Lighting Every Darkness in Two Pairs:

A Calibration-Free Pipeline for RAW Denoising

ICCV23 PARIS

Xin Jin*, Jia-Wen Xiao*, Ling-Hao Han, Chunle Guo#, Runxun Zhang, Xialei Liu, Chongyi Li

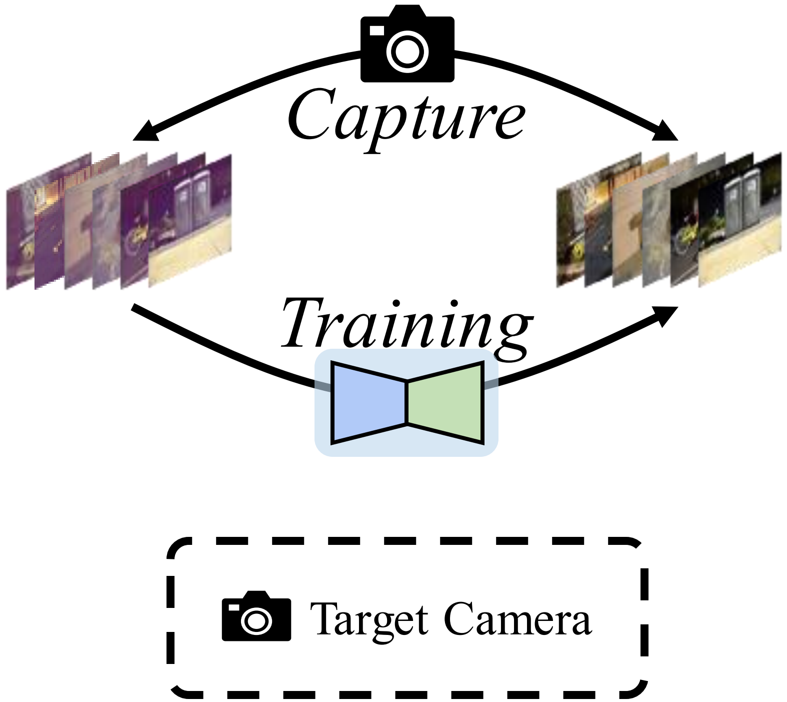
南開大學 Nankai University

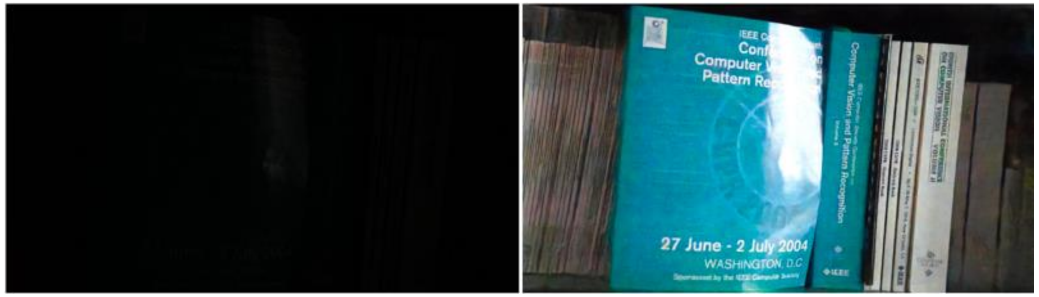
北京大學 PEKING UNIVERSITY

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Background

- **Task: Low-light RAW Image Denoising**
- **Previous work: SID[1]**





(a) Input (b) Output

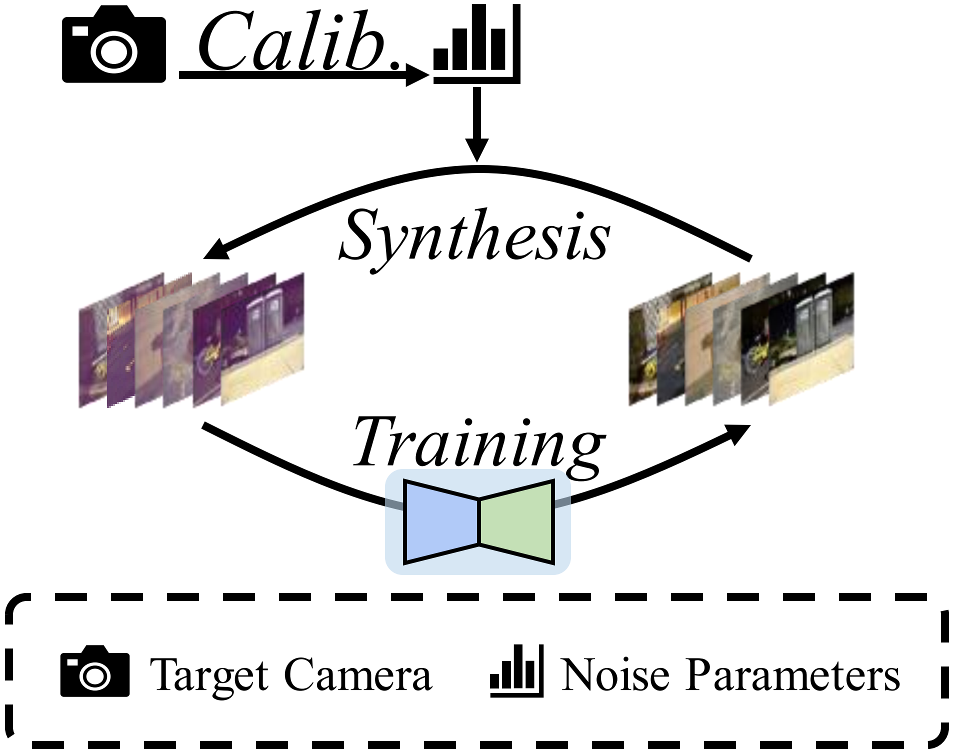
Why denoise in RAW?

The difference between sensors should be considered![2]
Dataset re-collection for each sensor is required!

[1] Chen, Chen, et al. "Learning to see in the dark." *CVPR*. 2018.
[2] Wei, Kaixuan, et al. "Physics-based noise modeling for extreme low-light photography." *IEEE TPAMI*. 2021.

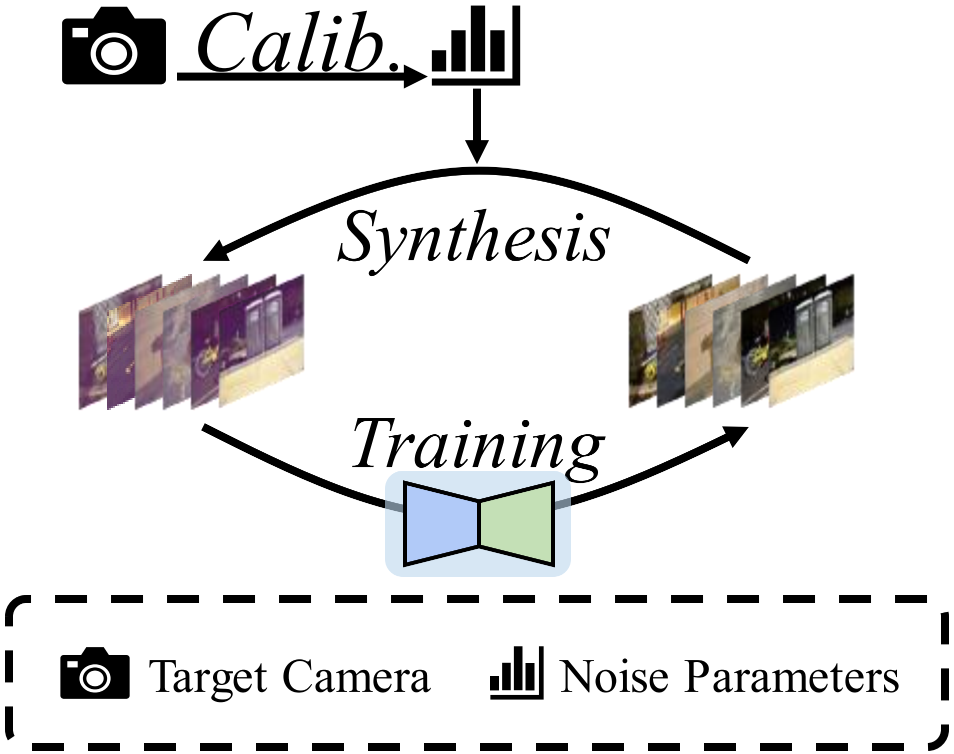
Background: Training with Synthesis Noise

- **Simulating real noise pairs with noise model.**
- **Existing SOTA methods[2][3][4][5] are all based on calibration.**

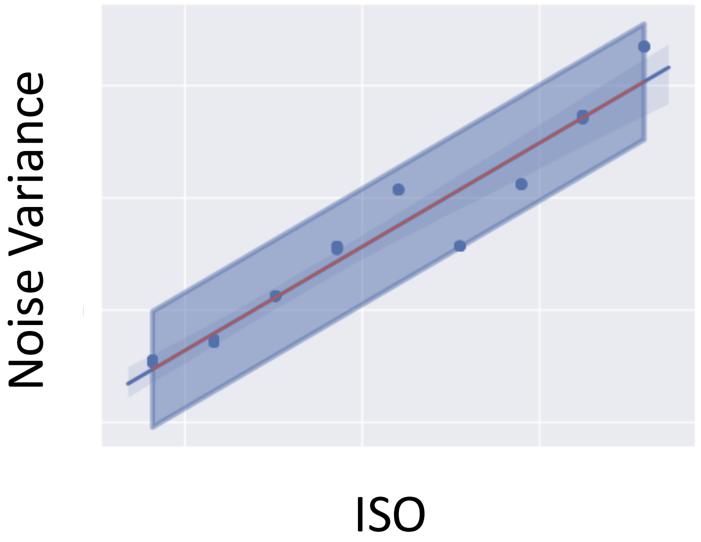


[2] Wei, Kaixuan, et al. "Physics-based noise modeling for extreme low-light photography." *IEEE TPAMI*. 2021.
[3] Zhang, Yi, et al. "Rethinking noise synthesis and modeling in raw denoising." *ICCV*. 2021.
[4] Feng, Hansen, et al. "Learnability enhancement for low-light raw denoising: Where paired real data meets noise modeling." *ACMMM.* 2022
[5] Cao, Yue, et al. "Physics-Guided ISO-Dependent Sensor Noise Modeling for Extreme Low-Light Photography." *CVPR*. 2023.

Background: Calibration



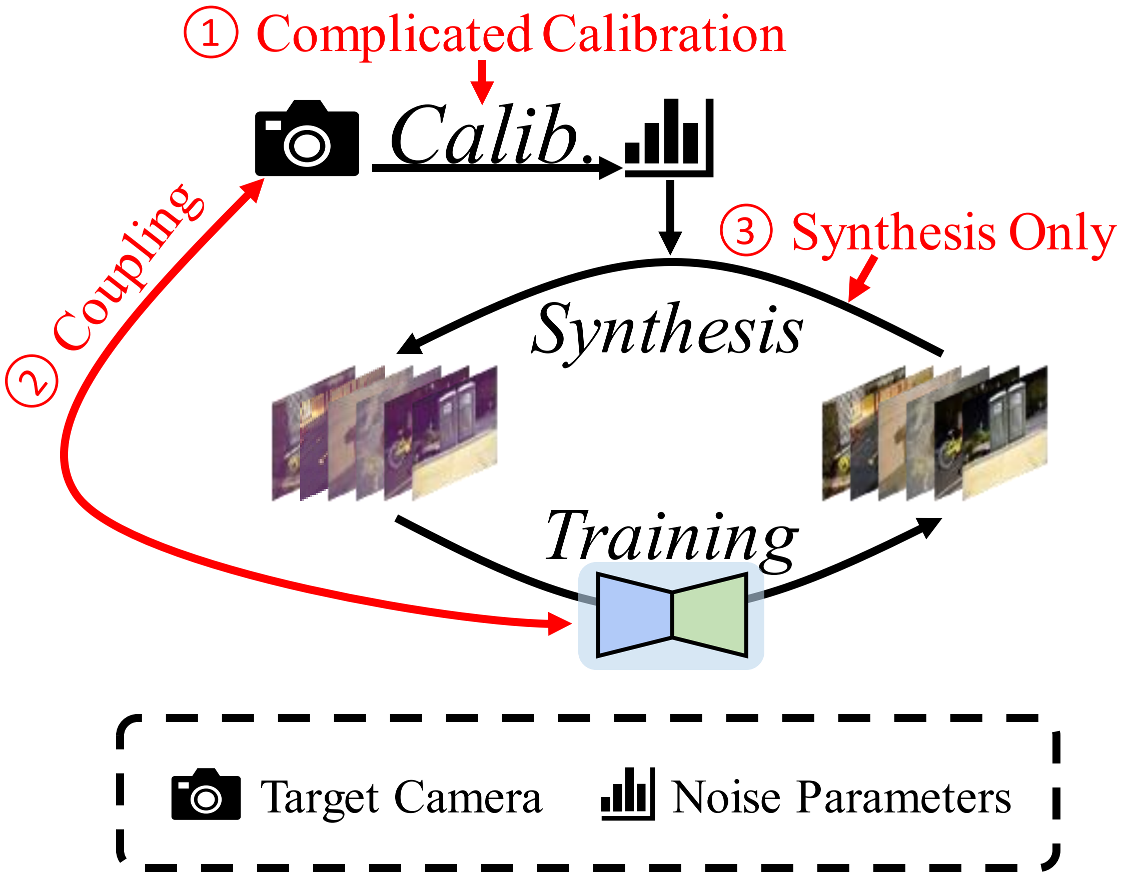
1. **Preparation:**
   pre-define noise model + data collection.
   $N = N_{shot} + N_{read} + N_{row} + N_{quant}$ [2]
2. **Calibration:** parameter estimation.



3. **Training:** train the network with synthesis data.

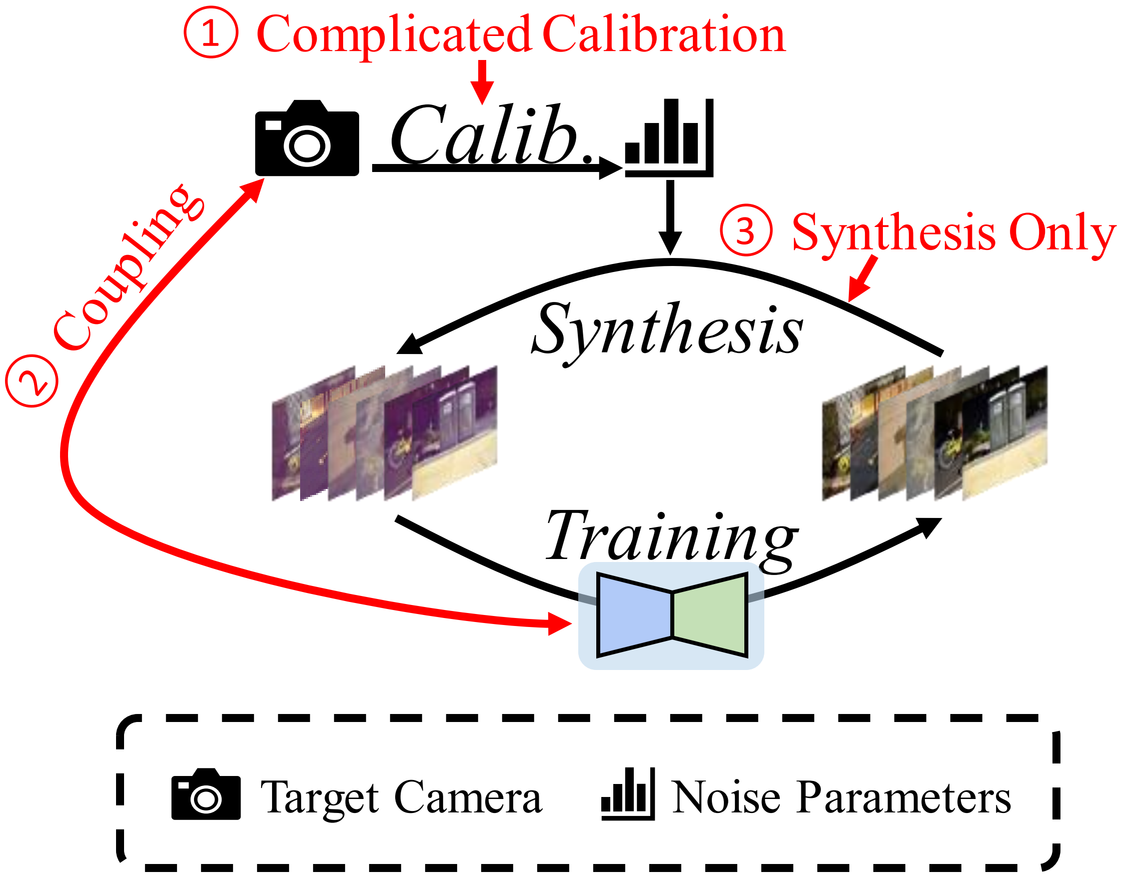
[2] Wei, Kaixuan, et al. "Physics-based noise modeling for extreme low-light photography." *IEEE TPAMI*. 2021.

Is Calibration Strategy Really Perfect?



- **Time-consuming and labor-intensive** calibration and noise parameter estimation.
- Neural networks are **tightly coupled** with the target camera sensor, requiring **retraining** for each camera type.
- During training, only synthetic noise is involved, and it **cannot generalize** to noise outside the noise model.

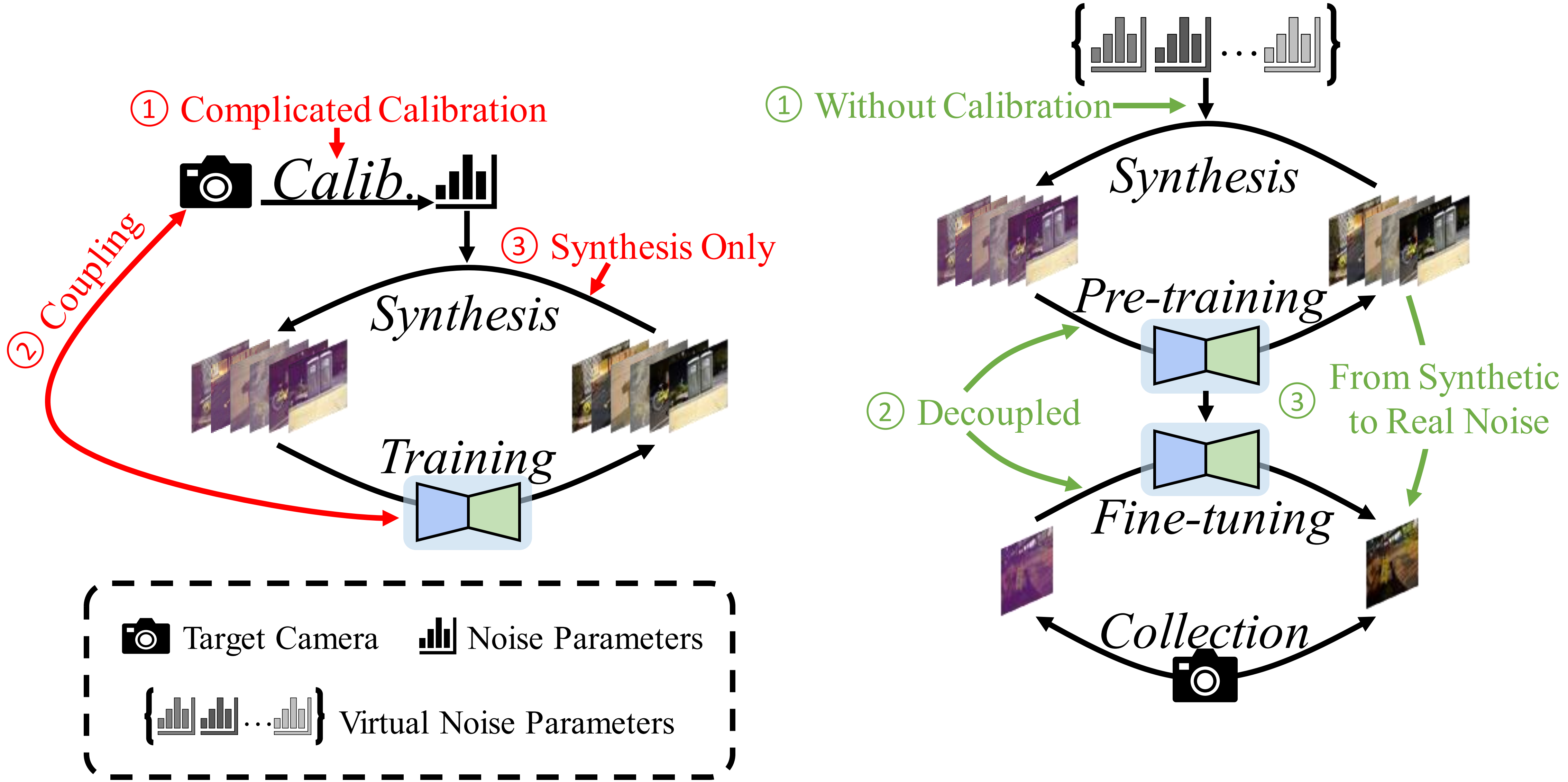
What Do We Want?



- **Simplify calibration**[6][7]**, even eliminate calibration procedure.**
- **The ability to rapidly deploy on new cameras.**
- **Generalization capability: Excellent generalization to real-world scenarios, bridging the domain gap between synthetic and real.**

[6] Zou, Yunhao, and Ying Fu. "Estimating fine-grained noise model via contrastive learning." CVPR. 2022.
[7] Monakhova, Kristina, et al. "Dancing under the stars: video denoising in starlight." CVPR . 2022.

LED: A Calibration-free Pipeline



What can LEDs achieve?

Table 1. Quantitative results on the SID [7] Sony subset. The best result is in **bold** whereas the second best one is in underlined. The extra data requiren ed methods require training noise gene at the AINDNet is pre-trained with ou ained with the same UNet architecture, while we keep the AINDNet the same as their paper with almost twice the number of parameters compared to the UNet.

6 Pairs + 1.5K iteration = SOTA performance!

Categories	Methods	Extra Data Requirements	Iterations (K)	×100		×250		×300	
				PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
DNN Model Based	Kristina *et al.* [37]	∼1800 noisy-clean pairs	327.6	38.7799	0.9120	34.4924	0.7900	31.2971	0.6990
	NoiseFlow [1]	∼1800 noisy-clean pairs	777.6	37.0200	0.8820	32.9457	0.7699	29.8068	0.6700
Calibration-Based	Calibrated P-G	∼300 calibration data	257.6	39.1576	0.8963	33.8929	0.7630	31.0035	0.6522
	ELD [46]	∼300 calibration data	257.6	41.8271	0.9538	38.8492	0.9278	35.9402	0.8982
	Zhang *et al.* [55]	∼150/∼150 for calib./database	257.6	40.9232	0.9488	38.4397	0.9255	35.5439	0.8975
Real Data Based	SID [7]	∼1800 noisy-clean pairs	257.6	41.7273	0.9531	39.1353	0.9304	**37.3627**	**0.9341**
	Noise2Noise [34]	∼12000 noisy pairs	257.6	39.2769	0.8993	34.1660	0.7824	31.0991	0.7080
	AINDNet [30]	∼300 noisy-clean pairs	**1.5**	40.5636	0.9194	36.2538	0.8509	32.2291	0.7397
	AINDNet*	∼300 noisy-clean pairs	**1.5**	39.8052	0.9350	37.2210	0.9101	34.5615	0.8856
	LED (Ours)	**6 noisy-clean pairs**	**1.5**	**41.9842**	**0.9539**	**39.3419**	**0.9317**	36.6728	0.9147

What can LEDs achieve?

PMN[4]'s training strategy

Reducing almost <u>a full day</u> of training time to just <u>4 minutes</u>!

```
2023-07-16 21:41:09,032 INFO: End of training. Time consumed: 23:42:27
2023-07-16 21:41:09,033 INFO: Save the latest model.
2023-07-16 21:41:15,461 INFO: Validation SIDSonyPaired100
     # psnr: 42.0811
     # ssim: 0.9550

2023-07-16 21:41:21,843 INFO: Validation SIDSonyPaired250
     # psnr: 39.4613
     # ssim: 0.9340

2023-07-16 21:41:29,654 INFO: Validation SIDSonyPaired300
     # psnr: 36.8701
     # ssim: 0.9203
```

Previous SOTA[2]

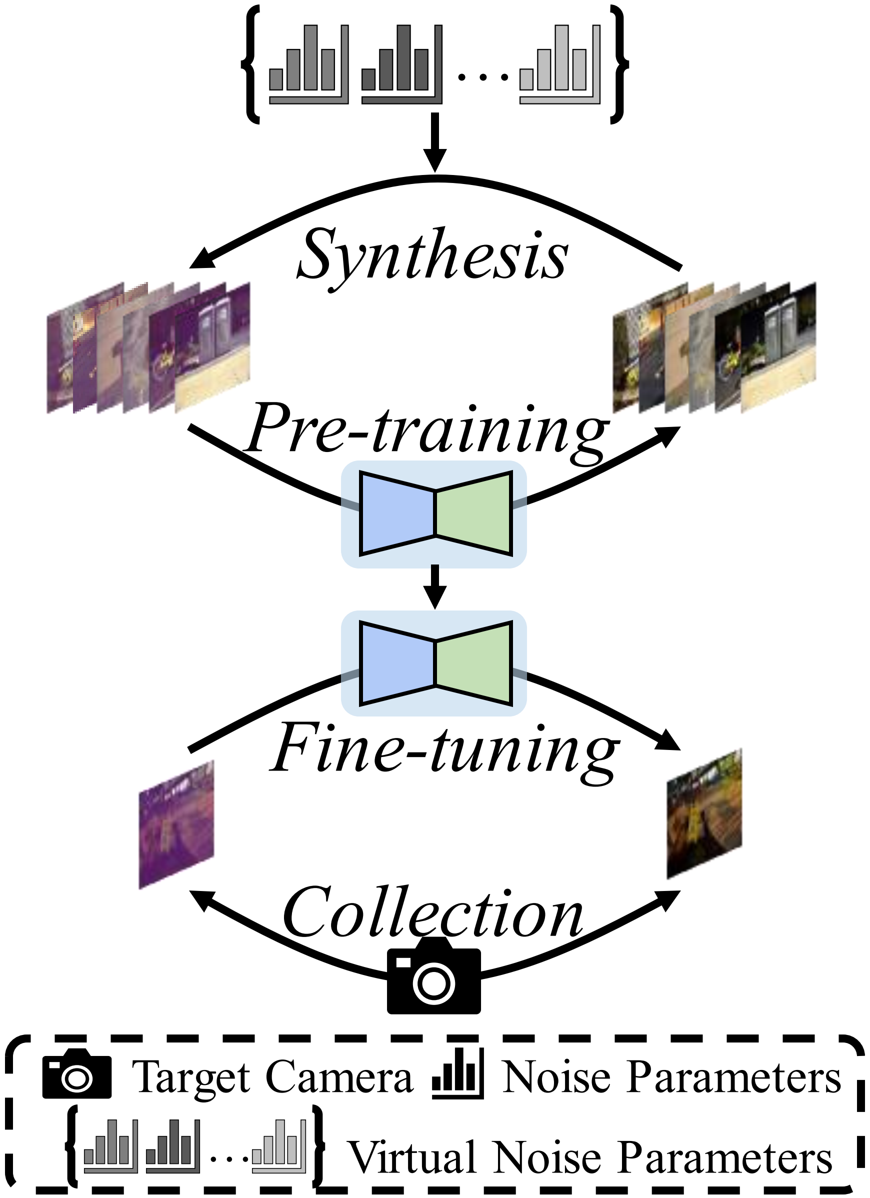
```
2023-07-16 21:26:03,931 INFO: End of training. Time consumed: 0:03:43
2023-07-16 21:26:03,931 INFO: Save the latest model.
2023-07-16 21:26:13,230 INFO: Validation SIDSonyPaired100
     # psnr: 42.3397 ↑0.2568
     # ssim: 0.9549

2023-07-16 21:26:20,423 INFO: Validation SIDSonyPaired250
     # psnr: 39.6064 ↑0.1451
     # ssim: 0.9370

2023-07-16 21:26:29,060 INFO: Validation SIDSonyPaired300
     # psnr: 36.9314 ↑0.0613
     # ssim: 0.9256
```

LED (Ours)

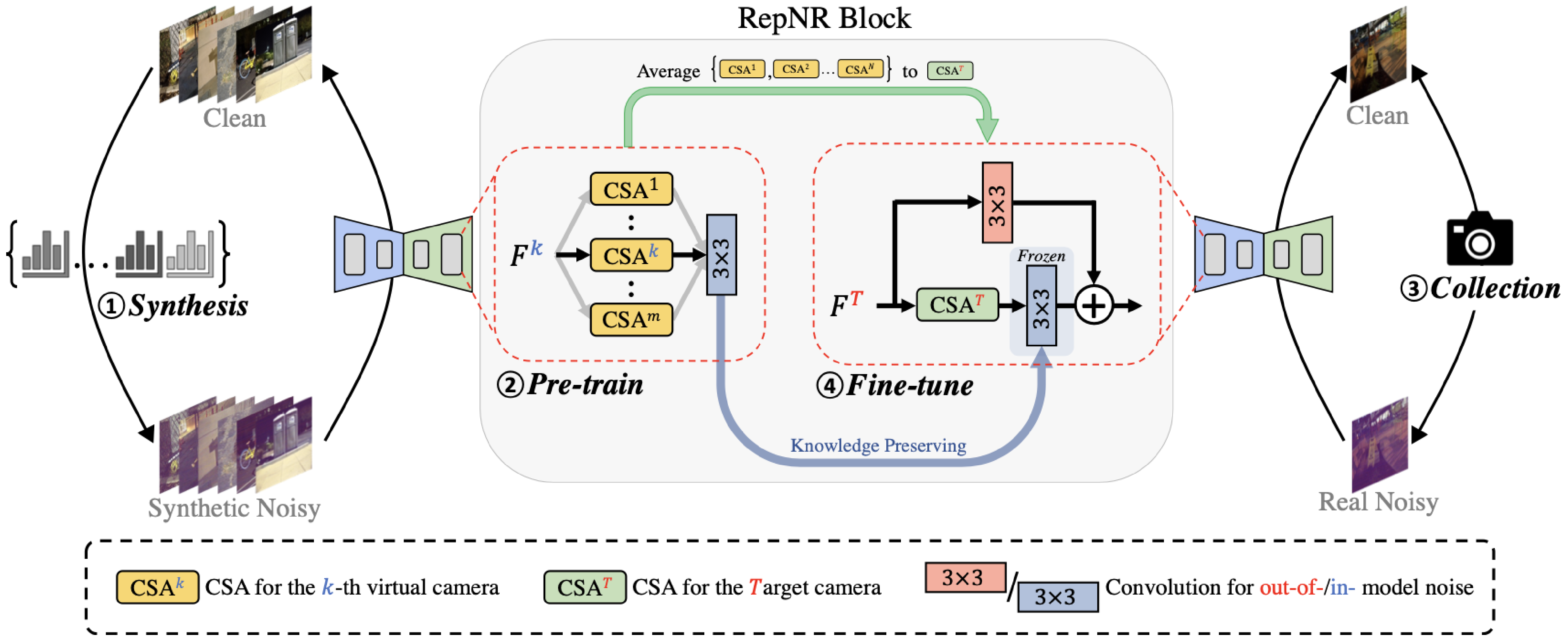
[2] Wei, Kaixuan, et al. "Physics-based noise modeling for extreme low-light photography." *IEEE TPAMI*. 2021.
[4] Feng, Hansen, et al. "Learnability enhancement for low-light raw denoising: Where paired real data meets noise modeling." *ACMMM*. 2022

LED: A Calibration-free Pipeline

1. **Pre-define noise model Φ, random sample noise parameters as “Virtual Cameras”.**
2. **Pre-train the network with synthesis noise.**
3. **Collect few-shot paired data with target camera.**
4. **Fine-tuning the network with data collected in 3.**



LED: A Calibration-free Pipeline



CSA^k $k * x + b$

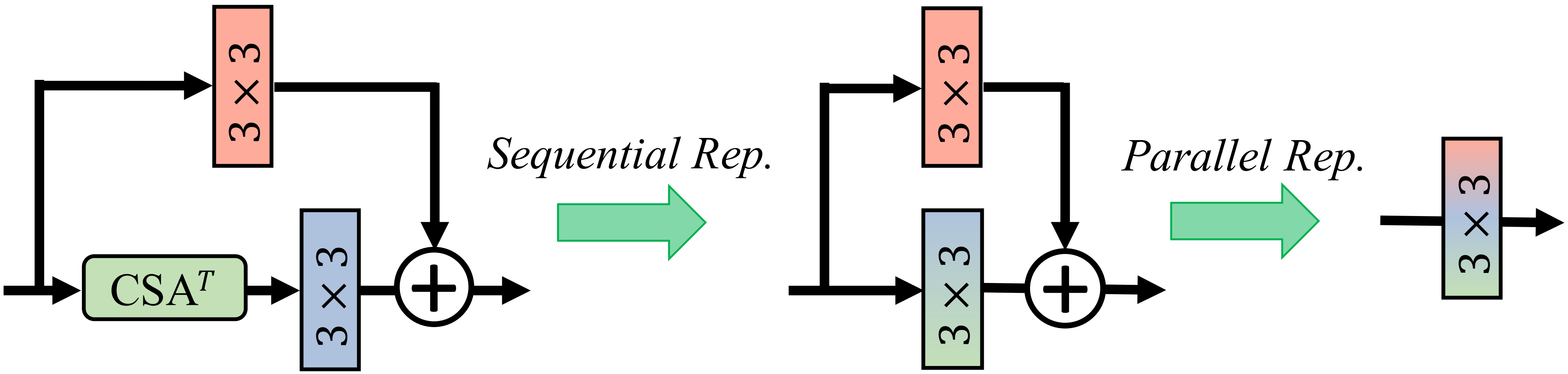
3×3 0 weight, 0 bias

Transfer Learning: rapidly deploy on new cameras

Continual Learning: generalize to real scenarios

LED: A Calibration-free Pipeline

- **Reparameterization: without any additional computational cost while deploying!**



Visualization

- **Out-Of-Model Noise**

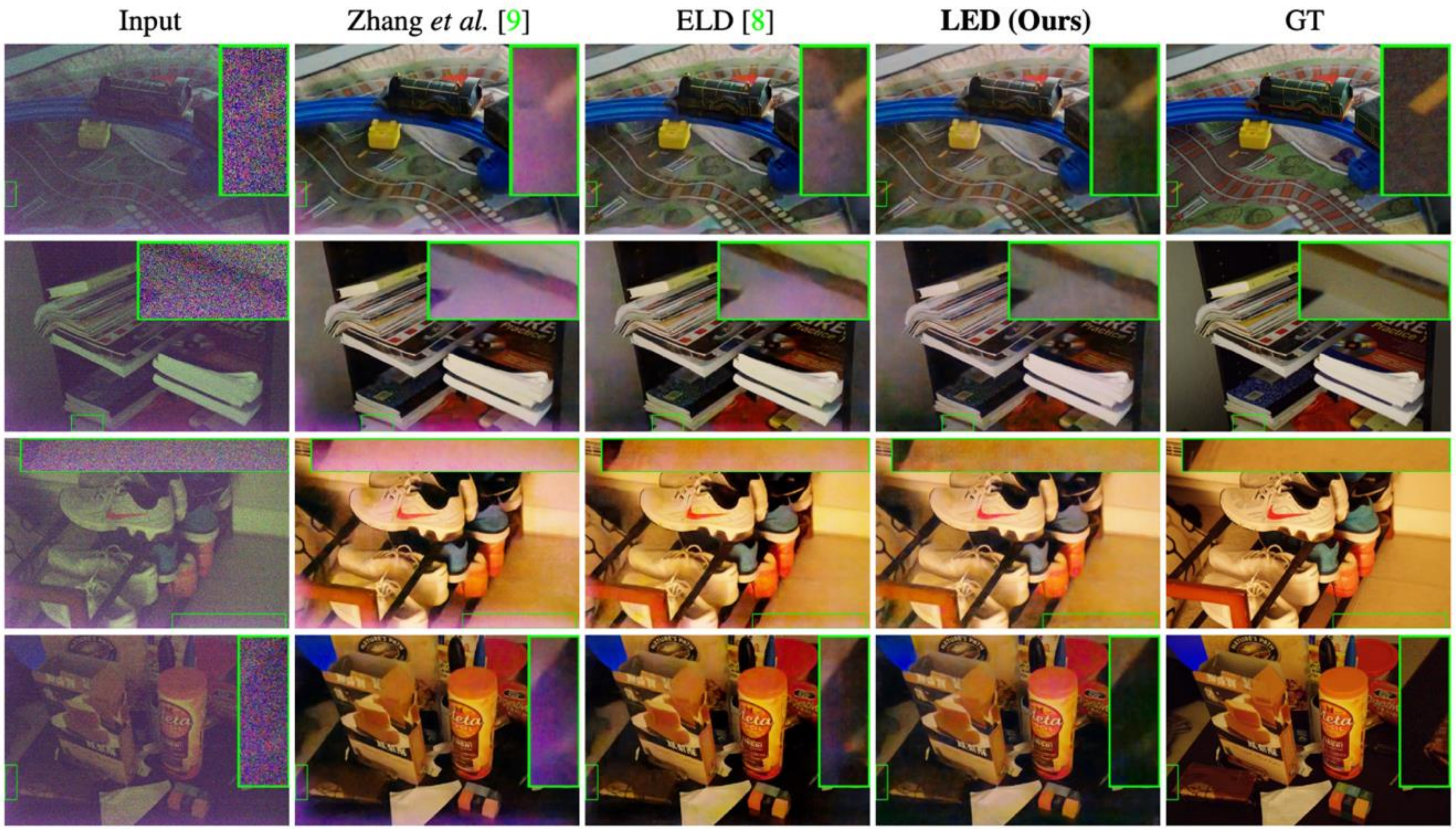
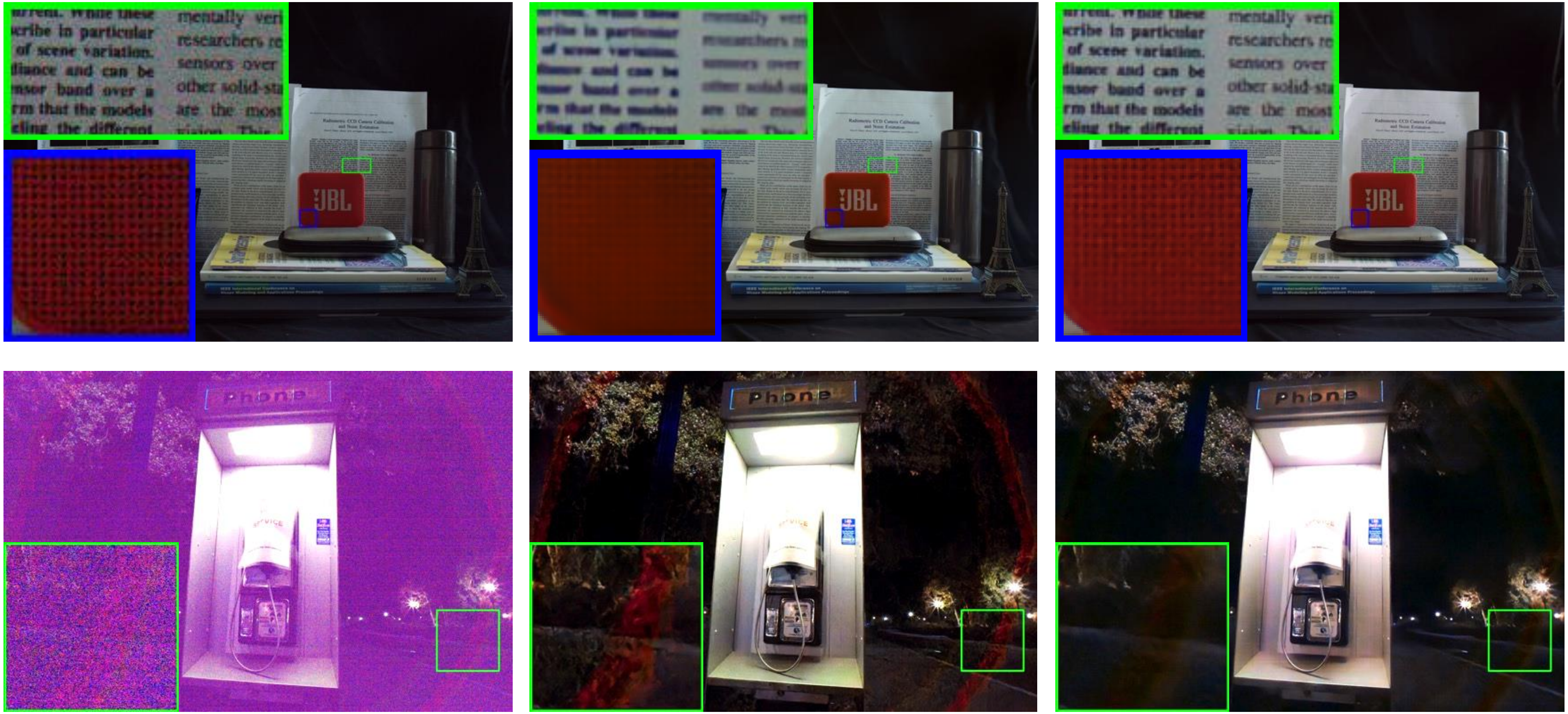


Figure 6. Compared with state-of-the-art calibration-based methods: ELD [8] and Zhang *et al.* [9], proposed LED is able to remove the out-of-model noise (*Zoom-in for best view*).

Visualization



Input ELD[2] **LED (Ours)**

[2] Wei, Kaixuan, et al. "Physics-based noise modeling for extreme low-light photography." *IEEE TPAMI*. 2021.

Discussion

- **Why with Two Pairs?**

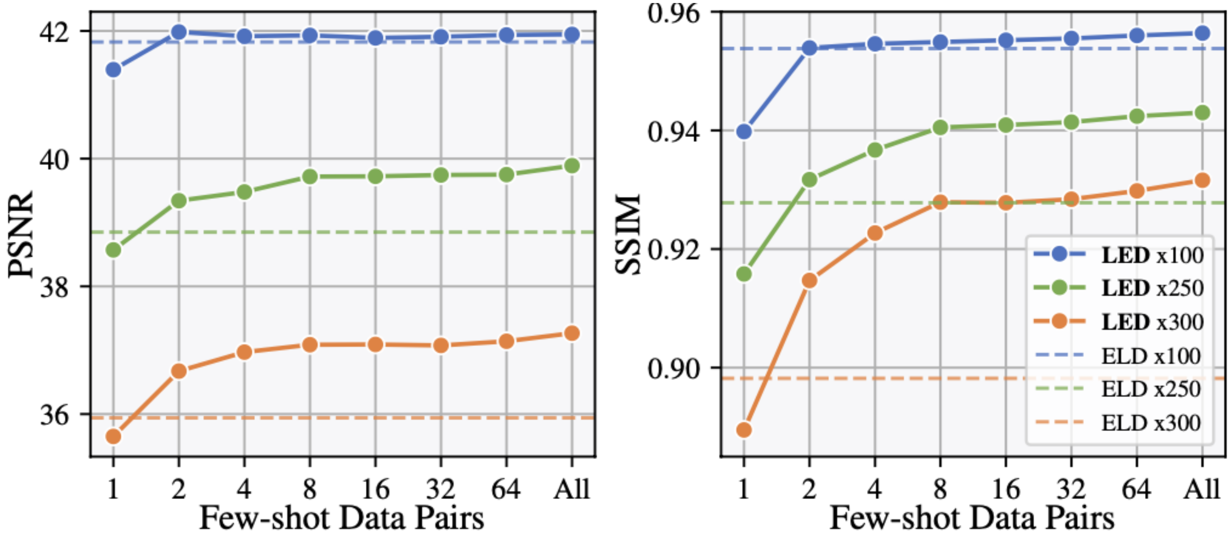


Figure 7. Ablation studies on the data amount for fine-tuning. LED achieves better performance with only 2 pairs for each ratio.

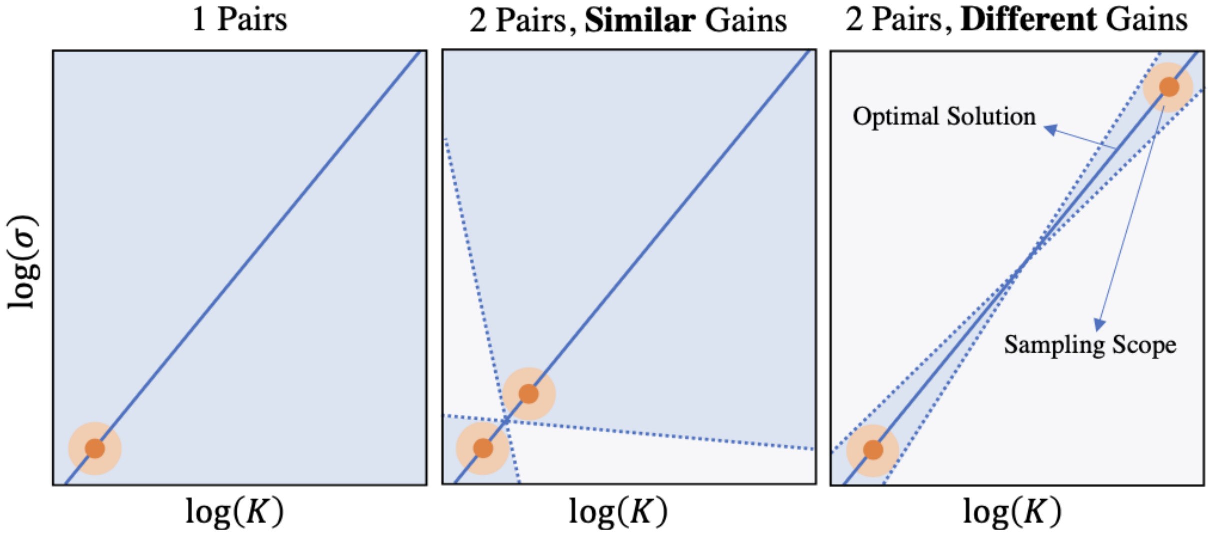


Figure 8. Illustration of the feasible solution space (blue area) of the linear relationship between the overall system gain $\log(K)$ and noise variance $\log(\sigma)$ under different sample strategies.

Table 6. Ablation studies on the pairs count for fine-tuning and testing on the synthetic dataset. N denotes fine-tuning with N pairs data of the similar overall system gain for each ratio. $N*$ denotes pairs data with marginally different overall system gains.

Ratio	1	2	4	2*
×100	41.295/0.9480	41.704/0.9523	41.432/0.9466	**43.795/0.9648**
×250	39.239/0.9350	39.410/0.9351	39.327/0.9367	**41.311/0.9457**
×300	38.314/0.9229	38.486/0.9216	38.499/0.9240	**39.190/0.9278**

Thanks!