

Exploiting Deblurring Networks for Radiance Fields

Haeyun Choi^{1,†}

Heemin Yang²

Janghyeok Han²

Sunghyun Cho²

¹KT

²POSTECH

Abstract

In this paper, we propose DeepDeblurRF, a novel radiance field deblurring approach that can synthesize high-quality novel views from blurred training views with significantly reduced training time. DeepDeblurRF leverages deep neural network (DNN)-based deblurring modules to enjoy their deblurring performance and computational efficiency. To effectively combine DNN-based deblurring and radiance field construction, we propose a novel radiance field (RF)-guided deblurring and an iterative framework that performs RF-guided deblurring and radiance field construction in an alternating manner. Moreover, DeepDeblurRF is compatible with various scene representations, such as voxel grids and 3D Gaussians, expanding its applicability. We also present BlurRF-Synth, the first large-scale synthetic dataset for training radiance field deblurring frameworks. We conduct extensive experiments on both camera motion blur and defocus blur, demonstrating that DeepDeblurRF achieves state-of-the-art novel-view synthesis quality with significantly reduced training time.

1. Introduction

Novel-view synthesis has seen significant advancements in recent years, leading to impressive improvements in photo-realistic rendering quality. One key development in this field is the Neural Radiance Field (NeRF) [25], which leverages neural networks to generate highly detailed images from novel viewpoints. Building on this, various works have focused on reducing training time and enhancing rendering quality [3, 12, 15, 27]. More recently, 3D Gaussian Splatting (3DGS) [17] has emerged, utilizing Gaussians as an explicit 3D representation. This technique enables high-quality scene reconstruction and real-time rendering through a differentiable rasterization method.

However, synthesizing sharp novel views from degraded training views remains a significant challenge. Capturing images in real-world conditions often leads to various degra-

dations such as blur and noise. Among them, blur makes it difficult to aggregate accurate 3D information from the training views, resulting in a blurry radiance field. To address blur, several works [20–22, 24, 31, 32, 39] have been proposed to reconstruct sharp radiance fields from training views with either camera motion blur or defocus blur. These radiance field deblurring methods jointly optimize blur kernels and the radiance field, and have shown promising results.

However, these methods still possess several limitations that hinder their effectiveness. First, existing approaches rely on linear blur models, describing blurred pixels in training views as linear combinations of sharp pixels. In contrast, real-world blurred images often suffer from non-linear outliers such as saturated pixels due to sensor clipping, noise from various sources, and non-linear in-camera processing [8, 35]. These violate the linear blur model, leading to deblurring artifacts and severely degrading the performance of current radiance field deblurring techniques. Second, existing approaches do not utilize priors on sharp images, relying solely on complementary information from different views to model blur kernels and estimate a sharp radiance field. Consequently, they tend to produce radiance fields with residual blur and may completely fail when all input views exhibit similar blur directions [24], akin to classical multi-frame deblurring approaches [5, 7, 33], due to the lack of sharp image priors. Third, existing ray-based methods [21, 22, 24, 31, 39] employ multiple ray samples per pixel to depict a blurry image, resulting in a substantial increase in rendering time for a single view from a radiance field, and consequently, a significant computation time for radiance field construction.

Apart from the radiance field deblurring methods, image deblurring has been extensively studied for decades [6, 14, 30, 40, 43]. Recently, a number of deep neural network (DNN)-based single-image deblurring methods have been proposed [4, 10, 19, 28, 38, 42]. These methods do not rely on linear blur models, but learn image deblurring from large-scale datasets. Thanks to this, they can effectively remove blur from a single image without the need for complementary information from other images, and also handle non-linearities such as saturated pixels [34]. Moreover, their feed-forward approach allows markedly reduced processing times,

[†]This work was done at POSTECH.

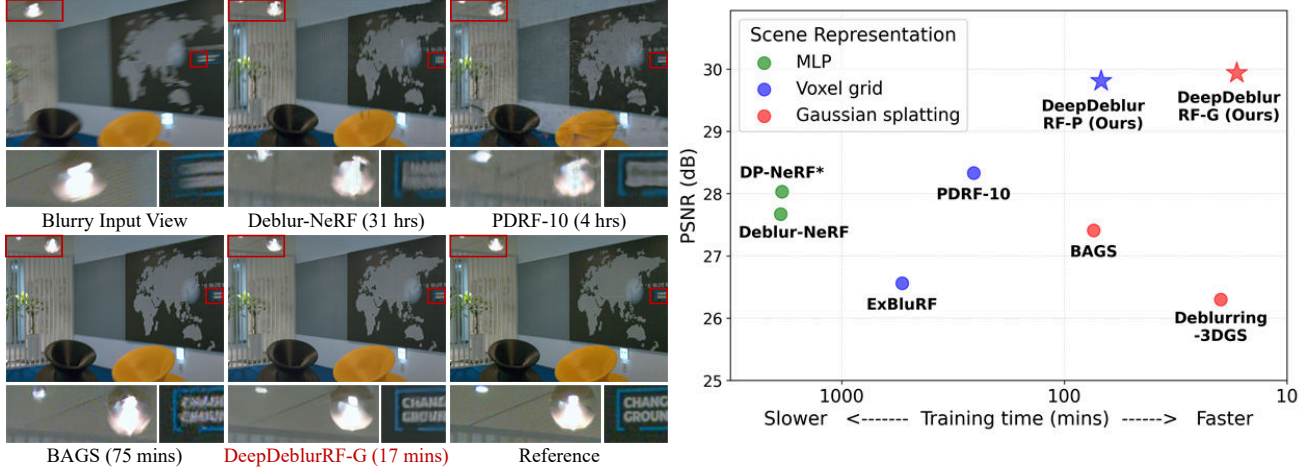


Figure 1. Given a set of multi-view blurry images even with non-linear outliers such as saturated pixels and noise, DeepDeblurRF performs high-quality novel-view synthesis with highly efficient training. DeepDeblurRF-P and DeepDeblurRF-G are the results of our framework, where radiance fields are constructed using Plenoxels [12] and 3D Gaussian Splatting [17], respectively. Note that DP-NeRF [21] was trained with two GPUs due to its memory demands, whereas other models were trained on a single NVIDIA TITAN RTX GPU.

unlike classical iterative optimization-based approaches.

Given the effectiveness of recent DNN-based deblurring approaches, a promising direction to construct a sharp radiance field from blurry images would be to combine a deblurring network and radiance field construction. However, a naïve combination of a deblurring network and radiance field construction, which performs single-image deblurring to each input blurry image and trains a radiance field using the deblurred images, results in unsatisfactory results as reported by Ma *et al.* [24]. This is because of the limited performance of single-image deblurring, which stems from the insufficient information available in a single image.

In this paper, we propose *DeepDeblurRF*, a novel radiance field deblurring framework that enables highly efficient training and high-quality novel-view synthesis from images blurred by motion blur or defocus blur, even in the presence of noise and saturated pixels, as shown in Fig. 1. Unlike existing radiance field deblurring approaches, our method leverages DNN-based deblurring modules to achieve superior performance and computational efficiency. Additionally, our approach is versatile, capable of handling various 3D representations such as voxel grids [12] and 3D Gaussians [17].

However, solely relying on deblurring networks to remove blur from input blurry images may yield unsatisfactory results as discussed above. To effectively combine DNN-based deblurring and radiance field construction, our framework adopts a novel radiance field (RF)-guided deblurring scheme, and an iterative framework that performs RF-guided deblurring and radiance field construction in an alternating manner. Specifically, our framework first deblurs input images, and then trains a radiance field using the deblurred images. Then, at the next iteration, it renders images corresponding to the input views from the trained radiance field. While the de-

blurred images may have limited qualities with inaccurately restored details, the rendered images provide higher-quality details as they are rendered from aggregated information from multiple deblurred images. Using the rendered images as guidance, our framework then performs RF-guided deblurring to the input blurred images. Thanks to the guidance, we can obtain higher-quality deblurred images, which we use for radiance field construction again. Finally, our framework iterates RF-guided deblurring and radiance field construction to gradually enhance the quality of the radiance field.

Training the deblurring modules of our framework requires a large-scale dataset of blurred images paired with ground-truth sharp images for radiance field reconstruction. However, no such datasets are available as existing radiance field deblurring approaches adopt linear blur model-based approaches, which do not require large-scale training datasets [20–22, 24, 31, 32, 39]. Consequently, the datasets provided by previous works, such as those from Deblur-NeRF [24] and ExBluRF [22], are small and primarily designed for evaluation rather than training. Furthermore, these datasets often overlook crucial factors like camera noise, in-camera processing, and other non-linear outliers, which are essential for realistic deblurring tasks.

To address these gaps, we propose *BlurRF-Synth*, a large-scale dataset for radiance field deblurring approaches. It includes 4,350 blurred-sharp image pairs across 150 scenes, with 2,175 pairs for each of 75 scenes, encompassing both camera motion blur and defocus blur. These images are carefully synthesized to reflect real-world camera degradations such as noise, saturated pixels, and in-camera processing pipelines. Additionally, we introduce a real-world dataset called *BlurRF-Real* for evaluation in non-ideal conditions. Unlike the Deblur-NeRF dataset, which has less noise and

adequate lighting, *BlurRF-Real* consists of five low-light indoor scenes captured with a machine vision camera.

We validate *DeepDeblurRF* on both synthetic and real-world datasets with two types of blur: camera motion blur and defocus blur. Experimental results demonstrate that our method achieves state-of-the-art novel-view synthesis performance with highly reduced processing times. Furthermore, we demonstrate the extensibility of our framework by constructing radiance fields using different scene representations, such as voxel grids and Gaussian Splatting. Our contributions can be summarized as follows:

- We propose *DeepDeblurRF*, a novel radiance field deblurring framework, which is the first approach that leverages DNN-based deblurring modules to overcome the limitations of the linear blur model.
- To this end, we present RF-guided deblurring, and an iterative framework that performs RF-guided deblurring and radiance field construction in an alternating manner to gradually enhance the quality of the radiance field.
- We also present the *BlurRF-Synth* dataset, the first large-scale dataset for training and evaluation of novel-view synthesis from blurry images.

2. Related Work

Radiance field deblurring Several works have recently been proposed to tackle radiance field deblurring [21, 22, 24, 31, 39]. Ma *et al.* presented Deblur-NeRF, the first framework to train a sharp radiance field from blurry training views [24]. Deblur-NeRF adopts an additional MLP to model blur kernels and jointly optimizes them with a radiance field. DP-NeRF [21] models blur kernels based on physical priors for rigid motions, while PDRF [31] uses a two-stage kernel modeling scheme where the first stage coarsely aggregates kernel points to reduce computational cost. All these methods support both camera motion blur and defocus blur by modeling them using multiple rays. On the other hand, BAD-NeRF [39] and ExBluRF [22] propose radiance field deblurring for camera motion blur, modeling camera motion blur using camera trajectories. Nevertheless, all these methods rely on linear blur models, and require multiple ray-tracing steps, resulting in limited deblurring quality and slow training speeds.

More recently, with the emergence of 3D Gaussian Splatting [17], 3D Gaussian-based radiance field deblurring methods have been proposed. Deblurring-3DGS [20] handles camera motion blur with small offsets in 3D Gaussian positions and defocus blur by adjusting the geometry of Gaussians. BAGS [32] models both types of blur using dense per-pixel blur kernels optimized in a coarse-to-fine manner. However, BAGS requires substantial computation time to optimize dense blur kernels. Moreover, these methods also rely on linear blur models and are limited in handling real-world blurred images.

Image deblurring Image deblurring has been widely studied to address blur that degrades image quality. Traditional methods, relying on linear blur models, often struggle with real-world blurred images containing non-linear outliers [6, 7, 9, 14, 30, 33, 40, 43]. Their iterative optimization processes are also computationally intensive, limiting their practical use. Recent advancements in deep learning have introduced DNN-based approaches, including single-image [4, 10, 19, 28, 38, 42] and multi-frame/video deblurring [2, 11, 13, 16, 29, 37, 41, 44, 45]. These methods can handle more complex blurs and reduce processing time with a feed-forward structure.

While applying DNN-based deblurring approaches might be a potential solution for constructing radiance fields from blurry images, naively using single-image deblurring techniques yields unsatisfactory results as discussed in Sec. 1. Another possible direction is to adopt multi-frame and video deblurring approaches that leverage complementary information from multiple images for higher-quality results. However, these methods assume that input images are captured from nearly the same viewpoints, a condition that does not hold for training views in radiance field construction.

3. DeepDeblurRF

Fig. 2-(a) illustrates the overall framework of DeepDeblurRF. It takes a set of M blurred images $\mathbf{B} = \{B_1, \dots, B_M\}$ of a scene and estimates a sharp radiance field that can synthesize a sharp novel view given an arbitrary pose. We assume that each blurred image B_m is obtained from its latent sharp image L_m through an unknown degradation process including blur and non-linear outliers such as saturated pixels and nonlinear in-camera processing. Based on this assumption, DeepDeblurRF first performs initial deblurring to the input blurred images, and obtains initial deblurred images. Then, our method iteratively performs radiance field construction using deblurred images and RF-guided deblurring to gradually enhance the quality of the radiance field and the deblurred images. At the last iteration, we perform only the radiance field construction step and obtain a final radiance field, from which we can synthesize sharp novel views. In the following, each step of DeepDeblurRF is described in more detail.

3.1. Initial Deblurring

The initial deblurring step removes blur from each blurry training view $B_m \in \mathbf{B}$ so that the following radiance field construction step can estimate the pose of each view more accurately, and more effectively aggregate information from different views. To this end, we adopt an off-the-shelf single-image deblurring network. Specifically, we adopt NAFNet [4], a state-of-the-art deblurring network, for its computational efficiency and performance. We denote the deblurred image from the initial deblurring step correspond-

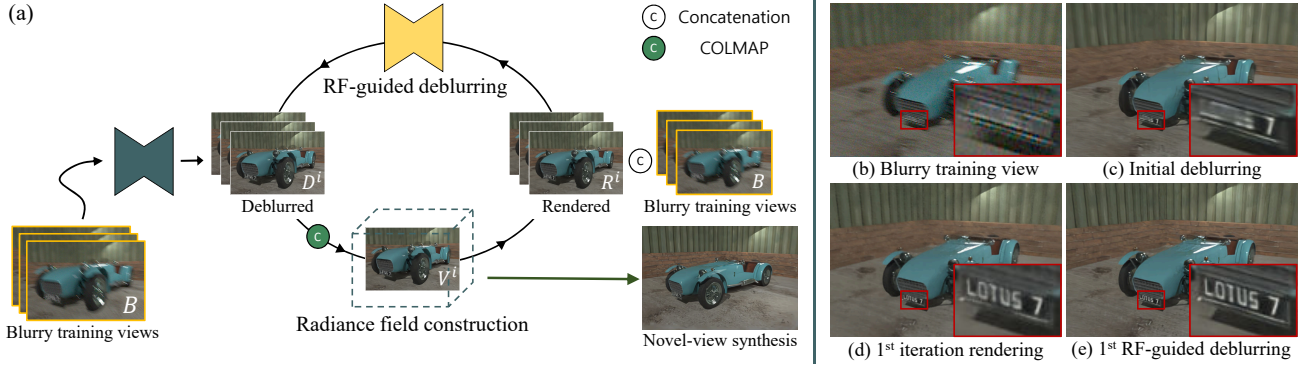


Figure 2. Overall framework and intermediate result of each step of DeepDeblurRF.

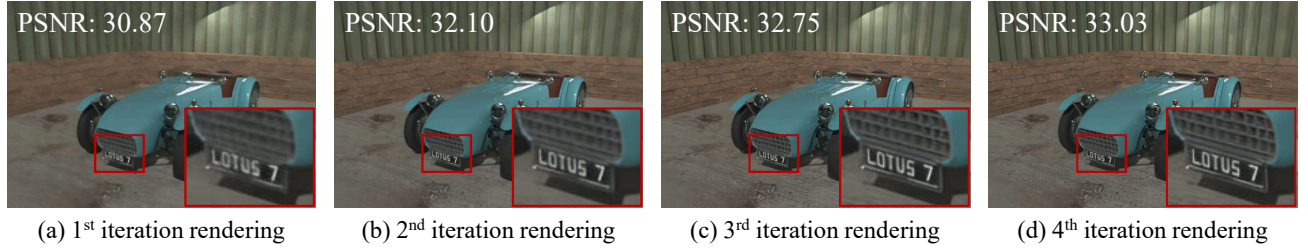


Figure 3. As iterations progress, the rendered images contain increasingly high-quality scene information, which subsequently improves the RF-guided deblurring network’s performance in the next iteration.

ing to B_m as D_m^0 , and the set of the deblurred images as $\mathbf{D}^0 = \{D_1^0, \dots, D_M^0\}$.

3.2. Radiance Field Construction

In the radiance field construction step at the i -th iteration, we first estimate the camera poses of the deblurred images from the previous step, denoted as $\mathbf{D}^{i-1} = \{D_1^{i-1}, \dots, D_M^{i-1}\}$ using COLMAP [36], where $i \in [1, N]$ is the index of our iterative process. Using the estimated poses and deblurred images, we aggregate information about the target 3D scene and construct a radiance field $\mathbf{V}^i$. As discussed in Sec. 1, our framework can adopt diverse radiance field representations. To demonstrate the extensibility of our framework, we adopt two different representations, Plenoxels [12] and 3D Gaussians [17], in our experiments. We refer to our framework using Plenoxels and 3D Gaussians as DeepDeblurRF-P and DeepDeblurRF-G, respectively.

Note that deblurred images $\mathbf{D}^{i-1}$ from the previous step may still contain residual blur and deblurring artifacts, which can introduce corrupted information into the aggregation process. Nevertheless, since the input images have overlapping regions, other deblurred images can provide information for those regions that appear corrupted in some images. By aggregating such information from different deblurred images on the same regions, the radiance field construction step can effectively suppress corrupted information and obtain a higher-quality radiance field, from which we can render higher-quality images compared to the input deblurred im-

ages at this step.

3.3. RF-guided Deblurring

The RF-guided deblurring step deblurs the input views $\mathbf{B}$ using the aggregated information in $\mathbf{V}^i$. Specifically, for each input view B_m , we render $\mathbf{V}^i$ to obtain a rendered image R_m^i . The rendered image R_m^i shares the same content as its corresponding deblurred image D_m^{i-1} , but contains fewer artifacts and finer details. Motivated by this, our RF-guided deblurring leverages R_m^i to obtain an updated deblurred image D_m^i from B_m .

To perform RF-guided deblurring, we employ a novel RF-guided deblurring network that takes both the rendered image R_m^i and the input blurred image B_m to guide the deblurring process with the aggregated information in R_m^i . For the deblurring network, we adopt NAFNet [4] and modify its first layer to take a concatenation of the two images. Once we obtain the updated deblurred images $\mathbf{D}^i$, we return to the radiance field construction step and carry out the process for the $(i + 1)$ -th iteration, gradually enhancing the quality of the radiance field and the deblurred images.

The intermediate results of each step are visualized in Fig. 2-(b) to Fig. 2-(e). Due to the severe blur in the input blurred image, its initial deblurring result retains some residual blur. In contrast, thanks to the information aggregation occurring in the radiance field construction step, the rendered image contains more accurate details. Finally, thanks to the guidance of the rendered image, the RF-guided deblurring

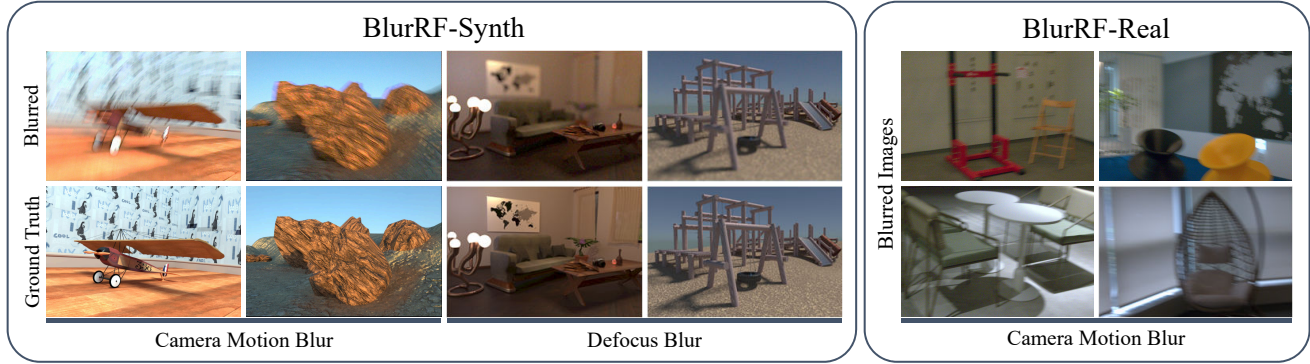


Figure 4. Examples of the BlurRF-Synth and BlurRF-Real datasets. The examples show blurred views in BlurRF-Real, while the top and bottom rows in BlurRF-Synth include blurred views and their corresponding sharp views.

result shows sharper and more precise details that cannot be observed in the intermediate results of previous steps.

Fig. 3 visualizes rendered images at different iterations. Although the rendered image at the first iteration Fig. 3-(a) contains blurry details, they are gradually refined at each iteration (Fig. 3 (b)-(d)) due to our iterative process, which incrementally improves the quality of the radiance fields and the deblurred images. The effectiveness of our iterative process is also demonstrated by the increasing PSNR score as iterations progress.

4. BlurRF Dataset

4.1. BlurRF-Synth

We propose the BlurRF-Synth dataset, the first large-scale dataset for training and evaluating novel-view synthesis from blurred images. The dataset comprises a train set and a test set for each type of blur: camera motion blur and defocus blur. The train set includes 65 scenes, each containing 29 pairs of synthetically blurred images and their corresponding sharp images, captured from different viewpoints. The test set includes 10 scenes, each with 29 blurred-sharp image pairs and five additional sharp images from different viewpoints for evaluating novel-view synthesis quality. Examples of blurred images and ground-truth sharp images are shown in Fig. 4. We synthetically generated the dataset using Blender models, collecting 95 models from Blendswap¹ under Creative Commons licenses and 5 models from the synthetic dataset of Deblur-NeRF [24]. For each model, we sampled 29 camera poses to capture different viewpoints.

Camera motion blur We simulated camera shakes by randomly sampling a camera trajectory for each pose during the exposure duration. To this end, we adopted Bézier interpolation to model 6-DOF camera motion and densely sampled 51 intermediate poses along this camera trajectory. We then rendered the scene using Blender at each intermediate camera pose to produce a series of sharp images, which we then

averaged to create a blurred image. All rendering and averaging were performed in the linear sRGB color space to accurately replicate the real-world image formation process. Among the sharp images, we sampled the temporally central image (i.e., the 26th image) as the ground-truth sharp image.

Defocus blur We also generated images with defocus blur using Blender. We adjusted the camera’s depth-of-field (DoF) in Blender to create defocus blur. We controlled the aperture size and randomly set the blade count between 7 and 9 to simulate realistic defocus effects. For each camera pose, we sampled a random focal distance within a predefined range based on the scene scale, shifting the focal plane to introduce variation in the defocus blur.

Realistic blur synthesis To reflect real-world image degradation, we adopted the blur synthesis pipeline of RSBlur [35]. Specifically, we generated camera motion blur by averaging sharp images in the linear sRGB space. Then, we synthesized saturated pixels to the blurred images and converted them to the camera RAW space. In the camera RAW space, we synthesized shot and read noise, and converted the images back to the camera sRGB space. For defocus blur, we rendered blurred images in the linear sRGB space and added noise in the same manner. For a more detailed pipeline, we refer the readers to [35]. The blur synthesis pipeline requires camera-specific noise parameters and a color correction matrix. We used a Sony A7R3 camera to estimate them. For more details about the BlurRF-Synth dataset, we refer the readers to the supplementary material.

4.2. BlurRF-Real

For evaluation under more realistic and challenging conditions, we propose BlurRF-Real, a real-world low-light camera motion blur dataset. The existing real datasets [22, 24] are captured in well-lit environments with minimal noise, making them less representative of camera motion blur in low-light conditions. Our dataset addresses this gap by providing data captured under challenging low-light scenarios. We collected five indoor scenes using a machine vision camera,

¹<https://blendswap.com/>

3D Representation	Model	Camera Motion Blur			Defocus Blur			Computation Time (Hr.)
		PSNR ($\uparrow$)	SSIM ($\uparrow$)	LPIPS ($\downarrow$)	PSNR ($\uparrow$)	SSIM ($\uparrow$)	LPIPS ($\downarrow$)	
MLP	Deblur-NeRF [24]	27.67	0.8340	0.1450	30.03	0.8727	0.1137	31.33
	BAD-NeRF [39]	21.74	0.5298	0.3969	-	-	-	24.68
	DP-NeRF [21]	28.03	0.8412	0.1267	30.15	0.8763	0.0991	30.00*
Voxel grid	ExBluRF [22]	26.56	0.7823	0.1955	-	-	-	8.92
	PDRF-10 [31]	28.33	0.8435	0.1495	30.03	0.8750	0.1225	4.26
	DeepDeblurRF-P	29.81	0.8668	0.1142	32.51	0.9058	0.0961	1.14
3D Gaussians	Deblurring-3DGS [20]	26.30	0.7729	0.1728	29.37	0.8545	0.1470	0.33
	BAGS [32]	27.41	0.8108	0.1382	29.90	0.8638	0.1152	1.25
	DeepDeblurRF-G	29.94	0.8681	0.1059	32.58	0.9060	0.0774	0.28

Table 1. Quantitative results of novel-view synthesis on BlurRF-Synth test scenes. We highlight the best metrics and the second best metrics. Note that the computation time of DP-NeRF [21] was measured with two GPUs due to its memory demands, whereas those of the other models were measured on a single GPU.

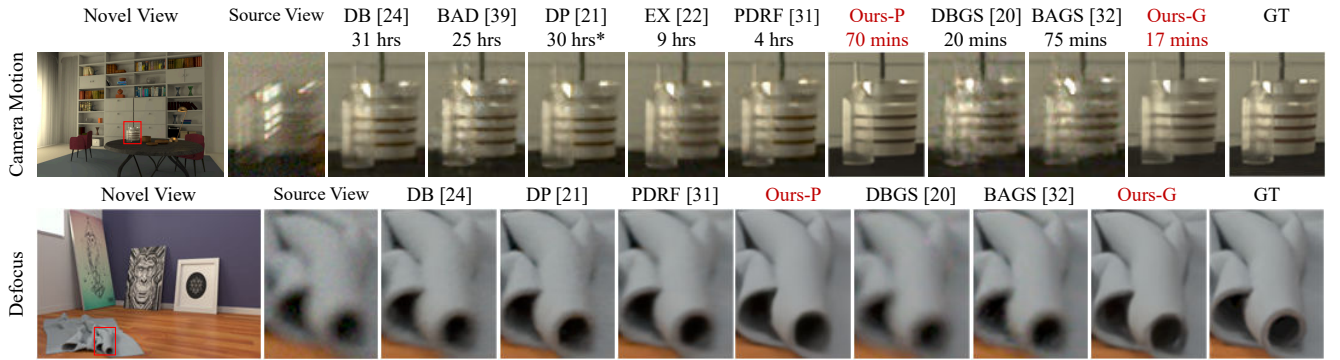


Figure 5. Qualitative results of novel-view synthesis on BlurRF-Synth test scenes.

each providing 20-40 multi-view blurry images, including 3-5 images for novel-view synthesis evaluation. Examples of blurred training views are shown in Fig. 4.

5. Experiments

Datasets The deblurring networks of DeepDeblurRF are trained with the train set of BlurRF-Synth. To evaluate DeepDeblurRF and other radiance field deblurring methods, we use the test set of BlurRF-Synth, the BlurRF-Real dataset, and the benchmark datasets from Deblur-NeRF [24]. The benchmark datasets include both synthetic and real-world scenes. The synthetic dataset consists of five scenes, while the real-world dataset contains ten scenes for each blur type. The real scenes of Deblur-NeRF were captured with a Canon EOS RP in manual exposure mode. The camera poses are estimated using COLMAP [36].

Implementation details We set the number of iterations N to 5. Thus, as detailed in Sec. 3, DeepDeblurRF includes one single-image deblurring network and four RF-guided deblurring networks for each blur type: camera motion blur and defocus blur. Note that our approach requires an additional training phase for the deblurring networks, unlike

previous radiance field deblurring methods. Once trained, our framework can construct radiance fields for arbitrary scenes and render novel views, similar to conventional approaches. To train the deblurring networks, we use the Adam optimizer [18] with $\beta_1 = 0.9$ and $\beta_2 = 0.9$. The learning rate is initially set to 10^{-3} and gradually reduced to 10^{-7} using the cosine annealing scheduler [23].

For constructing radiance fields in DeepDeblurRF-P, we follow the forward-facing experimental setup in Plenoxels, applying pruning and upsampling to the voxel grids every 38,400 iterations. For DeepDeblurRF-G, we adopt the original Gaussian Splatting method, incorporating the depth-based pruning and additional points strategy from DeblurGS [20] to address the sparse point clouds obtained from COLMAP on blurred images.

We measured the computation time for radiance field deblurring using the camera motion blur test set of BlurRF-Synth on a single NVIDIA TITAN RTX, except for DP-NeRF, which requires two GPUs due to its memory demands. Additional implementation details, including network configurations and the training settings for each iteration of DeepDeblurRF-P and DeepDeblurRF-G, are provided in the supplementary material.

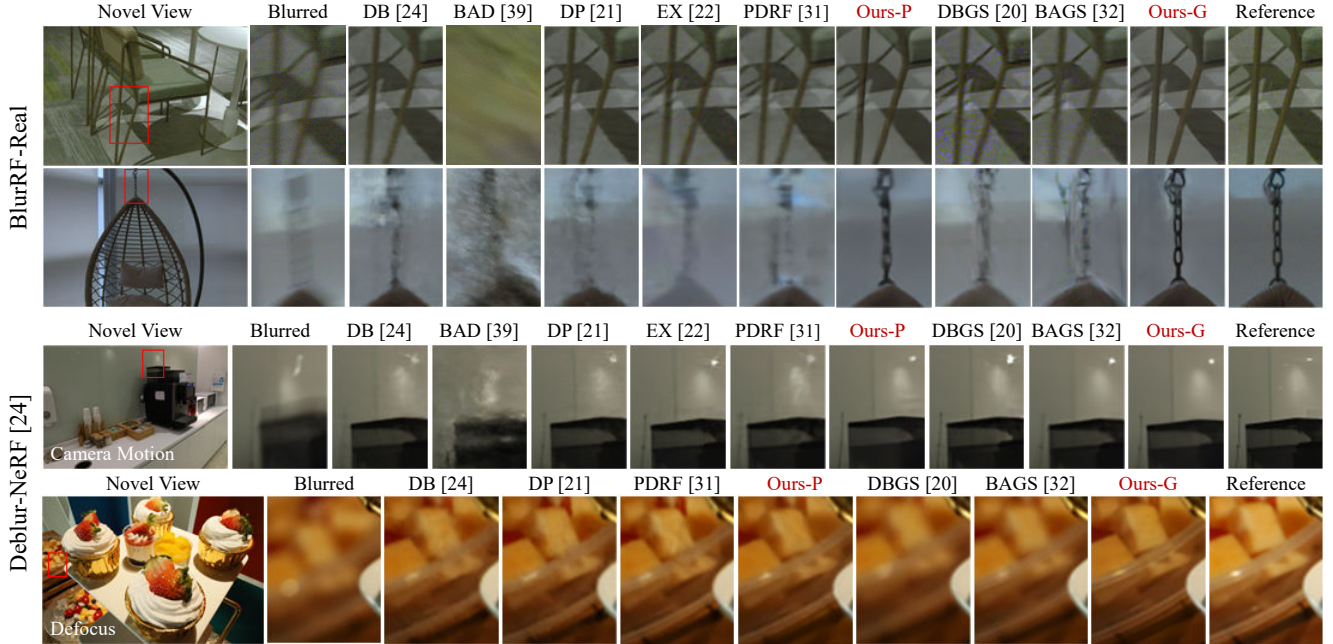


Figure 6. Qualitative results of novel-view synthesis on real-world datasets.

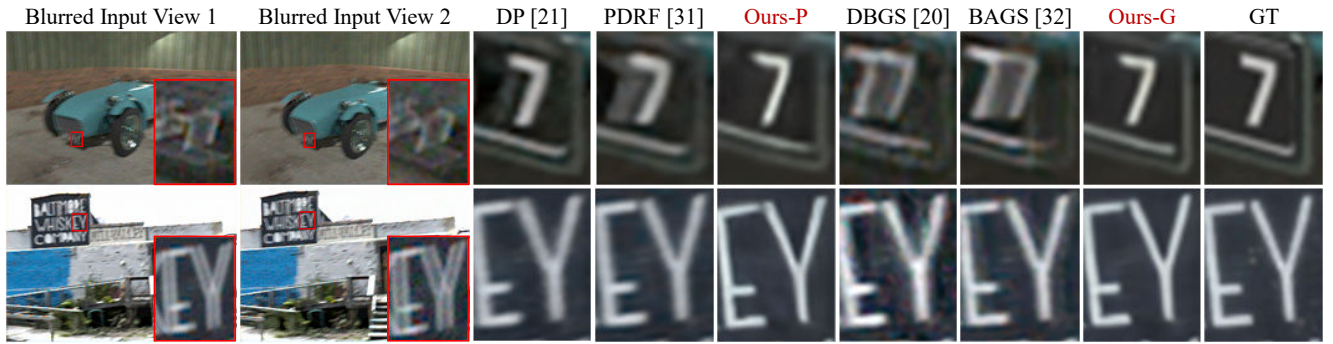


Figure 7. Qualitative results of novel-view synthesis on BlurRF-SB dataset.

Comparison using BlurRF-Synth We compare DeepDeblurRF with other radiance field deblurring methods [20–22, 24, 31, 32, 39]. Tab. 1 shows the quantitative results of DeepDeblurRF and the other methods on the test set of BlurRF-Synth. The table demonstrates that our approach outperforms the existing methods in terms of quality, regardless of the radiance field representation. Moreover, it achieves substantially shorter computation times. Specifically, DeepDeblurRF-P shows the shortest computation time among the approaches based on MLPs and voxel grids, while DeepDeblurRF-G is the fastest among all the methods. This superior performance can be attributed to its DNN-based deblurring modules, which handle blur efficiently and effectively, in contrast to the other methods constrained by linear blur models.

The superior performance of our approach is further demonstrated in Fig. 5, which shows that our method

achieves sharper results on the BlurRF-Synth dataset. Specifically, on the camera motion blur scene, all the other methods produce noisy results, whereas both DeepDeblurRF-P and DeepDeblurRF-G effectively remove the blur even in challenging conditions with noise and pixel saturation, achieving high-quality novel-view synthesis. Similarly, in the defocus blur scene, our methods render sharper novel views by minimizing residual blur than the other methods.

Comparison on real datasets We compare the novel-view synthesis performance of DeepDeblurRF against other approaches using both BlurRF-Real and the real scenes from Deblur-NeRF [24]. As detailed in Sec. 4.2, BlurRF-Real provides an evaluation of how well each method handles camera motion blur with non-linear outliers and noise in real-world low-light conditions. As shown in Tab. 2 and Fig. 6, our method not only outperforms the others in non-reference metrics [1, 26] but also achieves high-quality qualitative

Model	BlurRF-Real		BlurRF-SB		
	NIQE ($\downarrow$)	ARNIQA ($\uparrow$)	PSNR ($\uparrow$)	SSIM ($\uparrow$)	LPIPS ($\downarrow$)
MLP					
Deblur-NeRF [24]	6.341	0.279	25.43	0.7264	0.2319
BAD-NeRF [39]	9.908	0.237	21.97	0.5231	0.4731
DP-NeRF [21]	6.498	0.266	27.07	0.7940	0.1795
Voxel grid					
ExBluRF [22]	7.479	0.279	24.16	0.6649	0.2659
PDRF-10 [31]	6.243	0.266	26.66	0.7730	0.2124
DeepDeblurRF-P	5.829	0.323	29.45	0.8472	0.1677
3D Gaussians					
Deblurring-3DGS [20]	6.086	0.262	23.87	0.6459	0.2475
BAGS [32]	5.837	0.290	24.49	0.6748	0.2391
DeepDeblurRF-G	5.423	0.329	29.59	0.8548	0.1399

Table 2. Quantitative results of novel-view synthesis on BlurRF-Real and BlurRF-SB datasets. We highlight the best metrics and the second best metrics.

results, clearly surpassing the competing methods. It also shows superior performance on real scenes from Deblur-NeRF. Specifically, while other methods struggle with saturated pixels in the camera motion blur scene, our method effectively handles it. Likewise, in the defocus blur scene, it renders sharper novel views without residual blur.

While the real-world scenes of Deblur-NeRF provide blur-free reference images, they suffer from misalignment and exposure differences [24, 31, 32]. Similarly, BlurRF-Real also includes blur-free reference images, but they have not only similar limitations but also severe noise as they are captured under severe low-light conditions. Thus, we report quantitative evaluations against the reference images in the supplementary material. More experimental results are also provided in the supplementary material.

Input views with same blur directions As discussed in Sec. 1, unlike previous approaches, our method does not rely on the assumption that input blurred images contain complementary information from different blur directions. To verify this, we generated an additional synthetic test set, named BlurRF-SB, where the blurred images of the same scene share the same blur directions, while their blur magnitudes are different. BlurRF-SB was generated using five scenes from the BlurRF-Synth test set, following the same generation process.

Fig. 7 shows a qualitative comparison on BlurRF-SB. Due to the lack of complementary information in different training views, the other methods fail to restore sharp details. Conversely, our methods successfully restore sharp details. The quantitative comparison in Tab. 2 also demonstrates that our approach significantly outperforms the others for blurred images with the same blur directions. These results verify the benefit of using prior knowledge on sharp images from pre-trained deblurring networks.

Number of iterations We investigate the impact of the hyperparameter N , which sets the number of iterations in our framework. Tab. 3 shows that DeepDeblurRF-G achieves improved novel-view synthesis as N increases, highlighting

# Iter.	Camera motion			Defocus		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
$N = 1$	28.40	0.8175	0.1707	30.74	0.8862	0.1297
$N = 2$	29.23	0.8462	0.1281	31.22	0.8905	0.0955
$N = 3$	29.65	0.8588	0.1152	32.12	0.9003	0.0899
$N = 4$	29.83	0.8632	0.1108	32.30	0.9038	0.0860
$N = 5$	29.94	0.8681	0.1059	32.58	0.9060	0.0774
$N = 6$	29.97	0.8692	0.1034	32.75	0.9087	0.0736

Table 3. Quantitative results of ablation study on the number of iterations N using DeepDeblurRF-G on the test sets of BlurRF-Synth.



Figure 8. Qualitative results of ablation study on the number of iterations N using DeepDeblurRF-G on the real-world scenes [24].

the effectiveness of our iterative approach. However, gains diminish beyond $N = 5$, suggesting that $N = 5$ is sufficient. Fig. 8 further supports this with qualitative results on real scenes [24] (camera motion blur and defocus blur), showing similar trends. Similar results were observed for DeepDeblurRF-P, as detailed in the supplementary material.

6. Conclusion

In this paper, we proposed DeepDeblurRF, a novel radiance field deblurring approach that leverages DNN-based deblurring to overcome the limitations of the linear blur model of previous methods. To integrate DNN-based deblurring with radiance field construction, we presented a novel RF-guided deblurring scheme and the framework that alternates between RF-guided deblurring and radiance field training. We also presented BlurRF-Synth, the first large-scale dataset for training and evaluating radiance field deblurring frameworks. Experimental results showed that our method outperforms others, handling both camera motion blur and defocus blur effectively with highly reduced computation time.

Limitations and future work Our method leverages deblurring networks to remove blur from training views, but restoring images with more severe blur than in the training data remains challenging. Our RF-guided deblurring uses only rendered images from each input view for guidance. Imposing additional priors on other views and incorporating additional information, e.g., depth maps rendered from radiance fields, may further improve the deblurring performance.

Acknowledgements This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grants funded by the Korea government(MSIT) (RS-2024-00457882, AI Research Hub Project, RS-2019-II191906, Artificial Intelligence Graduate School Program(POSTECH), RS-2024-00437866, ITRC(Information Technology Research Center)).

References

- [1] Lorenzo Agnolucci et al. Arnika: Learning distortion manifold for image quality assessment. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 189–198, 2024. 7
- [2] Miika Aittala and Frédo Durand. Burst image deblurring using permutation invariant convolutional neural networks. In *Proceedings of the European conference on computer vision (ECCV)*, pages 731–747, 2018. 3
- [3] Anpei Chen, Zexiang Xu, Andreas Geiger, Jingyi Yu, and Hao Su. Tensorf: Tensorial radiance fields. In *European Conference on Computer Vision*, pages 333–350. Springer, 2022. 1
- [4] Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. Simple baselines for image restoration. In *European conference on computer vision*, pages 17–33. Springer, 2022. 1, 3, 4
- [5] Wei-Ge Chen, N Nandhakumar, and Worthy N Martin. Image motion estimation from motion smear—a new computational model. *IEEE transactions on pattern analysis and machine intelligence*, 18(4):412–425, 1996. 1
- [6] Sunghyun Cho and Seungyong Lee. Fast motion deblurring. *ACM Trans. Graph.*, 28(5):145:1–145:8, 2009. 1, 3
- [7] Sunghyun Cho, Yasuyuki Matsushita, and Seungyong Lee. Removing non-uniform motion blur from images. In *2007 IEEE 11th International Conference on Computer Vision*, pages 1–8. IEEE, 2007. 1, 3
- [8] Sunghyun Cho, Jue Wang, and Seungyong Lee. Handling outliers in non-blind image deconvolution. In *2011 international conference on computer vision*, pages 495–502. IEEE, 2011. 1
- [9] Sunghyun Cho, Jue Wang, and Seungyong Lee. Video deblurring for hand-held cameras using patch-based synthesis. *ACM Transactions on Graphics (TOG)*, 31(4):1–9, 2012. 3
- [10] Sung-Jin Cho, Seo-Won Ji, Jun-Pyo Hong, Seung-Won Jung, and Sung-Jea Ko. Rethinking coarse-to-fine approach in single image deblurring. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 4641–4650, 2021. 1, 3
- [11] Senyou Deng, Wenqi Ren, Yanyang Yan, Tao Wang, Fenglong Song, and Xiaochun Cao. Multi-scale separable network for ultra-high-definition video deblurring. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 14030–14039, 2021. 3
- [12] Sara Fridovich-Keil, Alex Yu, Matthew Tancik, Qinhong Chen, Benjamin Recht, and Angjoo Kanazawa. Plenoxels: Radiance fields without neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5501–5510, 2022. 1, 2, 4
- [13] Chunzhi Gu, Xuequan Lu, Ying He, and Chao Zhang. Blur removal via blurred-noisy image pair. *IEEE Transactions on Image Processing*, 30:345–359, 2020. 3
- [14] Michael Hirsch, Christian J Schuler, Stefan Harmeling, and Bernhard Schölkopf. Fast removal of non-uniform camera shake. In *2011 International Conference on Computer Vision*, pages 463–470. IEEE, 2011. 1, 3
- [15] Wenbo Hu, Yuling Wang, Lin Ma, Bangbang Yang, Lin Gao, Xiao Liu, and Yuewen Ma. Tri-miprf: Tri-mip representation for efficient anti-aliasing neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 19774–19783, 2023. 1
- [16] Tae Hyun Kim, Kyoung Mu Lee, Bernhard Scholkopf, and Michael Hirsch. Online video deblurring via dynamic temporal blending network. In *Proceedings of the IEEE international conference on computer vision*, pages 4038–4047, 2017. 3
- [17] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4):1–14, 2023. 1, 2, 3, 4
- [18] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 6
- [19] Orest Kupyn, Volodymyr Budzan, Mykola Mykhailych, Dmytro Mishkin, and Jiří Matas. Deblurgan: Blind motion deblurring using conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 8183–8192, 2018. 1, 3
- [20] Byeonghyeon Lee, Howoong Lee, Xiangyu Sun, Usman Ali, and Eunbyung Park. Deblurring 3d gaussian splatting, 2024. 1, 2, 3, 6, 7, 8
- [21] Dogyoon Lee, Minhyeok Lee, Chajin Shin, and Sangyoun Lee. Dp-nerf: Deblurred neural radiance field with physical scene priors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12386–12396, 2023. 1, 2, 3, 6, 8
- [22] Dongwoo Lee, Jeongtaek Oh, Jaesung Rim, Sunghyun Cho, and Kyoung Mu Lee. Exblurf: Efficient radiance fields for extreme motion blurred images. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 17639–17648, 2023. 1, 2, 3, 5, 6, 7, 8
- [23] Ilya Loshchilov and Frank Hutter. Sgdr: Stochastic gradient descent with warm restarts. *arXiv preprint arXiv:1608.03983*, 2016. 6
- [24] Li Ma, Xiaoyu Li, Jing Liao, Qi Zhang, Xuan Wang, Jue Wang, and Pedro V Sander. Deblur-nerf: Neural radiance fields from blurry images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12861–12870, 2022. 1, 2, 3, 5, 6, 7, 8
- [25] B Mildenhall, PP Srinivasan, M Tancik, JT Barron, R Ramamoorthi, and R Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. In *European conference on computer vision*, 2020. 1
- [26] Anish Mittal, Rajiv Soundararajan, and Alan C Bovik. Making a "completely blind" image quality analyzer. *IEEE Signal Processing Letters*, 20(3):209–212, 2012. 7

- [27] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Trans. Graph.*, 41(4):102:1–102:15, 2022. 1
- [28] Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee. Deep multi-scale convolutional neural network for dynamic scene deblurring. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3883–3891, 2017. 1, 3
- [29] Seungjun Nah, Sanghyun Son, and Kyoung Mu Lee. Recurrent neural networks with intra-frame iterations for video deblurring. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 8102–8111, 2019. 3
- [30] Jinshan Pan, Zhe Hu, Zhixun Su, and Ming-Hsuan Yang. Deblurring text images via l0-regularized intensity and gradient prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2901–2908, 2014. 1, 3
- [31] Cheng Peng and Rama Chellappa. Pdrf: progressively deblurring radiance field for fast scene reconstruction from blurry images. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 2029–2037, 2023. 1, 2, 3, 6, 7, 8
- [32] Cheng Peng, Yutao Tang, Yifan Zhou, Nengyu Wang, Xijun Liu, Deming Li, and Rama Chellappa. Bags: Blur agnostic gaussian splatting through multi-scale kernel modeling, 2024. 1, 2, 3, 6, 7, 8
- [33] Alex Rav-Acha and Shmuel Peleg. Two motion-blurred images are better than one. *Pattern recognition letters*, 26(3): 311–317, 2005. 1, 3
- [34] Jaesung Rim, Haeyun Lee, Jucheol Won, and Sunghyun Cho. Real-world blur dataset for learning and benchmarking deblurring algorithms. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXV 16*, pages 184–201. Springer, 2020. 1
- [35] Jaesung Rim, Geonung Kim, Jungeon Kim, Junyong Lee, Seungyong Lee, and Sunghyun Cho. Realistic blur synthesis for learning image deblurring. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2022. 1, 5
- [36] Johannes Lutz Schönberger and Jan-Michael Frahm. Structure-from-motion revisited. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016. 4, 6
- [37] Shuochen Su, Mauricio Delbracio, Jue Wang, Guillermo Sapiro, Wolfgang Heidrich, and Oliver Wang. Deep video deblurring for hand-held cameras. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1279–1288, 2017. 3
- [38] Xin Tao, Hongyun Gao, Xiaoyong Shen, Jue Wang, and Jiaya Jia. Scale-recurrent network for deep image deblurring. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 8174–8182, 2018. 1, 3
- [39] Peng Wang, Lingzhe Zhao, Ruijie Ma, and Peidong Liu. Bad-nerf: Bundle adjusted deblur neural radiance fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4170–4179, 2023. 1, 2, 3, 6, 7, 8
- [40] Oliver Whyte, Josef Sivic, Andrew Zisserman, and Jean Ponce. Non-uniform deblurring for shaken images. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 491–498, 2010. 1, 3
- [41] Patrick Wieschollek, Michael Hirsch, Bernhard Scholkopf, and Hendrik Lensch. Learning blind motion deblurring. In *Proceedings of the IEEE international conference on computer vision*, pages 231–240, 2017. 3
- [42] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Multi-stage progressive image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 14821–14831, 2021. 1, 3
- [43] Haichao Zhang and David Wipf. Non-uniform camera shake removal using a spatially-adaptive sparse penalty. *Advances in Neural Information Processing Systems*, 26, 2013. 1, 3
- [44] Shangchen Zhou, Jiawei Zhang, Jinshan Pan, Haozhe Xie, Wangmeng Zuo, and Jimmy Ren. Spatio-temporal filter adaptive network for video deblurring. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 2482–2491, 2019. 3
- [45] Shangchen Zhou, Jiawei Zhang, Wangmeng Zuo, Haozhe Xie, Jinshan Pan, and Jimmy S Ren. Davanet: Stereo deblurring with view aggregation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10996–11005, 2019. 3