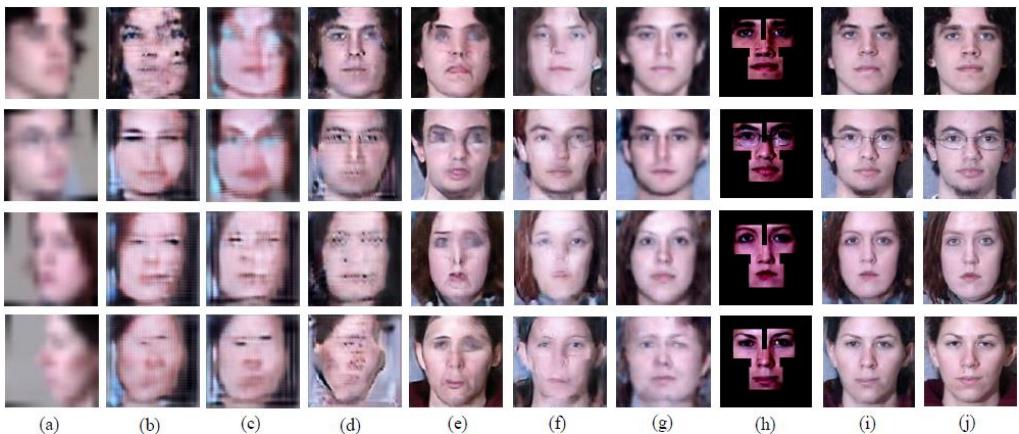


Fig. 7. Qualitative comparisons of some state-of-the-art **joint face frontalization methods** on the Multi-PIE database. Columns: (a) Unaligned LR inputs under various poses (Rows:  $+60^\circ$ ,  $+45^\circ$ ,  $-75^\circ$  and  $-90^\circ$ ). (b) Bicubic + [53] (c) [53] + [81]. (d) [53] + [18]. (e) [28] + [53]. (f) [70] + [53]. (g) [73]. (h) Fine-grained facial components. (i) VividGAN [76]. (j) Ground-truths.

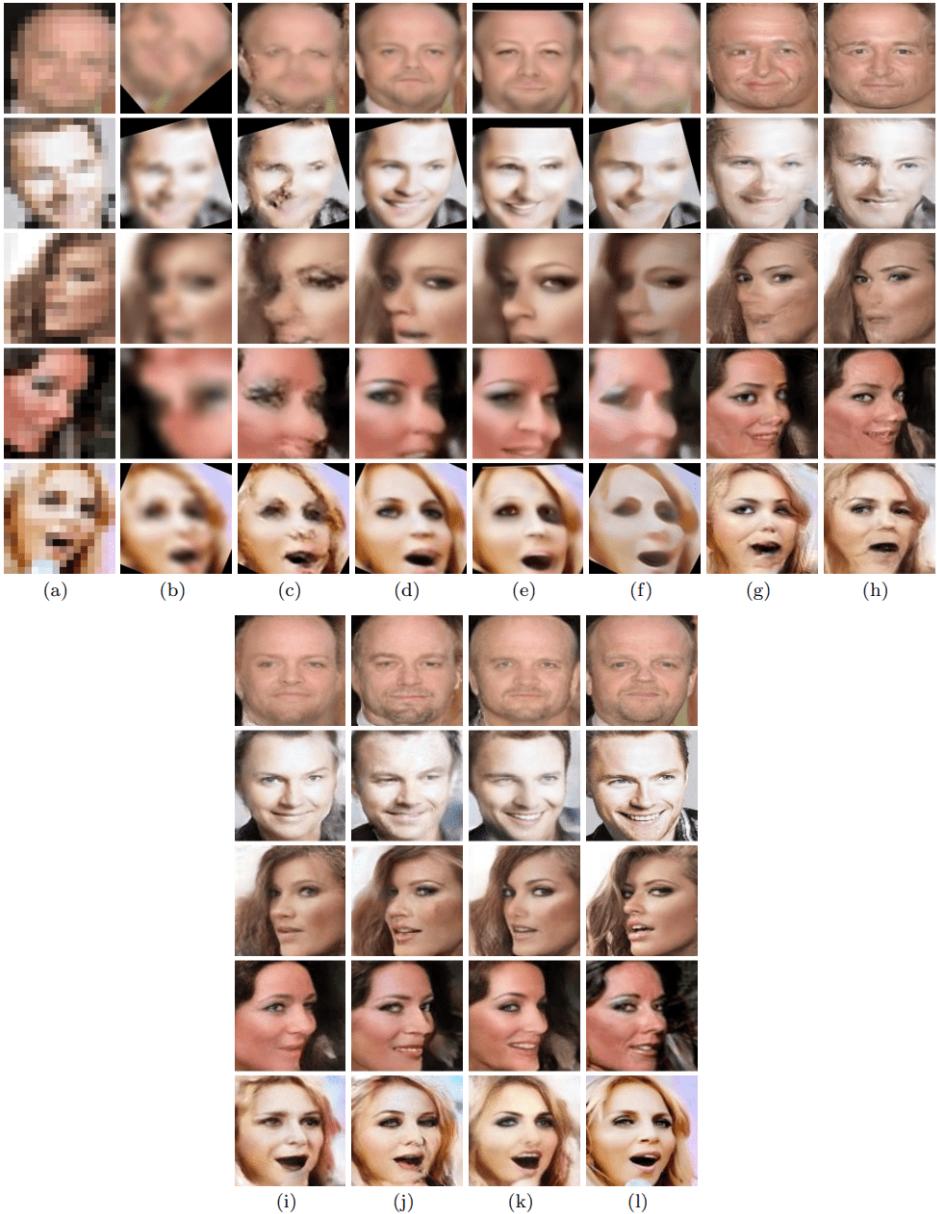


Fig. 8. Qualitative comparisons of some state-of-the-art **joint face alignment methods** on the input images of size  $16 \times 16$  pixels. The results are obtained in the scenario of first upsampling LR faces and then aligning the super-resolved faces by Bulat et al.'s method (Bulat and Tzimiropoulos, 2017). (a) Unaligned LR inputs. (b) Bicubic interpolation. (c) VDSR [24]. (d) SRGAN [28]. (e) CBN [81]. (f) FSRNet [6]. (g) TDN [70]. (h) TDAE [71]. (i) Yu et al. [68]. (j) Yu et al. [67]. (k) MTDN [72]. (l) Original HR.

Table 3. Comparison of face recognition rates evaluated by OpenFaces [1] for HR faces reconstructed by various **joint face recognition methods** on [65] by upscaling: (a) from  $8 \times 8$  to  $32 \times 32$ ; (b) from  $16 \times 16$  to  $64 \times 64$

Method	Top-1	Top-5	Top-10
HR ( $32 \times 32$ )	30.4%	51.2%	59.6%
LR ( $8 \times 8$ )	10.7%	19.5%	33.1%
Bicubic	10.8%	20.1%	34.4%
DFCG [52]	9.3%	17.7%	21.4%
UR-DGN [69]	9.9%	18.6%	22.7%
DCGAN [46]	4.6%	10.9%	16.8%
PRSR [7]	10.8%	18.8%	24.4%
SR-GAN [28]	8.8%	11.1%	19.4%
Wavelet-SRNet [17]	12.8%	20.2%	30.3%
SiGAN	15.8%	27.5%	40.4%

Method	Top-1	Top-5	Top-10
HR ( $64 \times 64$ )	36.8%	55.9%	63.8%
LR ( $16 \times 16$ )	12.4%	27.4%	37.1%
Bicubic	11.6%	27.5%	37.6%
DFCG [52]	9.6%	23.7%	34.8%
UR-DGN [69]	12.2%	29.0%	38.7%
DCGAN [46]	9.3%	24.9%	33.9%
PRSR [7]	13.3%	29.7%	40.1%
SR-GAN [28]	11.6%	23.2%	36.3%
Wavelet-SRNet [17]	12.0%	25.5%	38.8%
SiGAN	17.9%	32.9%	48.1%

Table 4. Comparison of face recognition rates evaluated by OpenFaces [1] for HR faces reconstructed by various **joint face recognition methods** on LFW [16] by upscaling: (a) from  $8 \times 8$  to  $32 \times 32$ ; (b) from  $16 \times 16$  to  $64 \times 64$

Method	Top-1	Top-5	Top-10
HR ( $32 \times 32$ )	32.2%	50.8%	56.7%
LR ( $8 \times 8$ )	9.3%	17.4%	30.9%
Bicubic	9.6%	17.7%	30.4%
DFCG [52]	9.3%	16.9%	27.5%
UR-DGN [69]	7.9%	16.8%	20.1%
DCGAN [46]	4.7%	9.9%	14.6%
PRSR [7]	10.3%	19.8%	26.1%
SR-GAN [28]	9.1%	13.3%	22.6%
Wavelet-SRNet [17]	13.1%	22.7%	32.0%
SiGAN (proposed)	14.5%	26.7%	39.2%

Method	Top-1	Top-5	Top-10
HR ( $64 \times 64$ )	35.4%	51.4%	60.1%
LR ( $16 \times 16$ )	14.8%	26.6%	35.3%
Bicubic	15.0%	26.4%	35.6%
DFCG [52]	13.2%	25.4%	34.7%
UR-DGN [69]	15.9%	30.2%	39.4%
DCGAN [46]	11.6%	24.3%	32.6%
PRSR [7]	18.3%	32.6%	45.5%
SR-GAN [28]	12.6%	26.5%	38.8%
Wavelet-SRNet [17]	15.1%	27.1%	40.2%
SiGAN (proposed)	21.5%	40.5%	50.2%



Fig. 9. Subjective visual quality comparison of some state-of-the-art **joint face recognition methods** for five faces with unknown identities selected from LFW [16] and CelebA [38]: (a) The LR face images ( $16 \times 16$ ). (b)–(i) are the reconstructed  $64 \times 64$  HR faces using (b) bicubic interpolation, (c) DFCG [52], (d) DCGAN [46], (e) UR-DGN [69], (f) PRSR [7], (g) SR-GAN [28], (h) Wavelet-SRNet [17], (i) SiGAN [15], and (j) the ground-truths ( $64 \times 64$ ).



Input    Shen *et al.* [49] Super-FAN [3]    UMSN [64] DeblurGANv2 [26] DFDNet [34] HiFaceGAN [60] NASFE [63] Ground-Truth

Fig. 10. Qualitative comparisons of some state-of-the-art **joint face illumination compensation methods** on synthetic test set *Test-BNL*.



Fig. 11. Qualitative comparisons of some state-of-the-art **joint face 3D reconstruction methods**. From left to right, we demonstrate the ground-truth image, the LQ input, and the results from L2R [79], Unsup3D [56], LAP [78], DF2Net [74] and DECA [10], respectively.

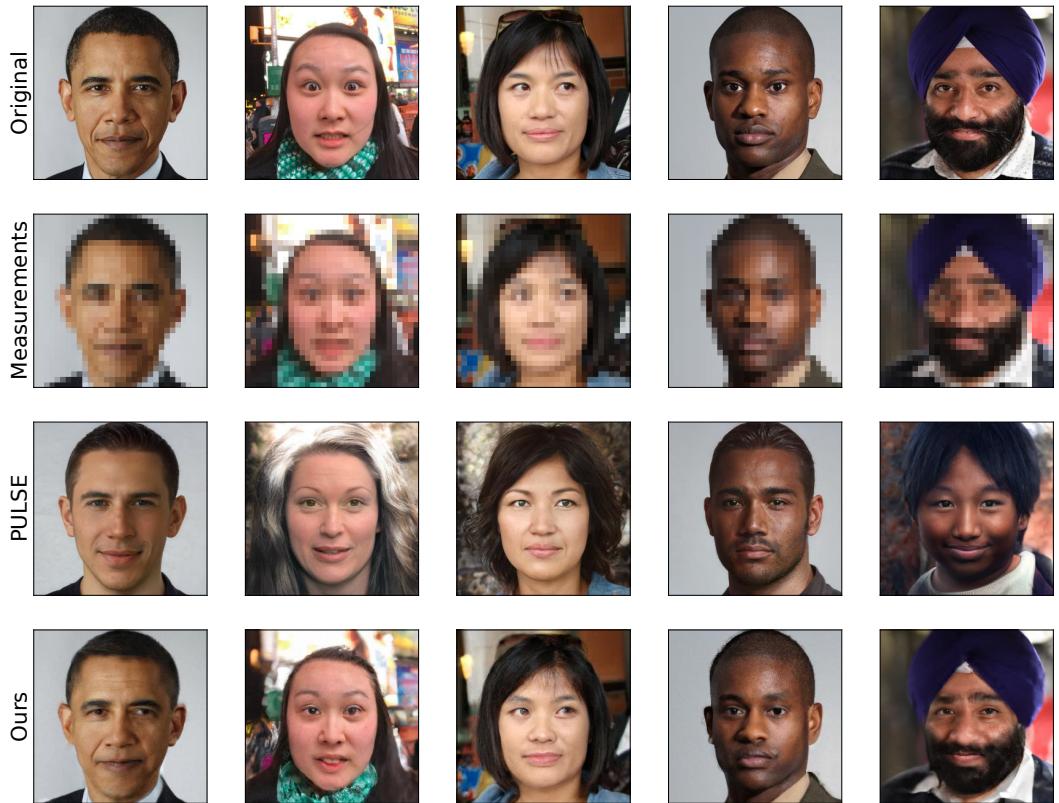


Fig. 12. **Joint face fairness** comparisons on Barack Obama and four faces from the FFHQ dataset. The top row shows original images, the second row shows what the algorithms observe: blurry measurements after downsampling by  $32\times$  in each dimension. The third row shows reconstructions by PULSE [41], and the last row shows reconstructions by Posterior Sampling via Langevin dynamics, the algorithm [19] are advocating for. These faces were chosen to compare performance on various ethnicities.

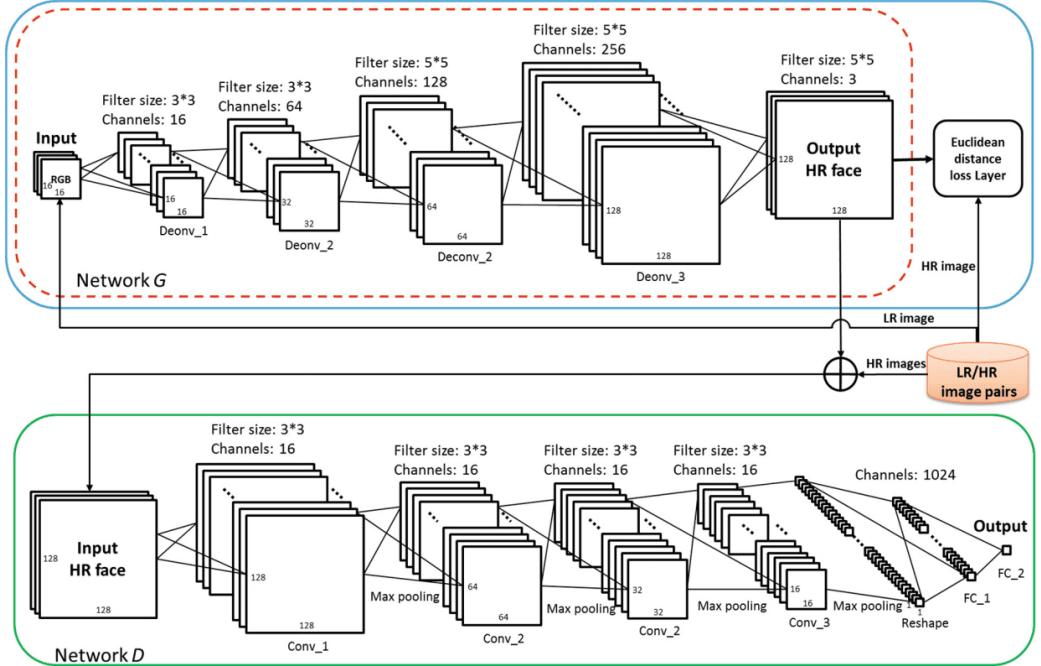


Fig. 13. The method URDGN [69] using the **general GAN structure**. In the testing phase, only the generative network in the red dashed block is employed. (Color figure online)

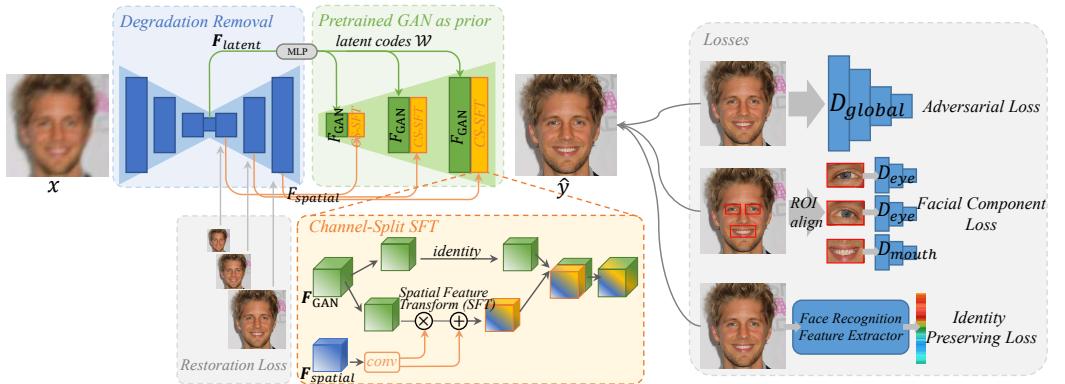
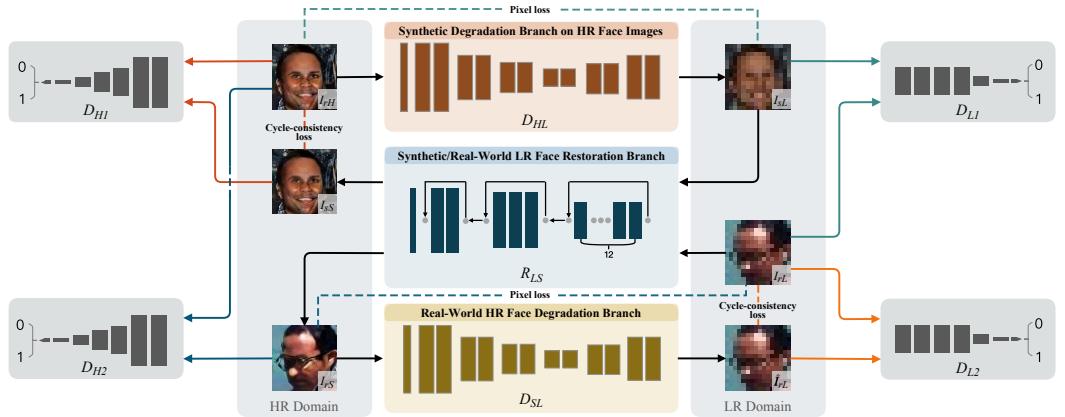
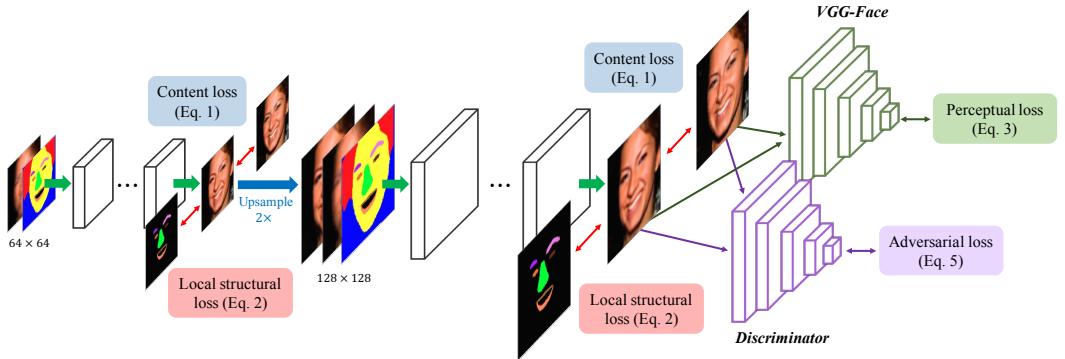


Fig. 14. The method GFPGAN [55] using the **Pre-trained GAN structure**. It consists of a degradation removal module (U-Net) and a pre-trained face GAN as a facial prior.

Fig. 15. The method SCGAN [14] using the **unpaired images GAN structure**.Fig. 16. The method DIC [39] using the **Pre-prior framework**.

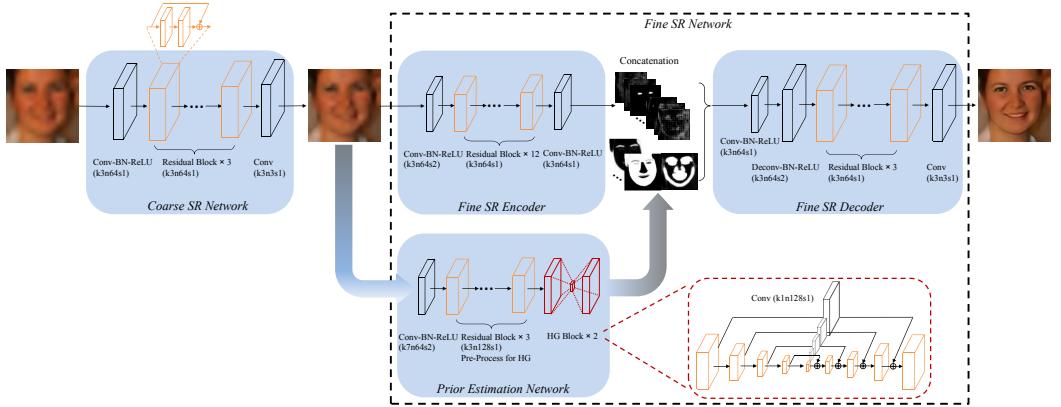


Fig. 17. The method FSRNet [6] using the **In-prior framework**.

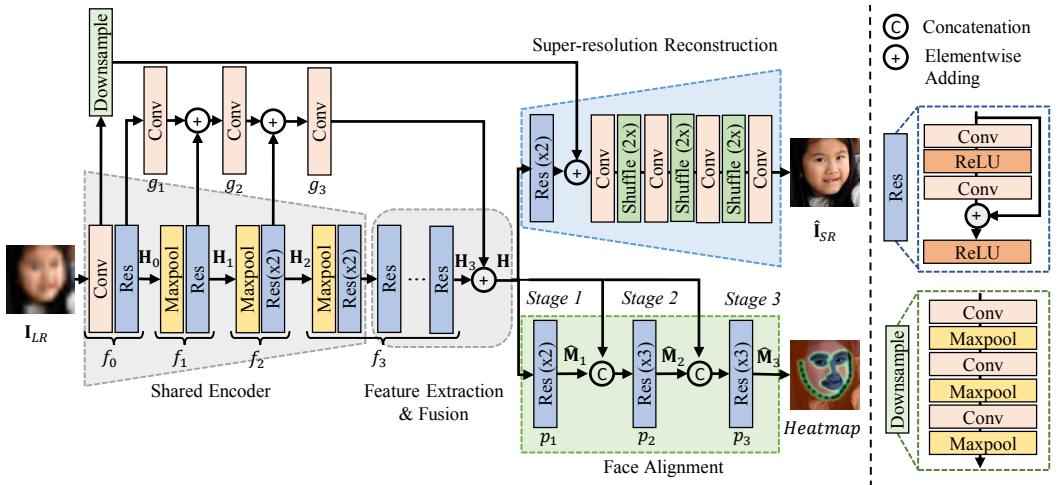


Fig. 18. The method JASRNet [66] using the **Parallel-prior framework**.

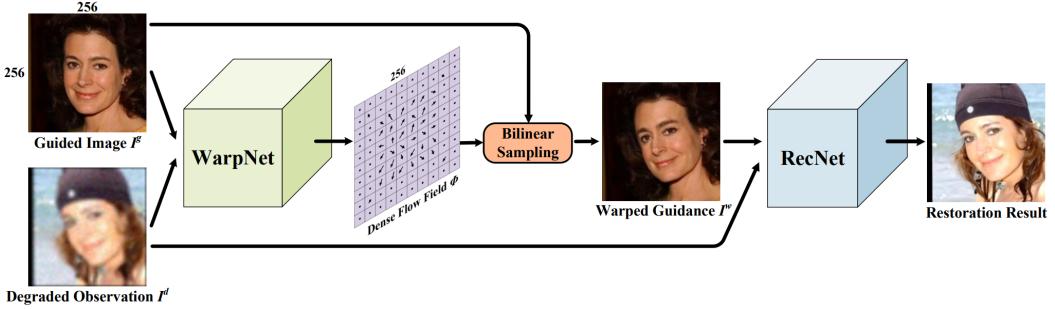


Fig. 19. The method GFRNet [36] using the **Input-prior framework**.

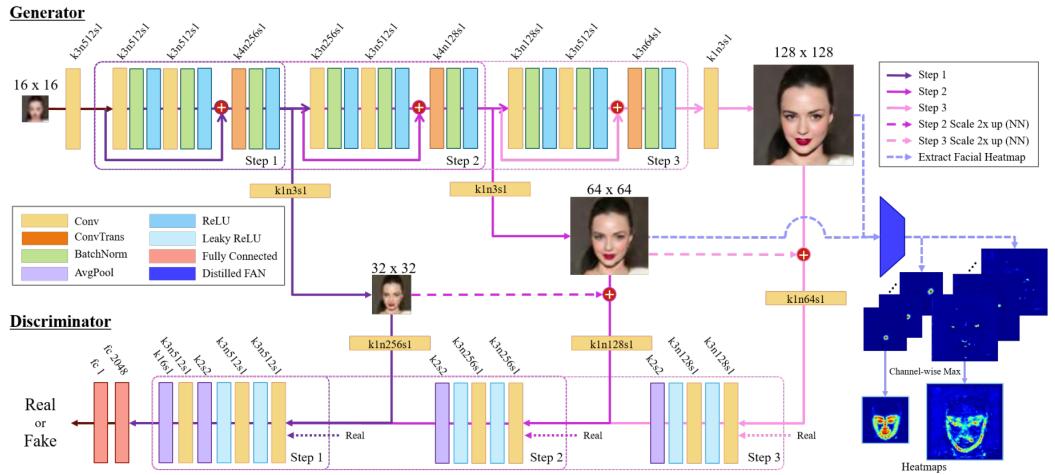


Fig. 20. The method FAN [3] using the **Post-prior framework**.

## REFERENCES

- [1] Brandon Amos, Bartosz Ludwiczuk, Mahadev Satyanarayanan, et al. 2016. Openface: A general-purpose face recognition library with mobile applications. *CMU School of Computer Science* 6, 2 (2016), 20.
- [2] Adrian Bulat and Georgios Tzimiropoulos. 2017. How far are we from solving the 2d & 3d face alignment problem?(and a dataset of 230,000 3d facial landmarks). In *ICCV*. 1021–1030.
- [3] Adrian Bulat and Georgios Tzimiropoulos. 2018. Super-fan: Integrated facial landmark localization and super-resolution of real-world low resolution faces in arbitrary poses with gans. In *CVPR*. 109–117.
- [4] Qiong Cao, Li Shen, Weidi Xie, Omkar M Parkhi, and Andrew Zisserman. 2018. Vggface2: A dataset for recognising faces across pose and age. In *FG*. IEEE, 67–74.
- [5] Chaofeng Chen, Xiaoming Li, Lingbo Yang, Xianhui Lin, Lei Zhang, and Kwan-Yee K Wong. 2021. Progressive semantic-aware style transformation for blind face restoration. In *CVPR*. 11896–11905.
- [6] Yu Chen, Ying Tai, Xiaoming Liu, Chunhua Shen, and Jian Yang. 2018. Fsrnet: End-to-end learning face super-resolution with facial priors. In *CVPR*. 2492–2501.
- [7] Ryan Dahl, Mohammad Norouzi, and Jonathon Shlens. 2017. Pixel recursive super resolution. In *ICCV*. 5439–5448.
- [8] Berk Dogan, Shuhang Gu, and Radu Timofte. 2019. Exemplar guided face image super-resolution without facial landmarks. In *CVPRW*. 0–0.

- [9] Chao Dong, Chen Change Loy, Kaiming He, and Xiaou Tang. 2015. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence* 38, 2 (2015), 295–307.
- [10] Yao Feng, Haiwen Feng, Michael J Black, and Timo Bolkart. 2021. Learning an animatable detailed 3D face model from in-the-wild images. *ACM Transactions on Graphics* 40, 4 (2021), 1–13.
- [11] Mislav Grgic, Kresimir Delac, and Sonja Grgic. 2011. SCface—surveillance cameras face database. *Multimedia Tools and Applications* 51 (2011), 863–879.
- [12] Ralph Gross, Iain Matthews, Jeffrey Cohn, Takeo Kanade, and Simon Baker. 2010. Multi-pie. *Image and Vision Computing* 28, 5 (2010), 807–813.
- [13] Shuhang Gu, Wangmeng Zuo, Qi Xie, Deyu Meng, Xiangchu Feng, and Lei Zhang. 2015. Convolutional sparse coding for image super-resolution. In *ICCV*. 1823–1831.
- [14] Hao Hou, Jun Xu, Yingkun Hou, Xiaotao Hu, Benzheng Wei, and Dinggang Shen. 2023. Semi-cycled generative adversarial networks for real-world face super-resolution. *IEEE Transactions on Image Processing* 32 (2023), 1184–1199.
- [15] Chih-Chung Hsu, Chia-Wen Lin, Weng-Tai Su, and Gene Cheung. 2019. Sigan: Siamese generative adversarial network for identity-preserving face hallucination. *IEEE Transactions on Image Processing* 28, 12 (2019), 6225–6236.
- [16] Gary B Huang, Marwan Mattar, Tamara Berg, and Eric Learned-Miller. 2008. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. In *Proceedings of Workshop on Faces in ‘Real-Life’ Images: Detection, Alignment, and Recognition*.
- [17] Huaibo Huang, Ran He, Zhenan Sun, and Tieniu Tan. 2017. Wavelet-srnet: A wavelet-based cnn for multi-scale face super resolution. In *ICCV*. 1689–1697.
- [18] Huaibo Huang, Ran He, Zhenan Sun, and Tieniu Tan. 2019. Wavelet domain generative adversarial network for multi-scale face hallucination. *International Journal of Computer Vision* 127, 6–7 (2019), 763–784.
- [19] Ajil Jalal, Sushrut Karmalkar, Jessica Hoffmann, Alex Dimakis, and Eric Price. 2021. Fairness for image generation with uncertain sensitive attributes. In *ICML*. PMLR, 4721–4732.
- [20] Junjun Jiang, Chenyang Wang, Xianming Liu, and Jiayi Ma. 2021. Deep learning-based face super-resolution: A survey. *Comput. Surveys* 55, 1 (2021), 1–36.
- [21] Ratheesh Kalarot, Tao Li, and Fatih Porikli. 2020. Component attention guided face super-resolution network: Cagface. In *WACV*. 370–380.
- [22] Tero Karras, Samuli Laine, and Timo Aila. 2019. A style-based generator architecture for generative adversarial networks. In *CVPR*. 4401–4410.
- [23] Deokyun Kim, Minseon Kim, Gihyun Kwon, and Dae-Shik Kim. 2019. Progressive face super-resolution via attention to facial landmark. *arXiv preprint arXiv:1908.08239* (2019).
- [24] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. 2016. Accurate image super-resolution using very deep convolutional networks. In *CVPR*. 1646–1654.
- [25] Martin Koestinger, Paul Wohlhart, Peter M Roth, and Horst Bischof. 2011. Annotated facial landmarks in the wild: A large-scale, real-world database for facial landmark localization. In *ICCVW*. 2144–2151.
- [26] Orest Kupyn, Tetiana Martyniuk, Junru Wu, and Zhangyang Wang. 2019. Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better. In *ICCV*. 8878–8887.
- [27] Vuong Le, Jonathan Brandt, Zhe Lin, Lubomir Bourdev, and Thomas S Huang. 2012. Interactive facial feature localization. In *ECCV*. 679–692.
- [28] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. 2017. Photo-realistic single image super-resolution using a generative adversarial network. In *CVPR*. 4681–4690.
- [29] Cheng-Han Lee, Ziwei Liu, Lingyun Wu, and Ping Luo. 2020. Maskgan: Towards diverse and interactive facial image manipulation. In *CVPR*. 5549–5558.
- [30] Jichun Li, Bahetiyaer Bare, Shili Zhou, Bo Yan, and Ke Li. 2021. Organ-branched CNN for robust face super-resolution. In *ICME*. IEEE, 1–6.
- [31] Jingyuan Li, Ning Wang, Lefei Zhang, Bo Du, and Dacheng Tao. 2020. Recurrent feature reasoning for image inpainting. In *CVPR*. 7760–7768.
- [32] Luying Li, Junshu Tang, Zhou Ye, Bin Sheng, Lijuan Mao, and Lizhuang Ma. 2021. Unsupervised face super-resolution via gradient enhancement and semantic guidance. *The Visual Computer* 37 (2021), 2855–2867.
- [33] Mengyan Li, Zhaoyu Zhang, Jun Yu, and Chang Wen Chen. 2020. Learning face image super-resolution through facial semantic attribute transformation and self-attentive structure enhancement. *IEEE Transactions on Multimedia* 23 (2020), 468–483.
- [34] Xiaoming Li, Chaofeng Chen, Shangchen Zhou, Xianhui Lin, Wangmeng Zuo, and Lei Zhang. 2020. Blind face restoration via deep multi-scale component dictionaries. In *ECCV*. Springer, 399–415.
- [35] Xiaoming Li, Wenyu Li, Dongwei Ren, Hongzhi Zhang, Meng Wang, and Wangmeng Zuo. 2020. Enhanced blind face restoration with multi-exemplar images and adaptive spatial feature fusion. In *CVPR*. 2706–2715.

- [36] Xiaoming Li, Ming Liu, Yuting Ye, Wangmeng Zuo, Liang Lin, and Ruigang Yang. 2018. Learning warped guidance for blind face restoration. In *ECCV*. 272–289.
- [37] Xiaoming Li, Shiguang Zhang, Shangchen Zhou, Lei Zhang, and Wangmeng Zuo. 2022. Learning Dual Memory Dictionaries for Blind Face Restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2022).
- [38] Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. 2015. Deep learning face attributes in the wild. In *ICCV*. 3730–3738.
- [39] Cheng Ma, Zhenyu Jiang, Yongming Rao, Jiwen Lu, and Jie Zhou. 2020. Deep face super-resolution with iterative collaboration between attentive recovery and landmark estimation. In *CVPR*. 5569–5578.
- [40] Brianna Maze, Jocelyn Adams, James A Duncan, Nathan Kalka, Tim Miller, Charles Otto, Anil K Jain, W Tyler Niggel, Janet Anderson, Jordan Cheney, et al. 2018. Iarpa janus benchmark-c: Face dataset and protocol. In *ICB*. IEEE, 158–165.
- [41] Sachit Menon, Alexandru Damian, Shijia Hu, Nikhil Ravi, and Cynthia Rudin. 2020. Pulse: Self-supervised photo upsampling via latent space exploration of generative models. In *CVPR*. 2437–2445.
- [42] Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee. 2017. Deep multi-scale convolutional neural network for dynamic scene deblurring. In *CVPR*. 3883–3891.
- [43] Kamyar Nazeri, Eric Ng, Tony Joseph, Faisal Z Qureshi, and Mehran Ebrahimi. 2019. Edgeconnect: Generative image inpainting with adversarial edge learning. *arXiv preprint arXiv:1901.00212* (2019).
- [44] Jinshan Pan, Zhe Hu, Zhixun Su, and Ming-Hsuan Yang. 2014. Deblurring face images with exemplars. In *ECCV*. Springer, 47–62.
- [45] Omkar Parkhi, Andrea Vedaldi, and Andrew Zisserman. 2015. Deep face recognition. In *BMVC*. British Machine Vision Association.
- [46] Alec Radford, Luke Metz, and Soumith Chintala. 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434* (2015).
- [47] Rasmus Rothe, Radu Timofte, and Luc Van Gool. 2015. Dex: Deep expectation of apparent age from a single image. In *ICCVW*. 10–15.
- [48] Christos Sagonas, Georgios Tzimiropoulos, Stefanos Zafeiriou, and Maja Pantic. 2013. 300 faces in-the-wild challenge: The first facial landmark localization challenge. In *ICCVW*. 397–403.
- [49] Ziyi Shen, Wei-Sheng Lai, Tingfa Xu, Jan Kautz, and Ming-Hsuan Yang. 2018. Deep semantic face deblurring. In *CVPR*. 8260–8269.
- [50] Ziyi Shen, Wei-Sheng Lai, Tingfa Xu, Jan Kautz, and Ming-Hsuan Yang. 2020. Exploiting semantics for face image deblurring. *International Journal of Computer Vision* 128 (2020), 1829–1846.
- [51] Yibing Song, Jiawei Zhang, Lijun Gong, Shengfeng He, Linchao Bao, Jinshan Pan, Qingxiong Yang, and Ming-Hsuan Yang. 2019. Joint face hallucination and deblurring via structure generation and detail enhancement. *International journal of computer vision* 127 (2019), 785–800.
- [52] Yibing Song, Jiawei Zhang, Shengfeng He, Linchao Bao, and Qingxiong Yang. 2017. Learning to hallucinate face images via component generation and enhancement. *arXiv preprint arXiv:1708.00223* (2017).
- [53] Luan Tran, Xi Yin, and Xiaoming Liu. 2018. Representation learning by rotating your faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 41, 12 (2018), 3007–3021.
- [54] Junke Wang, Shaoliang Chen, Zuxuan Wu, and Yu-Gang Jiang. 2022. Ft-tdr: Frequency-guided transformer and top-down refinement network for blind face inpainting. *IEEE Transactions on Multimedia* (2022).
- [55] Xintao Wang, Yu Li, Honglun Zhang, and Ying Shan. 2021. Towards real-world blind face restoration with generative facial prior. In *CVPR*. 9168–9178.
- [56] Shangzhe Wu, Christian Rupprecht, and Andrea Vedaldi. 2020. Unsupervised learning of probably symmetric deformable 3d objects from images in the wild. In *CVPR*. 1–10.
- [57] Li Xu, Shicheng Zheng, and Jiaya Jia. 2013. Unnatural l0 sparse representation for natural image deblurring. In *CVPR*. 1107–1114.
- [58] Xiangyu Xu, Deqing Sun, Jinshan Pan, Yujin Zhang, Hanspeter Pfister, and Ming-Hsuan Yang. 2017. Learning to super-resolve blurry face and text images. In *ICCV*. 251–260.
- [59] Chih-Yuan Yang, Sifei Liu, and Ming-Hsuan Yang. 2013. Structured face hallucination. In *CVPR*. 1099–1106.
- [60] Lingbo Yang, Shanshe Wang, Siwei Ma, Wen Gao, Chang Liu, Pan Wang, and Peiran Ren. 2020. Hifacegan: Face renovation via collaborative suppression and replenishment. In *ACMMM*. 1551–1560.
- [61] Shuo Yang, Ping Luo, Chen-Change Loy, and Xiaoou Tang. 2016. Wider face: A face detection benchmark. In *CVPR*. 5525–5533.
- [62] Yang Yang and Xiaojie Guo. 2020. Generative landmark guided face inpainting. In *PRCV*. Springer, 14–26.
- [63] Rajeev Yasarla, Hamid Reza Vaezi Joze, and Vishal M Patel. 2021. Network architecture search for face enhancement. *arXiv preprint arXiv:2105.06528* (2021).
- [64] Rajeev Yasarla, Federico Perazzi, and Vishal M Patel. 2020. Deblurring face images using uncertainty guided multi-stream semantic networks. *IEEE Transactions on Image Processing* 29 (2020), 6251–6263.

- [65] Dong Yi, Zhen Lei, Shengcai Liao, and Stan Z Li. 2014. Learning face representation from scratch. *arXiv preprint arXiv:1411.7923* (2014).
- [66] Yu Yin, Joseph Robinson, Yulun Zhang, and Yun Fu. 2020. Joint super-resolution and alignment of tiny faces. In *AAAI*, Vol. 34. 12693–12700.
- [67] Xin Yu, Basura Fernando, Bernard Ghanem, Fatih Porikli, and Richard Hartley. 2018. Face super-resolution guided by facial component heatmaps. In *ECCV*. 217–233.
- [68] Xin Yu, Basura Fernando, Richard Hartley, and Fatih Porikli. 2018. Super-resolving very low-resolution face images with supplementary attributes. In *CVPR*. 908–917.
- [69] Xin Yu and Fatih Porikli. 2016. Ultra-resolving face images by discriminative generative networks. In *ECCV*. Springer, 318–333.
- [70] Xin Yu and Fatih Porikli. 2017. Face hallucination with tiny unaligned images by transformative discriminative neural networks. In *AAAI*, Vol. 31.
- [71] Xin Yu and Fatih Porikli. 2017. Hallucinating very low-resolution unaligned and noisy face images by transformative discriminative autoencoders. In *CVPR*. 3760–3768.
- [72] Xin Yu, Fatih Porikli, Basura Fernando, and Richard Hartley. 2020. Hallucinating unaligned face images by multiscale transformative discriminative networks. *International Journal of Computer Vision* 128, 2 (2020), 500–526.
- [73] Xin Yu, Fatemeh Shiri, Bernard Ghanem, and Fatih Porikli. 2019. Can we see more? Joint frontalization and hallucination of unaligned tiny faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 42, 9 (2019), 2148–2164.
- [74] Xiaoxing Zeng, Xiaojiang Peng, and Yu Qiao. 2019. Df2net: A dense-fine-finer network for detailed 3d face reconstruction. In *ICCV*. 2315–2324.
- [75] Kaihao Zhang, Dongxu Li, Wenhuan Luo, Jingyu Liu, Jiankang Deng, Wei Liu, and Stefanos Zafeiriou. 2022. EDFace-Celeb-1 M: Benchmarking Face Hallucination with a Million-scale Dataset. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2022).
- [76] Yang Zhang, Ivor W Tsang, Jun Li, Ping Liu, Xiaobo Lu, and Xin Yu. 2021. Face hallucination with finishing touches. *IEEE Transactions on Image Processing* 30 (2021), 1728–1743.
- [77] Yunchen Zhang, Yi Wu, and Liang Chen. 2020. MSFSR: A multi-stage face super-resolution with accurate facial representation via enhanced facial boundaries. In *CVPRW*. 504–505.
- [78] Zhenyu Zhang, Yanhao Ge, Renwang Chen, Ying Tai, Yan Yan, Jian Yang, Chengjie Wang, Jilin Li, and Feiyue Huang. 2021. Learning to aggregate and personalize 3d face from in-the-wild photo collection. In *CVPR*. 14214–14224.
- [79] Zhenyu Zhang, Yanhao Ge, Ying Tai, Xiaoming Huang, Chengjie Wang, Hao Tang, Dongjin Huang, and Zhifeng Xie. 2022. Learning to Restore 3D Face from In-the-Wild Degraded Images. In *CVPR*. 4237–4247.
- [80] Lin Zhong, Sunghyun Cho, Dimitris Metaxas, Sylvain Paris, and Jue Wang. 2013. Handling noise in single image deblurring using directional filters. In *CVPR*. 612–619.
- [81] Shizhan Zhu, Sifei Liu, Chen Change Loy, and Xiaoou Tang. 2016. Deep cascaded bi-network for face hallucination. In *ECCV*. Springer, 614–630.