Task-aware Image Downscaling:
Supplementary Material

In this supplementary material, we first provide more results on TAD-TAU performance trade-off in Sec. A. In Sec. B, we compare our downscaling and upscaling framework with standard compression methods. Then, additional quantitative/qualitative results for multi-scale super-resolution (SR) with extreme scaling factors are shown in Sec. C. We visualize qualitative results for *task-aware downscaling (TAD) and upscaling (TAU) for SR* in Sec. D, followed by *TAD for colorization* results in the last Sec. E. All qualitative results are best viewed zoomed in.

Compared methods for SR are SRGAN [4] - which generates realistic textures - and EDSR+ [5] - which has the state-of-the-art PSNR performance. For image colorization, we compare with our trained colorization network baseline. For both SR and colorization, we report results for all or some subset of the widely used datasets: **Set5** [2], **Set14** [7], **B100** [6], **Urban100** [3], and the validation set of **DIV2K** [1].

A TAD-TAU Reconstruction Performance Trade-off

A.1 Quantitative Results for SR

Table 1: Performance (PSNR) trade-off for SR while varying λ (in dB)

	PSNR Pairs		λ						
Dataset	Scaling	GT	0	10^{-3}	10^{-2}	10^{-1}	1	10	10^{2}
DIV2K	TAU(×2)	Original	40.16	40.22	40.06	39.81	35.57	35.03	35.00
10 Imgs	TAD(×2)	Bicubic	19.97	24.68	32.97	40.69	46.30	50.75	51.18

A.2 Quantitative/Qualitative Results for Colorization

Table 2: Performance (PSNR) trade-off for colorization while varying λ (in dB)

	PSNR Pairs		λ						
Dataset	Scaling	GT	0	10^{-2}	10^{-1}	1	5	10	10^{2}
DIV2K	TAU	Original	40.07	40.40	40.57	38.37	35.74	33.61	20.38
10 Imgs	TAD	Y	17.94	27.26	37.27	44.10	48.48	49.78	49.89

For image resizing tasks, the relatively low PSNR of the downscaled image does not show much difference in the visual quality compared to bicubic-downsampled image. However, for image colorization task, the low PSNR in gray scale image shows some disturbing visual artifacts. We therefore put more weight on maintaining good-quality gray scale (Y channel from YCbCr) image, and chose $\lambda = 5$ for our experiments.

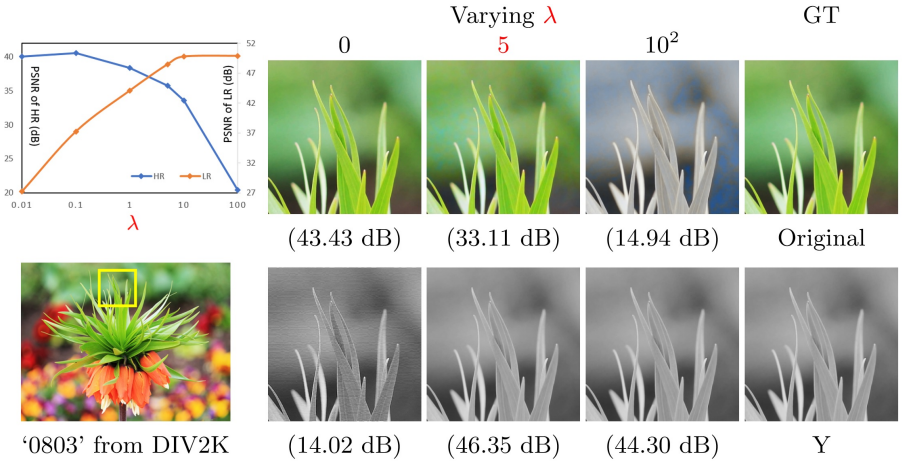

Fig. 1: TAD-TAU reconstruction performance trade-off - qualitative results for image colorization. Bottom row shows the TAD gray scale images, and upper row shows the reconstructed TAU color images.

A.3 Analysis on the LR image encoding process

In our framework, a residual to a guidance image is a distortion to an LR image. This can be seen in Figure 2, where the TAD LR image of the synthetic color grid example contains some structured additive noise. However, although the noise looks quite significant above, it becomes visually almost negligible at its original LR scale of 8x8 pixels. We assert that LR images need to be visually good enough only at low display resolution(e.g. yellow boxes in Figure 4 from the paper), not enlarged to higher resolution. For the quality of images at high display resolution, the LR images should be super-resolved where the distortions become clear image details.

For image colorization task, observing the difference of Gray and TAD Gray images in Figure 6 from the paper shows how the learned distortions (Figure 2, Gray - TAD Gray) that contain color information have little visual effect.

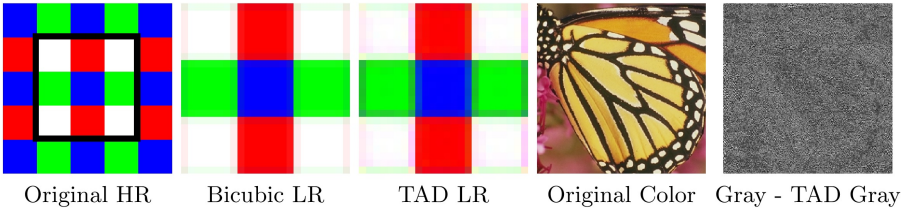

Fig. 2: Each colored box is 16 pixels tall at original HR and is downscaled with a scaling factor of 2. We also visualize the difference of TAD gray and gray guidance image of Figure 6 in the paper (multiplied by 50)

B Comparison with image compression algorithms

The purpose of image downscaling - *e.g.* reducing the resizing calculations for efficient display - is different from compression, so that direct performance comparison can be misleading. We still compare with standard image compression methods below, but due to the different number of pixels between HR and LR images, we report the memory storage of a benchmark dataset instead of bits per pixel. Note the PSNR gap between TAD-EDSR+ and Bicubic-EDSR+, though TAD LR images need more storage due to the additive noise. We further increase the compression and reconstruction performances with our TAD16 LR images from deeper networks.

Also note that our compression module, lossless image compression after differentiable rounding operation, obviously can be further improved by replacing to deep image compression methods.

Table 3: Quantitative comparison with image compression methods on DIV2K. We apply TAD16 network which consists of 16 residual blocks.

Downscaling	None			Bicubic		TAD		TAD16	
Compression	JPEG	JPEG2000	WebP	PNG	FLIF[1]	PNG	FLIF	PNG	FLIF
Storage(MB)	103.2	93.3	87.9	117.8	86.7	125.7	102.9	118.7	88.8
Upscaling	None			EDSR+		EDSR+		EDSR+	
PSNR(dB)	39.42	43.43	40.51	35.12		40.21		40.82	

C Multi-scale Extreme SR

In this section, we report the evaluated performance (PSNR) for recursive multi-scale SR with extreme scaling factors, and provide qualitative results. In addition to just comparing with a simple bicubic baseline, we also recursively apply EDSR [5], following the same procedure as our approach, and compare the results. Our upscaled (TAU) result with TAD (task-aware downscaling) shows the best performance both quantitatively in PSNR, and qualitatively in preserving the visual details and the global structure of the original image.

Table 4: PSNR results for images upscaled with extreme scaling factors.

	Scaling		Scale				
Dataset	Down	Up	×8	×16	×32	×64	×128
DIV2K	Bicubic	Bicubic	23.69 dB	21.35 dB	19.40 dB	17.68 dB	16.08 dB
		EDSR	25.50 dB	22.71 dB	20.33 dB	18.36 dB	16.54 dB
	TAD	TAU	26.77 dB	23.72 dB	21.25 dB	19.17 dB	17.27 dB

[1] https://flif.info/, state-of-the-art lossless image compression method.

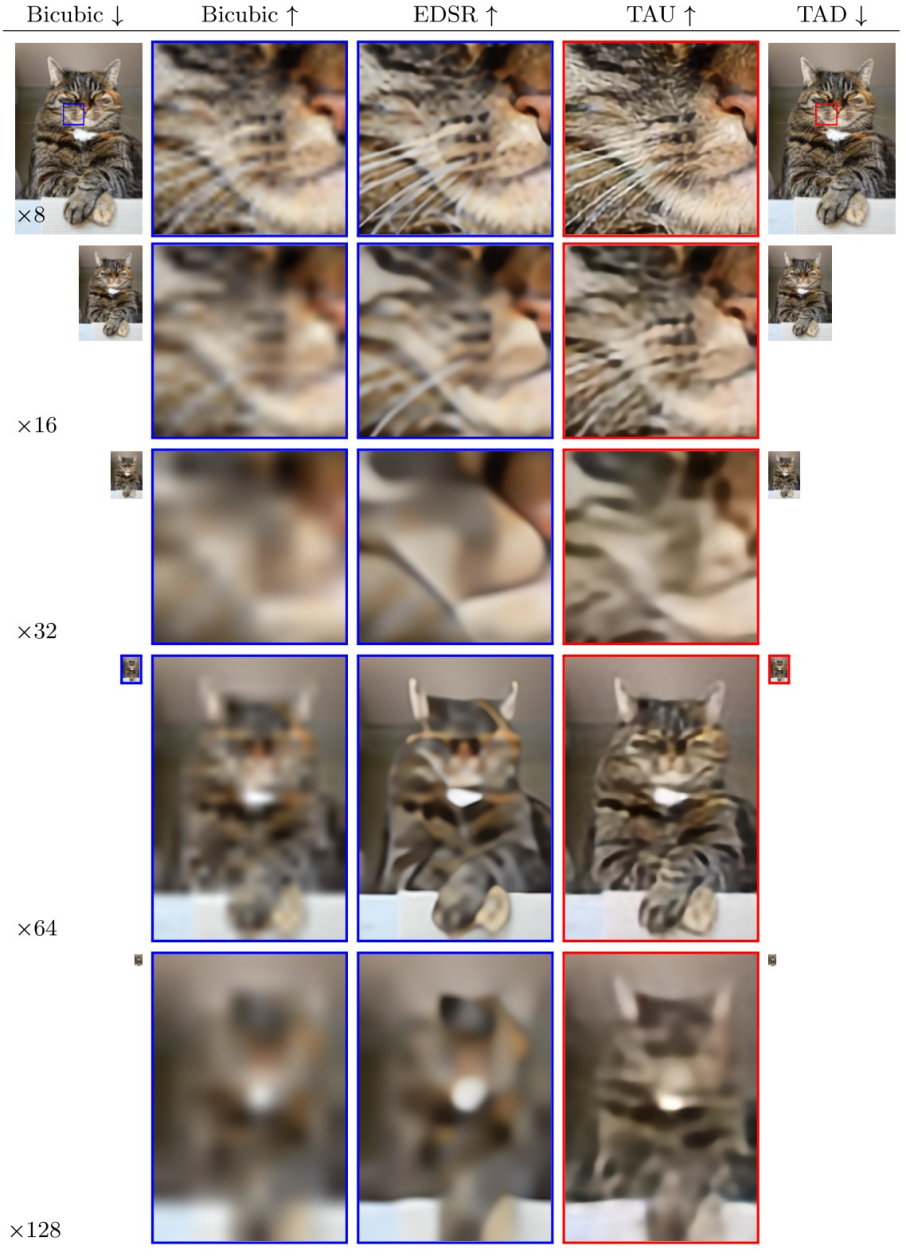

Fig. 3: Extreme SR results for “0869” image from DIV2K

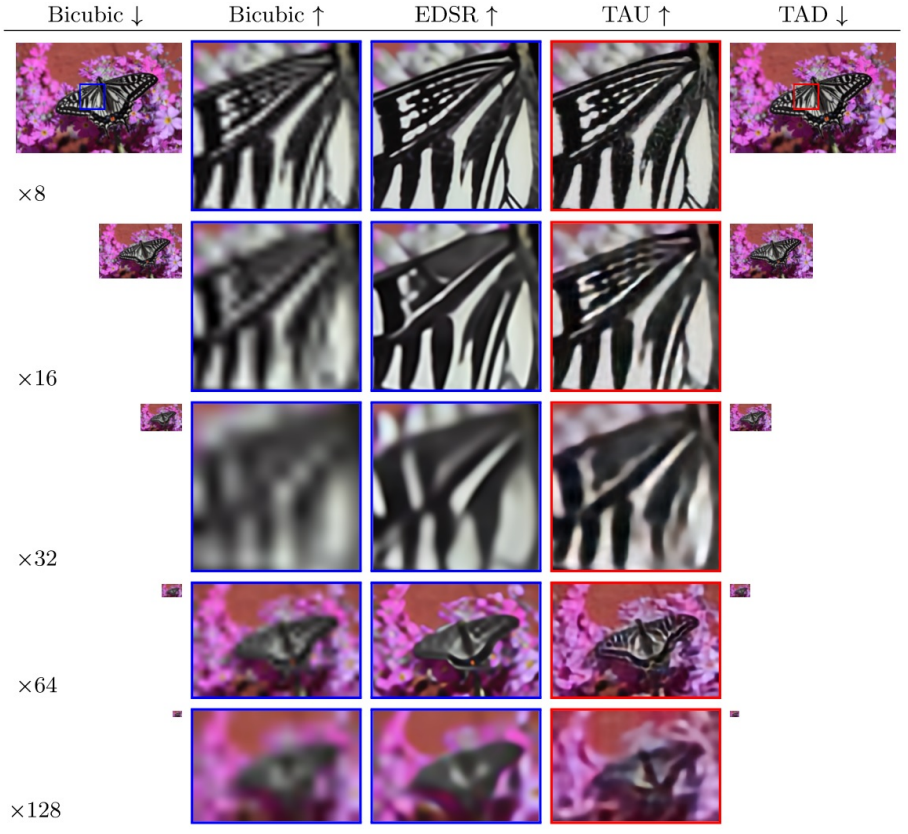

Fig. 4: Extreme SR results for “0882” image from DIV2K

D Qualitative Results for SR

For **Set5**, we provide results for scaling factors ×2 and ×4. Since all methods give good results for scaling factor ×2, only the results for scaling factor ×4 are shown for the other datasets.

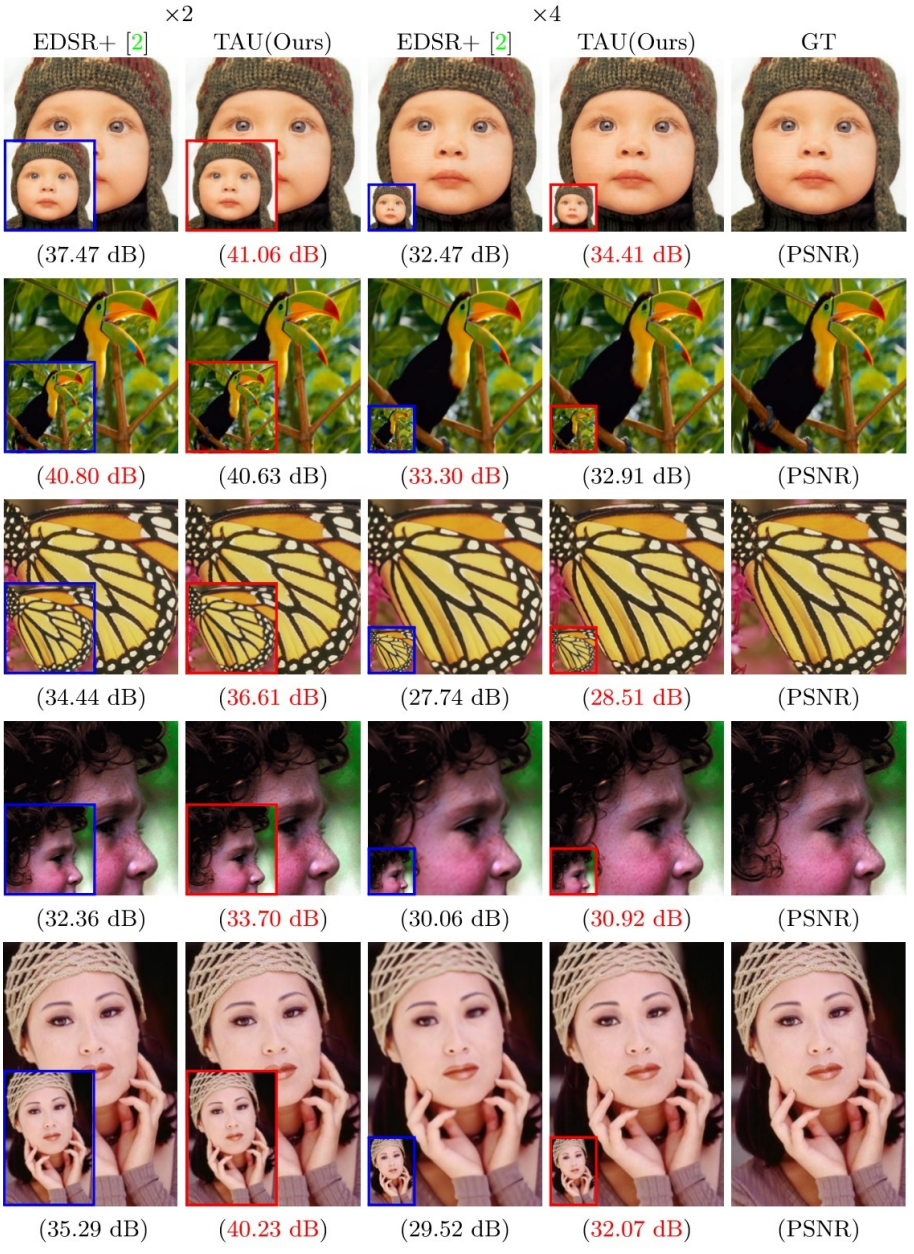

Fig. 5: SR results for Set5 [2].

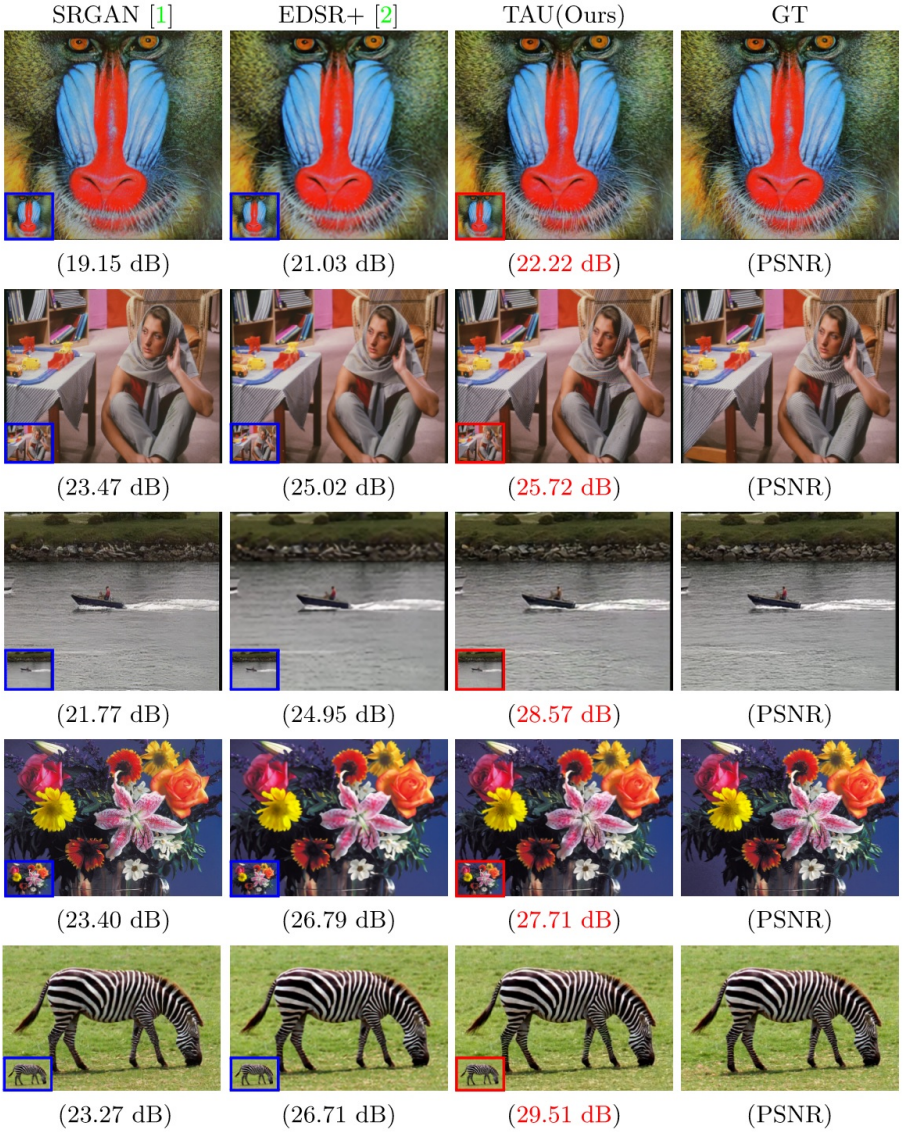

Fig. 6: SR results for 5 samples of Set14 [7]

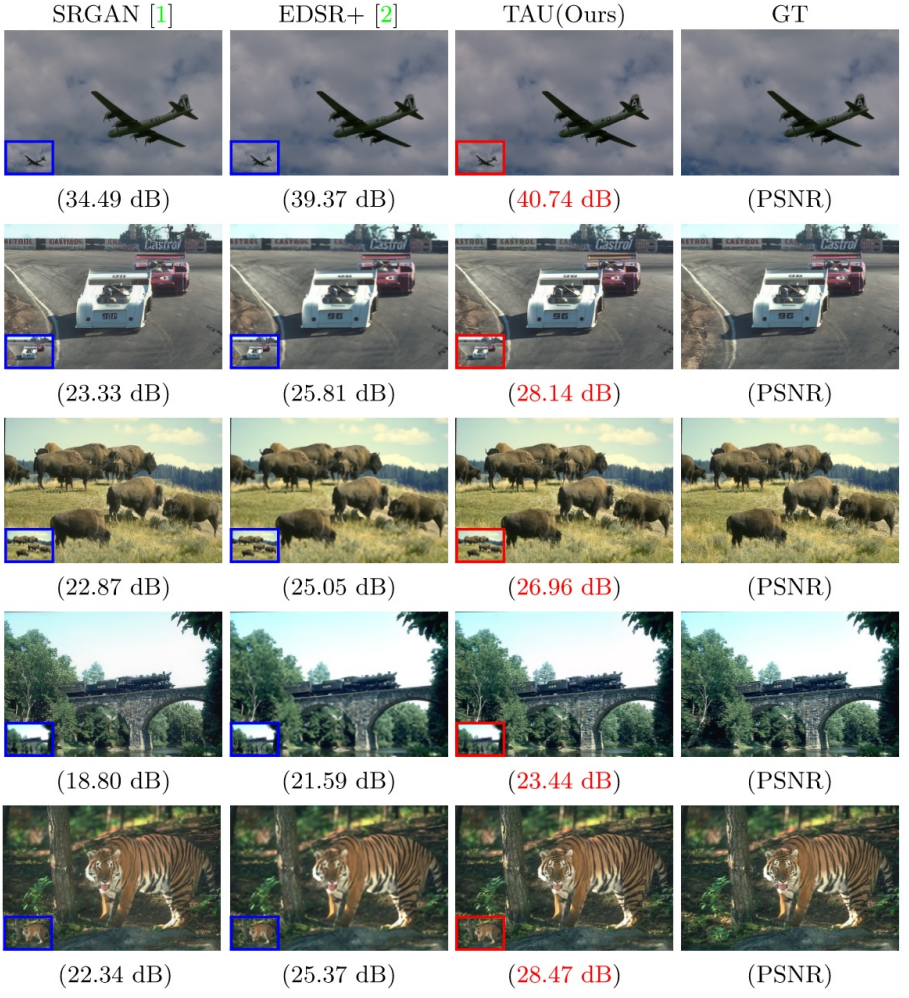

Fig. 7: SR results for 5 samples of B100 [6].

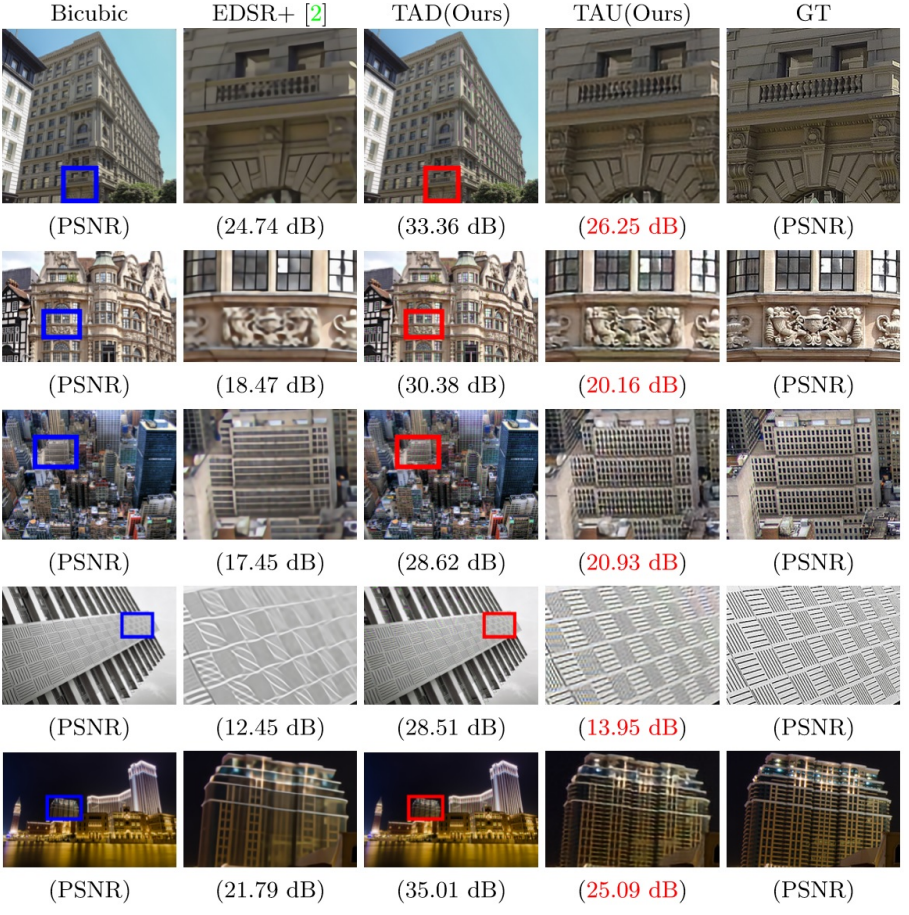

Fig. 8: SR results for 5 samples of Urban100 [3].

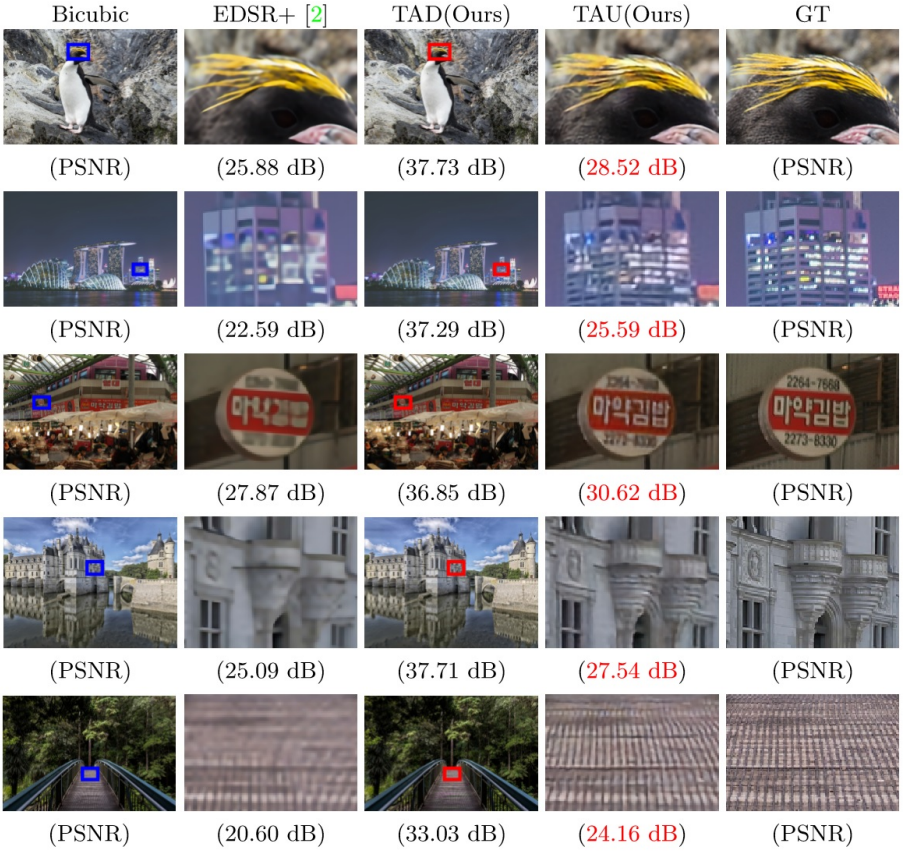

Fig. 9: SR results for 5 samples of DIV2K [1].

E Quantitative/Qualitative Results for Colorization

E.1 Quantitative Results

We evaluated our TAD for colorization model on 5 benchmark datasets: **Set5**, **Set14**, **B100**, **Urban100**, and **DIV2K**. Measured PSNR for gray scale images (TAD Gray) is with respect to the Y channel of YCbCr (converted from RGB) original image. PSNR for color images are measured w.r.t. the original RGB image.

Table 5: Quantitative PSNR results for image color space conversion.

	Model		
Dataset	Color (baseline)	TAD Gray	TAU Color
Set5	19.12 dB	45.16 dB	35.22 dB
Set14	21.14 dB	43.84 dB	32.67 dB
B100	24.21 dB	47.41 dB	32.73 dB
Urban100	23.29 dB	45.29 dB	30.98 dB
DIV2K	21.10 dB	45.61 dB	36.63 dB

E.2 Qualitative Results

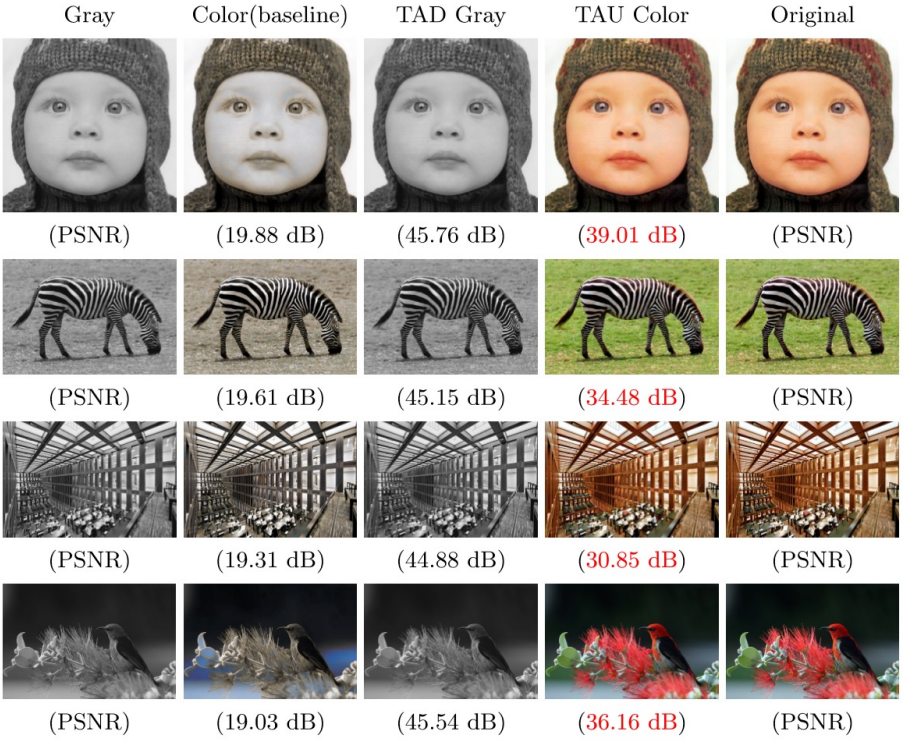

Fig. 10: Image color space conversion results. One image each from Set5, Set14, Urban100, and DIV2K are shown.

References

1. Agustsson, E., Timofte, R.: Ntire 2017 challenge on single image super-resolution: Dataset and study. In: The IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (2017)
2. Bevilacqua, M., Roumy, A., Guillemot, C., Alberi-Morel, M.L.: Low-complexity single-image super-resolution based on nonnegative neighbor embedding. In: British Machine Vision Conference (BMVC). BMVA press (2012)
3. Huang, J., Singh, A., Ahuja, N.: Single image super-resolution from transformed self-exemplars. In: CVPR. pp. 5197–5206. IEEE (2015)
4. Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., Shi, W.: Photo-realistic single image super-resolution using a generative adversarial network. In: CVPR (2017)
5. Lim, B., Son, S., Kim, H., Nah, S., Lee, K.M.: Enhanced deep residual networks for single image super-resolution. In: The IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (2017)
6. Martin, D.R., Fowlkes, C.C., Tal, D., Malik, J.: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: ICCV. pp. 416–425 (2001)
7. Zeyde, R., Elad, M., Protter, M.: On single image scale-up using sparse-representations. In: Curves and Surfaces. vol. 6920, pp. 711–730. Springer (2010)