Supplementary Material

ReRAW: RGB-to-RAW Image Reconstruction via Stratified Sampling for Efficient Object Detection on the Edge

Radu Berdan, Beril Besbinar, Christoph Reinders, Junji Otsuka, Daisuke Iso

1. Information content of RAW images

The difference in information content between RAW and RGB images is illustrated in Fig. S1. Taking the five different RAW image datasets, we compute the average entropy [9] of the RAW images and compare with the average entropy of the paired RGB images. In all cases, RAW images contain more information. The relationship between the RAW and RGB pixel values can also be visualized in Fig. S3, showcasing the variability of difficulty in the inverse reconstruction task.

2. Training on synthetic RAW images.

We perform a simple experiment to demonstrate the importance of the quality of converted synthetic RAW images (Fig. S1 (right)). A naive reverse ISP (which is a stripped-down version of ReRAW) is trained on a RAW dataset (NOD-Nikon) for 16 epochs, and a checkpoint is saved at each epoch. The PSNR of the conversion accuracy is measured for each. We then use each reverse ISP checkpoint to convert the ground-truth RGB images into a synthetic RAW set, and then train an object detector from scratch on each set. The evaluation is done on the ground-truth RAW test set. We observe a proportional relationship between reverse ISP conversion PSNR performance and the resulting detection performance of the object detector. The performance converges towards the ground truth case, where we trained the same detector on ground-truth RAW images. This highlights the importance of utilizing a high quality reverse ISP when creating synthetic RAW images.

Dataset	Concat.	Addition	Multiplication
NOD-Nikon	40.76	40.91	**41.00**
FIVEK-Canon	30.12	30.26	**30.45**

Table S1. Comparison on different global context incorporation methods on reconstruction performance (PSNR).

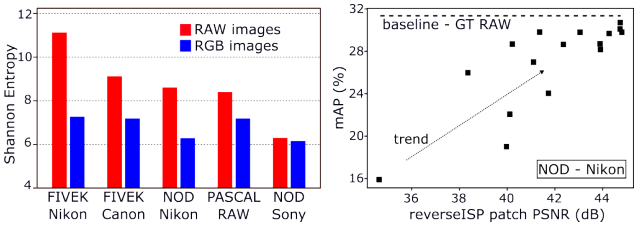

Figure S1. (Left) Shannon Entropy measured for RAW and their paired RGB images for 5 different RAW datasets [1, 5, 6]. The information content of RAW images is always higher than RGB. (right) One object detector trained from scratch on synthetic RAW images converted from original ground truth RGB images. Evaluation is done on ground-truth RAW images. The PSNR of the reverse ISP used to create the training sets shows correlation to the resulting detector performance.

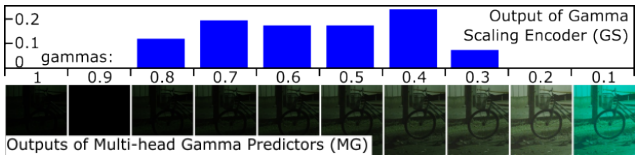

Figure S2. Examples of output of the Gamma Scaling Encoder for an example image from the NOD-Nikon dataset.

Method	Train len.	Patch size	# params (M)
CycleR2R [3]	100 ep.	512	15.8
U-Net [2]	100k iter.	256	1.9
SRISP [7]	800 ep.	256	1.38
InvISP [10]	300 ep.	256	1.4
ISPLess [4]	300 ep.	256	1.4
RAWDiffusion [8]	70k iter.	512	25.1
ReRAW	128 ep.	64	23.7

Table S2. Training parameters for the RGB2RAW task. Training methods have been kept identical to their published implementations.

References

[1] Vladimir Bychkovsky, Sylvain Paris, Eric Chan, and Frédo Durand. Learning photographic global tonal ad-

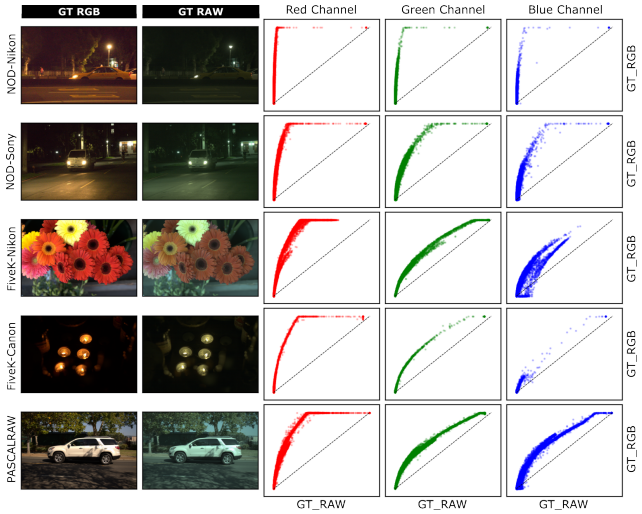

Figure S3. Examples of paired RAW and RGB image pixel color relationship. Columns represent: #1 Ground truth RGB image (GT_RGB); #2 Ground truth paired RAW image (GT_RAW); #3-5 plots of GT_RAW vs GT_RGB pixel values, per channel (RAW image green channels have been averaged, and resolution has been halved). Different cameras (corresponding to each dataset) produce different RAW-to-RGB conversion functions depending on image sensor type, ISP type and scene color/luminance characteristics. RAW images are displayed gamma corrected with $\gamma = 0.4$.

justment with a database of input / output image pairs. In *The Twenty-Fourth IEEE Conference on Computer Vision and Pattern Recognition*, 2011. 1, 3

[2] Marcos V Conde, Radu Timofte, Yibin Huang, Jingyang Peng, Chang Chen, Cheng Li, Eduardo Pérez-Pellitero, Fenglong Song, Furui Bai, Shuai Liu, et al. Reversed image signal processing and raw reconstruction. aim 2022 challenge report. In *European Conference on Computer Vision*, pages 3–26. Springer, 2022. 1

[3] Zhihao Li, Ming Lu, Xu Zhang, Xin Feng, M. Salman Asif, and Zhan Ma. Efficient visual computing with camera RAW snapshots. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–18, 2024. 1

[4] William Ljungbergh, Joakim Johnander, Christoffer Petersson, and Michael Felsberg. *Raw or Cooked? Object Detection on RAW Images*, page 374–385. Springer Nature Switzerland, 2023. 1

[5] Igor Morawski, Yu-An Chen, Yu-Sheng Lin, Shusil Dangi, Kai He, and Winston H Hsu. GenISP: Neural ISP for low-light machine cognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 630–639, 2022. 1, 3

[6] Ta David Omid-Zohoor, Alex and Boris. Murmann. Pascalraw: Raw image database for object detection. stanford digital repository. In *Available at: http://purl.stanford.edu/hq050zr7488*, 2014-2015. 1, 3

[7] Junji Otsuka, Masakazu Yoshimura, and Takeshi Ohashi. Self-Supervised Reversed Image Signal Processing via Reference-Guided Dynamic Parameter Selection. *arXiv preprint arXiv:2303.13916*, 2023. 1

[8] Christoph Reinders, Radu Berdan, Beril Besbinar, Junji Otsuka, and Daisuke Iso. Raw-diffusion: Rgb-guided diffusion models for high-fidelity raw image generation. In *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2025. 1

[9] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3): 379–423, 1948. 1

[10] Yazhou Xing, Zian Qian, and Qifeng Chen. Invertible image signal processing, 2021. 1

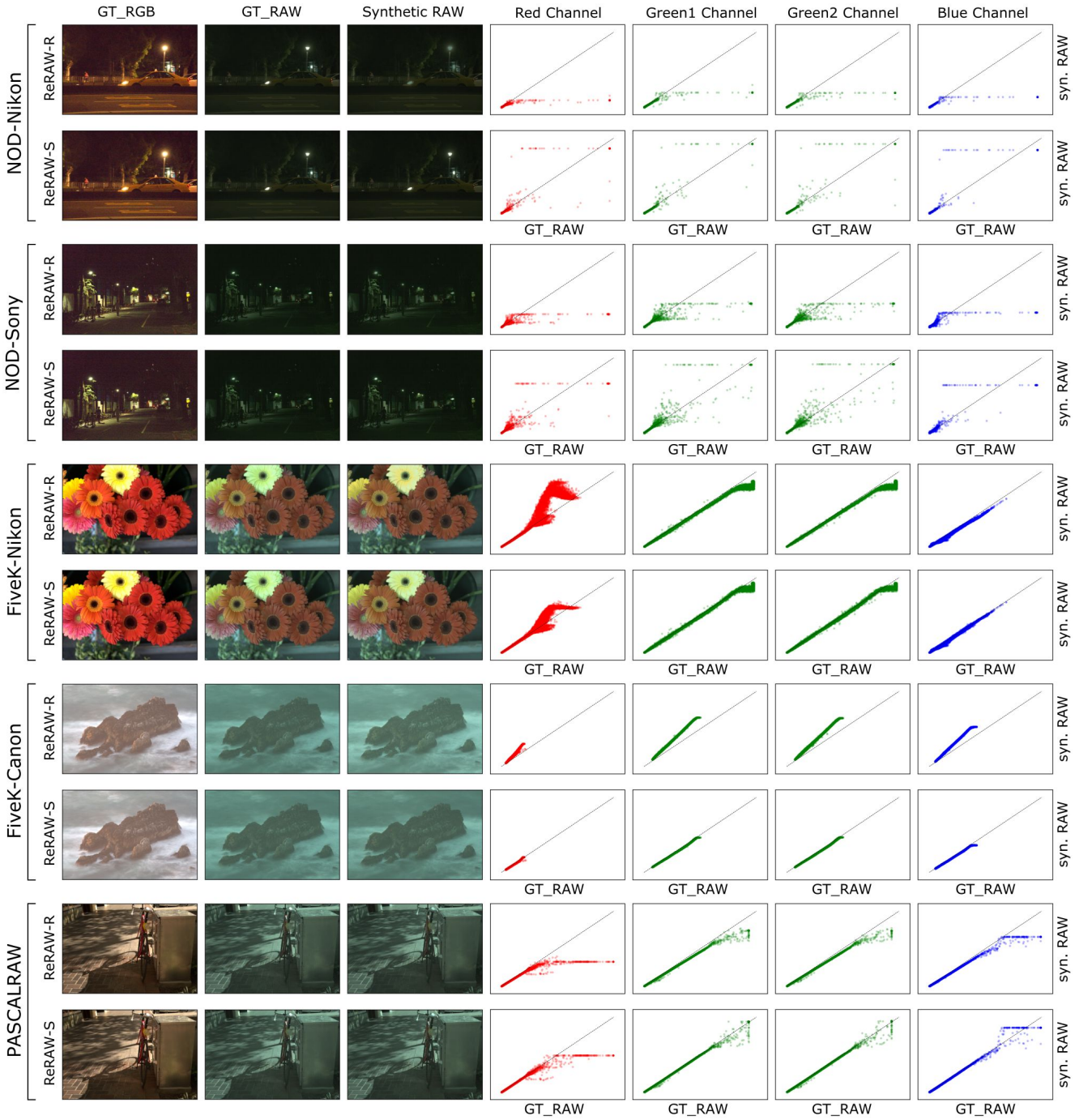

Figure S4. Qualitative and quantitative comparison of RGB-to-RAW conversions on five different datasets [1, 5, 6] between ReRAW-R (trained via random patch sampling) and ReRAW-S (trained via stratified patch sampling). Columns represent: #1 Ground truth RGB image (GT_RGB); #2; Ground truth RAW image (GT_RAW); #3 Synthetic RAW (converted). #4-7 GT_RAW vs synthetic RAW pixel value plot, per color channel. For all five tested datasets, ReRAW-S shows a better RAW conversion, especially in brighter pixel regions. Additionally, the resulting distribution between ground-truth RAW (GT_RAW) and synthetic pixel value is tighter in the ReRAW-S case. RAW images are displayed gamma corrected with $\gamma = 0.4$.

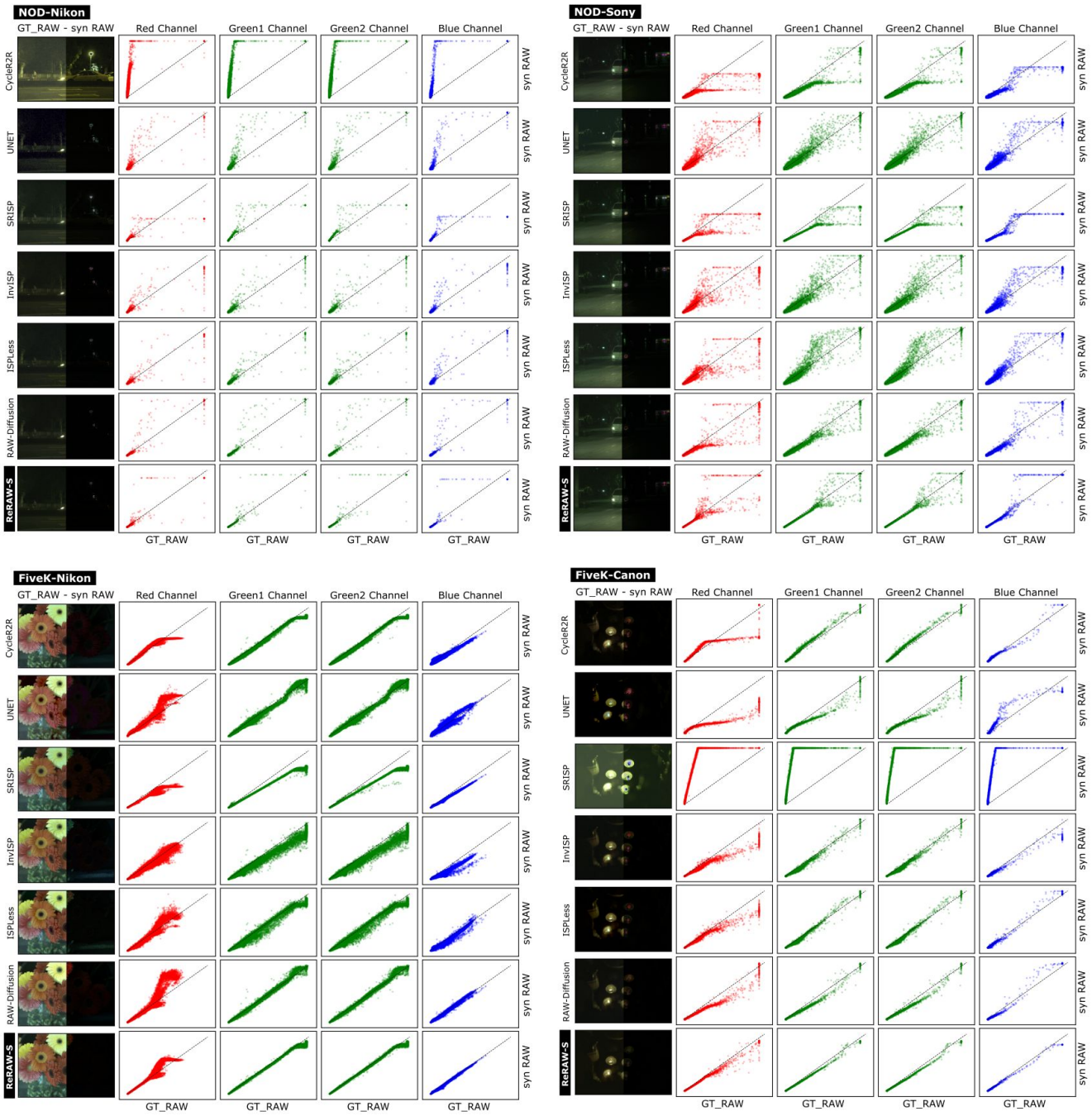

Figure S5. Qualitative and quantitative comparisons between ReRAW-S and state-of-the-art reverse ISPs, for four separate RAW datasets. Columns represent: #1 Synthetic (converted) RAW (half) and error vs ground truth RAW image (GT_RAW); #2-5 GT_RAW vs synthetic RAW pixel value plot, per color channel. ReRAW-S achieves tighter pixel value relationship vs ground truth RAW, better illumination estimation and better reconstruction of brighter RAW pixel values. RAW images are displayed gamma corrected with $\gamma = 0.4$.

Figure S6. Illustration of converted BDD subsets into daytime and nighttime BDD-ReRAW-S, facilitated by ReRAW trained on either PASCALRAW or NOD-Nikon. RAW images are displayed gamma corrected with $\gamma = 0.4$.