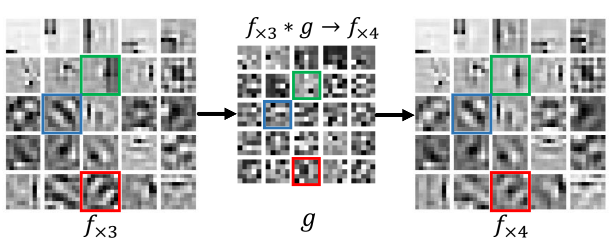
Modulating Image Restoration with Continual Levels via Adaptive Feature Modification Layers Supplementary File

Jingwen He[1,*] Chao Dong[1,*] Yu Qiao[1,2,†]
[1]ShenZhen Key Lab of Computer Vision and Pattern Recognition, SIAT-SenseTime Joint Lab,
Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China
[2]The Chinese University of Hong Kong

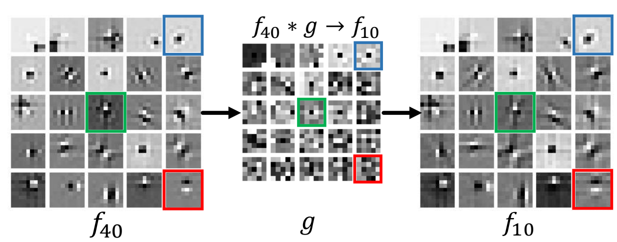
Abstract

In this supplementary file, we first show the filter visualization in super resolution, DeJPEG, and denoising. Then we add the curves showing performances with different filter sizes of AdaFM layers in the three restoration tasks. The artifacts produced by AdaBN and conditional instance normalization are presented in Section 3. In Section 4, we show the fitting curves and the corresponding fitting errors of modulation testing in the three restoration tasks. Finally, we provide more qualitative results about the modulation testing of our proposed AdaFM-Net.

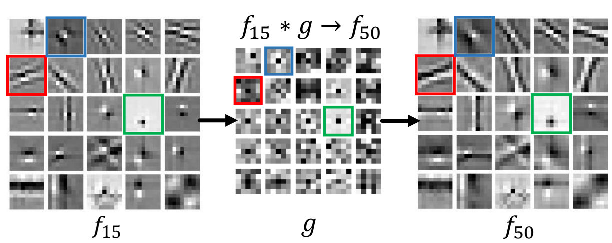
1. Filter Visualization



(a) Super Resolution: $\times3 \rightarrow \times4$



(b) DeJPEG: $q40 \rightarrow q10$



(c) Denoising: $\sigma15 \rightarrow \sigma50$

Figure 1: Filter Visualization

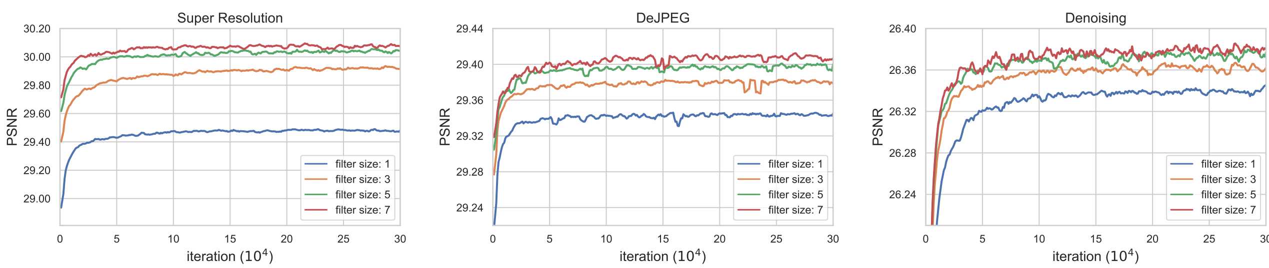
In the three restoration tasks, we find that the learned filters of restoration models trained with different restoration levels are pretty similar at visual patterns, as presented in Figure 1. In the problem of super resolution, the basic model is trained on the start level (upscaling factor 3) with SRCNN [3] structure, namely SRCNN-$\times3$. Then we finetune SRCNN-$\times3$ on the end level (upscaling factor 4) to obtain SRCNN-$\times4$. The filters of the first layers of both models are shown in Figure 1 (a). It can be obviously observed that the patterns of the corresponding filters are quite similar. Moreover, the mean consine distance

*The first two authors are co-first authors. (e-mail: jw.he@siat.ac.cn; dong.chao@siat.ac.cn).
†Corresponding author (e-mail: yu.qiao@siat.ac.cn).

between these two sets of filters is 0.11. In addition, the 5×5 filters that bridge filters of both ends are also visualized. Similarly, in DeJPEG, we obtain start ($q40$) and end level ($q10$) models based on ARCNN [2]. The corresponding filters of both ends as well as the 5×5 bridge filters are presented in Figure 1 (b). Figure 1 (c) shows the filters of denoising task which have been presented in the main paper.

2. Filter Size

In this section, we present the curves showing the performance with different filter sizes (1×1, 3×3, 5×5, and 7×7) in the three restoration tasks. The convergence curves in super resolution, DeJPEG and denoising on Set5 [1], LIVE1 [7], and CBSD68 [6] datasets are plotted in Figure 2, respectively. Generally, a larger filter size leads to better performance. Specially, in super resolution task $\times 3 \rightarrow \times 4$, as shown in Figure 2 (a), the PSNR gap between 1×1 and 3×3 is relatively large (more than 0.4 dB). However, in DeJPEG task $q80 \rightarrow q10$ and denoising $\sigma 15 \rightarrow \sigma 75$, shown in Figure 2 (b) (c), filter size 1×1 performs comparably with filter size 7×7 (PSNR gap less than 0.1dB).



(a) (b) (c)

Figure 2: The performance of adaptation with different filter sizes of AdaFM layers in the three restoration tasks.

3. Comparison With AdaBN and Conditional Instance Normalization

In this section, we present the results of AdaBN [5] and conditional instance normalization [4] in the three restoration tasks. As we can see from Figure 3, AdaBN is unable to achieve acceptable adaptation results in all three tasks, indicating that AdaBN is not suitable for image restoration. Conditional instance normalization performs well in super resolution but cannot deal with denoising and DeJPEG problems, as shown in Figure 3. In general, our proposed method outperforms these two methods in all the three restoration tasks.

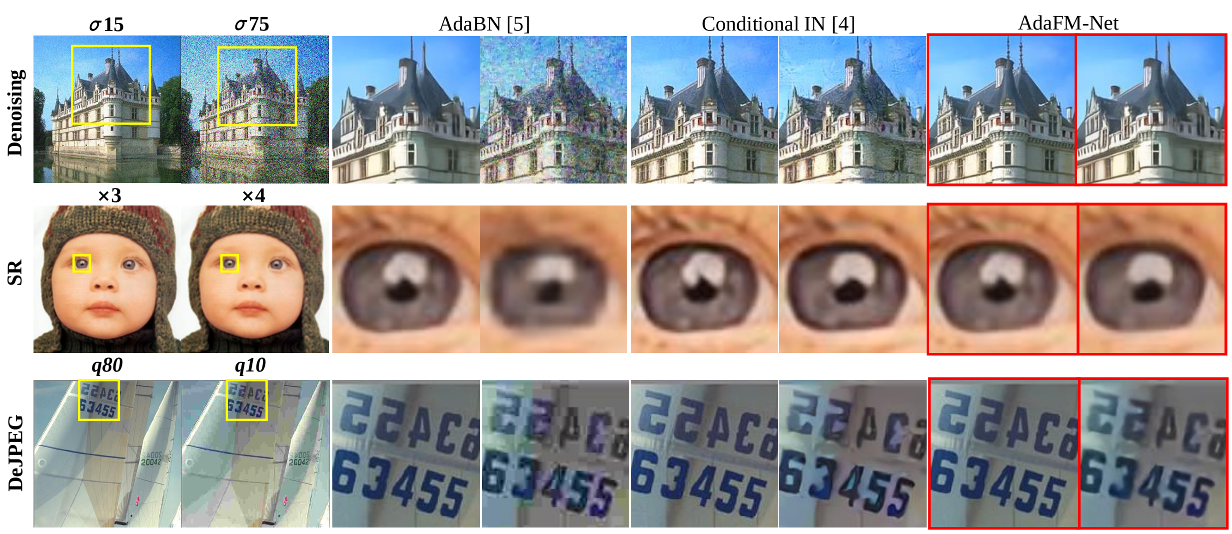
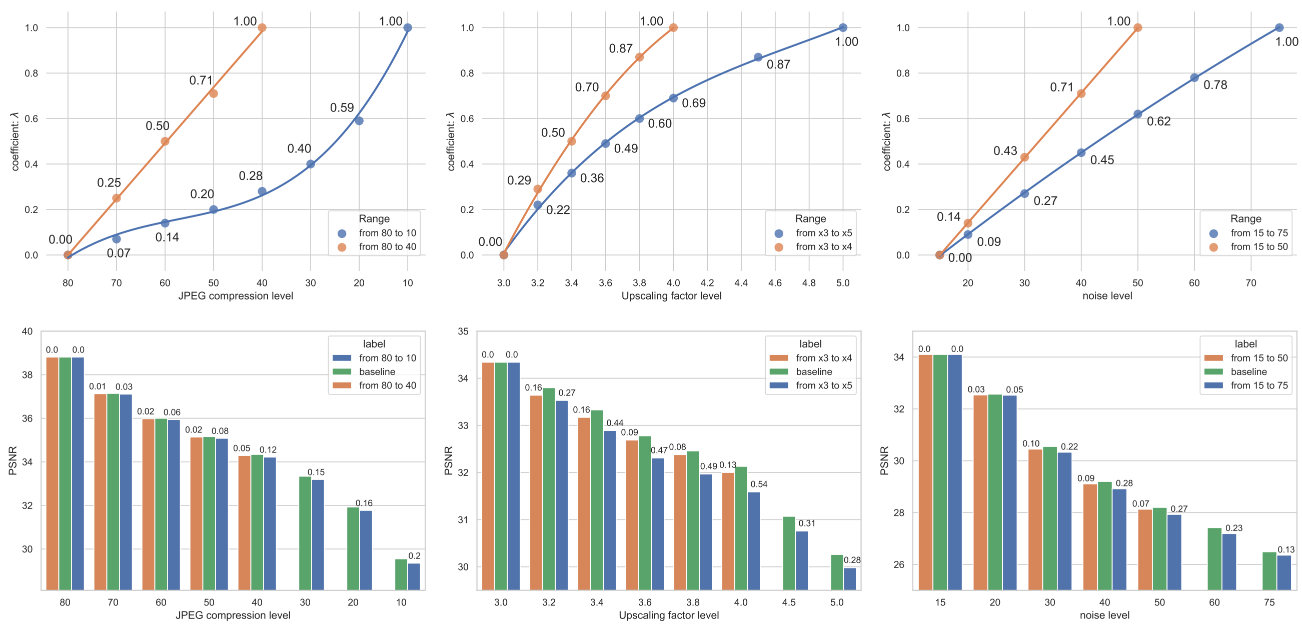


Figure 3: Artifacts on the output images produced by AdaBN and conditional instance normalization in the three restoration tasks.

4. Modulation Testing

In this section, we investigate the curving fitting with different ranges in DeJPEG, super resolution and denoising problems. The details are shown in Figure 4. In super resolution, we conduct experiments on two tasks: $\times 3 \rightarrow \times 4$ and $\times 3 \rightarrow \times 5$ with filter size 5×5. It is observed in Figure 4 (b) that the middle points in range $\times 3 \rightarrow \times 4$ can be fitted by a quadratic curve. As for range $\times 3 \rightarrow \times 5$, the curve is fitted by a cubic function: $\lambda = -9.42 + 6.01 L_c - 1.21 L_c^2 + 0.09 L_c^3$. However, the PSNR distances between the results of modulation testing and baselines are fairly large (almost 0.6dB) in $\times 3 \rightarrow \times 5$, indicating that AdaFM is unable to achieve feasible modulation within the range of 2 upscaling factors. Besides, the modulation in the range of $\times 3 \rightarrow \times 4$ yields results with PSNR distances no more than 0.16dB. In denoising tasks $\sigma 15 \rightarrow \sigma 50$ and $\sigma 15 \rightarrow \sigma 75$ (1×1), both curves in Figure 4 (c) can be closely fitted by linear functions, while the PSNR distances are within 0.1dB and 0.3dB, respectively.



(a) DeJPEG (b) Super Resolution (c) Denoising

Figure 4: Top: the curve fitting with different ranges in the three restoration tasks; Bottom: the value of PSNR distance is annotated above each bar.

As an alternative choice, we can also use the piece-wise linear function for approximation. Actually, when the range is small enough, the relationship between the coefficient λ and the middle level L_c is guaranteed to be linear. We can train a set of AdaFM-Nets on middle levels $\{L_c^i\}$. For a given level $L_c^i < L_c < L_c^{i+1}$, we can use the coefficient $\lambda = (L_c - L_c^i)/(L_c^{i+1} - L_c^i)$ to interpolate the parameters of AdaFM-Nets on L_c^i and L_c^{i+1}. This strategy needs to train and store more AdaFM-Nets on middle levels, but the adaptation accuracy is comparably higher due to the small range.

The proposed modulation strategy can also be accompanied with another λ estimation network, which automatically predicts the best λ for a given corrupted image. This refers to a blind image restoration problem, and is beyond the range of this work.

5. More Qualitative Results

In this section, we show more qualitative results of our modulation testing in the three restoration tasks (Figure 5, 6, 7). For denoising, the adaptation range is from noise level 15 to 75, while the modulation testing is applied for images with degradation $\sigma 30$, $\sigma 45$, and $\sigma 60$. In super resolution, we only consider a small adaptation range: $\times 3 \rightarrow \times 4$, and the $\times 3.2$, $\times 3.4$, $\times 3.6$ input images are used to perform the modulation testing. As for DeJPEG, the adaptation range is: $q80 \rightarrow q10$, while the degradation of images for modulation testing are $q20$, $q30$, and $q50$. Besides, we compare the results of modulation testing with those of baselines, and we find that they are fairly similar.

References

[1] Marco Bevilacqua, Aline Roumy, Christine Guillemot, and Marie Line Alberi-Morel. Low-complexity single-image super-resolution based on nonnegative neighbor embedding. 2012.

[2] Chao Dong, Yubin Deng, Chen Change Loy, and Xiaoou Tang. Compression artifacts reduction by a deep convolutional network. In *The IEEE International Conference on Computer Vision (ICCV)*, December 2015.

[3] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Learning a deep convolutional network for image super-resolution. In *European conference on computer vision*, pages 184–199. Springer, 2014.

[4] Vincent Dumoulin, Jonathon Shlens, and Manjunath Kudlur. A learned representation for artistic style. *Proc. of ICLR*, 2017.

[5] Yanghao Li, Naiyan Wang, Jianping Shi, Jiaying Liu, and Xiaodi Hou. Revisiting batch normalization for practical domain adaptation. *CoRR*, abs/1603.04779, 2016.

[6] Stefan Roth and Michael J Black. Fields of experts: A framework for learning image priors. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 860–867. IEEE, 2005.

[7] H. R. Sheikh, M. F. Sabir, and A. C. Bovik. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Transactions on Image Processing*, 15(11):3440–3451, Nov 2006.

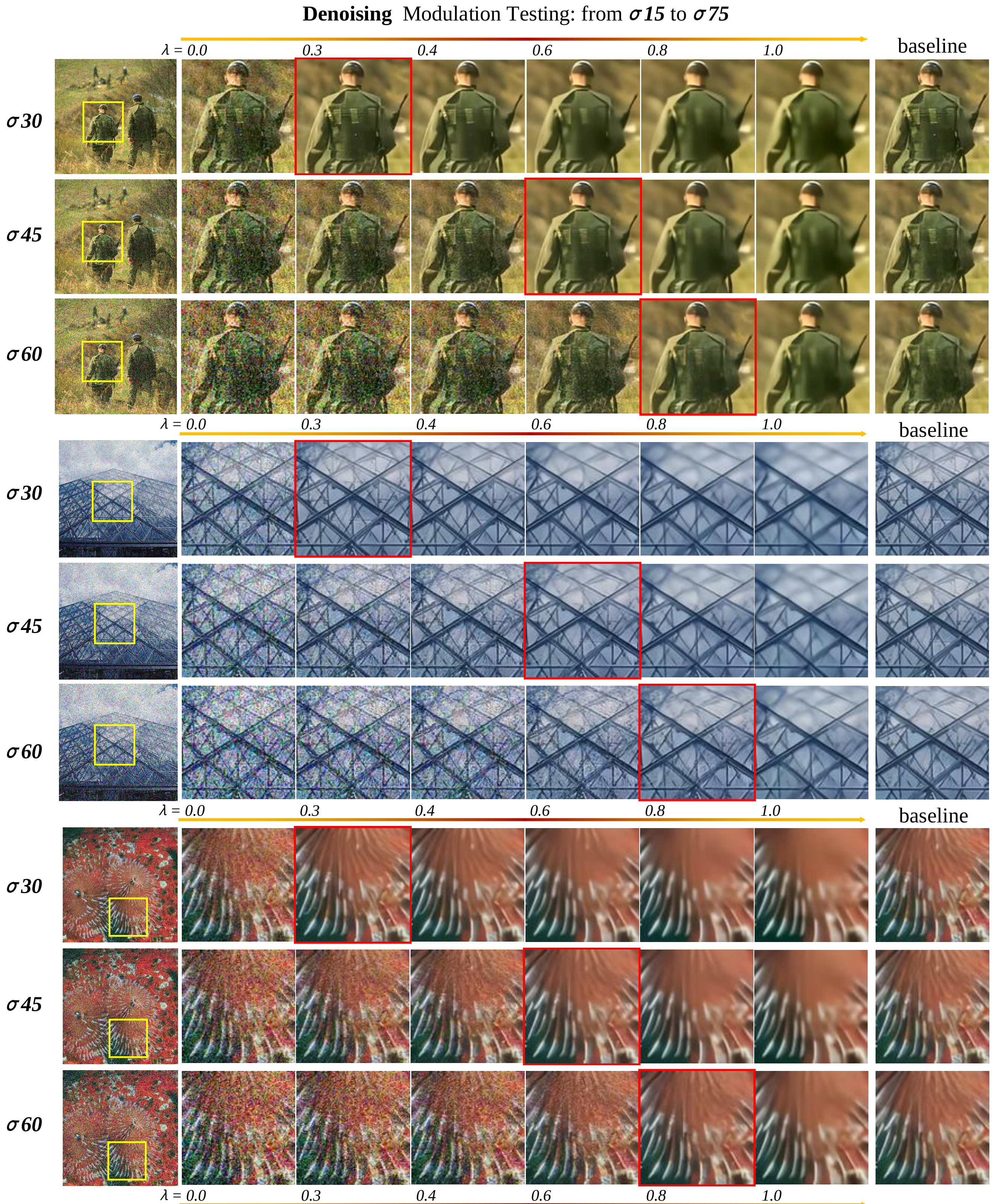


Figure 5: Qualitative results from modulation testing in denoising.

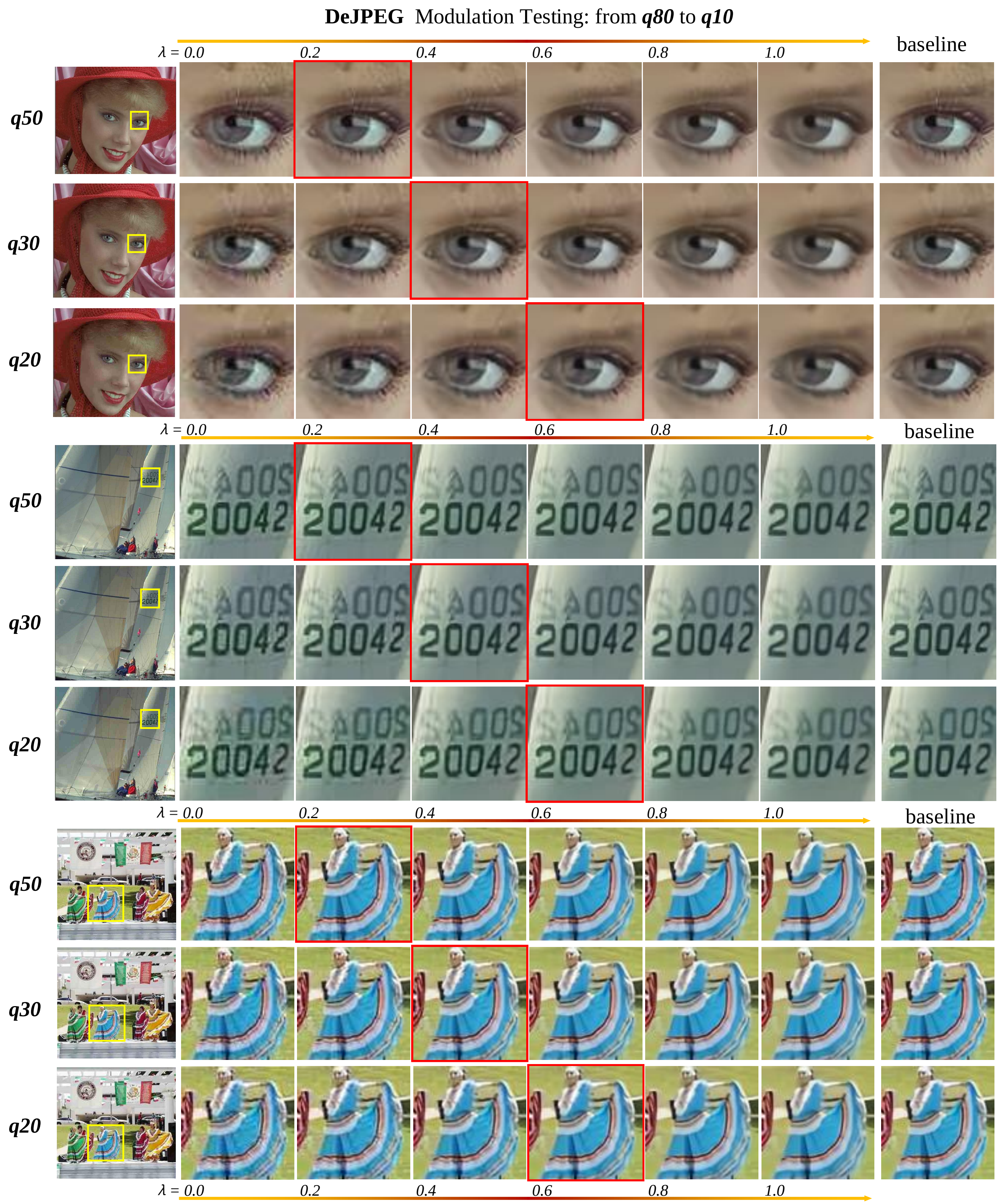


Figure 6: Qualitative results from modulation testing in DeJPEG.

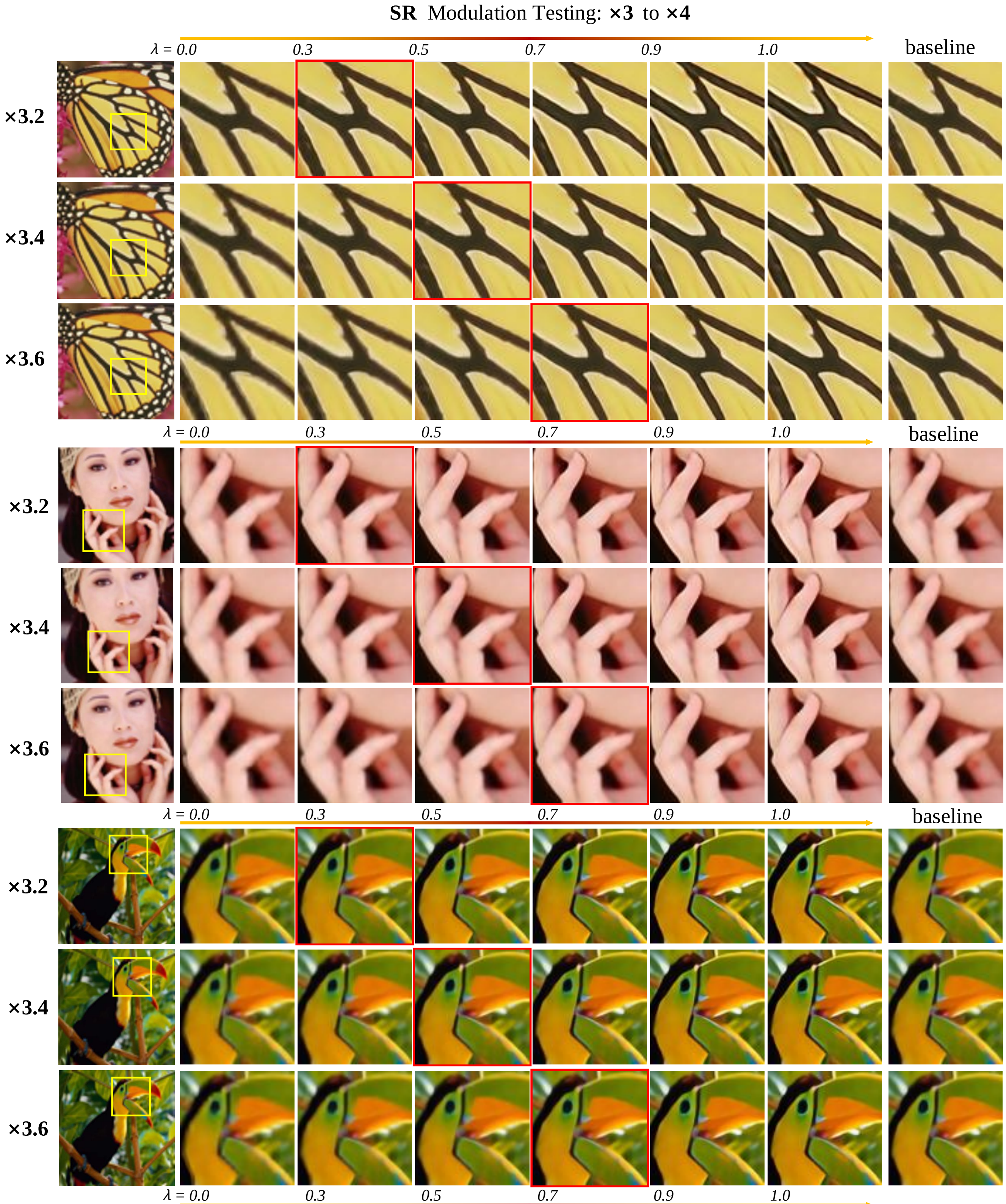


Figure 7: Qualitative results from modulation testing in super resolution.