PDA-RWSR: Pixel-Wise Degradation Adaptive Real-World Super-Resolution
Supplementary Material

Andreas Aakerberg[1], Majed El Helou[2], Kamal Nasrollahi[1,3], Thomas Moeslund[1]
[1] Aalborg University, Denmark, [2] ETH Zürich, Switzerland, [3] Milestone Systems, Denmark
anaa,tbm,kn@create.aau.dk, melhelou@ethz.ch

1. Spatially Variant Super-Resolution (SVSR) Dataset

This section presents additional details about the cameras and setup used to collect the SVSR dataset. The dataset is self-contained and does not contain other assets such as labels or metadata. We have manually screened the dataset to ensure that no images in the dataset enable the identification of individuals or contain content that could be considered offensive, insulting, or threatening. Examples of the images in the dataset can be seen in Figure 2 and Figure 3. Furthermore, Figure 1 shows examples of the spatially variant distribution of the noise in the dataset.

Hardware and settings: We use the following equipment to collect images for the dataset:

- Canon 600D, 18MPIX APC-C DSLR camera (Released in 2011). ISO range: 100 - 6400.
- Canon 1Ds Mark II, 16.6MPIX full-frame DSLR camera (Released in 2004). ISO range: 100–1600, plus 50 and 3200 as option.
- Canon 6D, 20MPIX full-frame DSLR camera(Released in 2012). ISO range: 100 – 25,600, plus 50, 51.200, and 65535 as option.
- Canon EF 28-135mm F3.5-5.6 IS USM lens
- Canon EF 70-300mm F4-5.6 L IS USM lens

We turn off any built-in noise reduction mechanisms in the cameras if possible and save the images in high-quality JPEG. The different lenses have different point-spread functions resulting in different blurring of the images. However, we also introduce more variance by capturing images with an equal balance of three different f-stop values, namely f/10, f/13, and f/16. For the 28-135mm lens, we obtain the $\times 1$, $\times 2$, and $\times 4$ scale differences at 28,56 and 112mm focal lengths, respectively. Similarly, for the 70-300mm lens, we use 70,140,280mm.

File naming: After pre-processing, we store and name the images in the dataset as PNG files using the following syntax:
`id_cameratype_lenstype_f-stop_domain_iso.png`
Where id refers to a sequence of Low-Resolution (LR) images captured with different ISO settings and the corresponding noise-free $\times 2$ and $\times 4$ High-Resolution (HR) images, and domain refers to the image being either a LR, X2- or X4-HR image.

Distribution: The dataset is publicly accessible via Zenodo under the Creative Commons Attribution Non Commercial Share Alike 4.0 International license at: https://doi.org/10.5281/zenodo.1004426. As such, the dataset is freely available to academic and non-academic entities for non-commercial purposes such as academic research, teaching, scientific publications, or personal experimentation.

2. Spatially Variant Degradation Model

This section describes the method and parameters used for generating masks for alpha blending the clean and noisy LR training images. We use four different mask types, namely linear, log, and radial gradients, and a thresholding-based mask. The linear gradients are generated with the Numpy.linspace function using a random start value between 0.0-1.0 and an end value of 1.0. Similarly, for the log gradient we use the Numpy.geomspace function with a random start value between 0.001 and 0.2 and an end value of 1.0. For the radial gradient, we generate two linear gradients orthogonal to each other. We use a starting value of -1.0 and random end values between 0.0-3.0. For masks based on thresholding, we perform binary greyscale

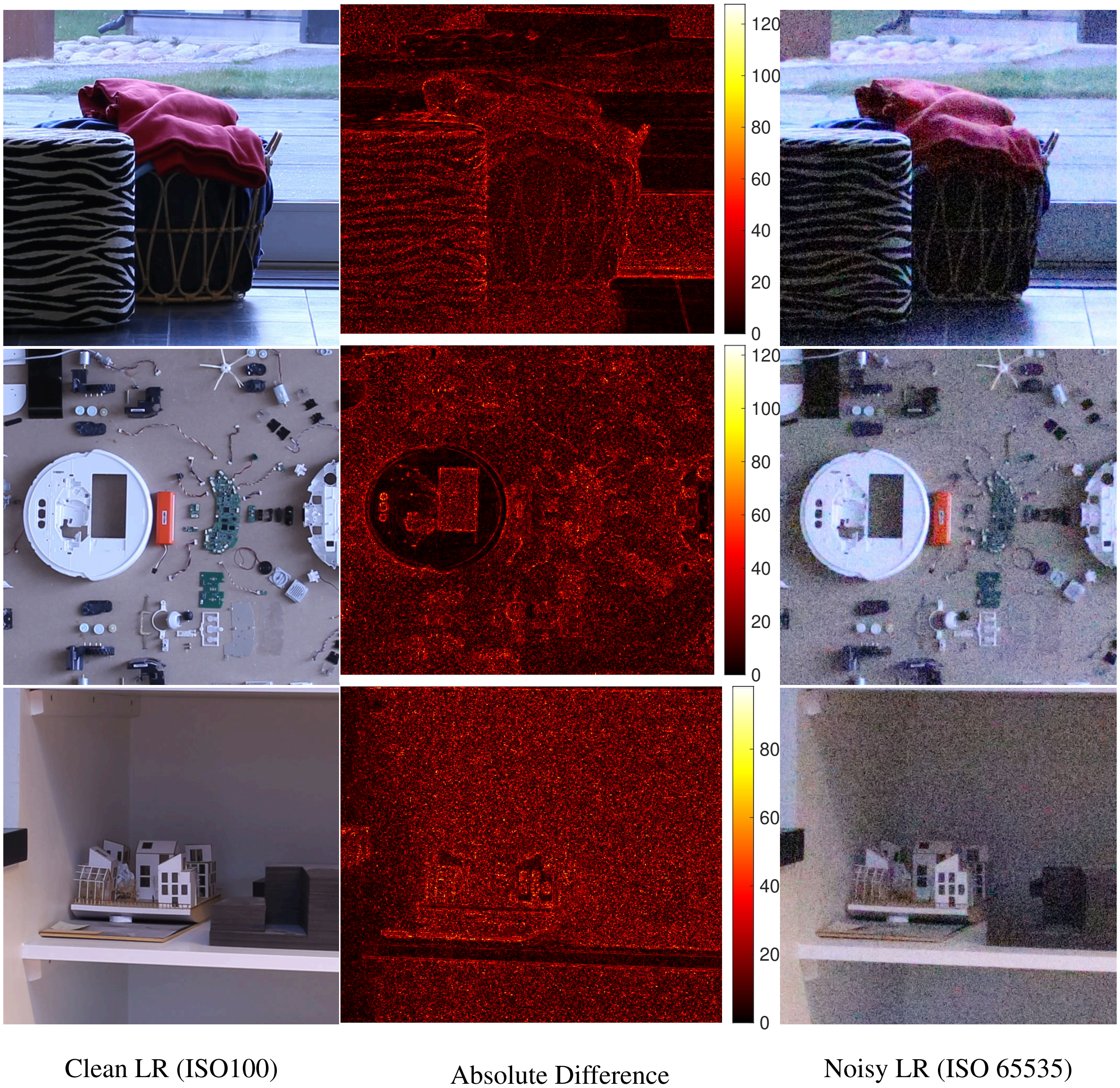


Figure 1. Visualization of the color-channel average absolute distance in LR space between a noisy and clean image pairs from the SVSR dataset. As seen, more noise is present in the darker regions of the noisy images.

thresholding with values between 0.1-0.5 combined with a uniform value between 0.3-0.8 to introduce spatial variance while ensuring that all parts of the image will contain some degree of noise. To smooth the masks, for a softer transition between low- and high-noise areas, we subsequently perform Gaussian blurring with a kernel size of 9×9. All masks are generated at 1.5 higher spatial resolution than the LR image and rotated at a random angle between 0-356 using a random selection between reflect, mirror, and nearest rotation modes with the Scipy.ndimage.rotate function. Finally, the masks are resized to match the LR image and used as pixel-wise alpha blending values.

3. Results

In Figure 4 and Figure 5 we show results of Super-Resolution (SR) of the synthetically degraded images used in the experimental section of our work. As seen, our method produces the most detail-rich reconstructions while removing the largest degree of undesired artifacts (noise).

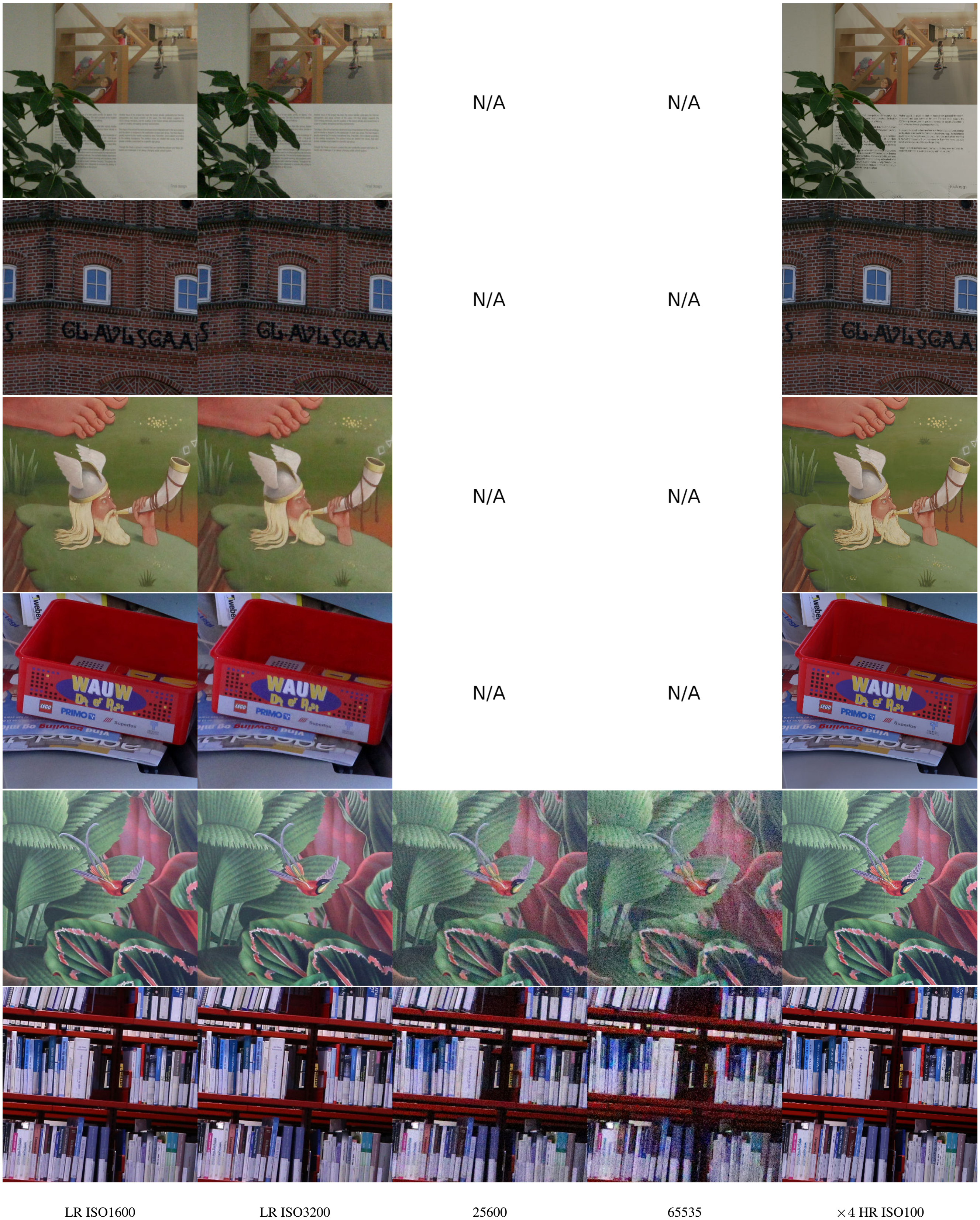
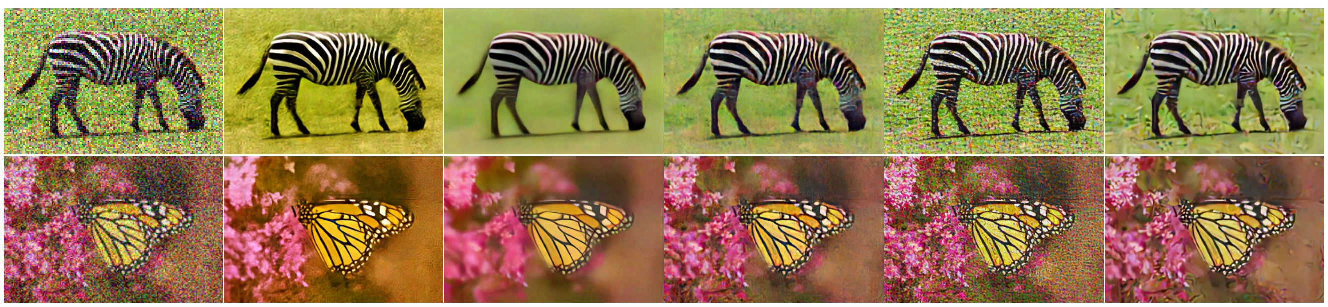


Figure 2. Examples of images from the SVSR dataset. Row 1-2: Samples from Canon 1DsII, Row 3-4: Samples from Canon 600D, Row 5-6: Samples from Canon 6D

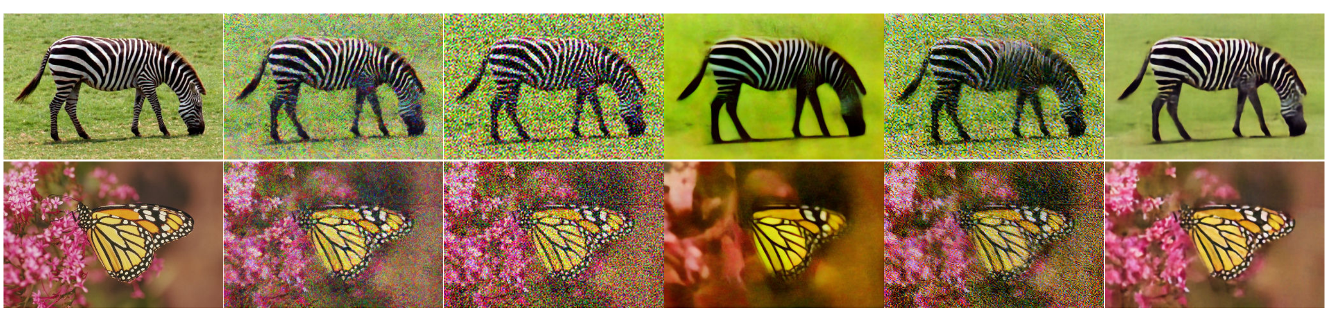
$\times 1$ LR $\quad$ $\times 2$ HR $\quad$ $\times 4$ HR

Figure 3. Examples of the scale difference in the SVSR dataset (ISO100)



LR $\quad$ SwinIRNet-L [2] $\quad$ RealESRNet [8] $\quad$ MM-RealSR [6] $\quad$ FeMASR [1] $\quad$ DASR [7]

Figure 4. Results on images from Set14 degraded with $\times 4$ downscaling and additive Gaussian noise with $\sigma = 50$.



GT $\quad$ DASR [3] $\quad$ DAN [5] $\quad$ BSRNet [9] $\quad$ PDM-SR [4] $\quad$ Ours

Figure 5. Results on images from Set14 degraded with $\times 4$ downscaling and additive Gaussian noise with $\sigma = 50$.

References

[1] Chaofeng Chen, Xinyu Shi, Yipeng Qin, Xiaoming Li, Xiaoguang Han, Tao Yang, and Shihui Guo. Real-world blind super-resolution via feature matching with implicit high-resolution priors. In *Proceedings of the 30th ACM International Conference on Multimedia*, pages 1329–1338, 2022. 4

[2] Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. Swinir: Image restoration using swin transformer. In *IEEE International Conference on Computer Vision Workshops*, 2021. 4

[3] Jie Liang, Hui Zeng, and Lei Zhang. Efficient and degradation-adaptive network for real-world image super-resolution. In *European Conference on Computer Vision*, 2022. 4

[4] Zhengxiong Luo, Yan Huang, , Shang Li, Liang Wang, and Tieniu Tan. Learning the degradation distribution for blind image super-resolution. In *CVPR*, 2022. 4

[5] Zhengxiong Luo, Yan Huang, Shang Li, Liang Wang, and Tieniu Tan. Unfolding the alternating optimization for blind super resolution. *Advances in Neural Information Processing Systems (NeurIPS)*, 33, 2020. 4

[6] Chong Mou, Yanze Wu, Xintao Wang, Chao Dong, Jian Zhang, and Ying Shan. Metric learning based interactive modulation for real-world super-resolution. In *European Conference on Computer Vision*, pages 723–740. Springer, 2022. 4

[7] Longguang Wang, Yingqian Wang, Xiaoyu Dong, Qingyu Xu, Jungang Yang, Wei An, and Yulan Guo. Unsupervised degradation representation learning for blind super-resolution. In *CVPR*, 2021. 4

[8] Xintao Wang, Liangbin Xie, Chao Dong, and Ying Shan. Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 1905–1914, 2021. 4

[9] Kai Zhang, Jingyun Liang, Luc Van Gool, and Radu Timofte. Designing a practical degradation model for deep blind image super-resolution. In *IEEE International Conference on Computer Vision*, pages 4791–4800, 2021. 4