Enhancing Quality of Compressed Images by Mitigating Enhancement Bias Towards Compression Domain

Qunliang Xing[1], Mai Xu[1,*] Shengxi Li[1], Xin Deng[1], Meisong Zheng, Huaida Liu, Ying Chen
[1]Beihang University

Abstract

Existing quality enhancement methods for compressed images focus on aligning the enhancement domain with the raw domain to yield realistic images. However, these methods exhibit a pervasive enhancement bias towards the compression domain, inadvertently regarding it as more realistic than the raw domain. This bias makes enhanced images closely resemble their compressed counterparts, thus degrading their perceptual quality. In this paper, we propose a simple yet effective method to mitigate this bias and enhance the quality of compressed images. Our method employs a conditional discriminator with the compressed image as a key condition, and then incorporates a domain-divergence regularization to actively distance the enhancement domain from the compression domain. Through this dual strategy, our method enables the discrimination against the compression domain, and brings the enhancement domain closer to the raw domain. Comprehensive quality evaluations confirm the superiority of our method over other state-of-the-art methods without incurring inference overheads.

1. Introduction

With the emergence of the big data era, we have been witnessing an explosive growth in images and videos. According to statistics from Domo [19], Instagram users shared approximately 66 thousand images every minute in 2022, marking an 18.3-fold increase over the past decade (*i.e.*, 3.6 thousand in 2013). Similar trends have been observed on other internet servers, including WeChat and Twitter. To efficiently store and transmit this vast volume of images, several lossy image compression standards have been developed, such as joint photographic experts group (JPEG) [40], JPEG 2000 [27], and high-efficiency video coding (HEVC) [37]. However, images compressed with these standards inevitably suffer from compression artifacts, such as the effects of ringing, blocking, and blur-
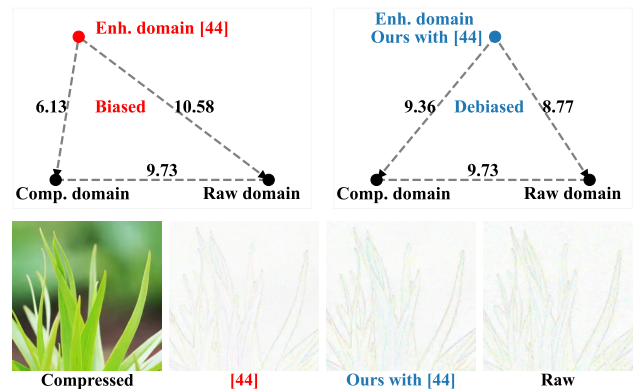
*Corresponding author.

Figure 1. *Top*: Frechet inception distance (FID) [18] scores between enhancement, compression, and raw domains on the DIV2K validation set [1]. *Bottom*: Visualization of residual to the compressed image. The results illustrate the enhancement bias towards the compression domain. Our method effectively mitigates this bias, bringing the enhancement domain closer to the raw domain.[1]

ring [34]. These artifacts may significantly degrade the quality of user experience (QoE) [20, 33].

To address the issue of quality degradation, a series of methods have been proposed to enhance the quality of compressed images. Unfortunately, the majority of these methods [11, 15, 42, 45, 47, 48, 54] focus on improving pixel-wise fidelity, limited in generating realistic textures as noted by Blau *et al*. [6]. In response to this limitation, several methods [12, 23] have ventured into enhancing perceptual quality by predominantly leveraging perception-driven super-resolution (SR) baselines that employ generative adversarial networks (GANs) [13]. By iteratively training a discriminator and a generator, the discriminator is optimized to discriminate between the enhancement and raw domains, and then guides the generator to produce realistic enhancement-domain images.

In fact, existing methods focus on aligning the enhancement domain with the raw domain, neglecting the crucial

---

[1]Three image domains correspond to the distributions of enhanced, raw, and compressed images. Distances between domains are represented in triangle plots, adhering to the triangle inequality [46] in our practice.

role of the compression domain in artifact reduction. To delve into the ramification of this oversight, we conduct a comprehensive investigation into these methods. Our findings highlight a pervasive enhancement bias towards the compression domain in two aspects. 1) *A deficiency in discriminating between the compression and raw domains*: Despite the presence of compression artifacts, these methods perceive the compression domain as more realistic than the raw domain. 2) *A tendency to generate an enhancement domain more akin to the compression domain than to the raw domain*: For existing methods, enhancement-domain images are more aligned with compression-domain images than with raw-domain images. Given these insights, it becomes clear that the compression domain should not be overlooked; instead, it is vital to harness its potential for enhancing image realism and quality.

Building upon these insights, we propose a simple yet effective method to mitigate the enhancement bias and enhance the quality of compressed images. On one hand, our method discerns the compression domain by employing a conditional discriminator [29] with the compressed image as a key condition. We theoretically prove that the domain discrimination in GAN-based quality enhancement can be facilitated by our improved discriminator. On the other hand, our method actively distinguishes the enhancement domain from the compression domain. This distancing is achieved by a domain-divergence regularization that promotes the disparity between the enhancement and compression domains. Through this dual strategy, our method enables the discrimination against the compression domain, and ultimately brings the enhancement domain closer to the raw domain to achieve superior image quality, as illustrated in Fig. 1. Finally, to verify the effectiveness of our proposed method, we conduct extensive experiments and affirm that our method proficiently addresses the problems highlighted in our observations, culminating in a substantial enhancement in image quality.

The main contributions of this paper are as follows.

1) We revisits the compression domain, an aspect often overlooked by existing methods. Through this exploration, we identify a significant enhancement bias towards the compression domain, which adversely affects the performance of quality enhancement.
2) We introduce a new method aimed at mitigating this bias. This method, characterized by its simplicity and effectiveness, is model-agnostic and yields a marked improvement in image quality without incurring any inference overhead to existing methods.
3) We conduct a series of comprehensive experiments including a thorough set of quality evaluations, demonstrating the effectiveness of the proposed method.

2. Related Works

In this section, we provide an overview of existing methods developed for enhancing the quality of compressed images. Most of these methods [11, 15, 42, 45, 47, 48, 54] have focused on improving compression fidelity using convolutional neural networks (CNNs). For example, Dong *et al*. [11] introduced a pioneering CNN-based method for enhancing the quality of JPEG-compressed images through a shallow four-layer artifacts reduction-CNN (AR-CNN). Following this, Zhang *et al*. [54] developed the denoising CNN (DnCNN) that combines a 20-layer deep network with cutting-edge techniques such as residual learning [17] and batch normalization [3]. Wang *et al*. [42] designed a 10-layer deep CNN-based auto decoder (DCAD), signifying the first CNN-based quality enhancement method for HEVC-compressed images. Then, Xing *et al*. [47] proposed the resource-efficient blind quality enhancement (RBQE) method suitable for both JPEG and HEVC-compressed images, which uniquely incorporates a dynamic inference structure for blind yet effective quality enhancement. Most recently, the defocus-aware quality enhancement (DAQE) method [48] emerged to initially discern region-wise quality differences via defocus estimation, followed by a region-wise divide-and-conquer enhancement. However, these fidelity-oriented methods, despite their advancements, still face challenges in generating realistic textures and often produce over-smoothed images [6].

To address the challenges of fidelity-oriented methods, Ghosh *et al*. [12] and Kim *et al*. [23] have focused on enhancing the perceptual quality of compressed images. They both leverage super-resolution (SR) methods [25, 43] as baselines by incorporating different GANs [13]. Specifically, Ghosh *et al*. [12] proposed the image enhancement GAN (IEGAN), a method grounded in the SRGAN method [25] with a standard GAN, for enhancing the quality of JPEG-compressed images. In [23], Kim *et al*. leveraged the enhanced SRGAN (ESRGAN) method [43] for quality enhancement by inheriting the relativistic average GAN (RaGAN) [21] and all loss functions. Recently, the Real-ESRGAN method [44] with a vanilla GAN regularized by the spectral normalization [31] has demonstrated its capability in enhancing real-world images that contain compression artifacts.

Unfortunately, the aforementioned quality enhancement methods neglect the compression domain for reducing artifacts, since they focus intently on aligning the enhancement domain with the raw domain, in a manner akin to the prevalent SR methods. Consequently, their enhanced images exhibit a strong resemblance to the compressed counterparts, leading to an enhancement bias towards the compression domain. We shall delve into this bias in the subsequent section.

Meth.	Data.	Raw	Enh.	Δ to raw	Comp.	Δ to raw
[44]	DIV2K	0.76	0.28	**-0.48**	0.84	**+0.08**
	Flickr2K	0.74	0.27	**-0.47**	0.84	**+0.10**
[43]	DIV2K	0.73	0.27	**-0.46**	0.76	**+0.03**
	Flickr2K	0.74	0.26	**-0.48**	0.78	**+0.04**

Table 1. Discriminator-evaluated realism scores for raw, enhanced, and compressed images. Higher scores indicate greater perceived realism. The results reveal that existing methods regard compressed images as more realistic than raw images.

3. Enhancement Bias Towards Compression Domain

In this section, we thoroughly examine the enhancement bias towards the compression domain, a phenomenon we consistently observe across existing perception-driven quality enhancement methods for compressed images. As aforementioned, these methods predominantly leverage perception-driven SR methods as baselines. We thus present the analysis results based on [43, 44], given their prominence in recent perception-driven quality enhancement endeavors. Note that these methods are substantially different: The former method [43] utilizes a VGG [35]-based RaGAN [21], while the latter method [44] adopts a U-Net [32]-based vanilla GAN [13] with spectral normalization regularization [31]. We re-train these methods by their default settings for quality enhancement using the images from the DIV2K training set [1]. Subsequently, these methods are evaluated using the images from the DIV2K validation and Flickr2K [39] datasets. All images are compressed by better portable graphics (BPG) [4] and JPEG [40].[2]

Observation 1: Existing quality enhancement methods fail to discriminate against compressed images, regarding them as more realistic than raw images.

Analysis. We measure the realism scores of enhanced, compressed, and raw images evaluated by discriminators. Specifically, we crop all images into patches with the size of 128×128 (in accordance with the patch size during training), which are then sent to the discriminators for evaluating realism scores. As shown in Tab. 1, the discriminators effectively discriminate between enhanced and raw images by assigning higher realism scores to the raw images. However, these discriminators struggle to distinguish between compressed and raw images; more surprisingly, the average realism score for compressed images even exceeds that of raw images. This reveals that the compression domain is perceived as more realistic, undermining the efforts to reduce compression artifacts and leading to the enhance-

---

[2]BPG implements the HEVC intra-frame compression [37]. Results for BPG-compressed images with quality parameter (QP) set to 37 are presented, while other results are comprehensively provided in the supplementary material.
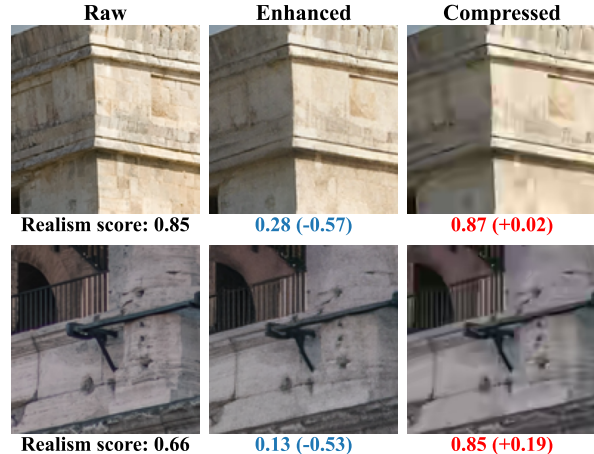

Figure 2. Visual demonstrations of realism scores for images as evaluated by [44]. The results highlight that existing methods overlook significant compression artifacts in realism evaluation.

ment bias towards the compression domain. Fig. 2 provides visual demonstrations to further illustrate our observation that the severe compression artifacts (*e.g.*, blurring) are neglected when evaluating image realism. We argue that these discriminators predominantly base their judgment on generative artifacts; as a result, the compressed image appears more realistic since the introduced compression artifacts differ significantly from generative artifacts. This insight wraps up our analysis for Observation 1.

Observation 2: For existing methods, enhancement-domain images are more aligned with compression-domain images than with raw-domain images.

Analysis. We analyze the similarity scores among compressed, raw, and enhanced images via utilizing the datasets and re-trained models above. Here, the similarity is measured by two widely-adopted metrics for assessing perceptual quality enhancement: the Frechet inception distance (FID) [18] and the learned perceptual image patch similarity (LPIPS) [55]. As illustrated in Fig. 3, there is a consistent trend where the enhancement domain exhibits closer alignment with the compression domain, compared to the raw domain. To quantitatively measure this trend, we calculate the horizontal deviation of each vertex relative to the centroid of its base, which can be equivalently expressed as:

$$\text{Deviation} := \frac{S_{\text{C,E}}^2 - S_{\text{R,E}}^2}{S_{\text{C,R}}^2} \times 100\%, \quad (1)$$

where $S_{\text{C,E}}$, $S_{\text{C,R}}$ and $S_{\text{R,E}}$ represent the similarity (in terms of either FID or LPIPS) between compression and enhancement domains, between compression and raw domains, and between raw and enhancement domains, respectively. The obtained results range from -5.71% to -98.43%, demonstrating the pervasive bias towards the compression domain, as observed consistently across datasets and similarity metrics.
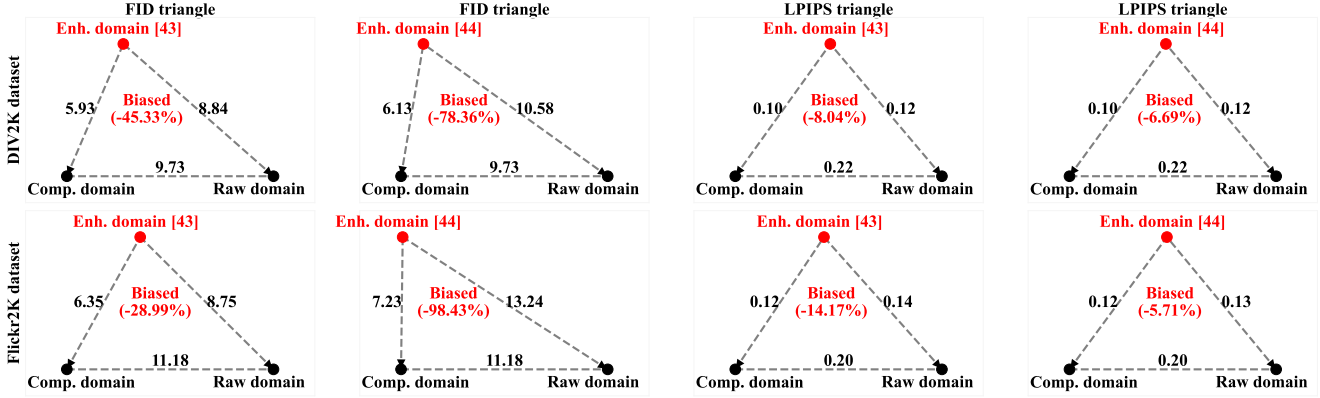
Figure 3. Similarity scores between compression, raw, and enhancement domains. Lower FID and LPIPS scores indicate greater similarity. The horizontal deviation of each vertex relative to the centroid of its base, calculated with float precision on original data, is also shown. The results underscore the enhancement bias, demonstrating a closer alignment of the enhancement domain with the compression domain than with the raw domain.

Fig. 4 provides visual demonstrations to further illustrate our observation. This comprehensive analysis brings us to the conclusion of Observation 2.

4. Mitigating Enhancement Bias for Quality Enhancement of Compressed Images

Given the above observations, we propose a simple yet effective method that aims at mitigating the enhancement bias to improve the quality of compressed images. First, our method discerns the compression domain by utilizing a conditional discriminator [29], wherein the compressed image serves as a pivotal condition. Second, our method actively distinguishes the enhancement domain from the compression domain via a proposed domain-divergence regularization. Employing this dual strategy, our method facilitates effective discrimination against the compression domain, and consequently aligns the enhancement domain more closely with the raw domain. Ultimately, by mitigating the enhancement bias, our method significantly enhances the quality of compressed images.

4.1. Discrimination against Compression Domain

Existing methods predominantly utilize GANs [13] to synthesize realistic images. Specifically, given a compressed image I_C, these methods focus on training a generative network $G(\cdot)$, which produces an enhanced image denoted as $I_E = G(I_C)$. Simultaneously, an image-realism discriminator $D(\cdot)$, designed to distinguish between the enhanced image I_E and the raw image I_R, is optimized via an adversarial loss function:

$$\max_D \mathcal{L}_D = \mathbb{E}[\log\left(1 - D(I_E)\right)] + \mathbb{E}[\log D(I_R)]. \quad (2)$$

Here, the discriminator aims to output 1 for I_R and 0 for I_E.

Figure 4. Residual comparisons between enhanced images [44] and their compressed/raw counterparts. These visualizations reveal a stronger resemblance of the enhanced images to their compressed counterparts, as indicated by the weaker residual.

However, we note that I_C is derived from compressing I_R and is subsequently used to produce I_E, suggesting that I_R and I_E have correlated distributions. Contrarily, most GANs are established on the principle of independent generated and real image distributions, typically with generated images sampled from random noise [13]. The violation of this independence may lead to mode collapse [2, 26], resulting in enhanced images that are deficient in realism. This outcome aligns with what we have identified as biased enhancement in our observations.

Indeed, I_R, I_C, and I_E constitute a basic probabilistic graphical model (PGM), comprising a three-node directed acyclic graph (DAG) [38] as depicted in Fig. 5. Recall that I_C results from compressing I_R and is later enhanced to form I_E. This underlying causal graph underpins the
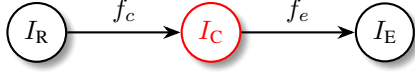
Figure 5. Causal relationships among the raw, compressed, and enhanced images, where f_c and f_e denote the image compression and quality enhancement respectively.

Data.	Raw	Enh.	Δ to raw	Comp.	Δ to raw
DIV2K	0.87	0.04	**-0.83**	0.05	**-0.82**
Flickr2K	0.86	0.03	**-0.83**	0.05	**-0.81**

Table 2. Realism scores evaluated by our method using [44] for raw, enhanced, and compressed images. Higher scores indicate greater perceived realism.

Bayesian D-separation relationship:

$$p(I_\mathrm{R}, I_\mathrm{E}|I_\mathrm{C}) = \underbrace{p(I_\mathrm{R}|I_\mathrm{C}) \cdot p(I_\mathrm{E}|I_\mathrm{R}, I_\mathrm{C})}_{\text{Chain rule}} = \underbrace{p(I_\mathrm{R}|I_\mathrm{C}) \cdot p(I_\mathrm{E}|I_\mathrm{C})}_{\text{Causality}},$$
(3)

thereby indicating I_R and I_E are conditionally independent given I_C, denoted as $I_\mathrm{R} \perp\!\!\!\perp I_\mathrm{E} \mid I_\mathrm{C}$.

This insight allows us to treat the enhanced and raw images as independent variables when I_C is provided for training both enhancement and discrimination networks. Thus, we deploy a conditional discriminator [29], incorporating the compression image as a key condition for quality enhancement:

$$\max_D \mathcal{L}_D = \mathbb{E}[\log(1 - D(I_\mathrm{E}|I_\mathrm{C}))] + \mathbb{E}[\log D(I_\mathrm{R}|I_\mathrm{C})]. \quad (4)$$

Empirical evidence strongly supports our method's effectiveness; providing the discriminator with concatenated inputs of compressed and enhanced/raw images markedly enhances realism evaluation, as evidenced in Tab. 2. More importantly, the enhanced discriminator effectively distinguishes the compression domain, which is achieved without modifying network architectures or optimization strategies.

4.2. Distancing from Compression Domain

Existing methods leverage a weighted combination of three distinct loss functions to optimize the generator $G(\cdot)$:

$$\mathcal{L}_G = \lambda_\mathrm{r}\mathcal{L}_{\text{recon.}} + \lambda_\mathrm{p}\mathcal{L}_{\text{percept.}} + \lambda_\mathrm{d}\mathcal{L}_{\text{discrim.}}. \quad (5)$$

In this formulation, $\mathcal{L}_{\text{recon.}}$ and $\mathcal{L}_{\text{percept.}}$ assess the pixel-wise and feature-wise mean absolute error (MAE) respectively, both between I_E and I_R; $\mathcal{L}_{\text{discrim.}}$ quantifies the discrimination loss associated with I_E. Notably, the compressed image I_C, indicative of the presence of compression artifacts, is largely overlooked. However, as outlined in Observation 2, enhancement-domain images in these methods are more aligned with compression-domain images than with raw-domain images.
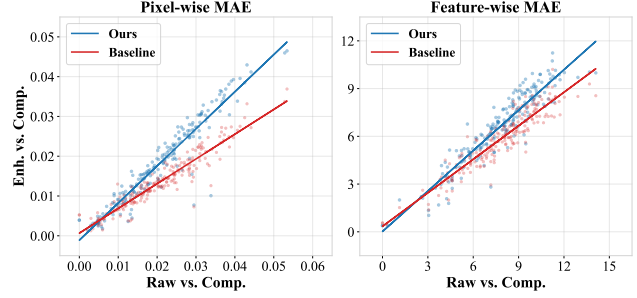

Figure 6. Visualization of domain disparity by our method using [44], as applied to randomly cropped patches from the DIV2K validation set. Linear regression lines are also illustrated.

To address this challenge, we incorporate a domain-divergence regularization that accounts for the disparities from the compression domain to both the raw and enhancement domains:

$$\mathcal{L}_R = \begin{cases} \mathcal{D}_{\mathrm{C,R}} - \mathcal{D}_{\mathrm{C,E}}, & \text{if } \mathcal{D}_{\mathrm{C,E}} < \mathcal{D}_{\mathrm{C,R}} \\ 0, & \text{otherwise} \end{cases}, \quad (6)$$

$$\mathcal{D}_{\mathrm{C,R}} = |\psi_l(I_\mathrm{R}) - \psi_l(I_\mathrm{C})|, \quad (7)$$

$$\mathcal{D}_{\mathrm{C,E}} = |\psi_l(I_\mathrm{E}) - \psi_l(I_\mathrm{C})|. \quad (8)$$

Here, ψ_l represents the final convolution layer within the l-th block of a pre-trained VGG-19 model [35]. During the training process, this regularization promotes a larger disparity between the enhancement and compression domains, especially when it is less pronounced than the disparity between the raw and compression domains; otherwise, this regularization exerts no influence on training. Consequently, the loss function evolves to:

$$\mathcal{L}_G = \lambda_\mathrm{r}\mathcal{L}_{\text{recon.}} + \lambda_\mathrm{p}\mathcal{L}_{\text{percept.}} + \lambda_\mathrm{d}\mathcal{L}_{\text{discrim.}} + \lambda_\mathcal{R}\mathcal{L}_R. \quad (9)$$

As substantiated in Fig. 6, this regularization effectively enlarges the disparity between the compression and enhancement domains, both pixel-wise and feature-wise, to a degree comparable with the compression-to-raw domain disparity.

Discussion. While many conditional-discriminator-based methods have shown advancements in tasks such as underwater enhancement [51], de-raining [53], interactive editing [8], and neural codecs [28, 50], this paper pioneers advances in the field of quality enhancement. It also provides a unique perspective to understand these advancements through our bias analysis. Our proposed method proficiently addresses the challenges of enhancement bias in a simple yet highly effective manner. Its model-agnostic nature facilitates seamless integration with a wide range of enhancement methods, without incurring inference overheads. Furthermore, an extensive array of SR methods can be enhanced and employed as baselines for perception-driven quality enhancement. We further investigate the efficacy of our method for quality enhancement in the subsequent section.

Metric	Comp.	[11]	[42]	[54]	[16]	[56]	[47]	[52]	[43]	Ours w/ [43]	[44]	Ours w/ [44]
↑ AHIQ [24]	.433	.443	.451	.453	.459	.450	.454	.454	.456	.463	.454	**.471**
↑ CLIP. [41]	.494	.538	.548	.553	.572	.551	.556	.558	.662	**.672**	.645	.660
↓ DISTS [10]	.126	.134	.143	.143	.140	.144	.142	.141	.060	**.058**	.064	.062
↓ FID [18]	9.73	10.3	10.4	10.7	10.7	10.6	10.3	10.3	8.84	**7.50**	10.6	8.77
↑ Hyper. [36]	.459	.507	.541	.549	.560	.544	.545	.553	.557	.576	.563	**.580**
↓ LPIPS [55]	.218	.215	.214	.212	.206	.213	.210	.208	.120	.114	.116	**.113**
↑ MUSIQ [22]	60.7	62.2	63.0	63.1	63.7	63.0	63.1	63.3	65.1	65.6	64.9	**66.2**
↓ NIQE [30]	4.22	4.49	4.61	4.60	4.63	4.65	4.61	4.65	2.87	**2.80**	3.14	2.92
↓ PI [7]	4.35	4.52	4.59	4.58	4.60	4.61	4.58	4.60	3.15	**3.12**	3.43	3.24
↑ PSNR (dB)	30.8	31.2	31.4	31.5	**31.7**	31.5	31.5	31.6	29.0	29.6	29.8	30.6
↑ TOPIQ [9]	.721	.716	.713	.713	.724	.711	.716	.720	.811	**.822**	.810	.821

Table 3. Objective quality of enhanced BPG-compressed images with QP set to 37. Results for other QPs and JPEG-compressed images, following similar trends, are detailed in the supplementary material. All results are to three significant figures.

5. Experiments

In this section, we present experimental results to verify the performance of our proposed method for the quality enhancement of compressed images.

5.1. Settings

Image dataset and compression. We utilize the high-quality raw images from the DIV2K dataset [1] for training and evaluation. Given the widespread adoption of BPG [4] and JPEG [40] as image compression codecs, we compress all images using these two codecs. Specifically, we employ five different compression settings for each codec: the QP for BPG is set to 27, 32, 37, 42, and 47, whereas the quality factor (QF) for JPEG is set to 10, 20, 30, 40, and 50. These settings are in alignment with the common practices in previous quality enhancement works [14, 49, 57].

Baselines and metrics. Our method is compared with a range of widely-used quality enhancement methods[3], including AR-CNN [11], DCAD [42], DnCNN [54], CBD-Net [16], RDN [56], RBQE [47], and MPRNet [52]. All of these methods are re-trained using our dataset for a fair comparison. In addition, GAN-based SR methods [43, 44][4] are adopted as baselines given their prominence in recent perception-driven quality enhancement endeavors. Accordingly, our training settings are consistent with these methods. Specifically, λ_r, λ_p, λ_d, and $\lambda_{\mathcal{R}}$ are set to $1e-2$, 1, $5e-3$, and $1e-1$ respectively for [43], and are set to 1, 1, $1e-1$, and $1e-1$ respectively for [44]. The evaluation is conducted using a comprehensive set of metrics[5], namely AHIQ [24], CLIP-IQA [41], DISTS [10], FID [18], HyperIQA [36], LPIPS [55], MUSIQ [22], NIQE [30], PI [7], PSNR, and TOPIQ [9].
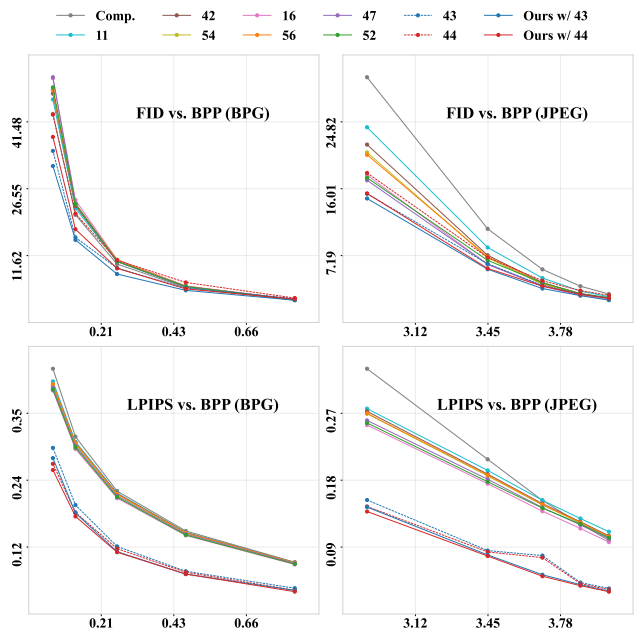

Figure 7. Rate-distortion curves comparing bits per pixel (BPP) against distortion measured by FID and LPIPS. Additional results using other metrics are provided in the supplementary material.

5.2. Evaluation

Objective quality evaluation. Tab. 3 presents the objective quality of enhanced compressed images. As detailed in Tab. 3, our method consistently enhances the perception-driven SR baseline methods [43, 44] across all metrics. For instance, the FID score by [44] stands at 10.6, which is even higher than the 9.73 baseline for compressed images, indicating a post-enhancement degradation in fidelity to raw images. On contrary, our proposed method effectively addresses this by lowering the FID score to 8.77. Furthermore, while these SR baseline methods surpass the traditional fidelity-oriented methods [11, 16, 42, 47, 52, 54, 56]

---

[3] https://github.com/ryanxingql/powerqe
[4] https://github.com/XPixelGroup/BasicSR
[5] https://github.com/chaofengc/IQA-PyTorch

Metric	[11]	[42]	[54]	[16]	[56]	[47]	[52]	[43]	Ours w/ [43]	[44]	Ours w/ [44]
AHIQ [24]	-5.63	-10.4	-12.0	-18.8	-11.6	-15.6	-15.6	-24.6	-29.3	-28.3	**-41.5**
CLIP. [41]	-29.5	-38.1	-40.8	-55.7	-42.9	-46.0	-47.2	-93.2	-93.5	-94.8	**-96.8**
DISTS [10]	+12.4	+26.5	+26.4	+21.6	+28.9	+26.2	+23.9	-75.5	**-75.8**	-69.5	-72.0
FID [18]	+6.73	+9.01	+11.5	+14.9	+11.0	+11.8	+10.7	-21.7	**-24.3**	+3.64	-13.7
Hyper. [36]	-46.3	-62.9	-65.9	-74.6	-65.8	-66.6	-68.9	-82.1	**-98.5**	-83.5	-95.3
LPIPS [55]	-5.44	-6.64	-7.92	-12.6	-6.90	-9.66	-11.4	-61.2	-63.5	-64.7	**-66.2**
MUSIQ [22]	-24.7	-34.1	-35.0	-42.5	-35.2	-36.3	-38.9	-71.9	**-81.6**	-64.2	-77.9
PSNR	-11.3	-15.3	-16.7	**-20.6**	-16.3	-18.3	-19.0	+58.7	+36.9	+30.7	+5.71
TOPIQ [9]	-1.46	-1.65	-2.03	-7.05	-1.48	-3.55	-5.19	-36.0	-39.8	-36.7	**-40.3**

Table 4. BD-BR (%) performance applied to BPG-compressed images. Negative values indicate a reduction in bit rate for the same quality. Results for JPEG-compressed images are available in the supplementary material. All results are to three significant figures.
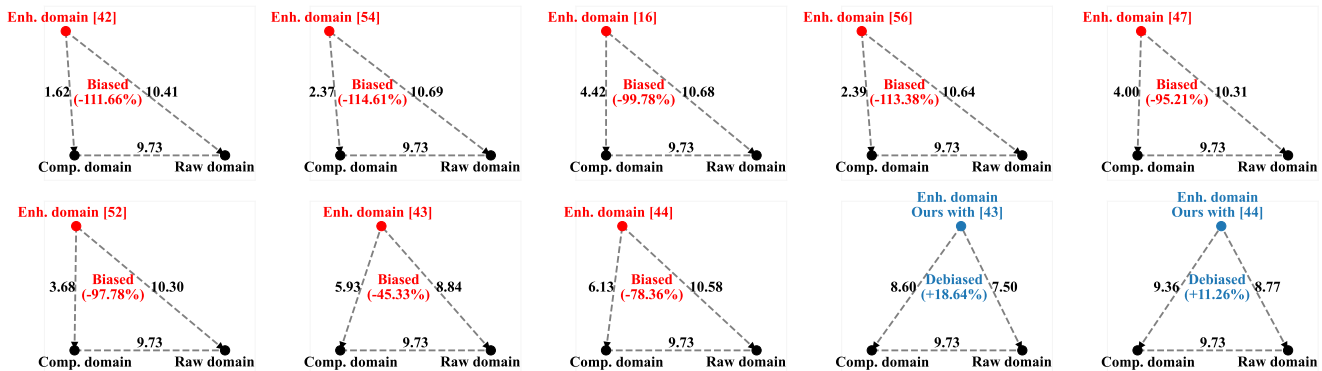

Figure 8. FID scores between three domains of the evaluation set compressed by BPG with QP set to 37. The horizontal deviation of each vertex relative to the centroid of its base, calculated with float precision on original data, is also shown.

in most metrics, they fall short in PSNR, which does not aligned with perception as noted by [6]. Nevertheless, our method also enhances the PSNR performance of these SR baseline methods by at least 0.6 dB. This advancement is particularly meaningful to recent perception-driven quality enhancement endeavors, which predominantly employ SR baseline methods. In summary, our method not only achieves state-of-the-art performance in enhancing the perceptual quality of compressed images, but also significantly improves the fidelity quality by perception-driven SR baseline methods.

Rate-distortion performance evaluation. In this section, we conduct a thorough evaluation of the rate-distortion performance of our and compared methods. Fig. 7 presents the rate-distortion curves for both BPG-compressed and JPEG-compressed images. Notably, the curves corresponding to our method are consistently lower than those of the compared methods, highlighting the enhanced image quality by our method across codecs and bitrates. Moreover, we quantify the rate-distortion performance through the Bjontegaard delta bit rate (BD-BR) [5] metric. As illustrated in Tab. 4, our method demonstrates a significant reduction in bitrate while maintaining the same perceptual quality compared to other methods, as evidenced by the minimal BD-

BR values. Even when considering PSNR, our method enhances the baseline performance by a minimum of 21.8%. In summary, our method significantly advances the state-of-the-art rate-perception performance, and also improves the rate-fidelity performance of perception-driven SR baseline methods.

Enhancement bias mitigation. The foremost objective of our method is to mitigate the enhancement bias, thereby enhancing the quality of compressed images. In this section, we examine the enhancement bias for both compared and our methods to validate the efficacy of our method in bias mitigation. Figs. 8 and 9 present the similarity scores between three domains, as measured by FID and LPIPS respectively. The results indicate that fidelity-oriented methods [16, 42, 47, 52, 54, 56] exhibit a pronounced enhancement bias towards the compression domain. This is evident as their enhancement domain closely resembles the compression domain rather than the raw domain, a phenomenon observable in the left-leaning triangles. In addition, employing perception-driven SR methods [43, 44], results in a modest reduction of the enhancement bias which remains noticeable. Conversely, our proposed method substantially diminishes this bias, aligning the enhancement domain more closely with the raw domain. Fig. 10 pro-
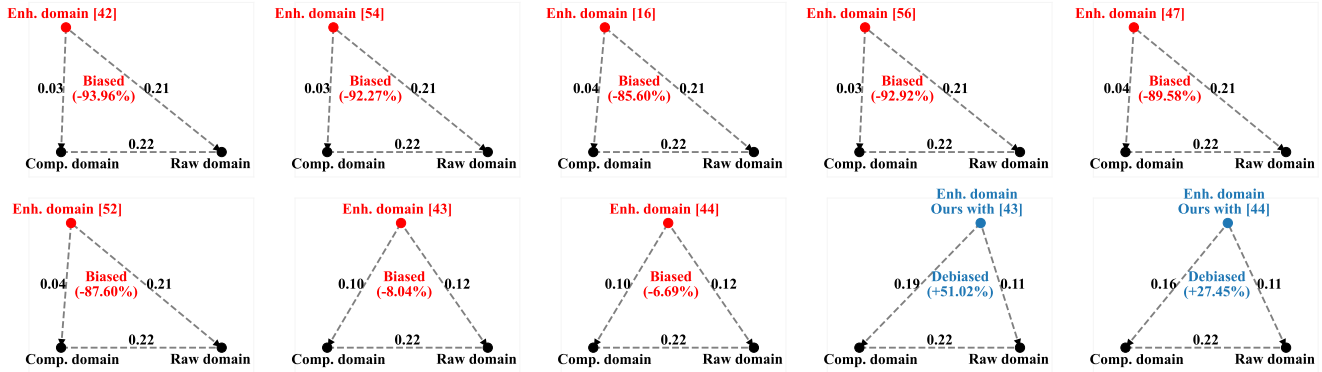
Figure 9. LPIPS scores between three domains on the evaluation set compressed by BPG with QP set to 37. The horizontal deviation of each vertex relative to the centroid of its base, calculated with float precision on original data, is also shown.
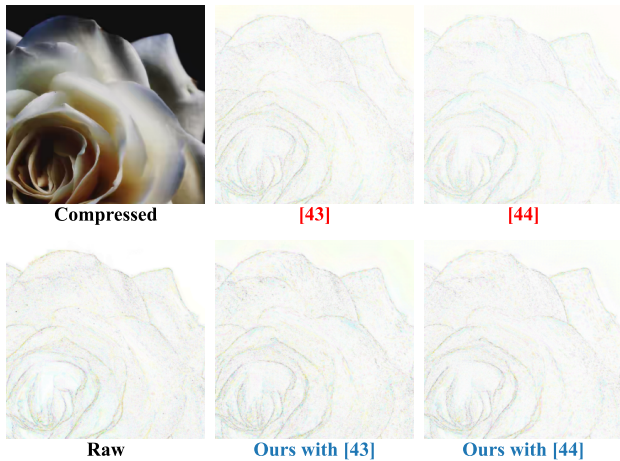

Figure 10. Visualization of residual to the compressed image. Additional examples are provided in the supplementary material.

Baseline	[43]	Δ to ori.	[44]	Δ to ori.
Vanilla	8.84/29.0	-	10.6/29.8	-
With (1)	8.05/29.4	-0.79/+0.38	9.21/30.0	-1.39/+0.24
With (2)	7.67/29.2	-1.17/+0.22	9.70/30.4	-0.90/+0.59
Ours (Full)	**7.50/29.6**	**-1.34/+0.63**	**8.77/30.6**	**-1.83/+0.79**

Table 5. FID/PSNR (dB) scores for enhanced BPG-compressed images with QP set to 37, using various configurations.

vides a visual demonstration, further illustrating the significant impact of our method in mitigating enhancement bias. In conclusion, our proposed method proficiently addresses the prevalent enhancement bias found in existing methods.

5.3. Ablation Study

In our ablation study, we focus on two pivotal mechanisms of our proposed method: (1) the conditional discriminator, which aids in discerning the compression domain, and (2) the domain-divergence regularization, designed to more distinctly separate the enhancement domain from the compression domain. As depicted in Tab. 5, the employment of either the conditional discriminator or the domain-divergence regularization leads to improvements in both FID and PSNR metrics across various baseline methods, highlighting their efficacy in enhancing the quality of compressed images. Moreover, our comprehensive method that integrates both mechanisms achieves the most significant advancements.

6. Conclusion

In this paper, we systematically explored and addressed a critical issue prevalent in existing quality enhancement methods: the enhancement bias towards the compression domain, which detrimentally affects the quality of enhanced compressed images. Our research introduced a simple yet effective method to mitigate this bias, thereby improving the quality of compressed images. Our method utilizes a conditional discriminator for discerning the compression domain, enabling a clear distinction between the raw and enhancement domains. This distinction is further reinforced by a domain-divergence regularization mechanism, effectively separating the enhancement domain from the compression domain. Experimental results validated the effectiveness of our approach, showing significant improvement in image quality by more accurately aligning with the raw domain. Our research not only achieves the state-of-the-art performance in quality enhancement, but also provides a novel perspective of evaluating the enhancement bias in future advancements of quality enhancement.

References

[1] Eirikur Agustsson and Radu Timofte. NTIRE 2017 challenge on single image super-resolution: Dataset and study. In *2017 IEEE conference on computer vision and pattern recognition workshops, CVPR workshops 2017, honolulu, HI, USA, july 21-26, 2017*, pages 1122–1131. IEEE Computer Society, 2017. tex.bibsource: dblp computer science bibliography, https://dblp.org tex.biburl: https://dblp.org/rec/conf/cvpr/AgustssonT17.bib tex.timestamp: Fri, 09 Apr 2021 18:48:31 +0200. 1, 3, 6

[2] Martin Arjovsky and Leon Bottou. Towards Principled Methods for Training Generative Adversarial Networks. In *International Conference on Learning Representations*, 2017. 4

[3] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E. Hinton. Layer Normalization, 2016. arXiv:1607.06450 [cs, stat]. 2

[4] Fabrice Bellard. Better portable graphics (BPG), 2018. 3, 6

[5] Gisle Bjontegaard. Calculation of average PSNR differences between RD-curves. *VCEG-M33*, 2001. 7

[6] Yochai Blau and Tomer Michaeli. The perception-distortion tradeoff. In *2018 IEEE/CVF conference on computer vision and pattern recognition*. IEEE, 2018. 1, 2, 7

[7] Yochai Blau, Roey Mechrez, Radu Timofte, Tomer Michaeli, and Lihi Zelnik-Manor. The 2018 PIRM Challenge on Perceptual Image Super-Resolution. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, 2018. 6

[8] Haoming Cai, Jingwen He, Yu Qiao, and Chao Dong. Toward Interactive Modulation for Photo-Realistic Image Restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 294–303, 2021. 5

[9] Chaofeng Chen, Jiadi Mo, Jingwen Hou, Haoning Wu, Liang Liao, Wenxiu Sun, Qiong Yan, and Weisi Lin. TOPIQ: A Top-down Approach from Semantics to Distortions for Image Quality Assessment, 2023. arXiv:2308.03060 [cs]. 6, 7

[10] 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*, pages 1–1, 2020. arXiv:2004.07728 [cs]. 6, 7

[11] Chao Dong, Yubin Deng, Chen Change Loy, and Xiaoou Tang. Compression artifacts reduction by a deep convolutional network. In *2015 IEEE international conference on computer vision (ICCV)*. IEEE, 2015. 1, 2, 6, 7

[12] Soumya Ghosh, Yang Hua, Sankha Subhra Mukherjee, and Neil Robertson. IEGAN: Multi-Purpose Perceptual Quality Image Enhancement Using Generative Adversarial Network. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 11–20, 2019. ISSN: 1550-5790. 1, 2

[13] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*. Curran Associates, Inc., 2014. 1, 2, 3, 4

[14] Zhenyu Guan, Qunliang Xing, Mai Xu, Ren Yang, Tie Liu, and Zulin Wang. MFQE 2.0: A New Approach for Multiframe Quality Enhancement on Compressed Video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(3):949–963, 2021. arXiv:1902.09707 [cs]. 6

[15] Jun Guo and Hongyang Chao. Building dual-domain representations for compression artifacts reduction. In *Computer vision – ECCV 2016*, pages 628–644. Springer International Publishing, 2016. 1, 2

[16] Shi Guo, Zifei Yan, Kai Zhang, Wangmeng Zuo, and Lei Zhang. Toward convolutional blind denoising of real photographs. In *2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*. IEEE, 2019. 6, 7

[17] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE conference on computer vision and pattern recognition (CVPR)*. IEEE, 2016. 2

[18] 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. In *Proceedings of the 31st international conference on neural information processing systems*, pages 6629–6640, Red Hook, NY, USA, 2017. Curran Associates Inc. Number of pages: 12 Place: Long Beach, California, USA. 1, 3, 6, 7

[19] Domo Inc. Data Never Sleeps 8.0: How much data is generated every minute?, 2020. 1

[20] International Telecommunication Union Communication Standardization Sector (ITU-T). P.10 : Vocabulary for performance, quality of service and quality of experience, 2017. 1

[21] Alexia Jolicoeur-Martineau. The relativistic discriminator: a key element missing from standard GAN. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net, 2019. 2, 3

[22] 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 (ICCV)*, pages 5148–5157, 2021. 6, 7

[23] Y. Kim, S. Cho, J. Lee, S. Jeong, J. Choi, and J. Do. Towards the Perceptual Quality Enhancement of Low Bit-rate Compressed Images. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 565–569, Los Alamitos, CA, USA, 2020. IEEE Computer Society. 1, 2

[24] Shanshan Lao, Yuan Gong, Shuwei Shi, Sidi Yang, Tianhe Wu, Jiahao Wang, Weihao Xia, and Yujiu Yang. Attentions Help CNNs See Better: Attention-Based Hybrid Image Quality Assessment Network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 1140–1149, 2022. 6, 7

[25] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In *2017 IEEE Conference*

on *Computer Vision and Pattern Recognition (CVPR)*, pages 105–114, Los Alamitos, CA, USA, 2017. IEEE Computer Society. ISSN: 1063-6919. 2

[26] Yang Li, Liangliang Shi, and Junchi Yan. IID-GAN: an IID Sampling Perspective for Regularizing Mode Collapse. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, pages 3929–3938, Macau, SAR China, 2023. International Joint Conferences on Artificial Intelligence Organization. 4

[27] M.W. Marcellin, M.J. Gormish, A. Bilgin, and M.P. Boliek. An overview of JPEG-2000. In *Proceedings DCC 2000. Data compression conference*. IEEE Comput. Soc, 2000. 1

[28] Fabian Mentzer, George Toderici, Michael Tschannen, and Eirikur Agustsson. High-Fidelity Generative Image Compression, 2020. arXiv:2006.09965 [cs, eess]. 5

[29] Mehdi Mirza and Simon Osindero. Conditional Generative Adversarial Nets, 2014. arXiv:1411.1784 [cs, stat]. 2, 4, 5

[30] Anish Mittal, Rajiv Soundararajan, and Alan C. Bovik. Making a "Completely Blind" Image Quality Analyzer. *IEEE Signal Processing Letters*, 20(3):209–212, 2013. Conference Name: IEEE Signal Processing Letters. 6

[31] Takeru Miyato, Toshiki Kataoka, Masanori Koyama, and Yuichi Yoshida. Spectral Normalization for Generative Adversarial Networks. In *International Conference on Learning Representations*, 2018. 2, 3

[32] E. Schonfeld, B. Schiele, and A. Khoreva. A U-Net Based Discriminator for Generative Adversarial Networks. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8204–8213, Los Alamitos, CA, USA, 2020. IEEE Computer Society. 3

[33] Kalpana Seshadrinathan, Rajiv Soundararajan, Alan Conrad Bovik, and Lawrence K. Cormack. Study of subjective and objective quality assessment of video. *IEEE Transactions on Image Processing*, 19(6):1427–1441, 2010. Publisher: Institute of Electrical and Electronics Engineers (IEEE). 1

[34] Mei-Yin Shen and C. C. Jay Kuo. Review of Postprocessing Techniques for Compression Artifact Removal. *Journal of Visual Communication and Image Representation*, 9(1): 2–14, 1998. 1

[35] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In *3rd international conference on learning representations, ICLR 2015, san diego, CA, USA, may 7-9, 2015, conference track proceedings*, 2015. tex.bibsource: dblp computer science bibliography, https://dblp.org tex.biburl: https://dblp.org/rec/journals/corr/SimonyanZ14a.bib tex.timestamp: Wed, 17 Jul 2019 10:40:54 +0200. 3, 5

[36] Shaolin Su, Qingsen Yan, Yu Zhu, Cheng Zhang, Xin Ge, Jinqiu Sun, and Yanning Zhang. Blindly Assess Image Quality in the Wild Guided by a Self-Adaptive Hyper Network. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3664–3673, 2020. ISSN: 2575-7075. 6, 7

[37] Gary J. Sullivan, Jens-Rainer Ohm, Woo-Jin Han, and Thomas Wiegand. Overview of the high efficiency video coding (HEVC) standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 22(12):1649–1668, 2012. Publisher: Institute of Electrical and Electronics Engineers (IEEE). 1, 3

[38] K. Thulasiraman and M.N.S. Swamy. *Graphs: Theory and Algorithms*. Wiley, 1992. 4

[39] Radu Timofte, Eirikur Agustsson, Luc Van Gool, Ming-Hsuan Yang, Lei Zhang, Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, Kyoung Mu Lee, Xintao Wang, Yapeng Tian, Ke Yu, Yulun Zhang, Shixiang Wu, Chao Dong, Liang Lin, Yu Qiao, Chen Change Loy, Woong Bae, Jae Jun Yoo, Yoseob Han, Jong Chul Ye, Jae-Seok Choi, Munchurl Kim, Yuchen Fan, Jiahui Yu, Wei Han, Ding Liu, Haichao Yu, Zhangyang Wang, Honghui Shi, Xinchao Wang, Thomas S. Huang, Yunjin Chen, Kai Zhang, Wangmeng Zuo, Zhimin Tang, Linkai Luo, Shaohui Li, Min Fu, Lei Cao, Wen Heng, Giang Bui, Truc Le, Ye Duan, Dacheng Tao, Ruxin Wang, Xu Lin, Jianxin Pang, Jinchang Xu, Yu Zhao, Xiangyu Xu, Jin-shan Pan, Deqing Sun, Yujin Zhang, Xibin Song, Yuchao Dai, Xueying Qin, Xuan-Phung Huynh, Tiantong Guo, Hojjat Seyed Mousavi, Tiep Huu Vu, Vishal Monga, Cristóvão Cruz, Karen O. Egiazarian, Vladimir Katkovnik, Rakesh Mehta, Arnav Kumar Jain, Abhinav Agarwalla, Ch V. Sai Praveen, Ruofan Zhou, Hongdiao Wen, Che Zhu, Zhiqiang Xia, Zhengtao Wang, and Qi Guo. NTIRE 2017 challenge on single image super-resolution: Methods and results. In *2017 IEEE conference on computer vision and pattern recognition workshops, CVPR workshops 2017, honolulu, HI, USA, july 21-26, 2017*, pages 1110–1121. IEEE Computer Society, 2017. tex.bibsource: dblp computer science bibliography, https://dblp.org tex.biburl: https://dblp.org/rec/conf/cvpr/TimofteAG0ZLSKN17.bib tex.timestamp: Thu, 21 Apr 2022 09:15:18 +0200. 3

[40] G.K. Wallace. The JPEG still picture compression standard. *IEEE Transactions on Consumer Electronics*, 38(1):xviii–xxxiv, 1992. Publisher: Institute of Electrical and Electronics Engineers (IEEE). 1, 3, 6

[41] Jianyi Wang, Kelvin C. K. Chan, and Chen Change Loy. Exploring CLIP for Assessing the Look and Feel of Images. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(2):2555–2563, 2023. Number: 2. 6, 7

[42] Tingting Wang, Mingjin Chen, and Hongyang Chao. A novel deep learning-based method of improving coding efficiency from the decoder-end for HEVC. In *2017 data compression conference (DCC)*. IEEE, 2017. 1, 2, 6, 7

[43] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, 2018. 2, 3, 6, 7, 8

[44] 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 (ICCV) Workshops*, pages 1905–1914, 2021. 2, 3, 4, 5, 6, 7, 8

[45] Zhangyang Wang, Ding Liu, Shiyu Chang, Qing Ling, Yingzhen Yang, and Thomas S. Huang. D3: Deep dual-domain based fast restoration of JPEG-Compressed images.

In *2016 IEEE conference on computer vision and pattern recognition (CVPR)*. IEEE, 2016. 1, 2

[46] Wikipedia contributors. Triangle inequality — Wikipedia, The Free Encyclopedia, 2023. 1

[47] Qunliang Xing, Mai Xu, Tianyi Li, and Zhenyu Guan. Early exit or not: Resource-efficient blind quality enhancement for compressed images. In *Computer vision - ECCV 2020 - 16th european conference, glasgow, UK, august 23-28, 2020, proceedings, part XVI*, pages 275–292. Springer, 2020. tex.bibsource: dblp computer science bibliography, https://dblp.org tex.biburl: https://dblp.org/rec/conf/eccv/XingXLG20.bib tex.timestamp: Thu, 17 Feb 2022 16:43:16 +0100. 1, 2, 6, 7

[48] Qunliang Xing, Mai Xu, Xin Deng, and Yichen Guo. DAQE: Enhancing the Quality of Compressed Images by Exploiting the Inherent Characteristic of Defocus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–17, 2023. Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence. 1, 2

[49] Ren Yang, Mai Xu, Zulin Wang, and Tianyi Li. Multi-frame quality enhancement for compressed video. In *2018 IEEE/CVF conference on computer vision and pattern recognition*. IEEE, 2018. 6

[50] Ren Yang, Radu Timofte, and Luc Van Gool. Perceptual Learned Video Compression with Recurrent Conditional GAN, 2022. arXiv:2109.03082 [cs, eess]. 5

[51] Xiaoli Yu, Yanyun Qu, and Ming Hong. Underwater-GAN: Underwater Image Restoration via Conditional Generative Adversarial Network. In *Pattern Recognition and Information Forensics*, pages 66–75, Cham, 2019. Springer International Publishing. 5

[52] Syed Waqas Zamir, Aditya Arora, Salman H. Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Multi-stage progressive image restoration. In *IEEE conference on computer vision and pattern recognition, CVPR 2021, virtual, june 19-25, 2021*, pages 14821–14831. Computer Vision Foundation / IEEE, 2021. tex.bibsource: dblp computer science bibliography, https://dblp.org tex.biburl: https://dblp.org/rec/conf/cvpr/ZamirA0HK0021.bib tex.timestamp: Mon, 30 Aug 2021 17:00:27 +0200. 6, 7

[53] He Zhang, Vishwanath Sindagi, and Vishal M. Patel. Image De-Raining Using a Conditional Generative Adversarial Network. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(11):3943–3956, 2020. Conference Name: IEEE Transactions on Circuits and Systems for Video Technology. 5

[54] K. Zhang, W. Zuo, S. Gu, and L. Zhang. Learning Deep CNN Denoiser Prior for Image Restoration. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2808–2817, Los Alamitos, CA, USA, 2017. IEEE Computer Society. ISSN: 1063-6919. 1, 2, 6, 7

[55] Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *2018 IEEE/CVF conference on computer vision and pattern recognition*. IEEE, 2018. 3, 6, 7

[56] Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu. Residual Dense Network for Image Super-Resolution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018. 6, 7

[57] Meisong Zheng, Qunliang Xing, Minglang Qiao, Mai Xu, Lai Jiang, Huaida Liu, and Ying Chen. Progressive Training of A Two-Stage Framework for Video Restoration. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1023–1030, New Orleans, LA, USA, 2022. IEEE. 6