

# Supplementary Material for Metric Learning based Interactive Modulation for Real-World Super-Resolution

Anonymous ECCV submission

Paper ID 968

In this supplementary material, we provide the following additional details to facilitate the understanding of our main paper:

## 1 Detailed Illustration of the Comparison between MM-RealSR and Existing Modulation-based Methods.

There are two main differences between our proposed MM-RealSR and existing modulation-based methods. **First**, existing methods are limited to be trained with simple degradations and known degradation levels. They cannot address the modulation problem for real-world images with complex and unknown degradations. **Second**, existing methods (*e.g.*, the recent CResMD [2] and CUGAN [1]) must be manually provided with degradation parameters for adjustment, which is over-complicated in processing abundant images. In this paper, we also propose an unsupervised estimation degradation module to automatically estimate a most suitable degradation level. The users are then able to modulate the restoration strength based on the estimated degradation levels. Equipped with the unsupervised degradation estimation module (UDEM), our MM-RealSR can flexibly switch between modulation and non-modulation modes for restoration.

To compare our MM-RealSR with the recent CUGAN on the RWSR task in our main paper, we choose the best result (lowest LPIPS) of CUGAN (with different degradation levels as inputs) on each image to make a comparison. Specifically, we uniformly divide the noise and blur degradation levels ( $S_n, S_b \in [0, 1]$ ) into 11 points and then traverse all cases (121 cases in total). By contrast, the input degradation scores ( $S_n, S_b$ ) of our MM-RealSR are generated by the UDEM. The quantitative comparison has been shown in *Tab. 1* in our main paper, presenting obvious performance gains. Here, we further present the visual comparison in Fig. 1, showing the better restoration quality of our MM-RealSR. Note that the first row is the set of low-quality images upsampled by the bicubic operation. The second row and the third row present the restoration results of CUGAN and our method, respectively.

## 2 More Comparisons with Non-modulation Methods

More visual comparisons between our MM-RealSR and several non-modulation methods on RWSR is presented in Fig. 2. Note that the “Bicubic  $\times 4$ ” in this



**Fig. 1.** Visual comparison between our proposed MM-RealSR (the third row) and CUGAN [1] (the second row) on the RWSR ( $\times 4$ ) task. The first row presents low-quality images upsampled by the bicubic operation.

figure represents upsampling the LR image with the Bicubic algorithm. We can observe that our MM-RealSR can achieve better restoration quality with vivid details, demonstrating that our approach can estimate satisfactory degradation scores to guide the restoration process in real-world scenarios.

### 3 Time Complexity

In our main paper, we demonstrate the excellent RWSR and interactive modulation performance of our proposed method. In this part, we present the time complexity of our method in Tab. 1. All inference speeds are evaluated on one Nvidia TESLA V100 GPU. We can find that our method has a comparable time complexity to top-performing methods. Note that these top-performing methods use the same backbone (RRDB), leading to the same time complexity.

**Table 1.** The inference time (s) that different methods take to process a  $512 \times 512$  RGB image.

Method	RealSR	ESRGAN	BSRGAN	Real-ESRGAN	Ours
Time (s)	0.66	0.66	0.66	0.66	0.70

### 4 UDEM Evaluation in High-order Degradation

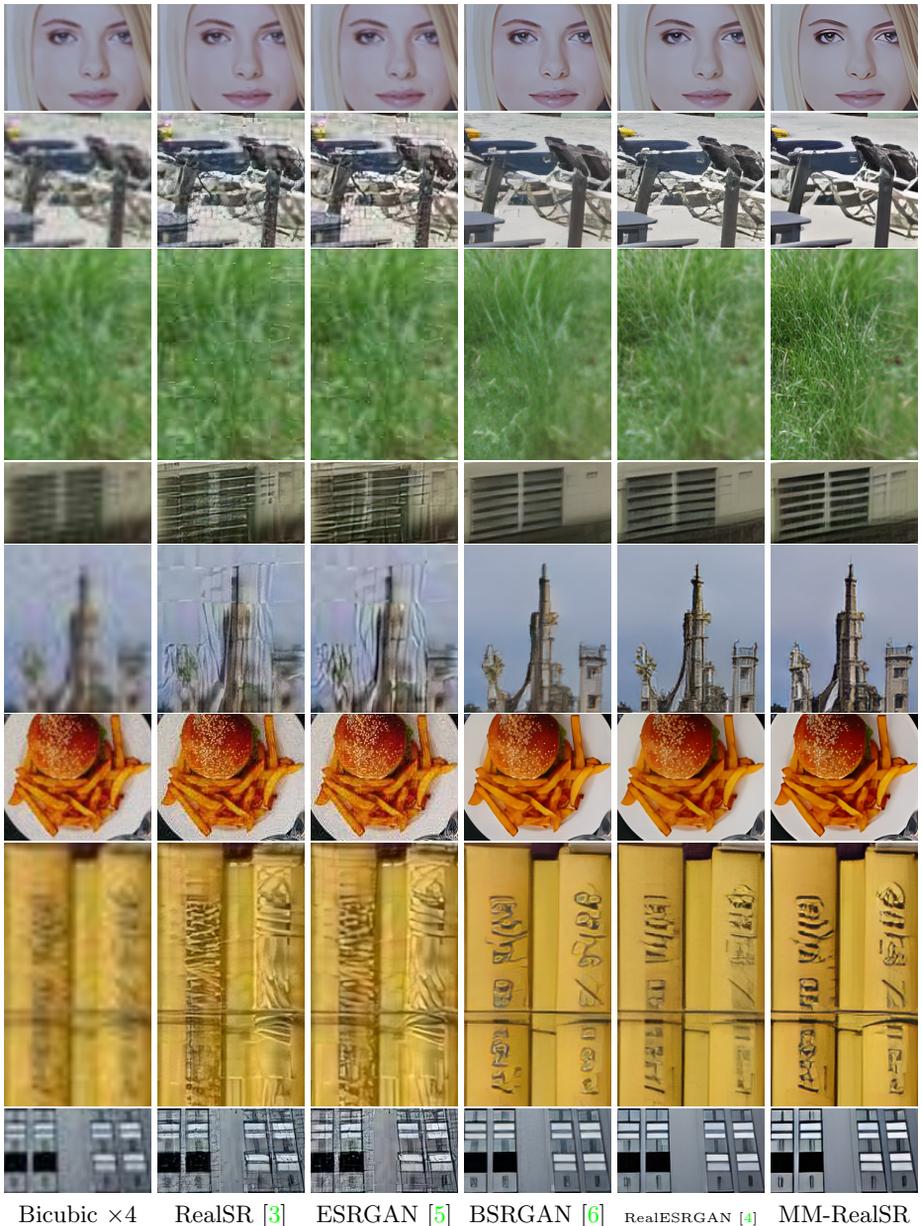
We evaluate UDEM under the high-order degradation, as shown in Fig. 3. The evaluation samples are generated by monotonically changing one degradation factor and fixing other factors in the high-order degradation. Note that the x-axis is unquantifiable, and the values just represent the index of 20 monotonic degradation points without real significance. Like the results in the first-order case, with the increase of degradation level, the estimated degradation score increases monotonously. We will add this analysis to our paper.

## 5 UDEM Evaluation with Resize Degradation

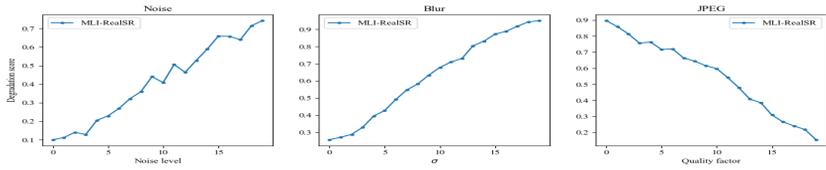
The score estimation of resizing-induced blur is shown in Fig. 4. The x-axis is the downsampling factor (*e.g.*, 0.5 is  $\times 2$  downsampling). The estimated degradation score increases monotonously with the increase of the resizing intensity.

## 6 More Visual Results of our Modulation Process

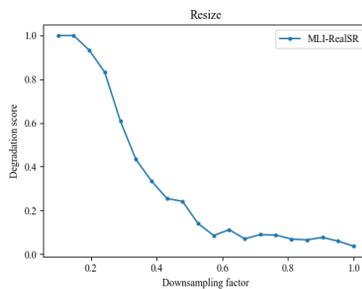
In this part, we provide more visualization results of the modulation process, which are presented in Fig. 5. Note that apart from evaluating the modulation performance within the default range  $S_n, S_b \in [0, 1]$  (labeled with yellow boxes) as mentioned in our main paper, we also show the modulation results beyond the default range (labeled with red boxes  $S_n, S_b \in [1, 2]$ ). From Fig. 5, we can find that our proposed MM-RealSR has excellent adjusting ability, which can provide several satisfactory results for user selection. Additionally, the direction of the modulation is well learned by our metric learning strategy. Concretely, when the degradation score  $S_n, S_b > 1$  the reconstruction results still conform to the expected direction of modulation.



**Fig. 2.** More visual comparisons between our proposed MM-RealSR and several recent methods on real-world super-resolution ( $\times 4$ ). The degradation scores are estimated by our UDEM without specific adjustment during inference. **Zoom in for best view**



**Fig. 3.** The evaluation of our UDEM in high-order degradation.



**Fig. 4.** The degradation estimation with resizing blur.

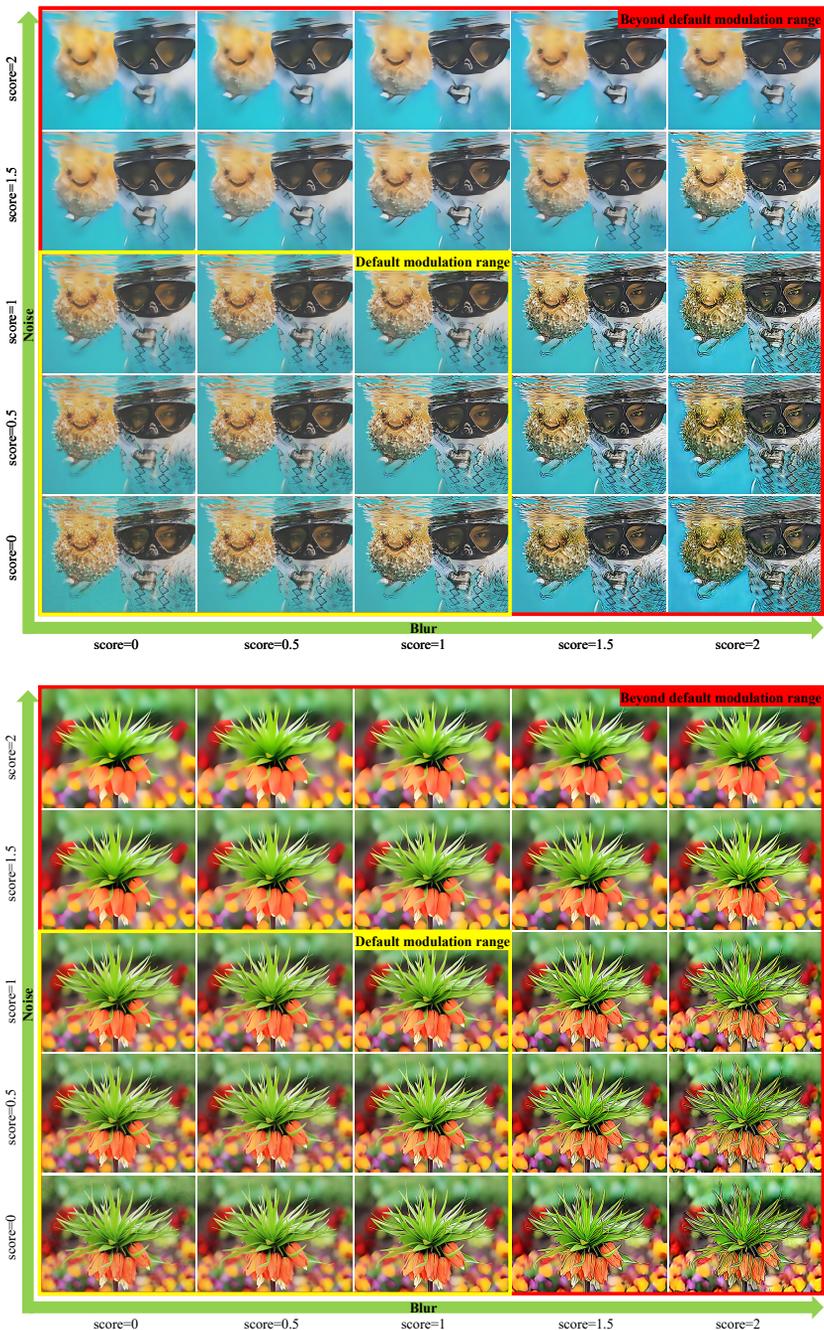


Fig. 5. More visual results of the modulation ability of our proposed MM-RealSR on real-world super-resolution ( $\times 4$ ). **Zoom in for best view**

## References

1. Cai, H., He, J., Qiao, Y., Dong, C.: Toward interactive modulation for photo-realistic image restoration. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 294–303 (2021)
2. He, J., Dong, C., Qiao, Y.: Interactive multi-dimension modulation with dynamic controllable residual learning for image restoration. In: Proceedings of the European Conference on Computer Vision (ECCV). pp. 53–68 (2020)
3. Ji, X., Cao, Y., Tai, Y., Wang, C., Li, J., Huang, F.: Real-world super-resolution via kernel estimation and noise injection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). pp. 466–467 (2020)
4. Wang, X., Xie, L., Dong, C., Shan, Y.: Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In: Proceedings of the International Conference on Computer Vision Workshops (ICCVW). pp. 1905–1914 (2021)
5. Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., Change Loy, C.: Esrgan: Enhanced super-resolution generative adversarial networks. In: Proceedings of the European Conference on Computer Vision Workshops (ECCVW). pp. 0–0 (2018)
6. Zhang, K., Liang, J., Van Gool, L., Timofte, R.: Designing a practical degradation model for deep blind image super-resolution. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV). pp. 4791–4800 (2021)