
Blind-Spot Guided Diffusion for Self-supervised Real-World Denoising

Shen Cheng
Dexmal

Haipeng Li
UESTC

Haibin Huang
Tele AI

Xiaohong Liu
SJTU

Shuaicheng Liu*
UESTC

Abstract

In this work, we present **Blind-Spot Guided Diffusion**, a novel self-supervised framework for real-world image denoising. Our approach addresses two major challenges: the limitations of blind-spot networks (BSNs), which often sacrifice local detail and introduce pixel discontinuities due to spatial independence assumptions, and the difficulty of adapting diffusion models to self-supervised denoising. We propose a dual-branch diffusion framework that combines a BSN-based diffusion branch, generating semi-clean images, with a conventional diffusion branch that captures underlying noise distributions. To enable effective training without paired data, we use the BSN-based branch to guide the sampling process, capturing noise structure while preserving local details. Extensive experiments on the SIDD and DND datasets demonstrate state-of-the-art performance, establishing our method as a highly effective self-supervised solution for real-world denoising. Code and pre-trained models will be released.

1 Introduction

As a long-standing and fundamental task, image denoising plays a crucial role in enhancing image quality for a wide range of downstream applications [64, 29]. The goal of image denoising is to recover a clean image x from a noisy observation y , defined as

$$y = x + n \quad (1)$$

where n represents the additive noise. This is an ill-posed problem as both the image x and noise n components are unknown and difficult to disentangle. Thus, effective denoising methods require robust image priors and noise models to estimate either the image or the noise from noisy observations.

Methods for image denoising fall into multiple categories, depending on the availability of paired noisy-clean images. In the presence of paired data, denoising can be formulated as a regression task, where deep learning models are trained to minimize the difference between their predictions and the clean ground-truth images. Recent deep networks based methods have demonstrated great success for this category [44, 6, 63, 18, 49, 31, 45]. However, acquiring such data at scale is resource-intensive and often impractical.

To address these challenges, self-supervised learning methods have emerged as promising alternatives by removing the dependency on paired data. Methods like Noise2Noise [25] enable models to learn from unpaired noisy images by leveraging the principle that multiple noisy observations of the same underlying clean image can be used to approximate clean data. Subsequent approaches, like Noise2Void [22] and Noise2Self [2], advance this paradigm by leveraging the intrinsic structure of images, allowing models to learn solely from noise using *only a collection of noisy images captured in a single instance*.

*Corresponding author

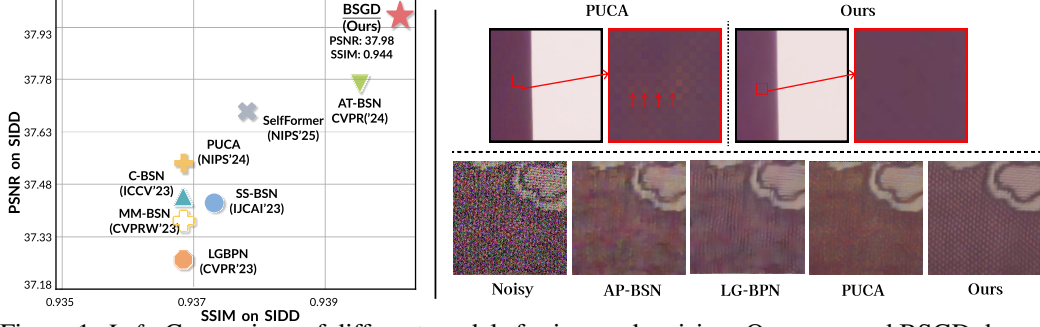


Figure 1: *Left*: Comparison of different models for image denoising. Our proposed BSGD demonstrates a significant improvement over previous methods. *Right Top*: Grid Pattern Effects. *Right Bottom*: SIDD Validation 0035-0029. Our results presented were denoised using a guidance of $w = 0.1$. All images are brightened for better visualization.

Recent advances in blind-spot networks (BSNs) [23] have extended self-supervised denoising capabilities by predicting pixel values from neighboring pixels while leaving the central pixel unseen, enabling learning from intrinsic image patterns. However, BSNs rely on an assumption of spatial noise independence, which often results in local detail loss and pixel discontinuities. Subsequent methods, such as AP-BSN [24], employ pixel shuffling to reduce spatial noise dependency, while networks like LG-BPN [47], MM-BSN [57], and PUCA [19] attempt to expand the receptive field through deeper, multi-scale features. Others, like SS-BSN and SelfFormer [55], use grid-based attention to improve contextual relationships. Nevertheless, due to pixel shuffling and blind-spot regions, these networks continue to experience losses in local details, causing pixel discontinuities, as illustrated in Fig. 1 right.

To preserve neighboring pixel continuity, this work employs a non-blind-spot diffusion model, avoiding blind-spot masking or pixel-shuffling operations. The design is driven by two key considerations: First, diffusion models demonstrate unique advantages in distribution modeling by implicitly learning hybrid noisy distributions that jointly characterize noise patterns and latent clean image structures without explicit denoising objectives. Second, building on the theoretical framework of Classifier-Free Guidance (CFG), we propose a semi-clean guidance mechanism that dynamically adjusts score estimates during sampling, enabling directional control over generation outcomes under noisy training distributions. The synergistic integration of these properties allows diffusion models to maintain pixel coherence while achieving controllable synthesis through sampling-phase guidance.

In this paper, we propose Blind-Spot Guided Diffusion (BSGD), a novel framework that leverages the generative strength of diffusion models to overcome the limitations of BSNs. Unlike prior approaches focused on architectural modifications, we explore a new property of diffusion models for the task of denoising: the ability to combine information from both blind-spot and non-blind-spot branches during the sampling process, effectively improving denoising results. Our method employs a dual-branch diffusion framework comprising a BSN-based diffusion branch, which generates **semi-clean** images, and a conventional diffusion branch, which learns the **noisy image distribution**. Since training a conventional diffusion model without paired data is challenging, we use the BSN diffusion as guidance during sampling to capture the noise structure effectively. The two branches are jointly optimized to produce robust noise estimates for real-world denoising, leading to a significant improvement in performance, as illustrated in Fig. 1 left. Our experiments demonstrate the effectiveness of our approach, achieving state-of-the-art results on the SIDD and DND datasets.

Contributions. Our contributions can be summarized as follows:

- We introduce the first dual-branch diffusion framework for self-supervised denoising, tailored for real-world applications.
- We combine BSN-based and conventional diffusion with blind-spot guidance and Random Replacement Sampling to overcome BSN limitations and enhance detail preservation.
- Our method achieves favorable performance against state-of-the-art self-supervised denoising approaches.

2 Related works

Supervised Methods Supervised methods for image denoising rely on paired data. Some approaches focus on reducing Gaussian noise through specialized network architectures. A pioneering study [3] utilized MLPs for image denoising, achieving results comparable to BM3D [10]. DnCNN [58] demonstrated the effectiveness of CNNs in denoising Gaussian noise, inspiring numerous studies [59, 31, 37, 39, 60] that proposed various architectures.

Subsequent works [56, 8] addressed real-world challenges by training on realistic paired noisy and clean data. However, obtaining such data is often labor-intensive. To address this, CBNet [41] proposed estimating image noise levels and explored the generalization capabilities of Gaussian denoisers. Additionally, Zhou *et al.* [63] revealed that real-world noise exhibits strong spatial correlation, and pixel-shuffle downsampling (PD) operators can disrupt this correlation, enabling networks trained solely on Gaussian noise to perform well on real-world noise.

Self-supervised Methods Lehtinen *et al.* [25] introduced Noise2Noise (N2N), which trains models using multiple independent noisy images of the same scenes. Subsequently, Noise2Void (N2V) [22] simplified this approach by employing a blind-spot scheme, allowing denoising from single images. Since then, more variants of the blind-spot approach have emerged, including Noise2Self [2], Noise-As-Clean [52], Neighbor2Neighbor [17], Noise2Same [50], and SASL [27].

Instead of masking input images, some methods focus on architectures known as Blind-Spot Networks (BSN). Laine *et al.* [23] proposed masked convolution, and Wu *et al.* [48] introduced stacked dilated convolution layers to integrate the blind-spot scheme within the model, improving convergence speed and performance.

However, BSN-based methods struggle with real-world noise due to its spatial correlation. To address this, AP-BSN [24] used pixel-shuffle downsampling (PD) to disrupt this correlation and incorporated BSN for self-supervised learning. Further developments focused on designing better BSNs, such as MM-BSN [57], which uses more center-masked convolutions, LGBPN [47], which employs dilated convolutions to expand the receptive field, SS-BSN [13], which utilizes grid-based attention, and PUCA [19], which designs a patch-shuffle downsampling method. However, these structures are based on grid receptive fields, leading to grid pattern artifacts, as shown in Fig. 1 (right-top part). In this work, we integrate BSN into diffusion models to effectively resolve these artifacts.

Diffusion Models Diffusion models [14] have become one of the most popular generative models due to their exceptional ability to model data distributions. They have demonstrated state-of-the-art performance across a wide range of tasks beyond image synthesis [38, 61]. In high-level applications, they excel in object detection [5], object tracking [30], 3D hand mesh synthesis [51], and pose estimation [16]. Additionally, for low-level tasks, diffusion models have shown impressive results in image alignment [26, 40], rectangular image stitching [62], depth estimation [21], dense matching [32], and image restoration [46, 12, 20]. Recently, Li *et al.* [28] employed Stable Diffusion [38] for zero-shot image denoising, while Yang *et al.* [53] achieved image denoising by generating noisy images. In this work, we propose the first diffusion framework for self-supervised denoising tasks.

3 Method

The proposed framework is depicted in Fig. 2. Our goal is to estimate the clean distribution $\mathbf{x}_0^c \sim q_0(\mathbf{x}^c)$ by modeling the real-world noisy data distribution $\mathbf{x}_0 \sim p_0(\mathbf{x})$ and generating clean samples through the reverse diffusion process. Towards this end, we introduce the Blind-Spot Guided Diffusion (BSGD) model, which incorporates blind-spot networks. To mitigate the detail loss often seen in BSNs, we also train a standard, non-blind-spot diffusion model. During sampling, BSGD guides this non-blind-spot model to improve sample efficacy. Additionally, we implement a Complementary Replacement Sampling module to further enhance denoising performance.

Diffusion models Diffusion models (DMs) operate by learning a sequence of transformations that progressively refine a simple noise distribution into the complex distribution of real-world images—here focusing on noisy images, represented as $\mathbf{x}_0 \sim p(\mathbf{x}_0)$. The process initiates with a forward diffusion step, where noise is systematically introduced into the data while sampling latent variables $\mathbf{x}_1, \dots, \mathbf{x}_T$ of identical dimensionality. This step is governed by a Gaussian function that is

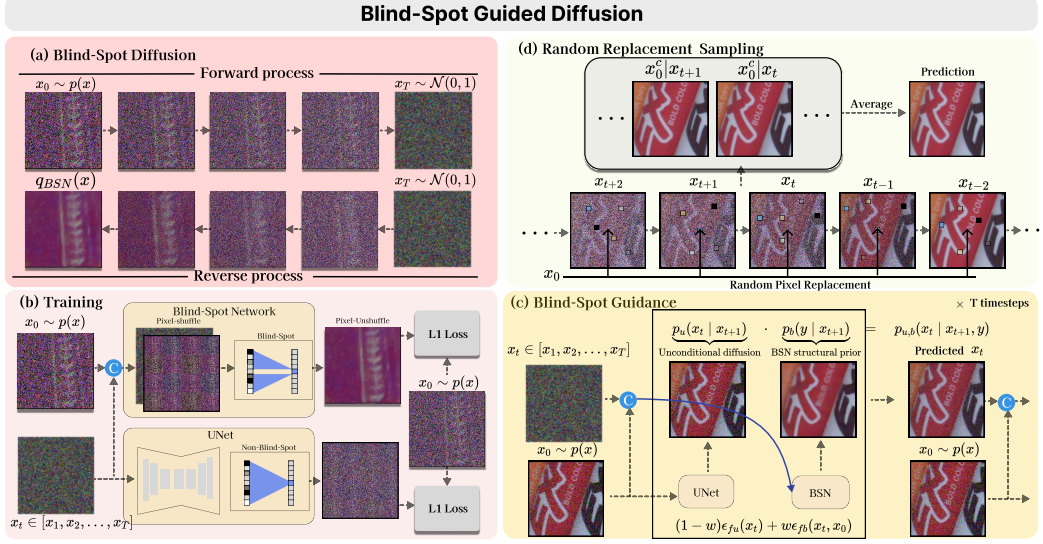


Figure 2: The Framework of Guided Blind-Spot Diffusion. (a) The training process incorporates real-world noise, while the reverse process generates clean images. (b) The framework consists of two diffusion models: one for estimating the noise-free image and the other for capturing local structures. (c) During the sampling phase, the Blind-Spot Network (BSN) serves as a guide for the UNet. (d) We implement a replacement strategy during sampling, averaging all estimated $[\mathbf{x}_0^c | \mathbf{x}_t]_{t=0}^T$.

contingent on the current state of the data and a predetermined noise schedule. The forward diffusion can be mathematically expressed as:

$$q(\mathbf{x}_t | \mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1 - \beta_t} \mathbf{x}_{t-1}, \beta_t \mathbf{I}), \quad (2)$$

where β_t adheres to a fixed schedule and $\mathbf{I}$ denotes the identity matrix.

The sampling component of DMs involves reversing the noise accumulation process. The model learns parameters to revert the noisy data back to the image, which is also modeled using a Gaussian distribution characterized by a learned mean and a fixed variance. Following established research in diffusion models [43, 14], we adopt the generation framework:

$$p_\theta(\mathbf{x}_{t-1} | \mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \mu_\theta(\mathbf{x}_t, t), \sigma_t^2 \mathbf{I}), \quad (3)$$

where μ_θ represents the learned mean and σ_t^2 is the fixed variance at time step t .

3.1 Blind-Spot Diffusion Guidance

Inspired by prior work [27, 69], we model the noisy observation as a perturbed version of the clean image, i.e., $p_0(x) \sim \mathcal{N}(x^c, \Sigma)$, where Σ is the covariance matrix. This assumption implies that the noisy data distribution lies in a neighborhood around the clean image manifold. Diffusion models, trained on noisy images only, inherently learn to approximate the clean image distribution through iterative denoising (see supplementary). However, unconditioned diffusion sampling often produces *uncontrolled* outputs, as the model lacks explicit guidance to prioritize plausible clean configurations. To address this, we draw inspiration from *classifier-free guidance (CFG)* [15], a standard process that trains both conditional and unconditional diffusion models and **mixes their output noise** during sampling. To achieve this, we implement two diffusion models: a conditional network utilizing a BSN architecture and an unconditional network based on U-Net. We leverage BSNs to provide a **coarse estimate** of x^c for CFG conditioning. Despite the limitations of BSNs, they encode structural coherence. This estimate acts as a “soft prior” to steer the diffusion process toward plausible clean configurations. BSGD integrates this prior by reparameterizing the CFG:

$$p_{u,b}(\mathbf{x}_t | \mathbf{x}_{t+1}, y) = Z \cdot \underbrace{p_u(\mathbf{x}_t | \mathbf{x}_{t+1})}_{\text{Unconditional diffusion}} \cdot \underbrace{p_b(y | \mathbf{x}_t)}_{\text{BSN structural prior}}, \quad (4)$$

where Z ensures normalization. The product of these terms prioritizes regions where the diffusion model’s noise distribution aligns with the BSN’s structural prior, effectively filtering out implausible artifacts while preserving global coherence.

Blind-Spot Diffusion To facilitate training within the diffusion framework, the network was adapted to be time-dependent, denoted as $f_b(\mathbf{x}_t, \mathbf{x}_0, t)$:

$$\min_{\theta} \sum_t \|f_b(\mathbf{x}_t, \mathbf{x}_0, t) - \mathbf{x}_0\|_1. \quad (5)$$

Due to the nature of the BSNs, once the model is trained, each estimation produced by the BSN corresponds to a clean image $\mathbf{x}_{0|t}^c = \mathbf{x}_0^c | \mathbf{x}_t$. Therefore, for each sampling instance, once the first sampling $\mathbf{x}_0^c | \mathbf{x}_T$ is done, the generation framework of the blind-spot diffusion can be modified as:

$$p_{\theta}(\mathbf{x}_{t-1} | \mathbf{x}_t, \mathbf{x}_{0|t}^c) = \mathcal{N}(\mathbf{x}_{t-1}; \mu_{\theta}(\mathbf{x}_t, \mathbf{x}_{0|t}^c, t), \sigma_t^2 \mathbf{I}), \quad (6)$$

Non-Blind-Spot Diffusion. The Non-blind diffusion refers to the standard diffusion framework introduced by [14], which exploits UNet architecture f_u , hence the training objective is:

$$\min_{\theta} \sum_t \|f_u(\mathbf{x}_t, t) - \mathbf{x}_0\|_1. \quad (7)$$

The UNet diffusion model is utilized to effectively capture the local structure of the noise distribution.

Blind-Spot Guidance. The two training objectives correspond to denoising score matching [43] and can be transformed between ϵ -prediction and $\mathbf{x}_0$ -prediction using the following equation:

$$\epsilon(\mathbf{x}_t) = \frac{\mathbf{x}_t - \sqrt{\alpha_t} \cdot f_{\theta}(\mathbf{x}_t, t)}{\sqrt{1 - \alpha_t}}, \theta \in [u, b]. \quad (8)$$

The ϵ -predictions for non-blind and blind-spot diffusion are denoted as ϵ_{fu} and ϵ_{fb} , respectively. Additionally, the diffusion models can be trained jointly by randomly omitting $\mathbf{x}_0$ from the input.

Once training is complete, image samples can be generated by simulating the reverse process:

$$\begin{aligned} \mathbf{x}_{t-1} = & \underbrace{\sqrt{1 - \alpha_{t-1} - \sigma_t^2} \cdot \epsilon(\mathbf{x}_t)}_{\text{direction to } \mathbf{x}_t} \\ & + \sqrt{\alpha_{t-1}} \underbrace{\frac{\mathbf{x}_t - \sqrt{1 - \alpha_t} \cdot \epsilon(\mathbf{x}_t)}{\sqrt{\alpha_t}}}_{\text{estimated } \mathbf{x}_0} + \sigma_t \mathbf{z}, \end{aligned} \quad (9)$$

For each sampling step, the blind-spot network guides the non-blind U-Net for conditional generation. The original diffusion score is modified to incorporate both networks. In our work, we employed a simple linear weighting, constraining the weights within the range of $0 \sim 1$ to ensure stability:

$$\epsilon(\mathbf{x}_t) = w \epsilon_{fb}(\mathbf{x}_t, \mathbf{x}_0) + (1 - w) \epsilon_{fu}(\mathbf{x}_t). \quad (10)$$

Next, we describe the whole sampling process and introduce the proposed Complementary Replacement Sampling.

3.2 Complementary Replacement Sampling

For simplicity, we denote the latent images sampled as $[\mathbf{x}_t]_{t=0}^T$ in Eq. 9. It is important to note that a predicted clean image corresponds to each sampling step, which we represent as $[\mathbf{x}_0^c | \mathbf{x}_t]_{t=0}^T$. These estimations serve as distinct observations of $\mathbf{x}_0^c$; therefore, we compute their average to obtain the final output, as illustrated in Fig. 2(d).

In each sampling step, we also implement a random replacement strategy. Specifically, pixels in the predicted images $[\mathbf{x}_0^c | \mathbf{x}_t]_{t=0}^T$ are selected with a probability p and replaced with pixels from the

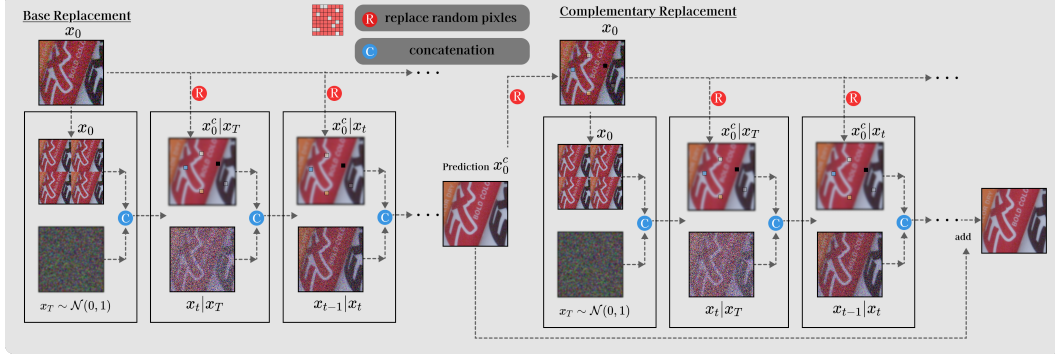


Figure 3: The Replacement Strategy. The Base Replacement is employed during the initial sampling process, while the Complementary Replacement is utilized for subsequent processes.

input noisy image x_0 . The resulting images, after replacement, are then utilized as conditions for the subsequent sampling steps. We employ two types of replacements during the sampling process: *Base Replacement* and *Complementary Replacement*.

Base Replacement. The replacement process is illustrated in Fig. 3. The pixel-shuffling technique is employed exclusively during the first sampling step and in the context of blind-spot diffusion. In this framework, the pixel-shuffled noisy images x_0 serve as the conditioning input, leading to the first estimation $x_0^c|x_T$. Subsequently, this estimation undergoes the replacement process.

It is noteworthy that the sampling procedure can be conceptualized as an ordinary differential equation (ODE) [43], allowing for sampling with larger steps. As established in [42], we can derive $x_t|x_T$ from $x_0^c|x_T$. For the subsequent steps, the modified $x_0^c|x_T$, post-replacement, is concatenated with the t -step sample $x_t|x_T$ to form the input $[x_0^c|x_T, x_t|x_T]$ for the diffusion process. Consequently, noisy images are utilized only in the initial step, while portions of the noisy pixels are retained for the remaining steps.

Complementary Replacement. We define a complete sampling process as a sampling round. Upon completion of the initial sampling round, we conduct several additional sampling rounds utilizing the complementary replacement strategy. In the base replacement approach, pixels in the predicted clean image are substituted with corresponding pixels from the input noisy image. This method directly integrates noise information into the predicted output.

Conversely, the complementary replacement technique involves first substituting the input noisy image x_0 with the previously predicted clean image. This modified image then serves as the conditioning input for subsequent sampling rounds. By leveraging the previously estimated clean image, the complementary replacement aims to refine the sampling process further, potentially enhancing the quality of the final output. This dual approach, Base and Complementary Replacement, facilitates a more robust exploration of the latent space, allowing for improved convergence towards a cleaner image representation.

4 Experiments

4.1 Training Settings

Dataset Preparation For benchmarking purposes, we evaluate the proposed method through experiments conducted on the **SIDD** and **DND** datasets. The correlation between adjacent noise pixels in RAW images is weaker compared to sRGB images, as sRGB contains numerous non-linear transformations introduced by the ISP. The Gaussian noise and **SIDD-Raw** datasets are not considered in our experiments.

The SIDD (Smartphone Image Denoising Dataset) comprises 30,000 noisy images captured across 10 distinct scenes under varying lighting conditions, utilizing five representative smartphone cameras. For benchmarking, SIDD extracts 1,280 color patches from 40 noisy images, which are utilized for validation and performance assessment. The benchmark results are reported on an online server,

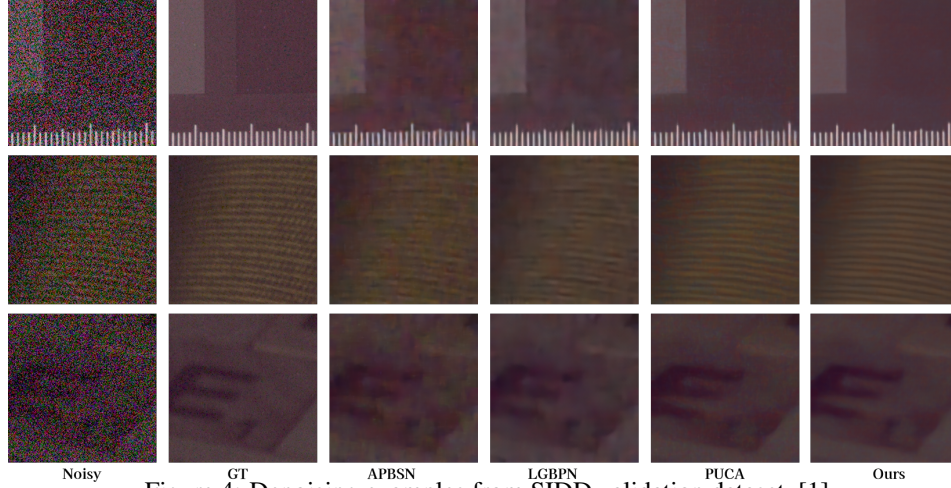


Figure 4: Denoising examples from SIDD validation dataset. [1]

Table 1: Quantitative comparisons on benchmark datasets are presented, with bold values indicating the best performance among self-supervised methods. A dash (–) denotes instances where data is not reported. (The SIDD benchmark server (updated Oct 2024) adopted a different SSIM calculation method. However, SOTA methods in the table lack open-source code, preventing SSIM comparisons under the new framework. We thus retain original metrics while providing full experimental results to support future benchmarking of subsequent methods.)

	Method	SIDD Benchmark		DND Benchmark	
		PSNR $\uparrow$	SSIM $\uparrow$	PSNR $\uparrow$	SSIM $\uparrow$
Supervised Real pairs	DnCNN [58]	37.64	0.937	32.43	0.790
	DANet [54]	39.47	0.918	39.59	0.955
	MIRNet [56]	39.72	0.959	39.88	0.956
	NBNet [8]	39.74	0.959	39.89	0.955
	NAFNet [?]]	40.30	0.961	-	-
Zero-Shot	NAC [52]	-	-	36.20	0.925
	ScoreDVI [7]	34.60	0.920	-	-
	self2self [36]	29.51	0.651	-	-
	MASH [9]	34.78	0.900	-	-
Self-supervised	N2V [22]	27.68	0.668	-	-
	Noise2Self [2]	29.56	0.808	-	-
	Laine <i>et al.</i> [17]	30.14	0.823	35.13	0.862-
	R2R [34]	34.78	0.898	-	-
	CVF-SID [33]	34.71	0.917	36.50	0.924
	AP-BSN [24]	36.91	0.931	38.09	0.937
	LG-BPN [47]	37.28	0.936	38.43	0.942
	MM-BSN [57]	37.37	0.936	38.46	0.940
	SASL [27]	37.41	0.934	38.18	0.938
	SS-BSN [13]	37.42	0.937	38.46	0.940
	C-BSN [11]	37.43	0.936	38.62	0.942
	Jang <i>et al.</i> [17]	37.53	0.936	38.56	0.9402
	PUCA [19]	37.54	0.936	38.83	0.942
	AT-BSN [4]	37.77	0.942	38.29	0.939
	SelfFormer [55]	37.69	0.937	38.92	0.943
	Ours	37.98	0.944	38.99	0.943

where ground-truth data is not accessible to researchers. The validation dataset, containing noisy-clean image pairs, is employed for our ablation studies. We utilize the SIDD-medium dataset for training, applying random cropping to generate image patches of size 128×128 .

The Darmstadt Noise Dataset (DND) [35] consists of 50 pairs of real noisy images and their corresponding ground-truth images, captured using consumer-grade cameras with varying sensor sizes. Each pair includes a source image taken at the base ISO level, while the noisy image is captured at a higher ISO with an appropriately adjusted exposure time. The reference image undergoes meticulous post-processing, which includes small camera shift adjustments, linear intensity scaling, and the removal of low-frequency bias. This post-processed image serves as the ground truth for the DND denoising benchmark. It is important to note that the DND benchmark results are also obtained by submitting predictions to the server for evaluation.

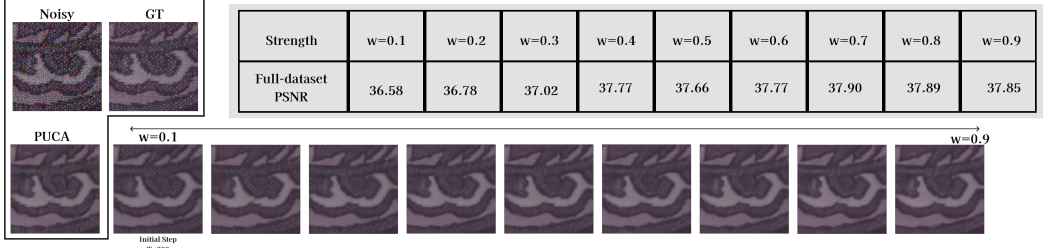


Figure 5: Ablation Study of Guidance Strength on SIDD Validation dataset. When the guidance strength w is small, finer details are preserved; however, a larger w results in blurriness. To achieve a balance, we select a higher strength for benchmarking purposes.

Table 2: Ablation Study of Key Modules on SIDD Validation dataset. PUCA serves as the baseline in this analysis.

Component	PUCA	PUCA ($R^3, T = 8$)	PUCA ($R^3, T = 64$)	Ours(w/o guidance)	Ours(w/o Replacement)	Ours
PSNR&SSIM	35.38/0.875	37.53/0.880	37.54/0.879	37.81/0.884	36.58/0.877	37.90/0.887

Implementation Details The proposed framework incorporates two distinct types of networks: a blind-spot network and a non-blind-spot UNet. The BSN architecture is derived from the PUCA model [19], which we have modified to create a time-dependent network utilizing a diffusion model. This modification introduces two additional inputs: the time step and the latent sampled image $\mathbf{x}_t$. Consequently, the PUCA model serves as our baseline for comparative analysis and ablation studies. In contrast, the UNet architecture, widely employed in diffusion models [14], features a non-blind perceptive field. We train both diffusion models under the following experimental conditions: a batch size of 8, a learning rate of 8×10^{-5} , and the Adam optimizer. The Blind-Spot diffusion model is trained for 200,000 iterations, while the UNet diffusion model undergoes training for 500,000 iterations. The diffusion steps for training are 1,000. During the sampling phase, we implement 8 steps for each sampling process, followed by 8 additional sampling processes. For the best results of benchmarking, we start sampling with the initial step $T = 300$.

4.2 Results on SIDD and DnD

Table 1 presents two real-world noise benchmarks, SIDD and DnD. The table compares the results of our method with those of other approaches. Self-supervised denoising methods have recently attracted significant research interest; however, their advancements over previous techniques have been relatively modest. This may be attributed to the limitations associated with the BSN and the pixel shuffle operator. In contrast, our method demonstrates a substantial enhancement compared to the SOTA methods, achieving a PSNR metric that approaches 38 dB on the SIDD dataset. Additionally, we have attained SOTA performance on the DnD benchmark.

The visual results are illustrated in Fig. 4, where we compare the outcomes of APBSN, LGBPN, and PUCA with our results. Notably, both the APBSN and LGBPN methods exhibit significant detail loss, whereas our approach retains finer details compared to PUCA, while also achieving a lower noise level. PUCA serves as a crucial baseline for our study, as our Blind-Spot Network (BSN) is built upon its architecture with some modifications to create a time-dependent model. Our approach demonstrates an improvement of 0.43 dB over PUCA.

4.3 Ablation Study

Components As discussed in Section 4.2, PUCA serves as a baseline, representing a variant of our method that does not employ the diffusion model during training. Consequently, the ablation studies incorporate the replicated PUCA method trained from scratch, with performance levels comparable to those reported in [19]. The results are presented in Table 2. PUCA [19] also employs the replacement strategy introduced in APBSN [24]. The optimal performance of PUCA was achieved with 8 iterations of replacement, while further iterations yielded negligible improvements. The data in the fourth column demonstrate that utilizing guidance effectively integrates non-blind-spot information, leading to performance enhancement. Additionally, the data in the fifth column indicate that the proposed random replacement sampling also contributes to a notable improvement in results.

Sampling steps and Sampling Rounds Diffusion-based methods are fundamentally reliant on the sampling process. Given that our approach requires multiple sampling steps and complete sampling rounds, we conducted separate analytical experiments on these two critical variables.

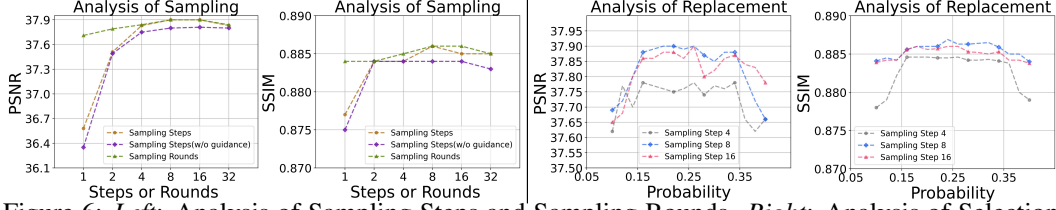


Figure 6: *Left: Analysis of Sampling Steps and Sampling Rounds. Right: Analysis of Selection Probability*

Initially, we investigated the impact of the sampling steps, both with and without guidance. The results are summarized in Fig. 6. As indicated in the figure, when the number of sampling steps is set to one, the performance is notably poor. This limitation arises because, at this stage, the replacement strategy cannot be effectively implemented; even with guidance, a single iteration fails to achieve satisfactory information fusion.

Performance improves significantly when steps increase to 2. Optimal results occur at 8 - 16 steps, but performance declines at 32 steps due to diffusion uncertainties. An ablation study on sampling iterations was done, with N set to 1 - 32. Results show the impact trend of iterations on performance resembles that of steps, with best results at 8 - 16 iterations and a decline at 32. The model performs well even with one sampling round as it completes the sampling process, effectively using guidance and replacement information. The results indicate that the trend in the impact of the number of iterations on performance mirrors that of the sampling steps, with optimal results observed at 8 and 16 iterations, while performance declines at 32 iterations. Notably, unlike the sampling steps, the model demonstrates competitive performance even with just one sampling round. This is attributed to the fact that, despite having only one sampling round, the model still undergoes a complete sampling process, allowing for effective utilization of both guidance and replacement information.

Guidance Strength The ablation study on guidance strength is presented in Fig. 5. We first conducted experiments on the overall performance across the entire dataset. The PSNR tends to decline when the guidance strength is close to 0 or 1, with a more pronounced decrease observed near 0. This is attributed to the predominant contribution of the unconditional non-blind-spot diffusion at that point. As the weight w increases, the PSNR gradually improves, achieving optimal results at $w = 0.7$ and $w = 0.8$, reaching a PSNR of 37.90 on the SIDD validation dataset.

Additionally, we performed visualization experiments. Notably, since sampling can commence from any step during the process, we chose to start sampling at the 300th step (out of a total of 1000 steps) to preserve image information. For example, the sampling trajectory is $[300 \rightarrow 263 \rightarrow 226 \rightarrow 189 \rightarrow 152 \rightarrow 115 \rightarrow 78 \rightarrow 41 \rightarrow 0]$. This choice ensures that even with weak guidance, random samples are not generated. It is evident that when using weak guidance, the model demonstrates strong detail recovery and denoising capabilities; however, the overall PSNR performance across the dataset is not optimal. This is due to the diffusion process introducing generative capabilities, which diminishes the reference value of the PSNR. Nevertheless, for the SIDD dataset, we selected the parameters that yielded the highest PSNR based on a compromise from our experiments.

Replacement probability We also conducted an experimental analysis on the probability of the replacement operation, with the results illustrated in Figure 6. Within the range of 0.15 to 0.35, the model’s performance shows minimal sensitivity to the probability of selecting pixels, as both PSNR and SSIM fluctuate within a narrow band. However, performance significantly declines when the probability is below 0.15 or above 0.35. Based on this experimental analysis, we recommend selecting a probability value around 0.25 for pixel selection.

5 Conclusion

In this paper, we revisit the challenge of self-supervised denoising for real-world applications and introduce a first diffusion-based framework tailored for this task. Our experiments demonstrate that non-blind-spot information can be effectively integrated through diffusion sampling, overcoming limitations of blind-spot networks and enhancing denoising performance. We hope this approach would inspire future advancements in self-supervised denoising.

References

- [1] Abdelrahman Abdelhamed, Stephen Lin, and Michael S Brown. A high-quality denoising dataset for smartphone cameras. In *Proc. CVPR*, pages 1692–1700, 2018.
- [2] Joshua Batson et al. Noise2self: Blind denoising by self-supervision. In *ICML*, pages 524–533. PMLR, 2019.
- [3] Harold C Burger, Christian J Schuler, and Stefan Harmeling. Image denoising: Can plain neural networks compete with bm3d? In *Proc. CVPR*, pages 1256–1272, 2012.
- [4] Shiyen Chen, Jiyan Zhang, Zhaoqi Yu, and Tiejun Huang. Exploring efficient asymmetric blind-spots for self-supervised denoising in real-world scenarios. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2814–2823, 2024.
- [5] Shoufa Chen, Peize Sun, Yibing Song, and Ping Luo. Diffusion-det: Diffusion model for object detection. In *Proc. CVPR*, pages 19830–19843, 2023.
- [6] Yunjin Chen and Thomas Pock. Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 39(6):1256–1272, 2017.
- [7] Jun Cheng, Tao Liu, and Shan Tan. Score priors guided deep variational inference for unsupervised real-world single image denoising. In *Proc. CVPR*, pages 12937–12948, 2023.
- [8] Shen Cheng, Yuzhi Wang, Haibin Huang, Donghao Liu, Haoqiang Fan, and Shuaicheng Liu. Nbnnet: Noise basis learning for image denoising with subspace projection. In *Proc. CVPR*, pages 4896–4906, 2021.
- [9] Hamadi Chihoui and Paolo Favaro. Masked and shuffled blind spot denoising for real-world images. In *Proc. CVPR*, pages 3025–3034, 2024.
- [10] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Trans. on Image Processing*, 16(8):2080–2095, 2007.
- [11] Linwei Fan, Jin Cui, Huiyu Li, Xiaoyu Yan, Hui Liu, and Caiming Zhang. Complementary blind-spot network for self-supervised real image denoising. *IEEE Trans. on Circuits and Systems for Video Technology*, 2024.
- [12] Sicheng Gao, Xuhui Liu, Bohan Zeng, Sheng Xu, Yanjing Li, Xiaoyan Luo, Jianzhuang Liu, Xiantong Zhen, and Baochang Zhang. Implicit diffusion models for continuous super-resolution. In *Proc. CVPR*, pages 10021–10030, 2023.
- [13] Young-Joo Han and Ha-Jin Yu. Ss-bsn: attentive blind-spot network for self-supervised denoising with nonlocal self-similarity. *arXiv preprint arXiv:2305.09890*, 2023.
- [14] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Proc. NeurIPS*, pages 6840–6851, 2020.
- [15] Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. *arXiv preprint arXiv:2207.12598*, 2022.
- [16] Karl Holmquist and Bastian Wandt. Diffpose: Multi-hypothesis human pose estimation using diffusion models. In *Proc. ICCV*, pages 15977–15987, 2023.
- [17] Tao Huang, Songjiang Li, Xu Jia, Huchuan Lu, and Jianzhuang Liu. Neighbor2neighbor: Self-supervised denoising from single noisy images. In *Proc. CVPR*, pages 14781–14790, 2021.
- [18] Viren Jain and Sebastian Seung. Natural image denoising with convolutional networks. In *Proc. NeurIPS*, pages 769–776, 2009.
- [19] Hyemi Jang, Junsung Park, Dahyun Jung, Jaihyun Lew, Ho Bae, and Sungroh Yoon. Puca: patch-unshuffle and channel attention for enhanced self-supervised image denoising. In *Proc. NeurIPS*, 2024.
- [20] Hai Jiang, Ao Luo, Haoqiang Fan, Songchen Han, and Shuaicheng Liu. Low-light image enhancement with wavelet-based diffusion models. *ACM Trans. on Graphics*, 42(6):1–14, 2023.
- [21] Bingxin Ke, Anton Obukhov, Shengyu Huang, Nando Metzger, Rodrigo Caye Daudt, and Konrad Schindler. Repurposing diffusion-based image generators for monocular depth estimation. In *Proc. CVPR*, 2024.

- [22] Alexander Krull, Tim-Oliver Buchholz, and Florian Jug. Noise2void-learning denoising from single noisy images. In *Proc. CVPR*, pages 2129–2137, 2019.
- [23] Samuli Laine et al. High-quality self-supervised deep image denoising. In *Proc. NeurIPS*, pages 6970–6980, 2019.
- [24] Wooseok Lee, Sanghyun Son, and Kyoung Mu Lee. Ap-bsn: Self-supervised denoising for real-world images via asymmetric pd and blind-spot network. In *Proc. CVPR*, pages 17725–17734, 2022.
- [25] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2noise: Learning image restoration without clean data. *arXiv preprint arXiv:1803.04189*, 2018.
- [26] Haipeng Li, Hai Jiang, Ao Luo, Ping Tan, Haoqiang Fan, Bing Zeng, and Shuaicheng Liu. Dmhomom: Learning homography with diffusion models. *ACM Trans. Graphics*, 2024.
- [27] Junyi Li, Zhilu Zhang, Xiaoyu Liu, Chaoyu Feng, Xiaotao Wang, Lei Lei, and Wangmeng Zuo. Spatially adaptive self-supervised learning for real-world image denoising. In *Proc. CVPR*, pages 9914–9924, 2023.
- [28] Tong Li, Hansen Feng, Lizhi Wang, Lin Zhu, Zhiwei Xiong, and Hua Huang. Stimulating diffusion model for image denoising via adaptive embedding and ensembling. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2024.
- [29] Hongyu Liu, Bin Jiang, Yi Xiao, and Chao Yang. Coherent semantic attention for image inpainting. In *ICCV*, pages 4170–4179, 2019.
- [30] Run Luo, Zikai Song, Lintao Ma, Jinlin Wei, Wei Yang, and Min Yang. Diffusiontrack: Diffusion model for multi-object tracking. In *Proc. AAAI*, volume 38, pages 3991–3999, 2024.
- [31] Xiaojiao Mao, Chunhua Shen, and Yu-Bin Yang. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. In *Proc. NeurIPS*, pages 2810–2818, 2016.
- [32] Jisu Nam, Gyuseong Lee, Sunwoo Kim, Hyeonsu Kim, Hyoungwon Cho, Seyeon Kim, and Seungryong Kim. Diffusion model for dense matching. In *The Twelfth International Conference on Learning Representations*, 2023.
- [33] Reyhaneh Neshatavar, Mohsen Yavartanoo, Sanghyun Son, and Kyoung Mu Lee. Cvf-sid: Cyclic multivariate function for self-supervised image denoising by disentangling noise from image. In *Proc. CVPR*, pages 17583–17591, 2022.
- [34] Tongyao Pang, Huan Zheng, Yuhui Quan, and Hui Ji. Recorrupted-to-recorrupted: Unsupervised deep learning for image denoising. In *Proc. CVPR*, pages 2043–2052, 2021.
- [35] Tobias Plötz and Stefan Roth. Benchmarking denoising algorithms with real photographs. In *Proc. CVPR*, pages 2750–2759, 2017.
- [36] Yuhui Quan, Mingqin Chen, Tongyao Pang, and Hui Ji. Self2self with dropout: Learning self-supervised denoising from single image. In *Proc. CVPR*, pages 1890–1898, 2020.
- [37] Haoyu Ren, Mostafa El-Khamy, and Jungwon Lee. Dn-resnet: Efficient deep residual network for image denoising. *arXiv preprint arXiv:1810.06766*, 2018.
- [38] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proc. CVPR*, pages 10684–10695, 2022.
- [39] Venkataraman Santhanam, Vlad I Morariu, and Larry S Davis. Generalized deep image to image regression. In *Proc. CVPR*, pages 5609–5619, 2017.
- [40] Saurabh Saxena, Charles Herrmann, Junhwa Hur, Abhishek Kar, Mohammad Norouzi, Deqing Sun, and David J Fleet. The surprising effectiveness of diffusion models for optical flow and monocular depth estimation. *arXiv preprint arXiv:2306.01923*, 2023.
- [41] Guo Shi, Yan Zifei, Zhang Kai, Zuo Wangmeng, and Zhang Lei. Toward convolutional blind denoising of real photographs. In *arXiv preprint arXiv:1807.04686*, 2018.
- [42] Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. *arXiv preprint arXiv:2010.02502*, 2020.
- [43] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.

- [44] Ying Tai, Jian Yang, Xiaoming Liu, and Chunyan Xu. Memnet: A persistent memory network for image restoration. In *Proc. ICCV*, pages 4539–4547, 2017.
- [45] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep image prior. In *Proc. CVPR*, pages 9446–9454, 2018.
- [46] Yinhuai Wang, Jiwen Yu, and Jian Zhang. Zero-shot image restoration using denoising diffusion null-space model. *arXiv preprint arXiv:2212.00490*, 2022.
- [47] Zichun Wang, Ying Fu, Ji Liu, and Yulun Zhang. Lg-bpn: Local and global blind-patch network for self-supervised real-world denoising. In *Proc. CVPR*, pages 18156–18165, 2023.
- [48] Xiaohe Wu et al. Unpaired learning of deep image denoising. In *Proc. ECCV*, pages 352–368. Springer, 2020.
- [49] Junyuan Xie, Linli Xu, and Enhong Chen. Image denoising and inpainting with deep neural networks. In *Proc. NeurIPS*, pages 341–349, 2012.
- [50] Yaochen Xie, Zhengyang Wang, and Shuiwang Ji. Noise2same: Optimizing a self-supervised bound for image denoising. In *Proc. NeurIPS*, volume 33, pages 20320–20330, 2020.
- [51] Hao Xu, Haipeng Li, Yinqiao Wang, Shuaicheng Liu, and Chi-Wing Fu. Handbooster: Boosting 3d hand-mesh reconstruction by conditional synthesis and sampling of hand-object interactions. In *Proc. CVPR*, pages 10159–10169, 2024.
- [52] Jun Xu, Yuan Huang, Ming-Ming Cheng, Li Liu, Fan Zhu, Zhou Xu, and Ling Shao. Noisy-as-clean: learning self-supervised denoising from corrupted image. *IEEE Trans. on Image Processing*, 29:9316–9329, 2020.
- [53] Cheng Yang, Cong Wang, Lijing Liang, and Zhixun Su. Real-world image denoising via efficient diffusion model with controllable noise generation. *Journal of Electronic Imaging*, 33(4):043003–043003, 2024.
- [54] Zongsheng Yue, Qian Zhao, Lei Zhang, and Deyu Meng. Dual adversarial network: Toward real-world noise removal and noise generation. *arXiv preprint arXiv:2007.05946*, 2020.
- [55] Quan Yuhui, Zheng Tianxiang, and Ji Hui. Pseudo-siamese blind-spot transformers for self-supervised real-world denoising. In *NIPS*, 2025.
- [56] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Learning enriched features for real image restoration and enhancement. *arXiv preprint arXiv:2003.06792*, 2020.
- [57] Dan Zhang, Fangfang Zhou, Yuwen Jiang, and Zhengming Fu. Mm-bsn: Self-supervised image denoising for real-world with multi-mask based on blind-spot network. In *Proc. CVPR*, pages 4189–4198, 2023.
- [58] 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. on Image Processing*, 26(7):3142–3155, 2017.
- [59] Kai Zhang, Wangmeng Zuo, Shuhang Gu, and Lei Zhang. Learning deep cnn denoiser prior for image restoration. In *Proc. CVPR*, pages 3929–3938, 2017.
- [60] Kai Zhang, Wangmeng Zuo, and Lei Zhang. Ffdnet: Toward a fast and flexible solution for cnn-based image denoising. *IEEE Trans. on Image Processing*, 27(9):4608–4622, 2018.
- [61] Lvmin Zhang, Anyi Rao, and Maneesh Agrawala. Adding conditional control to text-to-image diffusion models. In *Proc. ICCV*, pages 3836–3847, 2023.
- [62] Tianhao Zhou, Haipeng Li, Ziyi Wang, Ao Luo, Chen-Lin Zhang, Jiajun Li, Bing Zeng, and Shuaicheng Liu. Recdiffusion: Rectangling for image stitching with diffusion models. In *Proc. CVPR*, pages 2692–2701, 2024.
- [63] Yuqian Zhou, Jianbo Jiao, Haibin Huang, Yang Wang, Jue Wang, Honghui Shi, and Thomas Huang. When awgn-based denoiser meets real noises. In *Proc. AAAI*, pages 13074–13081, 2020.
- [64] Yuqian Zhou, Ding Liu, and Thomas Huang. Survey of face detection on low-quality images. In *2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018)*, pages 769–773. IEEE, 2018.

Blind-Spot Guided Diffusion for Self-supervised Real-World Denoising

SUPPLEMENTARY MATERIAL

Inference time. Our approach involves multiple sampling steps related to diffusion, as well as a multi-round sampling process. In our optimal experiments, we selected a complete sampling process consisting of 8 sampling iterations across 8 rounds, resulting in a total of 64 inference steps. Similarly, existing BSN methods, such as PUCA, also incorporate multiple inference processes. However, even when we allocate a sufficient number of inference steps to PUCA, there is no significant improvement in performance, as shown in Tab. 2. Additionally, we observe a substantial gap between current self-supervised methods and supervised approaches, indicating that the importance of enhancing performance outweighs that of model size. Conversely, even when our model performs only a single round of sampling with 8 inference steps, it significantly outperforms PUCA, as illustrated in Fig. 6.

Guidance. In our approach, BSN serves as a guide to direct the non-blind-spot UNet during the sampling process. Through multiple diffusion sampling iterations, the information from both components is synergistically integrated. As shown in Fig. 5, we observe that when the guidance is minimal, the PSNR performance across the entire dataset may be suboptimal; however, the denoised results exhibit significant detail recovery and are free of noise. This method of sampling for non-blind-spot information fusion presents a novel direction for self-supervised denoising approaches.

More Details. Our UNet-based diffusion model follows a self-supervised denoising framework without pre-trained weights. While L2 loss is standard for DMs, we opted for L1 loss for the non-BSN branch due to its very slightly better performance, which we attribute to its alignment with image denoising. In the revision, we included a comparison of L1 vs. L2 results. Additionally, we have detailed the computational costs (parameters, FLOPs) for both branches to address broader implementation concerns.

	ours-BSN	ours-UNet	SD-UNet
Params(M)&Macs(G)	18.1/137.23	9.2/30.92	865/880

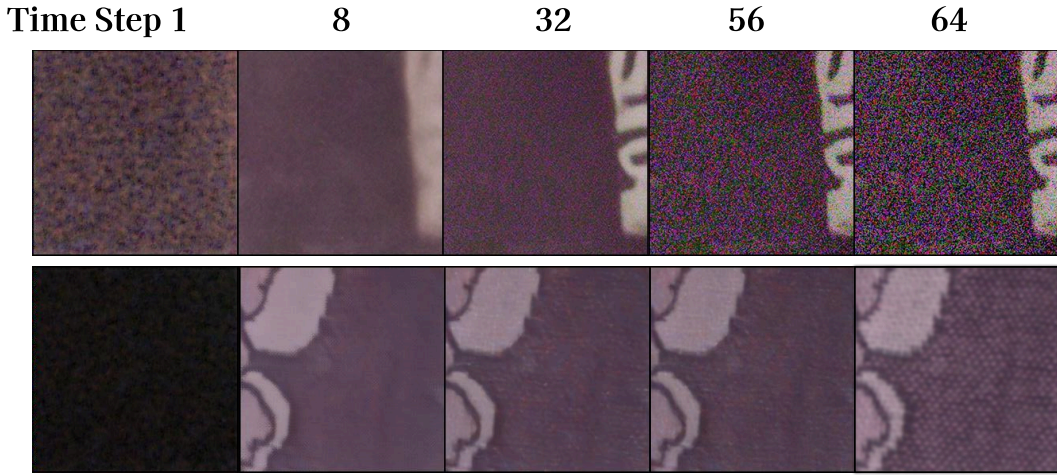


Figure 7: Output changes during the sampling process. Top: sampling without BSN prior. Bottom: sampling with BSN prior.

*Corresponding author

Algorithm 1 Training

```
1: repeat
2:    $\mathbf{x}_0 \sim p_0(\mathbf{x})$ 
3:    $t \sim \text{Uniform}(\{1, \dots, T\})$ 
4:    $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ 
5:   Take gradient descent step on
      $\nabla_{f_u} \|\mathbf{x}_0 - f_u(\sqrt{\alpha_t}\mathbf{x}_0 + \sqrt{1 - \alpha_t}\epsilon, t)\|$ 
      $\nabla_{f_b} \|\mathbf{x}_0 - f_b(\sqrt{\alpha_t}\mathbf{x}_0 + \sqrt{1 - \alpha_t}\epsilon, \mathbf{x}_0, t)\|$ 
6: until converged
```

Algorithm 2 Random Replacement Sampling

```
1:  $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ 
2: for  $t = T, \dots, 1$  do
3:    $z \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$  if  $t > 1$ , else  $z = \mathbf{0}$ 
4:   pixelshuffle  $\mathbf{x}_t$  if  $t = T$ 
5:    $\epsilon_\theta(\mathbf{x}_t) = \frac{\mathbf{x}_t - \sqrt{\alpha_t} \cdot f_\theta(\mathbf{x}_t, t)}{\sqrt{1 - \alpha_t}}, \theta \in [u, b]$ .
6:    $\epsilon(\mathbf{x}_t) = (1 - w)\epsilon_{f_b}(\mathbf{x}_t, \mathbf{x}_0) + w\epsilon_{f_u}(\mathbf{x}_t)$ .
7:    $\mathbf{x}_0|\mathbf{x}_t = \frac{\mathbf{x}_t - \sqrt{1 - \alpha_t} \cdot \epsilon(\mathbf{x}_t)}{\sqrt{\alpha_t}}$ 
8:   replace  $\mathbf{x}_0|\mathbf{x}_t$  with  $\mathbf{x}_0$ 
9:    $\mathbf{x}_{t-1} = \sqrt{1 - \alpha_{t-1} - \sigma_t^2} \cdot \epsilon(\mathbf{x}_t) + \sqrt{\alpha_{t-1}}\mathbf{x}_0|\mathbf{x}_t + \sigma_t\mathbf{z}$ ,
10:  pixelunshuffle  $\mathbf{x}_t$  if  $t < T$ 
11: end for
12: return Average  $\mathbf{x}_0|\mathbf{x}_T, \dots, \mathbf{x}_0|\mathbf{x}_1$ 
```

More Results. In the Fig. 8, we present additional subjective comparison results. These results primarily focus on images with weak texture richness. We observe that, in scenarios characterized by weak textures and relatively high signal-to-noise ratios (SNR), employing a lower guidance strength yields better outcomes. Conversely, for other types of images, a higher guidance strength is recommended.

Although we standardized the selection of a specific empirical value for the final weight w in our benchmarks, tailoring the guidance strength based on the SNR could further enhance overall performance. A key insight from this work is our discovery of a method to integrate information from non-BSN structures, which may potentially eliminate the need for BSN in future applications.

Algorithm 3 Complementary Replacement Sampling

```
1:  $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$   
2:  $\mathbf{x}_0|\mathbf{x}_R = \text{Sample\_proces}(\mathbf{x}_T)$   
3: for  $r = R - 1, \dots, 1$  do  
4:   replace  $\mathbf{x}_0$  with  $\mathbf{x}_0|\mathbf{x}_R$   
5:    $\mathbf{x}_0|\mathbf{x}_r = \text{Sample\_proces}(\mathbf{x}_T)$   
6: end for  
7: return Average  $\mathbf{x}_0|\mathbf{x}_T, \dots, \mathbf{x}_0|\mathbf{x}_1$ 
```

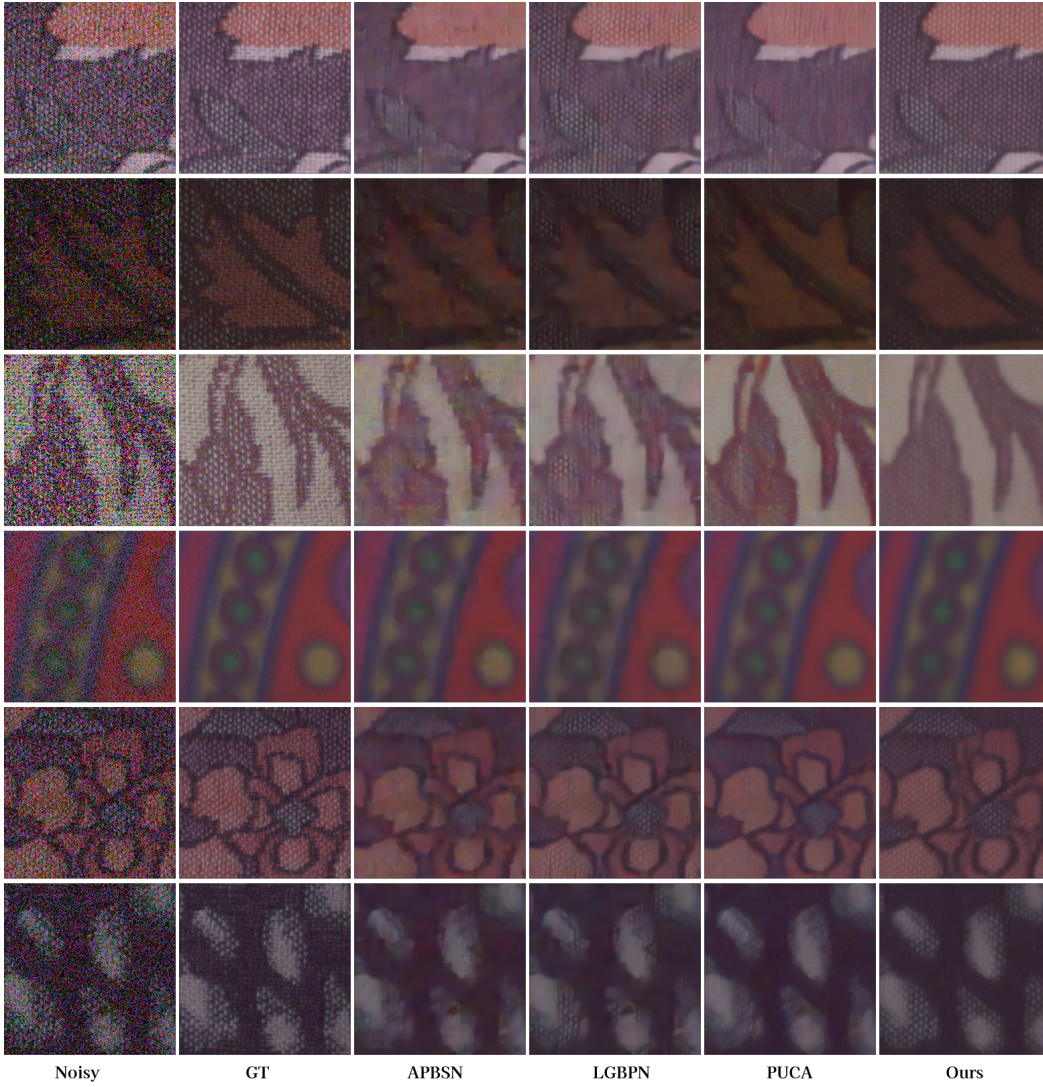


Figure 8: Our results presented were denoised using a guidance of $w = 0.1$