PerFuSe: Personalized Full-Image Restoration via Modular Fusion
Supplementary Material

Nazish Naeem[1] Zeeshan Nadir[2] Hamid R. Sheikh[2] SeokJun Lee[2]
[1]Massachusetts Institute of Technology [2]Samsung Research America
nazishn@mit.edu, {zeeshan.nadir, hr.sheikh, seokjun1.lee}@samsung.com

1. Implementation

Face Extraction. We follow a landmark-based alignment and cropping pipeline to ensure consistent facial geometry across all inputs. Specifically, we employ the dlib [2] 68-point facial landmark detector to locate key points corresponding to the eyes and mouth. Using these landmarks, we form a triangle spanning the two eyes and mouth center, and rotate the image such that the line joining the eyes becomes perfectly horizontal. To ensure sufficient face coverage within the crop, we measure the area occupied by the triangle connecting the eyes and mouth; if it covers less than 2% of the crop, the region is tightened until this threshold is met. The final crop and its corresponding coordinates are stored for later alignment and blending stages.
Mask Guided Context-Aware Blending. We use YOLOv8 [1] for instance segmentation and RANSAC-Flow [3] for fine alignment. Blending is performed using three level pyramids (Laplacian and Gaussian as mentioned in the main paper).
We train and test PerFuSe on H100 GPUs.

2. Additional Results

2.1. Performance of Mask Guided Context-Aware Blending

In this section, we present the qualitative performance of PerFuSe's mask-guided, context-aware blending approach. Figure S1a shows the low-quality input image, Figure S1b shows the restored PerFuSe output without context-aware blending, and Figure S1c shows the restored image with context-aware blending. As illustrated, the inclusion of contextual information enables seamless blending across the outputs of different restoration modules. In contrast, the absence of context leads to dark artifacts along the blending boundaries because the algorithm lacks information about the surrounding regions and attempts to blend in content that is not present in the source image.

Figure S1. **Mask-Guided Context-Aware Blending.** a) shows the low quality input image, b) shows restored image if PerFuSe is used without context-aware blending, and c) shows the results if PerFuSe is used with context-aware blending.

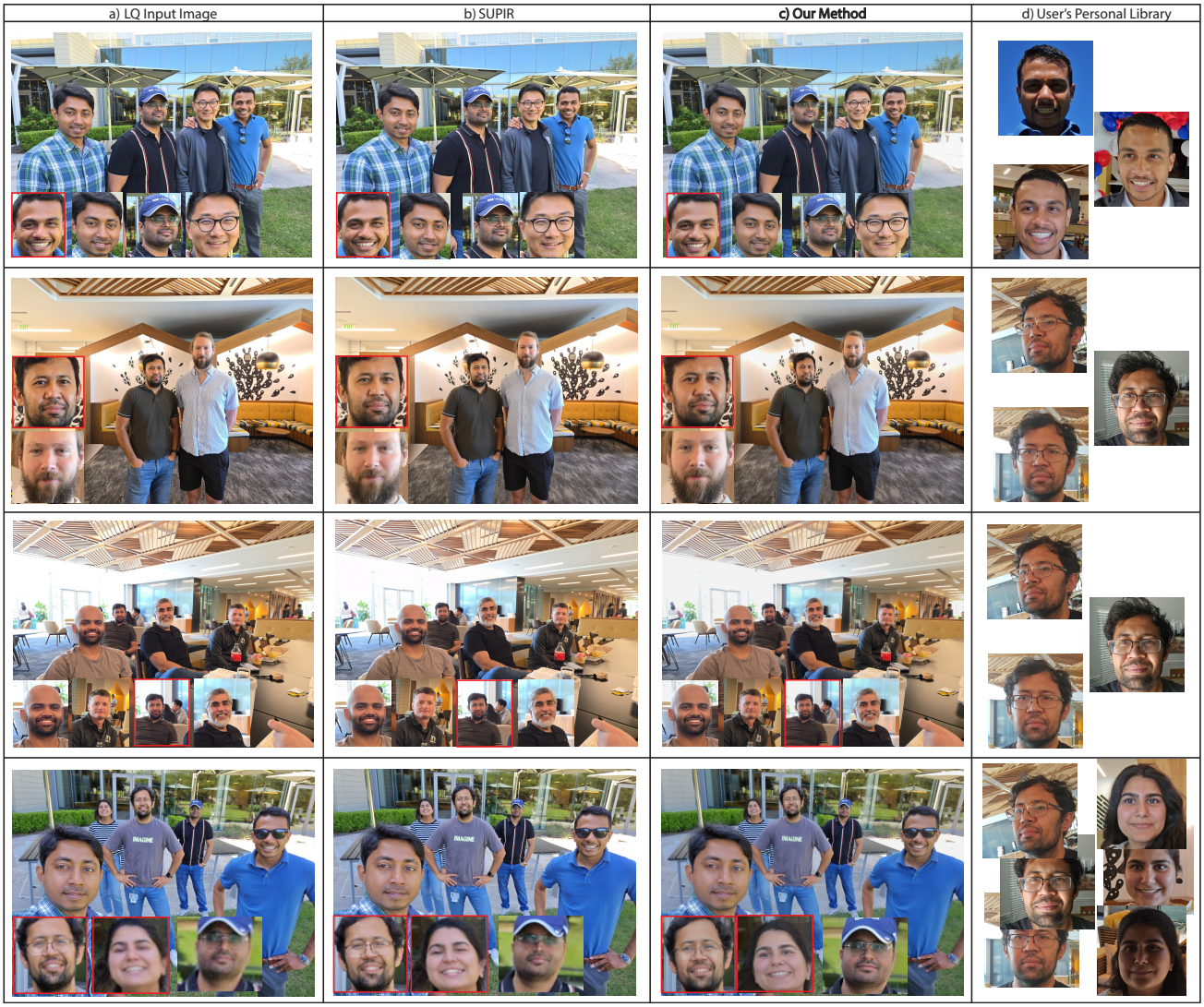

Figure S2. **Qualitative Results.** shows the end-to-end qualitative results on a 2× super resolution task. a) shows input low quality images, b) shows the reconstructed images from SUPIR [4], c) shows the reconstructed images from PerFuSe and d) shows the user's personal library used in each case.

2.2. Additional Full Image Qualitative Results

In this section, we present end-to-end qualitative results for the super-resolution task. Unlike the main paper, we include both zoomed-in crops and full-size images to facilitate direct comparison. Figure S2a shows the low-quality full input images alongside zoomed-in views of the faces. Figure S2b displays qualitative results from a state-of-the-art method [4]. Figure S2c presents the images restored by our method, and Figure S2d shows the corresponding example user libraries used for each case. Faces restored using the PFR modules are highlighted with red squares. Due to space constraints, we include only the strongest baseline from the main paper for comparison. These results clearly demonstrate the superior performance of PerFuSe compared with the state-of-the-art.

References

[1] Glenn Jocher, Ayush Chaurasia, and Jing Qiu. Ultralytics yolov8, 2023. 1

[2] Davis E. King. Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research*, 10:1755–1758, 2009. 1

[3] Xi Shen, François Darmon, Alexei A Efros, and Mathieu Aubry. Ransac-flow: generic two-stage image alignment. In *European Conference on Computer Vision*, pages 618–637. Springer, 2020. 1

[4] Fanghua Yu, Jinjin Gu, Zheyuan Li, Jinfan Hu, Xiangtao Kong, Xintao Wang, Jingwen He, Yu Qiao, and Chao Dong. Scaling up to excellence: Practicing model scaling for photo-realistic image restoration in the wild. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 25669–25680, 2024. 2