EvaGaussians: Event Stream Assisted Gaussian Splatting from Blurry Images

Supplementary Material

Contents

Please check our supplementary **project page** for video results.

A. Dataset Details

A.1. EvaGaussians-Blender Dataset

A.1.1. Dataset Overview

We use Blender [3] to craft nine indoor and outdoor 3D scenes, and further incorporate four 3D scenes from DeblurNeRF [7] and six 3D objects from the NeRF-Synthetic dataset [8] as our base scenes. We then design various camera trajectories to simulate motion-blurred images on these base scenes, and generate the corresponding event streams using V2E [6]. Visualization of our crafted scenes are shown in Figure. 1 and Figure. 3, and an overview of the design purposes of these crafted scenes is provided below:

Indoor Scenes

- *Classroom*: A typical classroom setting featuring desks, chairs, a blackboard, and educational posters. This scene is designed to simulate an academic environment, ideal for educational and surveillance applications.
- *Café*: A cozy café with tables, chairs, a counter, and various decorations. This scene mimics a social setting, providing a dynamic backdrop for testing social interaction algorithms and retail analytics.
- *Dormitory*: A student dormitory room equipped with beds, study desks, personal belongings, and typical dorm furniture. This scene represents a personal living space, useful for smart home and security applications.

Outdoor Scenes

- *Desert*: A vast, arid landscape with sand dunes and sparse vegetation. This scene is perfect for testing navigation and object detection in harsh, unstructured environments.
- *City Blocks*: Urban scenes featuring streets, buildings, vehicles, and pedestrians. This environment is essential for autonomous driving, urban planning, and smart city applications.
- *Lake*: A serene natural setting with dense forests surrounding a tranquil lake. This scene provides a complex environment for testing outdoor navigation, environmental monitoring, and wildlife tracking.
- *Forests*: A rugged terrain with forested areas and scattered boulders. This scene is useful for off-road navigation and geological survey applications.
- *Venice*: A picturesque representation of Venice with canals, bridges, and historic architecture. This scene offers a unique setting for cultural heritage preservation, tourism, and urban analytics.
- *London*: A bustling cityscape of London with iconic landmarks, streets, and a dynamic urban environment. This scene supports applications in tourism, traffic management, and city modeling.

A.1.2. Camera Settings

To render the base scenes and simulate motion blur, we configure the virtual camera in Blender with a resolution of 400×600, and set the scaling factor to 1.0. The virtual camera utilized a perspective model with a shutter speed of $1/180$ seconds. Subsequently, we developed a dedicated script to generate camera trajectory and motion blur. An example is shown in Figure. 4. Along each predefined virtual camera trajectory, we uniformly sampled 35 camera poses, adding a certain level of jitter to create the training set. We recorded the start and end time of the camera exposure time, the positions, and 20 intermediate frames during the exposure time (obtained through linear interpolation between the start and end positions). We then uniformly sample 100 camera poses along the same trajectory to form the test set. Using the event camera simulator from V2E [6], we simulate the event stream for each camera trajectory and synthesize the event bins from the event stream at the start and end of the exposure time.

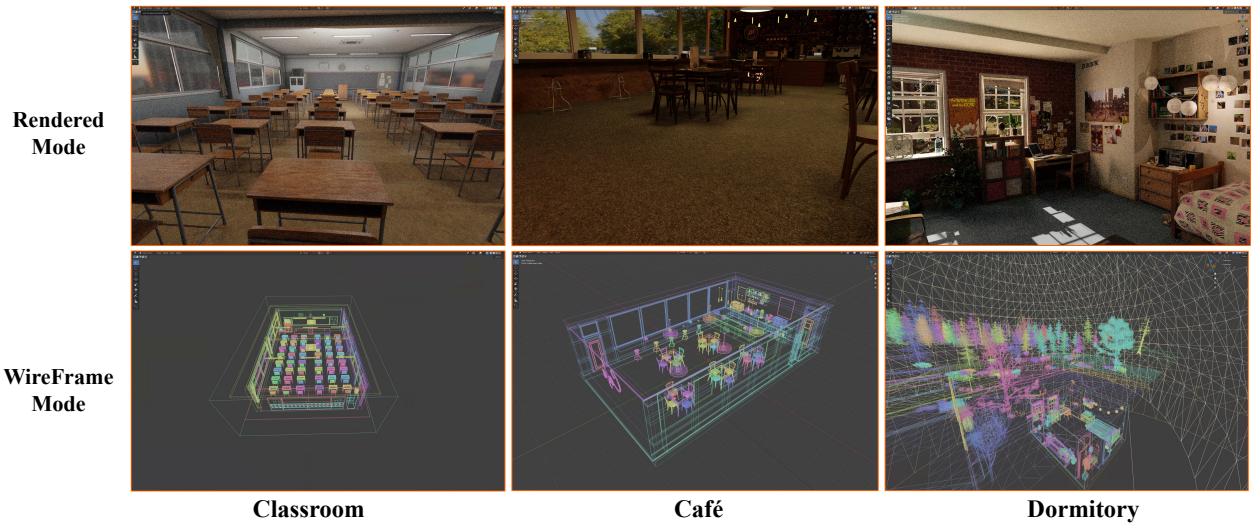


Figure 1. **Visualization of EvaGaussians-Blender Indoor Scenes.** The sizes of the *Café* and *Classroom* scenes are approximately $15 \times 7 \times 4$ meters, while the *Dormitory* scene is approximately $5 \times 5 \times 4$ meters (with an additional outdoor garden, making the overall scene size $20 \times 20 \times 6$ meters).

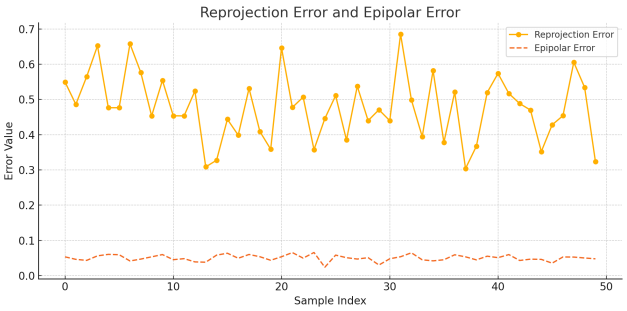


Figure 2. **Visualization of Reprojection Errors and Epipolar Errors.** The figure illustrates the 50 sets of reprojection errors and epipolar errors generated during the calibration process. The reprojection error e_r represents the average discrepancy between the observed points and the projected points, calculated as shown in Eq. 1. The epipolar error e_{epipolar} represents the average distance between points in one camera and the epipolar lines calculated from the other camera for each pair of images, calculated as shown in Eq. 2. As shown in the figure, the average reprojection error is approximately 0.5, and the average epipolar error is approximately 0.05, indicating a high level of accuracy in the calibration process.

A.2. EvaGaussians-DAVIS Dataset

We use the color DAVIS346 event camera [1] to record our real-world event and RGB sequences and utilize the default camera settings provided in the DV software that comes with the camera. We name the five captured scenes as *desk & chair*, *washroom*, *pokémon*, *pillow*, and *bag*.

Table 1. **Comparison with existing event-assisted deblurring reconstruction synthetic dataset.** The existing datasets suffer from limited data size and scene diversity, and none of them includes large-scale scene for evaluation.

Dataset	Scene				Event	
	Object	Medium	**Large**	Total	Event Map	Event Stream
E2NeRF[10]	7	0	**0**	7	✓	✗
EvDNeRF[2]	0	4	**0**	4	✗	✓
Ours	6	7	**6**	19	✓	✓

Table 2. **Comparison with existing event-assisted deblurring reconstruction real-world dataset.** The scenes in E2NeRF and EvDNeRF are captured within a limited area and more focus on the objects with a specific background.

Dataset	Scene			Camera Motion
	Object	**Medium**	Total	
E2NeRF[10]	5	**0**	5	*Diverse*
EvDNeRF[2]	5	**0**	5	*Naïve panning*
Ours	3	**2**	5	*Diverse*

A.2.1. Camera Calibration

We calibrated the event camera using the DV software provided by DAVIS. During the calibration process, we used a 6×9 checkerboard pattern with a square size of 30 mm. In the software configuration, we set the width to 9, height to 6, and square size to 30 mm. We then ran the calibration module and moved the calibration pattern in front of the camera. The software detected the pattern and collected images, highlighting the detected area in green, as shown in Figure. 5. We set the minimum detections parameter to 50 to ensure a sufficient number of samples and used the con-

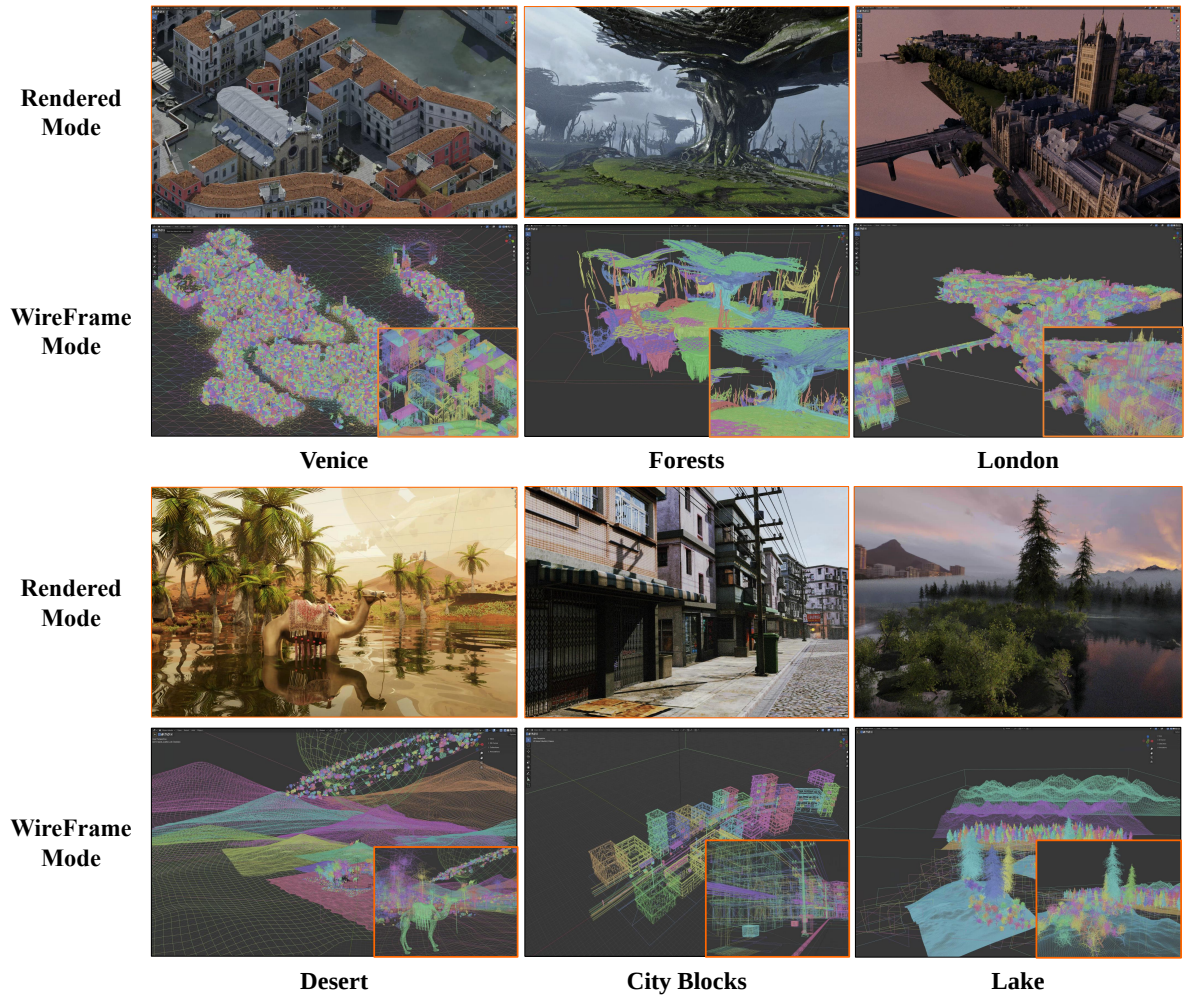


Figure 3. **Visualization of EvaGaussians-Blender Outdoor Scenes.** These scenes include rich details ane diverse components like sky, lake, river, desert, forest, cities, roads. All scenes cover an area of more than 1 square kilometer.

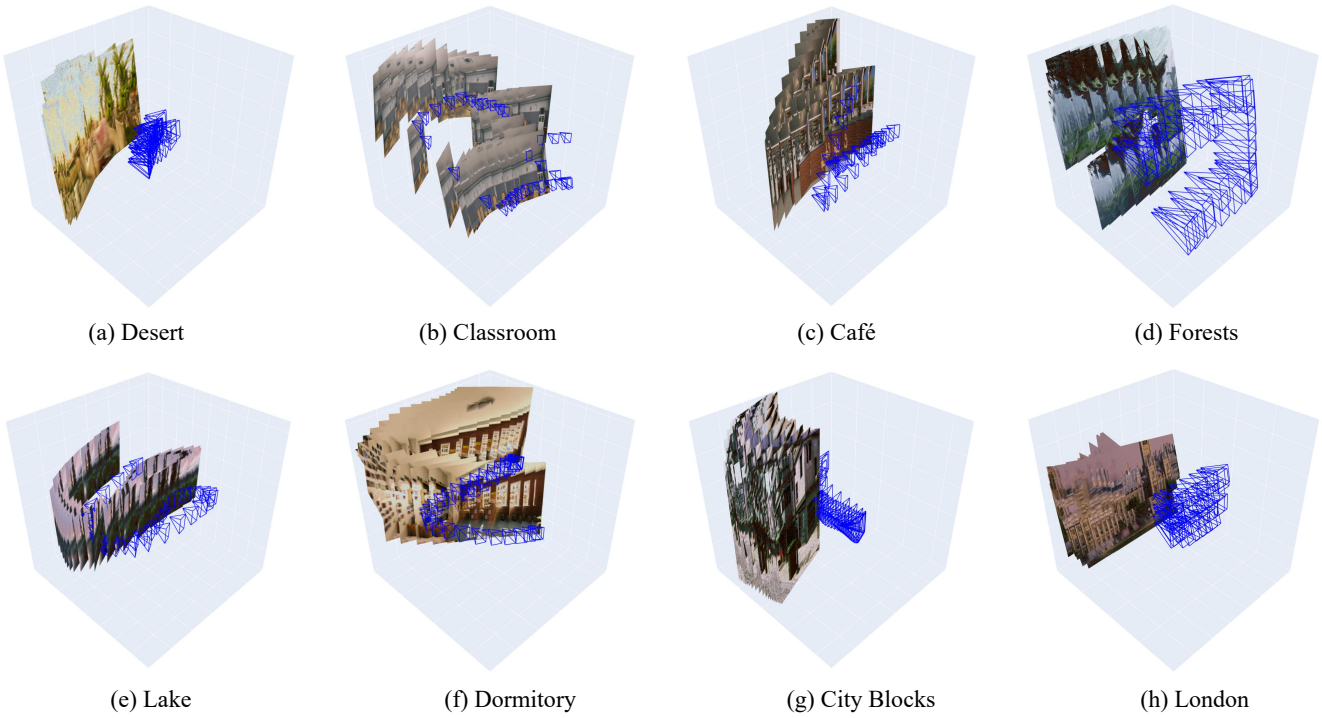
secutive detections parameter to ensure consistent pattern detection. Additionally, we enabled the image verification option to check the collected images in real-time, discarding inaccurately detected images and replacing them with new ones. We evaluate the calibration accuracy using the reprojection error e_r as Eq. 1 and the epipolar error e_{epipolar} as Eq. 2 in stereo calibration. The reprojection error is calculated as follows:

$$e_r = \frac{1}{n}\sum_{i=1}^{n} \|x_i - \hat{x}_i\| \tag{1}$$

where x_i represents the observed points and $\hat{x}_i$ represents the projected points. The epipolar error is calculated as the average epipolar error for each point in all collected images. For each pair of images, the error is calculated as the sum of the distances between the points in one camera and the epipolar lines calculated from the other camera (m is the number of acquired images, n is the number of points). The formula is as follows:

$$e_{\text{epipolar}} = \frac{1}{m \times n}\sum_{i=1}^{m}\sum_{j=1}^{n} [d(P1_{i,j}, l_{2,i,j}) + d(P2_{i,j}, l_{1,i,j})] \tag{2}$$

where $P1_{i,j}$ and $P2_{i,j}$ are the projection points of the jth point in the ith image in two cameras, and $l_{1,i,j}$ and $l_{2,i,j}$ are the epipolar lines corresponding to the jth point in the ith image calculated from the other camera. The maximum allowable error can be set under *Max Reprojection Error*. The stereo calibration also calculates the error caused by the epipolar constraint, which can be set under *Max Epipolar Error*. Once the calibration is successful, the results are



(a) Desert (b) Classroom (c) Café (d) Forests

(e) Lake (f) Dormitory (g) City Blocks (h) London

Figure 4. **Visualization of Camera Trajectory.** The trajectories depicted were manually configured within Blender [3] to ensure precise control over the camera paths. For the purpose of visualization, these trajectories have been normalized.

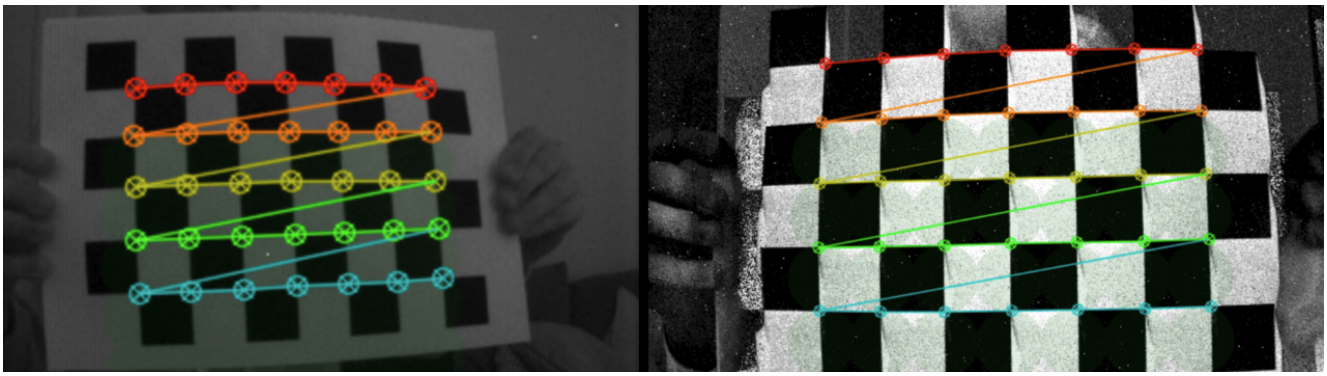


Figure 5. **Illustration of Camera Calibration.** The left panel shows the checkerboard pattern captured from various positions and angles, with detected corner points utilized for calibration. The right panel presents the calibrated checkerboard pattern, demonstrating the corresponding points and lines between two cameras, which reflect the geometric relationship and accuracy achieved after calibration. Different colored lines indicate the correspondences between points during the calibration process.

saved and the undistorted output is displayed. The visualization results of the two types of errors are shown in Figure. 2. This process ensures the accuracy of the calibration, thereby improving the measurement accuracy and stability in subsequent applications.

A.2.2. Camera Settings

We recorded the five real scenes using the calibration parameters obtained during the calibration process. By adjusting the indoor lighting and shooting angles, we ensured the richness of the recorded scene details. The adopted event camera has a spatial resolution of 346×260, a temporal resolution of 1 μs, a typical latency of less than 1 ms, a maximum throughput of 12 MEps, and a dynamic range of

Table 3. **Comparison with EvDNeRF [2] on EvDNeRF dataset and DSEC dataset.** It demonstrates the robustness of our method to different datasets.

	EvDNeRF [2] dataset		DSEC [5] dataset	
Metric	EvDNeRF	Ours	EvDNeRF	Ours
PSNR↑	31.75	**32.84**	31.89	**32.08**
SSIM↑	0.910	**0.922**	0.911	**0.915**
LPIPS↓	0.066	**0.064**	0.068	**0.064**

Table 4. **Comparison of the poses directly estimated from the blurry images with those estimated from the EDI-deblurred images.** Using COLMAP directly on the blurry images fails in 10 scenes and succeeds in only 9 scenes. For the successful cases, the ATE are presented below.

ATE ↓	Medium-scale	Large-scale
Poses (Blurry Images)	0.054 ± 0.014	0.086 ± 0.275
Poses (EDI-deblurred)	$\mathbf{0.046 \pm 0.013}$	$\mathbf{0.049 \pm 0.011}$

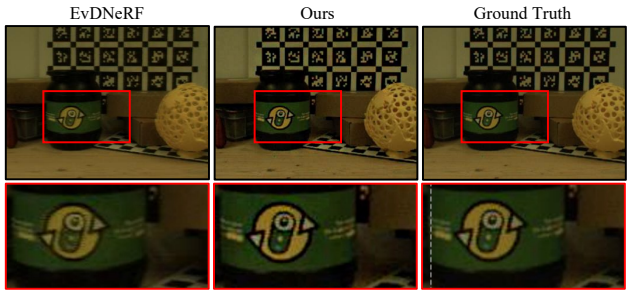


Figure 6. **Qualitative comparison with EvDNeRF [2] on the EvDNeRF dataset.** It demonstrates the robustness of our method to other datasets.

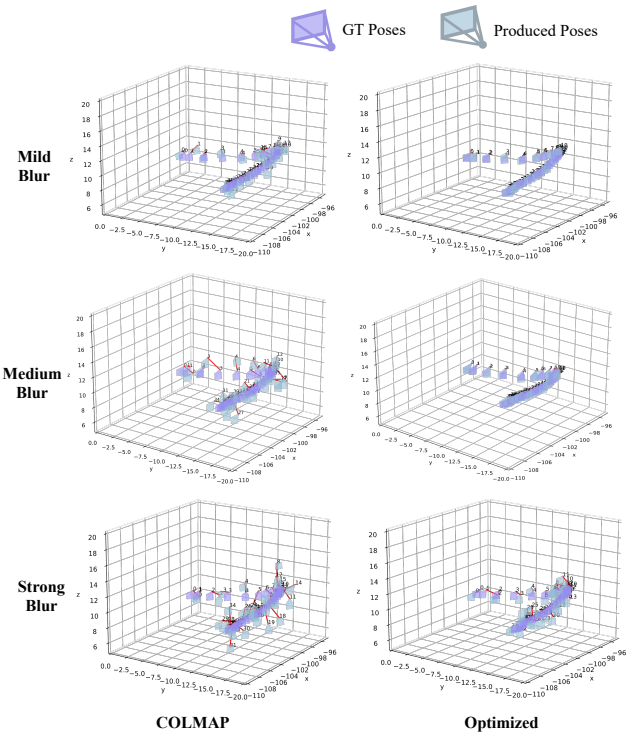


Figure 7. **Visualization of pose accuracy in different level of motion blur.** Images with mild blur are captured at $0.5\times$ camera motion speed, images with medium blur are captured at the default motion speed, and images with strong blur are captured at $2.0\times$ motion speed.

approximately 120 dB (with 50% of the pixels responding to 80% contrast changes under 0.1-100k lux conditions). The contrast sensitivity is 14.3% (ON) and 22.5% (OFF) (with 50% of the pixels responding). These parameters ensure that the event camera can stably and efficiently record scene information under various lighting conditions and dynamic ranges.

A.3. Comparison with Existing Datasets

The aim of contributing this dataset is to establish a more comprehensive evaluation benchmark for event-assisted deblurring reconstruction tasks. We observed that although E2NeRF [10] and EvDNeRF [2] have introduced similar datasets, their datasets suffer from limited scene and camera motion diversity, as well as insufficient data size. As shown in Table. 1 and Table. 2 , our proposed dataset includes more diverse scenes with various scales, making it better suited for a comprehensive evaluation.

B. Additional Ablations

B.1. Analysis on EDI Pre-processing

In our method, we use EDI [9] to pre-process the blurry RGB images, followed by COLMAP [11] to obtain the initial point cloud for 3D-GS optimization and the initial camera poses for camera trajectory optimization. Examples of the input blurry images and the EDI pre-process results are shown in Figure. 8. It shows that the EDI-deblur process produces sharper textures and clearer geometries than the original blurry input. However, as demonstrated by the 'EDI-GS' results in Table.1 of the main text, directly using them to train a 3D-GS model produces suboptimal results. This is because the images still exhibit poor texture details, and the deblurring process introduces inconsistencies in the images. For pose estimation, we compare the poses directly estimated from the blurry images with those estimated from the EDI-deblurred images on the **EvaGaussians-synthetic** dataset, which contains 19 scenes. Using COLMAP directly on the blurry images fails in 10 scenes and succeeds in only 9 scenes. For the successful cases, we report

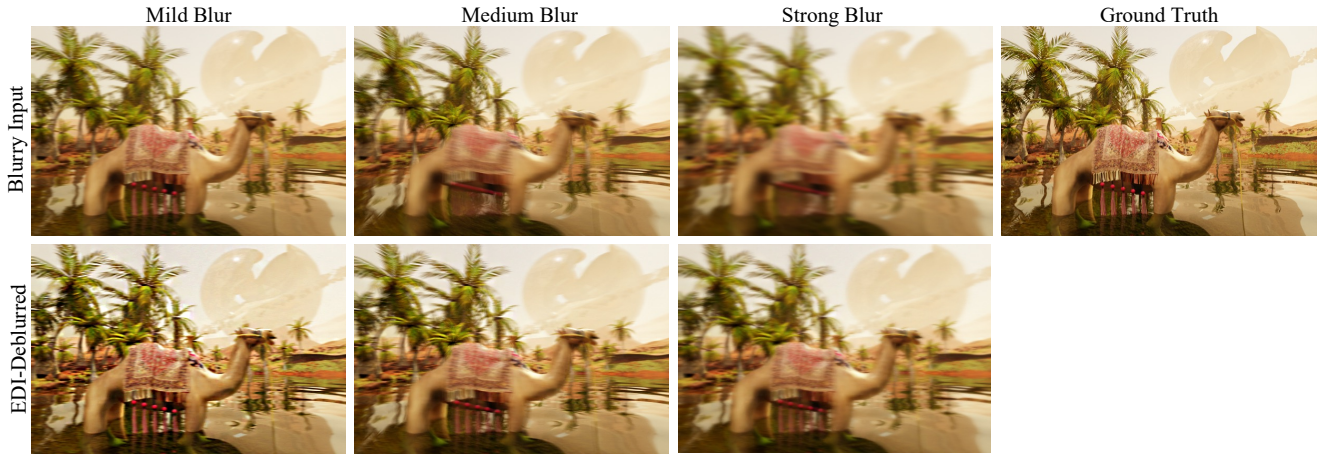


Figure 8. **Examples of blurry images and EDI pre-process results.** The EDI-deblur process produces sharper textures and clearer geometries than the original blurry input. However, as demonstrated by the 'EDI-GS' results in Table.1 of the main text, directly using them to train a 3D-GS model produces suboptimal results, because the images still exhibit poor texture details, and the deblurring process introduces inconsistencies in the images.

the ATE (Average Trajectory Error) on the large-scale and medium-scale scenes in Table 4.

B.2. Robustness to Different Datasets.

To validate the robustness of our method across different datasets, we perform a novel view synthesis comparison on another two real-world datasets: 1) the dataset provided by EvDNeRF [2], denoted as EvDNeRF dataset; 2) the *interlaken-00-f* scene from the DSEC [5] driving video dataset, which contains static clips for novel view synthesis evaluation. The results are reported in Table. 3 and Figure. 6. It demonstrates the robustness of our method to different datasets.

B.3. Robustness of Pose Optimization

We present an ablation on the robustness of pose optimization in different blur level. we set up three different camera speeds in the *City Blocks* scene of the synthetic dataset to obtain images with varying levels of blur. Images with mild blur are captured at $0.5\times$ camera motion speed, images with medium blur are captured at the default motion speed, and images with strong blur are captured at $2.0\times$ motion speed. We then compare the ATE (Average Trajectory Error between the initial EDI-produced COLMAP poses and the final poses optimized by our method. As shown in Figure. 7, with the motion blur becoming more severe, the accuracy of the COLMAP poses is significantly impacted, while the optimized poses maintain a higher level of accuracy. In a horizontal comparison, the optimized poses better match the ground truth across various levels of blur, demonstrating the effectiveness of pose optimization.

C. Detailed Comparisons

C.1. Synthetic Data Comparison

We present a detailed per-scene analysis of the novel view synthesis performance in medium and large scale scenes from the **EvaGaussians-Blender** dataset, to evaluate the effectiveness of our method across different challenging scenes.

Table. 5 shows the PSNR value of the NVS results, which demonstrates that our proposed method consistently outperforms other approaches across various scenes. The detailed metrics for SSIM and LPIPS in Table. 6 and Table. 7 further show that our model excels in maintaining structural integrity and perceptual quality in synthesized views. Specifically, in medium-scale scenes, our method exhibits robust performance, particularly in complex environments where maintaining detail and minimizing artifacts are challenging. This can be demonstrated in scenes such as *Cozyroom* and *Factory*, where our method achieves significant improvements in both PSNR and SSIM. For large-scale scenes, scenes like *Desert* and *City Blocks* highlight the model's capability to generalize across different scales and provide high-quality novel view synthesis.

We also present more qualitative results of novel view synthesis in Figure. 11 and Figure. 12. The results highlight our model's ability to reconstruct fine details and maintain high color accuracy beyond the comparison baselines.

C.2. Real-world Data Comparison

In this section, we present a comprehensive per-scene analysis of novel view synthesis results on the **EvaGaussians-DAVIS** dataset. As introudced in the main text, the adopted

Table 5. **The Novel View Synthesis Results of PSNR ↑ in the EvaGaussians-Blender Dataset**. The highest values in each category are highlighted in **bold** to indicate the best results.

Models	Medium							Large				
	Classroom	Dormitory	Café	Pool	Cozyroom	Factory	Tanabata	Desert	City Blocks	London	Forests	Lake
Blurry-NeRF	25.42	26.72	21.23	26.66	15.42	26.76	26.38	19.04	19.34	20.00	23.60	24.67
Blurry-GS	25.59	26.83	21.34	26.17	20.76	26.48	26.46	19.67	19.48	20.18	23.75	24.34
UFP-GS	28.64	29.68	21.54	26.79	21.62	27.60	28.77	20.78	20.85	16.61	22.96	25.60
EDI-GS	28.49	29.90	22.62	26.36	21.17	29.06	27.46	20.66	20.68	20.27	23.97	25.99
EFNET-GS	28.59	29.48	22.04	26.36	21.42	27.57	27.43	20.62	21.70	20.36	25.57	25.18
E^2NeRF	28.87	30.77	26.16	28.92	21.23	29.85	28.65	20.02	21.78	21.30	25.90	25.78
BAD-NeRF	30.28	31.23	27.18	28.71	21.68	30.28	29.85	21.09	21.91	22.91	26.35	26.99
BAD-Gaussians	30.28	31.23	27.18	28.72	21.68	30.28	29.85	21.10	21.93	22.93	26.35	26.98
BAGS	30.93	31.40	28.23	29.12	21.88	30.88	30.14	22.20	24.22	23.83	26.80	27.58
EvDNeRF	31.83	28.95	28.66	29.69	22.01	31.20	30.02	21.62	22.23	23.88	27.10	28.29
Ours	**34.38**	**32.97**	**30.41**	**30.26**	**22.71**	**31.85**	**30.71**	**24.88**	**23.71**	**23.99**	**27.62**	**29.90**

Table 6. **The Novel View Synthesis of SSIM ↑ in the EvaGaussians-Blender Dataset.** The highest values in each category are highlighted in **bold** to indicate the best results.

Models	Medium							Large				
	Classroom	Dormitory	Café	Pool	Cozyroom	Factory	Tanabata	Desert	City Blocks	London	Forests	Lake
Blurry-NeRF	0.7086	0.8281	0.5682	0.7442	0.5098	0.8567	0.8057	0.6023	0.6325	0.6732	0.7002	0.7823
Blurry-GS	0.7154	0.8312	0.5638	0.7265	0.7632	0.8519	0.8064	0.6386	0.6341	0.6819	0.7026	0.7807
UFP-GS	0.8701	0.9281	0.5706	0.7527	0.7729	0.8640	0.8569	0.6172	0.6144	0.5807	0.6911	0.7968
EDI-GS	0.8456	0.9291	0.5977	0.7288	0.7703	0.8963	0.8408	0.6567	0.5655	0.6882	0.7158	0.8012
EFNET-GS	0.8692	0.9259	0.5841	0.7288	0.7797	0.8643	0.8345	0.6164	0.6366	0.6069	0.7628	0.7901
E^2NeRF	0.8723	0.9319	0.7869	0.8795	0.7724	0.9245	0.8915	0.6169	0.6822	0.6734	0.7467	0.8139
BAD-NeRF	0.8978	0.9337	0.8041	0.8794	0.7764	0.9353	0.9271	0.6314	0.6867	0.6932	0.7583	0.8919
BAD-Gaussians	0.8992	0.9351	0.8031	0.8781	0.7755	0.9345	0.9266	0.6346	0.6851	0.6944	0.7578	0.8908
BAGS	0.9162	0.9471	0.8191	0.8931	0.7925	0.9465	0.9446	0.6496	0.7011	0.7074	0.7758	0.9048
EvDNeRF	0.9023	0.8935	0.8513	0.8885	0.7854	0.9454	0.9311	0.6589	0.7023	**0.7159**	0.7726	0.9129
Ours	**0.9402**	**0.9580**	**0.9033**	**0.9108**	**0.8052**	**0.9592**	**0.9382**	**0.8152**	**0.7405**	0.7145	**0.8152**	**0.9465**

metrics include BRISQUE, NIQE, PIQE, MetaIQA, and RankIQA, which can effectively assess the quality of synthesized views in a no-reference manner. As shown in Table. 8, our model achieves the best BRISQUE scores across all scenes, highlighting its ability to produce visually appealing and less distorted images. For NIQE, as presented in Table. 9, our approach significantly outperforms the baselines, achieving the lowest average NIQE score. This demonstrates our method's robustness in generating high-quality images with minimal perceptual artifacts. In terms of PIQE, Table. 10 shows that our model again leads in performance, achieving the lowest PIQE scores, which underscores the effectiveness of our model in preserving image details and reducing noise. Furthermore, our method excels in MetaIQA and RankIQA evaluations, as detailed in Tables. 11 and Table. 12, respectively. The highest MetaIQA scores and lowest RankIQA scores across most scenes affirm the overall better visual quality and fidelity of our synthesized views compared to baseline models. Overall, these results demonstrate the robustness of our method, particularly in handling complex scenes and maintaining high visual quality across diverse scenarios.

D. Broader Impacts

Our proposed **EvaGaussians** leverages event cameras to assist novel view synthesis from low-quality, blurred images. It has the potential to bring about both positive and negative societal impacts.

On the positive side, our method can improve the efficiency of surveillance systems by reconstructing clear 3D images from low-quality footage, enabling better identification of individuals and objects in challenging conditions.

Table 7. **The Novel View Synthesis of LPIPS ↓ in the EvaGaussians-Blender Dataset.** The highest values in each category are highlighted in **bold** to indicate the best results.

Models	Medium							Large				
	Classroom	Dormitory	Café	Pool	Cozyroom	Factory	Tanabata	Desert	City Blocks	London	Forests	Lake
Blurry-NeRF	0.3987	0.2998	0.4528	0.3848	0.5545	0.2063	0.2348	0.4447	0.5273	0.3971	0.3425	0.4127
Blurry-GS	0.3824	0.2873	0.4554	0.4198	0.2432	0.2091	0.2335	0.4231	0.4116	0.3925	0.3379	0.4206
UFP-GS	0.2838	0.1361	0.4511	0.3646	0.2331	0.1928	0.1856	0.3816	0.3342	0.4069	0.3383	0.4069
EDI-GS	0.2914	0.1267	0.4038	0.4102	0.2388	0.1446	0.1911	0.4062	0.3857	0.3901	0.3295	0.4001
EFNET-GS	0.2846	0.1373	0.4529	0.4102	0.2345	0.1926	0.1958	0.3912	0.3222	0.3932	0.2962	0.4128
E^2NeRF	0.2821	0.1165	0.2871	0.2048	0.2369	0.1073	0.1546	0.3938	0.3713	0.3928	0.3589	0.3588
BAD-NeRF	0.2365	0.1083	0.2715	0.2103	0.2278	0.0992	0.1224	0.3947	0.3618	0.3224	0.3467	0.3142
BAD-Gaussians	0.2384	0.1078	0.2695	0.2094	0.2262	0.0985	0.1215	0.3965	0.3604	0.3215	0.3452	0.3129
BAGS	0.2334	0.1076	0.2495	0.1994	0.2112	0.0985	0.1201	0.3765	0.3354	0.3015	0.3252	0.2979
EvDNeRF	0.2217	0.1455	0.2447	0.1926	0.1942	0.0768	0.1086	0.3823	0.3497	**0.2972**	0.3163	0.2941
Ours	**0.1927**	**0.0928**	**0.2311**	**0.1859**	**0.1873**	**0.0715**	**0.1023**	**0.2053**	**0.2835**	0.2983	**0.2857**	**0.2674**

Table 8. The **Novel View Synthesis** of BRISQUE in the **EvaGaussians-DAVIS** Dataset. The highest values in each category are highlighted in **bold** to indicate the best results.

Models	BRISQUE ↓					
	Desk & Chair	Washroom	Pokémon	Pillow	Bag	Average
B-NeRF	63.9428	102.7828	109.2711	97.4778	87.7699	92.2489
B-3DGS	51.1542	82.2262	87.4169	77.9823	70.2159	73.7991
UFP-GS	43.6932	69.7354	74.6684	66.6098	59.9711	62.9356
EDI-GS	43.4811	69.9923	74.3044	66.2849	59.6835	62.7492
EFN-GS	43.5235	69.8473	74.2875	66.3158	59.6523	62.7253
E^2NeRF	36.2148	64.4332	77.0112	67.1324	62.8063	61.5196
BAD-NeRF	42.6285	68.5219	72.8474	64.9852	58.5133	61.4993
BAD-GS	42.2065	67.8434	72.1261	64.3418	57.9339	60.8903
BAGS	36.8725	61.4523	71.3098	62.8746	60.3951	58.5809
EvDNeRF	34.5687	61.5044	73.5012	64.0809	59.4969	58.6304
Ours	**32.9225**	**56.5756**	**70.0211**	**61.0294**	**58.1876**	**53.9623**

This can bolster public safety and aid in criminal investigations. Additionally, the ability to reconstruct scenes from blurred inputs can enhance the performance of autonomous vehicles, drones, and robots, enabling them to navigate more accurately in poor visibility conditions, leading to safer and more efficient transportation and logistics. In situations where traditional cameras may struggle to capture clear images under extreme conditions, our method can provide valuable information for first responders and rescue teams, helping them make informed decisions and potentially saving lives. Furthermore, our technique can be applied to medical imaging, allowing for better visualization of internal structures and more accurate diagnoses, ultimately leading to improved patient outcomes.

On the negative side, the enhanced surveillance capabilities enabled by our method may raise privacy concerns. For example, our method could be used for malicious purposes, such as stalking or spying on individuals without their consent. It is important to establish regulations and guidelines to prevent such misuse.

Table 9. The **Novel View Synthesis** of NIQE in the **EvaGaussians-DAVIS** Dataset.

Models	NIQE ↓					
	Desk & Chair	**Washroom**	**Pokémon**	**Pillow**	**Bag**	**Average**
B-NeRF	11.4380	15.5386	16.9153	16.2316	14.8844	15.0016
B-3DGS	9.1504	12.4309	13.5323	12.9853	11.9075	12.0113
UFP-GS	7.7736	10.6170	11.3827	10.9266	10.1715	10.1743
EDI-GS	7.7778	10.5662	11.5024	11.0375	10.1214	10.2011
EFN-GS	7.8235	10.5412	11.4868	11.0743	10.1288	10.2109
E^2NeRF	6.8907	9.0924	11.6383	10.1830	9.3954	9.4400
BAD-NeRF	7.6253	10.3590	11.2769	10.8211	9.9229	10.0011
BAD-GS	7.5498	10.2565	11.1652	10.7139	9.8247	9.9020
BAGS	7.0145	9.0372	11.2487	9.8124	9.3176	9.2860
EvDNeRF	6.5775	8.6791	11.1092	9.7201	8.9684	9.0109
Ours	**6.1898**	**8.1117**	**10.2662**	**8.9455**	**8.3421**	**8.3711**

Table 10. The **Novel View Synthesis** of PIQE in the **EvaGaussians-DAVIS** Dataset.

Models	PIQE ↓					
	Desk & Chair	**Washroom**	**Pokémon**	**Pillow**	**Bag**	**Average**
B-NeRF	58.4566	68.8701	73.4701	54.0862	74.7257	65.9217
B-3DGS	46.7653	55.0961	58.7761	43.2689	59.7806	52.7374
UFP-GS	39.9453	47.0612	50.2026	36.9659	50.9631	45.0276
EDI-GS	39.7505	46.8317	49.9597	36.7786	50.8135	44.8268
EFN-GS	39.7748	46.7951	49.9346	36.7891	50.8275	44.8242
E^2NeRF	40.9825	49.2988	50.4868	39.7896	53.2203	46.7556
BAD-NeRF	38.9711	45.9134	48.9801	36.0574	49.8171	43.9478
BAD-GS	38.5852	45.4588	48.4951	35.7004	49.3239	43.5127
BAGS	38.5789	45.9635	46.5281	46.5281	48.7834	45.2744
EvDNeRF	39.1197	47.0580	48.1919	37.9810	50.8012	44.6304
Ours	**36.5717**	**44.3304**	**44.0641**	**35.3059**	**47.3749**	**41.5294**

Table 11. The **Novel View Synthesis** of MetaIQA in the **EvaGaussians-DAVIS** Dataset.

Models	MetaIQA ↑					
	Desk & Chair	**Washroom**	**Pokémon**	**Pillow**	**Bag**	**Average**
B-NeRF	0.1419	0.1272	0.1085	0.1122	0.1307	0.1241
B-3DGS	0.1621	0.1454	0.1240	0.1283	0.1494	0.1418
UFP-GS	0.1976	0.1774	0.1521	0.1564	0.1823	0.1732
EDI-GS	0.1986	0.1780	0.1518	0.1571	0.1830	0.1737
EFN-GS	0.1928	0.1835	0.1579	0.1643	0.1799	0.1757
E^2NeRF	0.1959	0.2020	0.1619	0.1723	0.1722	0.1809
BAD-NeRF	0.2027	0.1817	0.1549	0.1603	0.1867	0.1773
BAD-GS	0.2047	0.1835	0.1565	0.1620	0.1886	0.1790
BAGS	0.1987	0.2036	0.1819	0.2024	0.2012	0.1978
EvDNeRF	0.2067	0.2132	0.1709	0.1819	0.1817	0.1909
Ours	**0.2135**	**0.2188**	**0.1975**	**0.2188**	**0.2160**	**0.2129**

Table 12. The **Novel View Synthesis** of RankIQA in the **EvaGaussians-DAVIS** Dataset.

Models	**RankIQA ↓**					
	Desk & Chair	**Washroom**	**Pokémon**	**Pillow**	**Bag**	**Average**
B-NeRF	7.4896	10.4454	10.7921	9.6578	8.7541	9.4278
B-3DGS	5.9917	8.3563	8.6337	7.7262	7.0032	7.5422
UFP-GS	5.1153	7.1215	7.3741	6.6005	5.9829	6.4389
EDI-GS	5.0929	7.1029	7.3386	6.5673	5.9528	6.4109
EFN-GS	5.1046	7.0893	7.3158	6.5789	5.9682	6.4114
E^2NeRF	4.5141	6.4997	5.7764	5.7600	5.3166	5.5733
BAD-NeRF	4.9931	6.9636	7.1948	6.4385	5.8360	6.2852
BAD-GS	4.9436	6.8947	7.1235	6.3748	5.7783	6.2230
BAGS	4.6583	6.3721	5.8742	6.0157	5.2840	5.6410
EvDNeRF	4.3089	6.2043	5.5138	5.4981	5.0749	5.3200
Ours	**3.9369**	**5.6564**	**5.0867**	**5.1538**	**4.6422**	**4.8952**

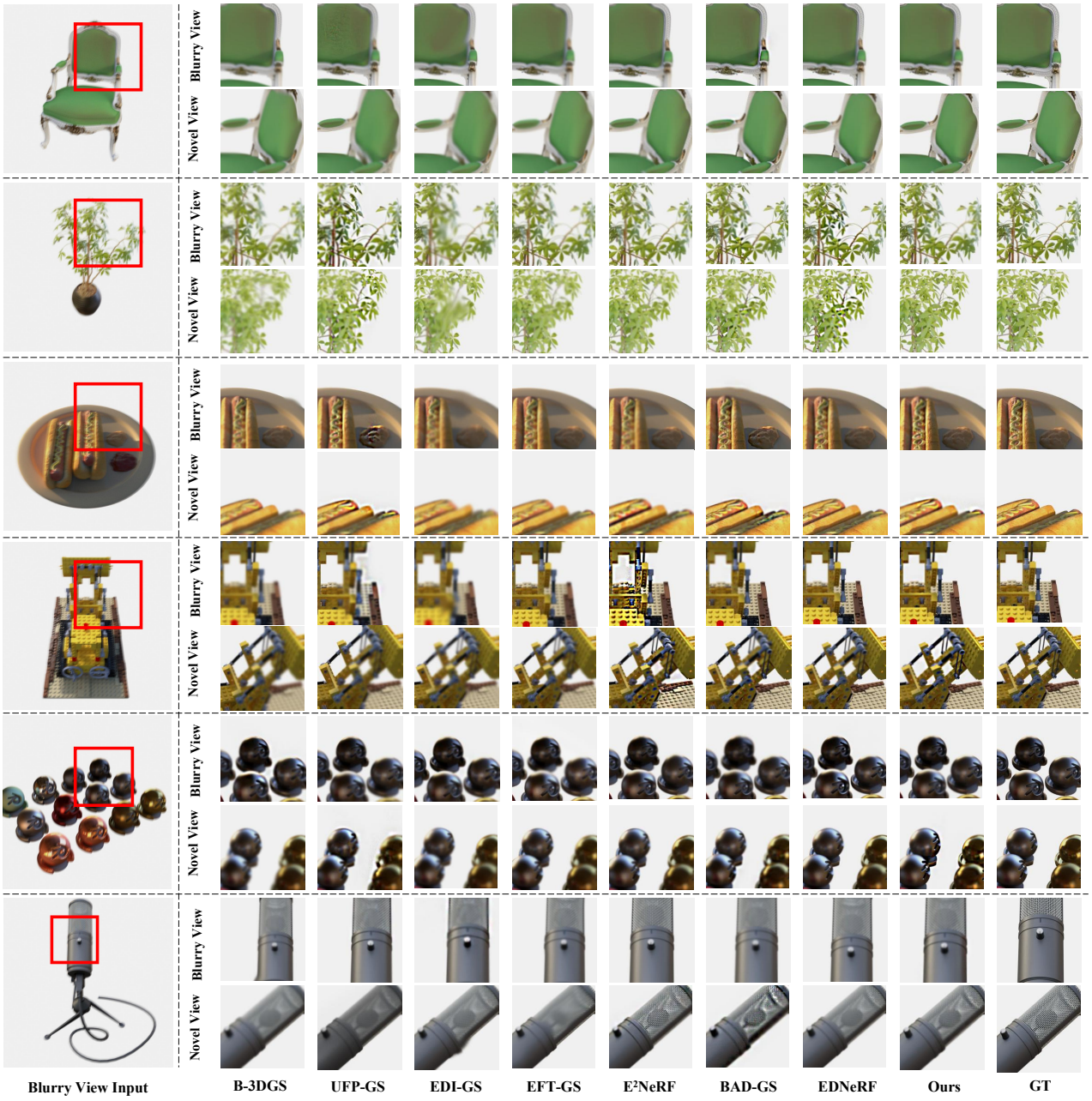


Figure 9. **Visualization of deblurring results and NVS results.** The deblurring results are highlighted in the red bouding box. The results demonstrate that our method consistently excels in reconstructing fine details and maintaining high color accuracy compared to the baselines.

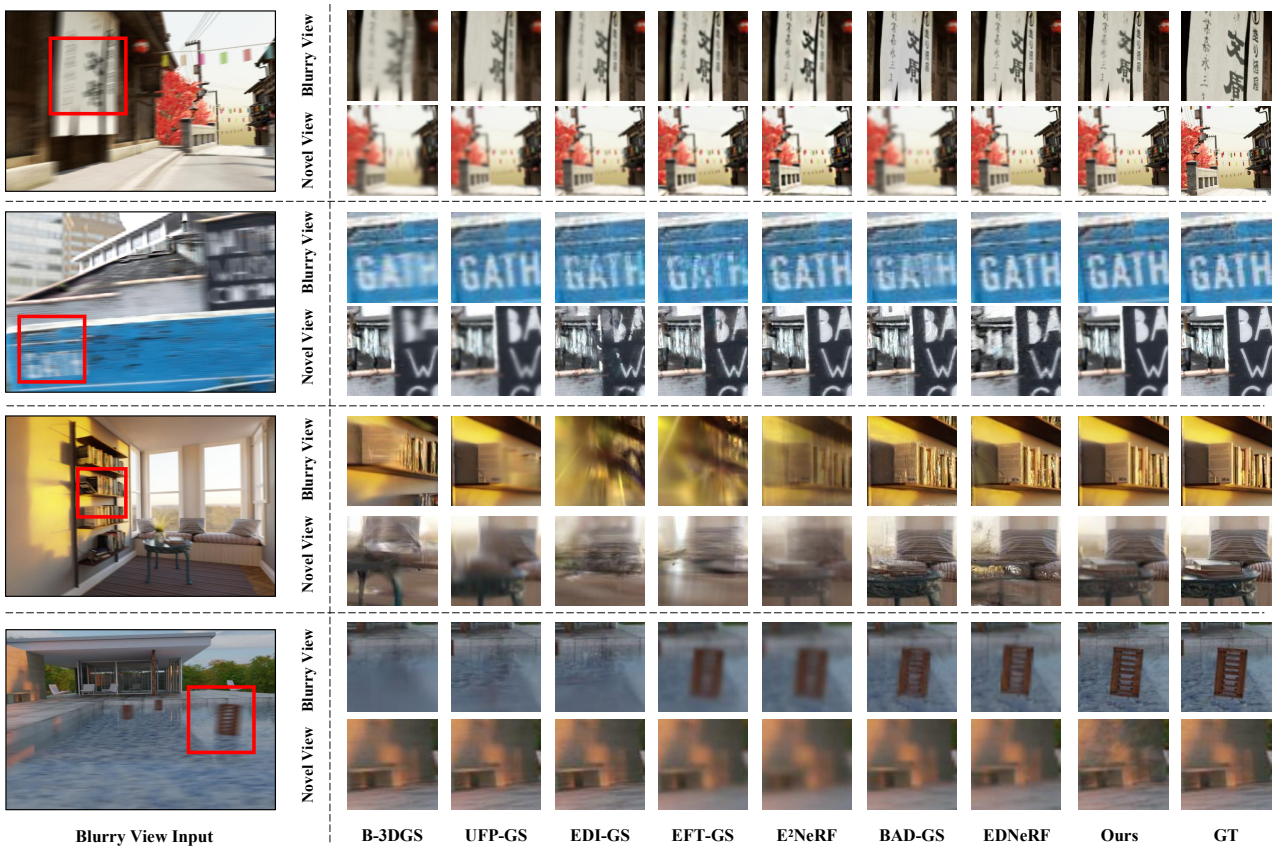


Figure 10. **Visualization of deblurring results and NVS results.** The deblurring results are highlighted in the red bouding box. The results demonstrate that our method consistently excels in reconstructing fine details and maintaining high color accuracy compared to the baselines.

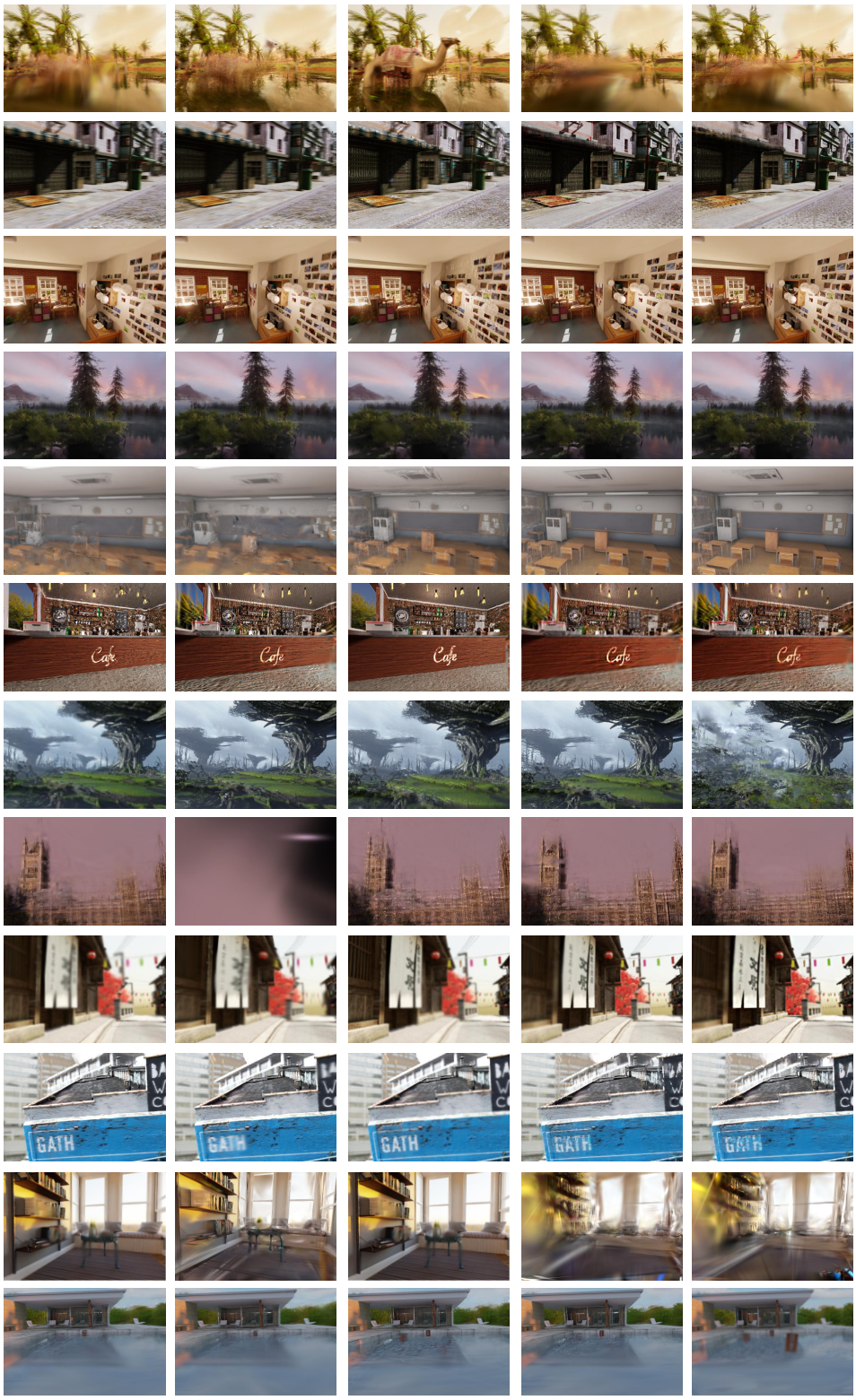


Figure 11. **Visualization of Novel View Synthesis.** We show results of B-NeRF, B-3DGS, EDI-GS [9], EFN-GS [12] and UFP-GS [4] in the EvaGaussians-Blender Dataset.

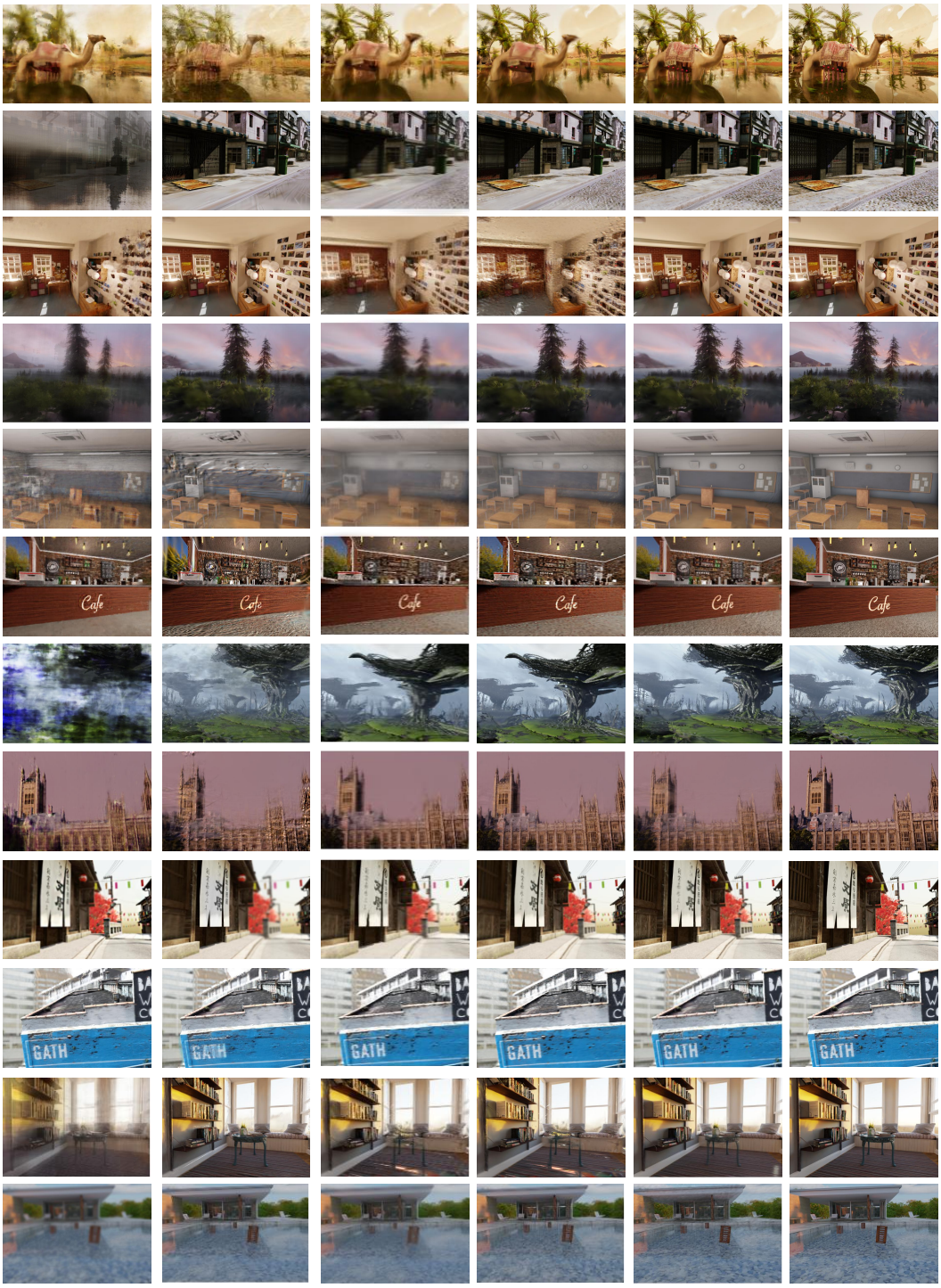


Figure 12. **Visualization of Novel View Synthesis**. We show results of E^{2}NeRF [10], BAD-GS [13], BAGS [13], EvDNeRF [2] and Ours in the EvaGaussian-Blender Dataset.

References

[1] Christian Brandli, Lorenz Muller, and Tobi Delbruck. Real-time, high-speed video decompression using a frame- and event-based davis sensor. In *2014 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 686–689, 2014. 2

[2] Marco Cannici and Davide Scaramuzza. Mitigating motion blur in neural radiance fields with events and frames. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024. 2, 5, 6, 14

[3] Blender Online Community. *Blender - a 3D modelling and rendering package*. Blender Foundation, Stichting Blender Foundation, Amsterdam, 2018. 1, 4

[4] Zhenxuan Fang, Fangfang Wu, Weisheng Dong, Xin Li, Jinjian Wu, and Guangming Shi. Self-supervised non-uniform kernel estimation with flow-based motion prior for blind image deblurring. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 18105–18114, 2023. 13

[5] Mathias Gehrig, Mario Millhäusler, Daniel Gehrig, and Davide Scaramuzza. E-raft: Dense optical flow from event cameras. In *International Conference on 3D Vision (3DV)*, 2021. 5, 6

[6] Yuhuang Hu, Shih-Chii Liu, and Tobi Delbruck. v2e: From video frames to realistic dvs events. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1312–1321, 2021. 1

[7] 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 *CVPR*, 2022. 1

[8] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In *ECCV*, 2020. 1

[9] Liyuan Pan, Cedric Scheerlinck, Xin Yu, Richard Hartley, Miaomiao Liu, and Yuchao Dai. Bringing a blurry frame alive at high frame-rate with an event camera. In *CVPR*, 2019. 5, 13

[10] Yunshan Qi, Lin Zhu, Yu Zhang, and Jia Li. E2nerf: Event enhanced neural radiance fields from blurry images. In *ICCV*, 2023. 2, 5, 14

[11] Johannes L Schonberger and Jan-Michael Frahm. Structure-from-motion Revisited. In *CVPR*, 2016. 5

[12] Lei Sun, Christos Sakaridis, Jingyun Liang, Qi Jiang, Kailun Yang, Peng Sun, Yaozu Ye, Kaiwei Wang, and Luc Van Gool. Event-based fusion for motion deblurring with cross-modal attention. In *European Conference on Computer Vision*, pages 412–428. Springer, 2022. 13

[13] Lingzhe Zhao, Peng Wang, and Peidong Liu. Bad-gaussians: Bundle adjusted deblur gaussian splatting. In *ECCV*, 2024. 14