Supplementary Materials: Personalized Restoration via Dual-Pivot Tuning

Pradyumna Chari, Sizhuo Ma, Daniil Ostashev,
Achuta Kadambi, Gurunandan Krishnan, Jian Wang, Kfir Aberman

Additional Qualitative Analysis

In this supplementary document, we discuss additional qualitative results, comparing with a range of baselines. We first qualitatively compare our proposed method with existing diffusion-based blind restoration techniques. Specifically, we compare with DifFace [1], DR2 [2] and DiffBIR [3]. Fig. 19 highlights these observations. We note that in our operating regime for degradation, both DR2 and DifFace result in inaccurate and unreliable restoration (in both these cases, the models trade off fidelity with restoration and need to be used accordingly depending on the degree of degradation). In contrast, DiffBIR, while worse in both detail and identity when compared with our method, proves to be a much more reliable blind restoration technique. We therefore use DiffBIR as our base model for our method, and use it as the representative diffusion-based blind restoration technique in our analysis.

We also show additional results in Fig. 20 and Fig. 22, to provide additional samples on top of the results in Figure 7 and Figure 8 of the main paper. These further reinforce our observations from the main paper on the superiority of the proposed method.

We next describe qualitative performance in comparision with GFPGAN [6] and CodeFormer [7] as blind image restoration methods, and MyStyle [8] as a method with a personalized generative prior, in additional detail to the paper. Fig. 21 shows results on synthetic degradations and corresponding restorations across six different identities. As can be seen, the blind restoration methods show clear drifts in identity across all six examples, while retaining fidelity to the input image. On the other hand, [8], while being able to retain a strong identity prior, sees a significant deviation in the restored image, from the input degraded image (potentially as a result of personalization on a small number of images (10)). In contrast, our method is able to achieve the best of both worlds: while maintaining a high degreee of faithfulness to the input image, we also see consistent identity retention across all examples. This leads to our results being closest to the reference image, despite the input having severe degradation in several cases.

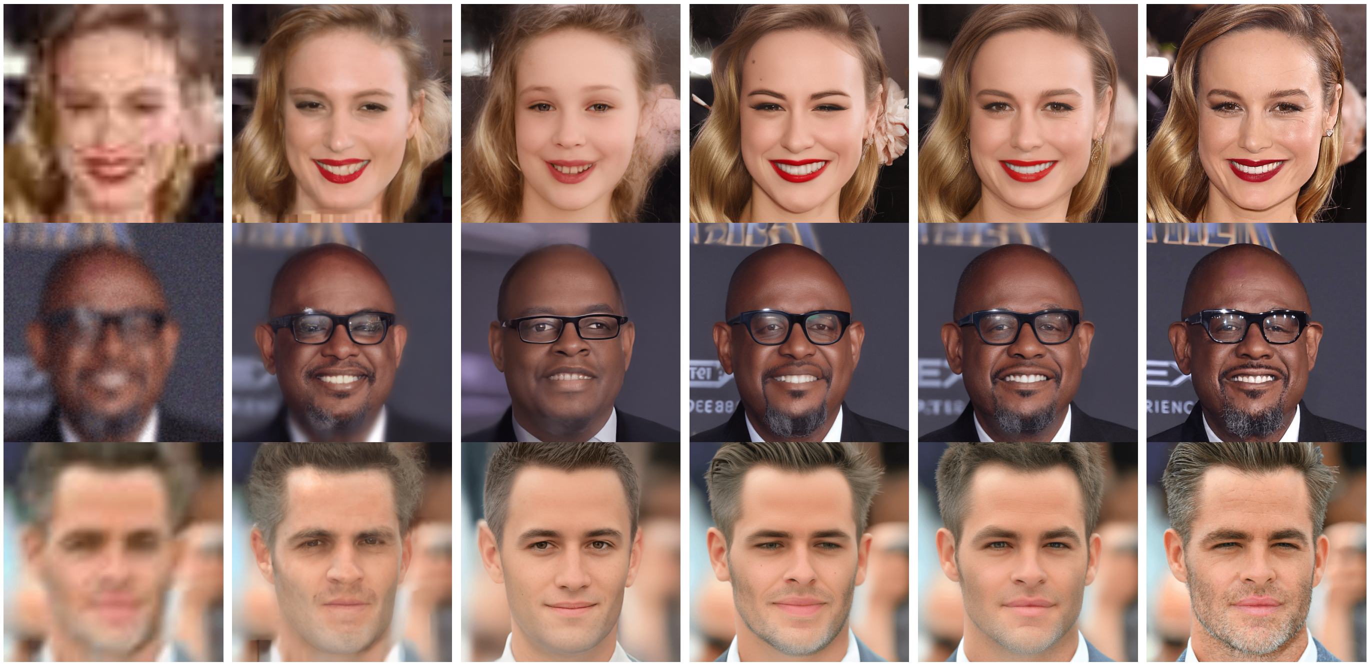
We next look at Fig. 23, for an analysis on real degraded images when compared with these additional baselines. Again, we note consistent observation. The blind restoration methods remain faithful to the input image, however they result in considerable identity drifts and artifacts in the restored images. MyStyle is able to continue retaining a strong personalized prior, however at the cost of losing all context relating to the input degraded image. Again, our method is able to retain identity, while being faithful to the input image.

Consistent restoration through personalized prior. To show the strength of our personalized prior, we conduct an experiment by applying different synthetic degradations on the same image and observe how each method restores them, as shown in Fig. 24. The upper row shows a blurring degradation, while the lower row shows a compression degradation. For ASFFNet and DMDNet, the quality of the restored images varies with the degradations being applied. DiffBIR generates higher quality images, but the identity remains sensitive to the input degraded image. For example, in Fig. 24(d), note the inconsistent eye and nose shape, when compared with Fig. 24(f). Our personalized model effectively restores the images with consistent perceptual quality and identity fidelity. This indicates the stability of the contextual prior and the reliability of the encoder conditioning, across degradation types.

Further, we supply additional synthetic (Fig. 25) as well as real (Fig. 26) degradation results across all baselines. We wish to highlight the robustness of our proposed method, across identites and degradations, over a larger number of image settings. These aspects can specifically be seen through identifying features in the participants, such as hair, teeth, ears, eye color and so on. The overall trends remain the same: prior methods are either able to retain strong identity without artifacts, or retaining faithfulness to the input image. It is through our personalzation regime that we achieve the best of both these worlds, getting identity-consistent restored images with high fidelity.

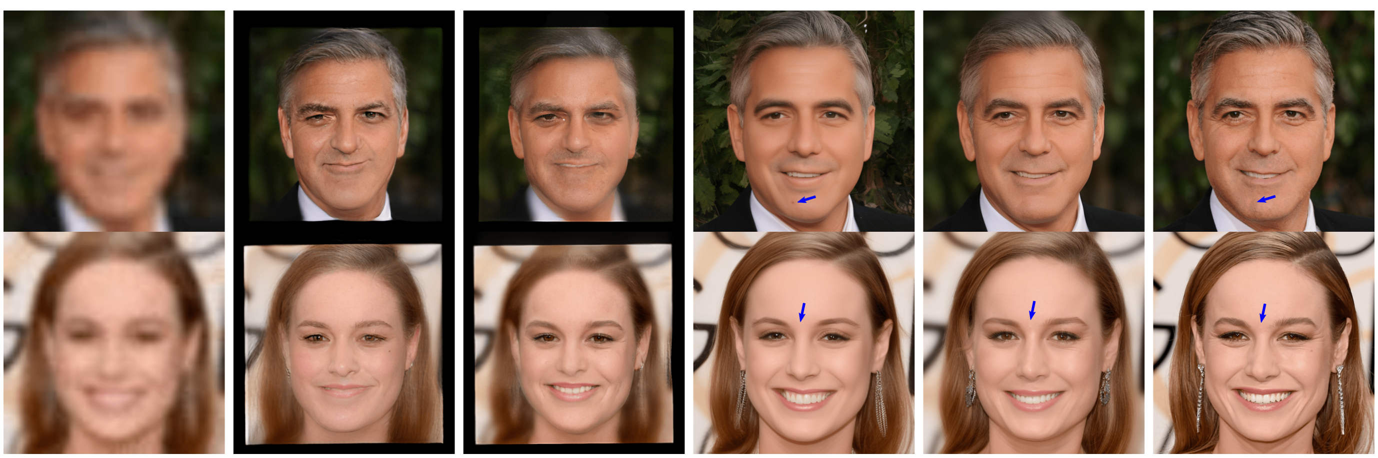
References

[1] Z. Yue and C. C. Loy, "Difface: Blind face restoration with diffused error contraction," *arXiv preprint arXiv:2212.06512*, 2022.
[2] Z. Wang, Z. Zhang, X. Zhang, H. Zheng, M. Zhou, Y. Zhang, and Y. Wang, "Dr2: Diffusion-based robust degradation remover for blind face restoration," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 1704–1713.
[3] X. Lin, J. He, Z. Chen, Z. Lyu, B. Fei, B. Dai, W. Ouyang, Y. Qiao, and C. Dong, "Diffbir: Towards blind image restoration with generative diffusion prior," *arXiv preprint arXiv:2308.15070*, 2023.
[4] X. Li, W. Li, D. Ren, H. Zhang, M. Wang, and W. Zuo, "Enhanced blind face restoration with multi-exemplar images and adaptive spatial feature fusion," in *CVPR*, 2020.
[5] X. Li, S. Zhang, S. Zhou, L. Zhang, and W. Zuo, "Learning dual memory dictionaries for blind face restoration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.
[6] X. Wang, Y. Li, H. Zhang, and Y. Shan, "Towards real-world blind face restoration with generative facial prior," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.
[7] S. Zhou, K. Chan, C. Li, and C. C. Loy, "Towards robust blind face restoration with codebook lookup transformer," *Advances in Neural Information Processing Systems*, vol. 35, pp. 30 599–30 611, 2022.



(a) DEGRADED IMAGE (b) DIFFACE [1] (c) DR2 [2] (d) DIFFBIR [3] (e) OURS (f) GROUND TRUTH

Fig. 19: **Comparison with existing diffusion-based blind image restoration methods.** We find that for the degree of degradation we deal with in our experiments, both DifFace [1] as well as DR2 [2] show unreliable performance, losing both identity and fidelity at different instances. DiffBIR [3], on the other hand, while not performing as well as our method in terms of identity and fidelity retention, is the best performer among the prior diffusion-based methods. We therefore choose DiffBIR as our base model and comparison benchmark for subsequent experiments.



(a) DEGRADED IMAGE (b) ASFFNET [4] (c) DMDNET [5] (d) DIFFBIR [3] (e) OURS (f) GROUND TRUTH

Fig. 20: **Additional Results on synthetically degraded images (in addition to main paper, Figure 7).**

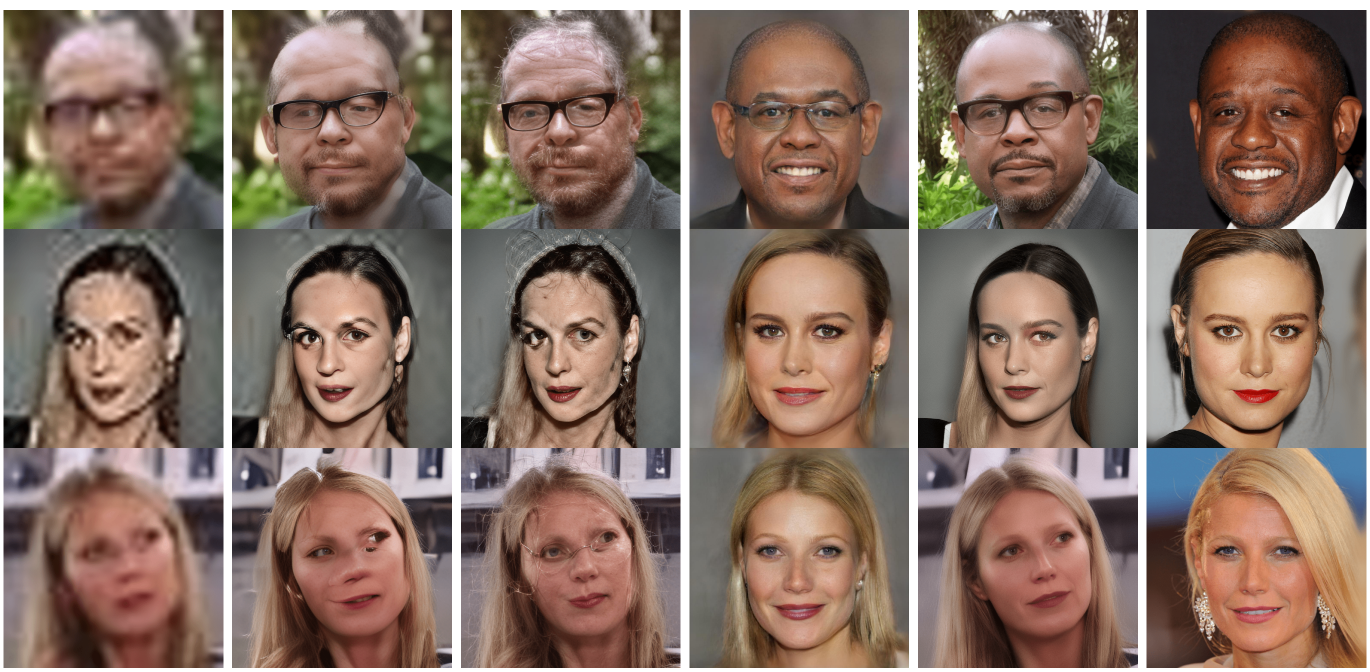
[8] Y. Nitzan, K. Aberman, Q. He, O. Liba, M. Yarom, Y. Gandelsman, I. Mosseri, Y. Pritch, and D. Cohen-Or, “Mystyle: A personalized generative prior,” *ACM Transactions on Graphics (TOG)*, vol. 41, no. 6, pp. 1–10, 2022.

(a) DEGRADED IMAGE (b) GFPGAN [6] (c) CODEFORMER [7] (d) MYSTYLE [8] (e) OURS (f) GROUND TRUTH

Fig. 21: **Additional baseline methods on synthetically degraded images (in addition to the results in the main paper).** Similar observations to the synthetically degraded results can be made. The blind image restoration methods suffer both on identity and quality, while MyStyle is able to retain identity at the cost of losing faithfulness to the input image. Our method is able to retain identity in the restoration, while being faithful to the input image. In terms of specifics for each row, note the eye and eyebrow shape in row 1, eye shape, bags under the eyes, and teeth in row 2, eye shape and ears (specifically, left ear) in row 3, mole on the left cheek and mark between the eyes in row 4, eyes and mouth expression in row 5, and eye color in row 6. Please zoom in to observe these details more easily.

(a) Degraded Image (b) ASFFNet [4] (c) DMDNet [5] (d) DiffBIR [3] (e) Ours (f) Id. Reference

Fig. 22: **Additional results on real degraded images (in addition to main paper, Figure 8).**



(a) Degraded Image (b) GFPGAN [6] (c) CodeFormer [7] (d) MyStyle [8] (e) Ours (f) Id. Reference

Fig. 23: **Additional baseline methods on real degraded images (in addition to the results in the main paper).** Comparison methods either result in artifacts or in an image that is not faithful to the input, while the proposed method is able to stably inject relevant identity information while retaining fidelity to the input degraded image.

(a) Degraded Image (b) ASFFNet [4] (c) DMDNet [5] (d) DiffBIR [3] (e) Ours (f) Ground Truth

Fig. 24: **Our approach is agnostic to different types of degradation, such as blur (top row), compression (middle row).** Prior face image restoration methods (b, c) have their estimates considerably affected by the nature of the degradation (note artifacts near eyes and mouth). Prior unconditional diffusion methods (d) have more consistent performance, but with lost identity information (note **eye shape**, **nose shape**). Our proposed method provides consistent restoration across a range of degradations, while retaining identity.

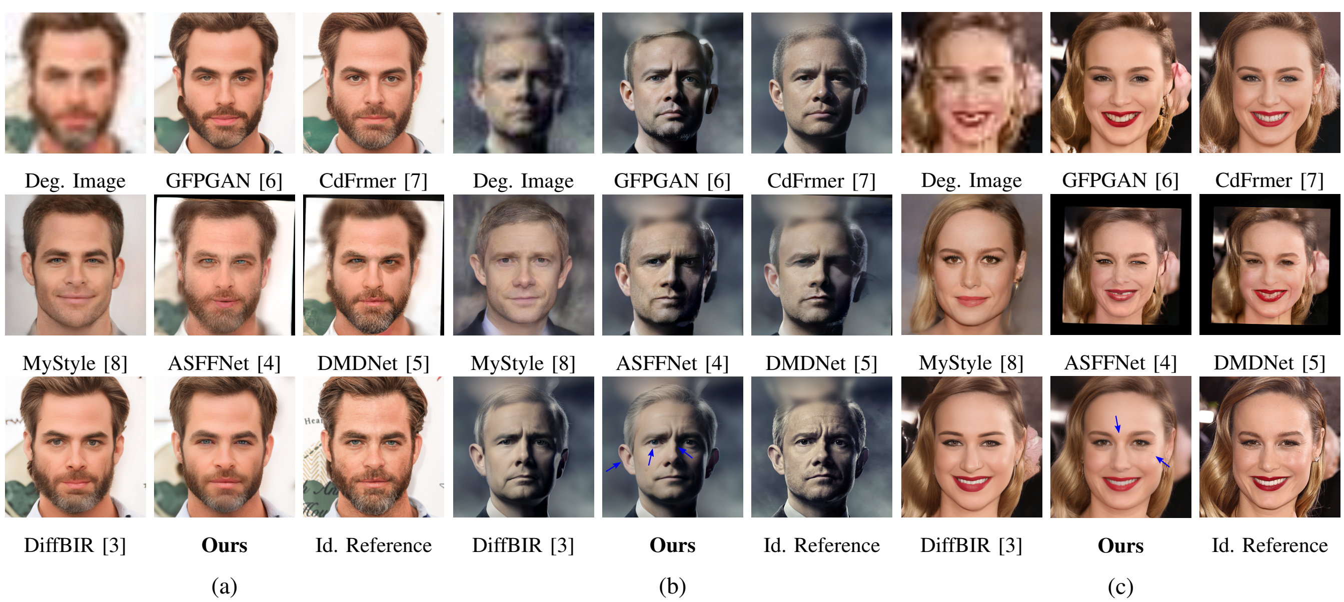


Fig. 25: **Additional results on synthetic degraded images, comparing with all discussed major comparison methods.** For identity (a), zooming in shows that all baseline methods either lead to considerable artifacts in the restored image, or through identity drifts (eye color) and lack of fidelity with the input image. For identity (b) in cases where the restored image does not have significant artifacts, identity drifts can be noted in the form of nose shape (as highlighted by the shadow on the nose), eye color and ear shape. For identity (c), the mark between the eyes and the mole on the left cheek are identifying features that show the superiority of the proposed method.

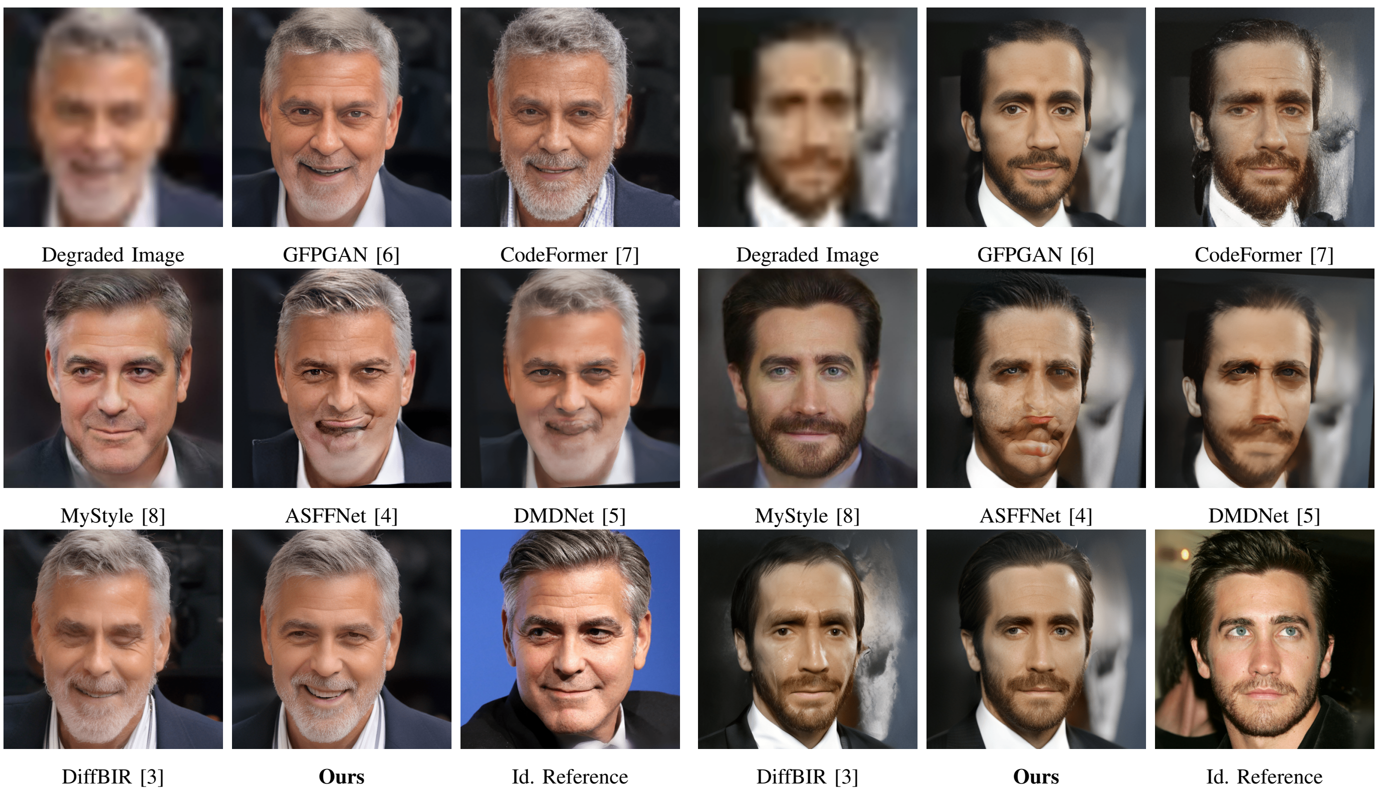


Fig. 26: **Additional results on real degraded images, comparing with all discussed major comparison methods.** For both identities, zooming in shows that comparison methods (both prior reference-based methods and blind restoration methods) result in strong artifacts, identity drifts, and in the case of MyStyle [8], lack of fidelity with the input degraded image. The proposed method is the only one consistently able to incorporate identity in the restoration in a stable manner. Please zoom in to the image for a clearer visualization.