Hyperrealistic Image Inpainting with Hypergraphs
Supplementary Material

Gourav Wadhwa[1] Abhinav Dhall[2,1] Subrahmanyam Murala[1] Usman Tariq[3]
Indian Institute of Technology, Ropar[1] Monash University[2] American University of Sharjah[3]
{2017eeb1206, murala}@iitrpr.ac.in abhinav.dhall@monash.edu utariq@aus.edu

1. Additional Results

Here, we give some additional results from our network. We give the additional results on all four publicly available datasets, including CelebA-HQ [3], Places2 [2], Paris Street View [1] and Facades Dataset [4]. The results are provided for both center and irregular mask for all the datasets.

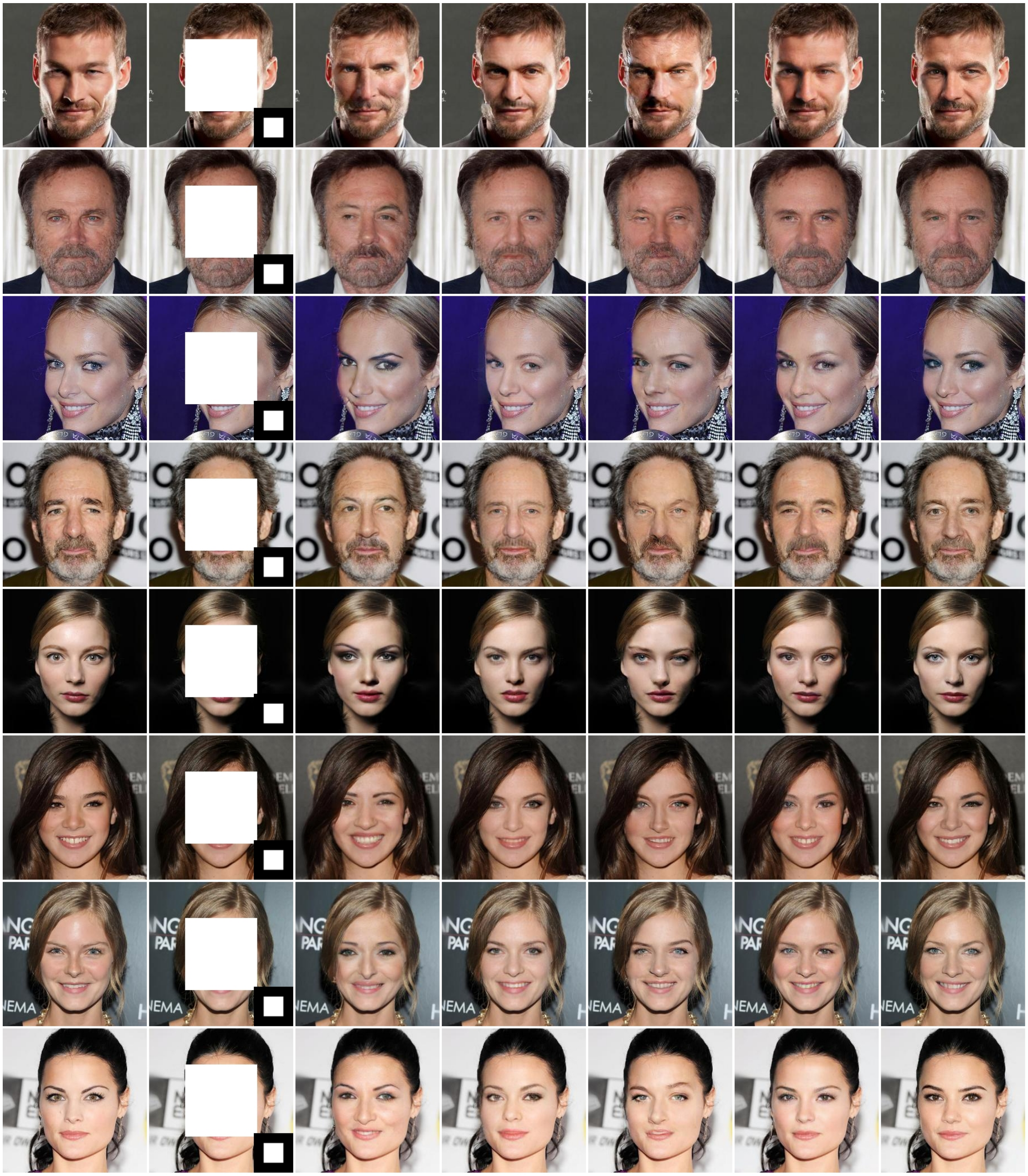
CelebA-HQ [3]: In Figure-1, we compare our method with the following state-of-the-art methods for center mask: a) pluralistic image completion (PICNet) [8], b) generative multi-column (GMCNN) [5], c) DeeFill-v2 [7], and d) Shift-Net (SN) [6]. Next, in Figure-2 and Figure-3 we compare our results for irregular mask..

Places2 [2]: In Figure-4, we evaluate our results on Places2 dataset for center mask. In Figure-5 and Figure-6 we compare our results with the state-of-the-art methods on random mask. We compare with a) pluralistic image completion (PICNet) [8], b) generative multi-column (GMCNN) [5], c) DeeFill-v2 [7], and d) Shift-Net (SN).

Paris Street View [1]: In Figure-7, we evaluate our results on Paris Street View dataset for center mask, and in Figure-8 we evaluate our results for irregular mask. The hole percentage in irregular mask is 40-60%.

Facades [4]: In Figure-9, we evaluate our results on Facades datasets with center mask, and in figure-10 we evaluate our results on irregular mask. Please note that we fine tune the model trained on Paris Street View datasets for evaluating on Facades Datasets.

We resize all the images to the scale of 256×256 for both training and evaluation purposes. The results are presented from the next page.



(a) Ground Truth (b) Input Image (c) PICNet (d) GMCNN (e) DeepFillV2 (f) SN (g) Ours

Figure 1. Qualitative Results on CelebA-HQ dataset with center mask. From left to right $a)$ Ground Truth, $b)$ Input Image, $c)$ Pluralistic (PICNet) [8], $d)$ GMCNN [5], $e)$ DeepFill-V2 [7], $f)$ Shift-Net (SN) [6], $g)$ Ours. All image are scaled to 256×256.

(a) Ground Truth (b) Input Image (c) PICNet (d) GMCNN (e) DeepFillV2 (f) SN (g) Ours

Figure 2. Qualitative Results on CelebA-HQ dataset with random mask. From left to right $a)$ Ground Truth, $b)$ Input Image, $c)$ Pluralistic (PICNet) [8], $d)$ GMCNN [5], $e)$ DeepFill-V2 [7], $f)$ Shift-Net (SN) [6], $g)$ Ours. All image are scaled to 256×256.

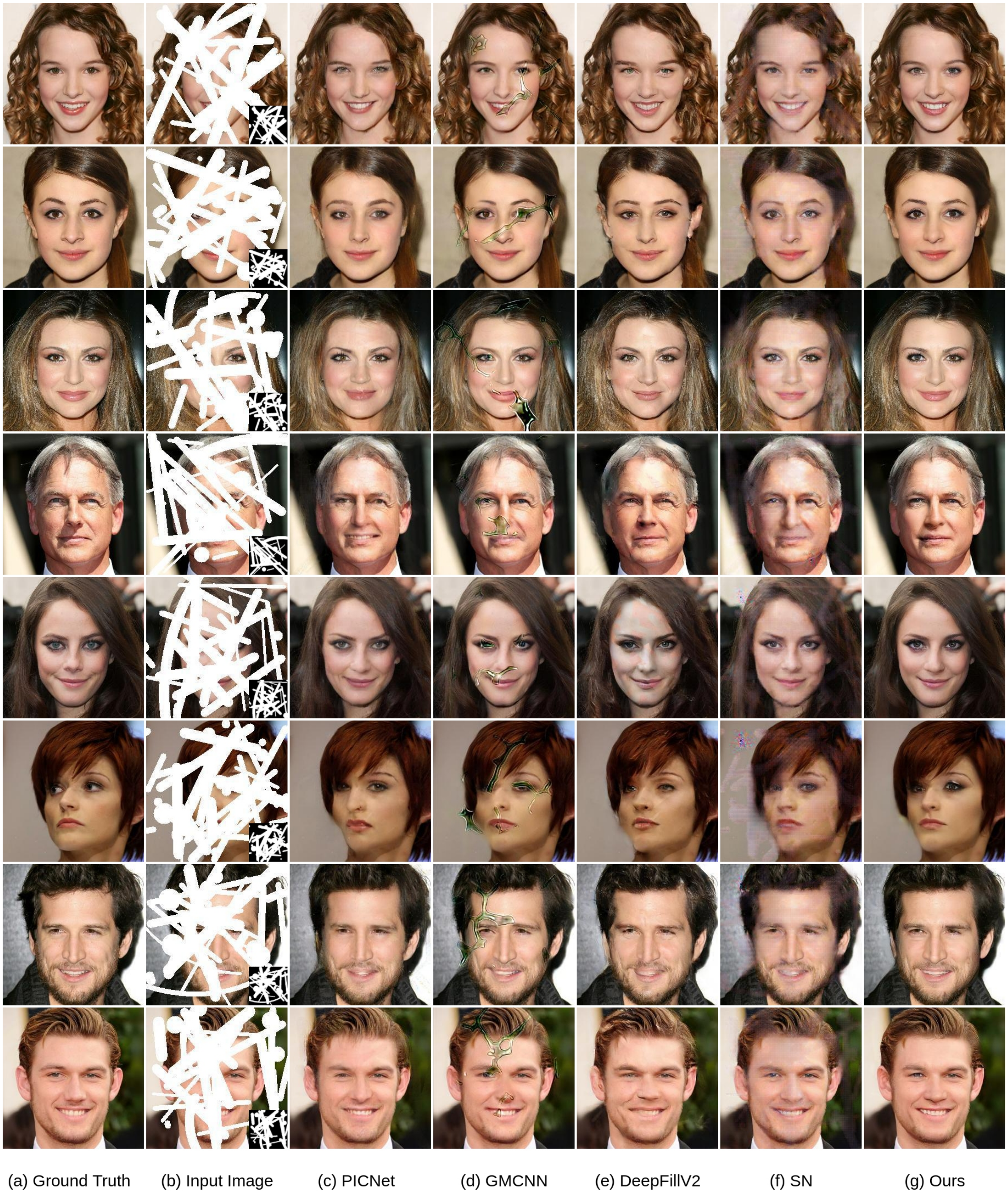
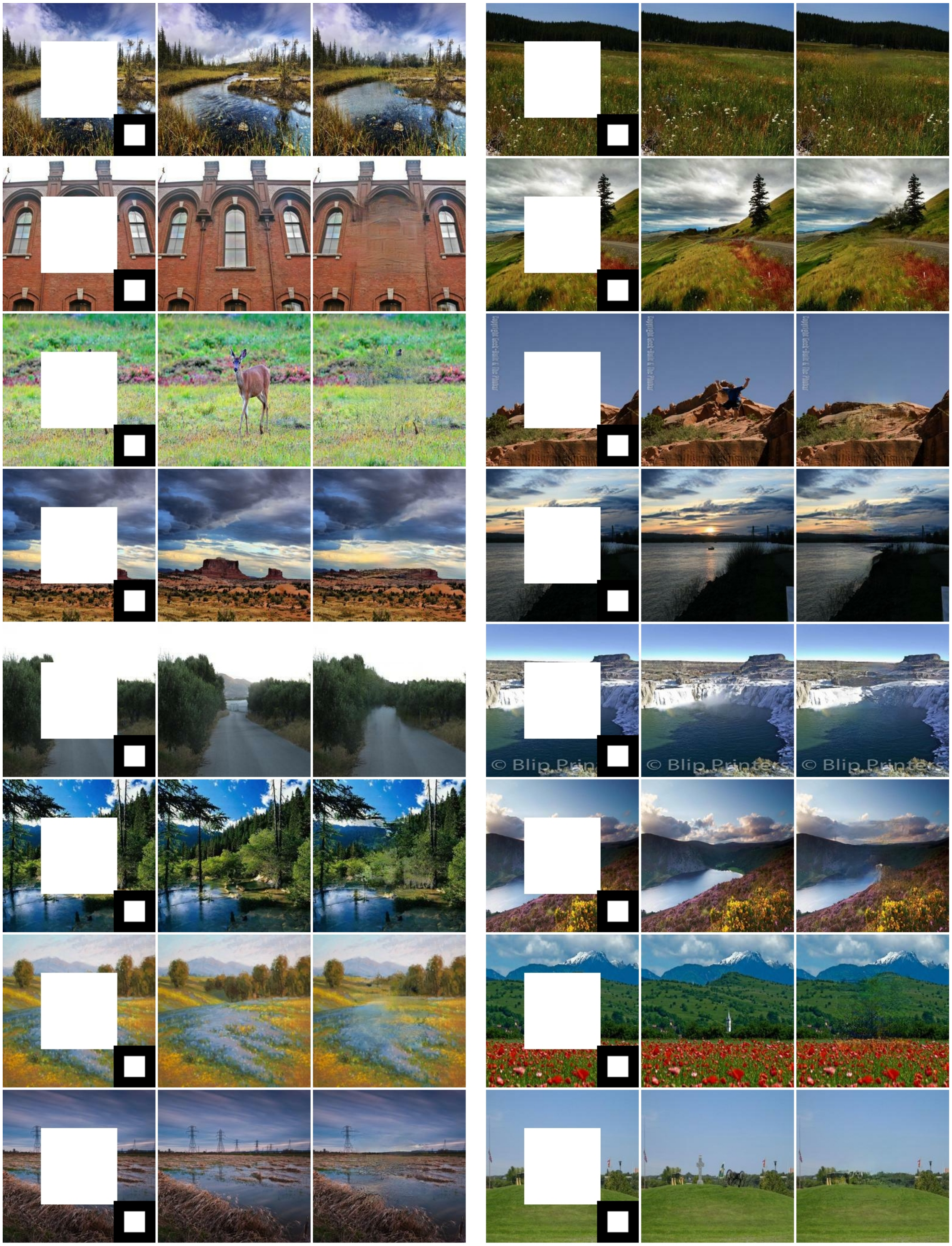
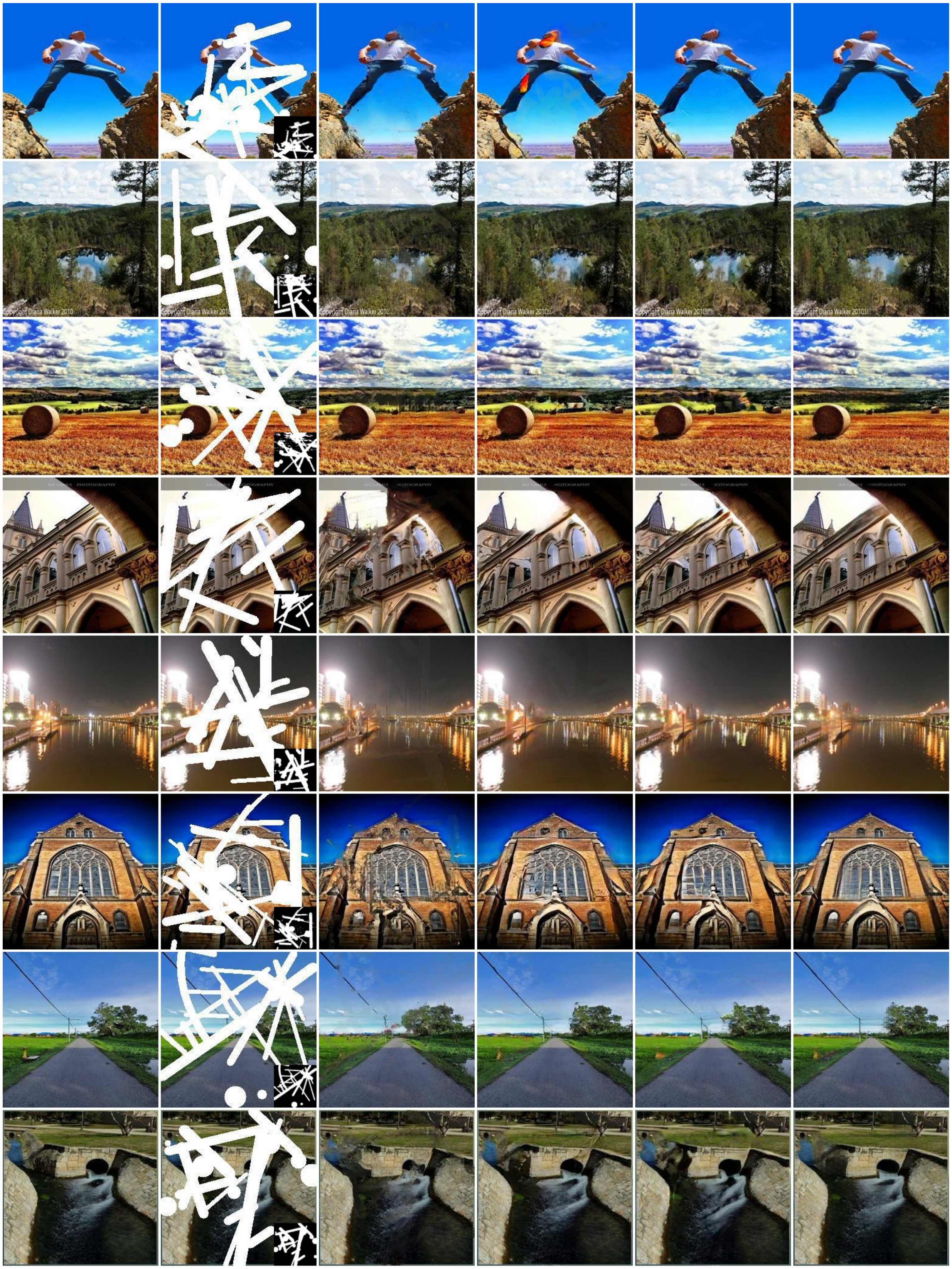


Figure 3. Qualitative Results on CelebA-HQ dataset with random mask. From left to right $a)$ Ground Truth, $b)$ Input Image, $c)$ Pluralistic (PICNet) [8], $d)$ GMCNN [5], $e)$ DeepFill-V2 [7], $f)$ Shift-Net (SN) [6], $g)$ Ours. All image are scaled to 256×256.



(a) Input Image (b) Ground Truth (c) Predicted (a) Input Image (b) Ground Truth (c) Predicted

Figure 4. Qualitative Results on Places2 dataset with center mask. From left to right $a)$ Input Image, $b)$ Ground Truth, $c)$ Ours. All image are scaled to 256×256.



a) GT b) Input Image c) PICNet d) GMCNN e) DeepFill-V2 f) Ours

Figure 5. Qualitative Results on Places2 dataset with random mask. From left to right $a)$ Ground Truth, $b)$ Input Image, $c)$ Pluralistic (PICNet) [8], $d)$ GMCNN [5], $e)$ DeepFill-V2 [7], $f)$ Ours. All image are scaled to 256×256.

Figure 6. Qualitative Results on Places2 dataset with random mask. From left to right $a)$ Ground Truth, $b)$ Input Image, $c)$ Pluralistic (PICNet) [8], $d)$ GMCNN [5], $e)$ DeepFill-V2 [7], $f)$ Ours. All image are scaled to 256×256.

(a) Input Image (b) Ground Truth (c) Predicted (a) Input Image (b) Ground Truth (c) Predicted

Figure 7. Qualitative Results on Paris Street View dataset with center mask. From left to right $a)$ Input Image, $b)$ Ground Truth, $c)$ Ours. All image are scaled to 256×256.

Figure 8. Qualitative Results on Paris Street View dataset with random mask. From left to right a) Input Image, b) Ground Truth, c) Ours. All image are scaled to 256×256.

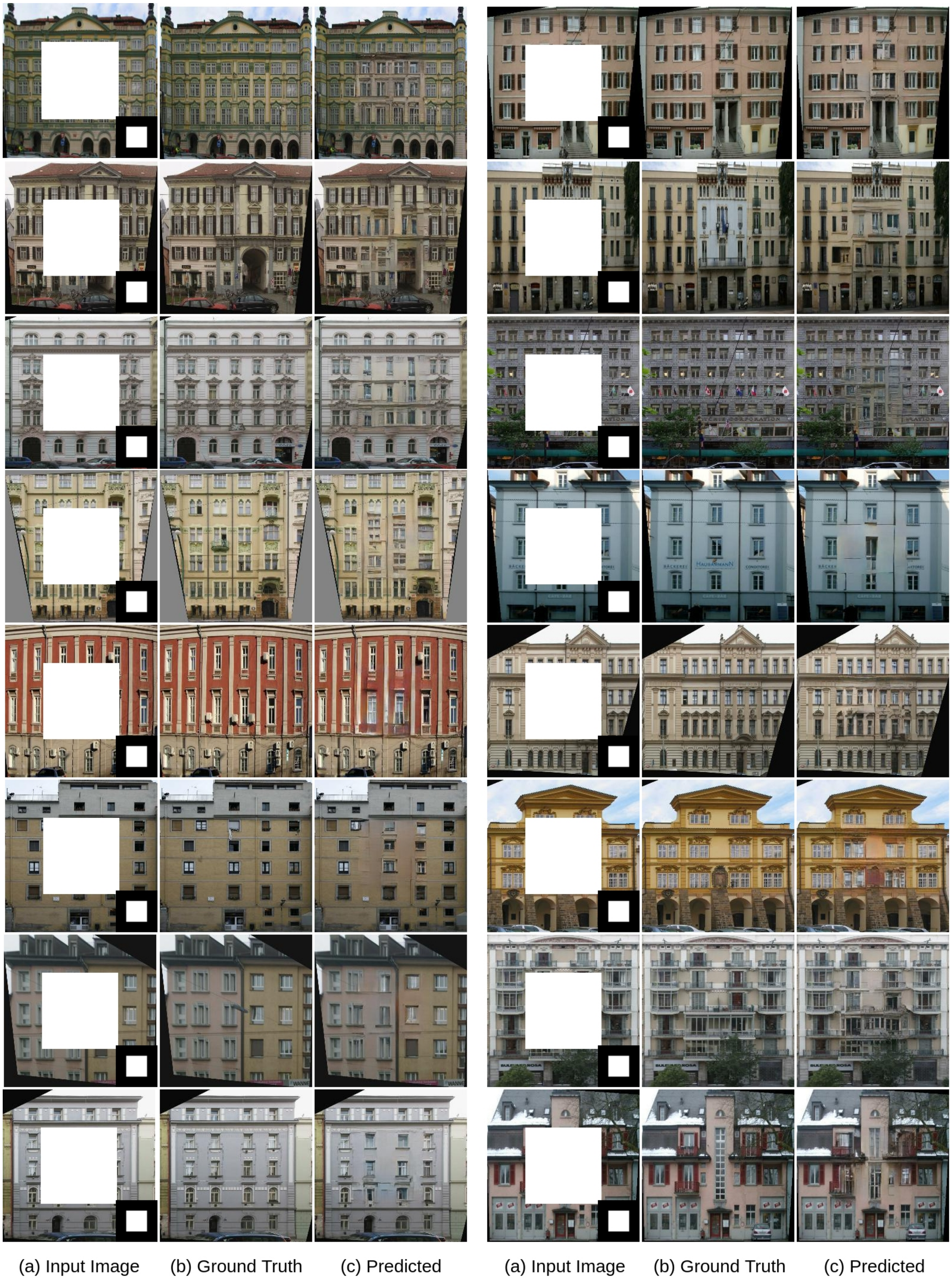


Figure 9. Qualitative Results on Facades dataset with center mask. From left to right $a)$ Input Image, $b)$ Ground Truth, $c)$ Ours. All image are scaled to 256×256.

(a) Input Image (b) Ground Truth (c) Predicted (a) Input Image (b) Ground Truth (c) Predicted

Figure 10. Qualitative Results on Facades dataset with random mask. From left to right $a)$ Input Image, $b)$ Ground Truth, $c)$ Ours. All image are scaled to 256×256.

References

[1] Zhou B., Lapedriza A., Khosla A., A. Oliva, and Torralba A. Places: A 10 million image database for scene recognition. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017.

[2] Doersch, A C. Singh, S. Gupta, Sivic J., and A Efros. What makes paris look like paris? In *ACM Transactions on Graphics 31(4)*, 2012.

[3] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of gans for improved quality stability, and variation. *arXiv preprint arXiv:1710.10196*, 2017.

[4] Radim Tyleček and Radim Šára. Spatial pattern templates for recognition of objects with regular structure. In *Proc. GCPR*, Saarbrucken, Germany, 2013.

[5] Yi Wang, Xin Tao, Xiaojuan Qi, Xiaoyong Shen, and Jiaya Jia. Image inpainting via generative multi-column convolutional neural networks. In *Advances in Neural Information Processing Systems*, pages 331–340, 2018.

[6] Zhaoyi Yan, Xiaoming Li, Mu Li, Wangmeng Zuo, and Shiguang Shan. Shift-net: Image inpainting via deep feature rearrangement. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.

[7] Jiahui Yu, Zhe Lin, Jimei Yang, Xiaohui Shen, Xin Lu, and Thomas S Huang. Free-form image inpainting with gated convolution. *arXiv preprint arXiv:1806.03589*, 2018.

[8] Chuanxia Zheng, Tat-Jen Cham, and Jianfei Cai. Pluralistic image completion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.