BokehMe: When Neural Rendering Meets Classical Rendering

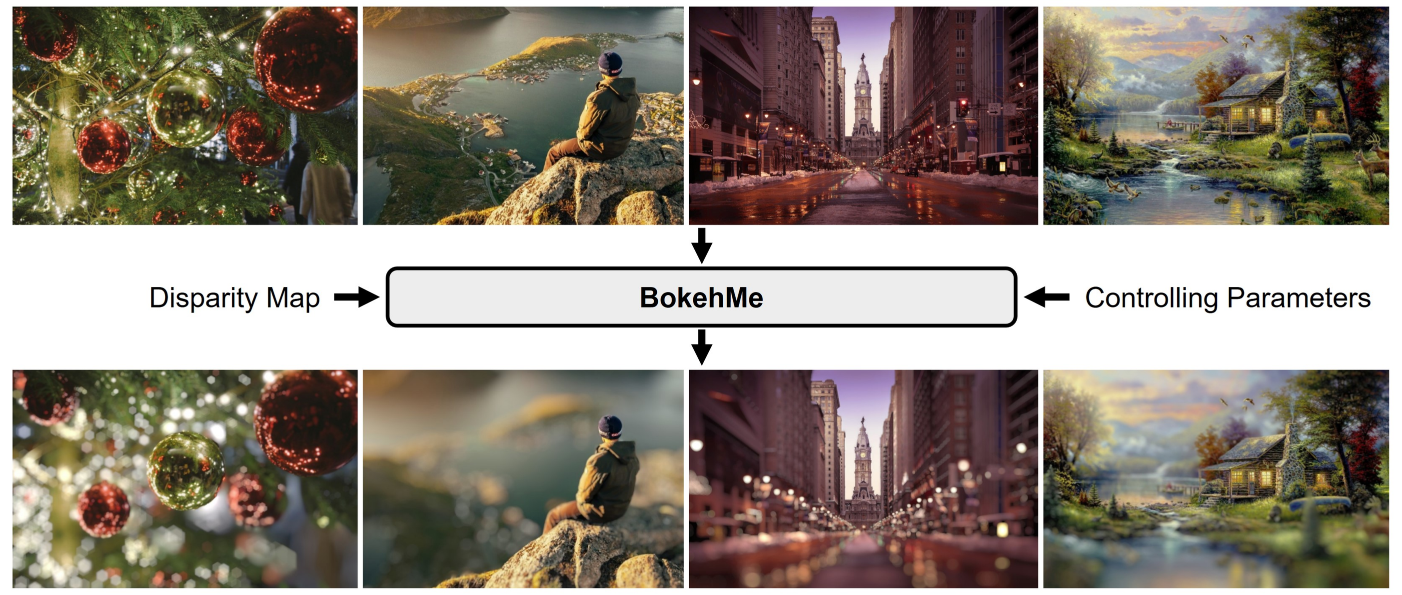
Juewen Peng[1], Zhiguo Cao[1], Xianrui Luo[1], Hao Lu[1], Ke Xian[1*], and Jianming Zhang[2]

[1]School of Artificial Intelligence and Automation, Huazhong University of Science and Technology [2]Adobe Research

CVPR JUNE 19-24 2022 NEW ORLEANS • LOUISIANA

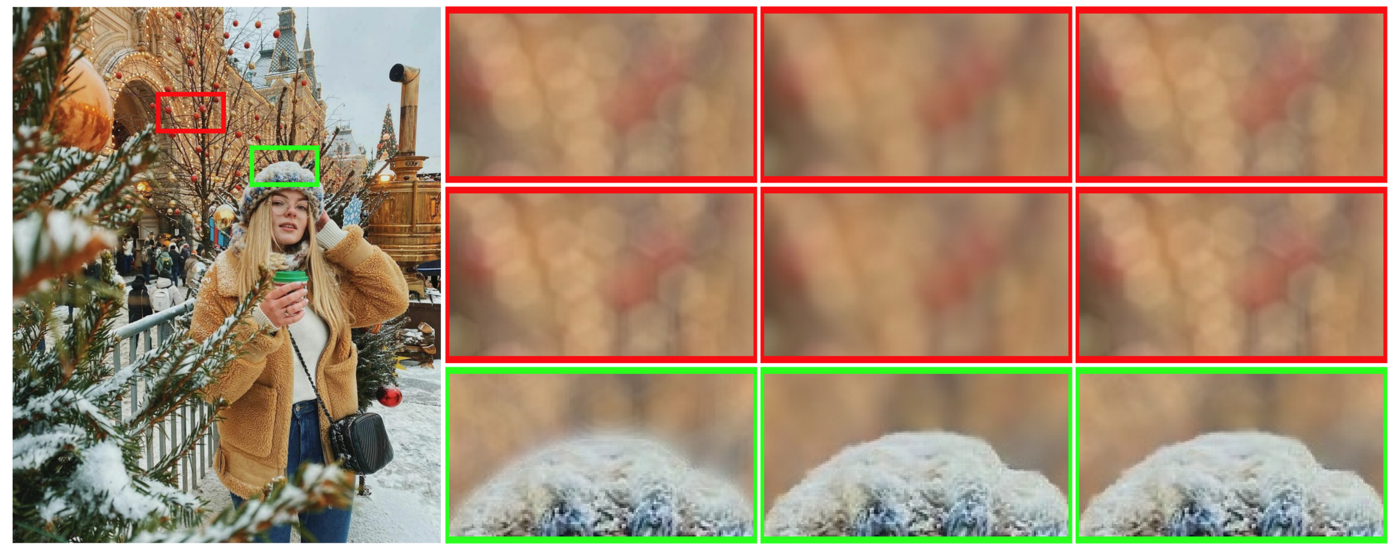
Problem Statement

Provided an all-in-focus image, a potentially imperfect disparity map (measured or predicted), and some controlling parameters (blur amount, refocused disparity, gamma, aperture shape), our goal is to render high-resolution, realistic and adjustable bokeh images.



Motivation

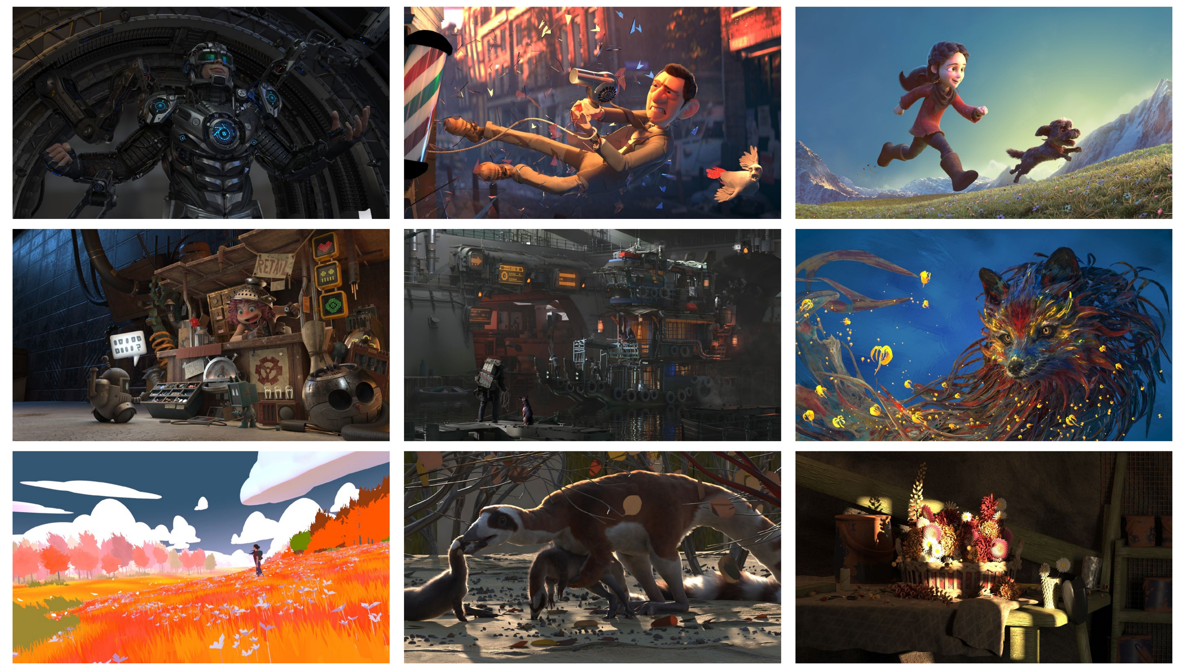
- Classical methods suffer from artifacts at depth discontinuities while neural methods have difficulty simulating real and variant bokeh balls.
- We make the best of the two worlds by fusing the results of a classical renderer (CR) and a neural renderer (NR).

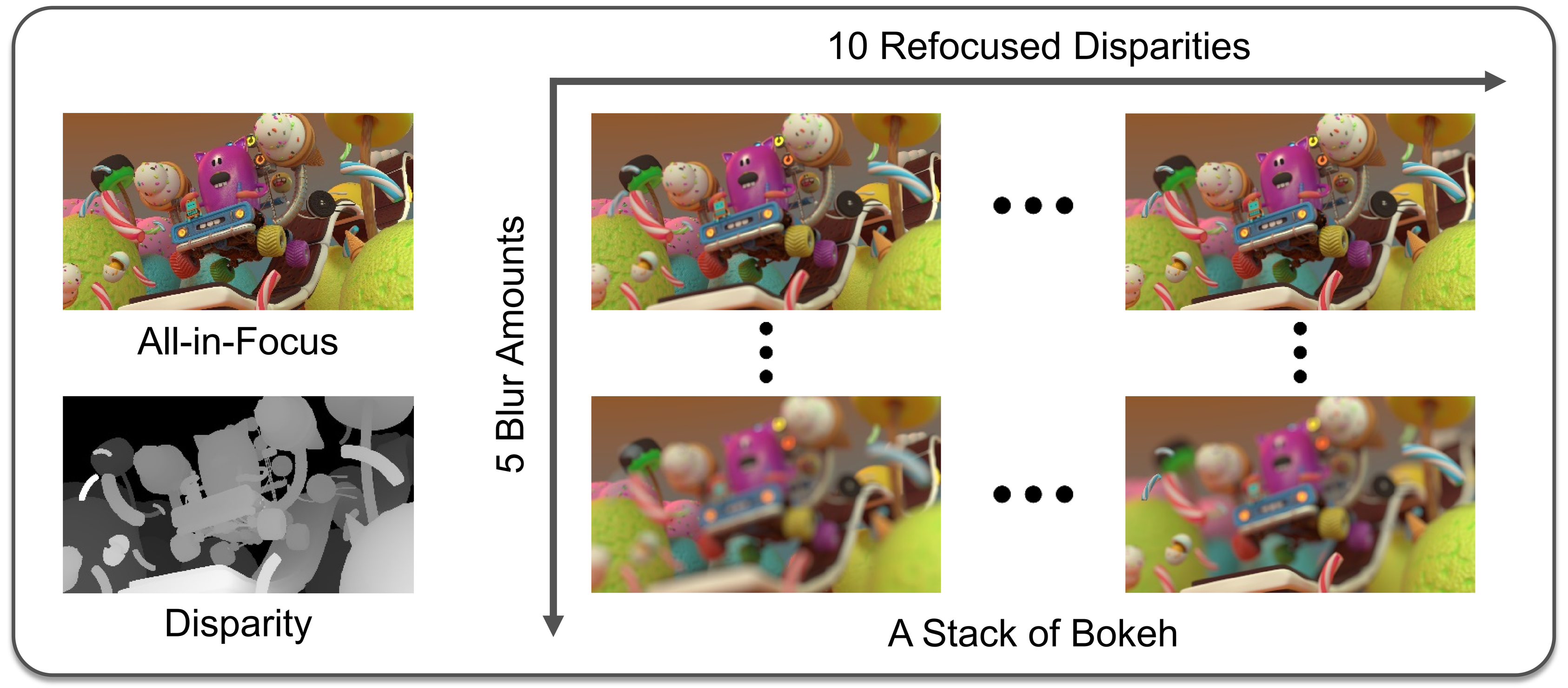


All-in-Focus Ours (CR) Ours (NR) Ours

BLB Dataset

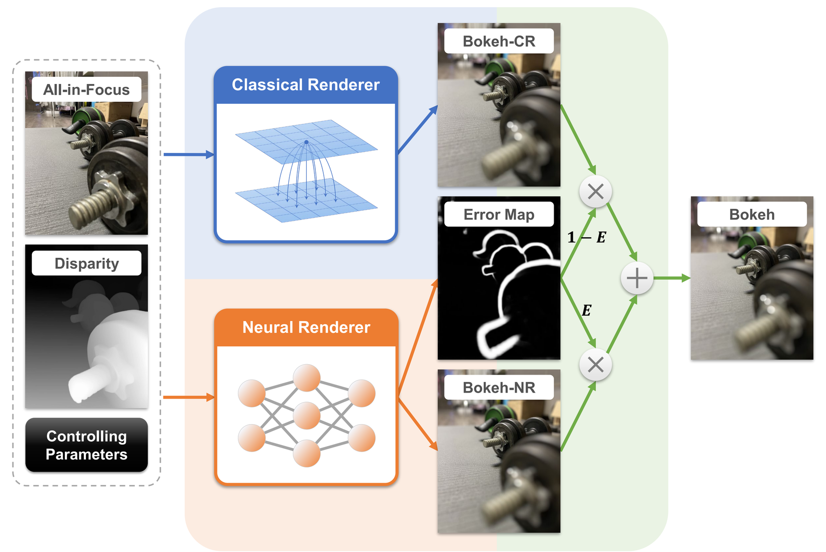
BLB dataset, which is synthesized by Blender 2.93, contains 10 scenes. Each scene consists of an all-in-focus image, a disparity map and a stack of bokeh images with 5 blur amounts and 10 refocused disparities.





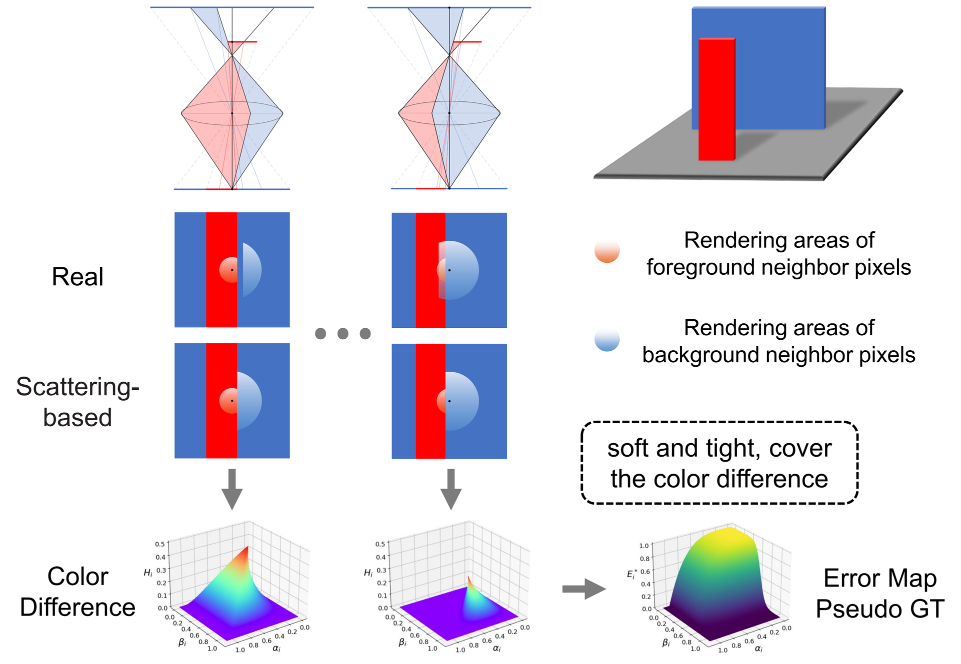
Method

Framework:



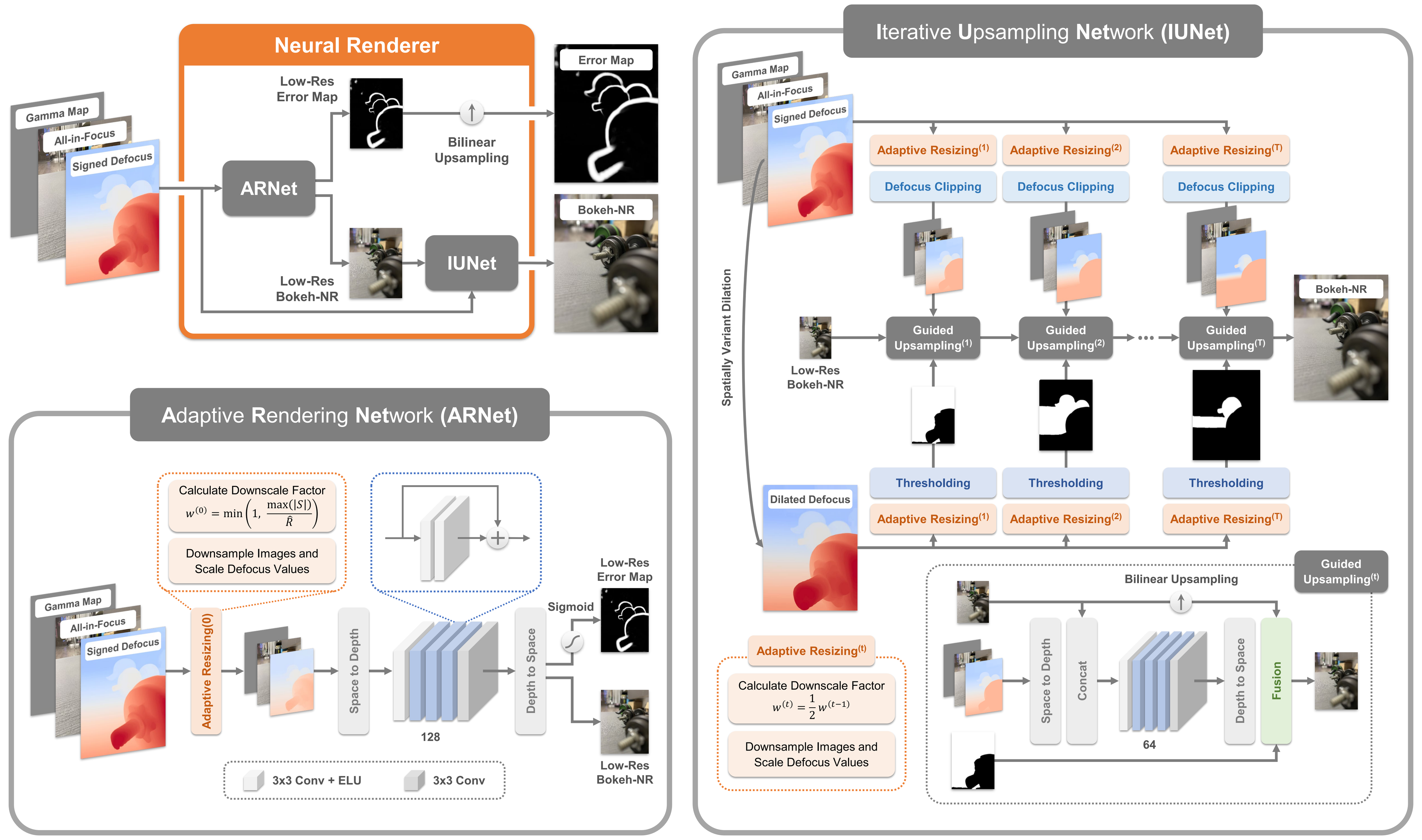
Classical Renderer and Error Analysis:

CR is implemented by a scattering-based method. To determine the areas rendered incorrectly by this method, we conduct an error analysis between scattering-based rendering and real rendering.



Neural Renderer:

- We propose an adaptive resizing layer in ARNet and an iterative guided upsampling mechanism in IUNet so that NR can handle arbitrarily large blur sizes.
- The input signed defocus map is defined by $S = K(D - d_f)$, where D is the disparity map, K is the blur parameter and d_f is the refocused disparity.



Paper, Code and Data: `https://juewenpeng.github.io/BokehMe/`

Experiments

Quantitative Results on BLB Dataset:

Methods	Level 1			Level 2			Level 3			Level 4			Level 5		
	PSNR	SSIM	Time(s)	PSNR	SSIM	Time(s)	PSNR	SSIM	Time(s)	PSNR	SSIM	Time(s)	PSNR	SSIM	Time(s)
VDSLR	41.13	0.9891	0.06	39.15	0.9848	0.23	37.64	0.9812	0.53	36.48	0.9783	0.97	35.57	0.9760	1.55
SteReFo	37.21	0.9831	0.13	35.28	0.9818	0.60	33.99	0.9813	1.69	32.94	0.9809	3.74	32.12	0.9805	6.87
RVR	32.35	0.9648	0.10	32.00	0.9321	0.43	28.36	0.9011	1.11	25.80	0.8775	2.30	23.94	0.8596	4.12
RVR†	37.15	0.9836	0.13	38.55	0.9880	0.62	35.56	0.9854	1.82	33.03	0.9815	3.97	31.15	0.9774	7.21
DeepLens	33.68	0.9679	0.14	31.43	0.9603	0.14	30.16	0.9564	**0.14**	29.30	0.9539	0.14	28.68	0.9521	0.14
DeepFocus	38.92	0.9900	0.71	36.13	0.9857	0.71	31.47	0.9623	0.71	25.55	0.9089	0.71	21.04	0.8227	0.71
DeepFocus†	38.92	0.9900	0.71	35.64	0.9855	0.49	34.08	0.9823	0.22	33.07	0.9797	**0.13**	32.29	0.9774	**0.09**
Ours (CR)	41.32	0.9900	**0.03**	39.51	0.9877	**0.10**	38.35	0.9868	0.20	37.53	0.9864	0.34	36.86	0.9862	0.52
Ours (NR)	40.41	0.9905	0.13	40.16	0.9904	0.13	39.21	0.9896	**0.14**	38.01	0.9884	0.16	37.20	0.9875	0.16
Ours	**43.30**	**0.9932**	0.16	**42.21**	**0.9924**	0.23	**41.02**	**0.9915**	0.34	**39.78**	**0.9906**	0.50	**38.80**	**0.9898**	0.68

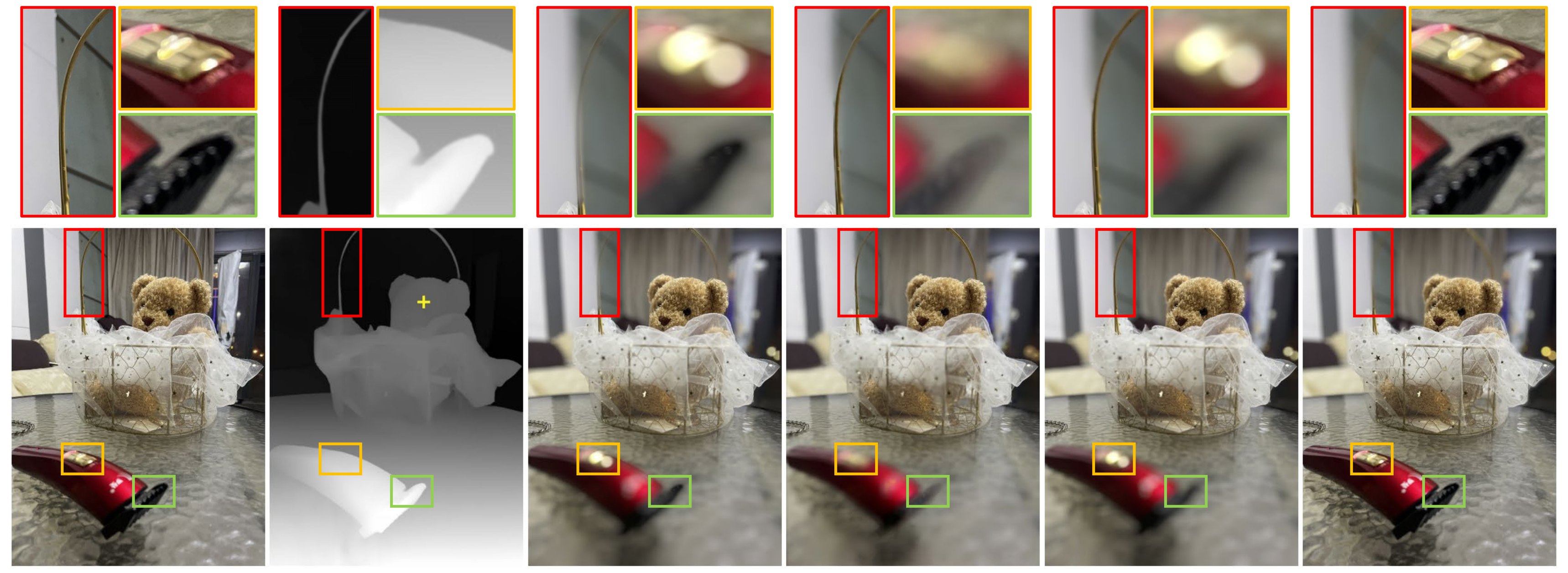
Quantitative Results on EBB400 Dataset:.

Methods	VDSLR	SteReFo	RVR†	DeepLens	DeepFocus†	Ours
PSNR	23.78	23.56	23.56	23.46	23.81	**23.85**
SSIM	0.8738	0.8674	0.8690	0.8707	0.8754	**0.8770**

User Study Results on IPB Dataset:

Methods	iPhone 12	VDSLR	DeepLens	Ours
Good (%)	19.3	26.6	26.3	**55.0**
Normal (%)	29.3	47.7	45.0	38.5
Bad (%)	51.4	25.7	28.7	**6.5**

Qualitative Results on IPB Dataset:



All-in-Focus Disparity VDSLR DeepLens Ours iPhone 12 Portrait