

# Blur, Noise, and Compression Robust Generative Adversarial Networks

Takuhiro Kaneko<sup>1</sup> Tatsuya Harada<sup>1,2</sup>

<sup>1</sup>The University of Tokyo <sup>2</sup>RIKEN

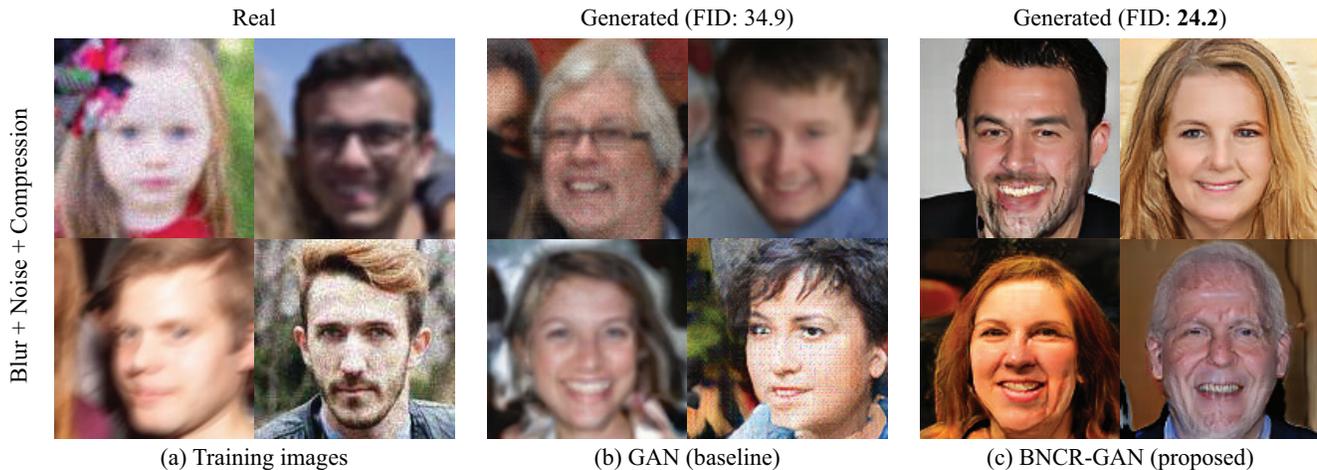


Figure 1. **Examples of blur, noise, and compression robust image generation.** Although recent GANs have shown remarkable results in image reproduction, they can recreate training images faithfully (b), despite degradation by blur, noise, and compression (a). To address this limitation, we propose *blur, noise, and compression robust GAN (BNCR-GAN)*, which can learn to generate clean images (c) even when trained with degraded images (a) and without knowledge of degradation parameters (e.g., blur kernel types, noise amounts, or quality factor values). The project page is available at <https://takuhirok.github.io/BNCR-GAN/>.

## Abstract

Generative adversarial networks (GANs) have gained considerable attention owing to their ability to reproduce images. However, they can recreate training images faithfully despite image degradation in the form of blur, noise, and compression, generating similarly degraded images. To solve this problem, the recently proposed noise robust GAN (NR-GAN) provides a partial solution by demonstrating the ability to learn a clean image generator directly from noisy images using a two-generator model comprising image and noise generators. However, its application is limited to noise, which is relatively easy to decompose owing to its additive and reversible characteristics, and its application to irreversible image degradation, in the form of blur, compression, and combination of all, remains a challenge. To address these problems, we propose blur, noise, and compression robust GAN (BNCR-GAN) that can learn a clean image generator directly from degraded images without knowledge of degradation parameters (e.g., blur kernel types, noise amounts, or quality factor values). Inspired by NR-GAN, BNCR-GAN uses a multiple-generator model composed of image, blur-kernel, noise, and quality-factor generators. However, in contrast to NR-GAN, to address irreversible characteristics, we introduce masking architectures adjusting degradation strength values in a data-driven manner using bypasses before and after degradation. Furthermore, to suppress uncertainty caused by the combina-

tion of blur, noise, and compression, we introduce adaptive consistency losses imposing consistency between irreversible degradation processes according to the degradation strengths. We demonstrate the effectiveness of BNCR-GAN through large-scale comparative studies on CIFAR-10 and a generality analysis on FFHQ. In addition, we demonstrate the applicability of BNCR-GAN in image restoration.

## 1. Introduction

Constructing generative models to generate images that are indistinguishable from real images is a fundamental problem in computer vision and machine learning. Recently, however significant advancements have been made in this regard, enabled by the emergence of deep generative models. Among them, generative adversarial networks (GANs) [22], which learn data distributions through adversarial training, have attracted considerable attention owing to their high image reproduction ability.

However, a persistent problem is that high-capacity GANs can replicate training images with high fidelity, even when the images are degraded, and they thus tend to replicate various forms of image degradation in their generated images. As shown in Figure 1, when standard GAN is trained with images degraded by blur, noise, and compression (i.e., JPEG) (Figure 1(a)), it produces similarly degraded images (Figure 1(b)) because standard GAN architectures do not consider such image degradation. This

is problematic when training images are collected in real-world scenarios (e.g., web crawling) because identifying degraded images is laborious. To address this problem, painstaking manual prescreening is often conducted.

One well-explored solution involves restoring images using an image restoration model, such as model-based image restoration methods [12, 23, 17, 57, 8, 56, 16, 54, 58, 18, 75, 42, 88, 67, 19], prior to the training of GANs. However, images restored by these methods tend to be either over- or under-restored owing to the gap between predefined and real priors.<sup>1</sup> To overcome this drawback, various learning-based methods have been developed. However, most of these methods require additional supervision for training, such as paired supervision (e.g., pairs of clean and degraded images) [87, 77, 74, 9, 65, 21, 45, 46, 28, 61, 79, 93, 94, 10, 25, 15, 78, 20, 49, 86] or set-level supervision (i.e., labels indicating whether images are degraded) [55, 52].<sup>2</sup>

AmbientGAN [4] was recently proposed as a different approach. This provides a promising solution by simulating image degradation on generated images and learning a discriminator that distinguishes a real *degraded* image from a *degraded* generated image. This formulation allows the learning of a clean image generator directly from degraded images without any pre-processing or paired/set-level supervision. However, it relies on a strong assumption that degradation parameters, such as blur kernel types, noise amounts, and quality factor values, are known in advance.

Motivated by these previous studies, we address the problem of developing a model to *learn a clean image generator directly from degraded images without knowledge of degradation parameters*. In particular, to apply the solution to real-world images, we aim to handle images degraded by a representative image degradation model [87], which addresses blur, noise, and compression in the same order as in a real image acquisition process (detailed in Equation 1). Based on this objective, we focus on blur, noise, and compression, and refer to the abovementioned problem of *blur, noise, and compression robust image generation*. We exemplify a solution using our proposed model, as shown in Figure 1(c). We aim to devise a model that can learn to generate clean images (Figure 1(c)), even when trained with blurred, noisy, and compressed images (Figure 1(a)).

Noise robust GAN (NR-GAN) [30], which uses a two-generator model consisting of noise and image generators, has provided a partial solution to this problem by demonstrating the ability to learn to generate clean images directly from noisy images. However, NR-GAN assumes that image information is lossless before and after degradation and utilizes this characteristic to decompose a degraded image into clean image and degradation components. Hence, its application is limited to noise, which has additive and reversible

<sup>1</sup>Deep image prior-based approaches [81, 68] can be used alternatively; however, they require optimization for each individual image. Pre-trained model-based approaches [68] provide another solution; however, they require the collection of clean images for training the pre-trained model.

<sup>2</sup>Self-supervised learning methods [44, 2, 48] have been also proposed; however, their application has been limited to denoising.

characteristics, and its application to irreversible degradation, in the form of blur, compression, and combination of all, remains a challenge.

To address these problems, we propose *blur, noise, and compression robust GAN (BNCR-GAN)*, that can learn a clean image generator directly from blurred, noisy, and compressed images. To solve the sub-problems, we first propose two variants: *blur robust GAN (BR-GAN)* and *compression robust GAN (CR-GAN)*, which are specific to blur and compression, respectively. Along the lines of NR-GAN, BR-GAN and CR-GAN learn a blur-kernel generator and a quality-factor generator, respectively, along with clean image generators, to learn a blur-kernel/quality-factor distribution jointly with an image distribution. However, in contrast to NR-GAN, to address the irreversible blur/compression characteristics, *masking architectures* adapting degradation strengths in a data-driven are introduced, using bypasses before and after image degradation. This architectural constraint is useful for conducting only the necessary changes through blur or compression while suppressing unnecessary changes.

The unique problem of BNCR-GAN, which is a unified model integrating BR-GAN, NR-GAN, and CR-GAN, is that it needs to handle the uncertainty caused by the combination of blur, noise, and compression. Thus, we incorporate novel losses called *adaptive consistency losses* that impose consistency between irreversible degradation processes according to the degradation strengths. This loss helps prevent the generated image from yielding unexpected artifacts, which can disappear and become unrecognizable after irreversible processes.

As the effects of blur, noise, and compression on GANs have not been sufficiently examined in previous studies, we first conducted large-scale comparative studies on CIFAR-10 [43], in which we compared diverse models under various degradation settings, in which we tested 134 total conditions. Moreover, following recent large-scale studies on GANs [47] and NR-GANs [30], we analyze a generality on a more complex dataset, that is, FFHQ [35].<sup>3</sup> Finally, we examined the applicability of BNCR-GAN in image restoration and demonstrated that, although BNCR-GAN is designed to be trained in an unsupervised manner, it is nonetheless competitive with two supervised models (i.e., CycleGAN with set-level supervision [97] and unsupervised adversarial image reconstruction (UNIR) with a pre-defined image degradation model [66]).

Our contributions are summarized as follows.

- We propose *blur, noise, and compression robust image generation*, wherein generation of clean images is learnt directly from degraded images without knowledge of degradation parameters.
- To address the sub-problems, we propose *BR-GAN* and *CR-GAN*, which train a blur-kernel generator and a

<sup>3</sup>We excluded LSUN BEDROOM [91], which was used in [47, 30], because its images were compressed with JPEG and ground-truth non-degraded images were not available.

quality-factor generator, respectively, along with clean image generators. In particular, we devise *masking architectures* to adjust the degradation strengths using bypasses before and after degradation.

- To handle all types of image degradation, we further propose *BNCR-GAN*, which unifies BR-GAN, NR-GAN, and CR-GAN as a single model. In particular, to address the uncertainty caused by the combination, we introduce *adaptive consistency losses*.
- We provide benchmark scores for these new problems through large-scale comparative studies on CIFAR-10 (in which we tested 134 conditions) and a generality analysis on FFHQ. We also demonstrate the applicability of BNCR-GAN in image restoration. The project page is available at <https://takuhirok.github.io/BNCR-GAN/>.

## 2. Related work

**Deep generative models.** Image generation is a fundamental problem in computer vision and machine learning. Recently, deep generative models, such as GANs [22], variational autoencoders [38, 72], autoregressive models [82], and flow-based models [13, 14], have garnered attention with promising results. All models have benefits and limitations. A common drawback of GANs is their training instability; however, this has been mitigated by recent progress [1, 60, 51, 3, 73, 24, 39, 63, 64, 33, 92, 6, 11, 35, 36, 96, 34]. Here, we target GANs because their design flexibility allows the incorporation of our core ideas, that is, a multiple-generator model. With regard to other models [83, 71, 62, 37], image reproducibility has improved; therefore, susceptibility to image degradation could be problematic. Applying our ideas to them remains as a potential direction for future work.

**Image restoration.** Image restoration, such as deblurring, denoising, and deblocking (or compression artifact removal), is also a fundamental problem and a large body of work exists. Typical methods are categorized as model-based methods [12, 23, 17, 57, 8, 56, 16, 54, 58, 18, 75, 42, 88, 67, 19] and learning-based methods [87, 77, 74, 9, 65, 21, 45, 46, 28, 61, 79, 93, 94, 10, 25, 15, 78, 20, 49, 86, 55, 52, 44, 2, 48]. Recently, learning-based methods have achieved better performance; however, as discussed in Section 1, most require additional supervision, such as paired or set-level supervision. In contrast, model-based methods (e.g., deblurring [18, 75, 42, 88, 67] and deblocking [19]) can be used without such supervision. However, the gap between predefined and real priors causes either over- or under-restoration and damages image fidelity. We confirm this through experiments in Section 7.

**Degradation robust image generation.** Two categories of studies on degradation robust image generation have recently become evident, including those addressing label degradation [32, 31, 29, 80] and those addressing image degradation [4, 66, 90, 50, 30]. This study belongs to the lat-

ter category. As mentioned in Section 1, AmbientGAN [4] was a pioneering model in this category; however, it is restricted by the assumption that the degradation parameters are predefined. UNIR [66] extends AmbientGAN to a conditional setting but suffers from the same restriction. Generative adversarial imputation nets [90] and MisGAN [50] addressed a similar problem in the context of data imputation. However, they relied on another strong assumption that both degraded images and the corresponding masks (i.e., blur kernel or quality factor in our case) could be obtained during training. NR-GAN [30] remedied these drawbacks by introducing a trainable noise generator; however, its application is limited to noise, which is relatively easy to decompose owing to its additive and reversible characteristics. To address these drawbacks and widen the field of degradation robust image generation, we introduce BNCR-GAN, which is applicable to irreversible image degradation without knowledge of the degradation parameters.

## 3. Notation and problem statement

We begin by defining some notation and a problem statement. Hereafter, we use superscripts  $r$  and  $g$  to denote the real and generated data, respectively. Let  $\mathbf{y}$  be a degraded image and  $\mathbf{x}$  be the corresponding clean image. Based on [87], we consider an image degradation model simulating an image acquisition process and including typical image degradation (i.e., blur, noise, and compression):<sup>4</sup>

$$\mathbf{y} = \psi((\mathbf{x} * \mathbf{k} + \mathbf{n}), q), \quad (1)$$

where  $\mathbf{k}$  is a blur kernel (e.g., an out-of-focus or motion blur kernel),  $*$  is a convolution operation,  $\mathbf{n}$  is an additive camera noise (e.g., read and shot noise), and  $\psi$  is a JPEG compression operator with a quality factor  $q \in [0, 100]$ . We aim to learn a clean image generator able to produce clean images such that  $p^g(\mathbf{x}) = p^r(\mathbf{x})$ , from partially or completely degraded training images.<sup>5</sup>

As discussed in Section 1, AmbientGAN [4] can solve this problem by simulating image degradation on generated images before passing them to a discriminator;<sup>6</sup> however, it is restricted by the need to predefine degradation simulation models (i.e.,  $\mathbf{k}^r \sim p^r(\mathbf{k})$ ,  $\mathbf{n}^r \sim p^r(\mathbf{n})$ , and  $q^r \sim p^r(q)$  must be predefined). NR-GAN [30] eliminates this restriction by making  $p^r(\mathbf{n})$  learnable. However, its

<sup>4</sup>In this study, we focus on Equation 1; however, our approach can be easily extended to more general settings by incorporating differentiable image processing (e.g., [7]). Even in this case, blur, noise, and compression remain the dominant degradation factors. Hence, we believe that our findings are not limited to a specific problem.

<sup>5</sup>For simplicity, in Sections 4–6, we only explain the case where all images are degraded (i.e., all images are sampled from  $p^r(\mathbf{y})$ ). However, as verified in Section 7, our models can be applied to images of which only parts are degraded (i.e., partial images are sampled from  $p^r(\mathbf{y})$ , while the remaining images are sampled from  $p^r(\mathbf{x})$ ), without modifying the model.

<sup>6</sup>Note that this is different from differentiable augmentation (DiffAugment) [96, 34]. DiffAugment applies augmentation to both real and generated images to learn *observable*  $p^r(\mathbf{x})$  from a few images  $\mathbf{x}^r \sim p^r(\mathbf{x})$ , whereas AmbientGAN applies image degradation *only to generated images* to learn *unobservable*  $p^r(\mathbf{x})$  from degraded images  $\mathbf{y}^r \sim p^r(\mathbf{y})$ .

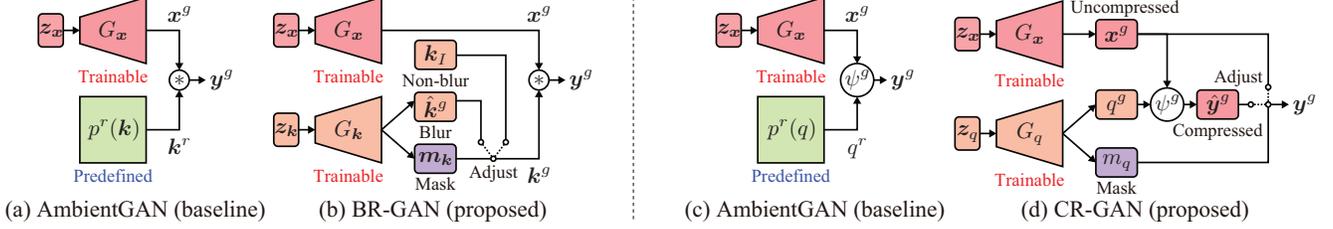


Figure 2. **Comparison of AmbientGANs (baseline), BR-GAN (proposed), and CR-GAN (proposed).** Because the discriminators are the same for all models, we only depict the generators. (a)(c) AmbientGANs assume that blur or compression simulation models are predefined. (b)(d) In order to eliminate this assumption, we introduce a blur-kernel generator  $G_k$  in BR-GAN (b) and a quality-factor generator  $G_q$  in CR-GAN (d), and train them along with clean image generators  $G_x$ . In BR-GAN (b), we introduce a masking architecture adjusting the balance between a generated kernel  $\hat{k}^g$  (i.e., blur) and an identity kernel  $k_I$  (i.e., non-blur) based on a mask  $m_k$ . Similarly, in CR-GAN (d), we incorporate a masking architecture adjusting the balance between a compressed image  $\hat{y}^g$  and an uncompressed image  $x^g$  based on a mask  $m_q$ . Note that every parameter (i.e.,  $\hat{k}^g$ ,  $m_k$ ,  $q^g$ , and  $m_q$ ) is trainable, and is determined in a data-driven manner.

application is limited to noise, where Equation 1 is simplified as  $\mathbf{y} = \mathbf{x} + \mathbf{n}$  (i.e.,  $\mathbf{x}$  and  $\mathbf{n}$  need to be decomposed additively). Considering this, first, we develop solutions to address the remaining two types of irreversible degradation, namely, blur (in which Equation 1 is rewritten as  $\mathbf{y} = \mathbf{x} * \mathbf{k}$ ) and compression (in which Equation 1 is replaced with  $\mathbf{y} = \psi(\mathbf{x}, q)$ ). Subsequently, we provide a solution for all types of degradation defined in Equation 1. Each solution is provided in Sections 4, 5, and 6.

#### 4. Blur robust GAN: BR-GAN

First, we provide a solution to *blur robust image generation*, which learns a clean image generator  $G_x$  from blurred images produced by  $\mathbf{y}^r = \mathbf{x}^r * \mathbf{k}^r$ . As discussed in Section 1, AmbientGAN [4] can solve this problem via incorporation of the blur simulation model  $\mathbf{k}^r \sim p^r(\mathbf{k})$  (Figure 2(a)). However, it needs  $p^r(\mathbf{k})$  to be predefined. To alleviate this issue and for  $p^r(\mathbf{k})$  to be learnt from the data, in BR-GAN, we replace the blur simulation model with a blur-kernel generator  $\mathbf{k}^g = G_k(\mathbf{z}_k)$  (Figure 2(b)) and train it along with  $G_x$  using the following objective function.

$$\begin{aligned} \mathcal{L}_{\text{BR-GAN}} = & \mathbb{E}_{\mathbf{y}^r \sim p^r(\mathbf{y})} [\log D_{\mathbf{y}}(\mathbf{y}^r)] \\ & + \mathbb{E}_{\mathbf{z}_x \sim p(\mathbf{z}_x), \mathbf{z}_k \sim p(\mathbf{z}_k)} [\log(1 - D_{\mathbf{y}}(G_x(\mathbf{z}_x) * G_k(\mathbf{z}_k)))] \end{aligned} \quad (2)$$

where  $D_{\mathbf{y}}$  is a discriminator distinguishing a real *blurred* image  $\mathbf{y}^r$  from a *blurred* generated image  $\mathbf{y}^g = \mathbf{x}^g * \mathbf{k}^g$  ( $\mathbf{x}^g = G_x(\mathbf{z}_x)$  and  $\mathbf{k}^g = G_k(\mathbf{z}_k)$ ).  $G_x$  and  $G_n$  are optimized by minimizing  $\mathcal{L}_{\text{BR-GAN}}$ , whereas  $D_{\mathbf{y}}$  is optimized by maximizing  $\mathcal{L}_{\text{BR-GAN}}$ .

In this formulation, it is challenging to adjust the strength of the blur kernel  $\mathbf{k}^g$  because  $\mathcal{L}_{\text{BR-GAN}}$  only regularizes a degraded image  $\mathbf{y}^g$  and an over- or under-blurred  $\mathbf{x}^g$  satisfies the solution conditions. Hence, we introduce a *masking architecture* (Figure 2(b)) adjusting  $\mathbf{k}^g$ 's strength using a bypass between blur and non-blur as follows.

$$\mathbf{k}^g = m_k \cdot \hat{\mathbf{k}}^g + (1 - m_k) \cdot \mathbf{k}_I, \quad (3)$$

where  $G_k$  is decomposed into a blur-kernel generator  $\hat{\mathbf{k}}^g = G_k^{\hat{\mathbf{k}}}(\mathbf{z}_k)$  and a mask generator  $m_k = G_k^{m_k}(\mathbf{z}_k)$ , and  $\mathbf{k}_I$  is

the identity kernel that does not affect any changes.  $m_k$  is a matrix of the same size as  $\hat{\mathbf{k}}^g$  and values within the range  $[0, 1]$ , controlling the balance between  $\hat{\mathbf{k}}^g$  (i.e., blur) and  $\mathbf{k}_I$  (i.e., non-blur). Both  $\hat{\mathbf{k}}^g$  and  $m_k$  are trainable, and their distributions are optimized in a data-driven manner.

This architecture design was inspired by the recent success of masks for adjusting generative components (e.g., [84] varies the foreground while preserving the background, and [70] varies facial expressions while preserving identity). In our case, the mask helps suppressing unnecessary changes (e.g., artifacts caused by over-/under-blurring) and allows only the necessary changes to be performed.

#### 5. Compression robust GAN: CR-GAN

Regarding *compression robust image generation*, which learns a clean image generator  $G_x$  from compressed images yielded by  $\mathbf{y}^r = \psi(\mathbf{x}^r, q^r)$ , we use roughly the same strategy as that described in Section 4. In CR-GAN, we replace the compression simulation model  $q^r \sim p^r(q)$  of AmbientGAN (Figure 2(c)) with a quality-factor generator  $q^g = G_q(\mathbf{z}_q)$  (Figure 2(d)) and train it along with  $G_x$  using the following objective function.

$$\begin{aligned} \mathcal{L}_{\text{CR-GAN}} = & \mathbb{E}_{\mathbf{y}^r \sim p^r(\mathbf{y})} [\log D_{\mathbf{y}}(\mathbf{y}^r)] \\ & + \mathbb{E}_{\mathbf{z}_x \sim p(\mathbf{z}_x), \mathbf{z}_q \sim p(\mathbf{z}_q)} [\log(1 - D_{\mathbf{y}}(\psi^g(G_x(\mathbf{z}_x), G_q(\mathbf{z}_q)))] \end{aligned} \quad (4)$$

where we use a differentiable JPEG  $\psi^g$  [76, 40], which approximates the non-differentiable rounding operation in JPEG using a differentiable continuous function. The optimization is performed similar to Equation 2.

In a typical setting, JPEG is lossy even at the best quality of 100 owing to chroma subsampling in the Y'CbCr domain and rounding in the discrete cosine transform domain. This can yield unexpected artifacts, which can disappear and become unrecognizable after compression. To alleviate this problem, we introduce a *masking architecture* (Figure 2(d)) that provides a bypass for producing a lossless image and adjusts the balance between compression and non-compression using the bypass as follows.

$$\mathbf{y}^g = m_q \hat{\mathbf{y}}^g + (1 - m_q) \mathbf{x}^g, \quad (5)$$

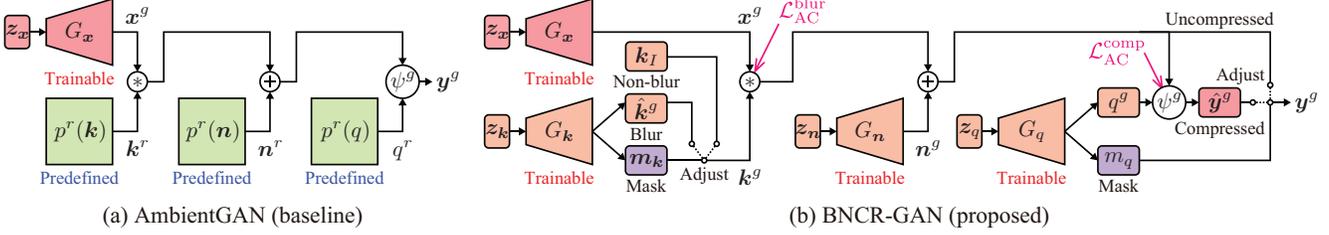


Figure 3. **Comparison of AmbientGAN (baseline) and BNCR-GAN (proposed).** We depict only the generators because the discriminators are the same for both models. (a) AmbientGAN assumes that blur, noise, and compression simulation models are predefined. (b) To reduce this assumption, we introduce a blur-kernel generator  $G_k$ , a noise generator  $G_n$ , and a quality-factor generator  $G_q$  and train them along with a clean image generator  $G_x$  with adaptive consistency losses (i.e.,  $\mathcal{L}_{AC}^{\text{blur}}$  and  $\mathcal{L}_{AC}^{\text{comp}}$ ), which impose consistency between before and after blur and compression according to the blur and compression strengths (i.e.,  $k^g$  and  $q^g$ ), respectively.

where  $G_q$  is decomposed into a quality-factor generator  $q^g = G_q^q(z_q)$  and a mask generator  $m_q = G_q^m(z_q)$ ,  $x^g = G_x(z_x)$  is an uncompressed image, and  $\hat{y}^g$  is a compressed image produced by  $\hat{y}^g = \psi^g(x^g, q^g)$ . Furthermore,  $m_q$  is a scalar within the range  $[0, 1]$ , adjusting the balance between  $\hat{y}^g$  (i.e., compression) and  $x^g$  (i.e., non-compression). In addition to  $q^g$ ,  $m_q$  is generated from  $G_q$ , and their distributions are optimized through training.

Similar to the masking architecture in BR-GAN (Section 4), this masking architecture is useful for suppressing unexpected artifacts, which are unrecognizable after compression, and allows only the necessary changes to be performed.

## 6. Blur, noise, and compression robust GAN: BNCR-GAN

Based on BR-GAN (Section 4), NR-GAN [30], and CR-GAN (Section 5), we consider *blur, noise, and compression robust image generation*, which learns a clean image generator directly from images exhibiting all types of degradation (Equation 1). To achieve this, we replace the predefined degradation simulation models of AmbientGAN (Figure 3(a)) with trainable generators (i.e., a blur-kernel generator  $G_k$ , a noise generator  $G_n$ , and a quality-factor generator  $G_q$ ) in BNCR-GAN (Figure 3(b)) and train them with a clean image generator  $G_x$  using the following objective function.

$$\begin{aligned} \mathcal{L}_{\text{BNCR-GAN}} = & \mathbb{E}_{y^r \sim p^r(y)} [\log D_y(y^r)] \\ & + \mathbb{E}_{z_x \sim p(z_x), z_k \sim p(z_k), z_n \sim p(z_n), z_q \sim p(z_q)} \\ & [\log(1 - D_y(\psi^g(G_x(z_x) * G_k(z_k) + G_n(z_n), G_q(z_q))))], \end{aligned} \quad (6)$$

The optimization is performed similar to that represented in Equation 2.

In this unified model, dealing with the uncertainty caused by combining multiple irreversible processes poses a challenge because  $\mathcal{L}_{\text{BNCR-GAN}}$  only regularizes the final output  $y^g$  and imposes no regularization on each process. To address this, we devised *adaptive consistency (AC) losses*  $\mathcal{L}_{AC} = \mathcal{L}_{AC}^{\text{blur}} + \mathcal{L}_{AC}^{\text{comp}}$ , suppressing the changes between irreversible blur and compression processes according to their

strengths.

$$\begin{aligned} \mathcal{L}_{AC}^{\text{blur}} = & \mathbb{E}_{z_x \sim p(z_x), z_k \sim p(z_k)} [e^{-\mu_k H(G_k(z_k))} \\ & \|G_x(z_x) - G_x(z_x) * G_k(z_k)\|^2], \end{aligned} \quad (7)$$

$$\begin{aligned} \mathcal{L}_{AC}^{\text{comp}} = & \mathbb{E}_{z_x \sim p(z_x), z_q \sim p(z_q)} [e^{-\mu_q \frac{100 - G_q(z_q)}{100}} \\ & \|G_x(z_x) - \psi^g(G_x(z_x), G_q(z_q))\|^2], \end{aligned} \quad (8)$$

where  $H(G_k(z_k))$  is the entropy of the blur kernel, and  $\mu_k$  and  $\mu_q$  are scale parameters. The weight term of  $\mathcal{L}_{AC}^{\text{blur}}$  (i.e.,  $e^{-\mu_k H(G_k(z_k))}$ ) increases as the generated kernel  $G_k(z_k)$  becomes closer to the identity (or non-blur) kernel. The weight term of  $\mathcal{L}_{AC}^{\text{comp}}$  (i.e.,  $e^{-\mu_q \frac{100 - G_q(z_q)}{100}}$ ) increases as the generated quality factor  $G_q(z_q)$  approaches 100. Namely,  $\mathcal{L}_{AC}$  weighs the consistency when blur and compression are weak. In our implementation, the gradients were propagated only for the left term  $G_x$  (an image before blur/compression) in  $\mathcal{L}_{AC}^{\text{blur}}$  and  $\mathcal{L}_{AC}^{\text{comp}}$ , and not for the right term (an image after blur/compression) because the right term could be regularized by the adversarial loss, whereas the left term could not be regularized in our training settings, in which clean images were not available for training.

## 7. Experiments

### 7.1. Experimental settings in comparative studies

To advance research on blur, noise, and compression robust image generation, we first conducted large-scale comparative studies on blur robust (Section 7.2), compression robust (Section 7.3), and blur, noise, and compression robust (Section 7.4) image generation. In this section, we describe the common experimental settings and provide the individual settings and results in Sections 7.2–7.4.

**Dataset.** In these studies, we used CIFAR-10 [43], which includes 60k  $32 \times 32$  natural images, split into 50k training and 10k test images, which are commonly used to examine the benchmark performance of generative models (it is also in the studies of AmbientGAN [4] and NR-GAN [30]). In addition, the image size was reasonable for a large-scale comparative study. Note that we examined the versatility on a more complex dataset (Section 7.5).

**Metrics.** Following the NR-GAN study [30], we used the Fréchet inception distance (FID) [27], which assesses the

No.	MODEL	(A)	(B)	(C)	(D)	(E)	(F)	(G)	No.	MODEL	(A)	(H)	(I)	(J)	(K)	(L)	(M)
		CLEAN	FOCUS	MOTION	(B)/(C)	$\frac{1}{4}$ (D)	$\frac{1}{2}$ (D)	$\frac{3}{4}$ (D)			CLEAN	60-80	80-100	60-100	$\frac{1}{4}$ (J)	$\frac{1}{2}$ (J)	$\frac{3}{4}$ (J)
1	GAN	<b>18.5</b>	47.9	36.1	43.1	22.4	27.4	34.1	9	GAN	<b>18.5</b>	41.7	27.8	33.3	21.0	24.7	29.0
2	AmbientGAN <sup>†</sup>	–	<b>25.5</b>	<b>23.0</b>	<b>24.2</b>	<b>18.9</b>	<b>20.7</b>	<b>21.8</b>	10	AmbientGAN <sup>†</sup>	–	43.9	36.6	40.5	20.6	22.8	<b>24.5</b>
3	P-AmbientGAN	<b>18.5</b>	48.9	35.7	41.5	22.5	27.5	34.1	11	P-AmbientGAN	35.7	42.8	36.5	42.0	34.1	30.2	36.9
4	BR-GAN	<b>17.6</b>	<b>24.8</b>	<b>22.9</b>	<b>23.4</b>	<b>18.4</b>	<b>20.2</b>	<b>21.7</b>	12	CR-GAN	18.8	<b>34.0</b>	<b>25.4</b>	<b>26.3</b>	<b>20.1</b>	<b>22.7</b>	<b>24.5</b>
5	BR-GAN w/o mask	21.9	27.6	25.4	26.6	22.5	22.9	24.0	13	CR-GAN w/o mask	34.6	41.0	35.9	38.6	33.9	31.1	34.2
6	BNCR-GAN	18.6	28.0	26.4	26.8	20.2	21.4	24.1	14	BNCR-GAN	<b>18.6</b>	<b>29.9</b>	<b>24.6</b>	<b>26.6</b>	<b>20.3</b>	<b>22.4</b>	<b>24.4</b>
7	BNCR-GAN w/o $\mathcal{L}_{AC}$	20.5	28.6	27.2	27.3	23.4	22.1	24.2	15	BNCR-GAN w/o $\mathcal{L}_{AC}$	20.5	36.1	30.6	33.4	22.4	24.1	30.6
8	Deblur+GAN <sup>†</sup>	–	31.4	24.8	28.0	19.8	22.8	25.2	16	Deblock+GAN	–	39.9	28.1	33.3	21.4	25.4	28.5

Table 1. Comparison of FID $\downarrow$  on CIFAR-10 under blur settings (left) and compression settings (right). Smaller values are preferable. We report the median score across three random seeds. Bold and bold italic fonts indicate the best and second-best scores, respectively. The symbol <sup>†</sup> indicates that the models were trained under advantageous conditions (i.e., trained using the ground-truth image degradation<sup>8</sup>).

distance between real and generative distributions. The validity of this measure has been shown in large-scale studies on GANs [53, 47]. Its sensitivity to image degradation has been also demonstrated [27, 30]. On this metric, a smaller value is preferred.

**Implementation.** We implemented GANs following the NR-GAN study [30]. Specifically, we used the ResNet architectures [26] for  $G_x$  and  $D_y$  and optimized them using a non-saturating GAN loss [22] with a real gradient penalty regularization [63]. We used similar architectures for  $G_x$  and  $G_n$ . With regard to  $G_k$  and  $G_q$ , we used three-layer multilayer perceptrons. Inspired by the findings in [30], we imposed diversity-sensitive regularization [89, 59] on  $G_k$ ,  $G_n$ , and  $G_q$ . As we aim to construct a generative model applicable to diverse degradation settings without specific tuning, we examined the performance for fixed hyperparameters values across all experiments.

## 7.2. Evaluation on blur robust image generation

**Degradation settings.** We tested three full-blur settings; here, all images were degraded by (B) out-of-focus blur with a disk kernel of radius  $r \in [0.5, 2]$ , (C) motion blur [5, 45] with a trajectory length of 5 and exposure time  $T \in \{0.1, 0.5, 1\}$ , and (D) either (B) or (C) with a selection rate of 0.5. Additionally, we tested three partial blur settings, in which one-quarter (E), one-half (F), and three-quarters (G) of images were blurred by setting (D), while the remaining were clean images. For reference, we also examined the performance of clean images (A).

**Compared models.** To examine the performance of our BR-GAN and BNCR-GAN, we compared them with three baseline GANs: standard GAN, which is agnostic to image degradation, AmbientGAN, with a ground-truth degradation simulation model, and P-AmbientGAN, which learns a degradation parameter (e.g.,  $k$  and  $q$ ) directly without using a generator. Furthermore, we conducted ablation studies on two models, including BR-GAN w/o mask and BNCR-GAN w/o  $\mathcal{L}_{AC}$ . We ablated the key components of our proposals (i.e., masking architectures and adaptive consistency losses). We also examined the performance of GAN trained with deblurred images (Deblur+GAN). We used a model-based deblurring method because typical learning-based methods cannot be trained in our setting, in which clean

images were not available. To examine the upper-bound performance, we used a non-blind deblurring method [41], which deblurred using the ground-truth kernel.

**Results.** Table 1 (left) summarizes the results. Our main findings are as follows. (i) *Comparison with baseline GANs (Nos. 1–4, 6).* As expected, AmbientGAN showed reasonable performance owing to ground-truth blur information. We found that BR-GAN, which must estimate such information through training, achieved competitive performance. BNCR-GAN, which should learn *no noise* and *non-compression* in this task, was inferior to these two models but outperformed standard GAN under all blur settings (B–G). P-AmbientGAN showed poor performance because it could not represent blur diversity, which is included in all blur settings (B–G). (ii) *Comparison with Deblur+GAN (Nos. 4, 6, 8).* We found that BR-GAN outperformed Deblur+GAN under all blur settings (B–G), and BNCR-GAN outperformed Deblur+GAN in the majority of cases (4/6). (iii) *Ablation study on BR-GAN (Nos. 4, 5).* We confirmed that the masking architecture was useful for boosting the performance under all settings (A–G). (iv) *Ablation study on BNCR-GAN (Nos. 6, 7).* We found that  $\mathcal{L}_{AC}$  was effective when parts or all of images were clean (A and E), and the negative effect did not exist across all settings (A–G).

## 7.3. Evaluation on compression robust image generation

**Degradation settings.** We tested three full compression settings; here, all images were compressed with a quality factor within the ranges of [60, 80] (H), [80, 100] (I), and [60, 100] (J). We also tested three partial compression settings, in which one-quarter (K), one-half (L), and three-quarters (M) of images were compressed by setting (J), while the remaining images were clean. For reference, we also examined the performance of clean images (A).

**Compared models.** In addition to the abovementioned models, we examined the performance of CR-GAN, CR-GAN w/o mask, and Deblock+GAN (GAN trained with deblocked images). In Deblock+GAN, we used a model-based deblocking method [19] for the same reason mentioned in Section 7.2.

<sup>8</sup>Specifically, the ground-truth information is linked to the individual image in Deblur+GAN, while this link is absent in AmbientGAN.

**Results.** Table 1 (right) lists the results. Our key findings are as follows. (i) *Comparison with baseline GANs (Nos. 9–12, 14).* CR-GAN and BNCR-GAN achieved the best or second-best scores under all compression settings (H–M). AmbientGAN was comparable to them when compression was applied to parts of images (K–M), while it underperformed them by a large margin (over 9.9) when compression was applied to all images (H–J). We consider that this occurred because of the lossy characteristics of JPEG (it remains lossy even at the best quality (Section 5)), allowing  $G_x$  to create unexpected artifacts, which were unrecognizable after the application of JPEG. In contrast, in CR-GAN, the masking architecture provides a bypass to a lossless image. This was useful for suppressing such artifacts. P-AmbientGAN showed poor performance in all compression settings (H–M) because it could not handle compression variety. (ii) *Comparison with Deblock+GAN (Nos. 12, 14, 16).* We confirmed that CR-GAN and BNCR-GAN outperformed Deblock+GAN under all compression settings (H–M). (iii) *Ablation study on CR-GAN (Nos. 12, 13).* In all settings (A and H–M), the masking architecture resulted in a positive effect. (iv) *Ablation study on BNCR-GAN (Nos. 14, 15).* We found that  $\mathcal{L}_{AC}$  improved performance under all settings (A and H–M).<sup>9</sup>

#### 7.4. Evaluation on blur, noise, and compression robust image generation

**Degradation settings.** We tested a full degradation setting (N) where all images were blurred by setting (D), noised by the setting of [7] (which consists of read and shot noise that simulated noise from a real noise dataset [69]), and compressed by setting (J). We also analyzed three partial degradation settings, in which one-quarter (O), one-half (P), and three-quarters (Q) of images were independently blurred, noised, and compressed by setting (N). Here, “independently” implies that the target image of each degradation (i.e., blur, noise, or compression) was randomly selected independently of the other degradation. For reference, we also tested clean images (A).

**Compared models.** In addition to the abovementioned models, we examined the performance of NR-GAN [30] to clarify its limitations.

**Results.** Table 2 summarizes the results. Our main findings are as follows. (i) *Comparison with baseline GANs (Nos. 17, 18, 22).* BNCR-GAN achieved the best or competitive scores in comparison to GAN and AmbientGAN. In particular, we found that BNCR-GAN outperformed AmbientGAN by a large margin (over 7.9) when the rate of degraded images was relatively high (N and Q). We consider that the

<sup>9</sup>We found that BNCR-GAN was better than CR-GAN in some cases (e.g., the setting (H)). We consider that  $\mathcal{L}_{AC}^{comp}$ , which was used in BNCR-GAN but not in CR-GAN, resulted in this behavior. To verify this statement, we examined the performance of CR-GAN with  $\mathcal{L}_{AC}^{comp}$  and found that in the setting (H), CR-GAN with  $\mathcal{L}_{AC}^{comp}$  achieves FID of 29.0, which was competitive to that obtained by BNCR-GAN (29.9) and was better than that obtained by CR-GAN without  $\mathcal{L}_{AC}^{comp}$  (34.0).

No.	MODEL	(A)	(N)	(O)	(P)	(Q)
		CLEAN	ALL	$\frac{1}{4}$ (N)	$\frac{1}{2}$ (N)	$\frac{3}{4}$ (N)
17	GAN	<b>18.5</b>	58.0	24.6	34.3	45.5
18	AmbientGAN <sup>†</sup>	–	46.8	<b>22.1</b>	<b>25.6</b>	37.3
19	BR-GAN	<b>17.6</b>	52.3	22.4	31.6	41.5
20	NR-GAN	20.0	58.1	26.1	34.6	44.3
21	CR-GAN	18.8	51.1	26.5	36.8	48.2
22	BNCR-GAN	18.6	<b>34.1</b>	<b>22.0</b>	<b>25.7</b>	<b>29.4</b>
23	BNCR-GAN w/o $\mathcal{L}_{AC}$	20.5	<b>41.0</b>	25.3	28.6	<b>34.6</b>

Table 2. **Comparison of FID↓ on CIFAR-10 under blur, noise, and compression settings.** The score calculation method and notation are the same as those in Table 1.

lossy characteristics of JPEG also affected the results on these settings. (ii) *Comparison with single-degradation robust GANs (Nos. 19–22).* In all degradation settings (N–Q), BNCR-GAN outperformed BR-GAN, NR-GAN, and CR-GAN. These results demonstrate the limitations of the single-degradation robust GANs on the datasets that contain several types of degradation. (iii) *Ablation study on BNCR-GAN (Nos. 22, 23).* We confirmed that  $\mathcal{L}_{AC}$  contributed to performance enhancement in all settings (A and N–Q).

**Summary.** Through the three comparative studies (Sections 7.2–7.4, in which we tested 134 total conditions), we found that BR-GAN, CR-GAN, and BNCR-GAN achieved better or competitive performance than AmbientGAN despite the disadvantageous training conditions. In particular, we confirmed that the key components of our proposals (i.e., masking architectures and adaptive consistency losses) contributed to achieving such high performance. In Section 7.4, we also demonstrate the limitations of single-degradation robust GANs and the importance of BNCR-GAN in the datasets exhibiting several types of degradation.

**Further analysis.** As a further analysis, we evaluated BNCR-GAN on noise robust image generation in the setting of [7]. We found that FID obtained by BNCR-GAN (20.5) was comparable to those obtained by AmbientGAN (20.0) and NR-GAN (20.3) and was better than those obtained by GAN (47.4) and BNCR-GAN without  $\mathcal{L}_{AC}$  (23.0).

#### 7.5. Evaluation on complex dataset

**Experimental settings.** Inspired by the findings of recent large-scale studies on GANs [47] and NR-GAN [30], we also investigated the performance of the proposed models on a more complex dataset, i.e., the  $128 \times 128$  version of FFHQ [35], which includes 70k face images, comprising 60k training and 10k test images.<sup>3</sup> To focus on the discussion, we compared the representative models in three representative settings in Sections 7.2–7.4 (D, J, and P).<sup>10</sup>

**Results.** Table 3 lists the results. We found that there was a similar tendency in this dataset: BR-GAN, CR-GAN, and BNCR-GAN achieved better or competitive performance

<sup>10</sup>According to the change in image size, we enlarged the blur size. We used out-of-focus blur with a disk kernel of radius  $r \in [0.5, 4]$  and motion blur with a trajectory length of 10 and exposure time  $T \in \{0.1, 0.5, 1\}$ .

MODEL	(D) BLUR	MODEL	(J) COMP.	MODEL	(P) ALL
GAN	34.1	GAN	45.5	GAN	34.9
AmbientGAN <sup>†</sup>	<b>27.7</b>	AmbientGAN <sup>†</sup>	<b>27.0</b>	AmbientGAN <sup>†</sup>	<b>24.5</b>
BR-GAN	<b>25.3</b>	CR-GAN	<b>25.7</b>	BNCR-GAN	<b>24.2</b>
Deblur+GAN <sup>†</sup>	32.9	Deblock+GAN	50.0		

Table 3. **Comparison of FID<sub>↓</sub> on FFHQ.** Smaller values are better. Experiments were conducted once owing to the time-consuming training. The notation is the same as that in Table 1.

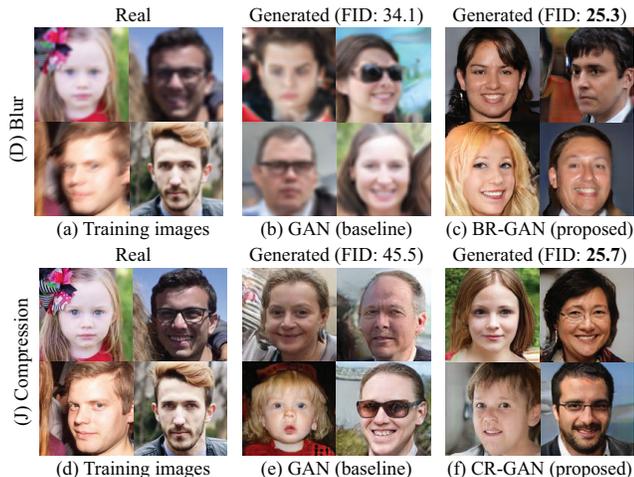


Figure 4. **Examples of blur robust image generation and compression robust image generation (best viewed at high resolution).** The upper and bottom rows show examples for the settings (D) and (J), respectively. We provide examples of blur, noise, and compression robust image generation (setting (P)) in Figure 1.

compared with AmbientGAN, which was trained under advantageous conditions, and outperformed the other baselines (i.e., GAN, Deblur+GAN, and Deblock+GAN).<sup>11</sup> We show examples of generated images for the settings (D and J) in Figure 4 and those for the setting (P) in Figure 1.

## 7.6. Application to image restoration

The recently proposed UNIR [66], a conditional extension of AmbientGAN [4], can learn an image restoration model directly from degraded images without paired or set-level supervision. However, it requires that degradation simulation models (i.e.,  $p^r(k)$ ,  $p^r(n)$ , and  $p^r(q)$ ) be predefined, similar to AmbientGAN. A reasonable solution was to utilize the blur-kernel, noise, and quality-factor generators obtained by BNCR-GAN, instead of the predefined degradation simulation models. In this section, we evaluated this approach (*BNCR-GAN+UNIR*).

**Experimental settings.** We trained the models under the same degradation settings used in Section 7.4 (settings (N–Q)). In the evaluation, we considered that test images were degraded in the most severe setting (setting (N)). With regard to BNCR-GAN+UNIR, we used the degradation gen-

<sup>11</sup>We found that Deblock+GAN was outperformed by GAN in this case. A common drawback of the model-based deblocking method is over-smoothing, and we suppose that this degraded FID more than JPEG compression.

MODEL	(O)	(P)	(Q)	(N)
	$\frac{1}{4}(N)$	$\frac{1}{2}(N)$	$\frac{3}{4}(N)$	ALL
UNIR <sup>†</sup>	<b>28.4 / 0.0204</b>	<b>22.4 / 0.0155</b>	<b>21.0 / 0.0137</b>	<b>24.8 / 0.0146</b>
BNCR-GAN+UNIR	<b>27.7 / 0.0184</b>	<b>21.5 / 0.0172</b>	<b>20.4 / 0.0152</b>	<b>24.8 / 0.0206</b>
CycleGAN <sup>‡</sup>	30.2 / 0.0226	27.1 / 0.0177	33.7 / 0.0259	N/A / N/A

Table 4. **Comparison of FID<sub>↓</sub> and LPIPS<sub>↓</sub> on CIFAR-10 in image restoration.** The results are listed as FID<sub>↓</sub> / LPIPS<sub>↓</sub>. We report the scores for the model that achieved the median FID over three training runs. Bold and bold italic fonts indicate the best and second-best scores, respectively. The symbols <sup>†</sup> and <sup>‡</sup> indicate that the models require additional supervision (i.e., predefined degradation models and set-level supervision, respectively).

erators that were obtained in the experiments described in Section 7.4. We compared BNCR-GAN+UNIR with UNIR with predefined degradation models [66] and CycleGAN with set-level supervision [97].<sup>12</sup>

**Results.** We report FID and learned perceptual image patch similarity (LPIPS) [95]<sup>13</sup> in Table 4. We found that the set-level supervision-based model (CycleGAN) was susceptible to data distribution, while degradation model-based models (i.e., UNIR and BNCR-GAN+UNIR) constantly performed better in all settings. Between UNIR and BNCR-GAN+UNIR, superiority or inferiority depended on the settings and metrics. These results indicate the potential use of BNCR-GAN+UNIR in image restoration, specifically when only degraded images are available for training.

## 8. Conclusion

We have presented a new variant of GANs called BNCR-GAN to achieve blur, noise, and compression robust image generation without knowledge of degradation parameters. The two sub-variants, BR-GAN and CR-GAN, used a two-generator model composed of an image and blur-kernel/quality-factor generators. In particular, we devised masking architectures to adjust the blur and compression strengths using bypasses before and after degradation. In BNCR-GAN, which is a unified model integrating BR-GAN, NR-GAN, and CR-GAN, we developed adaptive consistency losses to mitigate the uncertainty caused by the combination. We examined the effectiveness of the proposed models through comparative studies and a generality analysis on two benchmark datasets. We also demonstrated the applicability of BNCR-GAN in image restoration. We hope that our findings will facilitate the creation of generative models on real-world datasets that may contain several types of image degradation.

**Acknowledgment.** This work was partially supported by JST AIP Acceleration Research Grant Number JPMJCR20U3, JST CREST Grant Number JPMJCR2015, and JSPS KAKENHI Grant Number JP19H01115.

<sup>12</sup>We did not examine the performance of CycleGAN in the setting (N) because it cannot be used when clean images were not available.

<sup>13</sup>It is known that typical metrics, e.g., PSNR and SSIM [85], are inconsistent with human visual perception [95]. Hence, we used LPIPS, which was proposed as an alternative, along with FID.

## References

- [1] Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In *ICML*, 2017. 3
- [2] Joshua Batson and Loic Royer. Noise2Self: Blind denoising by self-supervision. In *ICML*, 2019. 2, 3
- [3] Marc G. Bellemare, Ivo Danihelka, Will Dabney, Shakir Mohamed, Balaji Lakshminarayanan, Stephan Hoyer, and Rémi Munos. The Cramer distance as a solution to biased Wasserstein gradients. *arXiv preprint arXiv:1705.10743*, 2017. 3
- [4] Ashish Bora, Eric Price, and Alexandros G. Dimakis. AmbientGAN: Generative models from lossy measurements. In *ICLR*, 2018. 2, 3, 4, 5, 8
- [5] Giacomo Boracchi and Alessandro Foi. Modeling the performance of image restoration from motion blur. *IEEE Trans. Image Process.*, 21(8):3502–3517, 2012. 6
- [6] Andrew Brock, Jeff Donahue, and Karen Simonyan. Large scale GAN training for high fidelity natural image synthesis. In *ICLR*, 2019. 3
- [7] Tim Brooks, Ben Mildenhall, Tianfan Xue, Jiawen Chen, Dillon Sharlet, and Jonathan T Barron. Unprocessing images for learned raw denoising. In *CVPR*, 2019. 3, 7
- [8] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. A non-local algorithm for image denoising. In *CVPR*, 2005. 2, 3
- [9] Ayan Chakrabarti. A neural approach to blind motion deblurring. In *ECCV*, 2016. 2, 3
- [10] Jingwen Chen, Jiawei Chen, Hongyang Chao, and Ming Yang. Image blind denoising with generative adversarial network based noise modeling. In *CVPR*, 2018. 2, 3
- [11] Ting Chen, Mario Lucic, Neil Houlsby, and Sylvain Gelly. On self modulation for generative adversarial networks. In *ICLR*, 2019. 3
- [12] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Trans. Image Process.*, 16(8):2080–2095, 2007. 2, 3
- [13] Laurent Dinh, David Krueger, and Yoshua Bengio. NICE: Non-linear independent components estimation. In *ICLR Workshop*, 2015. 3
- [14] Laurent Dinh, Jascha Sohl-Dickstein, and Samy Bengio. Density estimation using real NVP. In *ICLR*, 2017. 3
- [15] Chao Dong, Yubin Deng, Chen Change Loy, and Xiaoou Tang. Compression artifacts reduction by a deep convolutional network. In *ICCV*, 2015. 2, 3
- [16] Weisheng Dong, Lei Zhang, Guangming Shi, and Xin Li. Nonlocally centralized sparse representation for image restoration. *IEEE Trans. Image Process.*, 22(4):1620–1630, 2012. 2, 3
- [17] Michael Elad and Michal Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans. Image Process.*, 15(12):3736–3745, 2006. 2, 3
- [18] Rob Fergus, Barun Singh, Aaron Hertzmann, Sam T Roweis, and William T Freeman. Removing camera shake from a single photograph. *ACM Trans. Graph.*, 25(3):787–794, 2006. 2, 3
- [19] Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Pointwise shape-adaptive DCT for high-quality denoising and deblocking of grayscale and color images. *IEEE Trans. Image Process.*, 16(5):1395–1411, 2007. 2, 3, 6
- [20] Leonardo Galteri, Lorenzo Seidenari, Marco Bertini, and Alberto Del Bimbo. Deep generative adversarial compression artifact removal. In *ICCV*, 2017. 2, 3
- [21] Dong Gong, Jie Yang, Lingqiao Liu, Yaning Zhang, Ian Reid, Chunhua Shen, Anton Van Den Hengel, and Qinfeng Shi. From motion blur to motion flow: A deep learning solution for removing heterogeneous motion blur. In *CVPR*, 2017. 2, 3
- [22] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *NIPS*, 2014. 1, 3, 6
- [23] Shuhang Gu, Lei Zhang, Wangmeng Zuo, and Xiangchu Feng. Weighted nuclear norm minimization with application to image denoising. In *CVPR*, 2014. 2, 3
- [24] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron Courville. Improved training of Wasserstein GANs. In *NIPS*, 2017. 3
- [25] Shi Guo, Zifei Yan, Kai Zhang, Wangmeng Zuo, and Lei Zhang. Toward convolutional blind denoising of real photographs. In *CVPR*, 2019. 2, 3
- [26] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *CVPR*, 2016. 6
- [27] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, Günter Klambauer, and Sepp Hochreiter. GANs trained by a two time-scale update rule converge to a Nash equilibrium. In *NIPS*, 2017. 5, 6
- [28] Viren Jain and Sebastian Seung. Natural image denoising with convolutional networks. In *NIPS*, 2009. 2, 3
- [29] Takuhiro Kaneko and Tatsuya Harada. Label-noise robust multi-domain image-to-image translation. *arXiv preprint arXiv:1905.02185*, 2019. 3
- [30] Takuhiro Kaneko and Tatsuya Harada. Noise robust generative adversarial networks. In *CVPR*, 2020. 2, 3, 5, 6, 7
- [31] Takuhiro Kaneko, Yoshitaka Ushiku, and Tatsuya Harada. Class-distinct and class-mutual image generation with GANs. In *BMVC*, 2019. 3
- [32] Takuhiro Kaneko, Yoshitaka Ushiku, and Tatsuya Harada. Label-noise robust generative adversarial networks. In *CVPR*, 2019. 3
- [33] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of GANs for improved quality, stability, and variation. In *ICLR*, 2018. 3
- [34] Tero Karras, Miika Aittala, Janne Hellsten, Samuli Laine, Jaakko Lehtinen, and Timo Aila. Training generative adversarial networks with limited data. In *NeurIPS*, 2020. 3
- [35] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *CVPR*, 2019. 2, 3, 7
- [36] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of StyleGAN. In *CVPR*, 2020. 3
- [37] Diederik P. Kingma and Prafulla Dhariwal. Glow: Generative flow with invertible 1x1 convolutions. In *NeurIPS*, 2018. 3
- [38] Diederik P. Kingma and Max Welling. Auto-encoding variational bayes. In *ICLR*, 2014. 3
- [39] Naveen Kodali, Jacob Abernethy, James Hays, and Zsolt Kira. On convergence and stability of GANs. *arXiv preprint arXiv:1705.07215*, 2017. 3

- [40] Pawel Korus and Nasir Memon. Content authentication for neural imaging pipelines: End-to-end optimization of photo provenance in complex distribution channels. In *CVPR*, 2019. 4
- [41] Dilip Krishnan and Rob Fergus. Fast image deconvolution using hyper-Laplacian priors. In *NIPS*, 2009. 6
- [42] Dilip Krishnan, Terence Tay, and Rob Fergus. Blind deconvolution using a normalized sparsity measure. In *CVPR*, 2011. 2, 3
- [43] Alex Krizhevsky. Learning multiple layers of features from tiny images. *Technical report*, 2009. 2, 5
- [44] Alexander Krull, Tim-Oliver Buchholz, and Florian Jug. Noise2Void – Learning denoising from single noisy images. In *CVPR*, 2019. 2, 3
- [45] Orest Kupyn, Volodymyr Budzan, Mykola Mykhailych, Dmytro Mishkin, and Jiří Matas. DeblurGAN: Blind motion deblurring using conditional adversarial networks. In *CVPR*, 2018. 2, 3, 6
- [46] Orest Kupyn, Tetiana Martyniuk, Junru Wu, and Zhangyang Wang. DeblurGAN-v2: Deblurring (orders-of-magnitude) faster and better. In *ICCV*, 2019. 2, 3
- [47] Karol Kurach, Mario Lucic, Xiaohua Zhai, Marcin Michalski, and Sylvain Gelly. A large-scale study on regularization and normalization in GANs. In *ICML*, 2019. 2, 6, 7
- [48] Samuli Laine, Tero Karras, Jaakko Lehtinen, and Timo Aila. High-quality self-supervised deep image denoising. In *NeurIPS*, 2019. 2, 3
- [49] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2Noise: Learning image restoration without clean data. In *ICML*, 2018. 2, 3
- [50] Steven Cheng-Xian Li, Bo Jiang, and Benjamin Marlin. MisGAN: Learning from incomplete data with generative adversarial networks. In *ICLR*, 2019. 3
- [51] Jae Hyun Lim and Jong Chul Ye. Geometric GAN. *arXiv preprint arXiv:1705.02894*, 2017. 3
- [52] Boyu Lu, Jun-Cheng Chen, and Rama Chellappa. Unsupervised domain-specific deblurring via disentangled representations. In *CVPR*, 2019. 2, 3
- [53] Mario Lucic, Karol Kurach, Marcin Michalski, Sylvain Gelly, and Olivier Bousquet. Are GANs created equal? A large-scale study. In *NeurIPS*, 2018. 6
- [54] Florian Luisier, Thierry Blu, and Michael Unser. Image denoising in mixed Poisson-Gaussian noise. *IEEE Trans. Image Process.*, 20(3):696–708, 2010. 2, 3
- [55] Nimisha Thekke Madam, Sunil Kumar, and Rajagopalan A.N. Unsupervised class-specific deblurring. In *ECCV*, 2018. 2, 3
- [56] Julien Mairal, Francis Bach, Jean Ponce, Guillermo Sapiro, and Andrew Zisserman. Non-local sparse models for image restoration. In *ICCV*, 2009. 2, 3
- [57] Julien Mairal, Michael Elad, and Guillermo Sapiro. Sparse representation for color image restoration. *IEEE Trans. Image Process.*, 17(1):53–69, 2007. 2, 3
- [58] Markku Makitalo and Alessandro Foi. Optimal inversion of the generalized Anscombe transformation for Poisson-Gaussian noise. *IEEE Trans. Image Process.*, 22(1):91–103, 2012. 2, 3
- [59] Qi Mao, Hsin-Ying Lee, Hung-Yu Tseng, Siwei Ma, and Ming-Hsuan Yang. Mode seeking generative adversarial networks for diverse image synthesis. In *CVPR*, 2019. 6
- [60] Xudong Mao, Qing Li, Haoran Xie, Raymond Y.K. Lau, Zhen Wang, and Stephen Paul Smolley. Least squares generative adversarial networks. In *ICCV*, 2017. 3
- [61] Xiaojiao Mao, Chunhua Shen, and Yu-Bin Yang. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. In *NIPS*, 2016. 2, 3
- [62] Jacob Menick and Nal Kalchbrenner. Generating high fidelity images with subscale pixel networks and multidimensional upscaling. In *ICLR*, 2019. 3
- [63] Lars Mescheder, Andreas Geiger, and Sebastian Nowozin. Which training methods for GANs do actually converge? In *ICML*, 2018. 3, 6
- [64] Takeru Miyato, Toshiki Kataoka, Masanori Koyama, and Yuichi Yoshida. Spectral normalization for generative adversarial networks. In *ICLR*, 2018. 3
- [65] Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee. Deep multi-scale convolutional neural network for dynamic scene deblurring. In *CVPR*, 2017. 2, 3
- [66] Arthur Pajot, Emmanuel de Bezenac, and Patrick Gallinari. Unsupervised adversarial image reconstruction. In *ICLR*, 2019. 2, 3, 8
- [67] Jinshan Pan, Deqing Sun, Hanspeter Pfister, and Ming-Hsuan Yang. Blind image deblurring using dark channel prior. In *CVPR*, 2016. 2, 3
- [68] Xingang Pan, Xiaohang Zhan, Bo Dai, Dahua Lin, Chen Change Loy, and Ping Luo. Exploiting deep generative prior for versatile image restoration and manipulation. In *ECCV*, 2020. 2
- [69] Tobias Plotz and Stefan Roth. Benchmarking denoising algorithms with real photographs. In *CVPR*, 2017. 7
- [70] Albert Pumarola, Antonio Agudo, Alex M Martinez, Alberto Sanfeliu, and Francesc Moreno-Noguer. GANimation: Anatomically-aware facial animation from a single image. In *ECCV*, 2018. 4
- [71] Ali Razavi, Aäron van den Oord, and Oriol Vinyals. Generating diverse high-fidelity images with VQ-VAE-2. In *NeurIPS*, 2019. 3
- [72] Danilo Jimenez Rezende, Shakir Mohamed, and Daan Wierstra. Stochastic backpropagation and approximate inference in deep generative models. In *ICML*, 2014. 3
- [73] Tim Salimans, Han Zhang, Alec Radford, and Dimitris Metaxas. Improving GANs using optimal transport. In *ICLR*, 2018. 3
- [74] Christian J Schuler, Michael Hirsch, Stefan Harmeling, and Bernhard Schölkopf. Learning to deblur. *IEEE Trans. Pattern Anal. Mach. Intell.*, 38(7):1439–1451, 2015. 2, 3
- [75] Qi Shan, Jiaya Jia, and Aseem Agarwala. High-quality motion deblurring from a single image. *ACM Trans. Graph.*, 27(3):1–10, 2008. 2, 3
- [76] Richard Shin and Dawn Song. JPEG-resistant adversarial images. In *NIPS Workshop*, 2017. 4
- [77] Jian Sun, Wenfei Cao, Zongben Xu, and Jean Ponce. Learning a convolutional neural network for non-uniform motion blur removal. In *CVPR*, 2015. 2, 3
- [78] Pavel Svoboda, Michal Hradis, David Barina, and Pavel Zemcik. Compression artifacts removal using convolutional neural networks. In *WSCG*, 2016. 2, 3
- [79] Ying Tai, Jian Yang, Xiaoming Liu, and Chunyan Xu. MemNet: A persistent memory network for image restoration. In *ICCV*, 2017. 2, 3

- [80] Kiran Koshy Thekumparampil, Ashish Khetan, Zinan Lin, and Sewoong Oh. Robustness of conditional GANs to noisy labels. In *NeurIPS*, 2018. 3
- [81] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep image prior. In *CVPR*, 2018. 2
- [82] Aäron van den Oord, Nal Kalchbrenner, and Koray Kavukcuoglu. Pixel recurrent neural networks. In *ICML*, 2016. 3
- [83] Aäron van den Oord, Oriol Vinyals, and Koray Kavukcuoglu. Neural discrete representation learning. In *NIPS*, 2017. 3
- [84] Carl Vondrick, Hamed Pirsiavash, and Antonio Torralba. Generating videos with scene dynamics. In *NIPS*, 2016. 4
- [85] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.*, 13(4):600–612, 2004. 8
- [86] Zhihao Xia and Ayan Chakrabarti. Training image estimators without image ground truth. In *NeurIPS*, 2019. 2, 3
- [87] Li Xu, Jimmy SJ Ren, Ce Liu, and Jiaya Jia. Deep convolutional neural network for image deconvolution. In *NIPS*, 2014. 2, 3
- [88] Li Xu, Shicheng Zheng, and Jiaya Jia. Unnatural  $L_0$  sparse representation for natural image deblurring. In *CVPR*, 2013. 2, 3
- [89] Dingdong Yang, Seunghoon Hong, Yunseok Jang, Tianchen Zhao, and Honglak Lee. Diversity-sensitive conditional generative adversarial networks. In *ICLR*, 2019. 6
- [90] Jinsung Yoon, James Jordon, and Mihaela Van Der Schaar. GAIN: Missing data imputation using generative adversarial nets. In *ICML*, 2018. 3
- [91] 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. *arXiv preprint arXiv:1506.03365*, 2015. 2
- [92] Han Zhang, Ian Goodfellow, Dimitris Metaxas, and Augustus Odena. Self-attention generative adversarial networks. In *ICML*, 2019. 3
- [93] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Trans. Image Process.*, 26(7):3142–3155, 2017. 2, 3
- [94] Kai Zhang, Wangmeng Zuo, and Lei Zhang. FFDNet: Toward a fast and flexible solution for CNN-based image denoising. *IEEE Trans. Image Process.*, 27(9):4608–4622, 2018. 2, 3
- [95] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *CVPR*, 2018. 8
- [96] Shengyu Zhao, Zhijian Liu, Ji Lin, Jun-Yan Zhu, and Song Han. Differentiable augmentation for data-efficient GAN training. In *NeurIPS*, 2020. 3
- [97] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *ICCV*, 2017. 2, 8