DiET-GS

Diffusion Prior and Event Stream-Assisted Motion Deblurring 3DGS

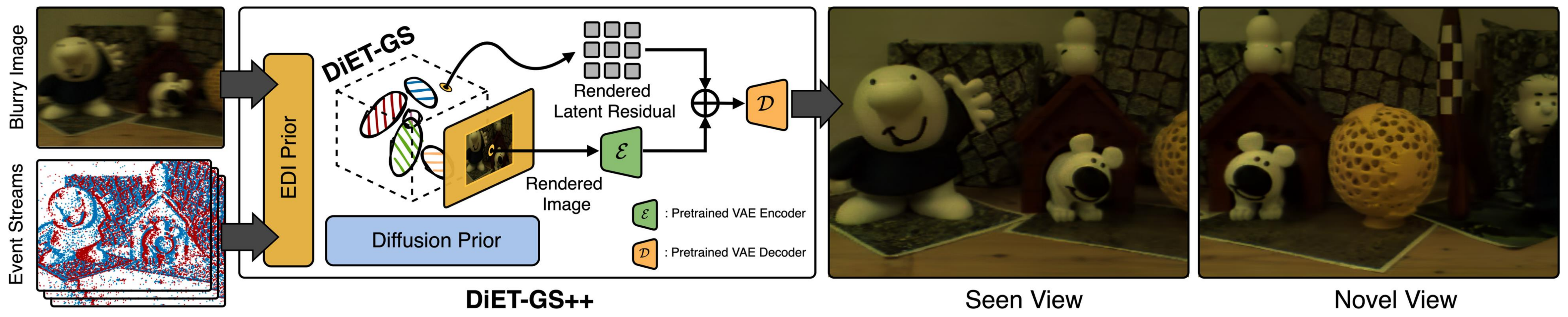
Seungjun Lee, Gim Hee Lee

Department of Computer Science, National University of Singapore

Project Page

Linkedin

Introduction



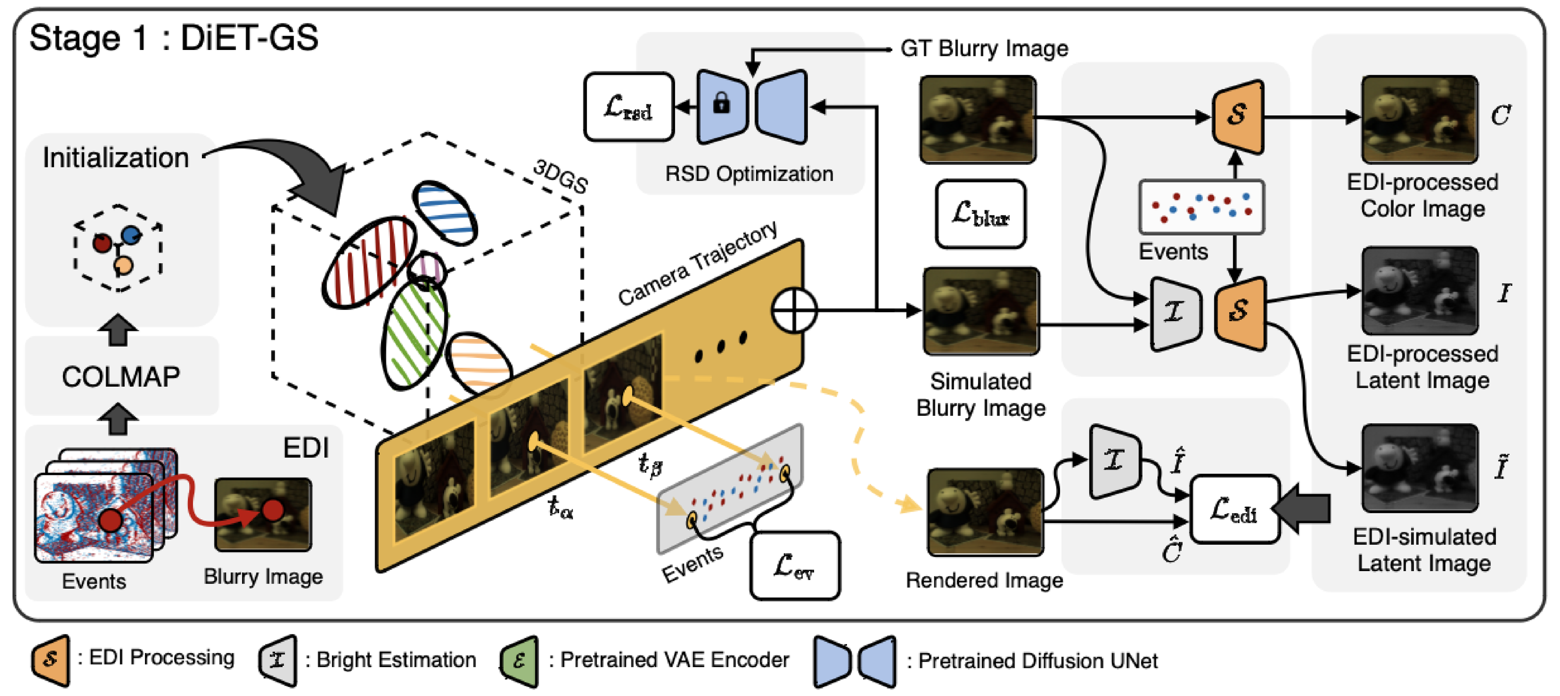
Motivation

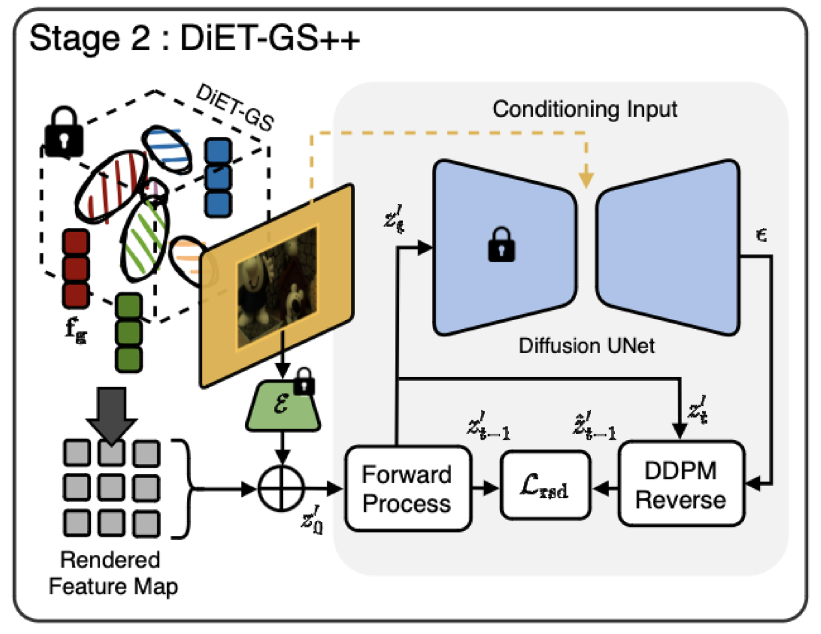
- Recent works attempt to enhance high-quality novel view synthesis from motion blur by leveraging event-based camera, benefiting from high dynamic range and microsecond temporal resolution.
- However, enhanced images with event streams often contain unnatural artifacts due to the noise accumulated from the events and unknown threshold of event trigger.
- Can we improve the image quality by both using event-stream and diffusion prior?

Contributions

- A novel framework to construct deblurring 3DGS by jointly leveraging event streams and the prior knowledge of a pretrained diffusion model.
- A two-stage training strategy to effectively utilize real-captured data and diffusion prior together. Once optimized, our method is capable of recovering well-defined details with accurate color from the input blurry images.

Overall Framework





Blur Reconstruction Loss

$$\hat{\mathbf{C}}_i^{\mathbf{B}} = \frac{1}{n}\sum_{j=0}^{n-1} g_\theta(p_{ij})$$

EDI Prior

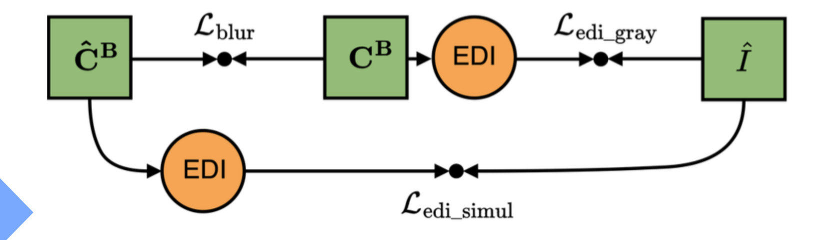
$$\mathbf{I}^{\mathbf{B}}(\mathbf{u},t) = \frac{I(\mathbf{u},t)}{\tau}\int_{t-\tau/2}^{t+\tau/2}\exp(\Theta\mathbf{E}(h))\,\mathrm{d}h$$

Can recover sharp image from blurry image

EDI Loss

Leverage EDI prior as:

1. Color Guidance
2. Guidance for fine-grained details
3. Regularization with cycle consistency



Stage 1: DiET-GS

RSD Loss: Diffusion prior

Leveraging diffusion prior to enhance the edge details of deblurred image, making more natural.

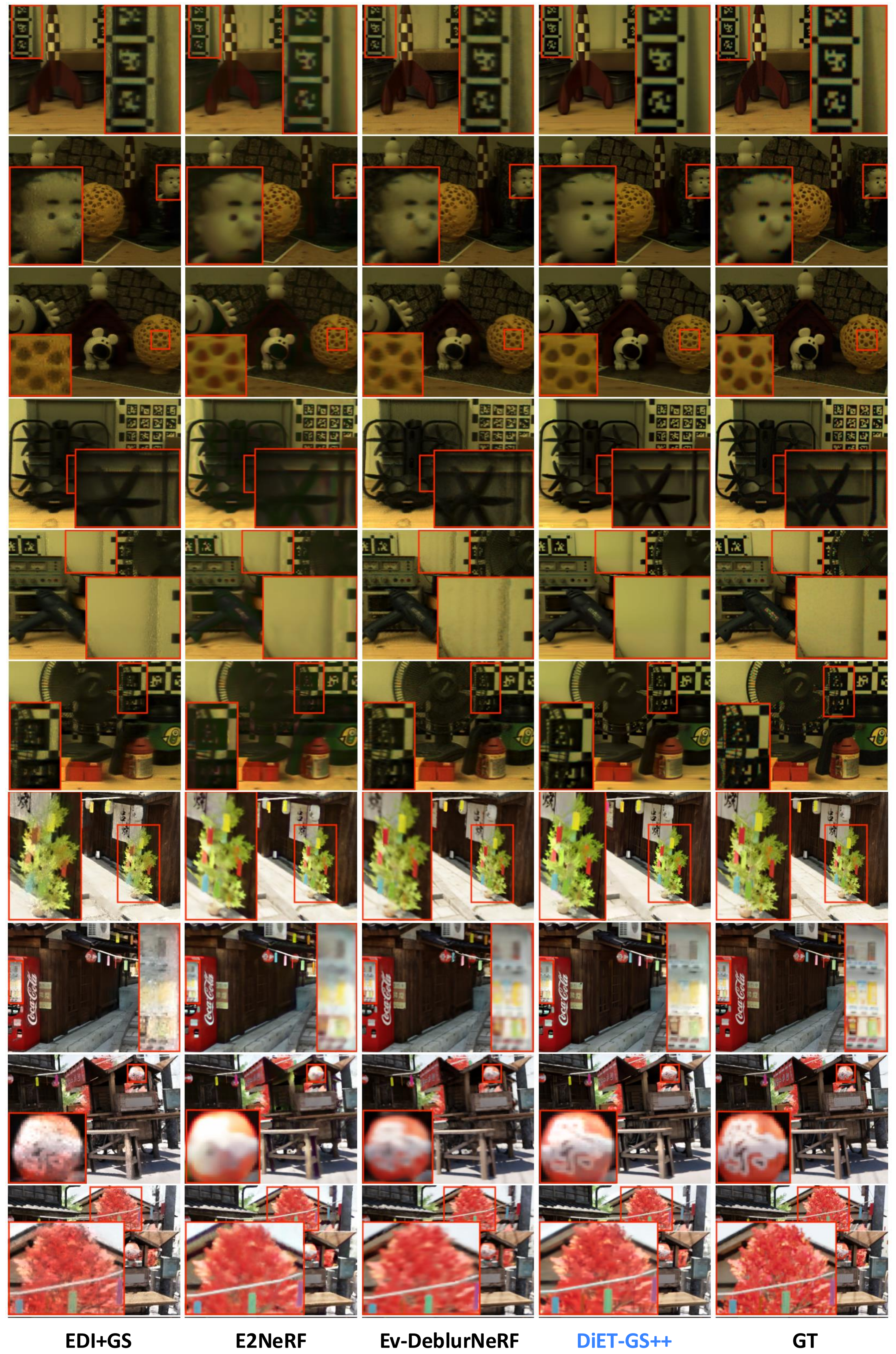
Key Idea: Modify RSD Loss proposed in DiSR-NeRF to deblurring setting.

- Use deblurred image from Stage 1 as conditional input.
- Declare learnable feature to each 3D Gaussian in pretrained DiET-GS, encoding diffusion prior into 3DGS.
- In the inference, rendered image from DiET-GS is combined to rendered latent feature:

$$\tilde{C} = \mathcal{D}(\mathbf{z}_0') = \mathcal{D}(\mathbf{f}_{2D} + \mathcal{E}(\hat{C}))$$

Stage 2: DiET-GS++

Check visualization results!



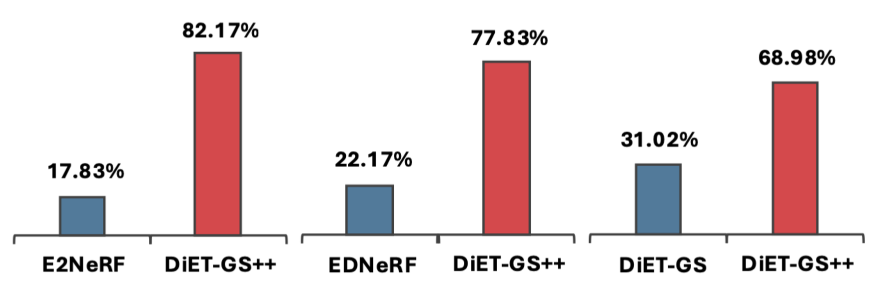
Experiments

Dataset	Metric	MPRNet+GS [65]	EDI+GS [35]	EFNet+GS [50]	BAD-NeRF [56]	BAD-GS [69]	E2NeRF [38]	Ev-DeblurNeRF [3]	DiET-GS (Ours)	DiET-GS++ (Ours)
EvDeblur-blender	PSNR↑	18.76	23.69	21.03	19.78	22.23	24.54	24.76	**26.69**	<u>26.23</u>
	SSIM↑	0.5912	0.7694	0.6413	0.6381	0.7213	0.7993	0.8038	**0.8607**	<u>0.8478</u>
	LPIPS↓	0.3545	0.1375	0.3214	0.2490	0.2012	0.1624	0.1788	<u>0.1064</u>	**0.1052**
	MUSIQ↑	24.12	55.13	35.13	23.63	32.43	47.31	42.38	<u>57.67</u>	**59.91**
	CLIP-IQA↑	0.2413	0.2751	0.2314	0.1888	0.1993	0.2129	0.2300	<u>0.2769</u>	**0.2960**
EvDeblur-CDAVIS	PSNR↑	27.51	32.95	30.97	28.47	29.12	31.54	32.30	**34.22**	<u>33.16</u>
	SSIM↑	0.7514	0.8922	0.8503	0.7981	0.8129	0.8687	0.8827	**0.9223**	<u>0.9039</u>
	LPIPS↓	0.2013	0.0790	0.1142	0.2526	0.2012	0.1059	0.0571	**0.0496**	<u>0.0502</u>
	MUSIQ↑	25.12	40.06	38.23	19.96	22.12	38.82	41.32	<u>45.80</u>	**50.44**
	CLIP-IQA↑	0.2134	0.2008	0.1934	0.1791	0.1812	<u>0.2235</u>	0.2211	0.2072	**0.2415**

Table 1. Quantitative comparisons on both synthetic and real-world dataset. The results are the average of every scenes within the dataset. The best results are in **bold** while the second best results are <u>underscored</u>.

Quantitative comparison. Our DiET-GS largely outperforms all baselines in PSNR, SSIM and LPIPS on both synthetic and real-world datasets. Furthermore, DiET-GS++ shows significant improvement in MUSIQ and CLIP-IQA metrics, achieving the best results.

Comparison between DiET-GS vs DiET-GS++. DiET-GS++ further refines the textures and edge details generated from DiET-GS, demonstrating the efficacy of fully leveraging diffusion prior in Stage 2.



User Study. User study is conducted by 60 evaluators with 30 pairs of samples. During the study, evaluators were asked to select the image with better quality between two presented options for every pairs. Our DiET-GS++ gains significantly higher votes against the baselines, showing at least 37.96% difference.