Expanding Synthetic Real-World Degradations for Blind Video Super Resolution –Supplementary Material–

Mehran Jeelani[1,2,*], Sadbhawna[4,*], Noshaba Cheema[2,3,5], Klaus Illgner-Fehns[1], Philipp Slusallek[2,5], and Sunil Jaiswal[1,†]

[1] K|Lens GmbH, Germany, [2] Saarland Informatics Campus, Germany, [3] MPI Informatics, Germany
[4] IIT Jammu, India, [5] German Research Center for Artificial Intelligence (DFKI), Germany

1. Blur Kernel Pool (B_{real}):

Recall that the proposed algorithm creates a blur kernel pool from real-world images. We propose to use KernelGAN [2] for this purpose and create a kernel pool of approximately 5000 kernels from a dataset of 5000 images. These 5000 images are collected from DF2K [8], K|Lens datasets [1, 4]. DF2K is an image dataset, whereas K|Lens datasets consist of videos.

Please note that in the training of our algorithm, for creating an LR-HR pair, we randomly selected one blur kernel out of these 5000 blur kernels, in addition to isotropic and anisotropic blur kernels for simulation.

2. Proposed K|Lens Dataset:

The unique optical lens developed by K|Lens [1, 4] enables any camera with exchangeable lenses to capture multiple perspectives of a scene with a single exposure as regular color images on the camera sensor. More specifically, it captures nine different perspectives of the same scene and also it can record videos. Please refer to [1, 4] for more details.

3. Quality Assessment Metrics:

We have used two no-reference quality assessment metrics for quantitative comparison in the paper, i.e., NRQM [5], and BRISQUE [6]. BRISQUE is a widely used metrics originally proposed for the quality evaluation of natural images. These methods rely upon natural scene statistical (NSS) features extracted from local image patches to calculate the quality of the distorted image. The BRISQUE is trained on features obtained from natural and distorted images and human judgments.

NRQM [5] is specifically designed to predict the quality scores of super-resolved images. They proposed to use three types of low-level statistical features in both spatial and frequency domains. These features are learned using a two-stage regression model to predict the quality scores of the super-resolved image without referring to ground-truth images. Extensive experimental analysis has been done in the [5] to compare NRQM with the existing no-reference IQA metrics, including BRISQUE metric. In terms of Spearman Rank Correlation Coefficient (SRCC), [7], the NRQM metric is better than all the compared metrics, as suggested in the [5]. Recall that we extensively compared different VSR algorithms in Tables 1, 2 and 3 of the manuscript, and we can observe that the proposed SRWD-VSR outperforms all the compared models in terms of NRQM.

4. Pre-trained weights of RealBasicVSR [3]

Recall that in the simulation results and experiments (Table 3 of the main paper), we have not shown the results [3] for 2 $\times$ scaling as we do not have the pre-trained weights. This is why visual results in Figure 8 of the main paper are only based on 4 $\times$ scaling.

References

[1] https://www.k-lens.de/ rgbdx - the 4th dimension in machine vision. 1

[2] S. Bell-Kligler, A. Shocher, and M. Irani. Blind super-resolution kernel estimation using an internal-gan. CoRR, Red Hook, NY, USA, 2019. 1

[3] K. Chan, S. Zhou, X. Xu, and C. C. Loy. Investigating tradeoffs in real-world video super-resolution. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2022. 1

[4] K. Illgner, J. Restrepo, S. P. Jaiswal, and I. Ihrke. Light-field imaging for industrial applications. In Masafumi Kimata, Joseph A. Shaw, and Christopher R. Valenta, editors, *SPIE Future Sensing Technologies*, volume 11525, page 1152526. International Society for Optics and Photonics, SPIE, 2020. 1

This work was partially funded by the German Ministry for Education and Research (BMBF) under the grant PLIMASC and Indo-German Science and Technology Center (IGSTC).

*Both authors contributed equally to the paper
†Corresponding author: sunil.jaiswal@k-lens.de

[5] C. Ma, C. Yang, X. Yang, and M. Yang. Learning a no-reference quality metric for single-image super-resolution. *Computer Vision and Image Understanding*, 12 2016. 1

[6] A. Mittal, A. K. Moorthy, and A. C. Bovik. Blind/referenceless image spatial quality evaluator. In *2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, pages 723–727, 2011. 1

[7] A. Craig R. Hogg, J. McKean. Introduction to mathematical statistics. *Pearson Education*, 2005. 1

[8] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. C. Loy. Esrgan: Enhanced super-resolution generative adversarial networks. In *The European Conference on Computer Vision Workshops (ECCVW)*, September 2018. 1