

Adversarial Diffusion Compression for Real-World Image Super-Resolution

Supplementary Material

Overview

Our main paper outlines the core idea and techniques of proposed method. It also demonstrates the effectiveness of our four main methodological contributions and adopted settings through experimental validation. In this **Supplementary Material**, we provide additional details, including the training and inference pseudocode of proposed ADC framework in Sec. A, more ablation studies in Sec. B, more comparison results and a user study analysis in Sec. C, as well as an efficiency evaluation of AdcSR and its SD-based one-step teacher OSEDiff [23] on a real mobile platform, which are not included in the main paper due to space constraints.

A. Pseudocode of Training and Inference

In this section, we present the training and inference procedures of our ADC framework, as summarized in Algo. 1. The training process consists of two stages: (1) pretraining the channel-pruned SD VAE decoder to restore its decoding ability, and (2) knowledge distillation with adversarial loss to compensate for performance degradation due to our compression. The inference of AdcSR is faster than complete SD [14, 15] models due to its compressed structure.

B. More Ablation Studies

Effect of Channel Pruning. Tab. B.1 compares employed channel pruning to other two alternative structural compression strategies: using a block-removed UNet [1, 8] and decoding by the pretrained tiny VAE [4, 16]. We observe that, with similar parameter numbers and inference speed, applying block removal results in a noticeable performance loss of 0.0083 and 0.0084 in LPIPS and DISTs, respectively. While the use of tiny VAE decoder can lead to reductions of 12M parameters and 0.01s in inference time, it substantially degrades performance by 0.0297 and 0.0161 in LPIPS and DISTs. This may be attributed to the reduced depth and the absence of global receptive field in tiny VAE, which relies solely on convolutions for decoding. These results validate the effectiveness of our adopted feature channel pruning.

Effect of Various LoRA Ranks, and Fully Finetuning the

Table B.1. Ablation study of structural compression on DRealSR.

Method (UNet / VAE Decoder)	LPIPS↓	DISTS↓	#Param.↓	Time↓
Block-Removed / Channel-Pruned	0.3129	0.2284	457	0.03
Channel-Pruned / Tiny VAE Dec.	0.3343	0.2361	444	0.02
Ch.-Pruned / Ch.-Pruned (Ours)	0.3046	0.2200	456	0.03

Table B.2. Ablation study of LoRA rank r and fully finetuning (FT.) the first convolution layer for discriminator on RealSR.

Method	DISTS↓	FID↓	MUSIQ↑	CLIPQA↑
Fully Finetuning the Discriminator	- (No Convergence)			
$r = 2$ (w/o Fully FT. the 1st Layer)	0.2182	121.76	62.27	0.6056
$r = 4$ (w/o Fully FT. the 1st Layer)	0.2171	119.67	68.85	0.6195
$r = 8$ (w/o Fully FT. the 1st Layer)	0.2182	120.94	68.76	0.6114
$r = 16$ (w/o Fully FT. the 1st Layer)	0.2191	120.33	68.72	0.6173
$r = 4$ (w/ Fully FT. 1st Lyr.) (Ours)	0.2129	118.41	69.90	0.6731

First Layer for the Discriminator. Tab. B.2 compares various finetuning settings for discriminator. Fully finetuning it can lead to unstable training without convergence. Compared to the rank of 2, a rank of 4 achieves notable quality gains of 0.0009, 2.09, 6.58, and 0.0139 in evaluation metrics DISTs, FID, MUSIQ, and CLIPQA, respectively. In contrast, higher ranks of 8 and 16 bring no evident improvements. Furthermore, based on the rank of 4, fully finetuning the first convolution layer further enhances performance by 0.0042, 1.26, 1.05, and 0.0536 in these four metrics. These results validate the effectiveness of our default ADC setting.

C. More Comparison Results on Benchmarks

C.1. More Quantitative Comparisons

In Tab. C.2, we compare the proposed AdcSR model against twelve state-of-the-arts, including four representative GAN-based approaches: BSRGAN [29], Real-ESRGAN [19], LDL [10], and FeMASR [3], as well as eight diffusion-based methods [11, 18, 20, 23, 24, 26–28] across three synthetic and real-world test datasets, evaluated using nine metrics [5–7, 17, 21, 25, 30, 31]. We observe that, firstly, the traditional GAN-based approaches generally perform well on reference-based metrics, particularly the fidelity measures PSNR and SSIM. Secondly, diffusion-based methods outperform these GANs in most perceptual quality met-

Algorithm 1: Training and Inference of ADC

Input: Pretrained one-step teacher; Pretrained SD models: VAE encoder \mathcal{E}_{SD} , VAE decoder \mathcal{D}_{SD} , UNet ϵ_{SD} ; Weighting factor λ_{adv} .

Stage 1: Pretraining Pruned VAE Decoder

Prune the SD VAE decoder \mathcal{D}_{SD} to obtain \mathcal{D}_{pruned} ;
Initialize \mathcal{D}_{pruned} and a discriminator as in [9, 14];

for number of training iterations **do**

Sample a batch of images \mathbf{x} ;
Obtain latent codes $\mathbf{z} = \mathcal{E}_{SD}(\mathbf{x})$;
Reconstruct images $\hat{\mathbf{x}} = \mathcal{D}_{pruned}(\mathbf{z})$;
Compute reconstruction loss [14] of \mathbf{x} and $\hat{\mathbf{x}}$;
Update \mathcal{D}_{pruned} using Adam optimizer;
Compute discriminator loss [14] of \mathbf{x} and $\hat{\mathbf{x}}$;
Update discriminator using Adam optimizer;

Stage 2: Adversarial Distillation

Prune the SD UNet ϵ_{SD} to obtain ϵ_{pruned} ;
Initialize the student model using ϵ_{pruned} and \mathcal{D}_{pruned} ;
Initialize a feature-space discriminator using ϵ_{SD} ;

for number of training iterations **do**

Sample a batch of LR-HR pairs $(\mathbf{x}_{LR}, \mathbf{x}_{HR})$;
Compute features $\mathbf{f}_{student}$ from student model;
Compute features $\mathbf{f}_{teacher}$ from teacher model;
Compute distillation loss:

$$\mathcal{L}_{distill} = \|\mathbf{f}_{student} - \mathbf{f}_{teacher}\|_1$$

Compute adversarial loss:

$$\mathcal{L}_{adv} = \text{Softplus}(-\text{Discriminator}(\mathbf{f}_{student}))$$

Compute total loss: $\mathcal{L} = \mathcal{L}_{distill} + \lambda_{adv}\mathcal{L}_{adv}$;
Update student model using Adam optimizer;
Compute features \mathbf{f}_{GT} using \mathbf{x}_{HR} ;
Compute discriminator loss:

$$\mathcal{L}_{disc} = \text{Softplus}(\text{Discriminator}(\mathbf{f}_{student})) \\ + \text{Softplus}(-\text{Discriminator}(\mathbf{f}_{GT}))$$

Update discriminator using Adam optimizer;

Inference:

Given LR image input \mathbf{x}_{LR} ;

return Super-resolved image $\hat{\mathbf{x}}_{HR} = \text{Student}(\mathbf{x}_{LR})$;

rics, showing their ability to better generate natural textures. Thirdly, AdcSR achieves competitive results, surpassing its teacher OSEDiff in most cases, which validates the effectiveness of ADC’s compression and adversarial distillation.



Figure C.1. 16 LR images from DIV2K-Val adopted in user study.

Table C.1. User study results of one-step diffusion-based methods.

Method	SinSR	OSEDiff	S3Diff	AdcSR (Ours)
Total Votes	35	149	168	160
Voting Rate (%)	7	29	33	31

C.2. More Qualitative Comparisons

Figs. C.2, C.3, and C.4 present visual comparisons across super-resolution images produced by these approaches. We observe that, firstly, GAN-based approaches generally show weaker generative capabilities than diffusion-based methods, recovering fewer details overall. Secondly, traditional multi-step SD-based methods generate rich details but may introduce artifacts, such as those observed on the spiky texture of the inflated pufferfish by StableSR, DiffBIR, SeeSR, and PASD. Thirdly, ResShift and SinSR tend to produce oversmoothed results in areas of the leaves and red flower petals, where the vein structures and textures are less distinct. This may be due to their lack of exploiting the powerful SD priors. Fourthly, AdcSR demonstrates competitive performance, generating natural and balanced details in the pufferfish and leaves, comparable to OSEDiff and S3Diff, which can subtly introduce an additional slight highlight effect on the cluster of leaves. These results comprehensively confirm the effectiveness of our approach in compressing SD-based models for Real-ISR while maintaining quality.

C.3. User Study

To further evaluate the effectiveness of our AdcSR, we conduct a user study comparing four one-step diffusion-based Real-ISR methods, including SinSR, OSEDiff, S3Diff, and AdcSR. We employ sixteen LR images from the DIV2K-Val dataset, shown in a thumbnail form in Fig. C.1. Thirty-two expert researchers are invited to choose the best super-resolution image for each test sample based on two equally weighted criteria: (1) perceptual quality, focusing on clarity, detail, and realism, and (2) content consistency with the LR input, including alignment in image structure and texture.

As reported in Tab. C.1, AdcSR achieves a high voting rate of 31%, comparable to those of 29% and 33% obtained by OSEDiff and S3Diff, both of which employ the complete SD models. Although SinSR has fewer total parameters, its super-resolution quality can be less favorable, as reflected by a lower voting rate of 7%. These results validate that our compressed diffusion-GAN hybrid maintains highly competitive Real-ISR performance while achieving $4.3\times$, $3.7\times$,

Table C.2. **Quantitative comparison among thirteen different GAN-based and diffusion-based Real-ISR approaches on both synthetic and real-world benchmarks.** “S” denotes the required number of sampling steps in the diffusion-based method.

Test Dataset	Method	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	DISTS \downarrow	FID \downarrow	NIQE \downarrow	MUSIQ \uparrow	MANIQA \uparrow	CLIPQA \uparrow
DIV2K-Val	BSRGAN	<u>24.58</u>	<u>0.6269</u>	0.3351	0.2275	44.23	4.75	61.20	0.5071	0.5247
	Real-ESRGAN	24.29	0.6371	0.3112	0.2141	37.64	<u>4.68</u>	61.06	0.5501	0.5277
	LDL	23.83	<u>0.6344</u>	0.3256	0.2227	42.29	4.86	60.04	0.5350	0.5180
	FeMASR	23.06	0.5887	0.3126	0.2057	35.87	4.74	60.83	0.5074	0.5997
	StableSR-S200	23.26	0.5726	0.3113	0.2048	<u>24.44</u>	4.76	65.92	0.6192	0.6771
	DiffBIR-S50	23.64	0.5647	0.3524	0.2128	30.72	<u>4.70</u>	65.81	0.6210	0.6704
	SeeSR-S50	23.68	0.6043	0.3194	<u>0.1968</u>	25.90	4.81	<u>68.67</u>	<u>0.6240</u>	<u>0.6936</u>
	PASD-S20	23.14	0.5505	0.3571	0.2207	29.20	4.36	68.95	0.6483	<u>0.6788</u>
	ResShift-S15	24.65	0.6181	0.3349	0.2213	36.11	6.82	61.09	0.5454	0.6071
	SinSR-S1	<u>24.41</u>	0.6018	0.3240	0.2066	35.57	6.02	62.82	0.5386	0.6471
	OSEDiff-S1	23.72	0.6108	<u>0.2941</u>	0.1976	26.32	4.71	67.97	0.6148	0.6683
	S3Diff-S1	23.52	0.5949	0.2581	0.1725	19.66	4.74	<u>68.01</u>	<u>0.6318</u>	0.7012
AdcSR-S1 (Ours)	23.74	0.6017	<u>0.2853</u>	<u>0.1899</u>	<u>25.52</u>	4.36	68.00	0.6090	0.6764	
DRealSR	BSRGAN	28.75	<u>0.8031</u>	<u>0.2883</u>	0.2142	155.63	6.52	57.14	0.4878	0.4915
	Real-ESRGAN	<u>28.64</u>	<u>0.8053</u>	<u>0.2847</u>	0.2089	147.62	6.69	54.18	0.4907	0.4422
	LDL	28.21	0.8126	0.2815	<u>0.2132</u>	155.53	7.13	53.85	0.4914	0.4310
	FeMASR	26.90	0.7572	0.3169	0.2235	157.78	<u>5.91</u>	53.74	0.4420	0.5464
	StableSR-S200	28.03	0.7536	0.3284	0.2269	148.98	6.52	58.51	0.5601	0.6356
	DiffBIR-S50	26.71	0.6571	0.4557	0.2748	166.79	6.31	61.07	0.5930	0.6395
	SeeSR-S50	28.17	0.7691	0.3189	0.2315	147.39	6.40	<u>64.93</u>	<u>0.6042</u>	0.6804
	PASD-S20	27.36	0.7073	0.3760	0.2531	156.13	5.55	<u>64.87</u>	0.6169	0.6808
	ResShift-S15	<u>28.46</u>	0.7673	0.4006	0.2656	172.26	8.12	50.60	0.4586	0.5342
	SinSR-S1	28.36	0.7515	0.3665	0.2485	170.57	6.99	55.33	0.4884	0.6383
	OSEDiff-S1	27.92	0.7835	0.2968	0.2165	<u>135.30</u>	6.49	64.65	0.5899	<u>0.6963</u>
	S3Diff-S1	27.39	0.7469	0.3129	<u>0.2108</u>	119.21	<u>6.17</u>	64.16	<u>0.6081</u>	0.7156
AdcSR-S1 (Ours)	28.10	0.7726	0.3046	0.2200	<u>134.05</u>	6.45	66.26	0.5927	<u>0.7049</u>	
RealSR	BSRGAN	26.39	0.7654	0.2670	<u>0.2121</u>	141.28	5.66	63.21	0.5399	0.5001
	Real-ESRGAN	25.69	<u>0.7616</u>	<u>0.2727</u>	<u>0.2063</u>	135.18	5.83	60.18	0.5487	0.4449
	LDL	25.28	<u>0.7567</u>	<u>0.2766</u>	<u>0.2121</u>	142.71	6.00	60.82	0.5485	0.4477
	FeMASR	25.07	0.7358	0.2942	0.2288	141.05	5.79	58.95	0.4865	0.5270
	StableSR-S200	24.70	0.7085	0.3018	0.2288	128.51	5.91	65.78	0.6221	0.6178
	DiffBIR-S50	24.75	0.6567	0.3636	0.2312	128.99	5.53	64.98	0.6246	0.6463
	SeeSR-S50	25.18	0.7216	0.3009	0.2223	125.55	<u>5.41</u>	<u>69.77</u>	<u>0.6442</u>	0.6612
	PASD-S20	25.21	0.6798	0.3380	0.2260	124.29	<u>5.41</u>	68.75	0.6487	0.6620
	ResShift-S15	<u>26.31</u>	0.7421	0.3460	0.2498	141.71	7.26	58.43	0.5285	0.5444
	SinSR-S1	<u>26.28</u>	0.7347	0.3188	0.2353	135.93	6.29	60.80	0.5385	0.6122
	OSEDiff-S1	25.15	0.7341	0.2921	0.2128	<u>123.49</u>	5.65	<u>69.09</u>	0.6326	<u>0.6693</u>
	S3Diff-S1	25.19	0.7315	0.2707	0.1994	110.34	5.33	67.92	<u>0.6398</u>	0.6761
AdcSR-S1 (Ours)	25.47	0.7301	0.2885	0.2129	<u>118.41</u>	<u>5.35</u>	69.90	0.6360	<u>0.6731</u>	

Table C.3. **Efficiency comparison on a flagship mobile device, Qualcomm SM8750 (Snapdragon 8 Gen 4),** for super-resolving an LR input image of size 128×128 with a scaling factor of 4.

Method	Latency (ms) \downarrow	Memory (MB) \downarrow	Storage (MB) \downarrow
OSEDiff	1647	1777	1693
AdcSR (Ours)	65	510	435
Reduction Rate (%)	97	71	74

and $9.3\times$ faster inference than SinSR, OSEDiff, and S3Diff, respectively, and reducing computation by 81%, 78%, and 81% in GMACs, thus verifying its appealing efficiency.

D. Efficiency Evaluation on Mobile Device

We conduct an efficiency comparison of the proposed AdcSR method against its teacher model, OSEDiff, on a flagship mobile platform, **Qualcomm SM8750 (Snapdragon 8 Gen 4)** [13], utilizing the Hexagon Digital Signal Processor (DSP). All models are evaluated using the Qualcomm AI Engine Direct Software Development Kit (SDK) [12] with 8-bit weights and 16-bit activations (W8A16) quantization for fair comparison. The results reported in Tab. C.3 demonstrate that AdcSR significantly outperforms OSEDiff in both speed and resource efficiency. Specifically, AdcSR achieves a $25\times$ acceleration in inference latency, reduces memory footprint by 71% (from 1.7GB to 0.5GB),

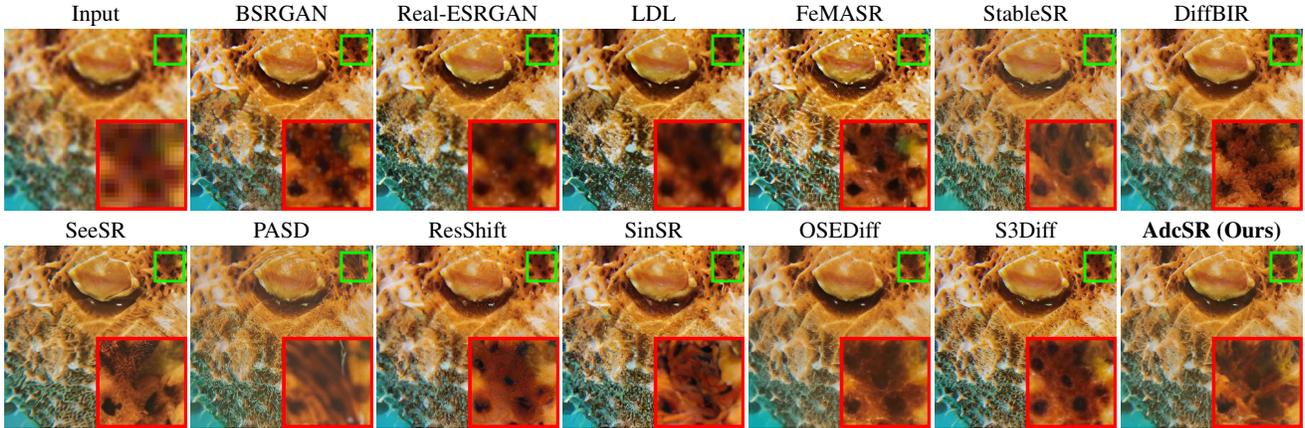


Figure C.2. **Qualitative comparison of different approaches** on an image named “0835_pch_00017” from the DIV2K-Val [18] dataset.

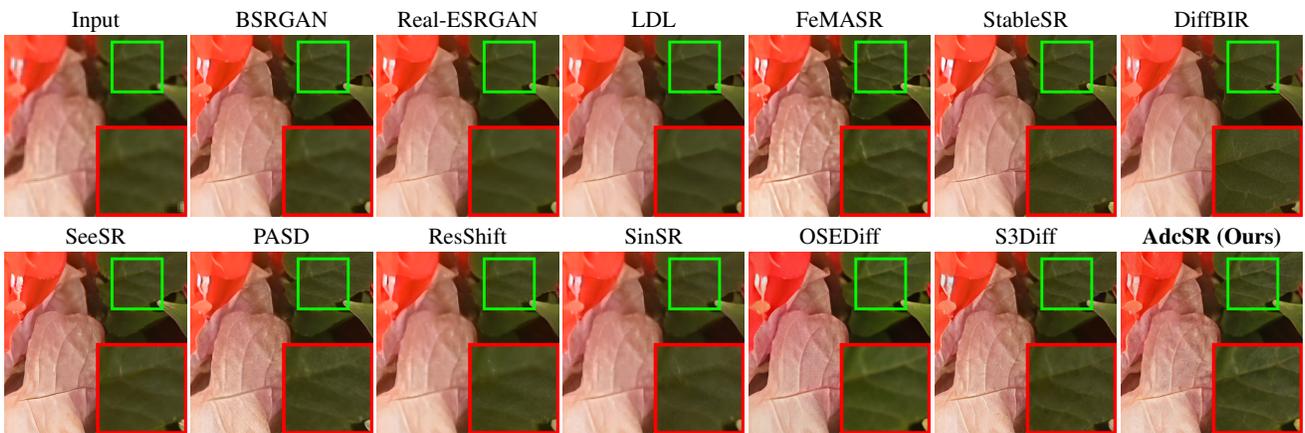


Figure C.3. **Qualitative comparison of different approaches** on a real-world image named “DSC_1599” from the DRealSR [22] dataset.

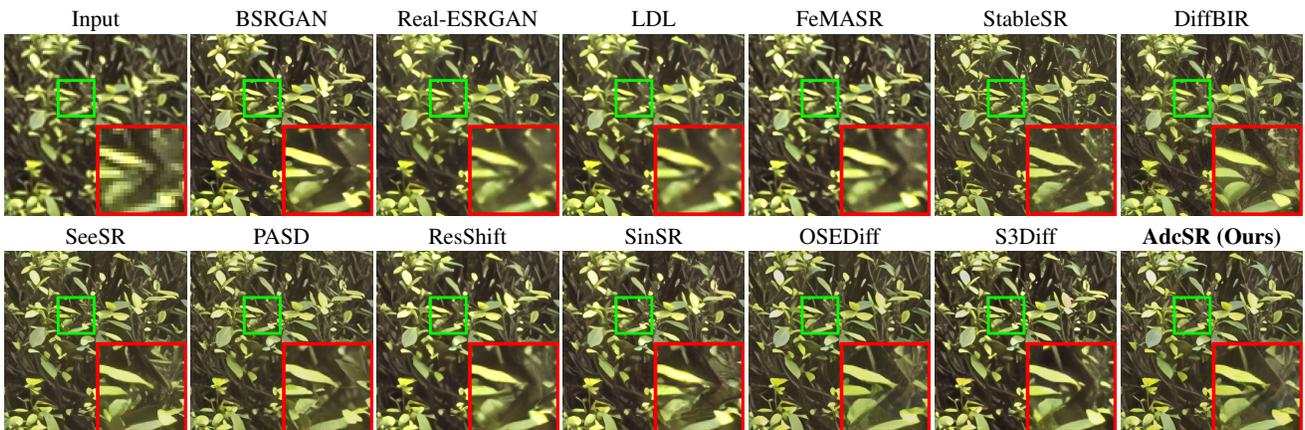


Figure C.4. **Qualitative comparison of different approaches** on a real-world image named “Nikon_013” from the RealSR [2] dataset.

and decreases storage requirements by 74% (from 1.7GB to 0.4GB). These savings are substantial for practical deployment on mobile devices, where resources are typically constrained. To summarize, AdcSR advances beyond previous SD-based one-step Real-ISR models, providing a more ef-

ficient, cost-effective solution for real mobile applications.

References

- [1] BK-SDM-v2-Small. <https://huggingface.co/nota-ai/bk-sdm-v2-small.1>

- [2] Jianrui Cai, Hui Zeng, Hongwei Yong, Zisheng Cao, and Lei Zhang. Toward real-world single image super-resolution: A new benchmark and a new model. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 3086–3095, 2019. 4
- [3] Chaofeng Chen, Xinyu Shi, Yipeng Qin, Xiaoming Li, Xiaoguang Han, Tao Yang, and Shihui Guo. Real-world blind super-resolution via feature matching with implicit high-resolution priors. In *Proceedings of the 30th ACM International Conference on Multimedia*, pages 1329–1338, 2022. 1
- [4] Trung Dao, Thuan Hoang Nguyen, Thanh Le, Duc Vu, Khoi Nguyen, Cuong Pham, and Anh Tran. Swiftbrush v2: Make your one-step diffusion model better than its teacher. In *European Conference on Computer Vision*, pages 176–192. Springer, 2024. 1
- [5] Keyan Ding, Kede Ma, Shiqi Wang, and Eero P Simoncelli. Image quality assessment: Unifying structure and texture similarity. *IEEE transactions on pattern analysis and machine intelligence*, 44(5):2567–2581, 2020. 1
- [6] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, 30, 2017.
- [7] Junjie Ke, Qifei Wang, Yilin Wang, Peyman Milanfar, and Feng Yang. Musiq: Multi-scale image quality transformer. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 5148–5157, 2021. 1
- [8] Bo-Kyeong Kim, Hyoung-Kyu Song, Thibault Castells, and Shinkook Choi. Bk-sdm: Architecturally compressed stable diffusion for efficient text-to-image generation. In *Workshop on Efficient Systems for Foundation Models@ ICML2023*, 2023. 1
- [9] Latent Diffusion Models. <https://github.com/CompVis/latent-diffusion>. 2
- [10] Jie Liang, Hui Zeng, and Lei Zhang. Details or artifacts: A locally discriminative learning approach to realistic image super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5657–5666, 2022. 1
- [11] Xinqi Lin, Jingwen He, Ziyang Chen, Zhaoyang Lyu, Bo Dai, Fanghua Yu, Wanli Ouyang, Yu Qiao, and Chao Dong. Diffbir: Towards blind image restoration with generative diffusion prior. *arXiv preprint arXiv:2308.15070*, 2023. 1
- [12] Qualcomm AI Engine Direct SDK. <https://www.qualcomm.com/developer/software/qualcomm-ai-engine-direct-sdk>. 3
- [13] Qualcomm Snapdragon Processors. <https://www.qualcomm.com/snapdragon>. 3
- [14] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10684–10695, 2022. 1, 2
- [15] Stability.ai. <https://stability.ai>. 1
- [16] Tiny AutoEncoder for Stable Diffusion. <https://huggingface.co/madebyollin/taesd>. 1
- [17] Jianyi Wang, Kelvin CK Chan, and Chen Change Loy. Exploring clip for assessing the look and feel of images. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 2555–2563, 2023. 1
- [18] Jianyi Wang, Zongsheng Yue, Shangchen Zhou, Kelvin CK Chan, and Chen Change Loy. Exploiting diffusion prior for real-world image super-resolution. *International Journal of Computer Vision*, pages 1–21, 2024. 1, 4
- [19] Xintao Wang, Liangbin Xie, Chao Dong, and Ying Shan. Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 1905–1914, 2021. 1
- [20] Yufei Wang, Wenhan Yang, Xinyuan Chen, Yaohui Wang, Lanqing Guo, Lap-Pui Chau, Ziwei Liu, Yu Qiao, Alex C Kot, and Bihan Wen. Sinsr: diffusion-based image super-resolution in a single step. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 25796–25805, 2024. 1
- [21] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004. 1
- [22] Pengxu Wei, Ziwei Xie, Hannan Lu, Zongyuan Zhan, Qixiang Ye, Wangmeng Zuo, and Liang Lin. Component divide-and-conquer for real-world image super-resolution. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VIII 16*, pages 101–117. Springer, 2020. 4
- [23] Rongyuan Wu, Lingchen Sun, Zhiyuan Ma, and Lei Zhang. One-step effective diffusion network for real-world image super-resolution. *arXiv preprint arXiv:2406.08177*, 2024. 1
- [24] Rongyuan Wu, Tao Yang, Lingchen Sun, Zhengqiang Zhang, Shuai Li, and Lei Zhang. Seesr: Towards semantics-aware real-world image super-resolution. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 25456–25467, 2024. 1
- [25] Sidi Yang, Tianhe Wu, Shuwei Shi, Shanshan Lao, Yuan Gong, Mingdeng Cao, Jiahao Wang, and Yujiu Yang. Maniqa: Multi-dimension attention network for no-reference image quality assessment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1191–1200, 2022. 1
- [26] Tao Yang, Rongyuan Wu, Peiran Ren, Xuansong Xie, and Lei Zhang. Pixel-aware stable diffusion for realistic image super-resolution and personalized stylization. *arXiv preprint arXiv:2308.14469*, 2023. 1
- [27] Zongsheng Yue, Jianyi Wang, and Chen Change Loy. Resshift: Efficient diffusion model for image super-resolution by residual shifting. *Advances in Neural Information Processing Systems*, 36, 2024.
- [28] Aiping Zhang, Zongsheng Yue, Renjing Pei, Wenqi Ren, and Xiaochun Cao. Degradation-guided one-step image super-resolution with diffusion priors. *arXiv preprint arXiv:2409.17058*, 2024. 1
- [29] Kai Zhang, Jingyun Liang, Luc Van Gool, and Radu Timofte. Designing a practical degradation model for deep blind

image super-resolution. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4791–4800, 2021. [1](#)

[30] Lin Zhang, Lei Zhang, and Alan C Bovik. A feature-enriched completely blind image quality evaluator. *IEEE Transactions on Image Processing*, 24(8):2579–2591, 2015. [1](#)

[31] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 586–595, 2018. [1](#)