PassionSR: Post-Training Quantization with Adaptive Scale in One-Step Diffusion based Image Super-Resolution *Supplemental Materials*

Libo Zhu[1], **Jianze Li**[1], **Haotong Qin**[2]*,
Yulun Zhang[1]*, **Yong Guo**[3], **Xiaokang Yang**[1]
[1]Shanghai Jiao Tong University, [2]ETH Zürich, [3]Max Planck Institute for Informatics

1. Additional Visual Comparison

In addition to the visual comparison images provided in the main text, we offer additional visual comparison figures in Fig. 1. It indicates that PassionSR gains better visual effects than other comparative previous methods.

2. UNet-only quantization

Considering that the previous methods only quantize the UNet, we design a UNet-only experiment to make the comparison more fair. The Unet-only refers to only quantizing the UNet while VAE keeps full precision. We aim to analyze the quantization effects of various quantization methods on Unet-only setting in this experiment.

Quantitative Results. In the UNet-only quantization experiment, as shown in Tab. 1, PassionSR demonstrates a significant improvement over previous methods under W8A8 and W6A6 settings. The experiments phenomena and results are similar to the UNet-VAE experiment. On every dataset, 8-bit PassionSR achieves results comparable to the full-precision OSEDiff and even surpasses it in certain cases. In contrast, other quantization methods show noticeable drops in quantitative performance. For 6-bit quantization, contrast methods such as LSQ and Q-Diffusion yield lower structural metrics (*e.g.*, PSNR and SSIM) but achieve higher scores in non-reference IQA metrics. PassionSR, however, delivers the best reference IQA results while maintaining relatively lower non-reference IQA scores compared to other methods. The quantitative experimental results indicate that PassionSR achieves relatively the best quantization performance on the UNet. Comparatively, under the same quantization bits setting, the UNet-only quantized model gains slightly better effects than UNet-VAE one. The little performance drop is brought by the precision reduction of VAE.

Visual Comparison. Figure 3 presents a visual comparison (×4) for the UNet-only quantization. To highlight the differences, we select several challenging examples for analysis. The visual comparison results are also similar to UNet-VAE quantization experiment results. Compared to existing methods, PassionSR produces sharper details and more realistic textures, with results closely comparative with those of the full-precision model. Remarkably, in some cases, PassionSR even outperforms its full-precision counterpart, PassionSR-FP. It indicates that PassionSR still gains better visual effects in the Unet-only quantization.

3. Implementation of Comparative methods

Considering that there aren't quantization works for One-Step Diffusion based Image Super-Resolution (OSDISR), we implement many comparative methods based on their released code. However, owing to the difference of full-precision backbones, we encounter some problems while performing model migration with the implementation code. We make appropriate modifications to address these issues.

3.1. Q-Diffusion [5]

Q-Diffusion is an impressing quantization method on multi-step diffusion models. It performs well on quantizing Latent Diffusion (LDM) [7] for higher resolution LSUN [10] experiments. In the Q-Diffusion calibration process, it uniformly samples UNet inputs at different timesteps for its calibration dataset. We utilize the calibration dataset constructed for our PassionSR in Q-Diffusion. Q-Diffusion calibrates the weight quantizers and the activation quantizers seperately. We find that GPU comsumption of calibrating the activation quantizers is obviously larger and memory overflow issue happens in RTX A6000. We originally intended to use the same parameters for calibration as in the Q-Diffusion implementation. But we find that calibration time is too long so we reduce the number of the calibration epochs while ensuring that the quantization process has converged. We adopt the same implementation on the quantization of VAE, calibrating them seperately.

[*] Corresponding authors: Haotong Qin and Yulun Zhang

3.2. EfficientDM [2]

EfficientDM is a newly proposed quantization method on multi-step diffusion models. Based on the QALoRA [9] to finetune the weight and quantization parameters, EfficientDM performs outstanding ability in minimizing quantization errors. When we apply EfficientDM to PassionSR-FP, the original quantization setting is kept. Because the code EfficientDM provides only supports the 2, 4 and 8 bits quantization, we made appropriate modifications to it. The main reason why only quantization with bit widths that are powers of 2 is supported lies behind that EfficientDM attempts to store quantized weight. The storing of tensor matrix in Pytorch only supports bit widths that are powers of 2. Considering that we only need fake quantization to simulate quantization errors, we remove this part so that 6 bit quantization we adopt can be supported by EfficientDM. The removed parts are unrelated to the quantization algorithms of EfficientDM, not affecting the quantization performance.

4. Detailed algorithm of PassionSR

Algorithm 1 PassionSR Calibration

```
Require: Pretrained full precision one-step diffusion based image super-resolution (OSDIR) model $[W]$
Require: Quantized low bits OSDIR model $[\hat{W}]$
Require: Dataset [D] consisted of low, high resolution image pairs randomly cropped from $W$'s training datasets
Initialize the quantization parameters of $\hat{W}$ with D
Q-list = [VAE encoder, UNet, VAE decoder] in $\hat{W}$
F-list = [VAE encoder, UNet, VAE decoder] in $W$
Initialize calibration dataset $D_q$ = D
for $\hat{M}$, $M$ in Q-list, F-list do
    for $n = 1, \dots, N$ epochs do
        for $i$ in $D_q$ do
            $Loss = \|\hat{M}(i) - M(i)\|_2$
            Update $\hat{M}$'s LET component with $Loss$
            Reinitialize M's LBQ component
        end for
    end for
    for $n = 1, \dots N$ epochs do
        for $i$ in $D_q$ do
            $Loss = \|\hat{M}(i) - M(i)\|_2$
            Update $\hat{M}$'s LBQ component with $Loss$
        end for
    end for
    for $i$ in $D_q$ do
        Update $D_q$ with $\hat{M}(i)$
    end for
end for
```

5. Visual Comparison of Ablation Study

To make the ablation study more convincing, we not only provide quantative results in the main text but also include additional visual results in Fig. 2. To illustrate the effects of each component, we seperate our PassionSR into three parts, LBQ (learnable boundary quantizer), LET (learnable equivalent transformation) and DQC (distributed quantization calibration). Based on the MaxMin [4] baseline, we sequentially incorporate LBQ, LET, and DQC to investigate the role of each component. Under the setting of W6A6, MaxMin maintains basic structural information but has a lot of color distortions. The introduction of LBQ and LET gradually solves the color distortions problem while the details in the images are more realistic. Althought DQC is used for stablizing the calibrating process, there are also slight performance increase after applying DQC.

References

[1] Steven K Esser, Jeffrey L McKinstry, Deepika Bablani, Rathinakumar Appuswamy, and Dharmendra S Modha. Learned step size quantization. In *ICLR*, 2020. 3, 4, 5

[2] Yefei He, Jing Liu, Weijia Wu, Hong Zhou, and Bohan Zhuang. Efficientdm: Efficient quantization-aware fine-tuning of low-bit diffusion models. In *ICLR*, 2024. 2, 3, 4, 5

[3] Benoit Jacob, Skirmantas Kligys, Bo Chen, Menglong Zhu, Matthew Tang, Andrew Howard, Hartwig Adam, and Dmitry Kalenichenko. Quantization and training of neural networks for efficient integer-arithmetic-only inference. In *CVPR*, 2018. 4

[4] Eli Kravchik, Fan Yang, Pavel Kisilev, and Yoni Choukroun. Low-bit quantization of neural networks for efficient inference. In *ICCVW*, 2019. 2, 3, 4, 5

[5] Xiuyu Li, Yijiang Liu, Long Lian, Huanrui Yang, Zhen Dong, Daniel Kang, Shanghang Zhang, and Kurt Keutzer. Q-diffusion: Quantizing diffusion models. In *ICCV*, 2023. 1, 3, 4, 5

[6] Xinqi Lin, Jingwen He, Ziyan Chen, Zhaoyang Lyu, Bo Dai, Fanghua Yu, Wanli Ouyang, Yu Qiao, and Chao Dong. Diffbir: Towards blind image restoration with generative diffusion prior. In *ECCV*, 2024. 3, 5

[7] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *CVPR*, 2022. 1

[8] Rongyuan Wu, Lingchen Sun, Zhiyuan Ma, and Lei Zhang. One-step effective diffusion network for real-world image super-resolution. In *NeurIPS*, 2024. 3, 4, 5

[9] Yuhui Xu, Lingxi Xie, Xiaotao Gu, Xin Chen, Heng Chang, Hengheng Zhang, Zhengsu Chen, Xiaopeng Zhang, and Qi Tian. Qa-lora: Quantization-aware low-rank adaptation of large language models. In *ICLR*, 2024. 2

[10] Fisher Yu, Ari Seff, Yinda Zhang, Shuran Song, Thomas Funkhouser, and Jianxiong Xiao. Lsun: Construction of a large-scale image dataset using deep learning with humans in the loop. In *CVPR*, 2015. 1

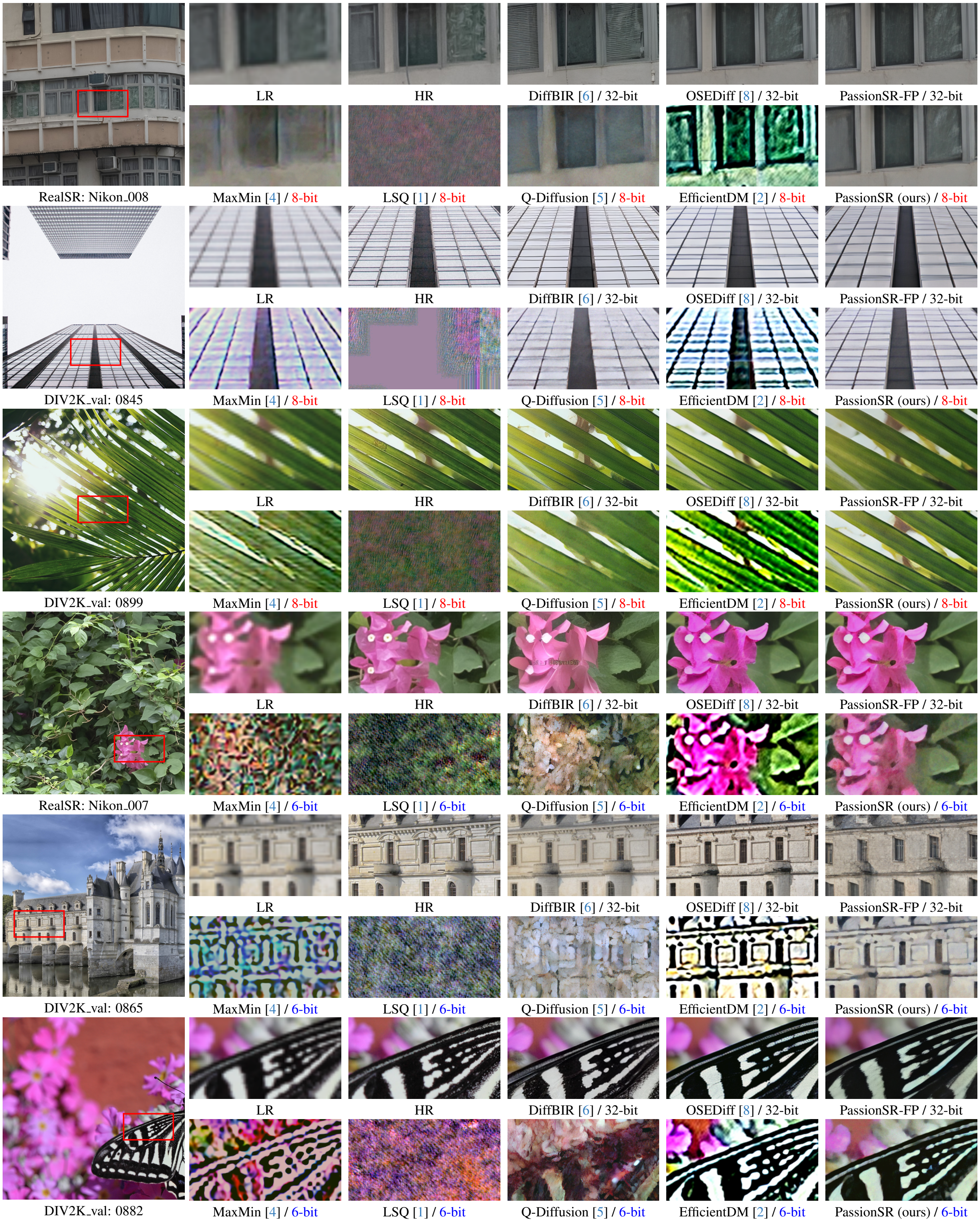

Figure 1. Visual comparison (×4) with high-resolution image, full-precision model's output and different quantization methods in some challenging cases at W8A8 and W6A6 **UNet-VAE quantization**. PassionSR gains significant visual advantages over other methods.

Datasets	Bits	Methods	PSNR↑	SSIM↑	LPIPS↓	DISTS↓	NIQE↓	MUSIQ↑	MANIQA↑	CLIP-IQA↑
RealSR	W32A32	OSEDiff [8]	25.27	0.7379	0.3027	0.1808	4.355	67.43	0.4766	0.6835
		PassionSR-FP	25.39	0.7460	0.2984	0.1813	4.453	67.05	0.4680	0.6796
	W8A8	MaxMin [3]	23.16	0.6875	0.5463	0.2879	7.932	32.92	0.1849	0.2363
		LSQ [1]	15.39	0.3375	0.9944	0.5427	10.08	50.11	0.3533	0.3173
		Q-Diffusion [5]	24.88	0.6967	0.4993	0.2696	8.437	44.69	0.2352	0.5604
		EfficientDM [2]	14.77	0.4253	0.5478	0.3462	7.526	44.75	0.2568	0.4000
		PassionSR (ours)	25.67	0.7499	0.3140	0.1932	5.654	65.88	0.4437	0.6912
	W6A6	MaxMin [3]	15.55	0.2417	0.8018	0.4449	9.263	42.15	0.2791	0.4174
		LSQ [1]	13.73	0.1081	1.0900	0.5450	8.430	53.61	0.3036	0.4396
		Q-Diffusion [5]	19.75	0.4727	0.6877	0.4024	7.381	56.46	0.4380	0.6439
		EfficientDM [2]	14.75	0.4386	0.5233	0.3451	7.497	42.97	0.2498	0.3740
		PassionSR (ours)	25.15	0.7196	0.4199	0.2592	8.618	44.43	0.2131	0.4612
DRealSR	W32A32	OSEDiff [8]	25.57	0.7885	0.3447	0.1808	4.371	37.22	0.4794	0.7540
		PassionSR-FP	26.70	0.7978	0.3339	0.1765	4.336	37.03	0.4686	0.7520
	W8A8	MaxMin [3]	24.97	0.7989	0.5091	0.2921	8.215	24.05	0.1846	0.3163
		LSQ [1]	14.56	0.1795	1.1661	0.592	10.19	29.07	0.4010	0.3970
		Q-Diffusion [5]	27.14	0.7184	0.4765	0.2895	9.861	26.44	0.2284	0.5608
		EfficientDM [2]	15.55	0.4183	0.6291	0.3555	6.859	28.61	0.2468	0.4150
		PassionSR (ours)	27.41	0.8146	0.3422	0.1918	6.070	33.56	0.4286	0.7554
	W6A6	MaxMin [3]	13.08	0.2291	0.8131	0.5077	10.51	35.83	0.2702	0.3864
		LSQ [1]	12.95	0.0934	1.1890	0.5833	8.591	26.39	0.2911	0.5600
		Q-Diffusion [5]	21.75	0.6096	0.7008	0.4039	6.854	24.39	0.4109	0.6696
		EfficientDM [2]	15.07	0.4287	0.6127	0.357	6.690	28.37	0.2351	0.3973
		PassionSR (ours)	26.62	0.7984	0.4429	0.2571	8.484	26.26	0.1824	0.4358
DIV2K_val	W32A32	OSEDiff [8]	24.95	0.7154	0.2325	0.1197	3.616	68.92	0.4340	0.6842
		PassionSR-FP	25.16	0.7221	0.2373	0.1185	3.573	69.27	0.4402	0.6958
	W8A8	MaxMin [3]	22.33	0.6618	0.5639	0.2731	7.563	33.68	0.1913	0.2818
		LSQ [1]	13.90	0.2537	0.9932	0.5515	9.578	48.11	0.3512	0.3246
		Q-Diffusion [5]	24.20	0.6813	0.3997	0.2400	7.955	51.95	0.2709	0.6243
		EfficientDM [2]	15.24	0.4954	0.6041	0.3374	6.856	48.78	0.2685	0.4235
		PassionSR (ours)	25.11	0.7199	0.2496	0.1277	4.424	67.92	0.3993	0.6939
	W6A6	MaxMin [3]	11.66	0.1606	0.8509	0.4966	11.30	45.47	0.2764	0.3523
		LSQ [1]	12.21	0.0858	1.0695	0.5424	8.564	52.74	0.2872	0.4692
		Q-Diffusion [5]	18.92	0.4939	0.6227	0.3718	6.162	51.50	0.3946	0.5814
		EfficientDM [2]	15.09	0.4991	0.5953	0.3292	6.900	46.01	0.2570	0.4007
		PassionSR (ours)	24.34	0.7097	0.3440	0.2075	7.039	51.19	0.2267	0.4802

Table 1. Quantitative **UNet-only** quantization experiments results. PassionSR-FP is used as full-precision backbones rather than original OSEDiff. W8A8 denotes 8 bits weight and 8 bits activation quantization. The best results in the same setting are colored with red.

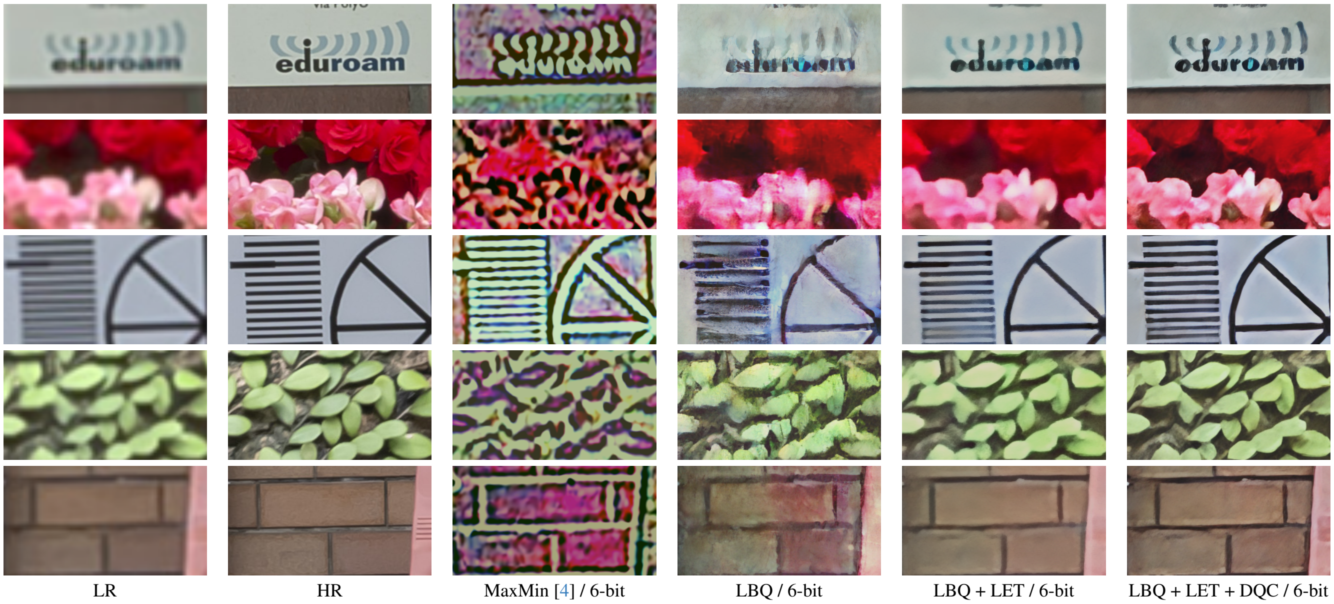

Figure 2. Visual comparison (×4) with low-resolution, high-resolution images and different quantization settings in ablation study.

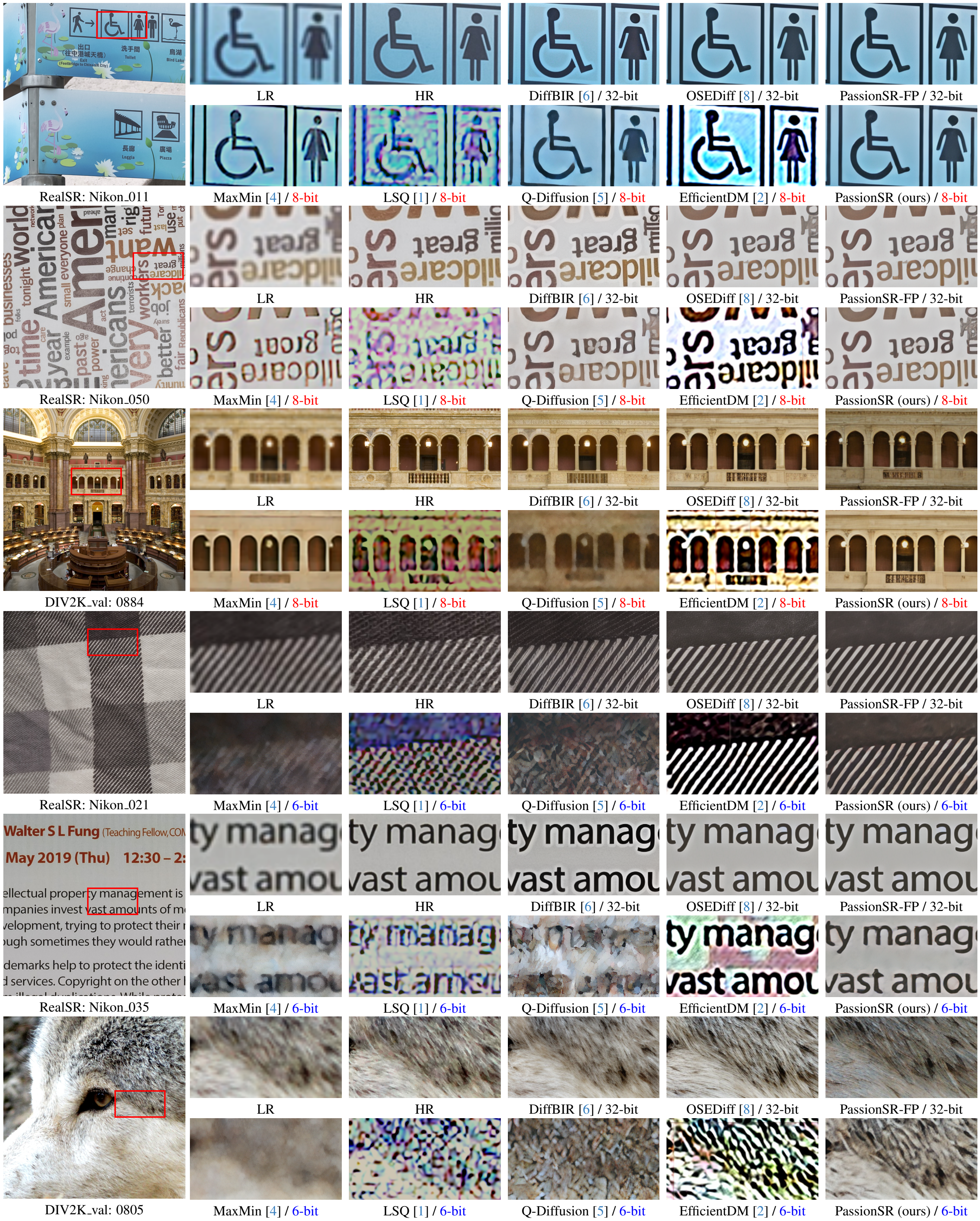

Figure 3. Visual comparison (×4) with high-resolution image, full-precision model's output and different quantization methods in some challenging cases at W8A8 and W6A6 **UNet-only quantization**. PassionSR gains significant visual advantages over other methods.